Chapter 12

The Space D

12.1. Introduction

This chapter is devoted to the function space $D \equiv D([0,T],\mathbb{R}^k)$ with the Skorohod M_1 topology, expanding upon the introduction in Sections 3.3 and 11.5 and the classic paper by Skorohod (1956). We omit most proofs here. Many are provided in Chapter 6 of the Internet Supplement.

Here is how the present chapter is organized: We start in Section 12.2 by discussing regularity properties of the function space D. A key property, which we frequently use, is the fact that any function in D can be approximated uniformly closely by piecewise-constant functions with only finitely many discontinuities.

In Section 12.3 we introduce the strong and weak versions of the M_1 topology on $D([0,T],\mathbb{R}^k)$, referred to as SM_1 and WM_1, and establish basic properties. We also discuss the relation among the non-uniform Skorohod topologies on D. In Section 12.4 we discuss local uniform convergence at continuity points and relate it to oscillation functions used to characterize different forms of convergence.

In Section 12.5 we provide several different alternative characterizations of SM_1 and WM_1 convergence. Some involve parametric representations of the completed graphs and others involve oscillation functions. It is significant that there are forms of the oscillation-function characterizations that involve considering one function argument t at a time. Consequently, the examples in Figure 11.2 tend to be more than illustrative: The topologies are characterized by the local behavior in the neighborhood of single discontinuities.

In Section 12.6 we discuss conditions that allow us to strengthen the mode of convergence from WM_1 to SM_1. The key condition is to have the coordinate limit functions have no common discontinuities. In Section 12.7 we study how SM_1 convergence in $D([0,T],\mathbb{R}^k)$ can be characterized by associated limits of mappings.

In Section 12.8 we exhibit a complete metric topologically equivalent to the incomplete metric inducing the SM_1 topology introduced earlier. As with the J_1 metric d_{J_1} in (3.2) of Section 3.3, the natural M_1 metric is incomplete, but there exists a topologically equivalent complete metric, so that D with the SM_1 topology is Polish (metrizable as a complete separable metric space).

In Section 12.9 we discuss extensions of the SM_1 and WM_1 topologies on $D([0,T],\mathbb{R}^K)$ to corresponding spaces of functions with non-compact domains. The principal example of such a non-compact domain is the interval $[0,\infty)$, but $(0,\infty)$ and $(-\infty,\infty)$ also arise.

In Section 12.10 we introduce the strong and weak versions of the M_2 topology, denoted by SM_2 and WM_2. In Section 12.11 we provide alternative characterizations of these topologies and discuss additional properties.

Finally, in Section 12.12 we discuss characterizations of compact subsets of D using oscillation functions. These characterizations are useful because they lead to characterizations of tightness for sequences of probability measures on D, which is a principal way to establish weak convergence of the probability measures; see Section 11.6.

12.2. Regularity Properties of D

Let $D \equiv D^k \equiv D([0,T],\mathbb{R}^k)$ be the set of all $\mathbb{R}^k$-valued functions $x \equiv (x^1,\ldots,x^k)$ on $[0,T]$ that are right continuous at all $t \in [0,T)$ and have left limits at all $t \in (0,T]$: If $x \in D$, then

$$\text{for } 0 \le t < T, \quad x(t+) \equiv \lim_{s\downarrow t} x(s) \quad \text{exists with} \quad x(t+) = x(t)$$

and

$$\text{for } 0 < t \le T, \quad x(t-) \equiv \lim_{s\uparrow t} x(s) \quad \text{exists} .$$

However, with the M_1 topology, we will be working with the completed graphs of the functions, which are obtained by adding segments joining the left and right limits to the graph at each discontinuity point. Thus the actual value of the function at discontinuity points does not matter, provided that the function value falls appropriately between the left and right limits. Such functions are said to have *discontinuities of the first kind.* In Chapter 15 we consider more general functions.

We use superscripts to designate coordinate functions, so that subscripts can index different functions in D. For example, x_3^2 denotes the second coordinate function in $D([0,T],\mathbb{R}^1)$ of $x_3 \equiv (x_3^1,\ldots,x_3^k)$ in $D([0,T],\mathbb{R}^k)$, where x_3 is the third element of the sequence $\{x_n : n \geq 1\}$. Let C be the subset of continuous functions in D.

Let $\|\cdot\|$ be the maximum (or l_∞) norm on $\mathbb{R}^k$ and the *uniform norm* on D; i.e., for each $b \equiv (b^1,\ldots,b^k) \in \mathbb{R}^k$, let

$$\|b\| \equiv \max_{1\leq i\leq k} |b^i| \tag{2.1}$$

and, for each $x \equiv (x^1,\ldots,x^k) \in D([0,T],\mathbb{R}^k)$, let

$$\|x\| \equiv \sup_{0\leq t\leq T} \|x(t)\| = \sup_{0\leq t\leq T} \max_{1\leq i\leq k} |x^i(t)| \ . \tag{2.2}$$

The maximum norm on $\mathbb{R}^k$ in (2.1) is topologically equivalent to the l_p norm

$$\|b\|_p \equiv \left(\sum_{i=1}^{k} (b^i)^p\right)^{1/p} .$$

For $p = 2$, the l_p norm is the Euclidean (or l_2) norm. For $p = 1$, the l_p norm is the sum (or l_1) norm. The uniform norm on D induces the uniform metric on D.

We first discuss regularity properties of D due to the existence of limits. Let $Disc(x)$ be the set of discontinuities of x, i.e.,

$$Disc(x) \equiv \{t \in (0,T] : x(t-) \neq x(t)\} \tag{2.3}$$

and let $Disc(x,\epsilon)$ be the set of discontinuities of magnitude at least ϵ, i.e.,

$$Disc(x,\epsilon) \equiv \{t \in (0,T] : \|x(t-) - x(t)\| \geq \epsilon\} \ . \tag{2.4}$$

The following is a key regularity property of D.

Theorem 12.2.1. (the number of discontinuities of a given size) *For each $x \in D$ and $\epsilon > 0$, $Disc(x,\epsilon)$ is a finite subset of* $[0,T]$.

Corollary 12.2.1. (the number of discontinuities) *For each $x \in D$, $Disc(x)$ is either finite or countably infinite.*

We say that a function x in D is *piecewise-constant* if there are finitely many time points t_i such that $0 \equiv t_0 < t_1 < \cdots < t_{m-1} \le t_m \equiv T$ and x is constant on the intervals $[t_{i-1}, t_i)$, $1 \le i \le m-1$, and $[t_{m-1}, T]$. Let D_c be the subset of piecewise-constant functions in D. Let $v(x; A)$ be the *modulus of continuity* of the function x over the set A, defined by

$$v(x; A) \equiv \sup_{t_1, t_2 \in A} \{\|x(t_1) - x(t_2)\|\} \tag{2.5}$$

for $A \subseteq [0, T]$. The following is a second important regularity property of D.

Theorem 12.2.2. (approximation by piecewise-constant functions) *For each $x \in D$ and $\epsilon > 0$, there exists $x_c \in D_c$ such that $\|x - x_c\| < \epsilon$.*

We can deduce other useful consequences from Theorem 12.2.2.

Corollary 12.2.2. (oscillation function property) *For each $x \in D$ and $\epsilon > 0$, there exist finitely many points t_i with $0 \equiv t_0 < t_1 < \cdots < t_{m-1} \le t_m \equiv T$ such that $v(x, [t_{i-1}, t_i)) < \epsilon$, $1 \le i \le m-1$, and $v(x, [t_{m-1}, T]) < \epsilon$.*

Corollary 12.2.3. (boundedness) *Each x in D is bounded, i.e., $\|x\| < \infty$.*

Corollary 12.2.4. (measurability) *Each x in D is a Borel measurable real-valued function on $[0, T]$.*

12.3. Strong and Weak M_1 Topologies

In this section we define strong and weak versions of the M_1 topology on the function space $D([0, T], \mathbb{R}^k)$, denoted by SM_1 and WM_1. The strong topology agrees with the standard topology introduced by Skorohod (1956). The strong and weak topologies coincide when $k = 1$ but differ for $k > 1$. We will show that the weak topology coincides with the product topology.

We consider functions with domain $[0, T]$, but our results can be applied to non-compact domains such as $[0, \infty)$, if as is customary we understand $x_n \to x$ as $n \to \infty$ in $D([0, \infty), \mathbb{R}^k)$ to mean that the restrictions of x_n to $[0, T]$ converge to the restriction of x to $[0, T]$ for all T that are continuity points of x. We discuss $D([0, \infty), \mathbb{R}^k)$ further in Section 12.9.

12.3.1. Definitions

The strong and weak topologies will be based on different notions of a segment in $\mathbb{R}^k$. For $a \equiv (a^1, \ldots, a^k)$, $b \equiv (b^1, \ldots, b^k) \in \mathbb{R}^k$, let $[a,b]$ be the *standard segment*, i.e.,

$$[a,b] \equiv \{\alpha a + (1-\alpha)b : 0 \leq \alpha \leq 1\} \tag{3.1}$$

and let $[[a,b]]$ be the *product segment*, i.e.,

$$[[a,b]] \equiv \mathop{\times}_{i=1}^{k} [a^i, b^i] \equiv [a^1, b^1] \times \cdots \times [a^k, b^k] \ , \tag{3.2}$$

where the one-dimensional segment $[a^i, b^i]$ coincides with the closed interval $[a^i \wedge b^i, a^i \vee b^i]$, with $c \wedge d = \min\{c,d\}$ and $c \vee d = \max\{c,d\}$ for $c, d \in \mathbb{R}$. Note that $[a,b]$ and $[[a,b]]$ are both subsets of $\mathbb{R}^k$. If $a = b$, then $[a,b] = [[a,b]] = \{a\} = \{b\}$; if $a^i \neq b^i$ for one and only one i, then $[a,b] = [[a,b]]$. If $a \neq b$, then $[a,b]$ is always a one-dimensional line in $\mathbb{R}^k$, while $[[a,b]]$ is a j-dimensional subset, where j is the number of coordinates i for which $a^i \neq b^i$. Always, $[a,b] \subseteq [[a,b]]$.

Remark 12.3.1. *More general range spaces.* We may want to consider the space D with a more general range space than $\mathbb{R}^k$. Generalizations of the M topologies are restricted by the linear structure in the definition of segments in (3.1) and (3.2). However, we can extend the M topologies to Banach-space valued functions. We use that extension to treat the workload process in the infinite-server queue in Section 10.3. ∎

We now define completed graphs of the functions: For $x \in D$, let the (standard) *thin graph* of x be

$$\Gamma_x \equiv \{(z,t) \in \mathbb{R}^k \times [0,T] : z \in [x(t-), x(t)]\} , \tag{3.3}$$

where $x(0-) \equiv x(0)$ and let the *thick graph* of x be

$$\begin{aligned} G_x &\equiv \{(z,t) \in \mathbb{R}^k \times [0,T] : z \in [[x(t-), x(t)]]\} \\ &= \{(z,t) \in \mathbb{R}^k \times [0,T] : z^i \in [x^i(t-), x^i(t)] \quad \text{for each } i\} \end{aligned} \tag{3.4}$$

for $1 \leq i \leq k$. Since $[a,b] \subseteq [[a,b]]$ for all $a, b \in \mathbb{R}^k$, $\Gamma_x \subseteq G_x$ for each x.

We now define *order relations* on the graphs Γ_x and G_x. We say that $(z_1, t_1) \leq (z_2, t_2)$ if either (i) $t_1 < t_2$ or (ii) $t_1 = t_2$ and $|x^i(t_1-) - z_1^i| \leq$

$|x^i(t_1-) - z_2^i|$ for all i. The relation $\leq$ induces a total order on Γ_x and a partial order on G_x.

It is also convenient to look at the ranges of the functions. Let the *thin range* of x be the projection of Γ_x onto $\mathbb{R}^k$, i.e.,

$$\rho(\Gamma_x) \equiv \{z \in \mathbb{R}^k : (z,t) \in \Gamma_x \quad \text{for some} \quad t \in [0,T]\} \tag{3.5}$$

and let the *thick range* of x be the projection of G_x onto $\mathbb{R}^k$, i.e.,

$$\rho(G_x) \equiv \{z \in \mathbb{R}^k : (z,t) \in G_x \quad \text{for some} \quad t \in [0,T]\} \ . \tag{3.6}$$

Note that $(z,t) \in \Gamma_x$ (G_x) for some t if and only if $z \in \rho(\Gamma_x)$ $(\rho(G_x))$. Thus a pair (z,t) cannot be in a graph of x if z is not in the corresponding range.

We now define strong (standard) and weak parametric representations based on these two kinds of graphs. A *strong parametric representation* of x is a continuous nondecreasing function (u,r) mapping $[0,1]$ onto Γ_x. A *weak parametric representation* of x is a continuous nondecreasing function (u,r) mapping $[0,1]$ into G_x such that $r(0)=0$, $r(1)=T$ and $u(1)=x(T)$. (For the parametric representation, "nondecreasing" is with respect to the usual order on the domain $[0,1]$ and the order on the graphs defined above.) Here it is understood that $u \equiv (u^1,\ldots,u^k) \in C([0,1],\mathbb{R}^k)$ is the spatial part of the parametric representation, while $r \in C([0,1],[0,T])$ is the time (domain) part. Let $\Pi_s(x)$ and $\Pi_w(x)$ be the sets of strong and weak parametric representations of x, respectively. For real-valued functions x, let $\Pi(x) \equiv \Pi_s(x) = \Pi_w(x)$. Note that $(u,r) \in \Pi_w(x)$ if and only if $(u^i,r) \in \Pi(x^i)$ for $1 \leq i \leq k$.

We use the parametric representations to characterize the strong and weak M_1 topologies. As in (2.1) and (2.2), let $\|\cdot\|$ denote the supremum norms in $\mathbb{R}^k$ and D. We use the definition $\|\cdot\|$ in (2.2) also for the $\mathbb{R}^k$-valued functions u and r on $[0,1]$.

Now, for any $x_1, x_2 \in D$, let

$$d_s(x_1,x_2) \equiv \inf_{\substack{(u_j,r_j)\in\Pi_s(x_j)\\ j=1,2}} \{\|u_1-u_2\| \vee \|r_1-r_2\|\} \tag{3.7}$$

and

$$d_w(x_1,x_2) \equiv \inf_{\substack{(u_j,r_j)\in\Pi_w(x_j)\\ j=1,2}} \{\|u_1-u_2\| \vee \|r_1-r_2\|\} \ . \tag{3.8}$$

Note that $\|u_1-u_2\| \vee \|r_1-r_2\|$ can also be written as $\|(u_1,r_1)-(u_2,r_2)\|$, due to definitions (2.1) and (2.2). Of course, when the range is $\mathbb{R}$, $d_s = d_w = d_{M_1}$ for d_{M_1} defined in (3.4) in Section 3.3.

We say that $x_n \to x$ in D for a sequence or net $\{x_n\}$ in the SM_1 (WM_1) topology if $d_s(x_n, x) \to 0$ ($d_w(x_n, x) \to 0$) as $n \to \infty$. We start with the following basic result.

12.3.2. Metric Properties

Theorem 12.3.1. (metric inducing SM_1) *d_s is a metric on D.*

Proof. Only the triangle inequality is difficult. By Lemma 12.3.2 below, for any $\epsilon > 0$, a common parametric representation $(u_3, r_3) \in \Pi_s(x_3)$ can be used to obtain

$$\|u_1 - u_3\| \vee \|r_1 - r_3\| < d_s(x_1, x_3) + \epsilon$$

and

$$\|u_2 - u_3\| \vee \|r_2 - r_3\| < d_s(x_1, x_3) + \epsilon$$

for some $(u_1, r_1) \in \Pi_s(x_1)$ and $(u_2, r_2) \in \Pi_s(x_2)$. Hence

$$d_s(x_1, x_2) \leq \|u_1 - u_2\| \vee \|r_1 - r_2\| \leq d_s(x_1, x_3) + d_s(x_3, x_2) + 2\epsilon \ .$$

Since ϵ was arbitrary, the proof is complete. ■

To prove Theorem 12.3.1, we use finite approximations to the graphs Γ_x. We first define an order-consistent distance between a graph and a finite subset. We use the notion of a finite ordered subset.

