# Chapter 6

# The Space D

## 6.1. Introduction

This chapter contains proofs omitted from Chapter 12 of the book, with the same title. For convenience, the theorems are restated here. The section and theorem numbers parallel Chapter 12 of the book, so the proofs should be easy to find.

*Here is how the present chapter is organized:* We start in Section 6.2 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 6.3 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 6.4 we discuss local uniform convergence at continuity points and relate it to oscillation functions used to characterize different forms of convergence.

In Section 6.5 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 of the book tend to be more than illustrative: The topologies are characterized by the local behavior in the neighborhood of single discontinuities.

In Section 6.6 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 6.7 we study how $SM_1$ convergence in $D([0,T],\mathbb{R}^k)$ can be characterized by associated limits of mappings.

In Section 6.8 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 equation (3.2) of Section 3.3 in the book, 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 6.9 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 6.10 we introduce the strong and weak versions of the $M_2$ topology, denoted by $SM_2$ and $WM_2$. In Section 6.11 we provide alternative characterizations of these topologies and discuss additional properties.

Finally, in Section 6.12 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 of the book.

## 6.2. Regularity Properties of $D$

Recall that $D \equiv D^k \equiv D([0,T],\mathbb{R}^k)$ is 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 $t \in (0,T]$:

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 $l_\infty$) 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 $\epsilon$, 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 6.2.1.** (the number of discontinuities of a given size) *For each $x \in D$ and $\epsilon > 0$, $Disc(x,\epsilon)$ is a finite subset of* $[0,T]$.

**Proof.** We will show that $Disc(x,\epsilon)$ being infinite contradicts the existence of limits from the left and right. If $Disc(x,\epsilon)$ were infinite, then there would exist $t \in [0,T]$ and a sequence $\{t_n : n \geq 1\}$ with $t_n \in Disc(x,\epsilon)$ for all $n$ and $t_n \downarrow t$ or $t_n \uparrow t$ as $n \to \infty$. Suppose that $t_n \downarrow t$; the other case is treated in the same way. Since $t_n \in Disc(x,\epsilon)$, we must have $\|x(t_n-) - x(t_n)\| \geq \epsilon$ for all $n$. Hence, there must exist another sequence $\{t'_n : n \geq 1\}$ such that $t_n > t'_n > t_{n+1} > t'_{n+1} > t$ for all $n$ and $\|x(t_n) - x(t'_n)\| > \epsilon/2$ for all $n$. However, that contradicts the existence of limits from the right at $t$. ∎

**Corollary 6.2.1.** (the number of discontinuities) *For each $x \in D$, $Disc(x)$ is either finite or countably infinite.*

**Proof.** Note that

$$Disc(x) = \bigcup_{n=1}^{\infty} Disc(x, n^{-1}) . \quad \blacksquare$$

We say that a function $x$ in $D$ is *piecewise-constant* if there are finitely many time points $t_i$ such that $0 \equiv t_0 < t_1 < \cdots < t_{m-1} \leq t_m \equiv T$ and $x$ is constant on the intervals $[t_{i-1}, t_i)$, $1 \leq i \leq m-1$, 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 6.2.2.** (approximation by piecewise-constant functions) *For each $x \in D$ and $\epsilon > 0$, there exists $x_c \in D_c$ such that $\|x - x_c\| < \epsilon$.*

**Proof.** We show how to construct $x_c$. Given $x$ and $\epsilon$, construct the subset $Disc(x, \epsilon)$, which is finite by Theorem 6.2.1. Due to the existence of limits, for each $t \in Disc(x, \epsilon)$ we can find $t_1 \equiv t_1(t)$ and $t_2 \equiv t_2(t)$ such that $t_1 < t < t_2$, $v(x, [t_1, t)) < \epsilon$, $v(x, [t, t_2]) < \epsilon$,

$$Disc(x, \epsilon) \cap [t_1, t) = \phi \quad \text{and} \quad Disc(x, \epsilon) \cap (t, t_2] = \phi\,.$$

For each $t \in Disc(x, \epsilon)$, let these points $t$, $t_1(t)$ and $t_2(t)$ all belong to $Disc(x_c)$; let $x_c(t') = x(t-)$ for $t' \in (t_1, t)$ and let $x_c(t') = x(t)$ for $t \in [t, t_2)$. Now let

$$A \equiv [0, T] - \bigcup_{t \in Disc(x,\epsilon)} (t_1(t), t_2(t))\ .$$

The set $A$ is a finite union of closed intervals. Consider any one of these intervals, say $[a, b]$. If $v(x; [a, b]) < \epsilon$, then it suffices to let $x_c(t) = x(t)$ for any $t \in [a, b]$, and not add any points to $Disc(x_c)$. Suppose that $v(x; [a, b]) \geq \epsilon$. For each $t \in [a, b]$, since $t \in Disc(x, \epsilon)^c$, it is possible to find an interval $(t_1(t), t_2(t))$, $[a, t_2(t))$ or $(t_1(t), b]$ containing $t$ such that $v(x, (t_1(t), t_2(t)) < \epsilon$. (The intervals $[a, t)$ and $(t, b]$ are open in the relative topology on $[a, b]$. Thus the collection of all these subintervals form an open cover of $[a, b]$.) Since $[a, b]$ is compact, there is a finite collection of these intervals covering $[a, b]$; i.e., there are points

$$a < t'_1 < t_1 < \cdots < t'_m < t_m < b$$

for $m \geq 1$ such that $[a, t_1)$, $(t'_1, t_2)$, $(t'_2, t_3), \ldots$, $(t'_{m-1}, t_m)$, $(t'_m, b]$ are in the finite collection. Necessarily, $t'_i < t_i$ for all $i$. It suffices to choose $t''_i \in (t'_i, t_i)$ for each $i$, $1 \leq i \leq m$, and let $t''_i \in Disc(x_c)$. We can let $x_c(t''_i) = x(t''_i)$ for each such $t''_i$. We have thus constructed $x_c \in D_c$ with $\|x - x_c\| < \epsilon$. ∎

## 6.3. Strong and Weak $M_1$ Topologies

### 6.3.1. Definitions

We start by making some definitions, repeating what is in the book. 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.,

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

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, $[a, b] \subseteq [[a, b]]$.

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 $1 \le i \le k$. 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 $\Gamma_x$ and $G_x$. We say that $(z_1, t_1) \le (z_2, t_2)$ if either (i) $t_1 < t_2$ or (ii) $t_1 = t_2$ and $|x^i(t_1-) - z_1^i| \le |x^i(t_1-) - z_2^i|$ for all $i$. The relation $\le$ induces a total order on $\Gamma_x$ 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 $\Gamma_x$ 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 $\Gamma_x$. 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 $1 \le i \le k$.

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 equation (3.4) in Section 3.3 of the book.

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.

### 6.3.2. Metric Properties

**Theorem 6.3.1.** (metric inducing $SM_1$) *$d_s$ is a metric on $D$.*

**Proof.** Only the triangle inequality is difficult. By Lemma 6.3.2 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 $\epsilon$ was arbitrary, the proof is complete. ∎

To prove Theorem 6.3.1, we use finite approximations to the graphs $\Gamma_x$. We first define an order-consistent distance between a graph and a finite subset. We use the notion of a finite ordered subset.

**Definition 6.3.1.** (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 $\Gamma_x$ 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 $\Gamma_x$ 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 show that finite ordered subsets $A$ can be chosen to make $\hat{d}(A, \Gamma_x)$ arbitrarily small.

**Lemma 6.3.1.** (finite approximations to graphs) *For any $x \in D$ and $\epsilon > 0$, there exists a finite ordered subset $A$ of $\Gamma_x$ such that $\hat{d}(A, \Gamma_x) < \epsilon$ for $\hat{d}$ in* (3.10).

**Proof.** First put finitely many points $(x(t_i), t_i)$ in $A$ to meet the requirement on the domain $[0, T]$, i.e., to have $0 = t_1 < t_2 < \cdots < t_m = T$ with $t_{i+1} - t_i < \epsilon$. We add additional points to account for the spatial component. For each $t \in Disc(x, \epsilon)$, choose the points $(x(t-), t)$, $(x(t), t)$ and finitely many points on the segment $[(x(t-), t), (x(t), t)]$ such that the distance between successive points is less than $\epsilon$. Since $x$ has left and right limits everywhere, there are open neighborhoods $(t_1, t)$ and $(t, t_2)$ of each $t \in Disc(x, \epsilon)$ such that

$$\sup\{\|x(t') - x(t'')\| : t_1 < t' < t'' < t\} < \epsilon$$

and

$$\sup\{\|x(t') - x(t'')\| : t < t' < t'' < t_2\} < \epsilon \ .$$

We thus can choose one more point, if needed, in each of the sets $\Gamma_x \cap [R^k \times (t_1, t)]$ and $\Gamma_x \cap [R^k \times (t, t_2)]$ to achieve the desired property over each open interval $(t_1, t_2)$ in $[0, T]$. The complement of the union of these finitely may open intervals in $[0, T]$ is a compact subset of $[0, T]$. Knowing that (i) all remaining discontinuities are of magnitude less than $\epsilon$ and (ii) limits exist everywhere from the left and right, we can conclude that there is a closed interval of positive length about each point in the compact set, where $x$ oscillates by less than $\epsilon$, i.e., $\sup\{\|x(t') - x(t'')\| < \epsilon$, where $t'$, $t''$ are points in the interval. However, by the compactness, only finitely many of these closed intervals cover the compact set. We add points $(x(t), t)$ to $A$ to ensure that there is at least one point $(z, t)$ for which $t$ is in one of these closed intervals. By this construction, $A$ is finite and $\hat{d}(A, \Gamma_x) < \epsilon$. ■

To complete the proof of Theorem 6.3.1, we need the following result, which we prove by applying Lemma 6.3.1.

**Lemma 6.3.2.** (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*

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

**Proof.** For $x_1, x_2 \in D$ and $\epsilon$ given, choose $(u_1', r_1') \in \Pi_s(x_1)$ and $(u_2', r_2') \in \Pi_s(x_2)$ such that

$$\|u_1' - u_2'\| \vee \|r_1' - r_2'\| < d_s(x_1, x_2) + \epsilon/4 \ . \tag{3.11}$$

Next apply Lemma 6.3.1 to find a finite ordered subsets $A_1 \subseteq \Gamma_{x_1}$ such that $\hat{d}(A_1, \Gamma_{x_1}) < \epsilon/4$. Next find a finite subset $S_1'$ of $[0, 1]$ of the same cardinality as $A_1$ such that $(u_1'(s), r_1'(s)) \in A_1$ for each $s \in S_1'$. Let $S_1$ be another finite subset of $[0,1]$ of the same cardinality as $A_1$ such that $(u_1(s), r_1(s)) \in A_1$ for each $s \in S_1$. Let $\lambda$ be a homeomorphism of $[0,1]$ such that $\lambda$ maps $S_1$ onto $S_1'$. Let $(u_2, r_2) = (u_2' \circ \lambda, r_2' \circ \lambda)$, where $\circ$ is the composition map. Trivially, by (3.11),

$$\|u_1' \circ \lambda - u_2' \circ \lambda\| \vee \|r_1' \circ \lambda - r_2' \circ \lambda\| < d_s(x_1, x_2) + \epsilon/4 \ .$$

Hence, it suffices to show that

$$\|u_1 - u_1' \circ \lambda\| \vee \|r_1 - r_1' \circ \lambda\| < 3\epsilon/4 \ . \tag{3.12}$$

First there is equality $u_1(s) = u_1'(\lambda(s))$ by construction at each $s \in S_1$. However, since $\hat{d}(A_1, \Gamma_x) < \epsilon/4$, (3.12) holds: For each $s \in [0,1]$, there is $s_i \in S_1$ such that $s_i \le s < s_{i+1}$ and

$$\begin{aligned} \|u_1(s) - u_1'(\lambda(s))\| &\le \|u_1(s) - u_1(s_i)\| + \|u_1(s_i) - u_1'(\lambda(s_i)) \\ &\quad + \|u_1'(\lambda(s_i)) - u_1'(\lambda(s))\| \le \epsilon/2 \, . \ \blacksquare \end{aligned}$$

We will show that the metric $d_s$ induces the standard $M_1$ topology defined by Skorohod (1956); see Theorem 6.5.1. 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 **??** 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

$$d_p(x_1, x_2) \equiv \max_{1 \le i \le k} d(x_1^i, x_2^i) \tag{3.13}$$

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), (3.13) 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 6.5 we will show that actually the $WM_1$ and product topologies coincide.

