Chapter 9

Nonlinear Centering and Derivatives

9.1. Introduction

In this chapter we continue to study the useful functions introduced in Section 3.5 of the book and investigated in Chapter 13 of the book. Now we consider supremum, reflection and inverse maps with nonlinear centering.

Following Mandelbaum and Massey (1995), we identify the limit of the properly scaled function as a derivative. We also show how the convergence-preservation results for the reflection map can be applied to establish heavy-traffic limits for nonstationary queues.

To explain the derivative representation, recall that our previous results on the preservation of convergence with linear centering started with the assumed convergence

$$c_n(x_n - e) \to y \quad \text{in} \quad D\,, \tag{1.1}$$

where $c_n \to \infty$ and e is the identity function, i.e., $e(t) = t$, $t \geq 0$. Given (1.1), we found conditions under which

$$c_n(\phi(x_n) - e) \to z \quad \text{in} \quad D \tag{1.2}$$

for various functions ϕ and we identified the limit z. We also obtained some extensions in which the linear centering function e in (1.1) is replaced by a nonlinear function x; i.e., instead of (1.1), we assumed that

$$c_n(x_n - x) \to y \quad \text{in} \quad D \quad \text{as} \quad n \to \infty\,, \tag{1.3}$$

where $c_n \to \infty$. In particular, see Theorems 13.3.2, 13.7.2 and 13.7.4 and Corollaries 13.4.1, 13.7.1 and 13.7.2 in the book. We now want to obtain some further results of this kind.

Given (1.3), we have as a consequence

$$x_n \to x \quad \text{in} \quad D\,. \tag{1.4}$$

Hence, for any continuous function ϕ, we have

$$\phi(x_n) \to \phi(x) \quad \text{in} \quad D\,. \tag{1.5}$$

Thus we want to find functions $z \in D$ and regularity conditions such that

$$c_n(\phi(x_n) - \phi(x)) \to z \quad \text{in} \quad D\,. \tag{1.6}$$

The previous results with centering by e were of this form, where $\phi(x) = x = e$. The M topologies play an important role, because the limit z in (1.6) may have discontinuities even when y, x and x_n are all continuous functions.

In a probability context, (1.6) is interesting because it corresponds to a FCLT refinement to a nonlinear FLLN. We may have scaled stochastic processes $\{X_n(t) : t \geq 0\}$ which obey a nonlinear FWLLN of the form

$$X_n \Rightarrow x \quad \text{in} \quad D\,, \tag{1.7}$$

where x is a nonlinear deterministic function, and a FCLT refinement of the form

$$c_n(X_n - x) \Rightarrow Y \quad \text{in} \quad D\,, \tag{1.8}$$

where $c_n \to \infty$. From the FWLLN (1.7) it follows directly that

$$\phi(X_n) \Rightarrow \phi(x) \quad \text{in} \quad D \tag{1.9}$$

for a continuous function ϕ. Our goal is to establish the FCLT refinement of (1.9), i.e.,

$$c_n(\phi(X_n) - \phi(x)) \Rightarrow Z \quad \text{in} \quad D\,. \tag{1.10}$$

As before, (1.10) follows from (1.8) when (1.6) follows from (1.3). Hence we focus on obtaining (1.6) from (1.3).

It is interesting that, under regularity conditions, z in (1.6) can be thought of as a derivative of the map ϕ, in particular, a directional derivative of ϕ in the direction y, evaluated at x. To see that, it is convenient to index the functions by ϵ in such a way that x_n becomes x_ϵ and c_n becomes ϵ^{-1}. (That is without loss of generality.) Then (1.3) is equivalent to

$$\epsilon^{-1}(x_\epsilon - x) \to y \quad \text{as} \quad \epsilon \downarrow 0\,. \tag{1.11}$$

Without being too precise, we can rewrite (1.11) as

$$x_\epsilon = x + \epsilon y + o(\epsilon) \quad \text{as} \quad \epsilon \downarrow 0 \ . \tag{1.12}$$

Now, assuming that the function $\phi : D \to D$ satisfies

$$\phi(\tilde{x}_\epsilon + o(\epsilon)) - \phi(\tilde{x}_\epsilon) = o(\epsilon) \quad \text{as} \quad \epsilon \downarrow 0 \tag{1.13}$$

for any $\tilde{x}_\epsilon$ with $\tilde{x}_\epsilon \to x$ in D as $\epsilon \downarrow 0$ (which is not automatic), we have

$$\phi(x_\epsilon) = \phi(x + \epsilon y) + o(\epsilon) \quad \text{as} \quad \epsilon \downarrow 0 \tag{1.14}$$

and, given the ϵ-analog of (1.6),

$$\phi(x + \epsilon y) = \phi(x) + \epsilon z + o(\epsilon) \quad \text{as} \quad \epsilon \downarrow 0 \ . \tag{1.15}$$

From (1.15), it is evident that z can be given the directional derivative interpretation. Moreover, (1.14) and (1.15) together imply that

$$\epsilon^{-1}(\phi(x_\epsilon) - \phi(x)) \to z \quad \text{as} \quad \epsilon \downarrow 0 \ . \tag{1.16}$$

Equivalently, (1.3), (1.13) and (1.16) imply the desired (1.6).

Here is how the present chapter is organized: In Section 2 we investigate when the convergence-preservation question (when (1.3) implies (1.6)) can be reduced to the derivative determination in (1.15). Unfortunately, we are not able to show that this can be done as generally as we would like. This step seems to be the weak link in our analysis in this chapter. Hopefully future research will provide further insights.

In Sections 9.3 – 9.5 we determine sufficient conditions for the derivatives of the supremum and reflection maps to exist and determine their form. As should be anticipated from Chapter 13 in the book, the reflection derivative can be expressed in terms of the supremum derivative. The M_1 topology plays an important role even if x and y in (1.3) are both continuous.

In Section 9.6 we apply the derivative calculation and convergence-preservation results for the reflection map to establish heavy-traffic limits for nonstationary queues. For example, these results cover the $\mathrm{M_t/M_t/1}$ queue with time-dependent arrival and service rates.

Finally, in Section 9.7 we consider the derivative of the inverse map.

9.2. Nonlinear Centering and Derivatives

In this section we investigate when the desired convergence-preservation (when (1.11) implies (1.16)) can be deduced by determining the derivative via (1.15). For any function $\phi : D \to D$, a general approach to establish the desired limit (1.16) for $\phi(x_\epsilon)$ is to exploit the triangle inequality:

$$d(\epsilon^{-1}[\phi(x_\epsilon) - \phi(x)], z) \leq d(\epsilon^{-1}[\phi(x + \epsilon y) - \phi(x)], z) + d(\epsilon^{-1}[\phi(x_\epsilon) - \phi(x)], \epsilon^{-1}[\phi(x + \epsilon y) - \phi(x)]) \quad (2.1)$$

for an appropriate metric d. A limit for the first term in (2.1) as $\epsilon \downarrow 0$ identifies z as the derivative of ϕ in the direction y evaluated at x. In addition to establishing the existence of this derivative, we must also show that the second term in (2.1) converges to 0 as $\epsilon \downarrow 0$. Surprisingly, the second term presents difficulties. However, we are able to show that it is negligible under regularity conditions. The results are in a good form when $y \in C$, but not so good when only $y \in D$. (Recall that the limit z in (1.16) may be discontinuous even if $y \in C$, so the case $y \in C$ is interesting and important.)

We now obtain results about the second term in (2.1) for general functions $\phi : (D_1, d_1) \to (D_2, d_2)$, where $D_i \equiv D([0, t_i], \mathbb{R}^{k_i})$ for $i = 1, 2$.

Theorem 9.2.1. (reduction of convergence preservation to the derivative) *Suppose that $\phi : (D_1, d_1) \to (D_2, d_2)$, where the metrics d_i satisfy the properties:*

$$d_i(cx_1, cx_2) = cd_i(x_1, x_2) \quad \textit{for all} \quad c > 0, \; i = 1, 2 \; , \quad (2.2)$$

$$d_i(x_1 + x_3, x_2 + x_3) = d_i(x_1, x_2), \; i = 1, 2, \quad (2.3)$$

$$d_2(\phi(x_1), \phi(x_2)) \leq K d_1(x_1, x_2) \quad \textit{for some} \quad K > 0 \; , \quad (2.4)$$

for all x_1, x_2, $x_3 \in D_i$. Then

$$d_2(\epsilon^{-1}[\phi(x_\epsilon) - \phi(x)], \epsilon^{-1}[\phi(x + \epsilon y) - \phi(x)]) \leq K d_1(\epsilon^{-1}(x_\epsilon - x), y) \; . \quad (2.5)$$

Proof. The conditions imply that

$$\begin{aligned} & d_2(\epsilon^{-1}[\phi(x_\epsilon) - \phi(x)], \epsilon^{-1}[\phi(x + \epsilon y) - \phi(x)]) = \epsilon^{-1} d_2(\phi(x_\epsilon), \phi(x + \epsilon y)) \\ & \leq \epsilon^{-1} K d_1(x_\epsilon, x + \epsilon y) \\ & = K d_1(\epsilon^{-1}(x_\epsilon - x), y) \; . \quad \blacksquare \end{aligned}$$

Notice that the uniform metric satisfies conditions (2.2) and (2.3). The following application of Theorem 9.2.1 is elementary.

Theorem 9.2.2. (reduction for the supremum and reflection maps with the uniform metric) *If d_1 and d_2 in Theorem* 9.2.1 *above are the uniform* *metric on* $D([0,t],\mathbb{R})$ *and* ϕ *is the supremum function in equation* (13.4.1) *in the book or the reflection map in equation* (13.5.1) *in the book, then the conditions of Theorem* 9.2.1 *above are satisfied, so that conclusion* (2.5) *holds.*

Proof. It is evident that the uniform metric on D satisfies conditions (2.2) and (2.3). The supremum and reflection functions also satisfy (2.4) with respect to the uniform metric by Lemmas 13.4.1 and 13.5.1 in the book.

Example 9.2.1. *The need for the map* ϕ *to be Lipschitz.* To see the need for $\phi : D \to D$ being Lipschitz in Theorem 9.2.1, let $\phi(x)(t) = \sqrt{x(1)}$, $t \geq 0$. If $\|x_\epsilon - x\|_t \to 0$ for $t > 1$, then $\|\phi(x_\epsilon) - \phi(x)\|_t \to 0$, but ϕ is not Lipschitz. Suppose that $x(t) = 0$, $y(t) = 1$ and $x_\epsilon(t) = x(t) + \epsilon y(t) = \epsilon$, $t \geq 0$. Then $\|\epsilon^{-1}(x_\epsilon - x) - y\| = 0$ for all ϵ,

$$\epsilon^{-1}[\phi(x_\epsilon) - \phi(x)](t) = \epsilon^{-1}[\sqrt{\epsilon} - 0] = \epsilon^{-1/2} \to \infty \quad \text{as} \quad \epsilon \downarrow 0\ . \quad \blacksquare \tag{2.6}$$

Unfortunately, for the non-uniform Skorohod metrics on D, which we will want to consider when $y \notin C$, we do not have properties (2.2) and (2.3) in Theorem 9.2.1.

Example 9.2.2. *Failure for nonuniform metrics.* Unlike with the uniform metric, we cannot conclude that $d(\epsilon^{-1}x_\epsilon, \epsilon^{-1}x + y) \to 0$ as $\epsilon \downarrow 0$ when $d(\epsilon^{-1}(x_\epsilon - x), y) \to 0$ as $\epsilon \downarrow 0$ if d is the J_1, M_1 or M_2 metric and y is not continuous. To see this, let $x(t) = tI_{[0,1)}(t) + (2-t)I_{[1,2]}(t)$, $y = I_{[0,1)} - I_{[1,2]}$ and $x_\epsilon = (x+\epsilon)I_{[0,1-\epsilon]} + (x-\epsilon)I_{[1-\epsilon,2]}$ in $D([0,2],\mathbb{R})$. Then $\epsilon^{-1}(x_\epsilon - x) = y \circ \lambda_\epsilon$, where $\lambda_\epsilon \in \Lambda$ with $\lambda_\epsilon(1) = 1 - \epsilon$, $\lambda_\epsilon(0) = 0$ and $\lambda_\epsilon(2) = 2$. Hence $d_{J_1}(\epsilon^{-1}(x_\epsilon - x), y) = \|\lambda_\epsilon - e\| = \epsilon \to 0$ as $\epsilon \downarrow 0$. However $\epsilon^{-1}x_\epsilon^\uparrow(2) = \epsilon^{-1}x_\epsilon^\uparrow(1-\epsilon) = \epsilon^{-1}$, while $(\epsilon^{-1}x + y)^\uparrow(2) = (\epsilon^{-1}x + y)(1-) = \epsilon^{-1} + 1$, so that $d_{M_2}(\epsilon^{-1}x_\epsilon, \epsilon^{-1}x + y) \geq 1$. $\blacksquare$

However, under regularity conditions, we can also establish results starting from J_1, M_1 and M_2 convergence. We state the following results for the strong SJ_1, SM_1 and SM_2 metrics on $D([0,t],\mathbb{R}^k)$. Corresponding results for the product metrics for Lemmas 9.2.1 and 9.2.2 below follow; just consider one coordinate at a time.

Recall that x is Lipschitz on $[0,t]$ if there is a constant K so that $|x(t_1) - x(t_2)| \leq K|t_1 - t_2|$ for $0 \leq t_1, t_2 \leq t$. This regularity condition is typically satisfied in applications, because x often satisfies an ordinary differential equation (ODE). If x is absolutely continuous with derivative $\dot{x}$, where $\dot{x} \in$

D, then for each $t > 0$, there exists K such that $|\dot{x}(s)| \le K$ for $0 \le s \le t$ and, for $0 \le t_1 < t_2 \le t$,

$$|x(t_2) - x(t_1)| \le \int_{t_1}^{t_2} |\dot{x}(s)| ds \le K|t_2 - t_1| \;, \tag{2.7}$$

so that x is Lipschitz.

Lemma 9.2.1. (subtracting a common Lipschitz function) *Suppose that x is Lipschitz in $[0,t]$ with Lipschitz constant K. If d_t is the SJ_1, SM_1 or SM_2 metric on $D([0,t],\mathbb{R}^k)$, then*

$$d_t(x_1 - x, x_2 - x) \le (1+K) d_t(x_1, x_2) \;. \tag{2.8}$$

Proof. First consider J_1. For all $\epsilon > 0$, there exist $\eta(\epsilon) > 0$ and increasing homeomorphisms λ_ϵ of $[0,t]$ such that

$$\|x_1 - x_2 \circ \lambda_\epsilon\|_t \vee \|\lambda_\epsilon - e\|_t \le (1 + \eta(\epsilon)) d_t(x_1, x_2) \;.$$

It follows that

$$\begin{aligned} \|(x_1 - x) - (x_2 - x) \circ \lambda_\epsilon\|_t &\le \|x_1 - x_2 \circ \lambda_\epsilon\|_t + \|x - x \circ \lambda_\epsilon\|_t \\ &\le (1 + \eta(\epsilon)) d_t(x_1, x_2) + K\|\lambda_\epsilon - e\|_t \\ &\le (1 + \eta(\epsilon) + K[1 + \eta(\epsilon)]) d_t(x_1, x_2) \;. \end{aligned}$$

Since $\eta(\epsilon)$ can be made arbitrarily small, the proof for J_1 is complete. Now consider M_1. For all $\epsilon > 0$ and $t > 0$, there exist $\eta(\epsilon) > 0$ and parametric representations $(u_1, r_{1\epsilon})$ of x_1 and $(u_{2\epsilon}, r_{1\epsilon})$ of x_2 such that

$$\|u_{1\epsilon} - u_{2\epsilon}\| \vee \|r_{1\epsilon} - r_{2\epsilon}\| \le (1 + \eta(\epsilon)) d_t(x_1, x_2) \;.$$

Since x is continuous, $(x \circ r_{1\epsilon}, r_{1\epsilon})$ and $(x \circ r_{2\epsilon}, r_{2\epsilon})$ are parametric representations of x, $(u_{1\epsilon} - x \circ r_{1\epsilon}, r_{1\epsilon})$ and $(u_{2\epsilon} - x \circ r_{2\epsilon}, r_{2\epsilon})$ are parametric representations of $x_1 - x$ and $x_2 - x$, and

$$\begin{aligned} \|(u_{1\epsilon} - x \circ r_{1\epsilon}) - (u_{2\epsilon} - x \circ r_{2\epsilon})\| &\le \|u_{1\epsilon} - u_{2\epsilon}\| + \|x \circ r_{1\epsilon} - x \circ r_{2\epsilon}\| \\ &\le (1 + \eta(\epsilon)) d_t(x_1, x_2) + K\|r_{1\epsilon} - r_{2\epsilon}\| \\ &\le (1 + \eta(\epsilon) + K[1 + \eta(\epsilon)]) d_t(x_1, x_2) \;. \end{aligned}$$

Since $\eta(\epsilon)$ can be arbitrarily small, the proof for M_1 is complete. Now consider M_2. let $(z_1, t_1) \in \Gamma_{x_1}$. If $(z_2, t_2) \in \Gamma_{x_2}$ is such that $\|(z_1, t_1) - (z_2, t_2\| < \delta$, then $(z_1 - x(t_1), t_1) \in \Gamma_{x_1 - x}$, $(z_2 - x(t_2), t_2) \in \Gamma_{x_2 - x}$ and

$$\begin{aligned} \|(z_1 - x(t_1), t_1) \vee (z_2 - x(t_2), t_2)\| &\le \|(z_1, t_1) - (z_2, t_2)\| + \|x(t_1) - x(t_2)\| \\ &\le \delta + K\|t_1 - t_2\| \le (1 + K)\delta \;. \quad \blacksquare \end{aligned}$$

Next we generalize (2.2).