Definition 12.3.1. (order-consistent distance) *For $x \in D$, let A be a finite ordered subset of the ordered graph $(\Gamma_x, \leq)$, i.e., for some $m \geq 1$, A contains $m+1$ points (z_i, t_i) from Γ_x such that*

$$(x(0), 0) \equiv (z_0, t_0) \leq (z_1, t_1) \leq \cdots \leq (z_m, t_m) \equiv (x(T), T) \ . \tag{3.9}$$

The order-consistent distance between A and Γ_x is

$$\hat{d}(A, \Gamma_x) \equiv \sup\{\|(z, t) - (z_i, t_i)\| \vee \|(z, t) - (z_{i+1}, t_{i+1})\|\} \ , \tag{3.10}$$

where the supremum is over all $(z_i, t_i) \in A$, $1 \leq i \leq m-1$, and all $(z, t) \in \Gamma_x$ such that

$$(z_i, t_i) \leq (z, t) < (z_{i+1}, t_{i+1}) \ ,$$

using the order on the graph. ■

We now observe that finite ordered subsets A can be chosen to make $\hat{d}(A, \Gamma_x)$ arbitrarily small. The missing proofs are in the Internet Supplement.

Lemma 12.3.1. (finite approximations to graphs) *For any $x \in D$ and $\epsilon > 0$, there exists a finite ordered subset A of Γ_x such that $\hat{d}(A, \Gamma_x) < \epsilon$ for $\hat{d}$ in* (3.10).

To complete the proof of Theorem 12.3.1, we need the following result, which we prove by applying Lemma 12.3.1.

Lemma 12.3.2. (flexibility in choice of parametric representations) *For any $x_1, x_2 \in D$, $(u_1, r_1) \in \Pi_s(x_1)$ and $\epsilon > 0$, it is possible to find $(u_2, r_2) \in \Pi_s(x_2)$ such that*

$$\|u_1 - u_2\| \vee \|r_1 - r_2\| \le d_s(x_1, x_2) + \epsilon \ .$$

We will show that the metric d_s induces the standard M_1 topology defined by Skorohod (1956); see Theorem 12.5.1. Since $\Pi_s(x) \subseteq \Pi_w(x)$ for all x, we have $d_w(x_1, x_2) \le d_s(x_1, x_2)$ for all x_1, x_2, so that the WM_1 topology is indeed weaker than the SM_1 topology. However, we show below in Example 12.3.2 that d_w in (3.8) is *not* a metric when $k > 1$.

For $x_1, x_2 \in D([0, T], \mathbb{R}^k)$, let d_p be a metric inducing the product topology, defined by

$$d_p(x_1, x_2) \equiv \max_{1 \le i \le k} d(x_1^i, x_2^i) \tag{3.11}$$

for $x_j \equiv (x_j^1, \ldots, x_j^k)$ and $j = 1, 2$. (Note that $d_s = d_w = d_p$ when the functions are real valued, in which case we use the notation d.) It is an easy consequence of (3.8), (3.11) and the second representation in (3.4) that the WM_1 topology is stronger than the product topology, i.e., $d_p(x_1, x_2) \le d_w(x_1, x_2)$ for all $x_1, x_2 \in D$. In Section 12.5 we will show that actually the WM_1 and product topologies coincide.

We now show that SM_1 is strictly stronger than WM_1. Let I_A denote the indicator function of a set A; i.e., $I_A(t) = 1$ if $t \in A$ and $I_A(t) = 0$ otherwise.

Example 12.3.1. *WM_1 convergence without SM_1 convergence.* To show that we can have $d_w(x_n, x) \to 0$ as $n \to \infty$ without $d_s(x_n, x) \to 0$ as $n \to \infty$, let $x \equiv (x^1, x^2) \in D([0, 2], \mathbb{R}^2)$ be defined by $x^1 = x^2 = 2I_{[1,2]}$ and let $x_n^1 = 2I_{[1-n^{-1},2]}$ and $x_n^2 = I_{[1-n^{-1},1)} + 2I_{[1,2]}$. The thin range of x is the set $\{(0,0), (2,2)\}$ plus the line segment $[(0,0), (2,2)]$ connecting those two points, while the thin range of x_n is the set $\{(0,0),(2,1),(2,2)\}$ plus the line segments $[(0,0),(2,1)]$ and $[(2,1),(2,2)]$. Since $(2,1) \in \Gamma_{x_n}$ for all n but $(2,1) \notin \Gamma_x$, we must have $d_s(x_n, x) \not\to 0$ as $n \to \infty$. On the other hand, the thick ranges of x and x_n, $n \geq 1$ all are $[0,2] \times [0,2]$. To demonstrate that $d_w(x_n, x) \to 0$ as $n \to \infty$, we construct suitable parametric representations. Let

$$r(0) = 0,\ r(1/3) = 1 = r(2/3),\ r(1) = 2$$

$$r_n(0) = 0,\ r_n(1/3) = 1 - n^{-1} = r_n((1-n^{-1})/2),$$

$$r_n((1+n^{-1})/2) = 1 = r_n(2/3),\ r_n(1) = 2$$

$$u^1(0) = 0 = u^1(1/3),\ u^1(1/2) = 2 = u^1(1)$$

$$u_n^1(0) = 0 = u_n^1(1/3),\ u_n^1((1-n^{-1})/2) = 2 = u_n^1(1)$$

$$u^2(0) = 0 = u^2(1/3),\ u^2(1/2) = 1,\ u^2(2/3) = 2 = u^2(1)$$

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

$$u_n^2(2/3) = 2 = u_n^2(1)$$

with r, r_n, u^1, u_n^1, u^2, u_n^2 defined by linear interpolation in the gaps. With this construction, $(u_n, r_n) \in \Pi_w(x_n)$ and $(u, r) \in \Pi_w(x)$, $\|r_n - r\| = n^{-1}$, $\|u_n^1 - u^1\| = 6n^{-1}$ and $\|u_n^2 - u^2\| = 3n^{-1}$. Hence,

$$d_w(x_n, x) \leq \|u_n - u\| \vee \|r_n - r\| = 6n^{-1} \to 0 \quad \text{as} \quad n \to \infty\ . \quad \blacksquare$$

Example 12.3.2. *d_w is not a metric.* We now show that d_w in (3.8) is not a metric. For this purpose, we use a minor modification of Example 12.3.1. Let $x^1 = x^2 = 2I_{[1,2]}$ as before. For even n, let $x_n^1 = 2I_{[1-n^{-1},2]}$ and $x_n^2 = I_{[1-n^{-1},1)} + 2I_{[1,2]}$ as before. Then let $x_{2n+1}^1 = x_{2n}^2$ and $x_{2n+1}^2 = x_{2n}^1$. We show that $d_w(x_{2n}, x_{2n+1}) \not\to 0$ as $n \to \infty$ even though $d_w(x_n, x) \to 0$ as $n \to \infty$, contradicting the triangle inequality property of a metric. The thick range of x_n is $([0,2] \times [0,1]) \cup (\{2\} \times [1,2])$ for n even and $([0,1] \times [0,2]) \cup ([1,2] \times \{2\})$ for n odd. The points $(2,1)$ and $(1,2)$ appear for n even and odd, respectively, but are distance 1 from the other thick range. Any parametric representation must pass through $(2,1,1-n^{-1})$ in $\mathbb{R}^2 \times [0,2]$ for n even and $(1,2,1-n^{-1})$ for n odd. However, for n odd (n even) all points on G_{x_n} are at least a distance 1 from $(2,1,1-n^{-1})$ $((1,2,1-n^{-1}))$. This example shows that we cannot find a constant K such that $d_w(x_1, x_2) \leq K d_p(x_1, x_2)$ for all $x_1, x_2 \in D$. $\blacksquare$

We now relate the metrics $d_{M_1} \equiv d_s$ and d_{J_1} for d_{J_1} in (3.2) of Section 3.3.

Theorem 12.3.2. (comparison of J_1 and M_1 metrics) *For each* $x_1, x_2 \in D$,

$$d_s(x_1, x_2) \le d_{J_1}(x_1, x_2) \ .$$

Remark 12.3.2. *Uses of the* M_1 *topology.* The M_1 topology has not been used extensively. It was used by Whitt (1971b, 1980, 2000b), Wichura (1974), Avram and Taqqu (1989, 1992), Kella and Whitt (1990), Chen and Whitt (1993), Mandelbaum and Massey (1995), Harrison and Williams (1996), Puhalskii and Whitt (1997, 1998), Resnick and van der Berg (2000), O'Brien (2000) and no doubt a few others. ■

12.3.3. Properties of Parametric Representations

We conclude this section by further discussing strong parametric representations. We first indicate how to construct a parametric representation (u, r) of Γ_x for any $x \in D$.

Remark 12.3.3. *How to construct a parametric representation.* Let t_j, $j \ge 1$, be a list of the discontinuity points of x (of which there are finitely or countably infinite many). For each j, select a subinterval $[a_j, b_j] \subseteq [0, 1]$ and let $r(s) = t_j$ for $a_j \le s \le b_j$, $u(a_j) = x(t_j-)$, $u(b_j) = x(t_j)$ and $u(\alpha a_j + (1-\alpha)b_j) = \alpha u(a_j) + (1-\alpha)u(b_j)$, $0 < \alpha < 1$. For successive discontinuities, do this in an order-preserving way; i.e., if $t_i < t_j < t_k$, then we require that $b_i < a_j < b_j < a_k$. Let this be done for all j. Next, suppose that t is not a discontinuity point but is the limit of discontinuity points. If $t_j \downarrow t$ as $j \to \infty$ where $t_j \in Disc(x)$, then let $r(a) = t$ and $u(a) = \lim_{j\to\infty} x(t_j-)$, where $a = \lim_{j\to\infty} a_j$ with $r(a_j) = t_j$. Similarly, if $t_j \uparrow t$ as $j \to \infty$ where $t_j \in Disc(x)$, then let $r(b) = t$ and $u(b) = \lim_{j\to\infty} x(t_j)$, where $b = \lim_{j\to\infty} b_j$ with $r(b_j) = t_j$. Finally, there may remain open intervals (a, b) over which (u, r) is undefined. Since (u, r) is already defined at the endpoints a and b, let $r(\alpha a + (1-\alpha)b) = \alpha r(a) + (1-\alpha)r(b)$ and $u(\alpha a + (1-\alpha)b) = x(r(\alpha a + (1-\alpha)b)$ for $0 < \alpha < 1$. This construction makes (u, r) a one-to-one function. This construction also makes r a generalization of piecewise linear; i.e., there are finite or countably many subintervals $[a_j, b_j]$ over which r is constant and there are finite or countably many intervals (b_k, a_k) over which r is linear. The union of all those points (where r is constant or linear) is dense in $[0, 1]$. The function r is extended to all other points by continuity. ■

Remark 12.3.4. *Parametric representations need not be one-to-one.* We do not require that a parametric representation be a one-to-one function. For example, even if x is continuous at t, we could have $r(s) = t$ for $a \le s \le b$.

Then, necessarily, $u(s) = x(t)$, $a \leq s \leq b$. However, we get the same metric if the parametric representations (u, r) are required to be one-to-one with r nondecreasing, e.g., as done by Wichura (1974); see Remark 12.5.2 in Section 5. Skorohod (1956) only originally required that r be nondecreasing instead of (u, r), without the one-to-one property, in his Definitions 2.2.4 and 2.2.5. However, from his remarks after 2.2.5, it is evident that he meant to require that (u, r) be nondecreasing as we have defined it. As stated, Skorohod's version of the M_1 topology with only r nondecreasing is actually the M_2 topology. ∎

Example 12.3.3. *Need for monotonicity.* To see the importance of requiring that the parametric representation be nondecreasing, using the order on the graphs, let $x = I_{[1,2]}$, $x_n(1) = x_n(1 - 2n^{-1}) = x_n(2) = 1$ and $x_n(0) = x_n(1 - 3n^{-1}) = x_n(1 - n^{-1}) = 0$, with x_n defined by linear interpolation elsewhere. For these functions, $x_n \to x$ as $n \to \infty$ in the M_2 topology but not in the M_1 topology. If we did not require that parametric representations of x be nondecreasing in our M_1 definitions, then we would have $x_n \to x$ as $n \to \infty$ in the M_1 topology. To see this, we exhibit parametric representations. Let $u_n = u$, $n \geq 4$, and let

$$\begin{aligned}
&r(0) = 0,\ r(1/5) = r(4/5) = 1,\ r(1) = 2 \\
&u(0) = u(1/5) = u(3/5) = 0,\ u(2/5) = u(4/5) = u(2) = 1 \\
&r_n(0) = 0,\ r_n(1/5) = 1 - 3n^{-1},\ r_n(2/5) = 1 - 2n^{-1}, \\
&r_n(3/5) = 1 - n^{-1},\ r_n(4/5) = 1,\ r_n(1) = 2
\end{aligned}$$

with r, u, r_n and u_n defined by linear interpolation elsewhere. It is easy to see that $(u_n, r_n) \in \Pi_s(x_n)$, and $\|(u_n, r_n) - (u, r)\| = \|r_n - r\| = 3n^{-1}$, but $(u, r) \notin \Pi_s(x)$ because (u, r) fails to be nondecreasing, since it backtracks on the graph at $t = 1$. If r were only required to be nondecreasing, then we would have $(u, r) \in \Pi_s(x)$. ∎

We now continue characterizing parametric representations. For $x \in D$, $t \in Disc(x)$ and $(u, r) \in \Pi_s(x)$, there exists a unique pair of points $s_l \equiv s_l(t, x)$ and $s_r \equiv s_r(t, x)$ such that $s_l < s_r$ and $r^{-1}(\{t\}) = [s_l, s_r]$, i.e.,

$$\begin{aligned}
&\text{(i) } r(s) < t \text{ for } s < s_l \\
&\text{(ii) } r(s) = t \text{ for } s_l \leq s \leq s_r \\
&\text{(iii) } r(s) > t \text{ for } s > s_r\ .
\end{aligned} \tag{3.12}$$

We will exploit the fact that a parametric representation (u, r) in $\Pi_s(x)$ is *jump consistent*: for each $t \in Disc(x)$ and pair $s_l \equiv s_l(t, x) < s_r \equiv$

$s_r(t,x)$ such that (3.12) holds, there is a continuous nondecreasing function β_t mapping $[0,1]$ onto $[0,1]$ such that

$$u(s) = \beta_t\left(\frac{s-s_l}{s_r-s_l}\right)u(s_r) + \left[1-\beta_t\left(\frac{s-s_l}{s_r-s_l}\right)\right]u(s_l) \quad \text{for} \quad s_l \le s \le s_r \ . \tag{3.13}$$

Condition (3.13) means that u is defined within jumps by interpolation from the definition at the endpoints s_l and s_r, consistently over all coordinates. In particular, suppose that $t \in Disc(x^i)$. (Since $t \in Disc(x)$, we must have $t \in Disc(x^i)$ for some coordinate i.) Suppose that $x^i(t-) < x^i(t)$. Then we can let

$$\beta_t(s) = \frac{u^i(s)-u^i(s_l)}{u^i(s_r)-u^i(s_l)} \ . \tag{3.14}$$

We see that (3.13) and (3.14) are consistent in that

$$u^i(s) = \beta_t\left(\frac{s-s_l}{s_r-s_l}\right)u^i(s_r) + \left[1-\beta_t\left(\frac{s-s_l}{s_r-s_l}\right)\right]u^i(s_l) \tag{3.15}$$

for β_t in (3.14). For another coordinate j, (3.13) and (3.14) imply that

$$u^j(s) = \left(\frac{u^i(s)-u^i(s_l)}{u^i(s_r)-u^i(s_l)}\right)u^j(s_r) + \left(\frac{u^i(s_r)-u^i(s)}{u^i(s_r)-u^i(s_l)}\right)u^j(s_l) \ . \tag{3.16}$$

It is possible that $t \notin Disc(x^j)$, in which case $u^j(s) = u^j(s_l) = u^j(s_r)$ for all s, $s_l \le s \le s_r$.