Example 12.3.1 of the book shows that $SM_1$ is strictly stronger than $WM_1$.

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

**Theorem 6.3.2.** (comparison of $J_1$ and $M_1$ metrics) *For each $x_1, x_2 \in D$,*

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

**Proof.** For any $x_1, x_2 \in D$ and $\lambda \in \Lambda$, we show how to define parametric representations $(u_j, r_j)$ in $\Pi_s(x_j)$ for $j = 1, 2$ such that

$$\|u_1 - u_2\| \vee \|r_1 - r_2\| = \|x_1 \circ \lambda - x_2\| \vee \|\lambda - e\| \ . \tag{3.14}$$

If, for any $\epsilon > 0$, we first choose $\lambda \in \Lambda$ so that

$$\|x_1 \circ \lambda - x_2\| \vee \|\lambda - e\| \le d_{J_1}(x_1, x_2) + \epsilon \ ,$$

the associated parametric representation yield

$$d_s(x_1, x_2) \le \|u_1 - u_2\| \vee \|r_1 - r_2\| \le d_{J_1}(x_1, x_2) + \epsilon \ .$$

Since $\epsilon$ is arbitrary, that will complete the proof. Suppose that

$$t_n \in Disc(x_1, x_2) \equiv Disc(x_1) \cup Disc(x_2) \ , \ \ n \ge 1 \ ,$$

where $t_n$ is ordered (indexed) first by the norm of the jump and then the location, with values closer to 0 occurring first. Associate with each time point $t_n$ a closed subinterval $[a_n, b_n]$ in $(0, 1)$ such that the subintervals are ordered, i.e., if $t_i < t_j < t_k$ are three points in $Disc(x_1, x_2)$, then $a_i < b_i < a_j < b_j < a_k < b_k$. Then let $r_2(s) = t_n$ for $a_n \le s \le b_n$. If $t \notin Disc(x_1, x_2)$ but $t_{n_k} \downarrow t$ as $n_k \to \infty$ for $t_{n_k} \in Disc(x_1, x_2)$, then let $r_2(s) = \lim_{n_k \to \infty} r_2(a_{n_k})$. Similarly, if $t \notin Disc(x_1, x_2)$ but $t_{n_k} \uparrow t$ as $n_k \to \infty$ for $t_{n_k} \in Disc(x_1, x_2)$, then let $r_2(s) = \lim_{n_k \to \infty} r_2(b_{n_k})$. Finally, let $r_2(s)$ be defined by linear interpolation in all remaining gaps. This makes $r_2$ continuous and nondecreasing. Having defined $r_2$, let $r_1 = \lambda \circ r_2$, $u_1(s) = (x_1 \circ r_1)(s)$ and $u_2(s) = (x_2 \circ r_2)(s)$ for all $s$, except $s \in (a_n, b_n)$ for some $n$. Within each subinterval $(a_n, b_n)$, let $u_1$ and $u_2$ be defined by linear interpolation from their values at the endpoints $a_n$ and $b_n$. This construction makes $(u_j, r_j) \in \Pi_s(x_j)$ for $j = 1, 2$ and yields (3.14), thus completing the proof. ■

### 6.3.3. Properties of Parametric Representations

We conclude this section by further discussing strong 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.,

$$\begin{aligned} &\text{(i) } r(s) < t \text{ for } s < s_l \\ &\text{(ii) } r(s) = t \text{ for } s_l \le s \le s_r \\ &\text{(iii) } r(s) > t \text{ for } s > s_r \ . \end{aligned} \tag{3.15}$$

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 (3.15) holds, there is a continuous nondecreasing function $\beta_t$ mapping $[0,1]$ onto $[0,1]$ such that

$$u(s) = \beta_t\left(\frac{s-s_l}{s_r-s_l}\right)u(s_r) + \left[1-\beta_t\left(\frac{s-s_l}{s_r-s_l}\right)\right]u(s_l) \quad \text{for} \quad s_l \le s \le s_r \ . \tag{3.16}$$

Condition (3.16) 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

$$\beta_t(s) = \frac{u^i(s) - u^i(s_l)}{u^i(s_r) - u^i(s_l)} \ . \tag{3.17}$$

We see that (3.16) and (3.17) are consistent in that

$$u^i(s) = \beta_t\left(\frac{s-s_l}{s_r-s_l}\right)u^i(s_r) + \left[1-\beta_t\left(\frac{s-s_l}{s_r-s_l}\right)\right]u^i(s_l) \tag{3.18}$$

for $\beta_t$ in (3.17). For another coordinate $j$, (3.16) and (3.17) imply that

$$u^j(s) = \left(\frac{u^i(s) - u^i(s_l)}{u^i(s_r) - u^i(s_l)}\right)u^j(s_r) + \left(\frac{u^i(s_r) - u^i(s)}{u^i(s_r) - u^i(s_l)}\right)u^j(s_l) \ . \tag{3.19}$$

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 (3.15) holds and

$$\begin{aligned} &\text{(i) } u(s) = u(s_l) \text{ for } s_l \le s \le s_l', \\ &\text{(ii) for each } i, \text{ either } u^i(s_l) < u^i(s) < u^i(s_r), \\ &\qquad \text{or } u^i(s_l) > u^i(s) > u^i(s_r) \text{ for } s_l' < s < s_r', \\ &\text{(iii) } u(s) = u(s_r) \text{ for } s_r' \le s \le s_r \ . \end{aligned} \tag{3.20}$$

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

**Proof.** Since $\Pi_s(x) \subseteq \Pi_w(x)$, we need to show that $(u,r) \in \Pi_w(x)$ is in $\Pi_s(x)$ for $x$ in $D^{(1)}$. Pick any $t \in Disc(x)$ and let $i$ be the coordinate of $x$ with a jump at $t$. We can then define the $\beta_t$ needed for (3.16) using (3.17). Since $u^j(s) = u^j(s_l) = u^j(s_r)$ for all $j$ with $j \neq i$, (3.19) and (3.16) are then satisfied. ■

**Corollary.** *For each $x \in D([0,T], \mathbb{R}^1)$, $\Pi_s(x) = \Pi_w(x)$.*

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

**Proof.** By the Corollary to Lemma 6.3.3, $\Pi_s(\eta x) = \Pi_w(\eta x)$. Hence, it suffices to show that $(\eta u, r) \in \Pi_w(\eta x)$. It is clear that $(\eta u, r)$ is continuous and nondecreasing. For $t \in Disc(\eta x)$, necessarily $t \in Disc(x)$. (We could have $t \in Disc(x)$ but $t \notin Disc(\eta x)$, but that does not concern us.) By (3.16), when $r(s) = t$,

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

which completes the proof. ■

## 6.4. 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 start by defining two basic *uniform-distance functions.* For $x_1, x_2 \in D$, $t \in [0,T]$ and $\delta > 0$, let

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

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

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

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

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 equation 5.3 in Section 11.5 of the book. 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 6.10 and 6.11. 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 6.4.1.** (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 $x$.

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 6.4.1 below.

**Theorem 6.4.1.** (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*

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

**Proof.** By Lemma 6.4.1, we have the implications (i) $\leftrightarrow$ (ii) $\leftrightarrow$ (iii) $\leftrightarrow$ (iv) and (ii) $\to$ (v) $\to$ (vi). Hence it suffices to show that (vi) $\to$ (i), which we now do. For $x, t \notin Disc(x)$ and $\epsilon > 0$ given, choose $\delta > 0$ so that $\bar{v}(x, t, \delta) < \epsilon$, which is possible since $t \notin Disc(x)$. Also let $\delta$ be sufficiently small so that $x_n(t_1') \to x(t_1')$ as $n \to \infty$ for all $t_1'$ in a dense subset of $[0 \vee (t-\delta), (t+\delta) \wedge T]$. Note that we can treat 0 and $T$ directly. For $t_1 \in (0 \vee (t-\delta), T \wedge (t+\delta))$ given, choose $t_1'$, $t_2'$ so that $0 \vee (t-\delta) < t_1' < t_1 < t_2' < (t+\delta) \wedge T$ and $x_n(t_j') \to x(t_j')$

as $n \to \infty$ for $j = 1, 2$. Then choose $n_0$ so that $\|x_n(t') - x(t'')\| < \epsilon$ for $t'' = 0$, $T$, $t_1'$ and $t_2'$ and $w_w(x_n, t, \delta) < \epsilon$ for $n \geq n_0$. Then, for $n \geq n_0$,

$$\begin{aligned}
\|x_n(t_1) - x(t_1)\| &\leq \|x_n(t_1) - x_n(t_1')\| + \|x_n(t_1') - x(t_1')\| + \|x(t_1') - x(t_1)\| \\
&\leq \|x_n(t_1) - x_n(t_1')\| + 2\epsilon \\
&\leq \|x_n(t_1) - [[x_n(t_1'), x_n(t_2')]]\| + \|x_n(t_1') - x_n(t_2')\| + 2\epsilon \\
&\leq w_w(x_n, t, \delta) + \|x_n(t_1') - x_n(t_2')\| + 2\epsilon \\
&\leq \|x_n(t_1') - x(t_1')\| + \|x(t_1') - x(t_2')\| \\
&\qquad + \|x(t_2') - x_n(t_2')\| + 3\epsilon \leq 6\epsilon \ .
\end{aligned}$$

It remains to consider $t = 0$ and $t = T$. The reasoning is the same for these two cases, so we consider only $t = 0$. For $t = 0$, note that

$$\|x_n(t_1) - x(t_1)\| \leq \|x_n(t_1) - x_n(0)\| + \|x_n(0) - x(0)\| + \|x(0) - x(t)\| \ . \quad (4.13)$$

The third term in (4.13) can be made small using the right continuity of $x$ at 0; the second term in (4.13) can be made small by the assumed convergence at 0; the first term in (4.13) can be made small by (4.12). ■

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

**Lemma 6.4.2.** (local uniform convergence everywhere in a compact interval) *If (4.8) holds for all $t \in [a, b]$, then*

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

**Proof.** By (4.8), for all $\epsilon > 0$ and $t \in [a, b]$, there exists $\delta(t)$ such that

$$\overline{\lim_{n \to \infty}} \ u(x_n, x, t, \delta(t)) < \epsilon \ .$$

For each $t$, there is thus uniform asymptotic closeness in the intervals $(0 \vee (t - \delta(t)), (t + \delta(t)) \wedge T)$. However, these intervals form an open cover of the interval $[a, b]$. Since $[a, b]$ is compact, there is a finite subcover. Hence, there is a $\delta' > 0$ such that

$$\overline{\lim_{n \to \infty}} \sup_{0 \vee (a-\delta')] \leq t \leq (b+\delta') \wedge T} \{\|x_n(t) - x(t)\|\} < \epsilon \ .$$

Since $\epsilon$ was arbitrary, this implies the desired conclusion. ■

## 6.5. Alternative Characterizations of $M_1$ Convergence

We now give alternative characterizations of $SM_1$ and $WM_1$ convergence.

### 6.5.1. $SM_1$ Convergence

We first establish alternative characterizations of $SM_1$ convergence or, equivalently, $d_s$-convergence. One characterization is 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 6.6. 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).

The following main result is proved in the book. It only remains to prove the supporting lemmas, which we do here.