Lemma 9.2.2. (deterministic scaling) *Let d_t be the SJ_1, SM_1 or SM_2 metric on $D([0,t],\mathbb{R}^k)$. For any $c>0$,*

$$d_{ct}(cx_1 \circ c^{-1}e, cx_2 \circ c^{-1}e) = cd_t(x_1,x_2) \tag{2.9}$$

or, equivalently,

$$d_t(cx_1, cx_2) = cd_{t/c}(x_1 \circ ce, x_2 \circ ce) \le (c \vee 1)d_t(x_1,x_2)\ . \tag{2.10}$$

Proof. First, for SJ_1, note that $\lambda \in \Lambda_t$ if and only if $c\lambda \circ c^{-1}e \in \Lambda_{ct}$ for $c>0$ and

$$\|c\lambda \circ c^{-1}e - e\|_{ct} = c\|\lambda - e\|_t\ .$$

Hence

$$\begin{aligned} &d_{ct}(cx_1 \circ c^{-1}e, cx_2 \circ c^{-1}e) \\ &\quad = \inf_{\lambda\in\Lambda_t}\{\|cx_1 \circ c^{-1}e - (cx_2 \circ c^{-1}e)\circ(c\lambda \circ c^{-1}e)\|_{ct} \vee \|c\lambda \circ c^{-1}e - e\|_{ct}\} \\ &\quad = \inf_{\lambda\in\Lambda_t}\{c\|x_1 - x_2\circ\lambda\|_t \vee c\|\lambda - e\|_t\} \\ &\quad = cd_t(x_1,x_2)\ . \end{aligned}$$

Next, for SM_2, note that $c\Gamma_{x_i}$ is the graph of $cx_i \circ c^{-1}e$ over $[0,ct]$ if and only if Γ_{x_i} is the graph of x_i over $[0,t]$. Hence (2.9) holds. Finally, for SM_1, note that (cu_i, cr_i) is a parametric representation of $cx_i \circ c^{-1}e$ over $[0,ct]$ if and only if (u_i,r_i) is a parametric representation of x_i over $[0,t]$. Hence (2.9) holds. ∎

Our next result goes beyond Theorem 9.2.1 by allowing the map ϕ to be Lipschitz with respect to the SJ_1, SM_1 or SM_2 metrics, but not the uniform metric.

Theorem 9.2.3. (Lipschitz functions with respect to non-uniform metrics) *Suppose that $y \in D([0,t_1],\mathbb{R}^{k_1})$ and x, x_ϵ, $x+\epsilon y$ all belong to a subset A of $D([0,t_1],\mathbb{R}^{k_1})$ for sufficiently small $\epsilon>0$. Suppose that $\phi : A \to D([0,t_2],\mathbb{R}^{k_2})$ is Lipschitz with respect to the metrics d_1 on A and d_2 on $D([0,t_2],\mathbb{R}^{k_2})$, i.e., there is a constant K such that*

$$d_2(\phi(x_1),\phi(x_2)) \le Kd_1(x_1,x_2) \tag{2.11}$$

for all $x_1,x_2 \in A$, where d_1 and d_2 are non-uniform Skorohod metrics (not necessarily the same). Suppose that x is Lipschitz on $[0,t_1]$ and $\phi(x)$ is Lipschitz on $[0,t_2]$. Then there is a constant K' such that

$$\begin{aligned} &d_2(\epsilon^{-1}[\phi(x_\epsilon)-\phi(x)], \epsilon^{-1}[\phi(x+\epsilon y)-\phi(x)]) \\ &\qquad \le K'\epsilon^{-1}d_1(x_\epsilon - x, \epsilon y) \\ &\qquad \le K'\|\epsilon^{-1}(x_\epsilon - x) - y\|_{t_1}\ . \end{aligned} \tag{2.12}$$

Proof. By Lemmas 9.2.2 and 9.2.1 and the assumptions, for $\epsilon < 1$, there are constants K_1, K_2 and K_3 such that

$$\begin{aligned} &d_2(\epsilon^{-1}[\phi(x_\epsilon) - \phi(x)], \epsilon^{-1}[\phi(x+\epsilon y) - \phi(x)]) \\ &\quad \le \epsilon^{-1} d_2(\phi(x_\epsilon) - \phi(x), \phi(x+\epsilon y) - \phi(x)) \\ &\quad \le K_1 \epsilon^{-1} d_2(\phi(x_\epsilon), \phi(x+\epsilon y)) \\ &\quad \le K_1 K_2 \epsilon^{-1} d_1(x_\epsilon, x+\epsilon y) \\ &\quad \le K_1 K_2 K_3 \epsilon^{-1} d_1(x_\epsilon - x, \epsilon y) \\ &\quad \le K_1 K_2 K_3 \epsilon^{-1} \|x_\epsilon - x - \epsilon y\|_{t_1} \\ &\quad \le K_1 K_2 K_3 \|\epsilon^{-1}(x_\epsilon - x) - y\|_{t_1} . \ \blacksquare \end{aligned} \tag{2.13}$$

The final upper bound in Theorem 9.2.3 does not help with the supremum and reflection maps because the supremum and reflection maps already have the required Lipschitz properties with respect to the uniform metric, by Theorem 9.2.2. In order to apply Theorem 9.2.3 without having to resort to the cruder uniform metric bound, we need to have

$$d_1(x_\epsilon - x, \epsilon y) = o(\epsilon) \quad \text{as} \quad \epsilon \downarrow 0 . \tag{2.14}$$

First, from this analysis, we see the need to be precise about what we mean about $o(\epsilon)$ terms in (1.12)–(1.15). Next, we observe that $d_1(\epsilon^{-1}[x_\epsilon - x], y) \to 0$ does not directly imply that $d_1(x_\epsilon - x, \epsilon y) = o(\epsilon)$ as $\epsilon \downarrow 0$, but that it is possible to have $d_1(x_\epsilon - x, \epsilon y) = o(\epsilon)$ as $\epsilon \downarrow 0$ without having $\|x_\epsilon - x - \epsilon y\|_{t_1} = o(\epsilon)$ as $\epsilon \downarrow 0$.

Example 9.2.3. *Condition* (2.14) *is weaker than the usual limit.* We would like to have $\epsilon^{-1} d_t(x_\epsilon - x, \epsilon y) \to 0$ as $\epsilon \downarrow 0$ whenever $d_t(\epsilon^{-1}(x_\epsilon - x), y) \to 0$ as $\epsilon \downarrow 0$, so that we could improve upon (2.12), but that implication is not valid. To see that, let $x = y = I_{[1,2]}$ in $D([0,2], \mathbb{R})$ and let $x_\epsilon = x + \epsilon I_{[1+\delta_\epsilon, 2]}$. Then $\epsilon^{-1}(x_\epsilon - x) = I_{[1+\delta_\epsilon, 2]}$ and $d_{J_1}(\epsilon^{-1}(x_\epsilon - x), y) = \delta_\epsilon$. On the other hand $\epsilon^{-1} d_{J_1}(x_\epsilon - x, \epsilon y) = \epsilon^{-1}(\epsilon \wedge \delta_\epsilon)$, which converges to 0 if and only if $\epsilon^{-1}\delta_\epsilon \to 0$ as $\epsilon \to 0$. Hence, we do not necessarily have $\epsilon^{-1} d_t(x_\epsilon - x, \epsilon y) \to 0$ as $\epsilon \downarrow 0$, given $d_t(\epsilon^{-1}(x_\epsilon - x, y) \to 0$, but we could have it, as is the case here when $\epsilon^{-1}\delta_\epsilon \to 0$ as $\epsilon \downarrow 0$. On the other hand, $\|\epsilon^{-1}(x_\epsilon - x) - y\| = 1$ for all $\epsilon > 0$. $\blacksquare$

Example 9.2.4. *A parametric family of examples.* Consider Example 9.2.2 modified by having

$$x_\epsilon = (x + \epsilon) I_{[0, 1-\epsilon^p)} + (x - \epsilon) I_{[1-\epsilon^p, 2]} . \tag{2.15}$$

Then $x_\epsilon - x = \epsilon y \circ \lambda_\epsilon$ for $\lambda_\epsilon(1) = 1 - \epsilon^p$, $\lambda_\epsilon(0) = 0$ and $\lambda_\epsilon(2) = 2$ with λ_ϵ defined by linear interpolation elsewhere. Thus

$$d_{J_1}(x_\epsilon - x, \epsilon y) = d_{J_1}(\epsilon^{-1}(x_\epsilon - x), y) = \|\lambda_\epsilon - e\| = \epsilon^p \ , \tag{2.16}$$

so that condition (2.14) holds if $p > 1$, but not if $0 < p \leq 1$.

9.3. Derivative of the Supremum Function

In this section we consider the derivative of the supremum function; i.e., we find conditions under which the limit (1.15) is valid and identify the limit z when $\phi : D \to D$ is the supremum function. The supremum function maps $x \in D \equiv D([0,T], \mathbb{R})$ into $x^\uparrow \in D$ for

$$x^\uparrow(t) \equiv \sup_{0 \leq s \leq t} x(s), \quad 0 \leq t \leq T \,. \tag{3.1}$$

In order to treat the derivatives, we will find it necessary to consider functions outside of D. Thus let D_{lim} be the set of functions with left and right limits everywhere, but without having to be either left continuous or right continuous at each discontinuity point. In general, we will only be able to conclude (in Theorem 9.3.2 below) that the derivative belongs to D_{lim}. In our definition of the derivative, we start by allowing one function to be in D_{lim}. For $x \in D$, $y \in D_{lim}$ and $\epsilon > 0$, let

$$z_\epsilon \equiv z_\epsilon(x, y) \equiv \epsilon^{-1}[(x + \epsilon y)^\uparrow - x^\uparrow] = (\epsilon^{-1}x + y)^\uparrow - \epsilon^{-1}x^\uparrow \,. \tag{3.2}$$

The derivative of the supremum function (in the direction y, evaluated at x) is the limit of z_ϵ as $\epsilon \downarrow 0$, if it exists. We will show that the limit does exist under regularity conditions and identify it. In this section we consider pointwise convergence for all t; in the next section we consider M_1 and M_2 convergence.

We start by stating two elementary lemmas; the second follows from the first.

Lemma 9.3.1. (the case of constant y) *If $y(s) = c$, $0 \leq s \leq t$, then $z_\epsilon(t) = c$ for all ϵ.*

For $z^\downarrow$ be the infimum function; i.e., $z^\downarrow = -(-z)^\uparrow$.

Lemma 9.3.2. (monotone bounds) *For all $\epsilon > 0$, $y^\downarrow \leq z_\epsilon \leq y^\uparrow$.*

Even though x is right-continuous, it can approach its supremum from the left ($x(s) = sI_{[0,t_0)}(s)$) or right ($x(s) = -sI_{[t_0,t_1)}(s)$). Let $\Phi_x^L(t)$ and $\Phi_x^R(t)$ be the subsets of time points in $[0,t]$ at which the left and right limits of x attain the supremum; i.e.,

$$\Phi_x^L(t) = \{s : 0 < s \le t, x(s-) = x^{\uparrow}(t)\} \tag{3.3}$$

and

$$\Phi_x^R(t) = \{s : 0 \le s \le t, x(s+) = x^{\uparrow}(t)\} \ . \tag{3.4}$$

Let $\Phi_x(t) = \Phi_x^L(t) \cup \Phi_x^R(t)$. When $x \in C$, $\Phi_x^L(t) = \Phi_x^R(t)$.

Example 9.3.1. *The possibility of empty sets.* It is possible for Φ_x^L or $\Phi_x^R(t)$ to be empty: Let $x(t) = tI_{[0,1)}(t)$, $t \ge 0$. Then, for $t \ge 1$, $\Phi_x^L(t) = \{1\}$, while $\Phi_x^R(t) = \phi$. However, $\Phi_x^L(t) \cup \Phi_x^R(t) \ne \phi$. ∎

These subsets need not be closed, but they have the following partial closure property.

Lemma 9.3.3. (partial closure property) *For any $x \in D$ and $t \ge 0$, $\Phi_x^L(t)$ is closed from the left, while $\Phi_x^R(t)$ is closed from the right; i.e., if $s_n \uparrow s$ in $[0,t]$ and $s_n \in \Phi_x^L(t)$ for all n, then $s \in \Phi_x^L(t)$; if $s_n \downarrow s$ and $s_n \in \Phi_x^R(t)$ for all n, then $s \in \Phi_x^R(t)$. Moreover, if $s_n \uparrow s$ in $[0,t]$ and $s_n \in \Phi_x^R(t)$ for all n, then $s \in \Phi_x^L(t)$; if $s_n \downarrow s$ in $[0,t]$ and $s_n \in \Phi_x^L(t)$ for all n, then $s \in \Phi_x^R(t)$.*

Corollary 9.3.1. (compactness of $\Phi_x(t)$) *For each $t > 0$, $\Phi_x(t)$ is a compact subset of $[0,t]$.*

We next show that z_ϵ is monotone in ϵ.

Lemma 9.3.4. (monotonicity in ϵ) *For z_ϵ in (3.2), $z_\epsilon(t)$ decreases as ϵ decreases for each t.*

Proof. We want to show that

$$(\epsilon_2^{-1}x + y)^{\uparrow} - \epsilon_2^{-1}x^{\uparrow} < (\epsilon_1^{-1}x + y)^{\uparrow} - \epsilon_1^{-1}x^{\uparrow}$$

for $\epsilon_1 > \epsilon_2$ or, equivalently,

$$(\epsilon_2^{-1}x + y)^{\uparrow} - (\epsilon_1^{-1}x + y)^{\uparrow} < (\epsilon_2^{-1} - \epsilon_1^{-1})x^{\uparrow} \ . \tag{3.5}$$

However, (3.5) follows from the relation

$$x_1^{\uparrow} - x_2^{\uparrow} \le (x_1 - x_2)^{\uparrow} \ . \quad ∎$$

We first establish pointwise convergence for z_ϵ in (3.2).

Theorem 9.3.1. (pointwise convergence) *For each* $x \in D, y \in D_{lim}$ *and* $t \geq 0$,

$$\lim_{\epsilon \downarrow 0} z_\epsilon(t) = z(t) \equiv \sup_{s \in \Phi_x^L(t)} y(s-) \vee \sup_{s \in \Phi_x^R(t)} \{y(s), y(s+)\} \ . \tag{3.6}$$

Proof. The convergence follows from the monotonicity established in Lemma 9.3.3. Lemma 9.3.2 above provides a lower bound, which implies that there is a proper limit for each t. For any $\delta > 0$, let $s_\epsilon(t)$ be a point in $[0, t]$ such that

$$(\epsilon^{-1}x + y)(s_\epsilon(t)) \geq (\epsilon^{-1}x + y)^{\uparrow}(t) - \delta \ . \tag{3.7}$$

(Since x and y need not be continuous, the supremum of $\epsilon^{-1}x + y$ need not be attained.) Then

$$\begin{aligned} y(s_\epsilon(t)) &\geq y(s_\epsilon(t)) + \epsilon^{-1}\{x[s_\epsilon(t)] - x^{\uparrow}(t)\} \\ &\geq y(s) + \epsilon^{-1}[x(s) - x^{\uparrow}(t)] - \delta \quad \text{for} \quad 0 \leq s \leq t \\ &\geq \begin{cases} y(s-) - \delta & \text{for} \quad s \in \Phi_x^L(t) \\ y(s) - \delta & \text{for} \quad s \in \Phi_x^R(t) \\ y(s+) - \delta & \text{for} \quad s \in \Phi_x^R(t) \ , \end{cases} \end{aligned} \tag{3.8}$$

implying that

$$\underline{\lim}_{n \to \infty} \ y(s_\epsilon(t)) \geq z(t), \quad t \geq 0 \ . \tag{3.9}$$

We now verify that

$$\overline{\lim}_{n \to \infty} \ y(s_\epsilon(t)) \leq z(t), \quad t \geq 0 \ . \tag{3.10}$$

Start by choosing $\{s_\epsilon(t)\}$ such that $y(s_\epsilon(t)) \to \overline{\lim}_{n \to \infty} y(s_\epsilon(t))$ as $\epsilon \downarrow 0$. Since $s_\epsilon(t) \in [0, t]$, any subsequence from $\{s_\epsilon(t)\}$ has a convergent subsequence $\{s_{\epsilon'}(t)\}$ as $\epsilon' \downarrow 0$. (Let $\epsilon' \downarrow 0$ through countably many values.) So suppose that $s_{\epsilon'}(t) \to s_0(t)$ as $\epsilon' \downarrow 0$. Without loss of generality, by taking a further subsequence if necessary, we can assume that either $s_{\epsilon'}(t) \uparrow s_0(t)$ with $s_{\epsilon'}(t) < s_0(t)$ for all $\epsilon' > 0$ or $s_{\epsilon'}(t) \downarrow s_0(t)$ with $s_{\epsilon'}(t) \geq s_0(t)$ for all $\epsilon' > 0$. Suppose that $s_{\epsilon'}(t) \uparrow s_0(t)$. Then $y(s_{\epsilon'}(t)) \to y(s_0(t)-)$. We can deduce from (3.8) that there is a constant K such that, for all ϵ',

$$-K \leq \epsilon^{-1}[x(s_{\epsilon'}(t)) - x^{\uparrow}(t)] \leq 0 \ , \tag{3.11}$$

implying that $x(s_{\epsilon'}(t)) \to x^{\uparrow}(t)$ as $\epsilon' \to 0$, so that $x(s_0(t)-) = x^{\uparrow}(t)$ and $s_0(t) \in \Phi_x^L(t)$. By this argument,

$$\overline{\lim_{n\to\infty}} \; y(s_\epsilon(t)) \leq \sup_{s\in\Phi_x^L(t)} y(s-) \; . \tag{3.12}$$