We can further characterize the behavior of a strong parametric representation at a discontinuity point. For $x \in D$, $t \in Disc(x)$ and $(u,r) \in \Pi_s(x)$, there exists a unique set of four points $s_l \equiv s_l(t,x) \le s_l' \equiv s_l'(t,x) < s_r' \equiv s_r'(t,x) \le s_r \equiv s_r(t,x)$ such that (3.12) holds and

$$\begin{aligned}
&\text{(i) } u(s) = u(s_l) \text{ for } s_l \le s \le s_l', \\
&\text{(ii) for each } i, \text{ either } u^i(s_l) < u^i(s) < u^i(s_r), \\
&\qquad \text{or } u^i(s_l) > u^i(s) > u^i(s_r) \text{ for } s_l' < s < s_r', \\
&\text{(iii) } u(s) = u(s_r) \text{ for } s_r' \le s \le s_r \ .
\end{aligned} \tag{3.17}$$

Let D_1 be the subset of D containing functions all of whose jumps occur in only one coordinate, i.e., the set of x such that, for each $t \in Disc(x)$ there exists one and only one $i \equiv i(t)$ such that $t \in Disc(x^i)$. (The coordinate i may depend on t.)

Lemma 12.3.3. (strong and weak parametric representations coincide on D_1) *For each $x \in D_1$, $\Pi_s(x) = \Pi_w(x)$.*

We now show that parametric representations are preserved under linear functions of the coordinates when $x \in \Pi_s(x)$. That is *not* true in $\Pi_w(x)$.

Lemma 12.3.4. (linear functions of parametric representations) *If* $(u, r) \in \Pi_s(x)$, *then* $(\eta u, r) \in \Pi_s(\eta x)$ *for any* $\eta \in \mathbb{R}^k$.

12.4. Local Uniform Convergence at Continuity Points

In this section we provide alternative characterizations of local uniform convergence at continuity points of a limit function. The non-uniform Skorohod topologies on D all imply local uniform convergence at continuity points of a limit function. They differ by their behavior at discontinuity points.

We first observe that pointwise convergence is weaker than local uniform convergence.

Example 12.4.1. *Pointwise convergence is weaker than local uniform convergence.* To see that pointwise convergence in D at all continuity points of the limit is strictly weaker than local uniform convergence at continuity points of the limit, let $x(t) = 0$, $0 \le t \le 2$, and $x_n = I_{[1+n^{-1}, 1+2n^{-1})}$, $n \ge 1$. Then $x_n(t) \to x(t) = 0$ as $n \to \infty$ for all t, but $x_n(1 + n^{-1}) = 1 \not\to 0$ as $n \to \infty$, so we do not have local uniform convergence at $t = 1$. We also do not have $x_n \to x$ as $n \to \infty$ in D in any of the Skorohod topologies. ■

We start by defining two basic *uniform-distance functions.*

For $x_1, x_2 \in D$, $t \in [0, T]$ and $\delta > 0$, let

$$u(x_1, x_2, t, \delta) \equiv \sup_{0 \vee (t-\delta) \le t_1 \le (t+\delta) \wedge T} \{\|x_1(t_1) - x_2(t_1)\|\} , \tag{4.1}$$

$$v(x_1, x_2, t, \delta) \equiv \sup_{0 \vee (t-\delta) \le t_1, t_2 \le (t+\delta) \wedge T} \{\|x_1(t_1) - x_2(t_2)\|\} , \tag{4.2}$$

We also define an *oscillation function.* For $x \in D$, $t \in [0, T]$ and $\delta > 0$, let

$$\bar{v}(x, t, \delta) \equiv \sup_{0 \vee (t-\delta) \le t_1 \le t_2 \le (t+\delta) \wedge T} \{\|x(t_1) - x(t_2)\|\} . \tag{4.3}$$

We next define oscillation functions that we will use with the M_1 topologies. They use the distance $\|z - A\|$ between a point z and a subset A in $\mathbb{R}^k$ defined in (5.3) in Section 11.5. The SM_1 and WM_1 topologies use the standard and product segments in (3.1) and (3.2). For each $x \in D$, $t \in [0,T]$ and $\delta > 0$, let

$$w_s(x,t,\delta) \equiv \sup_{0\vee(t-\delta)\leq t_1<t_2<t_3\leq(t+\delta)\wedge T} \{\|x(t_2) - [x(t_1), x(t_3)]\| \tag{4.4}$$

and

$$w_w(x,t,\delta) \equiv \sup_{0\vee(t-\delta)\leq t_1<t_2<t_3\leq(t+\delta)\wedge T} \{\|x(t_2) - [[x(t_1), x(t_3)]]\| \tag{4.5}$$

We now turn to the M_2 topology, which we will be studying in Sections 12.10 and 12.11. We define two uniform-distance functions. We use $\bar{w}$ as opposed to w to denote an M_2 uniform-distance function. Just as with the M_1 topologies, the SM_2 and WM_2 topologies use the standard and product segments in (3.1) and (3.2). For x_1, $x_2 \in D$, let

$$\bar{w}_s(x_1,x_2,t,\delta) \equiv \sup_{0\vee(t-\delta)\leq t_1\leq(t+\delta)\wedge T} \{\|x_1(t_1) - [x_2(t-), x_2(t)]\|\} \tag{4.6}$$

$$\bar{w}_w(x_1,x_2,t,\delta) \equiv \sup_{0\vee(t-\delta)\leq t_1\leq(t+\delta)\wedge T} \{\|x_1(t_1) - [[x_2(t-), x_2(t)]]\|\} \tag{4.7}$$

It is easy to establish the following relations among the uniform-distance and oscillation functions.

Lemma 12.4.1. (inequalities for uniform-distance and oscillation functions) *For all $x, x_n \in D$, $t \in [0,T]$ and $\delta > 0$,*

$$u(x_n,x,t,\delta) \leq v(x_n,x,t,\delta) \leq u(x_n,x,t,\delta) + \bar{v}(x,t,\delta) \ ,$$

$$w_w(x_n,t,\delta) \leq w_s(x_n,t,\delta) \leq \bar{v}(x_n,t,\delta) \leq 2v(x_n,x,t,\delta) + \bar{v}(x,t,\delta) \ ,$$

$$\bar{w}_w(x_n,x,t,\delta) \leq \bar{w}_s(x_n,x,t,\delta) \leq v(x_n,x,t,\delta) \leq 2\bar{w}_w(x_n,x,t,\delta) + \bar{v}(x,t,\delta) \ .$$

Since the M_1-oscillation functions $w_s(x_n,t,\delta)$ and $w_w(x_n,t,\delta)$ do not contain the limit x, their convergence to 0 as $n \to \infty$ and then $\delta \downarrow 0$ does not directly imply local uniform convergence at a continuity point of a prospective limit function x.

Example 12.4.2. *Characterizations of local uniform convergence at continuity points.* We show that it is possible to have $t \notin Disc(x)$, $x_n(t) \to x(t)$ as $n \to \infty$ and

$$\lim_{\delta\downarrow 0} \overline{\lim_{n\to\infty}} \ w_s(x_n,t,\delta) = 0$$

without having

$$\lim_{\delta\downarrow 0} \overline{\lim_{n\to\infty}} \; v(x_n, x, t, \delta) = 0\,.$$

That occurs for $t = 1$ when $x(t) = 0$, $0 \le t \le 2$, and $x_n = I_{[1+n^{-1},2]}$, $n \ge 1$. In this example, we have $\bar{v}(x, t, \delta) = 0$ and $w_s(x_n, t, \delta) = 0$ for all n, t and $\delta > 0$, but $v(x_n, x, 1, \delta) = 1$ for $n > 1/\delta$. ∎

We relate convergence of $w_s(x)n, t, \delta)$ and $w_w(x_n, t, \delta)$ to 0 as $n \to \infty$ and $\delta \downarrow 0$ to local uniform convergence by requiring pointwise convergence in a neighborhood of t; see (vi) in Theorem 12.4.1 below.

Theorem 12.4.1. (characterizations of local uniform convergence at continuity points) *If $t \notin Disc(x)$, then the following are equivalent:*

$$(i) \qquad \lim_{\delta\downarrow 0} \overline{\lim_{n\to\infty}} \; u(x_n, x, t, \delta) = 0 \;, \tag{4.8}$$

$$(ii) \qquad \lim_{\delta\downarrow 0} \overline{\lim_{n\to\infty}} \; v(x_n, x, t, \delta) = 0 \;, \tag{4.9}$$

$$(iii) \qquad \lim_{\delta\downarrow 0} \overline{\lim_{n\to\infty}} \; \bar{w}_s(x_n, x, t, \delta) = 0 \;, \tag{4.10}$$

$$(iv) \qquad \lim_{\delta\downarrow 0} \overline{\lim_{n\to\infty}} \; \bar{w}_w(x_n, x, t, \delta) = 0 \;, \tag{4.11}$$

(v) $x_n(t_1) \to x(t_1)$ for all t_1 in a dense subset of a neighborhood of t (including 0 if $t = 0$ or T if $t = T$) and

$$\lim_{\delta\downarrow 0} \overline{\lim_{n\to\infty}} \; w_s(x_n, t, \delta) = 0 \;,$$

(vi) $x_n(t_1) \to x(t_1)$ for all t_1 in a dense subset of a neighborhood of t (including 0 if $t = 0$ or T if $t = T$) and

$$\lim_{\delta\downarrow 0} \overline{\lim_{n\to\infty}} \; w_w(x_n, t, \delta) = 0 \;. \tag{4.12}$$

We now show that local uniform convergence at all points in a compact interval implies uniform convergence over the compact interval.

Lemma 12.4.2. (local uniform convergence everywhere in a compact interval) *If (4.8) holds for all $t \in [a, b]$, then*

$$\lim_{\delta\downarrow 0} \overline{\lim_{n\to\infty}} \sup_{0\vee(a-\delta)\le t\le(b+\delta)\wedge T} \{\|x_n(t) - x(t)\|\} = 0 \;.$$

12.5. Alternative Characterizations of M_1 Convergence

We now give alternative characterizations of SM_1 and WM_1 convergence.

12.5.1. SM_1 Convergence

We first give several alternative characterizations of SM_1-convergence (or, equivalently, d_s-convergence) in D, one being a minor variant of the original one involving an oscillation function established by Skorohod (1956). Another one – (v) below – involves only the local behavior of the functions. It helps us establish sufficient conditions to have $d_s((x_n, y_n), (x, y)) \to 0$ in $D([0,T], \mathbb{R}^{k+l})$ when $d_s(x_n, x) \to 0$ in $D([0,T], \mathbb{R}^k)$ and $d_s(y_n, y) \to 0$ in $D([0,T], \mathbb{R}^l)$; see Section 12.6. For the SM_1 topology, we define another oscillation function. For any $x_1, x_2 \in D$ and $\delta > 0$, let

$$w_s(x, \delta) \equiv \sup_{0 \le t \le T} w_s(x, t, \delta) \ , \tag{5.1}$$

for $w_s(x, t, \delta)$ in (4.4). We include the proof here, except for the supporting lemmas, which are proved in the Internet Supplement.

Theorem 12.5.1. (characterizations of SM_1 convergence) *The following are equivalent characterizations of convergence $x_n \to x$ as $n \to \infty$ in (D, SM_1):*

(i) For any $(u, r) \in \Pi_s(x)$, there exists $(u_n, r_n) \in \Pi_s(x_n)$, $n \ge 1$, such that

$$\|u_n - u\| \vee \|r_n - r\| \to 0 \quad as \quad n \to \infty \ . \tag{5.2}$$

(ii) There exist $(u, r) \in \Pi_s(x)$ and $(u_n, r_n) \in \Pi_s(x_n)$ for $n \ge 1$ such that (5.2) holds.

(iii) $d_s(x_n, x) \to 0$ as $n \to \infty$; i.e., for all $\epsilon > 0$ and all sufficiently large n, there exist $(u, r) \in \Pi_s(x)$ and $(u_n, r_n) \in \Pi_s(x_n)$ such that

$$\|u_n - u\| \vee \|r_n - r\| < \epsilon \ .$$

(iv) $x_n(t) \to x(t)$ as $n \to \infty$ for each t in a dense subset of $[0, T]$ including 0 and T, and

$$\lim_{\delta \downarrow 0} \ \overline{\lim_{n \to \infty}} \ w_s(x_n, \delta) = 0 \tag{5.3}$$

for $w_s(x, \delta)$ in (5.1) and $w_s(x, t, \delta)$ in (4.4).

(v) $x_n(T) \to x(T)$ *as* $n \to \infty$*; for each* $t \notin Disc(x)$,

$$\lim_{\delta \downarrow 0} \overline{\lim_{n \to \infty}} \; v(x_n, x, t, \delta) = 0 \tag{5.4}$$

for $v(x_1, x_2, t, \delta)$ *in* (4.2)*; and, for each* $t \in Disc(x)$,

$$\lim_{\delta \downarrow 0} \overline{\lim_{n \to \infty}} \; w_s(x_n, t, \delta) = 0 \tag{5.5}$$

for $w_s(x, t, \delta)$ *in* (4.4).

(vi) *For all* $\epsilon > 0$, , *there exist integers* m *and* n_1*, a finite ordered subset* A *of* Γ_x *of cardinality* m *as in* (3.9) and, for all $n \geq n_1$, finite ordered subsets A_n of Γ_{x_n} of cardinality m *such that, for all* $n \geq n_1$, $\hat{d}(A, \Gamma_x) < \epsilon$, $\hat{d}(A_n, \Gamma_{x_n}) < \epsilon$ *for* $\hat{d}$ *in* (3.10) *and* $d^*(A, A_n) < \epsilon$*, where*

$$d^*(A, A_n) \equiv \max_{1 \leq i \leq m} \{\|(z_i, t_i) - (z_{n,i}, t_{n,i})\| : (z_i, t_i) \in A, (z_{n,i}, t_{n,i}) \in A_n\} . \tag{5.6}$$

In preparation for the proof of Theorem 12.5.1, we establish some preliminary results. We first show that SM_1 convergence implies local uniform convergence at all continuity points.

Lemma 12.5.1. (local uniform convergence) *If* $d_s(x_n, x) \to 0$ *as* $n \to \infty$*, then* (4.9) *holds for each* $t \notin Disc(x)$.

We next relate the modulus w_s applied to x and the modulus applied to corresponding points on the graph Γ_x. The following lemma is established in the proof of Skorohod's (1956) 2.4.1.

Lemma 12.5.2. (extending the modulus from a function to its graph) *If* (z_1, t_1), (z_2, t_2), $(z_3, t_3) \in \Gamma_x$ *with* $0 \vee (t - \delta) \leq t_1 < t_2 < t_3 \leq (t + \delta) \wedge T$*, then* $\|z_2 - [z_1, z_3]\| \leq w_s(x, \delta)$.