**Theorem 6.5.1.** (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* $\Gamma_x$ *of cardinality* $m$ *as in* (3.9) and, for all $n \geq n_1$, finite ordered subsets $A_n$ of $\Gamma_{x_n}$ 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 6.5.1, we establish some preliminary results. We first show that $SM_1$ convergence implies local uniform convergence at all continuity points.

**Lemma 6.5.1.** (local uniform convergence) *If* $d_s(x_n, x) \to 0$ *as* $n \to \infty$, *then* (4.9) *holds for each* $t \notin Disc(x)$.

**Proof.** For $x$, $t \in Disc(x)^c$ and $\epsilon > 0$ given, choose $\delta > 0$ so that $\|x(t') - x(t)\| < \epsilon$ for $|t - t'| < \delta$. Then choose $n_0 \geq 4$, $(u_n, r_n) \in \Pi_s(x_n)$ and $(u, r) \in \Pi_s(x)$ such that

$$\|u_n - u\| \vee \|r_n - r\| < (\delta \wedge \epsilon)/4$$

for all $n \geq n_0$. Let $s_1$, $s_2$, $s_3$ be such that $r(s_1) = t - \delta/2$, $r(s_2) = t$ and $r(s_3) = t + \delta/2$. Then $r_n(s_1) < t < \delta/4$ and $r_n(s_3) > t + \delta/4$ for all $n \geq n_0$. Hence, for all $t' \in (t - \delta/4, t + \delta/4)$ and $n \geq n_0$ there exists $s_n$, $s_1 < s_n < s_3$, such that $(u_n(s_n), r_n(s_n)) = (x_n(t'), t')$. Hence,

$$\begin{aligned} \|x_n(t') - x(t')\| &= \|u_n(s_n) - u(s_2)\| + \|x(t) - x(t'')\| \\ &\leq \|u_n(s_n) - u(s_n)\| + \|u(s_n) - u(s_2)\| + \epsilon \\ &\leq (\delta \wedge \epsilon)/2 + 2\epsilon < 3\epsilon . \quad \blacksquare \end{aligned}$$

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

**Lemma 6.5.2.** (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* $0 \vee (t - \delta) \le t_1 < t_2 < t_3 \le (t + \delta) \wedge T$, *then* $\|z_2 - [z_1, z_3]\| \le w_s(x, \delta)$.

**Proof.** Suppose that $w_s(x, \delta) = \epsilon$. It suffices to show: (i) that $\|z_2 - [z_1, z_3]\| \le \epsilon$ when $\|z_2' - [z_1, z_3]\| \le \epsilon$, $\|z_2'' - [z_1, z_3]\| \le \epsilon$ and $z_2 \in [z_2', z_2'']$ and (ii) that $\|z_2 - [z_1, z_3]\| \le \epsilon$ when $\|z_2 - [z_1', z_3]\| \le \epsilon$, $\|z_2 - [z_1'', z_3\| \le \epsilon$ and $z_1 \in [z_1', z_1'']$. For (i), note that there exist $z'$, $z'' \in [z_1, z_3]$ such that $\|z_2' - z'\| \le \epsilon$ and $\|z_2'' - z''\| \le \epsilon$. Also there exists $\alpha$, $0 \le \alpha \le 1$ such that $z_2 = \alpha z_2' + (1-\alpha) z_2''$. Hence $\|z_2 - (\alpha z' + (1-\alpha) z'')\| \le \epsilon$, which implies that

$$\|z_2 - [z', z'']\| \le \|z_2 - [z_1, z_3]\| \le \epsilon \ .$$

For (ii), note first that there exist $z' \in [z_1', x_3]$ and $z'' \in [z_1'', z_3]$ such that $\|z_2 - z'\| \le \epsilon$ and $\|z_2 - z''\| \le \epsilon$. Hence, for any $z \in [z', z'']$, $\|z_2 - z\| \le \epsilon$. The desired $z$ lies on the intersection of $[z_1, z_3]$ and $[z', z'']$. That implies the desired conclusion. ■

**Lemma 6.5.3.** (asymptotic negligibility of the modulus) *For any* $x \in D$, $w_s(x, \delta) \downarrow 0$ *as* $\delta \downarrow 0$.

**Proof.** For any $\epsilon > 0$, choose $x_c \in D_c$ such that $\|x - x_c\| < \epsilon/2$, which is always possible by Theorem 6.2.2. Note that, for any $\delta > 0$,

$$w_s(x, \delta) \le w_s(x_c, \delta) + 2\|x - x_c\| \ ,$$

so that

$$w_s(x, \delta) \le w_s(x_c, \delta) + \epsilon \ .$$

Let $\eta$ be the minimum distance between successive discontinuities in $x_c$. Since $w_s(x_c, \delta) = 0$ when $\delta < \eta$, $w_s(x, \delta) < \epsilon$ when $\delta < \eta$. ■

**Proof of Theorem 6.5.1.** Contained in the book. ■

### 6.5.2. $WM_1$ Convergence

We now establish an analog of Theorem 6.5.1 for the $WM_1$ topology. Several alternative characterizations of $WM_1$ convergence will follow directly from Theorem 6.5.1 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

$$w_w(x, \delta) \equiv \sup_{0 \le t \le T} w_w(x, t, \delta) \tag{5.7}$$

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 6.3.1, 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

$$\hat{d}(A,G_x) \equiv \sup\{\|(z,t)-(z_i,t_i)\| \vee \|(z,t)-(z_{i+1},t_{i+1})\| : (z,t)\in G_x\} \quad (5.8)$$

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

**Theorem 6.5.2.** (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*

$$\lim_{\delta\downarrow 0}\ \overline{\lim_{n\to\infty}}\ w_w(x_n,\delta)=0\ . \quad (5.9)$$

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

$$\lim_{\delta\downarrow 0}\ \overline{\lim_{n\to\infty}}\ v(x_n,x,t,\delta)=0 \quad (5.10)$$

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

$$\lim_{\delta\downarrow 0}\ \overline{\lim_{n\to\infty}}\ w_w(x_n,t,\delta)=0 \quad (5.11)$$

*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* (5.8) *and* $d^*$ *in* (5.6).

**Proof.** (i)$\rightarrow$(ii). Since $d_p \le d_w$, (i)$\rightarrow$(ii) is immediate.
(ii)$\leftrightarrow$(iii). The implication (iii)$\rightarrow$(ii) is immediate, so we show (ii)$\rightarrow$(iii). By Lemma 6.5.1, $x_n^i(t) \rightarrow x^i(t)$ as $n \rightarrow \infty$ for each $t \in Disc(x^i)^c$, $1 \le i \le k$. That implies that $x_n(t) \rightarrow x(t)$ as $n \rightarrow \infty$ for each $t \in Disc(x)^c$. From Theorem 6.5.1, $d_p(x_n, x) \rightarrow 0$ as $n \rightarrow \infty$ also implies that

$$\lim_{\delta \downarrow 0} \overline{\lim_{n \rightarrow \infty}} \; w_s(x_n^i, \delta) = 0$$

for each $i$, $1 \le i \le k$, but that directly implies (5.9), because

$$\|x_n(t_2) - [[x_n(t_1), x_n(t_3)]]\| = \max_{1 \le i \le k} \|x_n^i(t_2) - [x_n^i(t_1), x_n^i(t_3)]\| \; , \quad (5.12)$$

so that

$$w_w(x_n, \delta) = \max_{1 \le i \le k} w_s(x_n^i, \delta) \; . \quad (5.13)$$

(iii)$\leftrightarrow$(iv). The equivalence between (iii) and (iv) holds by the same reasoning used to establish the equivalence of (iv) and (v) in Theorem 6.5.1.
(iii)$\rightarrow$(v). The proof of (iii)$\rightarrow$(v) parallels the proof of (iv)$\rightarrow$(vi) in Theorem 6.5.1, but requires some modifications. Paralleling the previous beginning, for $\epsilon > 0$ given, find $\eta < \epsilon/16$ and $n_0$ such that $w_w(x_n, \eta) < \epsilon/32$ for $n \ge n_0$. However, we do not next directly construct $A \in G_x$. Instead, just as with the $SM_1$ topology, we first construct the finite set $A$ of $\Gamma_x$ as before with the properties in the proof of Theorem 6.5.1. We denote this subset $A'$ to distinguish it from the desired subset $A$ of $G_x$. As before, for all $t_i \in S \cap A'$, let $n_1 \ge n_0$ be such that $\|x_n(t_i) - x(t_i)\| < \epsilon/32$ for all $i$, $1 \le i \le k$, and all $n \ge n_1$. We now want to construct the ordered subset $A_n$ in $G_{x_n}$. For $t \in S$, the construction is as before: $(z_{n,i}, t_{n,i}) = (x_n(t_i), t_i)$. Next suppose that (**??**) holds. Then $(z_{n,r}, t_{n,r})$ and $(z_{n,r+j+1}, t_{n,r+j+1})$ have been defined with respect to $A'$. We insert points into $A_n$ from $G_{x_n}$ appropriately spaced in between the two points. By construction specified before (but using the product segments),

$$\begin{aligned} &\|[[(x_n(t_r), t_r), (x_n(t_{r+j+1}), t_{r+j+1})]] \\ &-[[(x(t_r), t_r), (x(t_{r+j+1}), t_{r+j+1})]]\| < \epsilon/32 \end{aligned} \quad (5.14)$$

and

$$\|[[(x(t_r), t_r), (x(t_{r+j+1}), t_{r+j+1})]] - [[(x(t-), t), (x(t), t)]]\| < \epsilon/32 \; . \quad (5.15)$$

To simplify the discussion, suppose that $x^i(t-) \le x^i(t)$ for all $i$. (This is without loss of generality after redefining the order.) Consider an arbitrary nondecreasing (in the order on $G_{x_n}$) continuous curve in $G_{x_n}$ from

$(z_{n,r}, t_{n,r})$ to $(z_{n,r+j+1}, t_{n,r+j+1})$. Let $(z'_{n,r+1}, t'_{n,r+1})$ be the first point on this curve for which the $i^{\text{th}}$ coordinate first reaches $z^i_{n,r} + \epsilon/4$ for some $i$. Given $(z_{n,r+k}, t_{n,r+k})$, let $(z_{n,r+k+1}, t_{n,r_k+1})$ be the next point on the curve at which the $i^{\text{th}}$ coordinate first reaches $z^i_{n,r+k} + \epsilon/4$ for some $i$. Since $x^i(t-) \le x^i(t)$ for all $i$ and since $w_w(x_n, \eta) < \epsilon/32$, no coordinate of the curve in $G_{x_n}$ can decrease by more than $\epsilon/32$ over any subinterval, and thus from one points to the next in $A_n$. Continue in this manner for at most finitely many steps until the end point $(z_{n,r+j+1}, t_{n,r+j+1})$ is reached. The distance between successive points is $\epsilon/4$, while the distance between the last point inserted and $(z_{n,r+j+1}, t_{n,r+j+1})$ is less than $\epsilon/4$. Delete the first and last point inserted, so that all distances between successive points are between $\epsilon/4$ and $\epsilon/2$. In general, the number of inserted points is some finite number, not necessarily equal to $j$. These points are ordered, since they lie on the nondecreasing continuous curve through $G_{x_n}$. For each $t \in Disc(x, \epsilon/2)$, let $A_n$ contain these specified points. This construction yields $\hat{d}(A_n, G_{x_n}) < \epsilon/2$. For $t \notin Disc(x, \epsilon/2)$, let $A$ contain the points already constructed in $A'$. It remains to construct the points in $A$ for $t \in Disc(x, \epsilon/2)$. For this purpose, we use the points in $A_n$ associated with $t$. Again, to simplify the discussion, suppose that $x^i(t-) \le x^i(t)$ for all $i$. With this ordering, we let

$$z^i_{r+k} = x^i(t-) \vee \max_{1 \le l \le k} z^i_{n,r+l} \wedge x^i(t)$$

for each $k$ and $i$. This definition guarantees that the points $(z_{r+k}, t)$ belong to $G_x$ and are ordered. Moreover, $\hat{d}(A, G_x) < \epsilon$. Finally, we must have $d^*(A, A_n) < \epsilon$, because otherwise the condition $w_w(x_n, \eta) < \epsilon/32$ would be violated.