On the other hand, if $s_{\epsilon'}(t) \downarrow s_0(t)$, we can deduce by the same reasoning that

$$\overline{\lim_{n\to\infty}} \; y(s_\epsilon(t)) \leq \sup_{s\in\Phi_x^R(t)} \{y(s), y(s+)\} \; . \tag{3.13}$$

Since one of (3.12) or (3.13) must hold, we have established (3.10). Finally, from the first and last lines of (3.8),

$$0 \geq \epsilon^{-1}\{x(s_\epsilon(t)) - x^{\uparrow}(t)\} \geq z(t) - y(s_\epsilon(t)) \; . \tag{3.14}$$

Since $y(s_\epsilon(t)) \to z(t)$, $\epsilon^{-1}\{x(s_\epsilon(t)) - x^{\uparrow}(t)\} \to 0$ as $\epsilon \downarrow 0$, which implies that $z_\epsilon(t) \to z(t)$ as $\epsilon \downarrow 0$. ■

Corollary 9.3.2. (simplification under extra conditions) *Suppose that $x \in C$ and $y \in D_{lim}$. Then the limit z in* (3.6) *is*

$$z(t) = \sup_{s\in\Phi_x(t)} \{y(s-), y(s), y(s+)\} \, . \tag{3.15}$$

If, in addition, $y \in C$, then

$$z(t) = \sup_{s\in\Phi_x(t)} \{y(s)\} \; . \tag{3.16}$$

We now determine the structure of the limit function z in (3.6). Since $\Phi_x^L(t)$, $\Phi_x^R(t)$ and $\Phi_x(t)$ are subsets of $[0,t]$, we need a notion of convergence of sets. For subsets A_n and A of $\mathbb{R}$, we say that $A_n \to A$ if (i) for all $a_n \in A_n$, $n \geq 1$, $\{a_n\}$ has a convergent subsequence and the limits of all convergent subsequences belong to A, and (ii) for all $a \in A$, there exists $a_n \in A_n$, $n \geq 1$, such that $a_n \to a$ as $n \to \infty$. In our set limits involving $\Phi_x^L(t)$ and $\Phi_x^R(t)$, only three special cases arise: (i) A_n is independent of n for all sufficiently large n, (ii) the sequence $\{A_n\}$ is eventually monotone, i.e., either $A_n \subseteq A_{n+1}$ for all sufficiently large n or $A_n \supset A_{n+1}$ for all sufficiently large n, and (iii) $A = \{a\}$, i.e., the limit set contains a single point.

When we consider $\Phi_x(t) \equiv \Phi_x^L(t) \cup \Phi_x^R(t)$, we have compact subsets of $[0,t]$. Then the notion of set convergence above is induced by the Hausdorff metric on the space $\mathcal{C} \equiv \mathcal{C}([0,\infty))$ of compact subsets of $[0,\infty)$, defined in (2.8) in Chapter V.

However, even if x and $x^{\uparrow}$ are continuous in t, $\Phi_x(t)$ is in general *not* continuous in t. Moreover, at some time points, $\Phi_x(t)$ is neither left-continuous nor right-continuous.

Example 9.3.2. *Lack of continuity from left or right in $\Phi_x(t)$.* Suppose that $x(t) = (1-t)I_{[0,1)}(t) + (t-1)I_{[1,\infty)}(t)$. Then $\Phi_x(t) = \{0\}$, $0 \le t < 2$, $\Phi_x(2) = \{0,1\}$ and $\Phi_x(t-1) = \{t\}$, $t > 2$, so that Φ_x is neither left-continuous nor right-continuous at $t = 2$. However, $\Phi_x(2)$ is the union of the left and right limits $\Phi_x(2-)$ and $\Phi_x(2+)$. ∎

Example 9.3.3. *Neither left-continuous everywhere nor right-continuous everywhere.* We can extend Example 9.3.2 to show that the limit z need not be either a left-continuous function or a right-continuous function, even if x and y are both continuous. Let

$$x(t) = (1-t)I_{[0,1)}(t) + (t-1)I_{[1,3)}(t) + (5-t)I_{[3,4)}(t) + (t-3)I_{[4,\infty)}(t) \quad (3.17)$$

and

$$y(t) = -tI_{[0,2.5]} + 6(t-2.5)I_{[2.5,\infty)}(t)\,. \quad (3.18)$$

Then

$$\begin{aligned}
\Phi_x(t) &= \{0\}, \quad 0 \le t < 2, \quad \Phi_x(2) = \{0,2\} \\
\Phi_x(t) &= \{t\}, \quad 2 < t \le 3, \quad \Phi_x(t) = \{3\}, \quad 3 \le t < 5\,, \\
\Phi_x(5) &= \{3,5\}, \quad \Phi_x(t) = \{t\}, \quad t > 5\,, \\
z(2) &= 0 \quad \text{and} \quad z(5) = 15\,. \qquad (3.19)
\end{aligned}$$

Then z is discontinuous at $t = 2$ and $t = 5$, with z being left-continuous at 2 and right-continuous at 5. Hence z is neither left-continuous everywhere nor right-continuous everywhere. On the positive side, z is either left-continuous or right-continuous at each t and z is upper semicontinuous everywhere. ∎

Example 9.3.4. *Neither left-continuous nor right-continuous at one t.* We now show that the limit z in (3.8) need not be either left-continuous or right-continuous at a single argument t when $x \in C$ and $y \in D$ but $y \notin C$. We construct y and x so that y and Φ_x have only one common discontinuity. Let

$$y(t) = tI_{[0,1)}(t) + I_{[1,2)}(t), \quad t \ge 0\,, \quad (3.20)$$

and

$$x(t) = -tI_{[0,1)}(t) + (t-2)I_{[1,\infty)}(t), \quad t \ge 0\,, \quad (3.21)$$

so that

$$\Phi_x(t) = \{0\}, 0 \le t < 2, \Phi_x(2) = \{0, 2\} \quad \text{and} \quad \Phi_x(t) = t, \quad t > 2 \ . \tag{3.22}$$

Hence y and Φ_x are continuous everywhere except $t = 2$. Moreover,

$$z(2) = \sup_{s \in \{0,2\}} \{y(s)\} \vee \sup_{s \in \{2\}} \{y(s-)\} = 0 \vee 1 = 1 \ , \tag{3.23}$$

while $z(t) = 0$ for all other t. Hence the left and right limits coincide at $t = 2$ but do not equal $z(2)$, so that $z \notin D$. It is easy to see that $z_\epsilon(2) = 1$ and

$$z_\epsilon(t) = 0, \quad 0 \le t \le 2 - \epsilon \quad \text{and} \quad t \ge 2 + \epsilon \ ,$$

with z_ϵ defined by linear interpolation elsewhere. Hence, z_ϵ has slope ϵ^{-1} on $[2 - \epsilon, 2]$, slope $-\epsilon^{-1}$ on $[2, 2 + \epsilon]$ and is 0 elsewhere. Consistent with Theorem 9.3.1, z_ϵ converges pointwise to z. We will want to impose regularity conditions to prevent such pathological behavior. As an alternative, we could conclude that z_ϵ converges to a limit in one of the larger spaces E or F in Chapter X. ∎

We now introduce a regularity condition under which the limit z in (3.6) has left and right limits everywhere and is either left continuous or right continuous everywhere (without necessarily being right continuous everywhere). Let $D_{l,r}$ denote this space. We first define some subsets of $[0, \infty)$. (We could alternatively restrict attention to a subinterval $[0, T]$.) For any $x \in D$, let $Rinc(x)$ and $Linc(x)$ be the set of right-increase and left-increase points of x, let $Lconst(x)$ be the set of left-constant points of x, and let $Amax(x)$ be the argmax set of x, i.e., the set of arguments at which x equals its supremum, i.e.,

$$Rinc(x) \equiv \{t \ge 0 : x(t) < x(t + \epsilon) \quad \text{for all sufficiently small} \quad \epsilon > 0\} \tag{3.24}$$

$$Linc(x) \equiv \{t \ge 0 : x(t - \epsilon) < x(t) \quad \text{for all sufficiently small} \quad \epsilon > 0\} \tag{3.25}$$

$$Lconst(x) \equiv \{t \ge 0 : x(t - \epsilon) = x(t) \quad \text{for all sufficiently small} \quad \epsilon > 0\} \tag{3.26}$$

$$Amax(x) \equiv \{t \ge 0 : t \in \Phi_x^R(t)\} \ . \tag{3.27}$$

We will look at these sets for the functions x and $x^\uparrow$. Of course, $x^\uparrow$ is nondecreasing and right-continuous. Let $Disc(x)$ be the set of discontinuity points of x.

Theorem 9.3.2. (regularity properties of the limit z) *Suppose that x, $y \in D$. Then $z \in D_{lim}$, where z is the limit in* (3.6). *At all t not in the set*

$$Bad(x) \equiv Rinc(x^\uparrow) \cap Lconst(x^\uparrow) \cap Disc(x)^c \cap Linc(x) \cap Amax(x) \ , \tag{3.28}$$

z is either left-continuous or right-continuous. For $t \in Bad(x)$, $z(t+) = y(t)$, $z(t-)$ *is independent of* $\{y(t-), y(t)\}$ *and* $z(t) = \max\{z(t-), y(t-), y(t)\}$, *so that z is left-continuous at t if* $z(t-) \geq y(t-) \vee y(t)$, *right-continuous at t if* $y(t) \geq y(t-) \vee z(t-)$, *and neither left-continuous nor right-continuous if* $y(t-) > y(t) \vee z(t-)$. *If*

$$y(t-) \leq z(t-) \vee y(t) \tag{3.29}$$

for all $t \in Bad(x)$, *for which a sufficient condition is*

$$Disc(y) \cap Bad(x) = \phi \ , \tag{3.30}$$

then z is either left-continuous or right-continuous at all t, so that $z \in D_{l,r}$.

Corollary 9.3.3. (regularity for continuous y) *If* $x \in D$ *and* $y \in C$, *then* $z \in D_{l,r}$.

Remark 9.3.1. *Sufficient condition for having more than one point in the set.* Let $|\Phi_x(t)|$ be the cardinality of the set $\Phi_x(t)$. Note that $|\Phi_x(t)| \geq 2$ when $t \in Lconst(x^{\uparrow}) \cap Amax(x)$, i.e.,

$$Lconst(x^{\uparrow}) \cap Amax(x) \subseteq \{t : |\Phi_x(t)| \geq 2\} \ , \tag{3.31}$$

so that $t \in Bad(x)$ when $|\Phi_x(t)| \geq 2$ and $x(t-\epsilon) < x(t-) = x(t) = x^{\uparrow}(t) < x^{\uparrow}(t+\epsilon)$ for all suitably small $\epsilon > 0$. ∎

Remark 9.3.2. *The set* $Bad(x)$ *is at most countably infinite.* From (3.28), it follows that $Bad(x) \subseteq Disc(\Phi_x)$, where $\Phi_x \in D([0,\infty),(\mathcal{C},h))$. Therefore, $Bad(x)$ is a countable set. ∎

Corollary 9.3.4. (regularity properties of the limit Z when Y is a stochastic process) *Suppose that* $\{Y(t) : t \geq 0\}$ *is a stochastic process with sample paths in D. If* $x \in D$ *and if* $P(t \in Disc(Y)) = 0$ *for each* $t > 0$, *then* $P(Z \in D_{l,r}) = 1$, *where Z is the limiting stochastic process defined by applying* (3.6) *to* Y.

Proof. In Remark 9.3.2 it was noted that the set $Bad(x)$ in (3.28) is countable. Consequently,

$$P(Z \in D_{l,r}) = P(Disc(Y) \cap Bad(x) = \phi) = 1 \ . \ \blacksquare \tag{3.32}$$

Theorem 9.3.2 is proved by examining all relevant cases. We identify appropriate cases and results for those cases in the following theorem.

Theorem 9.3.3. (identification of relevant cases) *The following is a set of exhaustive and mutually exclusive cases and subcases when $x, y \in D$:*

1. $t \notin Amax(x)$, *i.e.,* $t \notin \Phi_x^R(t)$: *z is right-continuous with a left limit at t.*

2. $\Phi_x^R(t) = \Phi_x^L(t) = \{t\} : z(t) = y(t-) \vee y(t)$, *$z$ is either right-continuous or left-continuous at t.*

3. $\Phi_x^R(t) = \{t\}$, $\Phi_x^L(t) = \phi$: *z is right-continuous with a left limit at t.*

4. $t \in \Phi_x^R(t) \subseteq \Phi_x(t) \neq \{t\}$, *so that cases 1–3 do not hold;*

   (a) $t \notin Rinc(x^{\uparrow})$, *i.e.,* $\Phi_x(t) \subseteq \Phi_x(u)$ *for some* $u > t$: *z is right-continuous with a left limit at t.*

   (b) *Condition* (a) *does not hold and* $t \in Lconst(x^{\uparrow}) \cap Linc(x)^c$, *i.e., t is not isolated in* $\Phi_x(t)$: $z(t-) \geq y(t-)$ *and* $z(t+) = y(t)$, *so that z is left (right) continuous at t if* $z(t-) \geq (\leq)\ y(t)$.

   (c) *Condition* (a) *does not hold, t is isolated in* $\Phi_x(t)$ *and* $t \in Disc(x)$: $z(t+) = y(t)$ *and* $z(t) = \max\{z(t-), y(t)\}$, *so that z is left (right) continuous if* $z(t-) \geq (\leq)\ y(t)$. *(In this case $z(t)$ does not depend upon $y(t-)$.)*

   (d) *Condition* (a) *does not hold, t is isolated in* $\Phi_x(t)$ *and* $t \notin Disc(x)$, *i.e.,* $t \in Bad(x)$ *in* (3.28): $z(t+) = y(t)$, $z(t-)$ *is independent of* $\{y(t-), y(t)\}$ *and* $z(t) = \max\{z(t-), y(t-), y(t)\}$. *Hence z is neither left-continuous nor right-continuous at t if and only if* $y(t-) > z(t-) \vee y(t)$.

Proof. We prove Theorem 9.3.3 by examining all relevant subcases. We provide a further characterization below, but do not give all details. For this purpose, let

$$\Psi_x^L(t) = \{s : 0 < s \leq t, x(s-) = x^{\uparrow}(t-)\}\ , \quad (3.33)$$

$$\Psi_x^R(t) = \{s : 0 \leq s \leq t, x(s) = x^{\uparrow}(t-)\} \quad (3.34)$$

and $\Psi_x(t) = \Psi_x^L(t) \cup \Psi_x^R(t)$.

Case 1: In this case, $x(t) < x^{\uparrow}(t)$ and $x^{\uparrow}(t-) = x^{\uparrow}(t)$. Since x and $x^{\uparrow}$ are right-continuous, Φ_x^L and Φ_x^R are constant in $[t, t+\epsilon)$ for all suitably small $\epsilon > 0$, so that z is necessarily right-continuous. We identify three subcases:

(i) If $t \notin \Psi_x^L(t)$, then $x(t-) < x^{\uparrow}(t-)$, so that Φ_x^L and Φ_x^R are constant in $(t-\epsilon, t+\epsilon)$ for all suitably small $\epsilon > 0$, so that z is constant in the same subinterval. (ii) If $t \in \Psi_x^L(t) = \Phi_x^L(t)$ and $\Phi_x(t) \neq \{t\}$, then x jumps down at time t, so that $x(t-) = x^{\uparrow}(t-) = x^{\uparrow}(t) > x(t)$. Since $\Phi_x(t) \neq \{t\}$, $x^{\uparrow}$ must be constant in $(t-\epsilon, t]$ for all suitably small $\epsilon > 0$ and there must exist $s < t$ such that $x(s) = x^{\uparrow}(t)$ or $x(s-) = x^{\uparrow}(t)$. Hence for $s < t' < t$, $\Phi_x^L(t')$ and $\Phi_x^R(t')$ increase as t' increases. Since $t \notin \Phi_x^R(t)$, $\Phi_x^R(t') \uparrow \Phi_x^R(t)$ as $t' \uparrow t$, so that Φ_x^R is continuous at t. Since $\Phi_x^L(t')$ increases as t' increases, $\Phi_x^L(t')$ has a limit as $t' \uparrow t$, but this limit set may be separated from $t \in \Phi_x^L(t)$. Hence, in general z is right-continuous with a left limit at t, with $z(t)$ depending upon $y(t-)$ but not $y(t)$. In this case z is continuous at t if and only if $z(t-) \geq y(t-)$. (iii) If $t \in \Psi_x^L(t) = \Phi_x^L(t)$ and $\Phi_x(t) = \{t\}$, then again x jumps down at time t, $x(t-) = x^{\uparrow}(t-)$. Since $x^{\uparrow}$ is increasing from the left at t, there exists a sequence $\{t_n\}$ with $t_n \uparrow t$ as $n \to \infty$ such that $x(t_n\pm) = x^{\uparrow}(t_n)$ and $\Phi_x(t_n) = \{t_n\}$. Moreover, for any s with $t_n < s < t$, necessarily $\Phi_x(s) \subseteq [t_n, s]$. Hence, $\Phi_x(s) \to \Phi_x(t)$ as $s \uparrow t$. This implies that z is continuous at t with $z(t) = y(t-)$. We remark that the case $t \in \Psi_x^L(t)$ but $t \notin \Phi_x^L(t)$ cannot occur because it requires $x(t-) = x^{\uparrow}(t-) < x^{\uparrow}(t)$, which implies that x make a jump up to a new maximum at time t, i.e., $t \in \Phi_x^R(t)$, which contradicts our original assumption.