Lemma 12.5.3. (asymptotic negligibility of the modulus) *For any* $x \in D$, $w_s(x, \delta) \downarrow 0$ *as* $\delta \downarrow 0$.

Proof of Theorem 12.5.1. The implications (i)→(ii)→(iii) are trivial. We establish the others exploiting transitivity.

(iii)→(iv). First, the convergence $x_n(T) \to x(T)$ is assumed directly. Next, by Lemma 12.5.1, if $d_s(x_n, x) \to 0$ as $n \to \infty$, then $x_n(t) \to x(t)$ for all $t \in Disc(x)^c$, which is a dense subset of $[0,T]$. We now want to show that, for any $\epsilon > 0$, there exists n_0 and δ such that $w_s(x_n, \delta) < \epsilon$ for all $n \geq n_0$. For $x \in D$ and $\epsilon > 0$ given, start by choosing η so that $w_s(x, \eta) < \epsilon/2$, which we can do by Lemma 12.5.3. Then apply (iii) to choose n_0 so that $(u_n, r_n) \in \Pi_s(x_n)$, $(u, r) \in \Pi_s(x)$ and

$$\|u_n - u\| \vee \|r_n - r\| < (\epsilon \wedge \eta)/4 \quad \text{for} \quad n \geq n_0 \ .$$

Suppose that $(t - \delta) \vee 0 \leq t_1 < t_2 \leq t_3 < (t + \delta) \wedge T$. Let $s_{n,i}$ be such that $r_n(s_{n,i}) = t_i$ and $u_n(s_{n,i}) = x_n(t_i)$ for $i = 1, 2, 3$ and all n. Then, apply Lemma 12.5.2 to obtain, for $n \geq n_0$,

$$\begin{aligned}
\|x_n(t_2) - [x_n(t_1), x_n(t_3)]\| &= \|u_n(s_{n,2}) - [u_n(s_{n1}), u_n(s_{n,3})]\| \\
&\leq \|u(s_{n,2}) - [u(s_{n,1}), u(s_{n,3})]\| + 2\|u_n - u\| \\
&\leq w_s(x, \delta + 2(\eta \wedge \epsilon)/4)) + 2((\eta \wedge \epsilon)/4) \\
&\leq w_s(x, \delta + (\eta/2)) + \epsilon/2 \ ,
\end{aligned}$$

so that, for $\delta < \eta/2$ and $n \geq n_0$, $w_s(x_n, \delta) < \epsilon$.

(iv)→(vi). First, for $\epsilon > 0$ given, apply (iv) to find $\eta < \epsilon/16$ and n_0 such that $w_s(x_n, \eta) < \epsilon/32$ for $n \geq n_0$. Next find a finite set A of points (z_i, t_i) in Γ_x with

$$(x(0), 0) = (z_1, t_1) < (z_2, t_2) < \cdots < (z_m, t_m) = (x(T), T) \ ,$$

using the order defined on Γ_x below (3.3), where for each i, either $t_i \in Disc(x, \epsilon/2)$ or $t_i \in S$, with $Disc(x, \epsilon/2)$ being as in (2.4) and S being a subset of $[0,T]$ including 0, T and the points in $Disc(x, \epsilon/2)^c$ at which x_n converges pointwise to x. Use the left and right limits of x to include in A for each $t \in Disc(x, \epsilon/2)$ points $t' \equiv t'(t)$ and $t'' = t''(t)$ in S such that $t' < t < t''$, t' is greater than all elements of $Disc(x, \epsilon/2)$ less than t, t'' is less than all elements of $Disc(x, \epsilon/2)$ greater than t, $|t' - t| < \eta$, $|t'' - t| < \eta$, $\|x(t') - x(t-)\| < \epsilon/32$ and $\|x(t'') - x(t)\| < \epsilon/32$. In addition, assume that $|t_{i+1} - t_i| < \eta$ for all i and $\hat{d}(A, \Gamma_x) < \epsilon/2$, for which we apply Lemma 12.3.1. Moreover, if $t \in Disc(x, \epsilon/2)$ and

$$t_r < t_{r+1} = t = t_{r+2} = \cdots = t_{r+j} < t_{r+j+1} \ , \tag{5.7}$$

then we require that $\|z_{r+1}-x(t-)\| > \epsilon/4$, $\|z_{r+j}-x(t)\| > \epsilon/4$ and $\|z_{r+i+1}-z_{r+i}\| > \epsilon/4$ for all i, $1 \le i \le j-1$. Since $\hat{d}(A,\Gamma_x) < \epsilon/2$, we also have the upper bound $\|z_{r+i+1}-z_{r+i}\| < \epsilon/2$. For $t_i \in S \cap A$, let $z_i = x(t_i)$. Now, for all $t_i \in S \cap A$, let $n_1 \ge n_0$ be s uch that $\|x_n(t_i) - x(t_i)\| < \epsilon/32$ for all i and $n \ge n_1$, using (iv). We now want to construct the subset A_n of Γ_{x_n}. First for all $t_i \in S \cap A$, let $(z_{n,i}, t_{n,i}) = (x_n(t_i), t_i)$. Now we consider time points in $Disc(x, \epsilon/2)$. By the construction above, given (5.7),

$$\|[x(t_r), x(t_{r+j+1})] - [x_n(t_r), x_n(t_{r+j+1})]\| < \epsilon/32$$

and

$$\|[x(t_r), x(t_{r_j+1})] - [x(t-), x(t)]\| < \epsilon/32 \; . \tag{5.8}$$

Since $w_s(x_n,\eta) < \epsilon/32$, for each (r,i) there is a point $(z_{n,r+i}, t_{n,r+i}) \in \Gamma_{x_n}$ such that

$$\|z_{n,r+i} - z_{r+i}\| < 3\epsilon/32 \quad \text{and} \quad |t_{n,r+i} - t| < \eta < \epsilon/16 \; . \tag{5.9}$$

Moreover, we must have $(z_{n,r+i+1}, t_{n,r+i+1}) > (z_{n,r+i}, t_{n,r+i})$ for $0 \le i \le j$. For $i = 0$ and $i = j$, we can conclude that $t_r < t < t_{r+j+1}$. For other i, a reversal of order can occur only if $w_s(x_n, t, \eta) > \epsilon/16$ because the construction implies that $\|z_{n,r+i+1} - z_{n,r+i}\| > \epsilon/16$, but that is prohibited by the condition that $w_s(x_n, t, \eta) < \epsilon/32$. Hence, the set of points A_n is ordered properly. Moreover, the construction yields $d^*(A, A_n) < \epsilon/16$. Finally, it remains to bound $\hat{d}(A_n, \Gamma_{x_n})$ for $n \ge n_1$. Consider (z_n, t_n) such that $(z_{n,i}, t_{n,i}) < (z_n, t_n) < (z_{n,i+1}, t_{n,i+1})$. Since $\|z_{n,i} - z_i\| < 3\epsilon/32$ for all i and $\|z_{i+1} - z_i\| < \hat{d}(A, \Gamma_x) < \epsilon/2$, $\|z_{n,i+1} - z_{n,i}\| < 5\epsilon/8$ by the triangle inequality. Since $w_s(x_n, \eta) < \epsilon/32$, invoking Lemma 12.5.2, we have

$$\|(z_n, t_n) - [(z_{n,i}, t_{n,i}), (z_{n,i+1}, t_{n,i+1})]\| < \epsilon/32 \; ,$$

so that

$$\|z_n - z_{n,i}\| \vee \|z_n - z_{n,i+1}\| < 21\epsilon/32 < \epsilon$$

and $|t_{n,i} - t_{n,i+1}| < 2\eta < \epsilon/8$. Hence $\hat{d}(A_n, \Gamma_{x_n}) < \epsilon$ for $n \ge n_1$, so that the proof is complete.

(vi)$\rightarrow$(i). Suppose that the conditions in (vi) hold and $\epsilon > 0$ is given. Let $(u, r) \in \Pi_s(x)$ and $(u_n, r_n) \in \Pi_s(x_n)$, $n \ge 1$, be arbitrary parametric representations. Let $s_1 = 0 < s_2 < \cdots < s_m = 1$ and $s_{n,1} = 0 < s_{n,2} < \cdots < s_{n,m} = 1$ be points such that $(u(s_i), r(s_i)) = (z_i, t_i) \in A$ and $(u_n(s_{n,i}), r_n(s_{n,i})) = (z_{n,i}, t_{n,i}) \in A_n$ for $1 \le i \le m$. Let $\lambda_n : [0,1] \to [0,1]$ be a continuous nondecreasing function such that $\lambda_n(s_i) = s_{n,i}$ for each i

and n. We will show that $(u_n \circ \lambda_n, r_n \circ \lambda_n)$ is a parametric representation of Γ_{x_n} for each n such that

$$\|u_n \circ \lambda_n - u\| \vee \|r_n \circ \lambda_n - r\| < 3\epsilon \quad \text{for} \quad n \geq n_1 \ . \tag{5.10}$$

Property (5.10) holds because, for $s_i \leq s \leq s_{i+1}$, $\lambda_n(s_i) = s_{n,i} \leq \lambda_n(s) \leq s_{n,i+1} = \lambda_n(s_{i+1})$ and

$$\begin{aligned}
& \|u_n \circ \lambda_n(s) - u(s)\| \vee \|r_n \circ \lambda_n(s) - r(s)\| \\
\leq \ & \|(u_n \circ \lambda_n(s), r_n \circ \lambda_n(s)) - (u_n(s_{n,i}), r_n(s_{n,i}))\| \vee \|(u_n \circ \lambda_n(s), r_n \circ \lambda_n(s)) \\
- \ & (u_n(s_{n,i+1}), r_n(s_{n,i+1}))\| \\
+ \ & \|(u(s), r(s)) - (u(s_i), r(s_i))\| \vee \|(u(s), r(s)) - (u(s_{i+1}), r(s_{i+1}))\| \\
+ \ & \|(u_n(s_{n,i}), r_n(s_{n,i})) - (u(s_i), r(s_i))\| \\
\leq \ & \hat{d}(A_n, \Gamma_{x_n}) + \hat{d}(A, \Gamma_x) + d^*(A, A_n) \leq 3\epsilon \ .
\end{aligned}$$

(v)$\rightarrow$(iv). First, the convergence $x_n(T) \rightarrow x(T)$ is assumed directly. Next, (5.4) implies that $x_n(t) \rightarrow x(t)$ as $n \rightarrow \infty$ for each $t \notin Disc(x)$. Since $[0,T] - Disc(x)$ is a dense subset of $[0,T]$, the first part of (iv) is established. Condition (5.4) also implies that (5.5) holds for each $t \notin Disc(x)$ by Theorem 12.4.1. Finally, we show that (5.5) for each $t \in [0,T]$ implies (5.3). Condition (5.5) for each t implies that for each $\epsilon > 0$ and t, there is $\delta \equiv \delta(t)$ and $n(t, \epsilon, \delta)$ such that $w_s(x_n, t, \delta) < \epsilon$ for all $n \geq n(t, \epsilon, \delta)$. Now suppose that (5.3) does not hold. Then there must exist $\epsilon > 0$ such that for all $\delta > 0$ there is a sequence $\{t_k\}$ of points in $[0,T]$ and a sequence of integers n_k such that $n_k \rightarrow \infty$ and $w_s(x_{n_k}, t_k, \delta/2) > \epsilon$ for all k. However, the sequence $\{t_k\}$ has a subsequence $\{t_{k_j}\}$ with $t_{k_j} \rightarrow t \in [0,T]$ as $k_j \rightarrow \infty$. Thus, for all k_j suitably large,

$$w_s(x_{n_{k_j}}, t, \delta) > w_s(x_{n_{k_j}}, t_{n_{k_j}}, \delta/2) > \epsilon \ ,$$

which is a contradiction, so that (5.3) must in fact hold.
(iii)+(iv)$\rightarrow$(v). By Lemma 12.5.1, (iii) implies (5.4) for each $t \in Disc(x)^c$. Trivially, (iv) implies (5.3), which in turn implies (5.5). ■

Remark 12.5.1. *Connection to Skorohod (1956).* Part (iv) of Theorem 12.5.1 is essentially Skorohod's (1956) original characterization, established in his 2.4.1. Instead of (5.1) with (4.4), Skorohod (1956) actually considered (5.1) with $w_s(x, t, \delta)$ replaced by

$$w'_s(x, t, \delta) \equiv \sup_{0 \vee (t-\delta) \leq t_1 \leq t \leq t_3 \leq (t+\delta) \wedge T} \{\|x(t) - [x(t_1), x(t_3)]\|\} \ , \tag{5.11}$$

but when the supremum over $t \in [0,T]$ is applied, w_s and w'_s are equivalent. In particular, clearly $w'_s(x,t,\delta) \le w_s(x,t,\delta)$ for each t. On the other hand, if $w_s(x,t,\delta) > \epsilon$ for all t, then $w'_s(x,t,2\delta) > \epsilon$ for all t. Hence (iv) is equivalent to Skorohod's original characterization. We have introduced $w_s(x,t,\delta)$ in (4.4) in order to get characterization (v) in Theorem 12.5.1. We cannot use Skorohod's (5.11) instead of (4.4) in characterization (v) in Theorem 12.5.1, because it does not rule out multiple large oscillations on the same side of t. ■

Remark 12.5.2. *Possibility of using one-to-one parametric representations.* The proof of the implication (vi)→(i) shows that the SM_1 topology is unaltered if all the parametric representations are required to be one-to-one functions from $[0,1]$ onto the graph. In the proof we would then let the transformations λ_n be homeomorphisms of $[0,1]$, so that $(u_n \circ \lambda_n, r_n \circ \lambda_n)$ become one-to-one functions. ■

We can apply Theorem 12.5.1 to develop a simple criterion for M_1 convergence for monotone functions.

Corollary 12.5.1. (the case of monotone functions) *If x_n is monotone for each n, then $d_s(x_n, x) \to 0$ for $x \in D$ if and only if $x_n(t) \to x(t)$ for all t in a dense subset of $[0,T]$ including 0 and T.*

Proof. Apply Theorem 12.5.1 (iv). Note that condition (5.3) always holds for monotone functions. ■

12.5.2. WM_1 Convergence

We now establish an analog of Theorem 12.5.1 for the WM_1 topology. Several alternative characterizations of WM_1 convergence will follow directly from Theorem 12.5.1 because we will show that convergence $x_n \to x$ as $n \to \infty$ in WM_1 is equivalent to $d_p(x_n, x) \to 0$. To treat the WM_1 topology, we define another oscillation function. Let

$$w_w(x,\delta) \equiv \sup_{0 \le t \le T} w_w(x,t,\delta) \tag{5.12}$$

for $w_w(x,t,\delta)$ in (4.5). Recall that $w_w(x,t,\delta)$ in (4.5) is the same as $w_s(x,t,\delta)$ in (4.4) except it has the product segment $[[x(t_1), x(t_3)]]$ in (3.2) instead of the standard segment $[x(t_1), x(t_3)]$ in (3.1).

Paralleling Definition 12.3.1, let an ordered subset A of G_x of cardinality m be such that (3.9) holds, but now with the order being the order on G_x. Paralleling (3.10), let the *order-consistent distance* between A and G_x be

$$\hat{d}(A, G_x) \equiv \sup\{\|(z,t)-(z_i,t_i)\| \vee \|(z,t)-(z_{i+1},t_{i+1})\| : (z,t) \in G_x\} \quad (5.13)$$

with the supremum being over all $(z,t) \in G_x$ such that $(z_i,t_i) \leq (z,t) \leq (z_{i+1},t_{i+1})$ for all i, $1 \leq i \leq m-1$.