(v)$\rightarrow$(i). Suppose that the conditions in (v) hold and let $\epsilon > 0$ be given. Construct the finite subsets $A$ and $A_n$ with the specified properties. Let $(u, r)$ and $(u_n, r_n)$ be arbitrary parametric representations of $G_x$ and $G_{x_n}$ such that there are points $s_i$ in $S \subseteq [0, 1]$ such that both $(u(s_i), r(s_i)) = (z_i, t_i) \in A$ and $(u_n(s_i), r_n(s_i)) = (z_{n,i}, t_{n,i}) \in A_n$. Since $A$ and $A_n$ are ordered subsets of $G_x$ and $G_{x_n}$, respectively that construction is possible. Finally, for any $s$, $0 < s < 1$, there is $s_i \in S$ such that $s_i \le s < s_{i+1}$ and

$$\begin{aligned}
&\|u_n(s) - u(s)\| \vee \|r_n(s) - r(s)\| \le \|(u_n(s), r_n(s)) - (u_n(s_i), r_n(s_i))\| \\
&\quad + \|(u_n(s_i), r_n(s_i)) - u(s_i), r(s_i)\| + \|(u(s_i), r(s_i)) - u(s), r(s)\| \\
&\le \hat{d}(A_n, G_{x_n}) + d^*(A, A_n) + \hat{d}(A, G_x) \le 3\epsilon \ . \quad \blacksquare
\end{aligned}$$

## 6.6. Strengthening the Mode of Convergence

Section 12.6 of the book applies the characterizations of $M_1$ convergence in previouis sections to establish conditions under which the mode of convergence can be strengthened: We find conditions under which $WM_1$ convergence can be replaced by $SM_1$ convergence. Most of the material appears in the book.

We use the following Lemma.

**Lemma 6.6.1.** (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* $\delta > 0$,

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

**Proof.** For $(t-\delta) \vee 0 \le t_1 < t_2 < t_3 \le (t+\delta) \wedge T$,

$$\begin{aligned} \|(x_n, y_n)(t_2) &- [(x_n, y_n)(t_1), (x_n, y_n)(t_3)]\| \\ &\le \|(x_n, y_n)(t_2) - [(x_n(t_1), y(t)), (x_n(t_3), y(t))]\| \\ &\quad + (\|y_n(t_1) - y(t)\| \vee \|y_n(t_3) - y(t)\|) \\ &\le \|x_n(t_2) - [x_n(t_1), x_n(t_3)]\| \vee \|y_n(t_2) - y(t)\| \\ &\quad + (\|y_n(t_1) - y(t)\| \vee \|y_n(t_3) - y(t)\|) \\ &\le \|x_n(t_2) - [x_n(t_1), x_n(t_3)]\| + 2v(y_n, y, t, \delta)\,. \quad \blacksquare \end{aligned}$$

**Theorem 6.6.1.** (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*

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

The proof is in the book.

## 6.7. Characterizing Convergence with Mappings

In this section we focus on alternative characterizations of $SM_1$ convergence using mappings.

### 6.7.1. Linear Functions of the Coordinates

The strong topology $SM_1$ differs from the weak topology $WM_1$ by the behavior of linear functions of the coordinates. Example **??** 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 $\eta x$ is real valued.

**Theorem 6.7.1.** (Lipschitz property of linear functions of the coordinate functions) *For any $x_1$, $x_2 \in D([0,T], \mathbb{R}^k)$ and $\eta \in \mathbb{R}^k$,*

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

**Proof.** Pick an arbitrary $\epsilon > 0$ and choose $(u_j, r_j) \in \Pi_s(x_j)$ for $j = 1, 2$ such that

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

which is possible by the definition (3.7). Because $\eta u_j \in \Pi(\eta x_j)$ for $j = 1, 2$, by Lemma 6.3.4,

$$\begin{aligned} d(\eta x_1, \eta x_2) &\leq \|\eta u_1 - \eta u_2\| \vee \|r_1 - r_2\| \\ &\leq \|r_1 - r_2\| \vee \|u_1 - u_2\| \|\eta\| \\ &\leq (\|\eta\| \vee 1)(d_s(x_1, x_2) + \epsilon) \ . \end{aligned}$$

Since $\epsilon$ was arbitrary, (7.1) is established. ∎

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 6.7.1.** ($SM_1$-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 6.6.1 to get $d_s((x_n, y_n), (x, y)) \to 0$ in $D([0,T], \mathbb{R}^{2k})$. Then apply Theorem 6.7.1. ∎

**Remark 6.7.1.** *Measurability of addition.* The measurability of addition on $(D, d_s) \times (D, d_s)$ holds because the Borel $\sigma$-field coincides with the Kolmogorov $\sigma$-field. It also follows from part of the proof of Theorem 4.1 of Whitt (1980). ∎

In Theorem 6.7.1 we showed that linear functions of the coordinates are Lipschitz in the $SM_1$ metric. We now apply Theorem 6.5.1 to show that convergence in the $SM_1$ topology is characterized by convergence of all such linear functions of the coordinates.

**Theorem 6.7.2.** (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 $\eta \in \mathbb{R}^k$.*

**Proof.** One direction is covered by Theorem 6.7.1. Suppose that $x_n \not\to x$ as $n \to \infty$ in $SM_1$. Then apply part (v) of Theorem 6.5.1 to deduce that $\eta x_n \not\to \eta x$ as $\eta \to \infty$ for some $\eta$. Note that $\|a\| > 0$ for $a \in \mathbb{R}^k$ if and only if $|\eta a| > 0$ in $\mathbb{R}$ for some $\eta \in \mathbb{R}^k$. Also, $\|a - A\| > 0$ for $A \subseteq R^k$ if and only if $|\eta a - \eta A| > 0$ in $\mathbb{R}$ for some $\eta \in \mathbb{R}^k$, where $\eta A = \{\eta b : b \in A\}$. ∎

We can get convergence of sums under more general conditions than in Corollary 6.7.1. 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 6.7.3.** (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.** The proof is in the book.

### 6.7.2. Visits to Strips

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 6.7.2 to develop associated characterizations for $\mathbb{R}^k$-valued functions. For each $x \in D([0,T],\mathbb{R}^1)$, $0 \le t_1 < t_2 \le T$ 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 $t_1 < t_1' < \cdots < t_k' \le t_2$ 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

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

or

$$\text{(ii)} \quad \sup\{x(s) : t_1 \le s \le t_2\} \le x(t-) \quad (\inf\{x(s) : t_1 \le s \le t_2)\} \ge x(t-)) \ .$$

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

**Proof.** For each $n$, let $\{t_{n,i}\}$ be a finite collection of points in $[0,T]$, including 0 and $T$. Let $\{t_{n,i}\}$ be a subcollection of $\{t_{n+1},i\}$ for each $n$ and let the minimum distance between points in $\{t_{n,i}\}$ be $\epsilon_n$, where $\epsilon_n \downarrow 0$ as $n \to \infty$. Note that there is one local maximum value and one local minimum value of $x$ relative to the interval endpoints in each interval $[t_{n,i}, t_{n,i+1})$, where $t_{n,i}$ and $t_{n,i+1}$ are successive points in $\{t_{n,i}\}$. Hence the total number of extreme values of $x$ relative to $\{t_{n,i}\}$ is countably infinite. Next note that any extreme value of $x$ is contained in this set. To see this, suppose that $b$ is an extreme value of $x$ at $t$ relative to the interval $(t_1,t_2)$. Then, for sufficiently large $n$, there is an interval $(t_{n,i}, t_{n,i+1})$ such that $t_1 \le t_{n,i} < t < t_{n,i+1} \le t_2$, so that $b$ is an extreme value of $x$ within $(t_{n,i}, t_{n,i+1})$. ∎

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 6.7.2.** *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 6.7.4.** (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.*

**Proof.** By Theorem 6.7.2, it suffices to establish the result for $\mathbb{R}$-valued functions. First, suppose that $x_n \to x$ as $n \to \infty$ in $D([0, T], \mathbb{R}, M_1)$. Suppose that $a$ and $b$ are not local extreme values of $x$. Let $t_1, t_2 \in Disc(x)^c$ and suppose that $x(t_1), x(t_2) \notin \{a, b\}$. Then, for sufficiently large $n$, by Lemma 6.7.2, $v^{a,b}_{t_1,t_2}(x_n) = v^{a,b}_{t_1,t_2}(x)$. Since there are at most countably many "bad" $a, b$ for any $x$, $v^{a,b}_{t_1,t_2}(x_n) \to v^{a,b}_{t_1,t_2}(x)$ for almost all $a, b$ with respect to Lebesgue measure. On the other hand, suppose that $v^{a,b}_{t_1,t_2}(x_n) \to v^{a,b}_{t_1,t_2}(x)$ for all $t_1, t_2 \in Disc(x)^c$ and for almost all $a, b$. We will show that characterization (v) of $SM_1$ convergence in Theorem 6.5.1 holds. For $x, t$ and $\epsilon > 0$ given, find $\eta$ such that $v(x, [t - \eta, t)) < \epsilon/2$ and $v(x, [t, t + \eta]) < \epsilon/2$. First suppose that $t \in Disc(x)^c$. Then $v^{a,b}_{t_1,t_2}(x) = 0$ for $t_1$, $t_2 \in Disc(x)^c$, $t - \eta < t_1 < t < t_2 < t + \eta$ and all $(a, b)$ with $a < x(t) - \epsilon/2 < x(t) + \epsilon/2 < b$. By assumption, for all suitably large $n$, $v^{a'b'}_{t_1,t_2}(x_n) = 0$ for some $a', b'$ with

$$x(t) - \epsilon < a' < x(t) - \epsilon/2 < x(t) + \epsilon/2 < b' < x(t) + \epsilon \ .$$

By the argument above, we can show that, for a time interval before $t$, $x_n$ and $x$ are first in a neighborhood of $x(t-)$ and then leave. Afterwards, $x_n$ and $x$ enter the neighborhood of $x(t)$ and stay there for a short interval after $t$. To see this, let $t_1$ and $t_2$ be as above and then find $a_1$, $b_1$, $a_2$, $b_2$ such that

$$x(t-) - \epsilon < a_1 < x(t-) - \epsilon/2, \;\; x(t-) + \epsilon/2 < b_1 < x(t) + \epsilon$$
$$x(t) - \epsilon < a_2 < x(t) - \epsilon/2, \;\; x(t) + \epsilon/2 < b_2 < x(t) + \epsilon \ ,$$

$v_{t_1,t_2}^{a_1,b_1}(x_n) \to v_{t_1,t_2}^{a_1,b_1}(x) = 1$ and $v_{t_1,t_2}^{a_2,b_2}(x_n) \to v_{t_1,t_2}^{a_2,b_2}(x) = 1$. that implies that $v(x_n, x, t, \delta) < \epsilon$ for $\delta < \min\{|t-t_1|, |t-t_2|\}$. Next suppose that $t \in Disc(x)$. Let $t_1$, $t_2$ be as above. Find $a_1, b_1, a_2, b_2$ such that

$$x(t-) - \epsilon < a_1 < x(t-) - \epsilon/2 < x(t-) + \epsilon/2 < b_1 < x(t-) + \epsilon\,,$$

$$x(t) - \epsilon < a_2 < x(t) - \epsilon/2 < x(t) + \epsilon/2 < b < x(t) + \epsilon\,,$$

$v_{t_1,t_2}^{a_1,b_1}(x_n) \to v_{t_1,t_2}^{a_1,b_1}(x) = 1$ and $v_{t_1,t_2}^{a_2,b_2}(x_n) \to v_{t_1,t_2}^{a_2,b_2}(x) = 1$. It remains to show that $x_n$ cannot fluctuate significantly between $x(t-)$ and $x(t)$. To be definite, suppose that $x(t-) < x(t)$ and suppose that $\epsilon < x(t) - x(t-)$. Then for almost all $a, b$ with

$$x(t-) + \epsilon/2 < a < b < x(t) - \epsilon/2\ ,$$

$$v_{t_1,t_2}^{a,b}(x_n) \to v_{t_1,t_2}^{a,b}(x) = 2 \quad \text{as} \quad n \to \infty\ .$$

That implies that $w_s(x_n, x, t, \delta) \to 0$ as $n \to \infty$ for $\delta < \min\{|t_1 - t|, |t - t_2|\}$, which completes the proof. ■

## 6.8. 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 (3.13) 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 $G_\delta$ subset of a complete metric space, as noted in Section 11.2 of the book.