Case 2: $\Phi_x^R(t) = \Phi_x^L(t) = \{t\}$.

In this case $x(t-) = x(t) = x^{\uparrow}(t)$, so that x is continuous at t. Since $\Phi_x(t) = \{t\}$, $\Phi_x(u) \subseteq [t, u]$ for all $u > t$. Hence $\Phi_x(u) \to \Phi_x(t) = \{t\}$ as $u \downarrow t$, so that Φ_x is right-continuous and z has a limit from the right with $z(t+) = y(t)$. In this case $x^{\uparrow}$ is increasing at t, and $\Phi_x(s) \to \Phi_x^L(t)$ as $s \uparrow t$, so that Φ_x is continuous at t and z has the left limit $z(t-) = y(t-)$. Since $z(t) = y(t) \vee y(t-)$, z is either left-continuous or right-continuous at t; z is continuous at time t if and only if y is.

Case 3: $\Phi_x^R(t) = \{t\}$ and $\Phi_x^L(t) = \phi$.

In this case $x(t-) \neq x(t) = x^{\uparrow}(t)$, so that x is discontinuous at t. As in case 2 above, $\Phi_x(s) \to \Phi_x(t) = \{t\}$ as $s \downarrow t$, so that Φ_x is right-continuous at t and z has the right limit $z(t+) = y(t)$. Since $z(t) = y(t)$, z is right-continuous in this case. We identify three subcases: (i) If $t \notin \Psi_x^L(t)$, then $x(t-) < x^{\uparrow}(t-) < x^{\uparrow}(t)$, so that x jumps up to a new maximum at time t and Φ_x^L and Φ_x^R are constant in $(t-\epsilon, t)$ for all suitably small ϵ. Hence Φ_x^L, Φ_x^R and z have limits from the left, but may be discontinuous at t. (ii) If $\Psi_x(t) = \{t\}$, then $x(t-) = x^{\uparrow}(t-) < x^{\uparrow}(t)$. As in (ii), x jumps up to a new maximum at t. Since $\Phi_x(t) = \{t\}$, $x^\uparrow$ is increasing from the left at t. Hence, there exists a sequence $\{t_n\}$ with $t_n \uparrow t$ as $n \to \infty$ such that $x(t_n\pm) = x^\uparrow(t_n) \uparrow x^\uparrow(t-)$ and $\Phi_x(t_n) = \{t_n\}$. Hence $\Phi_x(s) \subseteq [t_n, s]$ for all s with $t_n < s < t$. Hence, $\Phi_x(s) \to \Psi_x^L(t) = \{t\}$ as $s \uparrow t$, so that Φ_x and z have limits from the left at t, with $z(t-) = y(t-)$. (iii) Suppose that $\Phi_x^L(t) = \phi$ and $t \in \Psi_x(t) \neq \{t\}$. This is similar to case (ii). Since $\Psi_x(t) \neq \{t\}$, $x^\uparrow$ is constant in $[t-\epsilon, t)$ for all suitably small ϵ. Thus, over $(t-\epsilon, t)$, $\Phi_x^L(s)$ and $\Phi_x^R(s)$ increase to $\Psi_x^L(t)$ and $\Psi_x^R(t)$ as $s \uparrow t$. Hence, z has a left limit at t. In general, z need not be continuous at t.

Case 4(a): In this case $x^\uparrow(t) = x^\uparrow(u)$ for some $u > t$. Hence $\Phi_x^L(u) \downarrow \Phi_x^L(t)$ and $\Phi_x^R(u) \downarrow \Phi_x^R(t)$ as $u \downarrow t$ so that z is right-continuous at t. If t is not isolated in $\Phi_x(t)$, as in Case 4(b), then there exists $t_n \uparrow t$ with $x(t_n-) = x^\uparrow(t)$ or $x(t_n) = x^\uparrow(t)$, so that $x^\uparrow$ is constant in $[t-\epsilon, t]$ for all suitably small ϵ. Moreover, $\Phi_x^L(s) \uparrow \Phi_x^L(t)$ and $\Phi_x^R(s) \uparrow \Phi_x^R(t)$ as $s \uparrow t$. Hence z has a left limit $z(t-) \geq y(t-)$. Moreover, Φ_x^L and Φ_x^R are continuous at t. If $y(t-) \leq z(t-) < y(t)$, then y is right-continuous but not continuous. On the other hand, if $z(t-) \geq y(t)$, then z is continuous at t. If instead t is isolated in $\Phi_x(t)$, as in Case 4(c), then $\Phi_x^L(s)$ and $\Phi_x^R(s)$ are constant in $(t-\epsilon, t)$ for all suitably small ϵ, but $\Phi_x^R(t) = \Phi_x^R(t-) \cup \{t\}$. Hence, Φ_x^L and Φ_x^R have limits from the left at t. Thus z has a limit from the left at t, which does not depend on $y(t-)$. If $z(t-) < y(t)$, then z is discontinuous at t; otherwise it is continuous.

Case 4(b): As in case 4(a), z has a left limit at t. If Case 4(a) does not hold, then $x^\uparrow(t) < x^\uparrow(t+\epsilon)$ for all sufficiently small ϵ. In this case, $\Phi_x(s) \to \{t\}$ as $s \downarrow t$, so that Φ_x and z have limits from the right with $z(t+) = y(t)$. However, since $\Phi_x(t) \neq \{t\}$ by assumption, Φ_x is not right-continuous. In this case z is left (right) continuous if $z(t-) \geq (\leq)\ y(t)$.

Case 4(c): In this case

$$t \in OK(x) \equiv Rinc(x^\uparrow) \cap Lconst(x^\uparrow) \cap Disc(x) \cap Linc(x) \cap Amax(x)\ . \quad (3.35)$$

Note that $OK(x)$ in (3.35) differs from $Bad(x)$ in (3.28) only by having $x(t-) < x(t)$. As noted for case 4(a) and 4(b), z has left limit $z(t-)$ and right limit $z(t+) = y(t)$ at t, with $z(t) = z(t-) \vee y(t)$. However, since $x(t-) < x(t) = x^\uparrow(t)$, $t \notin \Phi_x^L(t)$, so that $z(t)$ does not depend upon $y(t-)$. Hence z is either left-continuous or right-continuous at t.

Case 4(d): In this case $t \in Bad(x)$. Since $x(t-) = x(t) = x^{\uparrow}(t)$, $t \in \Phi_x^L(t)$ and $z(t) \geq y(t-)$. As in Case 4(c), z has left and right limits at t with $z(t+) = y(t)$ and $z(t) = \max\{z(t-), y(t-), y(t)\}$. ∎

Theorem 9.3.2 concluded that $z \in D_{lim}$ when $x, y \in D$. By the same reasoning, examining the cases in Theorem 9.3.3, we can obtain the same conclusion when $y \in D_{lim}$.

Theorem 9.3.4. (extension when $y \in D_{lim}$) *Suppose that $x \in D$ and $y \in D_{lim}$. Then $z \in D_{lim}$. At all t not in the set*

$$Bad(x,y) = [Bad_1(x) \cap Disc(y)] \cup Bad_2(y) \ , \tag{3.36}$$

where

$$Bad_1(x) \equiv Rinc(x^{\uparrow}) \cap Lconst(x^{\uparrow}) \cap Linc(x) \cap Amax(x) \tag{3.37}$$

and

$$Bad_2(y) \equiv \{t \in [0,T] : y(t) > y(t-), y(t+)\} \ , \tag{3.38}$$

z is either left-continuous or right-continuous. At $t \in Bad(x) \cap Disc(x)$, $z(t+) = y(t+)$, $z(t-)$ is independent of $y(t-)$ and $z(t) = z(t-) \vee y(t) \vee y(t+)$, so that z is left-continuous if $z(t-) \geq y(t) \vee y(t+)$, right-continuous if $y(t+) \geq y(t) \vee z(t-)$ and neither right-continuous nor left-continuous if $y(t) > z(t-) \vee y(t+)$. At $t \in Bad(x) \cap Disc(x)^c$, $z(t+) = y(t+)$, $z(t-)$ is independent of $y(t-)$ and $z(t) = z(t-) \vee y(t-) \vee y(t) \vee y(t+)$, so that z is left-continuous if $z(t-) \geq y(t-) \vee y(t) \vee y(t+)$, right-continuous if $y(t+) \geq z(t-) \vee y(t-) \vee y(t)$ and neither left-continuous nor right-continuous if $y(t-) \vee y(t) > z(t-) \vee y(t+)$.

We get extra regularity conditions if we assume that $x \in C$. Recall that z is upper semicontinuous at t if $\lim_{s \to t} z(s) \leq z(t)$; z is upper semicontinuous if it is upper semicontinuous at all t. Let D_{usc} be the subset of upper semicontinuous functions in D_{lim}.

Theorem 9.3.5. (upper-semicontinuity when $x \in C$) *Suppose that $x \in C$ and $y \in D_{lim}$, then $z \in D_{usc}$. Then $\Phi_x^L(t) = \Phi_x^R(t) = \Phi_x(t)$ for all $t > 0$ and*

$$z(t) = \sup_{s \in \Phi_x(t)} \{y(s-) \vee y(s) \vee y(s+)\} \ . \tag{3.39}$$

Proof. Since $x \in C$, the only relevant cases in Theorem 9.3.3 are: $1(i)$, 2 and 4. Formula (3.39) follows directly from formula (3.6). The upper semicontinuity follows from by considering the cases in Theorem 9.3.3. ∎

Remark 9.3.3. *The need for x to be continuous.* Without assuming that $x \in C$, we need not have z be upper semicontinuous. In Case 3 of Theorem 9.3.3, we can have $z(t-) > z(t) = z(t+) = y(t+)$. ∎

From the point of view of applications, the two most common cases are

$$\begin{aligned} &\text{(i)} \quad x \in C \quad \text{and} \quad y \in C \\ &\text{(ii)} \quad x \in C \quad \text{and} \quad y \in D\,. \end{aligned} \tag{3.40}$$

We thus summarize the situation in these two important cases.

First, with case (i) in (3.40) when both $x \in C$ and $y \in C$, we can apply Corollary 9.3.3 and Theorem 9.3.5 above to conclude that $z \in D_{l,r} \cap D_{usc}$, but Example 9.3.3 shows that we need not have $z \in D$. Indeed, we will always have $z \in D_{l,r} \cap D_{usc}$ instead. For $x \in D$, we have $x \in D_{usc}$ only if $x(t) \geq x(t-)$ for all t. So it is important to have the space $D_{l,r} \cap D_{usc}$.

Second, with Case (ii) in (3.40) when $x \in C$ but only $y \in D$, Theorem 9.3.2 shows that $z \in D_{lim}$, but Example 9.3.4 shows that we need not have $z \in D_{l,r}$ in general. However, under condition (3.29), which is implied by condition (3.30), Theorem 9.3.2 implies that we do have $z \in D_{l,r}$. Moreover Theorem 9.3.5 shows that $z \in D_{usc}$. So, in Case (ii) we should also have $z \in D_{l,r} \cap D_{usc}$, but we need to impose condition (3.30).

Because we assumed only that $y \in D_{lim}$ in Theorem 9.3.1, we can consider z playing the role of y. For example, we could start by considering $z_\epsilon(x_1, y)$ in (3.2) for some $x_1 \in D$ and obtain $z_1 = z(x, y)$ as $\epsilon \downarrow 0$. Then we could consider $z_\epsilon(x_2, z_1)$ in (3.2) for another $x_2 \in D$ and obtain $z_2 = z(x_2, z_1)$ as $\epsilon \downarrow 0$.

9.4. Extending Pointwise Convergence to M_1 Convergence

We now want to extend the pointwise convergence of z_ϵ to z as $\epsilon \downarrow 0$ in Theorem 9.3.1 to M_1 convergence. We first observe that monotone pointwise convergence of continuous functions in D does not by itself imply M_1 convergence.

Example 9.4.1. *Monotone pointwise convergence of continuous functions does not imply M_1 convergence.* To see that monotone pointwise convergence of continuous functions does not imply M_1 convergence in $D([0,2],\mathbb{R})$, let

$$x_{2^{-n}}(0) = x_{2^{-n}}(1-2^{-n}) = x_{2^{-n}}(1-2^{-(n+1)}) = 0$$

$$x_{2^{-n}}(1-3(2^{-(n+2)})) = x_{2^{-n}}(1-2^{-(n+1)}+2^{-(2n+1)}) = x_{2^{-n}}(2) = 1$$

for $n \geq 1$, with $x_{2^{-n}}$ defined by linear interpolation elsewhere. Clearly $x_{2^{-n}}$ is continuous for each n. Let $x_\epsilon = x_{2^{-n}}$ for $2^{-n} \geq \epsilon > 2^{-(n+1)}$, $n \geq 1$. It is easy to see that $x_{2^{-n}}(t) \geq x_{2^{-(n+1)}}(t) \downarrow x(t)$ as $n \to \infty$ for each $t \geq 0$, so that $x_\epsilon(t) \downarrow x(t)$ as $\epsilon \downarrow 0$ for each $t \geq 0$. Moreover $x_\epsilon \to x$ in D as $\epsilon \downarrow 0$ with the M_2 topology, but not in the M_1 topology, because, for any $\delta > 0$, x_ϵ crosses the strip $(1/3, 2/3)$ for t in $[1-\delta, 1+\delta]$ three times for all sufficiently small ϵ, whereas x crosses it only once; see Theorem 12.5.1 (v) in the book. ∎

In general (without continuity conditions) monotone pointwise convergence does not imply even M_2 convergence.

Example 9.4.2. *Monotone pointwise convergence without continuity does not imply M_2 convergence.* To see that M_2 convergence does not follow from monotone pointwise convergence in or $D_{l,r}$ when neither the limit nor the converging functions need be continuous, let $x = I_{[1,2]}$ and $x_n = 2I_{[1-n^{-1},1)} + I_{[1,2]}$, $n \geq 1$. ∎

However, we can obtain a positive result when the converging functions are continuous (without relying on the special structure associated with the supremum).

Theorem 9.4.1. (M_2 convergence from monotone pointwise convergence of continuous functions) *If $x \in D_{l,r}$, $x_\epsilon \in C$ for all ϵ and $x_\epsilon(t) \downarrow x(t)$ as $\epsilon \downarrow 0$ for all $t \geq 0$, then $x_\epsilon \to x$ in $(D_{l,r}, M_2)$ as $\epsilon \downarrow 0$.*

We can combine Theorems 9.3.1 and 9.4.1 above to obtain the following corollary.

Corollary 9.4.1. (M_2 convergence of the supremum derivative) *In the setting of Theorem 9.3.1, if x and y are both continuous, then $z_\epsilon \to z$ in $(D_{l,r}, M_2)$ as $\epsilon \downarrow 0$.*

However, by exploiting the special structure of the supremum function, we will actually establish the stronger M_1 convergence under weaker conditions. To prove Theorem 9.4.1, we exploit approximations by piecewise-constant functions see Section 12.2 in the book.

Proof of Theorem 9.4.1. Since the pointwise convergence is monotone, $x_\epsilon(t) \ge x(t)$ for all t and ϵ. For any u and $\delta > 0$, let $\tilde{x}$ be a piecewise-constant function in D with $\|x - \tilde{x}\|_u < \delta$. Then $x(t) \le \tilde{x}(t) + \delta$ for $0 \le t \le u$. Let $\hat{x}$ be the upper boundary (containing only vertical and horizontal pieces) of the δ neighborhood of the completed graph $\Gamma_{\tilde{x}+\delta}$ of $\tilde{x} + \delta$ for the time set $[0, t]$, using the Hausdorff metric, as depicted in Figure 9.1. Note that $\hat{x}(s) \ge x(s)$ for $0 \le s \le t$ and $h_t(\Gamma_x, \Gamma_{\hat{x}}) \le 3\delta$, where h_t is the Hausdorff metric applied to the graphs with time set $[0, t]$. It thus suffices to show that $x_\epsilon(s) \le \hat{x}(s)$ for all s, $0 \le s \le t$, for all sufficiently small ϵ.