Theorem 12.5.2. (characterizations of WM_1 convergence) *The following are equivalent characterizations of $x_n \to x$ as $n \to \infty$ in (D, WM_1):*

(i) $d_w(x_n, x) \to 0$ *as* $n \to \infty$.

(ii) $d_p(x_n, x) \to 0$ *as* $n \to \infty$.

(iii) $x_n(t) \to x(t)$ *as* $n \to \infty$ *for each* t *in a dense subset of* $[0,T]$ *including* 0 *and* T*, and*

$$\lim_{\delta \downarrow 0} \overline{\lim_{n\to\infty}} \; w_w(x_n, \delta) = 0 \ . \quad (5.14)$$

(iv) $x_n(T) \to x(T)$ *as* $n \to \infty$*; for each* $t \notin Disc(x)$,

$$\lim_{\delta \downarrow 0} \overline{\lim_{n\to\infty}} \; v(x_n, x, t, \delta) = 0 \quad (5.15)$$

for $v(x_n, x, t, \delta)$ *in* (4.2)*; and, for each* $t \in Disc(x)$,

$$\lim_{\delta \downarrow 0} \overline{\lim_{n\to\infty}} \; w_w(x_n, t, \delta) = 0 \quad (5.16)$$

for $w_w(x_n, t, \delta)$ *in* (4.5).

(v) for all $\epsilon > 0$ *and all* n *sufficiently large, there exist finite ordered subsets* A *of* G_x *(in general depending on* n*) and* A_n *of* G_{x_n} *of common cardinality such that* $\hat{d}(A, G_x) < \epsilon$, $\hat{d}(A_n, G_{x_n}) < \epsilon$ *and* $d^*(A, A_n) < \epsilon$ *for* $\hat{d}$ *in* (5.13) *and* d^* *in* (5.6).

Example 12.5.1. *Need for changing parametric representations.* In general, there is no analog of characterizations (i) and (ii) in Theorem 12.5.1 for the parametric representations in $\Pi_w(x)$ and $\Pi_w(x_n)$; i.e., if $d_w(x_n, x) \to 0$ as $n \to \infty$, there need not exist $(u,r) \in \Pi_w(x)$ and $(u_n, r_n) \in \Pi_w(x_n)$ such that (5.2) holds. To see this, let $x^1 = x^2 = I_{[1,2]}$, $x^1_{2n+1} = x^2_{2n} = I_{[1-n^{-1},2]}$ and $x^2_{2n+1} = x^1_{2n} = I_{[1+n^{-1},2]}$ for $n \geq 2$. Property (i) of Theorem 12.5.2 holds, but different parametric representations of x are needed for even and odd n. ■

12.6. Strengthening the Mode of Convergence

In this section we apply the characterizations of M_1 convergence in Sections 12.3 and 12.5 to establish conditions under which the mode of convergence can be strengthened: We seek conditions under which WM_1 convergence can be replaced by SM_1 convergence. We use the following Lemma.

Lemma 12.6.1. (modulus bound for (x_n, y_n)) *For* $x_n \in D([0,T], \mathbb{R}^k)$, $y_n, y \in D([0,T], \mathbb{R}^l)$, $t \in [0,T]$ *and* $\delta > 0$,

$$w_s((x_n, y_n), t, \delta) \le w_s(x_n, t, \delta) + 2v(y_n, y, t, \delta) .$$

Theorem 12.6.1. (extending SM_1 convergence to product spaces) *Suppose that* $d_s(x_n, x) \to 0$ *in* $D([0,T], \mathbb{R}^k)$ *and* $d_s(y_n, y) \to 0$ *in* $D([0,T], \mathbb{R}^l)$ *as* $n \to \infty$. *If*

$$Disc(x) \cap Disc(y) = \phi .$$

then

$$d_s((x_n, y_n), (x, y)) \to 0 \quad in \quad D([0,T], \mathbb{R}^{k+l}) \quad as \quad n \to \infty .$$

Proof. We use characterization (v) in Theorem 12.5.1. First, for each $t \notin Disc((x,y))$, $t \notin Disc(x) \cup Disc(y)$, (5.4) holds and

$$\lim_{\delta \downarrow 0} \overline{\lim_{n \to \infty}} \; v(y_n, y, \delta, t) = 0 \; , \tag{6.1}$$

which implies that

$$\lim_{\delta \downarrow 0} \overline{\lim_{n \to \infty}} \; v((x_n, y_n), (x, y), \delta, t) = 0 \; .$$

Now, for each $t \in Disc(x)$, (5.5) and (6.1) hold (because $Disc(x) \cap Disc(y) = \phi$). Thus, for those t, by Lemma 12.6.1,

$$\lim_{\delta \downarrow 0} \overline{\lim_{n \to \infty}} \; w_s((x_n, y_n), t, \delta) = 0 \; . \tag{6.2}$$

By the same reasoning (6.2) also holds for each $t \in Disc(y)$, so that (6.2) holds for all $t \in Disc((x,y)) = Disc(x) \cup Disc(y)$. ∎

Remark 12.6.1. *The discontinuity condition is not necessary.* The discontinuity condition $Disc(x) \cap Disc(y) = \phi$ in Theorem 12.6.1 is not necessary. To see that, note that if $x_n \to x$ as $n \to \infty$ in $D([0,T],\mathbb{R}^k)$, then $(x_n, x_n) \to (x, x)$ as $n \to \infty$ in $D([0,T],\mathbb{R}^{2k})$. However, some condition is needed, as can be seen from the fact that the WM_1 topology is strictly weaker than the SM_1 topology on $D([0,T],\mathbb{R}^k)$ for $k > 1$, as shown by Example 12.3.1. ∎

Remark 12.6.2. *The J_1 and M_2 analogs.* Analogs of Theorem 12.6.1 hold in the J_1 and M_2 topologies. For J_1, see Propositions 2.1 (a) and 2.2 (b) on p. 301 of Jacod and Shiryaev (1987). For M_2, see Theorem 12.11.3 below. ∎

As in Lemma 12.3.3, let $D_1 \equiv D_1([0,T],\mathbb{R}^k)$ be the subset of x in D with discontinuities in only one coordinate at a time; i.e., $x \in D_1$ if $x^i(t-) \neq x^i(t)$ for at most one coordinate i for each t. (The coordinate $i \equiv i(t)$ may depend upon t.)

Corollary 12.6.1. (from WM_1 convergence to SM_1 convergence when the limit is in D_1) *If $d_p(x_n, x) \to 0$ as $n \to \infty$ and $x \in D_1$, then $d_s(x_n, x) \to 0$.*

Example 12.3.3 shows that it is not enough to have $x \in D_s$ in Corollary 12.6.1.

12.7. Characterizing Convergence with Mappings

The strong topology SM_1 differs from the weak topology WM_1 by the behavior of linear functions of the coordinates. Example 12.3.1 shows that linear functions of the coordinates are not continuous in the product topology (there $(x_n^1 - x_n^2) \not\to (x^1 - x^2)$ as $n \to \infty$), but they are in the strong topology, as we now show. Note that there is no subscript on d on the left in (7.1) below because ηx is real valued.

Theorem 12.7.1. (Lipschitz property of linear functions of the coordinate functions) *For any x_1, $x_2 \in D([0,T],\mathbb{R}^k)$ and $\eta \in \mathbb{R}^k$,*

$$d(\eta x_1, \eta x_2) \leq (\|\eta\| \vee 1) d_s(x_1, x_2) \ . \tag{7.1}$$

Example 12.7.1. *Difficulties with the weak topology.* To see that $(\eta z, t)$ need not be an element of $\Gamma_{\eta x}$ when $(z,t) \in G_x$ and that $(\eta u, r)$ need not be an element of $\Pi(\eta x)$ when $(u, r) \in \Pi_w(x)$, let $x^1 = x^2 = I_{[1,2]}$ and consider $\eta x = x^1 - x^2$. The flexibility allowed by G_x allows $(z,t) \in G_x$ with $\eta z \neq 0$ and $(u,r) \in \Pi_w(x)$ with $u(s) \neq 0$. ∎

We now obtain a sufficient condition for addition to be continuous on $(D, d_s) \times (D, d_s)$, which is analogous to the J_1 result in Theorem 4.1 of Whitt (1980).

Corollary 12.7.1. (SM_1-continuity of addition) *If* $d_s(x_n, x) \to 0$ *and* $d_s(y_n, y) \to 0$ *in* $D([0,T], \mathbb{R}^k)$ *and*

$$Disc(x) \cap Disc(y) = \phi ,$$

then

$$d_s(x_n + y_n, x + y) \to 0 \quad in \quad D([0,T], \mathbb{R}^k) .$$

Proof. First apply Theorem 12.6.1 to get $d_s((x_n, y_n), (x, y)) \to 0$ in $D([0,T], \mathbb{R}^{2k})$. Then apply Theorem 12.7.1. ∎

Remark 12.7.1. *Measurability of addition.* The measurability of addition on $(D, d_s) \times (D, d_s)$ holds because the Borel σ-field coincides with the Kolmogorov σ-field. It also follows from part of the proof of Theorem 4.1 of Whitt (1980). ∎

In Theorem 12.7.1 we showed that linear functions of the coordinates are Lipschitz in the SM_1 metric. We now apply Theorem 12.5.1 to show that convergence in the SM_1 topology is characterized by convergence of all such linear functions of the coordinates.

Theorem 12.7.2. (characterization of SM_1 convergence by convergence of all linear functions) *There is convergence* $x_n \to x$ *in* $D([0,T], \mathbb{R}^k)$ *as* $n \to \infty$ *in the* SM_1 *topology if and only if* $\eta x_n \to \eta x$ *in* $D([0,T], \mathbb{R}^1)$ *as* $n \to \infty$ *in the* M_1 *topology for all* $\eta \in \mathbb{R}^k$.

We can get convergence of sums under more general conditions than in Corollary 12.7.1. It suffices to have the jumps of x^i and y^i have common sign for all i. We can express this property by the condition

$$(x^i(t) - x^i(t-))(y^i(t) - y^i(t-)) \geq 0 \tag{7.2}$$

for all t, $0 \leq t \leq T$, and all i, $1 \leq i \leq k$.

Theorem 12.7.3. (continuity of addition at limits with jumps of common sign) *If* $x_n \to x$ *and* $y_n \to y$ *in* $D([0,T], \mathbb{R}^k, SM_1)$ *and if condition* (7.2) *above holds, then*

$$x_n + y_n \to x + y \quad in \quad D([0,T], \mathbb{R}^k, SM_1) .$$

Proof. Apply the characterization of SM_1 convergence in Theorem 12.5.1 (v). At points t in $Disc(x)^c \cap Disc(y)^c$, use the local uniform convergence in Lemma 12.5.1 and Corollary 12.11.1. For other t not in $Disc(x) \cap Disc(y)$, use Theorem 12.6.1. For $t \in Disc(x) \cap Disc(y)$, exploit condition (7.2) to deduce that, for all $\epsilon > 0$, there exists δ and n_0 such that

$$w_s(x_n + y_n, t, \delta) \leq w_s(x_n, t, \delta) + w_s(y_n, t, \delta) + \epsilon$$

for all $n \geq n_0$. ■

In Sections (2.2.7)–(2.2.13) of Skorohod (1956), convenient characterizations of convergence in each topology are given for real-valued functions. We can apply Theorem 12.7.2 to develop associated characterizations for $\mathbb{R}^k$-valued functions. For each $x \in D([0,T], \mathbb{R}^1)$, $0 \leq t_1 < t_2 \leq T$ and, for each $a < b$ in $\mathbb{R}$, let $v_{t_1,t_2}^{a,b}(x)$ be the number of visits to the strip $[a,b]$ on the interval $[t_1, t_2]$; i.e., $v_{t_1,t_2}^{a,b}(x) = k$ if it is possible to find k (but not $k+1$) points t_i' such that $t_1 < t_1' < \cdots < t_k' \leq t_2$ such that either

$$x(t_1) \in [a,b], \; x(t_1') \notin [a,b], \; x(t_2') \in [a,b], \ldots,$$

or

$$x(t_1) \notin [a,b], \; x(t_1') \in [a,b], \; x(t_2') \notin [a,b], \ldots$$

We say that $x \in D([0,T], \mathbb{R})$ has a *local maximum (minimum) value at t relative to* (t_1, t_2) in $(0,T)$ if $t_1 < t < t_2$ and either

$$\text{(i)} \quad \sup\{x(s) : t_1 \leq s \leq t_2\} \leq x(t) \quad (\inf\{x(s) : t_1 \leq s \leq t_2\} \geq x(t))$$

or

$$\text{(ii)} \quad \sup\{x(s) : t_1 \leq s \leq t_2\} \leq x(t-) \quad (\inf\{x(s) : t_1 \leq s \leq t_2)\} \geq x(t-)) \; .$$

We say that x has a *local maximum (minimum) value* at t if it has a local maximum (minimum) value at t relative to some interval (t_1, t_2) with $t_1 < t < t_2$. We call local maximum and minimum values *local extreme values.*

Lemma 12.7.1. (local extreme values) *Any $x \in D([0,T], \mathbb{R})$ has at most countably many local extreme values.*

If b is not a local extreme value of x, then x crosses level b whenever x hits b; i.e., if b is not a local extreme value and if $x(t) = b$ or $x(t-) = b$, then for every t_1, t_2 with $t_1 < t < t_2$ there exist t_1', t_2' with $t_1 < t_1'$, $t_2' < t_2$ such that $x(t_1') < b$ and $x(t_2') > b$. This property implies the following lemma.

Lemma 12.7.2. *Consider an interval* $[t_1, t_2]$ *with* $0 < t_1 < t_2 < T$. *If* $x(t_i) \notin \{a, b\}$ *for* $i = 1, 2$ *and* a, b *are not local extreme values of* x, *then* x *crosses one of the levels* a *and* b *at each of the* $v^{a,b}_{t_1,t_2}(x)$ *visits to the strip* $[a, b]$ *in* $[t_1, t_2]$.

Theorem 12.7.4. (characterization of SM_1 convergence in terms of convergence of number of visits to strips) *There is convergence* $d_s(x_n, x) \to 0$ *as* $n \to \infty$ *in* $D([0, T], \mathbb{R}^k)$ *if and only if*

$$v^{a,b}_{t_1,t_2}(\eta x_n) \to v^{a,b}_{t_1,t_2}(\eta x) \quad as \quad n \to \infty$$

for all $\eta \in \mathbb{R}^k$, *all points* $t_1, t_2 \in \{T\} \cup Disc(x)^c$ *with* $t_1 < t_2$ *and almost all* a, b *with respect to Lebesgue measure.*

12.8. Topological Completeness

In this section we exhibit a complete metric topologically equivalent to the incomplete metric d_s in (3.7) inducing the SM_1 topology. Since a product metric defined as in (3.11) inherits the completeness of the component metrics, we also succeed in constructing complete metrics inducing the associated product topology. We make no use of the complete metrics beyond showing that the topology is topologically complete. Another approach to topological completeness would be to show that D is homeomorphic to a G_δ subset of a complete metric space, as noted in Section 11.2.