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 $\delta$. 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 $\delta$, $\hat{w}_s(x, z)$ 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 $\lambda$ is known to be a complete metric inducing the topology of pointwise convergence at all continuity points of the limit; i.e.,

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

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

Thus, our new metric is

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

**Theorem 6.8.1.** (a complete $SM_1$ metric) *The metric $\hat{d}_s$ on $D$ in* (8.3) *is complete and topologically equivalent to $d_s$.*

**Proof.** To show topological equivalence of $\hat{d}_s$ and $d_s$, it suffices to show that $\lambda(\hat{w}_s(x_n, \cdot), \hat{w}_s(x, \cdot)) \to 0$ as $n \to \infty$ whenever $d_s(x_n, x) \to 0$ as $n \to \infty$. However, if $d_s(x_n, x) \to 0$ as $n \to \infty$, then $w_s(x_n, \delta) \to w_s(x, \delta)$ as $n \to \infty$ at all $\delta$ which are continuity points of $w_s(x, \delta)$. (See Lemma 6.8.1 below.) That in turn implies that $\hat{w}_s(x_n, z) \to \hat{w}_s(x, z)$ as $n \to \infty$ for all $z$ which are continuity points of $\hat{w}_s(x, z)$. However, such convergence is equivalent to convergence under $\lambda$. Next, suppose that a sequence $\{x_n\}$ is fundamental under $\hat{d}_s$, i.e., $\hat{d}_s(x_m, x_n) \to 0$ as $m, n \to \infty$. It follows that $\{x_n(t) : 0 \le t \le T, n \ge 1\}$ is compact. Hence, there exists a countable dense set $N$ of $[0, T]$, including 0 and $T$, and a subsequence $\{x_{n_k}\}$ such that $x_{n_k}(t) \to x(t)$ as $n_k \to \infty$ for all $t \in N$, where $x$ is some $\mathbb{R}^k$-valued function on $[0, T]$. At the same time, since $\lambda$ is known to be a complete metric, there must exist a distribution function $F$ such that

$$\lim_{n\to\infty} \lambda(\hat{w}_s(x_n, \cdot), F) = 0 \ ,$$

which implies that

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

However, Theorem **??** and Corollary **??** imply that there exists $\bar{x} \in D$ (with $\bar{x}$ not necessarily $x$) such that $d_s(x_{n_k}, \bar{x}) \to 0$ as $n_k \to \infty$. Since $d_s(x_n, \bar{x}) \le d_s(x_n, x_{n_k}) + d_s(x_{n_k}, \bar{x})$ and $d_s(x_m, x_n) \to 0$ as $m, n \to \infty$, $d_s(x_n, \bar{x}) \to 0$ as $n \to \infty$. ∎

To complete the proof of Theorem 6.8.1, we need the following lemma.

**Lemma 6.8.1.** (continuity of $SM_1$ modulus) *If $d_s(x_n, x) \to 0$ as $n \to \infty$, then $w_s(x_n, \delta) \to w_s(x, \delta)$ as $n \to \infty$ for each $\delta$ that is a continuity point of $w_s(x, \delta)$.*

**Proof.** Let $\delta$ be a continuity point of $w_s(x,\delta)$. Then, for each $\epsilon_1 > 0$, there is $\epsilon_2 > 0$ such that $w_s(x,\delta-\epsilon_2) \geq w_s(x,\delta)-\epsilon_1$. For $\delta$, $\epsilon_1$ and $\epsilon_2$ given, it is possible to choose continuity points $t$, $t_1$, $t_2$ and $t_3$ of $x$ such that

$$(t-\delta)\vee 0 \leq t_1 \leq t_2 \leq t_3 \leq (t+\delta)\wedge T \tag{8.4}$$

and

$$\|x(t_2)-[x(t_1),x(t_3)]\| \geq w_s(x,\delta-\epsilon_2)-\epsilon_1 \geq w_s(x,\delta)-2\epsilon_1 \ .$$

Since $d_s(x_n,x) \to 0$ as $n\to\infty$, $x_n(t_j) \to x(t_j)$ as $n \to \infty$ for $j = 1,2,3$. Hence, there exists $n_0$ such that, for all $n \geq n_0$,

$$\|x_n(t_2)-[x_n(t_1),x_n(t_3)]\| \geq w_s(x,\delta)-3\epsilon_2 \ .$$

However,

$$w_s(x_n,\delta) \geq \|x_n(t_2)-[x_n(t_1),x_n(t_3)]\| \ ,$$

so that $w_s(x_n,\delta) \geq w_s(x,\delta)-3\epsilon_2$. Since $\epsilon_2$ can be made arbitrarily small,

$$\varliminf_{n\to\infty} \ w_s(x_n,\delta) \geq w_s(x,\delta) \ . \tag{8.5}$$

We now establish an inequality in the other direction. Since $\delta$ is a continuity point of $w_s(x,\delta)$, for any $\epsilon_1 > 0$ there exists $\epsilon_2 > 0$ so that $w_s(x,\delta+\epsilon_2) \leq w_s(x,\delta)+\epsilon_1$. We can choose $t_n$, $t_{n1}$, $t_{n2}$ and $t_{n3}$ so that

$$(t_n-\delta)\vee 0 \leq t_{n1} \leq t_{n2} \leq t_{n3} \leq (t_n+\delta)\wedge T$$

and

$$\|x_n(t_{n2})-[x_n(t_{n1}),x_n(t_{n3})]\| \geq w_s(x_n,\delta)-\epsilon_2$$

for all $n$. There thus exists a subsequence $\{n_k\}$ such that $t_{n_k} \to t$ and $t_{n_kj} \to t_j$, $j = 1,2,3$, (8.4) holds and $\|x_{n_k}(t_{n_kj})-z_j\| \to 0$ as $n_k \to \infty$. Moreover, since $x$ and $x_n$, $n \geq 1$, are right-continuous for all $n$, we can have $t_1$, $t_2$ and $t_3$ be continuity points of $x$ with

$$(t-(\delta+\epsilon_2))\vee 0 \leq t_1 \leq t_2 \leq t_3 \leq (t+(\delta+\epsilon_2))\wedge T \ .$$

Then $\|x_{n_k}(t_{n_kj})-x(t_j)\| \to 0$ as $n_k \to \infty$. Hence, there is $n_0$ such that, for all $n_k \geq n_0$,

$$\begin{aligned} \|x(t_2)-[x(t_1),x(t)_3)]\| &\geq \|x_{n_k}(t_{n_k2})-[x_{n_k}(t_{n_k1}),x_{n_k}(t_{n_k3})]\|-\epsilon_2 \\ &\geq w_s(x_n,\delta)-2\epsilon_2 \ . \end{aligned} \tag{8.6}$$

However,

$$w_s(x,\delta) + \epsilon_1 \geq w_s(x,\delta+\epsilon_2) \geq \|x(t_2) - [x(t_1), x(t_3)]\| \ . \tag{8.7}$$

Combining (8.6) and (8.7), we obtain

$$w_s(x,\delta) \geq w_s(x_n,\delta) - \epsilon_1 - 2\epsilon_2 \ .$$

Since $\epsilon_1$ and $\epsilon_2$ can be made arbitrarily small,

$$\overline{\lim_{n\to\infty}} \ w_s(x_n,\delta) \leq w_s(x,\delta) \ . \tag{8.8}$$

Combining (8.5) and (8.8) completes the proof. ■

## 6.9. 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 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 6.9.1.** (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 6.9.2.** *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 6.9.1.** (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 6.9.1.** (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.*

**Proof.** Suppose that $x_n \to x$ as $n \to \infty$ in the specified topology. That means that $r_{t_m}(x_n) \to r_{t_m}(x)$ as $n \to \infty$ for some sequence $\{t_m\}$ with $t_m \to \infty$, possibly depending on $x$. Since $f$ has continuous restrictions,

$$r_{t_m}(f(x_n)) = f_{t_m}(r_{t_m}(x_n)) \to f_{t_m}(r_{t_m}(x)) = r_{t_m}(f(x))$$

as $n \to \infty$ for all $m$, which implies that $f(x_n) \to f(x)$ as $n \to \infty$ in the specified topology. ∎

No more material has been deleted from Section 12.9 of the book.

## 6.10. Strong and Weak $M_2$ Topologies

We now define strong and weak versions of Skorohod's $M_2$ topology. In Section 6.11 we will show that it is possible to define the $M_2$ topologies by a minor modification of the definitions in Section 6.3, 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 $\Gamma_x$. 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 equation 5.2 of Section 11.5 in the book. As in equation (5.4) in Section 11.5 of the book, 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 **??**, which showed that $WM_1$ is strictly weaker than $SM_1$ also shows that $WM_2$ is strictly weaker than $SM_2$. Example **??** shows that the $SM_2$ topology is strictly weaker than the $SM_1$ topology.

Note that $\mu_s$ in (10.1) is *not* symmetric in its two arguments. Example 12.10.1 of the book shows 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 $n \to \infty$.

## 6.10.1. The Hausdorff Metric Induces the $SM_2$ Topology

We now show that $m_s$ induces the $SM_2$ topology.

**Theorem 6.10.1.** (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 $n \to \infty$.*

**Proof.** Our proof will exploit lemmas below. Suppose that $\mu_s(x_n, x) \to 0$ but $\mu_s(x, x_n) \not\to 0$ as $n \to \infty$. Since $\mu_s(x, x_n) \not\to 0$, there exists $(z, t) \in \Gamma_x$ for which it is not possible to find $(z_n, t_n) \in \Gamma_{x_n}$ for $n \ge 1$ such that $(z_n, t_n) \to (z, t)$ as $n \to \infty$, but that contradicts Lemma 6.10.4 below. ∎

In order to complete the proof of Theorem 6.10.1, we prove the following four lemmas.