Consequently, it suffices to show that $x_\epsilon(s) \vee \hat{x}(s)$ converges uniformly to $\hat{x}(s)$ for $0 \le s \le t$ as $\epsilon \downarrow 0$. However, $\hat{x}$ has only finitely many discontinuities. Since $x_\epsilon \vee \hat{x}$ is continuous and nonincreasing in ϵ, we can apply Dini's theorem to get uniform convergence in any compact subset of $[0, t]$ excluding arbitrarily small open neighborhoods of each of the finitely many discontinuities. To treat the discontinuities, we need to carefully treat the neighborhood to the left (right) of a jump up (down). On the other side, the limit function constrains $x_\epsilon(s) \vee \hat{x}(s)$ as $\epsilon \downarrow 0$. Now suppose that t is one of the finitely many discontinuities of $\hat{x}$. Then there is $\epsilon_0(t)$ such that $|x_\epsilon(t) - x(t)| < \delta/2$ for all $\epsilon < \epsilon_0(t)$ by the pointwise convergence. Let ϵ_0 be the minimum of the finitely many $\epsilon_0(t)$. For any $\epsilon \le \epsilon_0$ given, the continuity of x_{ϵ_0} implies that, for each discontinuity point t, there is an $\eta(t) \equiv \eta(t, \epsilon) > 0$ such that $|x_\epsilon(t) - x_\epsilon(s)| < \delta/2$ for all s with $|s - t| < \eta(t)$. Thus, $|x_\epsilon(s) - x(t)| < \delta$ for $|s - t| < \eta(t)$. On the critical side of each discontinuity, the monotonicity implies that

$$x_{\epsilon'}(s) \le x_\epsilon(s) \le x_\epsilon(t) + \delta/2$$

for all $\epsilon' \le \epsilon$. Let the open neighborhood about t be $(t - \eta(t)/4, t + \eta(t)/4)$. Outside the finite union of those open intervals, we have the uniform convergence; inside those intervals we have established that $x_\epsilon(s) \vee \hat{x}(s) < \hat{x}(s) + \delta$. Hence Γ_{x_ϵ} is contained in the 4δ-neighborhood of Γ_x for suitably small ϵ, which implies the M_2 convergence. ∎

We will want to approximate $y \in D$ by $y \in D_c$. For this purpose, it is important to understand how z_ϵ and z in (3.2) and (3.6) depend upon y.

Lemma 9.4.1. (uniform Lipschitz property of z_ϵ as a function of y) *For any $\epsilon > 0$, $t > 0$, $x \in D$ and y_1, $y_2 \in D$,*

$$\|z_\epsilon(x, y_1) - z_\epsilon(x, y_2)\|_t \le \|y_1 - y_2\|_t \qquad (4.1)$$

and

$$\|z(x, y_1) - z(x, y_2)\|_t \le \|y_1 - y_2\|_t \ . \qquad (4.2)$$

Proof. Property (4.2) follows immediately from (3.6). For (4.1), note that

$$\begin{aligned} \|z_\epsilon(x,y_1) - z_\epsilon(x,y_2)\|_t &= \epsilon^{-1}\|(x+\epsilon y_1)^\uparrow - (x+\epsilon y_2)^\uparrow\|_t \\ &\leq \epsilon^{-1}\|(x+\epsilon y_1) - (x+\epsilon y_2)\|_t \\ &= \|y_1 - y_2\|_t \ . \ \blacksquare \end{aligned}$$

We also employ the following elementary, but useful, lemma.

Lemma 9.4.2. ($z \in D_c$ when $y \in D_c$) *Suppose that x, $y \in D$. If, in addition, $y \in D_c$ and*

$$Disc(y) \cap Bad(x) = \phi \tag{4.3}$$

for $Bad(x)$ in (3.28), *then $z \in D_c$. If y has k discontinuity points in $(0,t)$, then z has at most k discontinuity points in $[0,t]$.*

Proof. We use Theorem 9.3.2 to show that $z \in D$. Since $y \in D_c$, for any given interval $[0,t]$, there are time points $t_0 = 0 < t_1 < \cdots < t_k = t$ such that y is constant on $[t_{j-1}, t_j)$ and $[t_{k-1}, t]$ for $1 \leq j \leq k$. Note that $z(t) = y(0)$ for $t \in [0, t_1)$. From (3.6), it is obvious that z can only assume one of the k values $y(t_{j-1})$, $1 \leq j \leq k$. The function z may change to $y(t_{j-1})$ in the interval $[t_{j-1}, t_j)$, but it can only do so once. Transitions from $z(t_{j-1}-) < y(t_{j-2})$ to $y(t_{j-2}) < y(t_{j-1})$ to $y(t_{j-1})$ at t_{j-1} are ruled out by condition (4.3). ■

Theorem 9.4.2. (M_1 convergence of the supremum derivative) *Suppose that $x, y \in D$ and* (4.3) *holds for $Bad(x)$ in* (3.28). *Then*

$$z_\epsilon \to z \quad in \quad (D_{lr}, M_1) \quad as \quad \epsilon \downarrow 0$$

for z_ϵ in (3.2) *and z in* (3.6).

Proof. Lemmas 9.4.1 and 9.4.2 imply that it suffices to consider $y \in D_c$ in order to establish the M_1 convergence. By Theorem 9.3.2 and Example 9.3.2, the discontinuity condition (4.3) is necessary and sufficient to have $z \in D$. Under condition (4.3), it is possible to choose the piecewise-constant approximation to y so that it too satisfies (4.3). So, henceforth, assume that $y \in D_c$ and satisfies (4.3). By Lemma 9.4.2, $z \in D_c$ as well. Now, by applying mathematical induction over the successive discontinuities of z, it is not difficult to show that, for all sufficiently small $\epsilon > 0$, $z_\epsilon(t) = z(t)$ for all t outside a union of open neighborhoods of the discontinuities of z. (We strongly exploit D_c at this step.) For given discontinuities of y and z, by making ϵ suitably small, these neighborhoods can be chosen to be disjoint with the property that z_ϵ is monotone on each interval. The monotonicity together with the pointwise convergence established in Theorem 9.3.1 implies the local characterization of M_1 convergence in Theorem 12.5.1 in the book. ■

Example 9.4.3. *The need for M_1 convergence.* It is possible to have $z_\epsilon = z$ at a discontinuity point of z: For $x(t) = 0$, $t \ge 0$, $z_\epsilon(t) = z(t) = y^\uparrow(t)$ for all $t \ge 0$. Then z_ϵ and z have the discontinuities of $y^\uparrow$. A typical case requiring the M_1 convergence is $y = I_{[1,\infty)}$ and $x(t) = -tI_{[0,1)}(t) + (t-2)I_{[1,\infty)}(t)$. Then

$$z_\epsilon(t) = \epsilon^{-1}(2 - t + \epsilon)I_{[2-\epsilon,2)}(t) + I_{[2,\infty)}(t) \to z(t) = I_{[2,\infty)}(t) \quad \text{in} \quad (D, M_1) \ .$$

Finally, we can combine Theorems 9.2.3, 9.4.2 and the triangle inequality (2.1) to obtain a preservation-of-convergence result for the supremum function.

Theorem 9.4.3. (convergence preservation for the supremum map with nonlinear centering) *For $\epsilon > 0$, let x_ϵ, $y \in D$ and let x be a Lipschitz function in C. If*

$$d_{M_1}(x_\epsilon - x, \epsilon y) = o(\epsilon) \quad as \quad \epsilon \downarrow 0 \ , \tag{4.4}$$

for which a sufficient condition is

$$\|\|\epsilon^{-1}(x_\epsilon - x) - y\|_t \to 0 \quad as \quad \epsilon \downarrow 0 \quad for\ all \quad t > 0 \ , \tag{4.5}$$

and if (4.3) holds for $Bad(x)$ in (3.28), then

$$\epsilon^{-1}(x_\epsilon^\uparrow - x^\uparrow) \to z \quad in \quad (D_{l,r}, M_1) \quad as \quad \epsilon \downarrow 0 \tag{4.6}$$

for z in (3.6).

Corollary 9.4.2. (convergence preservation starting with the standard initial limit (4.5)) *For $\epsilon > 0$, let $x_\epsilon \in D$ and $x, y \in C$ with x being Lipschitz. If (4.5) holds, then (4.6) holds for z in* (3.39) *and $z \in D_{usc} \cap D_{l,r}$.*

9.5. Derivative of the Reflection Map

Now we consider the reflection map $\phi : D \to D$ defined by

$$\phi(x) \equiv x + (-x \vee 0)^\uparrow \ ; \tag{5.1}$$

see Section 13.4 in the book.

Results for the reflection map ϕ in (5.1) above follow from the results for the supremum map in Sections 9.3 and 9.4 above, because

$$\phi_\epsilon(x,y) \equiv \epsilon^{-1}[\phi(x+\epsilon y) - \phi(x)] = y + m_\epsilon(-x,-y) \ , \tag{5.2}$$

where

$$m_\epsilon(x,y) = \epsilon^{-1}[((x+\epsilon y)^\uparrow \vee 0 - (x^\uparrow \vee 0)] \ . \tag{5.3}$$

Note that $m_\epsilon(x,y)$ in (5.3) differs from $z_\epsilon(x,y)$ in (3.2) only by the extra maximum with respect to 0. In most applications, we will have $x(0) = y(0) = 0$, in which case the extra maximum $\vee 0$ is superfluous; then $m_\epsilon(x,y) = z_\epsilon(x,y)$. Thus, in this common case we can immediately apply the results in Section 9.3 to obtain corresponding results for the reflection map.

Theorem 9.5.1. (derivative of the reflection map in the common case) *Suppose that $x \in D$, $y \in D$ and $x(0) = y(0) = 0$. Then, for each $t > 0$,*

$$\lim_{\epsilon \downarrow 0} \phi_\epsilon(x,y)(t) = \dot{\phi}(t) \ , \tag{5.4}$$

where

$$\begin{aligned} \dot{\phi}(t) &\equiv \dot{\phi}(x,y)(t) \\ &\equiv y(t) - \left(\inf_{s \in \Phi^L_{-x}(t)} \{y(s-)\} \wedge \inf_{s \in \Phi^R_{-x}(t)} \{y(s)\} \right) \end{aligned} \tag{5.5}$$

and $\dot{\phi} \in D_{lim}$. If, in addition,

$$Disc(y) \cap Bad(-x) = \phi \ , \tag{5.6}$$

then $\dot{\phi} \in D_{l,r}$ and

$$\phi_\epsilon(x,y) \to \dot{\phi}(x,y) \quad in \quad (D_{l,r}, M_1) \quad as \quad \epsilon \downarrow 0 \, . \tag{5.7}$$

If, in addition, $x \in C$, then $\dot{\phi} \in D_{usc}$. If, in addition, x is Lipschitz and $y \in C$, then there is convergence preservation: If

$$\|\epsilon^{-1}(x_\epsilon - x) - y\|_t \to 0 \quad as \quad \epsilon \downarrow 0 \quad for\ all \quad t \tag{5.8}$$

then

$$\epsilon^{-1}(\phi(x_\epsilon) - \phi(x)) \to \dot{\phi}(x,y) \quad in \quad (D_{l,r}, M_1) \quad as \quad \epsilon \downarrow 0 \, . \tag{5.9}$$

for

$$\dot{\phi}(t) = y(t) - \inf_{s \in \Phi_{-x}(t)} \{y(s)\} \, . \tag{5.10}$$

where

$$\Phi_{-x}(t) = \{s : 0 \le s \le t, x(s) = x^\downarrow(t)\}, \quad t \ge 0 \, . \tag{5.11}$$

Proof. The pointwise limit in (5.4) follows from Theorem 9.3.1, noting that $-(-y)^{\uparrow} = y^{\downarrow}$. The fact that $\dot{\phi} \in D_{lim}$ follows from Theorem 9.3.2. The stronger conclusion that $\dot{\phi} \in D_{l,r}$ under condition (5.6) also follows from Theorem 9.3.2, exploiting condition (3.30). The M_1 convergence in (5.7) follows from Theorem 9.4.2. Finally, the convergence preservation ((5.8) implies (5.9)) follows from Corollary 9.4.3. ∎

We now return to the general case. For that purpose, let

$$t_l \equiv t_l(x) \equiv \inf\{t > 0 : x^{\uparrow}(t) = 0\} \tag{5.12}$$

and

$$t_u = t_u(x) \equiv \sup\{t > 0 : x^{\uparrow}(t) = 0\} \ , \tag{5.13}$$

with $t_l = t_u = \infty$ if $x^{\uparrow}(t) < 0$ for all t. In many applications we will have $x(0) = 0$; then $t_l = 0$ and $t_u = \infty$. It is easy to see that for any t, $0 \le t < t_l$, $m_\epsilon(x, y)(t) = 0$ for all sufficiently small positive ϵ. Similarly, for any t, $t_u < t < \infty$, $m_\epsilon(x, y)(t) = z_\epsilon(x, y)(t)$ for all sufficiently small positive ϵ. We need to examine the interval $(t_l - \epsilon, t_u + \epsilon)$ more carefully. To do so, we exploit the following analog of Lemma 9.4.1, which is proved in the same way.

Lemma 9.5.1. (uniform Lipschitz property for m_ϵ as a function of y) *For any $\epsilon > 0$, $t > 0$, $x \in D$ and $y_1, y_2 \in D$,*

$$\|m_\epsilon(x, y_1) - m_\epsilon(x, y_2)\|_t \le \|y_1 - y_2\|_t \ .$$

Our analog of Theorems 9.3.1, 9.3.2, 9.3.5 and 9.4.2 for m_ϵ is the following.

Theorem 9.5.2. (the derivative in the general case) *Suppose that $x, y \in D$. For each $t \ge 0$, $m_\epsilon(x, y)(t)$ is decreasing in ϵ and*

$$\lim_{\epsilon \downarrow 0} m_\epsilon(x, y)(t) = m(x, y)(t) \equiv \begin{cases} 0, & t < t_l \\ y(t-) \vee y(t) \vee 0, & t = t_l \\ z(t) \vee 0, & t_l < t < t_u \\ z(t-) \vee 0 \vee y(t), & t = t_u \\ z(t), & t > t_u \end{cases} \tag{5.14}$$

for m_ϵ in (5.3), t_l in (5.12), t_u in (5.13) and $z(t)$ in (3.6). The limit $m(x, y)$ in (5.14) has limits from the left and right at all t. If $x \in C$, then z is given by (3.39) and z and m are upper semicontinuous. At all t not in the set

$$B(x) \equiv \{t_l\} \cup (Bad(x) \cap (t_l, \infty)) \tag{5.15}$$

for Bad(x) in (3.28), *m is either left-continuous or right-continuous. At* $t = t_l$, *m is left-continuous if* $y(t-) \vee y(t) \le 0$, *m is right-continuous if* $y(t) \ge y(t-) \vee 0$, *and neither left-continuous nor right-continuous if* $y(t-) > y(t) \vee 0$. *If*

$$(i)\ y(t-) \le z(t-) \vee y(t) \vee 0 \quad \text{for} \quad t \in B(x) \cap [t_l, t_u] \tag{5.16}$$

and

$$(ii)\ y(t-) \le z(t-) \vee y(t) \quad \text{for} \quad t \in B(x) \cap (t_u, \infty)\ , \tag{5.17}$$

for which a sufficient condition is

$$Disc(y) \cap B(x) = \phi\ , \tag{5.18}$$

then m is either left-continuous or right-continuous at all t, so that $m \in D_{l,r}$. *Then*

$$m_\epsilon(x, y) \to m(x, y) \quad \text{in} \quad (D_{l,r}, M_1) \quad \text{as} \quad \epsilon \downarrow 0\ .$$

Proof. First, for any $\delta > 0$ and $T > 0$, $m_\epsilon(x, y)(t) = 0$ in $[0, (0 \vee (t_l - \delta)) \wedge T]$ and $m_\epsilon(x, y)(t) = z_\epsilon(x, y)(t)$ in $[(t_u + \delta) \wedge T, T]$ for all sufficiently small positive ϵ. We apply Theorems 9.3.1, 9.3.2 and 9.4.2 to treat the subinterval $[(t_u + \delta) \wedge T, T]$. Hence it suffices to focus on the subinterval $(t_l - \delta, t_u + \delta)$. By Lemmas 9.4.1 and 9.5.1, it suffices to assume that $y \in D_c$. The argument then is as for Theorems 9.3.1, 9.3.2, 9.3.5 and 9.4.2. ∎

Corollary 9.5.1. (convergence) *If* $x, y \in D$, *then*

$$\phi_\epsilon(x, y)(t) \downarrow y(t) + m(-x, -y)(t) \quad \text{as} \quad \epsilon \downarrow 0$$

for ϕ_ϵ *in* (5.2), *each* $t \ge 0$ *and m in* (5.14). *If in addition* (5.18) *holds, then*

$$\phi_\epsilon(x, y) \to y + m(-x, -y) \quad \text{in} \quad (D_{l,r}, M_1) \quad \text{as} \quad \epsilon \downarrow 0\ .$$

Finally, paralleling Theorem 9.4.3 for the supremum function, we can combine Theorems 9.2.3, 9.5.2 and the triangle inequality in (2.1) to obtain a preservation-of-convergence result for the reflection map.

Theorem 9.5.3. (M_1 convergence for the reflection derivative) *For* $\epsilon > 0$, *let* $x_\epsilon, y \in D$ *and let x be a Lipschitz function in C. If condition (4.4) holds, for which a sufficient condition is (4.5), and if* (5.18) *holds, then*

$$\epsilon^{-1}(\phi(x_\epsilon) - \phi(x)) \to y + m(-x, -y) \quad \text{in} \quad (D_{l,r}, M_1) \quad \text{as} \quad \epsilon \downarrow 0 \tag{5.19}$$

for m in (5.14).