In our construction of complete metrics, we follow the argument used by Prohorov (1956, Appendix 1) to show that the J_1 topology is topologically complete; we incorporate an oscillation function into the metric. For M_1, we use $w_s(x, \delta)$ in (5.1). Since $w_s(x, \delta) \to 0$ as $\delta \to 0$ for each $x \in D$, we need to appropriately "inflate" differences for small δ. For this purpose, let

$$\hat{w}_s(x, z) \equiv \begin{cases} w_s(x, e^z), & z < 0 \\ \\ w_s(x, 1), & z \geq 1 \ . \end{cases} \tag{8.1}$$

Since $w_s(x, \delta)$ is nondecreasing in δ, $\hat{w}_s(x, z)$ is nondecreasing in z. Note that $\hat{w}_s(x, z)$ as a function of z has the form of a cumulative distribution function (cdf) of a finite measure. On such cdf's, the Lévy metric λ is known to be a complete metric inducing the topology of pointwise convergence at all continuity points of the limit; i.e.,

$$\lambda(F_1, F_2) \equiv \inf\{\epsilon > 0 : F_2(x - \epsilon) - \epsilon \leq F_1(x) \leq F_2(x + \epsilon) + \epsilon\} \ . \tag{8.2}$$

The Helly selection theorem, p. 267 of Feller (1971), can be used to show that the metric λ is complete.

Thus, our new metric is

$$\hat{d}_s(x_1, x_2) \equiv d_s(x_1, x_2) + \lambda(\hat{w}_s(x_1, \cdot), \hat{w}_s(x_2, \cdot)) \; . \tag{8.3}$$

Theorem 12.8.1. (a complete SM_1 metric) *The metric $\hat{d}_s$ on D in* (8.3) *is complete and topologically equivalent to d_s.*

Example 12.8.1. *The counterexample for d_s is not fundamental under $\hat{d}_s$.* Recall that Example 12.10.1 was used to show that the metric d_s is not complete. That example has $x_n = I_{[1,1+n^{-1})}$, so that $d_s(x_m, x_n) \to 0$ as $m, n \to \infty$, i.e., the sequence $\{x_n\}$ is fundamental for d_s even though it does not converge. Note that $w_s(x_n, \delta) = 1$ for $\delta > 1/2n$ and $w_s(x_n, \delta) = 0$ otherwise. Hence, $\hat{w}_s(x_n, z) = 1$ for $z > \log(1/2n) = -\log(2n)$ and $\hat{w}_s(x_n, z) = 0$ otherwise. Note that $\hat{w}_s(x_n, \cdot)$ corresponds to the cdf of a unit point mass at $-\log(2n)$. Consequently, $\hat{d}_s(x_m, x_n) \not\to 0$ as $m, n \to \infty$.

Remark 12.8.1. *An alternative complete metric.* An alternative complete metric topologically equivalent to d_s is

$$d_s^{\dagger}(x_1, x_2) = m_s(x_1, x_2) + \lambda(\hat{w}_s(x_1, \cdot), \hat{w}_s(x_2, \cdot)) \; , \tag{8.4}$$

where $m_s \equiv d_{M_2}$ is the M_2 metric in (5.4) of Section 11.5. That is actually what Prohorov did for J_1 (with $\hat{w}_s$ in (8.4) replaced by the J_1 oscillation function). ∎

12.9. Non-Compact Domains

It is often convenient to consider the function space $D([0, \infty), \mathbb{R}^k)$ with domain $[0, \infty)$ instead of $[0, T]$. More generally, we may consider the function space $D(I, \mathbb{R}^k)$, where I is a subinterval of the real line. Common cases besides $[0, \infty)$ are $(0, \infty)$ and $(-\infty, \infty) \equiv \mathbb{R}$.

Given the function space $D(I, \mathbb{R}^k)$ for any subinterval I, we define convergence $x_n \to x$ with some topology to be convergence in $D([a, b], \mathbb{R}^k)$ with that same topology for the restrictions of x_n and x to the compact interval $[a, b]$ for all points a and b that are elements of I and either boundary points of I or are continuity points of the limit function x. For example, for $I = [c, d)$ with $-\infty < c < d < \infty$, we include $a = c$ but exclude $b = d$; for $I = [c, d]$, we include both c and d.

For simplicity, we henceforth consider only the special case in which $I = [0, \infty)$. In that setting, we can equivalently define convergence $x_n \to x$ as $n \to \infty$ in $D([0, \infty), \mathbb{R}^k)$ with some topology to be convergence $x_n \to x$ as $n \to \infty$ in $D([0, t], \mathbb{R}^k)$ with that topology for the restrictions of x_n and x to $[0, t]$ for $t = t_k$ for each t_k in some sequence $\{t_k\}$ with $t_k \to \infty$ as $k \to \infty$, where $\{t_k\}$ can depend on x. It suffices to let t_k be continuity points of the limit function x; for the J_1 topology, see Stone (1963), Lindvall (1973), Whitt (1980) and Jacod and Shiryaev (1987). We will discuss only the SM_1 topology here, but the discussion applies to the other non-uniform topologies as well. We also will omit most proofs.

As a first step, we consider the case of closed bounded intervals $[t_1, t_2]$. The space $D([t_1, t_2], \mathbb{R}^k)$ is essentially the same as (homeomorphic to) the space $D([0, T], \mathbb{R}^k)$ already studied, but we want to look at the behavior as we change the interval $[t_1, t_2]$. For $[t_3, t_4] \subseteq [t_1, t_2]$, we consider the restriction of x in $D([t_1, t_2], \mathbb{R}^k)$ to $[t_3, t_4]$, defined by

$$r_{t_3,t_4} : D([t_1, t_2], \mathbb{R}^k) \to D([t_3, t_4], \mathbb{R}^k)$$

with $r_{t_3,t_4}(x)(t) = x(t)$ for $t_3 \le t \le t_4$. Let d_{t_1,t_2} be the metric d_s on $D([t_1, t_2], \mathbb{R}^k)$. We want to relate the distance $d_{t_1,t_2}(x_1, x_2)$ and convergence $d_{t_1,t_2}(x_n, x) \to 0$ as $n \to \infty$ for different domains. We first state a result enabling us to go from the domains $[t_1, t_2]$ and $[t_2, t_3]$ to $[t_1, t_3]$ when $t_1 < t_2 < t_3$.

Lemma 12.9.1. (metric bounds) *For $0 \le t_1 < t_2 < t_3$ and x_1, $x_2 \in D([t_1, t_3], \mathbb{R}^k)$,*

$$d_{t_1,t_3}(x_1, x_2) \le d_{t_1,t_2}(x_1, x_2) \vee d_{t_2,t_3}(x_1, x_2) \ .$$

We now observe that there is an equivalence of convergence provided that the internal boundary point is a continuity point of the limit function.

Lemma 12.9.2. *For $0 \le t_1 < t_2 < t_3$ and x, $x_n \in D([t_1, t_3], \mathbb{R}^k)$, with $t_2 \in Disc(x)^c$, $d_{t_1,t_3}(x_n, x) \to 0$ as $n \to \infty$ if and only if $d_{t_1,t_2}(x_n, x) \to 0$ and $d_{t_2,t_3}(x_n, x) \to 0$ as $n \to \infty$.*

For $x \in D([0, T], \mathbb{R}^k)$ and $0 \le t_1 < t_2 \le T$, let $r_{t_1,t_2} : D([0, T], \mathbb{R}^k) \to D([t_1, t_2], \mathbb{R}^k)$ be the restriction map, defined by $r_{t_1,t_2}(x)(s) = x(s)$, $t_1 \le s \le t_2$.

Corollary 12.9.1. (continuity of restriction maps) *If $x_n \to x$ as $n \to \infty$ in $D([0, T], \mathbb{R}^k, SM_1)$ and if $t_1, t_2 \in Disc(x)^c$, then*

$$r_{t_1,t_2}(x_n) \to r_{t_1,t_2}(x) \quad as \quad n \to \infty \quad in \quad D([t_1, t_2], \mathbb{R}^k, SM_1) \ .$$

Let $r_t : D([0,\infty),\mathbb{R}^k) \to D([0,t],\mathbb{R}^k)$ be the *restriction map* with $r_t(x)(s) = x(s),\ 0 \le s \le t$. Suppose that $f : D([0,\infty),\mathbb{R}^k) \to D([0,\infty),\mathbb{R}^k)$ and $f_t : D([0,t],\mathbb{R}^k) \to D([0,t],\mathbb{R}^k)$ for $t > 0$ are functions with

$$f_t(r_t(x)) = r_t(f(x))$$

for all $x \in D([0,\infty),\mathbb{R}^k)$ and all $t > 0$. We then call the functions f_t *restrictions of the function* f.

Theorem 12.9.1. (continuity from continuous restrictions) *Suppose that $f : D([0,\infty),\mathbb{R}^k) \to D([0,\infty),\mathbb{R}^l)$ has continuous restrictions f_t with some topology for all $t > 0$. Then f itself is continuous in that topology.*

We now consider the extension of Lipschitz properties to subsets of $D([0,\infty),\mathbb{R}^k)$. For this purpose, suppose that μ_t is one of the M_1 metrics on $D([0,t],\mathbb{R}^k)$ for $t > 0$. As in Section 2 of Whitt(1980), an associated metric μ can be defined on $D([0,\infty),\mathbb{R}^k)$ by

$$\mu(x_1,x_2) = \int_0^\infty e^{-t}[\mu_t(r_t(x_1),r_t(x_2)) \wedge 1]dt. \tag{9.1}$$

The following result implies that the integral in (9.1) is well defined.

Theorem 12.9.2. (regularity of the metric $\mu_t(x_1,x_2)$ as a function of t) *Let μ_t be one of the M_1 metrics on $D([0,t],\mathbb{R}^k)$. For all $x_1, x_2 \in D([0,\infty),\mathbb{R}^k)$, $\mu_t(x_1,x_2)$ as a function of t is right-continuous with left limits in $(0,\infty)$ and has a right limit at 0. Moreover, $\mu_t(x_1,x_2)$ is continuous at $t > 0$ whenever x_1 and x_2 are both continuous at t.*

We also have the following result, paralleling Lemma 2.2 and Theorem 2.5 of Whitt (1980). For (iii), we exploit Theorem 12.5.1 (i).

Theorem 12.9.3. (characterizations of SM_1 convergence with domain $[0,\infty)$) *Suppose that μ and μ_t, $t > 0$ are the SM_1 (or WM_1) metrics on $D([0,\infty),\mathbb{R}^k)$ and $D([0,t],\mathbb{R}^k)$. Then the following are equivalent for x and x_n, $n \ge 1$, in $D([0,\infty),\mathbb{R}^k)$.*

(i) $\mu(x_n,x) \to 0$ *as* $n \to \infty$*;*

(ii) $\mu_t(r_t(x_n),r_t(x)) \to 0$ *as* $n \to \infty$ *for all* $t \notin Disc(x)$*;*

(iii) there exist parametric representations (u, r) *and* (u_n, r_n) *of* x *and* x_n *mapping* $[0, \infty)$ *into the graphs such that*

$$\|u_n - u\|_t \vee \|r_n - r\|_t \to 0 \quad as \quad n \to \infty$$

for each $t > 0$.

We now show that the Lipschitz property extends from $D([0, t], \mathbb{R}^k)$ to $D([0, \infty), \mathbb{R}^k)$.

Theorem 12.9.4. (functions with Lipschitz restrictions are Lipschitz) *If a function*

$$f : D([0, \infty), \mathbb{R}^k) \to D([0, \infty), \mathbb{R}^k)$$

has restrictions

$$f_t : D([0, T], \mathbb{R}^k) \to D([0, T], \mathbb{R}^k)$$

satisfying

$$\mu_t^2(f_t(r_t(x_1)), f_t(r_t(x_2))) \le K\mu_t^1(r_t(x_1), r_t(x_2)) \quad for\ all \quad t > 0 \ ,$$

where K *is independent of* t, *then*

$$\mu^2(f(x_1), f(x_2)) \le (K \vee 1)\mu^1(x_1, x_2).$$

Proof. By (9.1) and the conditions,

$$\begin{aligned}
\mu^2(f(x_1), f(x_2)) &= \int_0^\infty e^{-t}[\mu_t^2(r_t(f(x_1)), r_t(f(x_2))) \wedge 1]dt \\
&= \int_0^\infty e^{-t}[\mu_t^2(f_t(r_t(x_1)), f_t(r_t(x_2))) \wedge 1]dt \\
&\le \int_0^\infty e^{-t}[K\mu_t^1(r_t(x_1), r_t(x_2)) \wedge 1]dt \\
&\le (K \vee 1)\int_0^\infty e^{-t}[\mu_t^1(r_t(x_1), r_t(x_2)) \wedge 1]dt \\
&\le (K \vee 1)\mu^1(x_1, x_2) \ . \quad \blacksquare
\end{aligned}$$

12.10. Strong and Weak M_2 Topologies

We now define strong and weak versions of Skorohod's M_2 topology. In Section 12.11 we will show that it is possible to define the M_2 topologies by a minor modification of the definitions in Section 12.3, in particular, by simply using parametric representations in which only r is nondecreasing instead of (u, r), but now we will use Skorohod's (1956) original approach, and relate it to the Hausdorff metric on the space of graphs.

The weak topology will be defined just like the strong, except it will use the thick graphs G_x instead of the thin graphs Γ_x. In particular, let

$$\mu_s(x_1, x_2) \equiv \sup_{(z_1,t_1)\in\Gamma_{x_1}} \inf_{(z_2,t_2)\in\Gamma_{x_2}} \{\|(z_1, t_1) - (z_2, t_2)\|\} \tag{10.1}$$

and

$$\mu_w(x_1, x_2) \equiv \sup_{(z_1,t_1)\in G_{x_1}} \inf_{(z_2,t_2)\in G_{x_2}} \{\|(z_1, t_1) - (z_2, t_2)\|\} \ . \tag{10.2}$$

Following Skorohod (1956), we say that $x_n \to x$ as $n \to \infty$ for a sequence or net $\{x_n\}$ in the strong M_2 topology, denoted by SM_2 if $\mu_s(x_n, x) \to 0$ as $n \to \infty$. Paralleling that, we say that $x_n \to x$ as $n \to \infty$ in the weak M_2 topology, denoted by WM_2, if $\mu_w(x_n, x) \to 0$ as $n \to \infty$. We say that $x_n \to x$ as $n \to \infty$ in the product topology if $\mu_s(x_n^i, x^i) \to 0$ (or equivalently $\mu_w(x_n^i, x^i) \to 0$) as $n \to \infty$ for each i, $1 \le i \le k$.

We can also generate the SM_2 and WM_2 topologies using the Hausdorff metric in (5.2) of Section 11.5. As in (5.4) in Section 11.5, for $x_1, x_2 \in D$,

$$m_s(x_1, x_2) \equiv m_H(\Gamma_{x_1}, \Gamma_{x_2}) = \mu_s(x_1, x_2) \vee \mu_s(x_2, x_1) \ , \tag{10.3}$$

$$m_w(x_1, x_2) \equiv m_H(G_{x_1}, G_{x_2}) = \mu_w(x_1, x_2) \vee \mu_w(x_2, x_1) \tag{10.4}$$

and

$$m_p(x_1, x_2) \equiv \max_{1\le i\le k} m_s(x_1^i, x_2^i) \ . \tag{10.5}$$

We will show that the metric m_s induces the SM_2 topology.

That will imply that the metric m_p induces the associated product topology. However, it turns out that the metric m_w does *not* induce the WM_2 topology. We will show that the WM_2 topology coincides with the product topology, so that the Hausdorff metric can be used to define the WM_2 topology via m_p in (10.5).

Closely paralleling the d or M_1 metrics, we have $m_p \le m_s$ on $D([0, T], \mathbb{R}^k)$ and $m_p = m_w = m_s$ on $D([0, T], \mathbb{R}^1)$. Just as with d, we use m without subscript when the functions are real valued. Example 12.3.1, which showed that WM_1 is strictly weaker than SM_1 also shows that WM_2 is strictly weaker than SM_2. Example 12.3.3 shows that the SM_2 topology is strictly weaker than the SM_1 topology.