**Lemma 6.10.1.** *Suppose that $\mu_s(x_n, x) \to 0$ as $n \to \infty$. If $(z_n, t_n) \in \Gamma_{x_n}$ for $n \ge 1$, then there exists a subsequence $\{(z_{n_k}, t_{n_k})\}$ with $(z_{n_k}, t_{n_k}) \to (z, t)$ as $n_k \to \infty$ for some $(z, t) \in \Gamma_x$. Moreover, the limits of all convergent subsequences must be in $\Gamma_x$.*

**Proof.** Suppose that $\mu_s(x_n, x) \to 0$ as $n \to \infty$ and consider any sequence $\{(z_n, t_n)\}$ with $(z_n, t_n) \in \Gamma_{x_n}$ for $n \geq 1$. By the definition of $\mu_s$, there must exist $(z'_n, t'_n) \in \Gamma_x$ such that $\|(z_n, t_n) - (z'_n, t'_n)\| \to 0$ as $n \to \infty$. Since $\Gamma_x$ is compact, there exists a convergent subsequence of the sequence $\{(z'_n, t'_n)\}$; i.e., there exists $\{(z'_{n_k}, t'_{n_k})\}$ such that $(z'_{n_k}, t'_{n_k}) \to (z, t)$ for some $(z, t) \in \Gamma_x$. By the triangle inequality, we must also have $(z_{n_k}, t_{n_k}) \to (z, t)$ as $n_k \to \infty$. Finally, suppose $(z_{n_k}, t_{n_k})$ is an arbitrary convergent subsequence of $\{(z_n, t_n)\}$. By the argument above, there exists $(z, t) \in \Gamma_x$ such that a subsequence $(z_{n_{k_j}}, t_{n_{k_j}}) \to (z, t)$ as $n_{k_j} \to \infty$. This implies that $(z, t)$ must be the limit of the convergent subsequence $\{(z_{n_k}, t_{n_k})\}$. ■

**Lemma 6.10.2.** *Suppose that $\mu_s(x_n, x) \to 0$ as $n \to \infty$, $t \notin Disc(x)$ and $(z_n, t) \in \Gamma_{x_n}$ for $n \geq 1$. Then $z_n \to x(t)$ as $n \to \infty$.*

**Proof.** By Lemma 6.10.1, there is a subsequence $(z_{n_k}, t) \to (z, t) \in \Gamma_x$, but $z = x(t)$ for $(z, t) \in \Gamma_x$ because $t \notin Disc(x)$. Since all convergent subsequences must have the same limit, $z_n \to z = x(t)$ as $n \to \infty$. ■

**Corollary 6.10.1.** *If $t \notin Disc(x)$ and $\mu_s(x_n, x) \to 0$ as $n \to \infty$, then $x_n(t) \to x(t)$ and $x_n(t-) \to x(t)$ in $\mathbb{R}^k$ as $n \to \infty$.*

**Lemma 6.10.3.** *If $\mu_s(x_n, x) \to 0$ as $n \to \infty$ and $(z, t) \in \Gamma_x$, then for any $i$, $1 \leq i \leq k$, there exist $(z_n, t_n) \in \Gamma_{x_n}$ for $n \geq 1$ such that $|z_n^i - z^i| \vee |t_n - t| \to 0$.*

**Proof.** The conclusion follows from Corollary 6.10.1 if $t \notin Disc(x)$, so suppose that $t \in Disc(x)$. Then $z$ belongs to the segment $[x(t-), x(t)]$. First choose $t'_m > t$ with $t'_m \notin Disc(x)$ for all $m$ and $t'_m \downarrow t$ as $m \to \infty$. By Lemma 6.10.2, there exist $(z'_{m,n}, t'_m) \in \Gamma_{x_n}$ such that $z'_{m,n} \to x(t'_m)$ as $n \to \infty$. Next choose $t''_m < t$ with $t''_m \notin Disc(x)$ for all $m$ and $t''_m \uparrow t$ as $m \to \infty$. By Lemma 6.10.2 again, there exist $(z''_{m,n}, t''_m) \in \Gamma_{x_n}$ such that $z''_{m,n} \to x(t''_m)$ as $n \to \infty$. The diagonal sequences $(z'_{n,n}, t'_n)$ and $(z''_{n,n}, t''_n)$ thus belong to $\Gamma_{x_n}$ and satisfy $t'_n \downarrow t$, $t''_n \uparrow t$, $z'_{n,n} \to x(t)$ and $z''_{n,n} \to x(t-)$ as $n \to \infty$. Since $\Gamma_{x_n^i}$ is a continuous real-valued curve, every value in the segment $[z'^i_{n,n}, z''^i_{n,n}]$ is realized for some $t'''_n$ with $t''_n \leq t'''_n \leq t'_n$. Hence, for any $(z, t) \in \Gamma_x$, there exists $(z'''_n, t'''_n) \in \Gamma_{x_n}$ such that $(z'''^i_n, t'''_n) \to (z^i, t)$ as $n \to \infty$. ■

**Lemma 6.10.4.** *If $\mu_s(x_n, x) \to 0$ as $n \to \infty$ and $(z, t) \in \Gamma_x$, then there exist $(z_n, t_n) \in \Gamma_{x_n}$ for $n \geq 1$ such that $\|(z_n, t_n) - (z, t)\| \to 0$ as $n \to \infty$.*

**Proof.** If $t \notin Disc(x)$, then we can take $(x_n(t), t) \in \Gamma_{x_n}$ or $(x_n(t-), t) \in \Gamma_{x_n}$ by Corollary 6.10.1. Hence it suffices to assume that $t \in Disc(x)$. Then, by the first part of the proof of Lemma 6.10.3, it suffices to consider $(z, t)$ with $z \neq x(t)$ and $z \neq x(t-)$. For at least one coordinate $i$, either $x^i(t-) < z < x^i(t)$ or $x^i(t) > z > x^i(t)$. Consider one such coordinate. By Lemma 6.10.3, there is $(z_n, t_n) \in \Gamma_{x_n}$ such that $t_n \to t$ and $z_n^i \to z^i$ as $n \to \infty$. Moreover, since $\mu_s(x_n, x) \to 0$, given $(z_n, t_n) \in \Gamma_{x_n}$, we must have $(z_n', t_n') \in \Gamma_x$ such that $\|z_n - z_n'\| \vee |t_n - t_n'| \to 0$. Since $t_n \to t$, we must also have $t_n' \to t$. Since $z_n^i \to z^i$ and $\Gamma_x$ contains the line joining $(x(t-), t)$ and $(x(t), t)$, we must have $z_n' \to z$ as well, which implies that $z_n \to z$, establishing the desired conclusion. ∎

### 6.10.2. $WM_2$ is the Product Topology

We now observe that $m_p$ induces the $WM_2$ topology.

**Theorem 6.10.2.** ($WM_2$ is the product topology) *$\mu_w(x_n, x) \to 0$ as $n \to \infty$ for $\mu_w$ 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.*

**Proof.** First, if $\mu_w(x_n, x) \to 0$ as $n \to \infty$, then $\mu_w(x_n^i, x^i) \to 0$ for each $i$, but $\mu_w(x_n^i, x^i) = \mu_s(x_n^i, x^i)$, so that $\mu_s(x^i, x_n^i) \to 0$ and $m_p(x_n, x) \to 0$ by Theorem 6.10.1. Conversely, suppose that $m_p(x_n, x) \to 0$ as $n \to \infty$. Lemma 6.10.1 implies that $\cup_{n \geq 1} \Gamma_{x_n^i}$ is compact for each $i$, $1 \leq i \leq k$. That in turn implies that $\cup_{n \geq 1} G_{x_n}$ is compact. Hence, if $(z_n, t_n) \in G_{x_n}$ for $n \geq 1$, then every subsequence necessarily has a convergent subsubsequence. To have $\mu_w(x_n, x) \not\to 0$, we must have a subsequence of $\{(z_n, t_n)\}$ converge to a limit not in $G_x$. We will show that is not possible. Consider $(z_n, t_n) \in G_{x_n}$, $n \geq 1$. Since $t_n \in [0, T]$ for all $n$, there exists a subsequence $(z_{n_k}, t_{n_k})$ such that $t_{n_k} \to t$ for some $t$, $0 \leq t \leq T$. Since $m_p(x_n, x) \to 0$ as $n \to \infty$, there is a subsequence $\{(z_{n_{k_j}}, t_{n_k})\}$ such that $z_{n_{k_j}}^i \to z^i$ for some $z^i$ where $(z^i, t) \in \Gamma_{x^i}$. Moreover, there are such subsequences for all $i$, $1 \leq i \leq k$, so that $z_n^i \to z^i$ for all $i$ along the final subsequence. Moreover, $(z^i, t) \in \Gamma_{x^i}$ for all $i$, but this implies that $(z, t) \in G_x$. Hence every subsequence of $(z_n, t_n)$ has a convergent subsubsequence and every convergent subsequence of $\{(z_n, t_n)\}$ has limit $(z, t) \in G_x$. That implies that $\mu_w(x_n, x) \to 0$ as $n \to \infty$. ∎

## 6.11. Alternative Characterizations of $M_2$ Convergence

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

### 6.11.1. $M_2$ 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 6.3. 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 $\Gamma_x$ 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 $\Gamma_x$ 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.

### 6.11.2. $SM_2$ 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

$0 \vee (t - \delta) \le t_1 \le [0 \vee (t - \delta)] + \delta/2$ 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 $x_n(T) \to x(T)$.

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

**Theorem 6.11.1.** (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* $\mu_s$ *in* (10.1).

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

$$\lim_{\delta \downarrow 0} \overline{\lim_{n \to \infty}} \ \bar{w}_s(x_n, x, \delta) = 0 \ . \tag{11.8}$$

*(v) For each* $t$, $0 \le t \le T$,

$$\lim_{\delta \downarrow 0} \overline{\lim_{n \to \infty}} \ \bar{w}_s(x_n, x, t, \delta) = 0 \tag{11.9}$$

*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 $\Gamma_x$ and $A_n$ of $\Gamma_{x_n}$, as in* (3.9) *where $(z_1, t_1) \le (z_2, t_2)$ if $t_1 \le t_2$, 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*

$$\lim_{\delta \downarrow 0} \overline{\lim_{n\to\infty}} \; \bar{w}_s^*(x_n, \delta) = 0 \tag{11.10}$$

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

**Proof.** We already have shown the equivalence (ii)$\leftrightarrow$(iii) in Theorem 11.10.1.
(i)$\to$(ii). Suppose that (i) holds with $\epsilon$ and $n$ given. Since the parametric representations in $\Pi_{s,2}(x)$ map onto the graph $\Gamma_x$, for any $(z_n, t_n) \in \Gamma_{x_n}$, we can find $s \in [0, 1]$ such that $(u_n(s), r_n(s)) = (z_n, t_n)$. For that $s$, $(u(s), r(s)) = (z, t)$ for some $(z, t) \in \Gamma_x$ and

$$\|(z_n, t_n) - (z, t)\| \le \|u_n - u\| \vee \|r_n - r\| < \epsilon \; . \tag{11.11}$$

By the same reasoning, for any $(z, t) \in \Gamma_x$, there exists $(z_n, t_n) \in \Gamma_{x_n}$ such that (11.11) holds.
(ii)$\to$(v). For $x$, $t$ and $\epsilon$ given, find $\delta$ such that $v(x, [t - \delta, t)) < \epsilon/2$ and $v(x, [t, t + \delta)) < \epsilon/2$ for $v$ in (2.5). Then apply (ii) to find $n_0$ such that $m_s(x_n, x) < \eta \equiv (\epsilon \wedge \delta)/2$ for $n \ge n_0$. Then, for each $t'$ with $0 \vee (t - \eta) \le t' \le (t + \eta) \wedge T$, there must exist $(\bar{z}, \bar{t}) \in \Gamma_x$ such that

$$\|(x_n(t'), t') - (\bar{z}, \bar{t})\| < \eta \quad \text{for} \quad n \ge n_0 \; .$$

Since $|\bar{t} - t| \le |\bar{t} - t'| + |t' - t| < 2\eta < \delta$,

$$\|(\bar{z}, \bar{t}) - [x(t-), x(t)]\| < \epsilon/2 \; .$$

Consequently, for $n \ge n_0$,

$$\|x_n(t') - [x(t-), x(t)]\| < \eta + \epsilon/2 < \epsilon \; .$$

Since $t'$ was arbitrary,

$$w_s(x_n, x, t, \delta) < \epsilon \; .$$

(v)$\leftrightarrow$(iv). Characterization (iv) clearly implies (v), so that it suffices to show that (v) implies (iv). We will show that if (iv) fails, then so does (v). Hence suppose that (iv) does not hold. Then there must exist $\epsilon > 0$, such that for any $\delta > 0$ there is a subsequence $\{n_k\}$ such that $n_k \to \infty$ and $\bar{w}_s(x_{n_k}, x, \delta) > \epsilon$ for all $n_k$. Hence, there is an associated sequence $t_{n_k}$ such that

$$\bar{w}_s(x_{n_k}, x, t_{n_k}, \delta) > \epsilon/2$$

for all $n_k$. However, $\{t_{n_k}\}$ has a convergent subsequence $\{t_{n_{k_j}}\}$ with $t_{n_{k_j}} \to t$ as $n_{k_j} \to \infty$ for some $t$. Note that, if $z_n \in [x(t_n-), x(t_n)]$ for all $n$, where $t_n \to t$, and if $z_n \to z$, then necessarily $z \in [x(t-), x(t)]$. Hence,

$$\bar{w}_s(x_{n_{k_j}}, x, t, 2\delta) > \epsilon/2$$

for all sufficiently large $n_{k_j}$. That implies that (11.9) does not hold, so that (v) fails.
(iv)$\to$(vi). We construct the desired finite subsets $A$ of $\Gamma_x$ and $A_n$ of $\Gamma_{x_n}$ by considering two kinds of points in $\Gamma_x$. For $\epsilon > 0$ given, we let $A$ contain at least one point $(z, t)$ for each $t \in Disc(x, \epsilon/2)$. The other points have $t \in Disc(x)^c$. We first construct $A$ for $t$ outside a finite union of neighborhoods of points in $Disc(x, \epsilon/2)$. We then construct $A_n$ and finally we complete the definition of $A$ by adding appropriate points $(z, t)$ for $t \in Disc(x, \epsilon/2)$, which depend on $A_n$. Thus the set $A$ ultimately depends upon $A_n$ and thus upon $x_n$ and $n$.