Corollary 9.5.2. *For* $\epsilon > 0$, *let* $x_\epsilon \in D$ *and* $x, y \in C$ *with x being Lipschitz. If* (4.5) *holds, then* (5.19) *holds for m in* (5.14), *where* $m \in D_{usc} \cap D_{l,r}$.

9.6. Heavy-Traffic Limits for Nonstationary Queues

In this section we apply the convergence-preservation results in the last section to establish heavy-traffic limits for nonstationary queues. We assume that the queue-length process can be represented directly as the reflection map applied to a net-input process, which is the difference of two nondecreasing processes admitting nonstationary rates.

As background, note that the queue-length process $\{Q(t) : t \geq 0\}$ in the M/M/1 queue starting empty with arrival rate λ and service rate μ has such a representation. In particular, for the M/M/1 queue,

$$Q(t) = \phi(X)(t), \quad t \geq 0 \,, \tag{6.1}$$

where X is the net-input process, satisfying

$$X(t) = X^+(\Lambda^+(t)) - X^-(\Lambda^-(t)) \,, \tag{6.2}$$

with X^+ and X^- being rate-1 Poisson processes and

$$\Lambda^+(t) = \lambda t \quad \text{and} \quad \Lambda^-(t) = \mu t, \quad t \geq 0 \,. \tag{6.3}$$

Then $X^+ \circ \Lambda^+$ is a rate-λ Poisson process.

Similarly, for the $M_t/M_t/1$ queue with (integrable) time-dependent arrival-rate function $\lambda(t)$ and service-rate function $\mu(t)$, (6.1) and (6.2) remain valid with Λ^+ and Λ^- redefined as

$$\Lambda^+(t) = \int_0^t \lambda(s)ds \quad \text{and} \quad \Lambda^-(t) = \int_0^t \mu(s)ds \,. \tag{6.4}$$

It is easy to see that there are many generalizations. First, we obtain the queue-length process in an MMPP/MMPP/1 queue with independent Markov modulated Poisson process (MPPP) arrival and service processes if Λ^+ and Λ^- are independent stationary versions of finite-state continuous-time Markov chains. (We then assume that X^+, X^-, Λ^+ and Λ^- are mutually independent. We obtain the queue-length process in a more general $\mathrm{MMPP_t/MMPP_t}/1$ queue with independent time-dependent MMPP arrival and service processes if Λ^+ and Λ^- are independent time-dependent finite-state CTMCs, governed by time-dependent transition functions.

We construct associated fluid queue models by letting X^+ and X^- be other Lévy processes instead of Poisson processes. Without loss of generality, these again can be rate-1 processes. For nodes in a communication network with fixed bandwidth, it is natural to let $X^-(t) = t$, $t \geq 0$, but generalizations are possible.

We now establish limits for a sequence of models indexed by n. For each n, we have the four-tuple of stochastic processes $(X_n^+, X_n^-, \Lambda_n^+, \Lambda_n^-)$ with sample paths in D^4. We then form the associated scaled stochastic processes by letting

$$\begin{aligned}
\mathbf{X}_n^+(t) &\equiv c_n^{-1}[X_n^+(nt) - nx^+(t)] \\
\mathbf{X}_n^-(t) &\equiv c_n^{-1}[X_n^-(nt) - nx^-(t)] \\
\mathbf{\Lambda}_n^+(t) &\equiv c_n^{-1}[\Lambda_n^+(t) - ny^+(t)] \\
\mathbf{\Lambda}_n^-(t) &\equiv c_n^{-1}[\Lambda_n^-(t) - ny^-(t)] \\
\mathbf{X}_n(t) &\equiv c_n^{-1}[X_n^+(\Lambda_n^+(t)) - X_n^-(\Lambda_n^-(t)) - nx^+(y^+(t)) - x^-(y^-(t))] \\
\hat{\mathbf{X}}_n(t) &\equiv n^{-1}[X_n^+(\Lambda_n^+(t)) - X_n^-(\Lambda_n^-(t))]. \ t \geq 0 ,
\end{aligned} \tag{6.5}$$

We think of the centering terms x^+, x^-, y^+ and y^- as deterministic functions, but that is not necessary.

The following limit for the net-input process is a direct consequence of Theorem 13.3.2 in the book.

Theorem 9.6.1. (FLLN and FCLT for the net-input process) *Suppose that*

$$(\mathbf{X}_n^+, \mathbf{X}_n^-, \mathbf{\Lambda}_n^+, \mathbf{\Lambda}_n^-) \Rightarrow (\mathbf{U}^+, \mathbf{U}^-, \mathbf{V}^+, \mathbf{V}^-) \quad in \quad (D^4, WM_1) \tag{6.6}$$

for the processes in (6.5)*, where x^+ and x^- have continuous derivatives $\dot{x}^+$ and $\dot{x}^-$, y^+ and y^- are continuous nonnegative and strictly increasing, $c_n \to \infty$, $n/c_n \to \infty$ and*

$$\begin{aligned}
&Disc(\mathbf{U}^+ \circ y^+) \cap Disc(\mathbf{V}^+) = \phi \\
&Disc(\mathbf{U}^- \circ y^-) \cap Disc(\mathbf{V}^-) = \phi \\
&Disc(\mathbf{U}^+ \circ y^+ + (\dot{x}^+ \circ y^+)\mathbf{V}^+) \cap \\
&\quad Disc(\mathbf{U}^- \circ y^- + (\dot{x}^- \circ y^-)\mathbf{V}^-) = \phi .
\end{aligned} \tag{6.7}$$

Then

$$\hat{\mathbf{X}}_n \Rightarrow x \quad in \quad (D, M_1) \tag{6.8}$$

and

$$\mathbf{X}_n \Rightarrow \mathbf{X} \quad in \quad (D, M_1) \ , \tag{6.9}$$

for $\hat{\mathbf{X}}_n$ and $\mathbf{X}_n$ in (6.5)*, where*

$$x \equiv x^+ \circ y^+ - x^- \circ y^- \tag{6.10}$$

and

$$\mathbf{X} \equiv \mathbf{U}^+ \circ y^+ + (\dot{x}^+ \circ y^+)\mathbf{V}^+ - \mathbf{U}^- \circ y^- - (\dot{x}^- \circ y^-)\mathbf{V}^- . \tag{6.11}$$

Proof. As usual, start by applying the Skorohod representation theorem to replace the convergence in distribution in (6.6) by convergence w.p.1 for special versions, without introducing new notation for the special versions. Then apply Theorem 13.3.2 in the book, after rewriting $\mathbf{X}_n^+$ as

$$\mathbf{X}_n^+(t) \equiv (n/c_n)[n^{-1}X_n^+(nt) - x^+(t)], \quad t \ge 0\,, \tag{6.12}$$

and similarly for the other functions. That yields

$$\begin{aligned} &c_n^{-1}(X_n^+ \circ \Lambda_n^+ - nx^+ \circ y^+, X_n^- \circ \Lambda_n^- - nx^- \circ y^-) \\ &\quad \Rightarrow (U^+ \circ y^+ + (\dot{x}^+ \circ y^+)V^+, U^- \circ y^- + (\dot{x}^- \circ y^-)V^-) \end{aligned} \tag{6.13}$$

in (D^2, WM_1). Multiply by c_n/n in (6.13) to get

$$n^{-1}(X_n^+ \circ \Lambda_n^+, X_n^- \circ \Lambda_n^-) \Rightarrow (x^+ \circ y^+, x^- \circ y^-) \quad \text{in} \quad (D^2, WM_1) \tag{6.14}$$

Finally, given the last condition in (6.7), we can apply addition to go from (6.13) and (6.14) to (6.9) and (6.8). ∎

We now apply Theorem 9.5.1 to obtain a corresponding result for the queue-length processes. Let

$$\mathbf{Q}_n(t) \equiv c_n^{-1}(Q_n(nt) - nq(t)), \quad t \ge 0\,. \tag{6.15}$$

and

$$\hat{\mathbf{Q}}_n(t) \equiv n^{-1}Q_n(nt), \quad t \ge 0\,. \tag{6.16}$$

Theorem 9.6.2. (FLLN and FCLT for the queue-length process) *If, in addition to the assumptions of Theorem 9.6.1, y^+ and y^- are Lipschitz continuous, $x(0) = 0$, $P(\mathbf{X}(0) = 0) = 1$ and*

$$P((\mathbf{U}^+, \mathbf{U}^-, \mathbf{V}^+, \mathbf{V}^-) \in C^4) = 1\ , \tag{6.17}$$

then

$$\hat{\mathbf{Q}}_n \Rightarrow q \quad in \quad (D, M_1) \tag{6.18}$$

and

$$\mathbf{Q}_n \Rightarrow \mathbf{Q} \quad in \quad (D_{l,r}, M_1) \tag{6.19}$$

for $\hat{\mathbf{Q}}_n$ in (6.16) and $\mathbf{Q}_n$ in (6.15), where

$$q = \phi(x) \tag{6.20}$$

for x in (6.10) and

$$\mathbf{Q} = \mathbf{X} + z(-x, -\mathbf{X}) \tag{6.21}$$

for x in (6.10), $\mathbf{X}$ in (6.11) and z in (3.16). The limit process $\mathbf{Q}$ then has upper semicontinuous sample paths.

Example 9.6.1. *The* $\mathrm{M_t/M_t/1}$ *queue.* Now let us examine the special case of the $\mathrm{M_t/M_t/1}$ queue in more detail. For the $\mathrm{M_t/M_t/1}$ queue, $c_n = \sqrt{n}$, $x^+ = x^- = e$ and U^+, U^- are independent Brownian motions. It is natural to have

$$\Lambda_n^{\pm}(t) = \int_0^t \lambda_n^{\pm}(s)ds \quad \text{and} \quad y^{\pm}(t) = \int_0^t \lambda^{\pm}(s)ds \tag{6.22}$$

where $\lambda_n^{\pm}$ and $\lambda^{\pm}$ are deterministic functions. We can then have

$$n^{-1/2}(\lambda_n^{\pm}(t) - n\lambda^{\pm}(t)) \to \gamma^{\pm}(t) \quad \text{as} \quad n \to \infty \tag{6.23}$$

uniformly in $[0,T]$, where γ^+ and γ^- are deterministic, which implies that

$$\mathbf{\Lambda}_n^{\pm}(t) \to \int_0^t \gamma^{\pm}(s)ds \equiv \mathbf{V}^{\pm}\ . \tag{6.24}$$

Thus the assumptions of Theorems 9.6.1 and 9.6.2 are satisfied and

$$x(t) = \int_0^t [\lambda^+(s) - \lambda^-(s)]ds, \quad t \ge 0\,, \tag{6.25}$$

while

$$\begin{aligned} \mathbf{X}(t) &= \mathbf{U}^+\left(\int_0^t \lambda^+(s)ds\right) \\ &\quad -\mathbf{U}^-\left(\int_0^t \lambda^-(s)ds\right) + \int_0^t [\gamma^+(s) - \gamma^-(s)]ds \end{aligned} \tag{6.26}$$

where $\mathbf{U}^+$ and $\mathbf{U}^-$ are independent standard Brownian motions and the rest involves continuous deterministic functions. It is easy to see that X is equal in distribution (on D) to

$$U\left(\int_0^t [\lambda^+(s) + \lambda^-(s)]ds\right) + \int_0^t [\gamma^+(s) - \gamma^-(s)]ds, \quad t \ge 0\,, \tag{6.27}$$

where U is a standard Brownian motion.

The FWLLN limits x and q can be regarded as the net-input and buffer-content processes, respectively, in a fluid-queue model with time-dependent deterministic input rate $\lambda^+(t)$ and time-dependent deterministic potential output rate $\lambda^-(t)$. Then

$$-(-x)^{\downarrow} = -\min_{0 \le s \le t}\left\{\int_0^s [\lambda^-(r) - \lambda^+(r)]dr\right\} \tag{6.28}$$

represents the cumulative potential output that is lost (i.e., does not occur during the interval $[0, t]$ because of insufficient input. Then

$$\Phi_{-x}(t) = \{s : 0 \le s \le t, q(s) = 0, -(-x)^{\downarrow}(s) = -(-x)^{\downarrow}(t)\} \qquad (6.29)$$

i.e., $\Phi_{-x}(t)$ is the set of times s at which the buffer is empty ($q(s) = 0$) and there is no potential output los over $[s, t]$.

An important special case is when λ_n^+ and λ_n^- in (6.22) are independent of n. Then $\gamma^+(t) = \gamma^-(t) = 0$ for all $t \ge 0$ and the deterministic function $\int_0^t [\gamma^+(s) - \gamma^-(s)]ds$ in (6.27) is identically 0. Then the limit for the queue-length process has one of three forms over subintervals: time-scaled standard Brownian motion (BM), time-scaled canonical reflected Brownian motion (RBM) and the zero function. There can be discontinuities in the sample path when the set function $\Phi_{-x}(t)$ is discontinuous in t. We display possible sample paths of (λ^+, λ^-), $(-x, (-x)^{\uparrow})$, $\Phi_{-x}(t)$, q and $\mathbf{Q}$ when λ^- is the constant function in Figure 9.2 below. We identify nine intervals associated with nine time points $t_0 \equiv 0 < t_1 < \cdots < t_8$.

In this example, the fluid rates start out ordered by $\lambda^+(t) < \lambda^-(t)$. Thus $-x(t) \equiv \int_0^t [\lambda^-(s) - \lambda^+(s)]ds$ is initially increasing, which implies that $\Phi_{-x}(t) = \{t\}$. Thus $Q(t) = q(t) = 0$ for these t. At time t_1, the ordering switches to $\lambda^+(t) > \lambda^-(t)$. Thus after t_1, $-x$ is decreasing, so that $\Phi_{-x}(t) = \{t_1\}$. At time t_2, the ordering switches back to $\lambda^+(t) < \lambda^-(t)$, but $-x(t)$ does not reach $(-x)^{\uparrow}(t) = (-x)(t_1)$ and $q(t)$ does not return to 0 until $t = t_3$. In the interval (t_1, t_3), q is positive and Q is time-scaled BM.

At time t_3, there is a discontinuity in the set-valued function Φ_{-x} and a corresponding jump in the stochastic process Q. In the interval (t_3, t_4), $-x$ is still increasing and $\Phi_{-x}(t) = \{t\}$, so that $q(t) = Q(t) = 0$, just as in $[0, t_1)$. In the interval (t_4, t_5), $\lambda^+(t) = \lambda^-$, so that $-x$ is constant and $\Phi_{-x}(t) = [t_4, t]$, $t_4 \le t \le t_5$. In the interval (t_4, t_5), Q evolves as RBM. At t_5, λ^+ increases, so that $-x$ decreases and $\Phi_{-x}(t) = \Phi_{-x}(t_5) = [t_4, t_5]$ for $t_5 \le t < t_7$. At t_6, λ^+ starts to decrease again and at t_7 $q(t) = 0$ for the first time. Hence, Q evolves as BM in the interval (t_5, t_7).

At t_7, there is a second discontinuity in Φ_{-x} and a corresponding jump in Q. In the subsequent interval $[t_7, t_8]$, $\lambda^+(t) = \lambda^-$, so that $-x$ remains constant. Then $\Phi_{-x}(t) = [t_4, t_5] \cup [t_7, t]$ for $t_7 \le t < t_8$. During the interval $[t_7, t_8]$, $q(t) = 0$ and Q evolves as RBM. At t_8, λ^+ starts to decrease and thereafter remains below λ^-. Hence, Φ_{-x} has another discontinuity at t_8. After t_8, $\Phi_{-x}(t) = \{t\}$ and $q(t) = Q(t) = 0$.

We conclude this section by relating the three possible kinds of heavy-traffic limits for the case of the $\mathrm{M_t/M_t/1}$ queue with fixed arrival and service rate functions $\lambda^+(t)$ and $\mu^-(t)$ to the values of a *time-dependent traffic intensity*, defined by

$$\rho^*(t) \equiv \sup_{0 \le s \le t} \left\{ \int_0^t \lambda^+(r)dr / \int_s^t \lambda^-(r)dr \right\}, \quad t \ge 0 . \tag{6.30}$$

Notice that the buffer-content deterministic fluid limit q satisfies

$$\begin{aligned} q(t) &= x(t) - \inf_{0 \le s - t} x(s) \\ &= \sup_{0 \le s \le t} \{x(t) - x(s)\} \\ &= \sup_{0 \le s \le t} \left\{ \int_s^t [\lambda^+(r) - \lambda^-(r)]dr \right\} , \end{aligned} \tag{6.31}$$

so that $q(t) > 0$ if and only if $\rho^*(t) > 1$.

Moreover, we can have $q(t) = 0$ but $P(Q(t) = 0) = 0$ for all t in an interval (a, b) if and only if $\rho^*(t) = 1$ in (a, b). First, we must have $\rho^* \le 1$ since $q(t) = 0$. However, in this region we must also have

$$\int_s^t [\lambda^+(r) - \lambda^-(r)]dr = 0 \tag{6.32}$$

for some s suitably chose to t. For that s,

$$\int_s^t \lambda^+(r)dr / \int_s^t \lambda^-(r)dr = 1 \tag{6.33}$$

which implies that $\rho^*(t) \ge 1$. Since both $\rho^*(t) \le 1$ and $\rho^*(t) \ge 1$, we must have $\rho^*(t) = 1$.