Note that μ_s in (10.1) is *not* symmetric in its two arguments. We first show that if $\mu_s(x, x_n) \to 0$ as $n \to \infty$, we need not have $\mu_s(x_n, x) \to 0$ as $n \to \infty$.

Example 12.10.1. *Lack of symmetry of μ_s in its arguments.* To see that we can have $\mu_s(x, x_n) \to 0$ as $n \to \infty$ without $\mu_w(x_n, x) \to 0$ or $\mu_s(x_n, x) \to 0$ as $n \to \infty$, let $x(t) = 0$, $0 \le t \le 2$, and let $x_n = I_{[1,1+n^{-1})}$ in $D([0,2], \mathbb{R}^1)$. Clearly $m_w(x_n, x) \not\to 0$, but for any $(0, t) \in \Gamma_x = G_x$, we can find $(0, t_n) \in \Gamma_{x_n} = G_{x_n}$ such that $|t_n - t| \to 0$. ■

We now observe that m_s induces the SM_2 topology.

Theorem 12.10.1. (the Hausdorff metric m_s induces the SM_2 topology) *If $\mu_s(x_n, x) \to 0$ as $n \to \infty$, then $\mu_s(x, x_n) \to 0$ as $n \to \infty$. Hence, $\mu_s(x_n, x) \to 0$ as $n \to \infty$ if and only if $m_s(x_n, x) \to 0$ as $n \to \infty$.*

It may seem natural to consider a weak M_2 topology defined by the metric $m_w(x_1, x_2)$ in (10.4), but this does not yield a desirable topology.

Example 12.10.2. *Deficiency of the m_w metric.* To see a deficiency of the m_w metric in (10.4), we show that convergence $d_s(x_n, x) \to 0$ as $n \to \infty$, which implies $m_s(x_n, x) \to 0$, does not imply $\mu_w(x, x_n) \to 0$ or $m_w(x_n, x) \to 0$ as $n \to \infty$. For this purpose, consider x and x_n, $n \ge 1$, in $D([0,2], \mathbb{R}^2)$ defined by $x^1 = x^2 = I_{[1,2]}$ and $x_n^1(t) = x_n^2(t) = n(t-1)I_{[1,1+n^{-1})}(t) + I_{[1+n^{-1},2]}(t)$ for $n \ge 1$. Then $d_s(x_n, x) \to 0$ as $n \to \infty$ and thus $m_s(x_n, x) \to 0$ as $n \to \infty$, but the thick ranges of the graphs of x and x_n are $\rho(G_x) = [0,2] \times [0,2]$ and $\rho(G_{x_n}) = \{\alpha(0,0) + (1-\alpha)(2,2) : 0 \le \alpha \le 1\}$, so that $\mu_w(x, x_n) \not\to 0$ and $m_w(x_n, x) \not\to 0$ as $n \to \infty$. in this case, $\mu_w(x_n, x) \to 0$ as $n \to \infty$.

We now observe that m_p induces the WM_2 topology.

Theorem 12.10.2. (WM_2 is the product topology) *$\mu_w(x_n, x) \to 0$ as $n \to \infty$ for μ_w in* (10.2) *if and only if $m_p(x_n, x) \to 0$ as $n \to \infty$ for m_p in* (10.5), *so that the WM_2 topology on $D([0,T], \mathbb{R}^k)$ coincides with the product topology.*

We conclude this section by summarizing the relations among the primary distances under consideration in the following theorem.

Theorem 12.10.3. (comparison of distances) *For each* $x_1, x_2 \in D$,

$$d_p \leq d_w \leq d_s \leq d_{J_1} \leq \| \cdot \| ,$$

$$m_p \leq d_p \quad \text{and} \quad m_p \leq m_s \leq d_s .$$

Remark 12.10.1. *Relating the J and M topologies.* The J_i topologies were related to the M_i topologies in a revealing way in Pomarede (1976). The J_2 topology is induced by the Hausdorff metric on the space of incomplete graphs; that shows that J_2 is stronger than M_2. Similarly, the J_1 topology can be defined in terms of a metric applied to parametric representations of the incomplete graphs; that shows that J_1 is stronger than M_1. ■

12.11. Alternative Characterizations of M_2 Convergence

We now give alternative characterizations of the SM_2 and WM_2 topologies.

12.11.1. M_2 Parametric Representations

We first observe that the SM_2 and WM_2 topologies can be defined just like the SM_1 and WM_1 topologies in Section 12.3. For this purpose, we say that a *strong* M_2 (SM_2) *parametric representation* of x is a continuous function (u, r) mapping $[0, 1]$ onto Γ_x such that r is nondecreasing. A *weak* M_2 (WM_2) *parametric representation* of x is a continuous function mapping $[0, 1]$ into G_x such that r is nondecreasing with $r(0) = 0$, $r(1) = T$ and $u(1) = x(T)$. The corresponding M_1 parametric representations are nondecreasing using the order defined on the graphs Γ_x and G_x in Section 2. In contrast, only the component function r is nondecreasing in the M_2 parametric representations. Let $\Pi_{s,2}(x)$ and $\Pi_{w,2}(x)$ be the sets of all SM_2 and WM_2 parametric representations of x.

Paralleling (3.7) and (3.8), define the distance functions

$$d_{s,2}(x_1, x_2) \equiv \inf_{\substack{(u_j, r_j) \in \Pi_{s,2}(x_j) \\ j=1,2}} \{\|u_1 - u_2\| \vee \|r_1 - r_2\|\} \tag{11.1}$$

and

$$d_{w,2}(x_1, x_2) \equiv \inf_{\substack{(u_j, r_j) \in \Pi_{w,2}(x_j) \\ j=1,2}} \{\|u_1 - u_2\| \vee \|r_1 - r_2\|\} . \tag{11.2}$$

We then can say that $x_n \to x$ as $n \to \infty$ for a sequence or net $\{x_n\}$ if $d_{s,2}(x_n, x) \to 0$ or $d_{w,2}(x_n, x) \to 0$ as $n \to \infty$. A difficulty with this approach, just as for the WM_1 topology, is that neither $d_{s,2}$ nor $d_{w,2}$ is a metric.

Example 12.11.1. *Neither $d_{s,2}$ nor $d_{w,2}$ is a metric.* To see that neither $d_{s,2}$ nor $d_{w,2}$ is a metric, consider real-valued functions, so that $d_{s,2} = d_{w,2} = d_2$. Let $x = 2I_{[1,2]}$, $x_{2n+1} = 2I_{[1-2n^{-1},1-n^{-1})} + 2I_{[1,2]}$ and $x_{2n} = I_{[1-n^{-1},1)} + 2I_{[1,2]}$ in $D([0,2],\mathbb{R})$ for $n \ge 3$. For each n, it is possible to choose parametric representations of x_n and x such that $d_2(x_{2n+1}, x) \le 2n^{-1}$ and $d_2(x_{2n}, x) \le n^{-1}$. However, $d_2(x_{2n}, x_{2n+1}) \ge 1$ for all n. We cannot simultaneously match the points in $\{2\} \times [1-2n^{-1}, 1-n^{-1}] \subseteq \Gamma_{2x_n+1}$ to $\{2\} \times [1,2] \subseteq \Gamma_{x_{2n}}$ and the points in $\{0\} \times [(1-n^{-1}),1) \subseteq \Gamma_{x_{2n+1}}$ to $\{0\} \times [0, 1-n^{-1}) \subseteq \Gamma_{x_{2n}}$ because the times are inconsistently ordered. ■

12.11.2. SM_2 Convergence

We now establish the equivalence of several alternative characterizations of convergence in the SM_2 topology. To have a characterization involving the local behavior of the functions, we use the uniform-distance function $\bar{w}_s(x, x_2, t, \delta)$ in (4.6). We also use the related uniform-distance functions

$$\bar{w}_s(x_1, x_2, \delta) \equiv \sup_{0 \le t \le T} \bar{w}(x_1, x_2, t, \delta) \ . \tag{11.3}$$

$$\bar{w}_s^*(x_1, x_2, t, \delta) \equiv \|x_1(t) - [x_2((t-\delta) \vee 0), x_2((t+\delta) \wedge T)]\| \tag{11.4}$$

$$\bar{w}_s^*(x_1, x_2, \delta) \equiv \sup_{0 \le t \le T} \bar{w}_s^*(x_1, x_2, t, \delta) \ . \tag{11.5}$$

We now define new oscillation functions. The first is

$$\bar{w}_s^*(x, t, \delta) \equiv \sup\{\|x(t) - [x(t_1), x(t_2)]\|\} \ , \tag{11.6}$$

where the supremum is over

$$0 \vee (t-\delta) \le t_1 \le [0 \vee (t-\delta)] + \delta/2 \text{ and } [T \wedge (t+\delta)] - \delta/2 \le t_2 \le (t+\delta) \wedge T.$$

The second is

$$\bar{w}_s^*(x, \delta) \equiv \sup_{0 \le t \le T} \bar{w}_s^*(x, t, \delta) \ . \tag{11.7}$$

The uniform-distance function $\bar{w}_s^*(x_1, x_2, \delta)$ in (11.5) and the oscillation function $\bar{w}_s^*(x, \delta)$ in (11.7) were originally used by Skorohod (1956).

As before, T need not be a continuity point of x in $D([0,T],\mathbb{R}^k)$. Unlike for the M_1 topology, we can have $x_n \to x$ in (D, M_2) without having $x_n(T) \to x(T)$.

Example 12.11.2. *M_2 convergence does not imply pointwise convergence at the right endpoint.* To see that M_2 convergence does not imply that $x_n(T) \to x(T)$, let $x(0 = x(T-) = 0$, $x(T) = 1$,

$$x_n(0) = x_n(T - 2n^{-1}) = x_n(T) = 0$$

and $x_n(T - n^{-1}) = 1$ for $n \geq 1$ with x and x_n defined by linear interpolation elsewhere. It is easy to see that $x_n \to x$ (M_2), but $x_n(T) \not\to x(T)$.

Let $v(x, A)$ represent the oscillation of x over the set A as in (2.5).

Theorem 12.11.1. (characterizations of SM_2 convergence) *The following are equivalent characterizations of $x_n \to x$ as $n \to \infty$ in (D, SM_2):*

(i) $d_{s,2}(x_n, x) \to 0$ *as* $n \to \infty$ *for* $d_{s,2}$ *in* (11.1)*; i.e., for any* $\epsilon > 0$ *and* n *sufficiently large, there exist* $(u, r) \in \Pi_{s,2}(x)$ *and* $(u_n, r_n) \in \Pi_{s,2}(x_n)$ *such that* $\|u_n - u\| \vee \|r_n - r\| < \epsilon$.

(ii) $m_s(x_n, x) \to 0$ *as* $n \to \infty$ *for the metric* m_s *in* (10.3).

(iii) $\mu_s(x_n, x) \to 0$ *as* $n \to \infty$ *for* μ_s *in* (10.1).

(iv) Given $\bar{w}_s(x_1, x_2, \delta)$ *defined in* (11.3),

$$\lim_{\delta \downarrow 0} \overline{\lim_{n \to \infty}} \; \bar{w}_s(x_n, x, \delta) = 0 \ .$$

(v) For each t, $0 \leq t \leq T$,

$$\lim_{\delta \downarrow 0} \overline{\lim_{n \to \infty}} \; \bar{w}_s(x_n, x, t, \delta) = 0$$

for $\bar{w}_s(x_1, x_2, t, \delta)$ *in* (4.6).

(vi) For all $\epsilon > 0$ *and all* n *sufficiently large, there exist finite ordered subsets* A *of* Γ_x *and* A_n *of* Γ_{x_n}, *as in* (3.9) *where* $(z_1, t_1) \leq (z_2, t_2)$ *if* $t_1 \leq t_2$, *of the same cardinality such that* $\hat{d}(A, \Gamma_x) < \epsilon$, $\hat{d}(A_n, \Gamma_{x_n}) < \epsilon$ *and* $d^*(A, A_n) < \epsilon$ *for* $\hat{d}$ *in* (3.10) *and* d^* *in* (5.6).

(vii) Given $\bar{w}_s^*(x_1, x_2, \delta)$ *defined in* (11.5),

$$\lim_{\delta \downarrow 0} \overline{\lim_{n \to \infty}} \; \bar{w}_s^*(x_n, x, \delta) = 0 \ .$$

(viii) $x_n(t) \to x(t)$ *as* $n \to \infty$ *for each* t *in a dense subset of* $[0, T]$ *including* 0 *and*

$$\lim_{\delta \downarrow 0} \overline{\lim_{n \to \infty}} \ \bar{w}_s^*(x_n, \delta) = 0$$

for $\bar{w}_s^*(x, \delta)$ *in* (11.7).

Remark 12.11.1. The equivalence (iii)↔(vii)↔(viii) was established by Skorohod (1956). ■

Remark 12.11.2. There is no analog to characterization (v) involving $\bar{w}_s^*(x_n, x, t, \delta)$ in (11.4) instead of $\bar{w}_s(x_n, x, t, \delta)$. For $t \in Disc(x)^c$,

$$\lim_{\delta \downarrow 0} \overline{\lim_{n \to \infty}} \ \bar{w}_s^*(x_n, x, t, \delta) = 0$$

implies pointwise convergence $x_n(t) \to x(t)$, but not the local uniform convergence in Theorem 12.4.1. ■

12.11.3. WM_2 Convergence

Corresponding characterizations of WM_2 convergence follow from Theorem 12.11.1 because the WM_2 topology is the same as the product topology, by Theorem 12.10.2. Let

$$\bar{w}_w(x_1, x_2, \delta) \equiv \sup_{0 \le t \le T} \bar{w}_w(x_1, x_2, t, \delta) \tag{11.8}$$

for $\bar{w}_w(x_1, x_2, t, \delta)$ in (4.7).

Theorem 12.11.2. (characterizations of WM_2 convergence) *The following are equivalent characterizations of* $x_n \to x$ *as* $n \to \infty$ *in* (D, WM_2)*:*

(i) $d_{w,2}(x_n, x) \to 0$ *as* $n \to \infty$ *for* $d_{w,2}$ *in* (11.2)*; i.e., for any* $\epsilon > 0$ *and all* n *sufficiently large, there exist* $(u, r) \in \Pi_{w,2}(x)$ *and* $(u_n, r_n) \in \Pi_{w,2}(x_n)$ *such that* $\|u_n - u\| \vee \|r_n - r\| < \epsilon$.

(ii) $m_p(x_n, x) \to 0$ *as* $n \to \infty$ *for the metric* m_p *in* (10.5).

(iii) Given $\bar{w}_w(x_1, x_2, \delta)$ *defined in* (11.8),

$$\lim_{\delta \downarrow 0} \overline{\lim_{n \to \infty}} \ \bar{w}_w(x_n, x, \delta) = 0 \ .$$

(iv) For each t, $0 \le t \le T$,

$$\lim_{\delta \downarrow 0} \overline{\lim_{n \to \infty}} \ \bar{w}_w(x_n, x, t, \delta) = 0 \ .$$

(v) For all $\epsilon > 0$ *and all sufficiently large* n*, there exist finite ordered subsets* A *of* G_x *and* A_n *of* Γ_{x_n}*, of common cardinality* m *as in* (3.9) *with* $(z_1, t_1) \leq (z_2, t_2)$ *if* $t_1 \leq t_2$*, such that* $\hat{d}(A, G_x) < \epsilon$*,* $\hat{d}(A_n, \Gamma_{x_n}) < \epsilon$ *and* $d^*(A, A_n) < \epsilon$ *for all* $n \geq n_0$*, for* $\hat{d}$ *in* (5.13) *and* d^* *in* (5.6).

Theorem 12.11.2 and Section 12.4 show that all forms of M convergence imply uniform convergence to continuous limit functions.