Let $t(A)$ denote the set of $t$ for which there is at least one pair $(z, t)$ from $\Gamma_x$ in $A$. We first identify $t(A)$. We include $Disc(x, \epsilon/2)$ in $t(A)$. Use (11.8) to find an $\eta$ and an $n_0$ such that $\bar{w}_s(x_n, x, \eta) < \epsilon/4$ for all $n \geq n_0$. Let $t_1 < \cdots < t_m$ be the ordered set of points in $Disc(x, \epsilon/2) - \{T\}$; let $t_0 = 0$ and $t_{m+1} = T$. Use the existence of left and right limits for $x$ to identify points, for $1 \leq i \leq m$, points $t_i'$ and $t_i''$ in $Disc(x)^c$ such that $t_{i-1}'' < t_i' < t_i < t_i'' < t_{i+1}'$, $|t_i - t_i'| < \eta$, $|t_i - t_i''| < \eta$, $v(x, [t_i', t_i)) < \epsilon/4$ and $v(x, [t_i, t_i'']) < \epsilon/4$ for $v(x, B)$ in (2.5). We include these points $t_i'$ and $t_i''$ in $t(A)$. We also include in $A$ points $t_0''$ and $t_{m+1}'$ from $Disc(x)^c$ such that $t_0 = 0 < t_0'' < t_1'$, $t_m'' < t_{m+1}' < t_{m+1} = T$, $v(x, [0, t_0'']) < \epsilon/4$ and $v(x, [t_{m+1}', T]) < \epsilon/4$. We also include the points 0 and $T$ in $t(A)$. Moreover, we include the points $(x(t_i'), t_i')$, $(x(t_i''), t_i'')$, $(x(0), 0)$ and $(x(T), T)$ in $A$ itself. (Except possibly for $T$, these are the only possibilities since $t_i'$, $t_i''$, $0 \in Disc(x)^c$.) We next define $A$ for $t$ in the compact set

$$C \equiv [0, T] - \bigcup_{i=1}^{m} (t_i', t_i'') - [0, t_0'') - (t_{m+1}', T] \,. \tag{11.12}$$

The set $C$ is a finite union of the closed intervals $[t_i'', t_{i+1}']$, $0 \le i \le m-1$. For each $t$ in $C$ not a boundary point of one of these subintervals, it is possible to find $t'$ and $t''$ in the same subinterval as $t$ such that $t' < t < t''$, $|t-t'| < \eta/4$, $|t - t''| < \eta/4$ and $v(x, [t', t'']) < \epsilon/2$. (Recall that $C \subseteq Disc(x, \epsilon/2)^c$.) For the boundary points $t_i'$ and $t_i''$, include intervals $(\bar{t}_i, t_i']$ and $[t_i'', t_i^*)$ with the same properties; these intervals are open in the relative topology on $C$. Also include intervals $[0, t^*)$ and $(\bar{t}, T]$ with the same properties; these intervals again are open in the relative topology on $C$. These open intervals form an open cover of $C$. Since $C$ is compact, there exists a finite subcover. We let $t(A)$ contain one point $t$ in $Disc(x)^c$ from each subinterval in the finite subcover; we also put $(x(t), t)$ into $A$. Let the set $A$ be ordered according to the time points; i.e., $(z_1, t_1) \le (z_2, t_2)$ if $t_1 \le t_2$. So far, $A$ contains points $(x(t), t)$ for $t \in Disc(x)^c$, including the boundary points $t_i'$ and $t_i''$ of $C$. We have completed the definition of $t(A)$, which includes $Disc(x, \epsilon/2)$. If $\{t_i\}$ is the ordered set of points in $t(A)$, then the construction above implies that $|t_{i+1} - t_i| < \eta$ for all $i$ (where $\eta$ has been chosen so that $\bar{w}_s(x_n, x, \eta) < \epsilon/4$).

We now construct the set $A_n$. By Theorem 11.4.1, condition (11.8) implies that $x_n(t) \to x(t)$ for each $t \in Disc(x)^c$. For each $t \in t(A) - Disc(x, \epsilon/2)$, let $t \in t(A_n)$ and $(x_n(t), t) \in A_n$. Since each such $t$ belongs to $Disc(x)^c$, there is $n_1 \ge n_0$ such that $\|x_n(t) - x(t)\| < \epsilon/4$ for all $t \in t(A) - Disc(x, \epsilon/2)$ and for all $n \ge n_1$. Hence we have established $d^*(A, A_n) < \epsilon/4$ for $n \ge n_1$ over $C$ (outside the neighborhoods of $Disc(x, \epsilon/2)$). We complete the definition of $A_n$ by adding finitely many points $(z, t)$ for $t$ in the open interval $(t_i', t_i'')$ where $t_i'$ and $t_i''$ are the adjacent points in $t(A)$ to $t_i \in Disc(x, \epsilon/2)$. We also do this for the interval $(t_{m+1}', T]$ if $T \in Disc(x, \epsilon/2)$. We do this for all $t_i \in Disc(x, \epsilon/2)$ so that overall $\hat{d}(A_n, \Gamma_{x_n}) < \epsilon/2$. This is always possible by Lemma 6.3.1. We next complete the definition of $A$ by including a point $(z, t_i)$ for each point $(z, t)$ in $A_n$ with $t \in (t_i', t_i'')$. This ensures that $A_n$ and $A$ have the same cardinality. Since $d(A_n, \Gamma_{x_n}) \le \epsilon/2$, $\bar{w}_s(x_n, x, \eta) < \epsilon/4$,

$$\|x_n(t_i') - x(t_i')\| < \epsilon/4, \|x_n(t_i'') - x(t_i'')\| < \epsilon/4 \ ,$$

$$\|x(t_i') - x(t_i-)\| < \epsilon/4 \quad \text{and} \quad \|x(t_i'') - x(t_i)\| < \epsilon/4$$

for $n \ge n$, we can choose points in $A$ so that $d^*(A_n, A) \le \epsilon/2$ for $n \ge n_1$ and $\hat{d}(A, \Gamma_x) \le \epsilon$, which completes the proof.

(vi)$\to$(i). Suppose that $\epsilon$ is given and the sets $A$ and $A_n$ in (vi) have points $(z_i, t_i)$ and $(z_{n,i}, t_{n,i})$, $0 \le i \le m$, where $t_0 = 0$ and $t_m = T$. Construct arbitrary parametric representations of $(u, r)$ of $x$ and $(u_n, r_n)$ of $x_n$ such that

$$r(i/m) = t_i, \quad u(i/m) = z_i$$

and

$$r_n(i/m) = t_{n,i}, \quad u_n(i/m) = z_{n,i} \ .$$

Since $d^*(A_n, A) \le \epsilon$,

$$\max_{0 \le i \le m} \{|r(i/m) - r_n(i/m)| \vee \|u(i/m) - u_n(i/m)\|\} < \epsilon \ .$$

Since $\hat{d}(A, \Gamma_x) < \epsilon$ and $\hat{d}(A_n, \Gamma_{x_n}) < \epsilon$ too, by the triangle inequality,

$$\|r - r_n\| \vee \|u_n - u\| < 3\epsilon \ .$$

(iv)$\leftrightarrow$(vii). Suppose that $0 \le t \le T$. If $x$ is constant in the intervals $(0 \vee (t - 2\delta), t)$ and $[t, (t + 2\delta) \wedge T)$, then

$$[x(0 \vee (t' - \delta)), x((t' + \delta) \wedge T)] = [x(t-), x(t)]$$

for all $t'$ with $0 \vee (t - \delta) < t' < (t + \delta) \wedge T$. Consequently, in that situation

$$\begin{aligned} & \sup_{0 \vee (t-\delta) < t' < (t+\delta) \wedge T} \{\|x_n(t') - [x(0 \vee (t' - \delta)), x((t' + \delta) \wedge T)]\|\} \\ = & \sup_{0 \vee (t-\delta) < t' < (t+\delta) \wedge T} \{\|x_n(t') - [x(t-), x(t)]\|\} \ . \end{aligned} \tag{11.13}$$

Thus if $x$ is piecewise constant with the distance between successive discontinuities at least $\delta$, then $\bar{w}_s^*(x_n, x, \delta/2) = \bar{w}_s(x_n, x, \delta/2)$. Hence, for $\epsilon$ given suppose that we can choose $\eta$ to make $\bar{w}_s(x_n, x, \eta) < \epsilon/3$. Then approximate $x$ by $x_c \in D_c$ such that $\|x - x_c\| < \epsilon/3$. For that $x_c$, let $\alpha$ be the minimum distance between successive discontinuities. Then, for $\delta < \eta \wedge (\alpha/2)$,

$$\begin{aligned} \bar{w}_s^*(x_n, x, \delta) & \le \bar{w}_s^*(x_n, x_c, \delta) + \epsilon/3 \\ & \le \bar{w}_s(x_n, x_c, \delta) + \epsilon/3 \\ & \le \bar{w}_s(x_n, x, \eta) + 2\epsilon/3 \le \epsilon \ . \end{aligned} \tag{11.14}$$

Alternatively, for $\epsilon$ given, suppose that we can choose $\eta$ to make $\bar{w}_s^*(x_n, x, \eta) < \epsilon/3$. Following the same reasoning,

$$\begin{aligned} \bar{w}_s(x_n, x, \delta) & \le \bar{w}_s(x_n, x_c, \delta) + \epsilon/3 \\ & \le \bar{w}_s^*(x_n, x_c, \delta) + \epsilon/3 \\ & \le \bar{w}_s^*(x_n, x, \eta) + 2\epsilon/3 \le \epsilon \ . \end{aligned} \tag{11.15}$$

Hence (iv) is equivalent to (vii).
(ii)→(viii). By Theorem 11.4.1 and (ii)↔(v), (ii) implies that $x_n(t) \to x(t)$ for each $t \in Disc(x)^c$. It remains to show that (ii)→(11.10). For $\epsilon > 0$ given, first pick a piecewise-constant $x_c$ such that $\|x - x_c\| \le \epsilon/4$, which is possible by Lemma 6.3.1. Let $\gamma$ be the $\mathbb{R}^k$-valued function with $\gamma^i(t) = 1$, $0 \le t \le T$, $1 \le i \le k$. Then $x_c - (\epsilon/4)\gamma \le x \le x_\epsilon + (\epsilon/4)\gamma$, i.e.,

$$x_c(t) - \epsilon/4 \le x(t) \le x_c(t) + \epsilon/4 \quad \text{for} \quad 0 \le t \le T \ .$$