We thus say that the queue is *overloaded*, *critically loaded* or *underloaded* in an open interval (a, b) if $\rho^*(t) > 1$, $\rho^*(t) = 1$ or $\rho^*(t) < 1$ throughout the interval (a, b). In Figure 9.2 above, in the intervals $(0, t_1)$, (t_1, t_3), (t_3, t_4), (t_4, t_5), (t_5, t_7), (t_7, t_8) and (t_8, T), we have successively $\rho^*(t) < 1$, > 1, < 1, $= 1$, > 1, $= 1$ and < 1.

9.7. Derivative of the Inverse Map

In this section we obtain convergence-preservation results for the inverse map

$$x^{-1}(t) \equiv \inf\{s \ge 0 : x(s) > t\}, \quad t \ge 0 , \tag{7.1}$$

defined on the subset D_u of functions unbounded above in $D \equiv D([0,\infty),\mathbb{R})$, as in Section 13.6 of the book. As in previous sections here, we approach convergence preservation through a derivative representation.

To determine the derivative of the inverse map, we introduce yet another topology on D. Recall that we introduced the M_1' topology on $D([0,t],\mathbb{R})$ by appending a segment to the graphs, i.e., by letting

$$\Gamma_x' = \Gamma_x \cup \{(\alpha x(0), 0) : 0 \le \alpha \le 1\} \ , \tag{7.2}$$

where Γ_x is the graph of x, i.e.,

$$\begin{aligned}\Gamma_x \equiv \{&(z,s) \in \mathbb{R} \times [0,t] : \\ &z = \alpha x(s-) + (1-\alpha)x(s) \quad \text{for some } \alpha,\ 0 \le \alpha \le 1\}\,.\end{aligned} \tag{7.3}$$

We now construct a similar M_1'' topology on $D([0,t),\mathbb{R})$ by also appending the vertical line at t to the graph, i.e., by setting

$$\Gamma_x'' = \Gamma_x' \cup (\mathbb{R} \times \{t\}) \tag{7.4}$$

for Γ_x' in (7.2). Note that the function value at the right endpoint t plays no role in the M_1'' topology.

As done before for the graph Γ_x in (7.3), we define a lexicographic *order relation* on the graph Γ_x'' by saying that $(z_1,s_1) \le (z_2,s_2)$ if either (i) $s_1 < s_2$ or (ii) $s_1 = s_2$ and $|x(s_1-) - z_1| \le |x(s_1-) - z_2|$. The definition makes the relation $\le$ a total order on the graph Γ_x''. A parametric representation of the graph Γ_x'' or the function x is a continuous nondecreasing function (u,r) mapping $[0,1]$ *into* the graph Γ_x'' such that $r(0) = 0$, $u(0) = 0$ and $r(1) = t$. We allow the parametric representation of Γ_x'' to cover only part of the vertical line at t. If $r(s) < t$ for all $s < 1$, then the parametric representation (u,r) covers only the single point $(x(t-),t)$. If $r(s) = t$ for $a \le s \le 1$, then (u,r) covers a compact subinterval of either $\{(z,t) : z \ge x(t-)\}$ or $\{(z,t) : z \le x(t-)\}$. (Since (u,r) maps $[0,1]$ into Γ_x'', we must have $(u(1),r(1)) \in \Gamma_x''$, which implies that $|u(1)| < \infty$.) Let $\Pi''(x)$ be the set of all parametric representations of Γ_x''.

A metric d_t'' inducing the M_1'' topology on $D([0,t),\mathbb{R})$ is defined by letting

$$d_t''(x_1,x_2) = \inf_{\substack{(u_i,r_i)\in\Pi''(x_i)\\ i=1,2}} \{\|u_1 - u_2\|_1 \vee \|r_1 - r_2\|_1\} \ . \tag{7.5}$$

We have the following lemma linking the M_1' and M_1'' topologies with bounded function domains.

Lemma 9.7.1. *Let $x, x_n \in D([0,\infty),\mathbb{R})$. If $x_n \to x$ as $n \to \infty$ for the restrictions in $D([0,t_2),\mathbb{R},M_1'')$ for $0 < t_2 < \infty$, then $x_n \to x$ as $n \to \infty$ for the restrictions in $D([0,t_1],\mathbb{R},M_1')$ for each $t_1 \in Disc(x)^c$ with $0 < t_1 < t_2$.*

As before, we say that $x_n \to x$ in $D([0,\infty),R)$ with any of the topologies M_1, M_1' or M_1'' if $x_n \to x$ for the restrictions in $D([0,t],\mathbb{R})$ ($D([0,t),\mathbb{R})$ for M_2'') with the same topology for all t in a sequence $\{t_k\}$ with $t_k \to \infty$ as $k \to \infty$. (The boundary points t_k can be taken from $Disc(x)^c$.) We obtain the following result from Lemma 9.7.1.

Lemma 9.7.2. *The M_1' and M_1'' topologies coincide on $D([0,\infty),\mathbb{R})$.*

We can combine Lemma 9.7.2 here and Theorem 13.6.3 in the book to obtain the following connection between M_1'' and M_1.

Lemma 9.7.3. *If*

$$x_n \to x \quad in \quad D([0,\infty),\mathbb{R},M_1'') \ ,$$

where $x(0) = 0$, then

$$x_n \to x \quad in \quad D([0,\infty),\mathbb{R},M_1) \ .$$

A metric d'' inducing the M_1'' topology on $D([0,\infty),\mathbb{R})$ is defined by letting

$$d''(x_1,x_2) = \int_0^\infty e^{-t}[1 \wedge d_t''(x_1,x_2)]dt \ , \tag{7.6}$$

where $d_t''(x_1,x_2)$ is understood to be the d_t'' metric applied to the restrictions of x_1 and x_2 to $[0,t)$. There is convergence $d''(x_n,x) \to 0$ if and only if there exist parametric representations (u,r) of x and (u_n,r_n) of x_n, $n \geq 1$ with domains $[0,\infty)$, such that $\|u_n - u\|_t \vee \|r_n - r\|_t \to 0$ as $n \to \infty$ for each t.

To apply the approach in Section 9.2, we need the inverse map to be Lipschitz. The Lipschitz property is valid on an appropriate subset of D with an appropriate choice of metrics. Recall that D_u is the subset of functions x in $D \equiv D([0,\infty),\mathbb{R})$ that are unbounded above and have $x(0) \geq 0$. For positive t_1, t_2, let $D_u(t_1,t_2)$ be the subset of x in D_u with $x^{\uparrow}(t_1) \geq t_2$. Clearly $D_u(t_1,t_2)$ is a closed subset of D_u. Moreover,

$$D_u = \cap_{m=1}^{\infty} \cup_{k=1}^{\infty} D(k,m) \ . \tag{7.7}$$

We now show that the inverse map from $D_u(t_1,t_2) \subseteq D_u([0,t_1],\mathbb{R},M_1)$ to $D([0,t_2),\mathbb{R},M_1'')$ is Lipschitz.

Lemma 9.7.4. *For $t > 0$, let d''_t be the M''_1 metric on $D([0,t),\mathbb{R})$ and let d_t be the M_1 metric on $D([0,t],\mathbb{R})$. If x_1, $x_2 \in D_u(t_1,t_2)$, then*

$$d''_{t_2}(x_1^{-1}, x_2^{-1}) \le d_{t_1}(x_1^{\uparrow} \wedge t_2, x_2^{\uparrow} \wedge t_2) \le d_{t_1}(x_1^{\uparrow}, x_2^{\uparrow}) \le d_{t_1}(x_1, x_2) \ . \qquad (7.8)$$

where $(x_i^{\uparrow} \wedge t_2)(s) = x_i^{\uparrow}(s) \wedge t_2$, $0 \le s \le t_1$.

Proof. For $x_i \in D_u(t_1,t_2)$, let (u_i, r_i) be an arbitrary M_1 parametric representation of $x_i^{\uparrow} \wedge t_2$ over $[0,t_1]$. Then (r_i, u_i) is an M''_1 parametric representation of x_i^{-1} over $[0,t_2)$ with the special property that $u_i(1) = t_1$. Hence

$$d''_{t_2}(x_1^{-1}, x_2^{-1}) \le d_{t_1}(x_1^{\uparrow} \wedge t_2, x_2^{\uparrow} \wedge t_2) \ . \qquad (7.9)$$

It is not difficult to see that

$$d_{t_1}(x_1^{\uparrow} \wedge t_1, x_2^{\uparrow} \wedge t_2) \le d_{t_1}(x_1^{\uparrow}, x_2^{\uparrow}) \le d_{t_1}(x_1, x_2) \ .$$

Hence the proof is complete. ∎

Lemmas 9.7.2 and 9.7.4 imply that the inverse map from $D_u([0,\infty),\mathbb{R},M_1)$ to $D_u([0,\infty),\mathbb{R},M'_1)$ is continuous, which is weaker than Theorem 13.6.2 in the book. We now want to establish an analog of Theorem 9.2.3. For that purpose, we need both x and x^{-1} to be Lipschitz on $[0,t]$ for all $t > 0$. The following lemma provides natural conditions.

Lemma 9.7.5. (conditions for both x and x^{-1} to be Lipschitz) *If $x \in D_u$ is absolutely continuous, i.e., $x(t) = \int_0^t \dot{x}(s)ds$ for $t > 0$, with $\dot{x} \in D$ and with $l(t) \le \dot{x}(t) \le u(t)$ for all $t \ge 0$ where $0 < l^{\downarrow}(t) < u^{\uparrow}(t) < \infty$ for all t, then*

$$x^{-1}(t) = \int_0^t [1/\dot{x}(x^{-1}(s))]ds \quad \text{for all} \quad t > 0 \qquad (7.10)$$

and x and x^{-1} are both Lipschitz on $[0,t]$ for all $t > 0$, with

$$\dot{x}^{-1}(t) \equiv \frac{d}{dt}(x^{-1})(t) = 1/\dot{x}(x^{-1}(t)) \ . \qquad (7.11)$$

Proof. Clearly x is strictly increasing and continuous, so that x is a homeomorphism of $[0,\infty)$ and $x \circ x^{-1} = e$, where $\circ$ is the composition map. Thus

$$x(x^{-1}(t)) = \int_0^{x^{-1}(t)} \dot{x}(s)ds = t, \quad t \ge 0 \ ,$$

which implies that

$$x^{-1}(t) = \int_0^t [1/\dot{x}(x^{-1}(s))]ds, \quad t \geq 0 .$$

The Lipschitz properties hold because

$$|x(t_2) - x(t_1)| = \int_{t_1}^{t_2} \dot{x}(s)ds \leq u^{\uparrow}(t_2)|t_2 - t_1|$$

and

$$|x^{-1}(t_2) - x^{-1}(t_1)| = \int_{t_1}^{t_2} [1/\dot{x}(x^{-1}(s))]ds \leq |t_2 - t_1|/l^{\downarrow}(t_2) . \quad \blacksquare$$

We now want to establish an analog of Theorem 9.2.3. Since the M_1'' analog of Lemma 9.2.2 is evident, we only establish the M_1'' analog of Lemma 9.2.1.

Lemma 9.7.6. (reduction of convergence to the derivative with the M_1'' topology) *Suppose that x is Lipschitz on $[0,t]$ with Lipschitz constant K. Let d_t'' be the M_1'' metric on $D([0,t),\mathbb{R})$. Then*

$$d_t''(x_1 - x, x_2 - x) \leq (1+K)d_t''(x_1, x_2) .$$

Proof. For all $\epsilon > 0$, there exists $\eta(\epsilon) > 0$ and parametric representations $(u_{i,\epsilon}, r_{i,\epsilon}) \in \Pi_t''(x_i)$ such that

$$\|u_{1,\epsilon} - u_{2,\epsilon}\| \vee \|r_{1,\epsilon} - r_{2,\epsilon}\| \leq (1+\eta(\epsilon))d_t''(x_1, x_2) . \tag{7.12}$$

We now want natural modifications of the parametric representations of x_i to serve as parametric representations of x and $x_i - x$. To obtain such parametric representations for x, we need to allow for the line segment joining $(x(0), 0)$ to $(0,0)$. Hence we first modify the parametric representations of x_i. Let $(u_{i,\epsilon}', r_{i,\epsilon}') \in \Pi_t''(x_i)$ be scaled versions of the parametric representations $(u_{i,\epsilon}, r_{i,\epsilon})$ on $[\delta, 1]$ with $(u_{i,\epsilon}'(s), r_{i,\epsilon}'(s)) = (0,0)$, $0 \leq s \leq \delta$, i.e.,

$$(u_{i,\epsilon}'(\delta+s), r_{i,\epsilon}'(\delta+s)) = (u_{i,\epsilon}((1-\delta)^{-1}s), r_{i,\epsilon}((1-\delta)^{-1}s)), \quad 0 \leq s \leq 1-\delta . \tag{7.13}$$

Then

$$\|u_{1,\epsilon}' - u_{2,\epsilon}'\| \vee \|r_{1,\epsilon}' - r_{2,\epsilon}'\| = \|u_{1,\epsilon} - u_{2,\epsilon}\| \vee \|r_{1,\epsilon} - r_{2,\epsilon}\| . \tag{7.14}$$

Since $x \in C$, $(u''_{i,\epsilon}, r'_{i,\epsilon}) \in \Pi''(x)$ for $i = 1, 2$, if

$$u''_{i,\epsilon}(s) = \begin{cases} x \circ r'_{i,\epsilon}, & \delta \le s \le 1 \\ 0, & s = 0 \end{cases}$$

with $u''_{i,\epsilon}$ defined by linear interpolation on $(0, \delta)$. Then $(u'_{i,\epsilon} - u''_{i,\epsilon}, r'_{i,\epsilon}) \in \Pi''(x_i - x)$ and

$$\begin{aligned} d''_t(x_1 - x, x_2 - x) &\le \|(u'_{1,\epsilon} - u''_{1,\epsilon}) - (u'_{2,\epsilon} - u''_{2,\epsilon})\| \vee \|r'_{1,\epsilon} - r'_{2,\epsilon}\| \\ &\le (\|u'_{1,\epsilon} - u'_{2,\epsilon}\| + \|x \circ r_{1,\epsilon} - x \circ r_{2,\epsilon}\|) \vee \|r'_{1,\epsilon} - r'_{2,\epsilon}\| \\ &\le (1 + K)(1 + \eta(\epsilon)) d''_t(x_1, x_2) . \end{aligned}$$

Since $\eta(\epsilon) \to 0$ as $\epsilon \to 0$, the proof is complete. ∎

We now obtain the M''_1-analog of Theorem 9.2.3. By Lemma 9.7.2, the M'_1 and M''_1 topologies agree on $D([0, \infty), \mathbb{R})$.

Theorem 9.7.1. *Suppose that $x, x_\epsilon \in D_u([0, \infty), \mathbb{R})$ and that x satisfies the condition of Lemma 9.7.5. If $d_t(x_\epsilon - x, \epsilon y) = o(\epsilon)$ as $\epsilon \to 0$ for t in a sequence $\{t_k\}$ with $t_k \to \infty$ as $k \to \infty$, for which a sufficient condition is $\|\epsilon^{-1}(x_\epsilon - x) - y\|_t \to 0$ as $\epsilon \downarrow 0$ for all $t > 0$, then*

$$d'(\epsilon^{-1}[x_\epsilon^{-1} - x^{-1}], \epsilon^{-1}[(x + \epsilon y)^{-1} - x^{-1}]) \to 0 \quad as \quad \epsilon \downarrow 0 , \tag{7.15}$$

where d' is the M'_1 metric on $D([0, \infty, \mathbb{R})$.

Proof. For any $t_2 > 0$, choose t_1 such that $d_{t_1}(x_\epsilon - x, \epsilon y) = o(\epsilon)$ as $\epsilon \downarrow 0$ and $x^{-1}(t_2) < t_1$. The assumptions imply that $\|x_\epsilon - x\|_{t_1} \to 0$ and $\|\epsilon y\|_{t_1} \to 0$ as $\epsilon \downarrow 0$. Hence, for all sufficiently small ϵ, $x_\epsilon, x + \epsilon y \in D_u(t_1, t_2)$. On $D_u(t_1, t_2)$, we can apply Lemmas 9.7.4 and 9.7.6, and the M''_1 analog of Lemma 9.2.2 to conclude for $\epsilon \le 1$ that there are constants K_1 and K_2 such that

$$\begin{aligned} & d''_{t_2}(\epsilon^{-1}[x_\epsilon^{-1} - x^{-1}], \epsilon^{-1}[(x + \epsilon y)^{-1} - x^{-1}]) \\ & \quad \le \epsilon^{-1} d''_{t_2}(x_\epsilon^{-1} - x^{-1}, (x + \epsilon y)^{-1} - x^{-1}) \\ \le & \quad K_1 \epsilon^{-1} d''_{t_2}(x_\epsilon^{-1}, (x + \epsilon y)^{-1}) \\ \le & \quad K_1 \epsilon^{-1} d_{t_1}(x_\epsilon, x + \epsilon y) \\ \le & \quad K_1 K_2 \epsilon^{-1} d_{t_1}(x_\epsilon - x, \epsilon y) \\ \le & \quad K_1 K_2 \|(\epsilon^{-1}(x_\epsilon - x) - y\|_{t_1} . \end{aligned} \tag{7.16}$$

This argument applies for arbitrarily large t_2 provided that we increase t_1 appropriately. ∎

We now focus on the derivative of the inverse map. Let

$$z_\epsilon \equiv z_\epsilon(x,y) \equiv \epsilon^{-1}[(x+\epsilon y)^{-1} - x^{-1}] \; . \qquad (7.17)$$

We first observe that z_ϵ in (7.17) is monotone decreasing in y.