Corollary 12.11.1. (from WM_2 convergence to uniform convergence) *Suppose that* $m_p(x_n, x) \to 0$ *as* $n \to \infty$.

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

$$\lim_{\delta \downarrow 0} \overline{\lim_{n \to \infty}} \; v(x_n, x, t, \delta) = 0 \; .$$

(ii) If $x \in C$*, then* $\lim_{n \to \infty} \|x_n - x\| = 0$.

Proof. For (i) combine Theorems 12.4.1 and 12.11.2. For (ii) add Lemma 12.4.2. ∎

Convergence in WM_2 has the advantage that jumps in the converging functions must be inherited by the limit function.

Corollary 12.11.2. (inheritance of jumps) *If* $x_n \to x$ *in* (D, WM_2)*,* $t_n \to t$ *in* $[0, T]$ *and* $x_n^i(t_n) - x_n^i(t_n-) \geq c > 0$ *for all* n*, then* $x^i(t) - x^i(t-) \geq c$.

Proof. Apply Theorem 12.11.2 (iv). ∎

Let $J(x)$ be the maximum magnitude (absolute value) of the jumps of the function x in D. We apply Corollary 12.11.2 to show that J is upper semicontinuous.

Corollary 12.11.3. (upper semicontinuity of J) *If* $x_n \to x$ *in* (D, M_2)*, then*

$$\overline{\lim_{n \to \infty}} \; J(x_n) \leq J(x) \; .$$

Proof. Suppose that $x_n \to x$ in (D, WM_2) and there exists a subsequence $\{x_{n_k}\}$ such that $J(x_{n_k}) \to c$. Then there exist further subsubsequences $\{x_{n_{k_j}}\}$ and $\{t_{n_{k_j}}\}$, and a coordinate i, such that $t_{n_{k_j}} \to t$ for some $t \in [0, T]$ and $|x^i_{n_{k_j}}(t_{n_{k_j}}) - x^i_{n_{k_j}}(t_{n_{k_j}}-)| \to c$. Then Corollary 12.11.2 implies that $|x^i(t) - x^i(t-)| \geq c$. ∎

12.11.4. Additional Properties of M_2 Convergence

We conclude this section by discussing additional properties of the M_2 topologies. First we note that there are direct M_2 analogs of the M_1 results in Theorems 12.6.1, 12.7.1, 12.7.2 and 12.7.3.

Theorem 12.11.3. (extending SM_2 convergence to product spaces) *Suppose that $m_s(x_n, x) \to 0$ in $D([0,T], \mathbb{R}^k)$ and $m_s(y_n, y) \to 0$ in $D([0,T], \mathbb{R}^l)$ as $n \to \infty$. If*

$$Disc(x) \cap Disc(y) = \phi\,,$$

then

$$m_s((x_n, y_n), (x, y)) \to 0 \quad in \quad D([0,T], \mathbb{R}^{k+l}) \quad as \quad n \to \infty\,.$$

Corollary 12.11.4. (from WM_2 convergence to SM_2 convergence when the limit is in D_1) *If $m_p(x_n, x) \to 0$ as $n \to \infty$ and $x \in D_1$, then $m_s(x_n, x) \to 0$ as $n \to \infty$.*

Theorem 12.11.4. (Lipschitz property of linear functions of the coordinate functions) *For any x_1, $x_2 \in D([0,T], \mathbb{R}^k)$ and $\eta \in \mathbb{R}^k$,*

$$m(\eta x_1, \eta x_2) \le (\|\eta\| \vee 1) m_s(x_1, x_2)\ .$$

We have an analog of Corollary 12.7.1 for the M_2 topology.

Corollary 12.11.5. (SM_2-continuity of addition) *If $m_s(x_n, x) \to 0$ and $m_s(y_n, y) \to 0$ in $D([0,T], \mathbb{R}^k)$ and*

$$Disc(x) \cap Disc(y) = \phi\,,$$

then

$$m_s(x_n + y_n, x + y) \to 0 \quad in \quad D([0,T], \mathbb{R}^k)\,.$$

Theorem 12.11.5. (characterization of SM_2 convergence by convergence of all linear functions of the coordinates) *There is convergence $x_n \to x$ in $D([0,T], \mathbb{R}^k)$ as $n \to \infty$ in the SM_2 topology if and only if $\eta x_n \to \eta x$ in $D([0,T], \mathbb{R}^1)$ as $n \to \infty$ in the M_2 topology for all $\eta \in \mathbb{R}^k$.*

Just as with the M_1 topology, we can get convergence of sums under more general conditions than in Corollary 12.11.5. It suffices to have the jumps of x^i and y^i have common sign for all i. We can express this property by the condition (7.2).

Theorem 12.11.6. (continuity of addition at limits with jumps of common sign) *If* $x_n \to x$ *and* $y_n \to y$ *in* $D([0,T], \mathbb{R}^k, SM_2)$ *and if condition* (7.2) *holds, then*

$$x_n + y_n \to x + y \quad in \quad D([0,T], \mathbb{R}^k, SM_2) \ .$$

We now apply Theorem 12.11.5 to extend a characterization of convergence due to Skorohod (1956) to $\mathbb{R}^k$-valued functions. For each $x \in D([0,T], \mathbb{R}^1)$ and $0 \leq t_1 < t_2 \leq T$, let

$$M_{t_1,t_2}(x) \equiv \sup_{t_1 \leq t \leq t_2} x(t) \ . \tag{11.9}$$

The proof exploits the SM_2 analog of Corollary 12.9.1.

Theorem 12.11.7. (characterization of SM_2 convergence in terms of convergence of local extrema) *There is convergence* $m_s(x_n, x) \to 0$ *as* $n \to \infty$ *in* $D([0,T], \mathbb{R}^k)$ *if and only if*

$$M_{t_1,t_2}(\eta x_n) \to M_{t_1,t_2}(\eta x) \quad as \quad n \to \infty$$

for all $\eta \in \mathbb{R}^k$ *and all points* $t_1,\ t_2 \in \{T\} \cup Disc(x)^c$ *with* $t_1 < t_2$.

We can apply the characterization of M_2 convergence in Theorem 12.11.7 to show the preservation of convergence under bounding functions in the M_2 topology.

Corollary 12.11.6. (preservation of WM_2 convergence within bounding functions) *Suppose that*

$$y_n^i(t) \leq x_n^i(t) \leq z_n^i(t)$$

for all $t \in [0,T]$, $1 \leq i \leq k$, *and all* n. *If* $m_p(y_n, x) \to 0$ *and* $m_p(z_n, x) \to 0$ *as* $n \to \infty$, *then* $m_p(x_n, x) \to 0$ *as* $n \to \infty$.

Example 12.11.3. *Failure with other topologies.* To see that there is no analog of Corollary 12.11.6 for the M_1 and J_1 topologies, for $n \geq 1$, let $x = I_{[1,2]}$, $y_n = I_{[1+n^{-1},2]}$, $z_n = I_{[1-n^{-1},2]}$,

$$x_n(0) = x_n(1 - n^{-1}) = x_n(1 - (3n)^{-1}) = x_n(1 - (5n)^{-1}) = 0$$

and

$$x_n(1 - (2n)^{-1}) = x_n(1 - (4n)^{-1}) = x_n(1) = x_n(2) = 1 \ ,$$

with x_n defined by linear interpolation elsewhere. Then $y_n(t) \leq x_n(t) \leq z_n(t)$ for all t and n, $y_n \to x$ and $z_n \to x$ as $n \to \infty$ in $D([0,2], \mathbb{R})$ with the J_1 topology, while $x_n \to x$ with the M_2 topology, but not with the M_1, J_2 and J_1 topologies.

12.12. Compactness

We now characterize compact subsets in $D \equiv D([0,T],\mathbb{R}^k)$ in the M topologies, closely following Section 2.7 of Skorohod (1956). To do so, we define new oscillation functions that include more control of the behavior of the functions at the interval endpoints 0 and T. First let

$$\bar{w}_w^*(x,\delta) \equiv \max_{1\le i\le k} \bar{w}_s^*(x^i,\delta) \tag{12.1}$$

for $\bar{w}_s^*$ in (11.7). Given $w_s(x,\delta)$ in (5.1), $w_w(x,\delta)$ in (5.12), $\bar{w}_s^*(x,\delta)$ in (11.7), $\bar{w}_w^*(x,\delta)$ in (12.1) and $\bar{v}(x,t,\delta)$ in (4.3), let

$$\begin{aligned} w_s'(x,\delta) &\equiv w_s(x,\delta) \vee \bar{v}(x,0,\delta) \vee \bar{v}(x,T,\delta)\ , && (12.2)\\ w_w'(x,\delta) &\equiv w_w(x,\delta) \vee \bar{v}(x,0,\delta) \vee \bar{v}(x,T,\delta)\ , && (12.3)\\ \bar{w}_s'(x,\delta) &\equiv \bar{w}_s^*(x,\delta) \vee \bar{v}(x,0,\delta) \vee \bar{v}(x,T,\delta)\ , && (12.4)\\ \bar{w}_w'(x,\delta) &\equiv \bar{w}_w^*(x,\delta) \vee \bar{v}(x,0,\delta) \vee \bar{v}(x,T,\delta)\,. && (12.5)\end{aligned}$$

Since

$$\bar{w}_w^*(x,\delta) \le \bar{w}_s^*(x,\delta) \quad \text{and} \quad \bar{w}_w^*(x,\delta) \le w_w(x,\delta) \le w_s(x,\delta)$$

for all $x \in D$ and $\delta > 0$,

$$\bar{w}_w'(x,\delta) \le \bar{w}_s'(x,\delta) \quad \text{and} \quad \bar{w}_w'(x,\delta) \le w_w'(x,\delta) \le w_s'(x,\delta)$$

for all $x \in D$ and $\delta > 0$.

We start by stating a characterization of WM_2 convergence. The proof draws on Theorem 12.11.1.

Theorem 12.12.1. (another characterization of WM_2 convergence) *If $\{x_n\}$ is a sequence in D such that $x_n(t)$ converges as $n \to \infty$ for all t in a dense subset of $[0,T]$ including 0 and T and*

$$\lim_{\delta\downarrow 0}\ \overline{\lim_{n\to\infty}}\ \bar{w}_w'(x_n,\delta) = 0 \tag{12.6}$$

for $\bar{w}'$ in (12.5), then there exists $x \in D$ such that $m_p(x_n,x) \to 0$.

Example 12.12.1. *Need for the $\bar{v}$ terms.* To see the need for the terms $\bar{v}(x,0,\delta)$ and $\bar{v}(x,T,\delta)$ in $\bar{w}_w'(x,\delta)$, let $x_n(0) = 1$, $x_n(n^{-1}) = x_n(1) = 0$ with x_n defined by linear interpolation elsewhere on $[0,1]$. Then $\bar{w}_s^*(x_n,\delta) = 0$ for all n and δ, but $\{x_n : n \ge 1\}$ does not converge and is not compact in $D([0,1],\mathbb{R},M_2)$. Since $\sup_n \bar{v}(x_n,0,\delta) = 1$ for all $\delta > 0$, (12.6) fails.

Corollary 12.12.1. (new characterizations of convergence in other topologies) *If the conditions of Theorem 12.12.1 hold with $\bar{w}'_w$ in* (12.5) *replaced by $\bar{w}'_s$ in* (12.4)*, w'_w in* (12.3) *or w'_s in* (12.2)*, then the convergence can be strengthened to SM_2, WM_1 or SM_1, respectively.*

Theorem 12.12.2. (characterizations of compactness) *A subset A of D has compact closure in the SM_1, WM_1, SM_2 or WM_2 topology if*

$$\sup_{x \in A}\{\|x\|\} < \infty \tag{12.7}$$

and

$$\lim_{\delta \downarrow 0} \sup_{x \in A}\{w'(x, \delta)\} < \infty \ , \tag{12.8}$$

where w' is w'_s in (12.2) *for SM_1, w'_w in* (12.3) *for WM_1, $\bar{w}'_s$ in* (12.4) *for SM_2 and $\bar{w}'_w$ in* (12.5) *for SM_2. The conditions are necessary for SM_1 and WM_1.*

Example 12.12.2. *The conditions are not necessary for M_2.* To see that the conditions in Theorem 12.12.2 are not necessary for the M_2 topologies, for $s \in [1/4, 1/2]$, let

$$x_s = I_{[s, 1/4+s/2)} + I_{[1/2,1]}$$

in $D([0,1], \mathbb{R})$. The set $\{x_s : 1/4 \le s \le 1/2\}$ is clearly M_2 compact, but

$$\sup_{1/4 \le s \le 1/2} \bar{w}_w(x_s, \delta) = 1$$

for all δ, $0 < \delta < 1/4$. ■

Compactness characterizations on D translate into tightness characterizations for sets of probability measures on D. Recall from Chapter 11 that a set A of probability measures on a metric space (S, m) is said to be tight if for all $\epsilon > 0$ there exists a compact subset K of (S, m) such that

$$P(K) > 1 - \epsilon \quad \text{for all} \quad P \in A \ .$$

Theorem 12.12.3. (characterizations of tightness) *A sequence $\{P_n : n \ge 1\}$ of probability measures on D with the SM_1, WM_1, SM_2 or WM_2 topology is tight if the following two conditions hold:*

(i) For each $\epsilon > 0$, there exists c such that

$$P_n(\{x \in D : \|x\| > c\}) < \epsilon, \quad n \geq 1 .$$

(ii) For each $\epsilon > 0$ and $\eta > 0$, there exists $\delta > 0$ such that

$$P_n(\{x \in D : w'(x, \delta) \geq \eta) < \epsilon, \quad n \geq 1 ,$$

for w' being the appropriate oscillation function in (12.2)–(12.5). *The conditions are also necessary for the SM_1 and WM_1 topologies.*

Proof. Suppose that conditions (i) and (ii) hold, where w' is w'_s in (12.2) for SM_1 w'_w in (12.3) for WM_1, $\bar{w}'_s$ in (12.4) for SM_2 and $\bar{w}'_w$ in (12.5) for WM_2. For $\epsilon > 0$ given, choose c and δ_k such that $P_n(A_k^c) < \epsilon 2^{-(k+1)}$, $k \geq 0$, where

$$A_0 = \{x \in D : \|x\| \leq c\} \tag{12.9}$$

and

$$A_k = \{x \in D : w'(x, \delta_k) < k^{-1}\}, \quad k \geq 1 . \tag{12.10}$$

Then let $A = \cap_{k \geq 0} A_k$. By the construction,

$$P_n(A^c) = P_n\left(\cup_{k \geq 0} A_k^c\right) \leq \sum_{k=0}^{\infty} P_n(A_k^c) \leq \epsilon . \tag{12.11}$$

Since $A \subseteq A_0$ and

$$\lim_{\delta \downarrow 0} \sup_{x \in A} w'(x, \delta) = 0 , \tag{12.12}$$

the set A has compact closure by Theorem 12.12.1. Going the other way, assume that the topology is SM_1 or WM_1 and suppose that $\{P_n : n \geq 1\}$ is tight, so that for any $\epsilon > 0$ there exists a compact subset K of D such that $P_n(K) > 1 - \epsilon$. By Theorem 12.12.2, for any $\eta > 0$ given, $K \subseteq \{x : \|x\| \leq c\}$ for some c and $K \subseteq \{x : w'(x, \delta) \leq \eta\}$ for small enough δ; by the monotonicity of $w'(x, \delta)$ in δ for the SM_1 and WM_1 topologies. Hence conditions (i) and (ii) hold for all n. ∎

For an alternative characterization of M_1 tightness in $D([0, T], \mathbb{R})$, see Avram and Taqqu (1989).