Lemma 9.7.7. *For any $x \in D_u$ and $y \in D$, if $y_1(t) \le y_2(t)$ for all t, then $z_\epsilon(x,y_1)(t) \ge z_\epsilon(x,y_2)(t)$ for all ϵ and t, where z_ϵ is defined in* (7.17).

We now show that it suffices to consider piecewise-constant functions y, because under regularity conditions, $z_\epsilon(x,y)$ as a function of y is Lipschitz. Hence, for x and y given, we can replace y by $y_c \in D_c$.

Lemma 9.7.8. *Suppose that $x \in D_u$, $\dot{x} \in D$, $y_1 \in D$, $t_1 = x^{-1}(t_2)+1$, $0 < a \le \|\dot{x}\|_{t_1} < \infty$ and $\|y_1\|_{t_1} \le K$. If $\|y_1 - y_2\|_{t_1} < 1$, then*

$$\|z_\epsilon(x,y_1) - z_\epsilon(x,y_2)\|_{t_2} \le (2/a)\|y_1 - y_2\|_{t_1}$$

provided that $\epsilon \le a/[K+1]$.

Proof. By the monotonicity established in Lemma 9.7.7,

$$z_\epsilon(x, y_1 - \delta) \ge z_\epsilon(x,y_1), z_\epsilon(x,y_2) \ge z_\epsilon(x, y_1+\delta)$$

on $[0,t]$ provided that $\|y_1 - y_2\|_{t_1} \le \delta$ for a suitably large t_1. For the given t_1 and $\delta \le 1$,

$$\begin{aligned}(x+\epsilon y_i)^{-1}(t) &\le (x+\epsilon(y_1-\delta))^{-1}(t)\\ &\le (x-\epsilon(K+\delta))^{-1}(t)\\ &\le x^{-1}(t) + \frac{\epsilon(K+\delta)}{a} \le t_1\end{aligned}$$

provided that $\delta \le 1$ and $\epsilon \le a/(K+1)$. Hence, if $\|\dot{x}\|_{t_1} \ge a$ and $\|y_1\|_{t_1} \le K$ for that t_1, the inverses are all contained in $[0,t_1]$. Then, for $\|y_1 - y_2\|_{t_1} \le \delta \le 1$,

$$\begin{aligned}\|z_\epsilon(x,y_1) - z_\epsilon(x,y_2)\|_{t_2} &\le \|z_\epsilon(x,y_1+\delta) - z_\epsilon(x,y_1-\delta\|_{t_2}\\ &= \epsilon^{-1}\|(x+\epsilon(y_1-\delta))^{-1} - (x+\epsilon(y_1+\delta))^{-1}\|_{t_2}\\ &\le x^{-1}(t_2) + 2\delta/a \; . \quad \blacksquare\end{aligned}$$

We now establish pointwise convergence. For this purpose, let

$$Pos(x) = \{t \ge 0 : x(t) > 0\} \; . \qquad (7.18)$$

We obtain the following result by examining the indicated cases.

Theorem 9.7.2. *If $y \in D$ and $x \in D_u$ satisfies the condition of Lemma 9.7.5, then*

$$z_\epsilon(t) \equiv \epsilon^{-1}[(x+\epsilon y)^{-1}(t) - x^{-1}(t)] \to z(t) \quad in \quad \mathbb{R} \quad as \quad \epsilon \downarrow 0$$

for each t, where

$$(i) \quad z(t) = \frac{-y(x^{-1}(t)-)}{\dot{x}(x^{-1}(t)-)} < 0 \tag{7.19}$$

if $y(x^{-1}(t)-) > 0$;

$$(ii) \quad z(t) = \frac{-y(x^{-1}(t))}{\dot{x}(x^{-1}(t))} > 0 \tag{7.20}$$

if $y(x^{-1}(t)-) < 0$ and $y(x^{-1}(t)) < 0$ or if $y(x^{-1}(t)-) = 0$, $\sup\{Pos(y \circ x^{-1}) \cap [0,t)\} < t$ and $y(x^{-1}(t)) < 0$;

$$(iii) \quad z(t) = 0 \tag{7.21}$$

otherwise: if one of: (a) $y(x^{-1}(t)-) = 0$ and $\sup\{Pos(y \circ x) \cap [0,t)\} = t$, (b) $y(x^{-1}(t)-) < 0$ and $y(x^{-1}(t)) = 0$, (c) $y(x^{-1}(t)-) = 0$, $\sup\{Pos(y \circ x) \cap [0,t)\} < t$ and $y(x^{-1}(t)) = 0$, or (d) $y(x^{-1}(t)-) < 0 < y(x^{-1}(t))$.

Consequently, z is either left-continuous or right-continuous at t unless $y(x(t)-) < 0 < y(x(t))$, in which case $z(t-) > z(t) > z(t+)$.

Proof. It is elementary that $z_\epsilon(t)$ converges pointwise to $z(t)$ for $z(t)$ in (7.20) when both $\dot{x}$ and y are continuous at $x^{-1}(t)$, so that z is continuous at t. For the other cases, we apply Lemma 9.7.8 to approximate y by a piecewise-constant function. We then exploit Lemma 9.7.7 and the fact that $\dot{x}$ and y are elements of D. We obtain the conclusions by examining the different cases. ■

Remark 9.7.1. In order to have the pointwise convergence in Theorem 9.7.2, at a single t, it suffices to have the conditions on x and y hold only in a neighborhood of $x^{-1}(t)$. Then x need not be absolutely continuous or strictly increasing everywhere.

Remark 9.7.2. We have difficulty at some t if x is only an increasing homeomorphism of $[0,\infty)$. Then we can have $\dot{x}(x^{-1}(t)) = 0$ and $\dot{x}^{-1}(t)) = \infty$ for some t, so that $z_\epsilon(t) \to \infty$ as $\epsilon \downarrow 0$.

We now want to establish M_1' convergence in D. However, first we note that the limit z does not necessarily belong to D, because it may be neither left-continuous nor right-continuous at discontinuity points.

Example 9.7.1. *We need not have $z \in D$.* To see that we need not have $z \in D$, even if $\dot{x} \in C$, let $x = e$ and let $y = -I_{[0,1)} + I_{[1,\infty)}$. Then

$$z_\epsilon(t) = I_{[0,1-\epsilon)}(t) + \epsilon^{-1}(1-t)I_{[1-\epsilon,1+\epsilon)}(t) - I_{[1+\epsilon,\infty)}(t) \tag{7.22}$$

and

$$z = I_{[0,1)} - I_{(1,\infty)} \tag{7.23}$$

so that $z(1) = 0$, but $z(1-) = 1$ and $z(1+) = -1$. However, $z(1)$ is in between $z(1-)$ and $z(1+)$. ∎

Since $z(t)$ lies between $z(t-)$ and $z(t+)$ for all t, the space D^* of such functions with the M_1 and M', topologies is equivalent to D because functions in D and D^* have the same graphs.

Theorem 9.7.3. (conditions for convergence to the right-continuous version) *If $y \in D$ and $x \in D_u$ satisfies the condition of Lemma 9.7.5 with $\dot{x} \in D$, then*

$$z_\epsilon \to z_+ \quad in \quad (D, M_1') \quad as \quad \epsilon \downarrow 0$$

for z_ϵ in (7.17) *and z_+ the right-continuous version of z, i.e. $z_+(t) = z(t+)$, $t \geq 0$ and z in* (7.19). *If $z_+(0) = 0$, the convergence is in M_1.*

Proof. First, for x and y given, with $\dot{x}$ satisfying the conditions of Lemma 9.7.5, the conditions of Lemma 9.7.8 are satisfied. Since $\dot{x} \in D$ and $y \in D$, $z \in D^*$ for z defined in (7.19). Start by replacing z by its right-continuous version, which has the same graph. Invoking Lemma 9.4.1, for any $t > 0$, let $\tilde{z} \in D_c$ be such that $\|z - \tilde{z}\|_t \leq \delta_1$. Suppose that $x^{-1}(t_1)$ and $x^{-1}(t_2)$ are two successive discontinuity points of y (where $t_1, t_2 < t$), regarded as an element of D_c. Suppose that $y(s) = c > 0$ in $[x^{-1}(t_1), x^{-1}(t_2))$. Then, for any $\delta_2 > 0$, $z_\epsilon(s) \uparrow z(s)$ in $(t_1 + \delta_2, t_2 - \delta_2)$. Since z_ϵ and $\tilde{z} + \delta_1$ are both continuous in $(t_1 + \delta_2, t_2 - \delta_2)$, we can apply Dini's theorem to conclude that $z_\epsilon(s) \wedge \tilde{z}(s) - \delta_1$ converges uniformly to $\tilde{z}(s) - \delta_1$ in $(t_1 + \delta_2, t_2 - \delta_2)$. Similarly, if $y(s) = c < 0$ in $[t_1, t_2)$, then we can conclude that $z_\epsilon(s) \vee (\tilde{z}(s) + \delta_1)$ converges uniformly to $\tilde{z}(s) + \delta_1$ in $(t_1 + \delta_2, t_2 - \delta_2)$. It thus suffices to establish local M_1 convergence at each of the isolated discontinuity points of $\tilde{z}$; see Theorem **??**. However, z_ϵ is monotone in a neighborhood of each of these discontinuity points for all sufficiently small ϵ. Together with the pointwise convergence at all continuity points established in Theorem 9.7.2, this implies the required local M_1 convergence. To get the strengthened convergence to M_1, apply Theorem 13.6.3 in the book. ∎

The derivative result in Theorem 9.7.3 holds for arbitrary $y \in D$. By applying Theorem 9.7.1, we obtain a corresponding preservation result, but only under the extra condition of uniform convergence of $\epsilon^{-1}(x_\epsilon - x)$ to y as $\epsilon \downarrow 0$, which holds if $y \in C$.

Below let U be the topology on $D([0,\infty),\mathbb{R})$ of uniform convergence over compact subsets.

Corollary 9.7.1. *Suppose that* x_ϵ, $x \in E$. *Under the conditions of Theorem* 9.7.3, *if* $\|\epsilon^{-1}(x_\epsilon - x) - y\|_t \to 0$ *as* $\epsilon \downarrow 0$ *for all* $t > 0$, *then*

$$\epsilon^{-1}(x_\epsilon^{-1} - x^{-1}) \to z_+ \quad in \quad (D, M_1') \quad as \quad \epsilon \downarrow 0$$

for z_+ *as in Theorem* 9.7.3.

9.8. Chapter Notes

As indicated at the outset, this chapter is largely based on Mandelbaum and Massey (1995). They formulate convergence preservation in terms of the directional derivative. We focus on the second term of the triangle inequality in (2.1). Thus The results in Section 9.2 here are new. It would be nice if the upper bound $K\epsilon^{-1}d_1(x_\epsilon - x, y)$ in Theorem 9.2.3 could be replaced by $Kd_1(\epsilon^{-1}(x_\epsilon - x), y)$ under reasonable regularity conditions. The existing bound in terms of $K\epsilon^{-1}d_1(x_\epsilon - x, y)$ may be suitable for applying strong approximations. It thus also would be nice to develop such strong approximations to apply with Theorem 9.2.3 here.

Section 9.3 on the derivative of the supremum function is also based on Mandelbaum and Massey (1995). We provide extensions allowing the functions x and y appearing in $z_\epsilon(x,y)$ in (3.2) to be discontinuous. We also do not require that the limit z have only finitely many discontinuities in each finite interval. The arguments are quite a bit more complicated as a result. Some simplification is achieved here by exploiting approximations by piecewise-constant functions. In particular, for establishing M_1 convergence, Lemma 9.4.2 is key.

Given the intimate connection between the reflection and supremum maps, most of the work on the derivative of the reflection map in Section 9.5 is done in Sections 9.3 and 9.4. The application of Sections 9.3 – 9.5 in Section 9.6 to obtain heavy-traffic limits for nonstationary queues also follows Mandelbaum and Massey (1995). They focused on the $\mathrm{M_t/M_t/1}$ queue with fixed arrival-rate and service-rate functions $\lambda^+(t)$ and $\lambda^-(t)$, drawing on the strong approximation for Poisson processes. We show how the results can be generalized by applying convergence-preservation results for the composition function with nonlinear centering in Chapter 13 of the book.

Section 9.7 on the derivative of the inverse function is new. The M_1'' topology extends the M_1' topology introduced in Puhalskii and Whitt (1997).

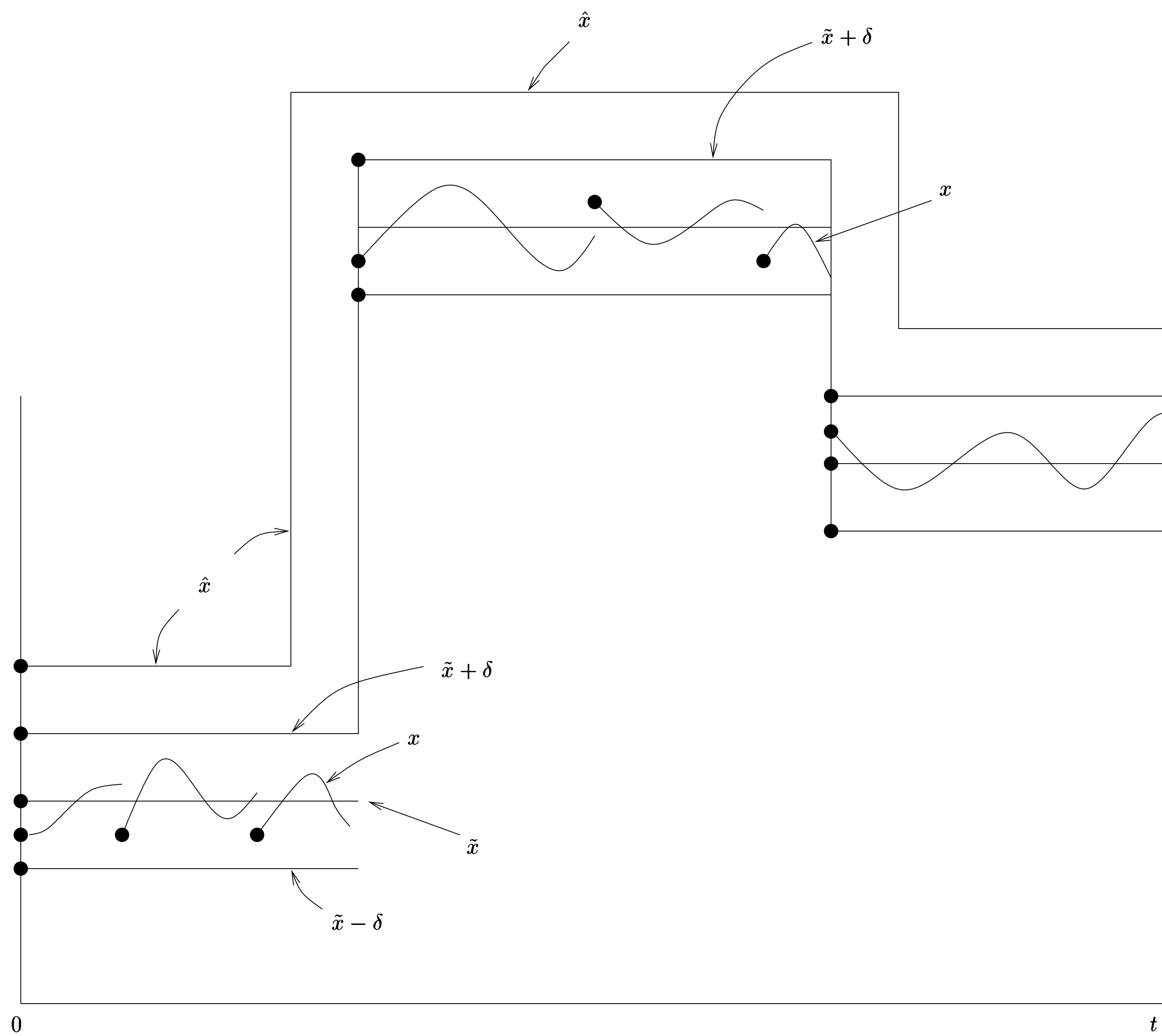

Figure 9.1: A possible function x, piecewise-constant approximation $\tilde{x}$, upper bound $\tilde{x}+\delta$ and upper boundary $\hat{x}$ of the δ-neighborhood of the graph $\Gamma_{\tilde{x}+\delta}$ used in the proof of Theorem 9.4.1.

Figure 9.2: Graphs of the time-dependent arrival-rate and service-rate functions $(\lambda^+(t), \lambda^-(t))$ with λ^- constant, the functions $(-x, (-x)^\uparrow)$, the set-valued function Φ_{-x} and the limits q and Q for a typical realization of the $\mathrm{M_t/M_t/1}$ queue.