Let the $(\alpha, \beta)$-neighborhood of $x \in D$ be

$$N_{\alpha,\beta}(x) \equiv \{[[x(t) - \alpha\gamma, x(t) + \alpha\gamma]] \times [0 \vee (t - \beta), (t + \beta) \wedge T] : 0 \le t \le T\} \ . \tag{11.16}$$

Thus, $x \in N_{\epsilon/4,0}(x_c)$ and $x_c \in N_{\epsilon/4,0}(x)$. Now let $\alpha$ be the minimum distance between successive discontinuities in $x_c$, or to 0 or $T$ for the leftmost and rightmost discontinuity points. Given (ii), choose $n_0$ so that $m_s(x_n, x) = \eta_n < \eta < (\epsilon \wedge \alpha)/4$ for $n \ge n_0$. Then $x_n \in N_{\eta+\epsilon/4,\eta}(x_c)$. Suppose that $\{t_i : 1 \le i \le m-1\}$ is the set of discontinuities of $x_c$, with $t_0 = 0$ and $t_m = T$. By the construction above, the open intervals $(t_i - \eta, t_i + \eta)$ are disjoint, $1 \le i \le m-1$. Now let $\delta = 2\eta$. Hence, if $t' \in (t_i - \eta, t_i + \eta)$ for $t_i \in Disc(x_c)$, then

$$t_{i-1}+\eta < t'-\delta < t'-\delta/2 < t_i-\eta < t_i+\eta < t'+\delta/2 < t'+\delta < t_{i+1}-\eta \tag{11.17}$$

for all $i$, $1 \le i \le m-1$. On the other hand, if $t' \in [t_{i-1} + \eta, t_i - \eta] = B_{i,n}$, then necessarily either $(t' + \delta/2, t' + \delta)$ intersects $B_{i,n}$ or $(t' - \delta, t' - \delta/2)$ intersects $B_{i,n}$. Thus, for $n \ge n_0$ and each $t' \in [0, T]$, there exists $t_1 \in [0 \vee (t' - \delta), 0 \vee (t' - \delta) + \delta/2)$ and $t_2 \in (T \wedge (t' + \delta) - \delta/2, T \wedge (t' + \delta)]$ such that

$$\|x_n(t') - [x_n(t_1), x_n(t_2)]\| \le 2((\epsilon/4) + \eta) < \epsilon \ ; \tag{11.18}$$

i.e., $\bar{w}_s^*(x_n, \delta) < \epsilon$.
(viii)→(v). For $x, t$ and $\epsilon$ given, choose $\eta$ so that $0 < t - \eta < t < t + \eta < T$, $v(x, [t - \eta, t)) < \epsilon/4$ and $v(x, [t, t + \eta]) < \epsilon/4$. Now choose $\delta < \eta$ and $t' \in (t - \delta/2, t + \delta/2)$. For $\delta$ and $t'$ given, find $t_1$, $t_2$ in $Disc(x)^c$ such that $t_1 < t < t_2$, $t' - \delta < t_1 < t' - \delta/2$ and $t' + \delta/2 < t_2 < t' + \delta$. Then choose $n_0$ so that $\|x_n(t_i) - x(t_i)\| < \epsilon/4$ for $i - 1, 2$ and $n \ge n_0$. Apply (viii) to choose $n_1 \ge n_0$ so that $\bar{w}_s^*(x_n, \delta) \le \epsilon/2$. Then

$$\begin{aligned} \|x_n(t') - [x(t-), x(t)]\| &\le \|x_n(t') - [x(t_1), x(t_2)]\| + \epsilon/4 \\ &\le \|x_n(t') - [x_n(t_1), x_n(t_2)]\| + \epsilon/2 \\ &\le \bar{w}_s^*(x_n, \delta) + \epsilon/2 \le \epsilon \quad \text{for} \quad n \ge n_1 \end{aligned} \tag{11.19}$$

Since $t'$ is arbitrary in $(t - \delta/2, t + \delta/2)$,

$$\bar{w}_s(x_n, x, t, \delta/2) \le \epsilon \quad \text{for} \quad n \ge n_1 \ ,$$

which implies (v). ∎

**Remark 6.11.1.** The equivalence (iii)↔(vii)↔(viii) was established by Skorohod (1956). ∎

**Remark 6.11.2.** 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 6.4.1. ∎

### 6.11.3. $WM_2$ Convergence

Corresponding characterizations of $WM_2$ convergence follow from Theorem 6.11.1 because the $WM_2$ topology is the same as the product topology, by Theorem 6.10.2. Let

$$\bar{w}_w(x_1, x_2, \delta) \equiv \sup_{0 \le t \le T} \bar{w}_w(x_1, x_2, t, \delta) \tag{11.20}$$

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

**Theorem 6.11.2.** (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 (11.20),*

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

*(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 $\Gamma_{x_n}$, of common cardinality $m$ as in (3.9) with $(z_1, t_1) \le (z_2, t_2)$ if $t_1 \le t_2$, such that $\hat{d}(A, G_x) < \epsilon$, $\hat{d}(A_n, \Gamma_{x_n}) < \epsilon$ and $d^*(A, A_n) < \epsilon$ for all $n \ge n_0$, for $\hat{d}$ in (5.8) and $d^*$ in (5.6).*

**Proof.** (i)$\to$(ii). Clearly, $d_{w,2}(x_n, x) \to 0$ as $n \to \infty$ implies that $d_{s,2}(x_n^i, x^i) \to 0$ as $n \to \infty$ for each $i$. By Theorem 6.11.1, that implies $m_s(x_n^i, x^i) \to 0$ as $n \to \infty$ for each $i$, which implies (ii).
(ii)$\leftrightarrow$(iii). By Theorem 6.11.1, (ii) is equivalent to

$$\lim_{\delta \downarrow 0} \overline{\lim_{n \to \infty}} \ \bar{w}_s(x_n^i, x^i, \delta) = 0 \tag{11.21}$$

for each $i$, but that is equivalent to (iii) because

$$\max_{1 \le i \le k} \bar{w}_s(x_n^i, x^i, \delta) = \bar{w}_w(x_n, x, \delta) \ . \tag{11.22}$$

(iii)$\leftrightarrow$(iv). By Theorem 6.11.1, (iii) is equivalent to

$$\lim_{\delta \downarrow 0} \overline{\lim_{n \to \infty}} \ \bar{w}_s(x_n^i, x^i, t, \delta) = 0 \tag{11.23}$$

for each $i$, but that is equivalent to (iv) because

$$\max_{1 \le i \le k} \bar{w}_s(x_n^i, x^i, t, \delta) = \bar{w}_w(x_n, x, t, \delta) \ . \tag{11.24}$$

(iii)$\to$(v). Follow the proof of (iv)$\to$(vi) in Theorem 6.11.1. Use (**??**) to find an $\eta$ and an $n_0$ such that $\bar{w}_w(x_n, x, \eta) < \epsilon/4$ for all $n \ge n_0$. Define $t(A)$ as before, first by including $Disc(x, \epsilon/2)$ and then by adding points from $Disc(x)^c$ in the complement of the union of the intervals about the points in $Disc(x, \epsilon/2)$. Let $A$ be defined for $t \in t(A) - Disc(x, \epsilon/2)$ just as before. Let $A_n$ be defined just as before. We complete the definition of $A$ by including a point $(z_i, t_i)$ for each point $(z, t)$ in $A_n$ with $t \in (t_i', t_i'')$. This ensures that $A$ and $A_n$ have the same cardinality. Since $d(A_n, \Gamma_{x_n}) \le \epsilon/2$, $\bar{w}_w(x_n, x, \eta) < \epsilon/4$, $\|x_n(t_i') - x(t_i')\| < \epsilon/4$, $\|x_n(t_i'') - x(t_i'')\| < \epsilon/4$, $\|x(t_i') - x(t_i'-)\| < \epsilon/4$ and $\|x(t_i'') - x(t_i)\| < \epsilon/4$ for $n \ge n_1$, we can choose these points to add to $A$ so that $d^*(A_n, A) \le \epsilon/2$ for $n \ge n_1$ and $\hat{d}(A, G_x) \le \epsilon$. (Unlike in the proof of Theorem 6.11.1, here we cannot conclude that $\hat{d}(A, \Gamma_x) \le \epsilon$.)
(v)$\to$(i). Paralleling the proof of (v)$\to$(i) in Theorem 11.5.2, suppose that the conditions of (v) hold and $A$, $A_n$ and $\epsilon$ are given. Let $(u, r)$ and $(u_n, r_n)$

be parametric representations of $x$ and $x_n$ such that

$$\begin{aligned} u(i/m) &= z_i, \ r(i/m) = t_i \quad \text{for} \quad (z_i, t_i) \in A \\ u_n(i/m) &= z_{n,i}, \ r_n(i/m) = t_{n,i} \quad \text{for} \quad (z_{n,i}, t_{n,i}) \in A_n \ . \end{aligned}$$

For any $s \in [0,1]$ there is $i$ such that $s_i \le s \le s_{i+1}$ and

$$\begin{aligned} &\|u_n(s) - u(s)\| \vee \|r_n(s) - r(s)\| \le \|(u_n(s), r_n(s)) - u_n(s_i), r_n(s_i)\| \\ &+\|(u_n(s_i), r_n(s_i)) - (u(s_i), r(s_i))\| + \|(u(s_i), r(s_i)) - (u(s), r(s))\| \\ &\le \hat{d}(A_n, G_{x_n}) + d^*(A_n, A) + \hat{d}(A, G_x) \le 3\epsilon \ . \quad \blacksquare \end{aligned}$$

Theorem 6.11.2 and Section 6.4 show that all forms of $M$ convergence imply uniform convergence to continuous limit functions.

**Corollary 6.11.1.** (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 6.4.1 and 6.11.2. For (ii) add Lemma 6.4.2. $\blacksquare$

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

**Corollary 6.11.2.** (inheritance of jumps) *If $x_n \to x$ in $(D, WM_2)$, $t_n \to t$ in $[0,T]$ and $x_n^i(t_n) - x_n^i(t_n-) \ge c > 0$ for all $n$, then $x^i(t) - x^i(t-) \ge c$.*

**Proof.** Apply Theorem 6.11.2 (iv). $\blacksquare$

Let $J(x)$ be the maximum magnitude (absolute value) of the jumps of the function $x$ in $D$. We apply Corollary 8.5.1 to show that $J$ is upper semicontinuous.

**Corollary 6.11.3.** (upper semicontinuity of $J$) *If $x_n \to x$ in $(D, M_2)$, then*

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

**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 8.5.1 implies that $|x^i(t) - x^i(t-)| \geq c$. ■

### 6.11.4. Additional Properties of $M_2$

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 6.6.1, 6.7.1, 6.7.2 and 6.7.3.

**Theorem 6.11.3.** (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*

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

**Proof.** We use characterization (v) in Theorem 6.11.1. Using the discontinuity condition, it is easy to show that (11.9) holds for $[(x_n, y_n), (x, y)]$ when it holds separately for $[x_n, x]$ and $[y_n, y]$, because i.e., at most one of the segments $[(x(t-), x(t)]$ and $[y(t-), y(t)]$ contains more than a single point. ■

**Corollary 6.11.4.** (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 6.11.4.** (Lipschitz property of linear functions of the coordinate functions) *For any $x_1$, $x_2 \in D([0,T], \mathbb{R}^k)$ and $\eta \in \mathbb{R}^k$,*

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

**Proof.** For (**??**), the key property is that

$$\Gamma_{\eta x} = \{(\eta z, t) : (z, t) \in \Gamma_x\} .$$

It suffices to show that for all $\epsilon > 0$ and $(z_1', t_1) \in \Gamma_{\eta x_1}$ there exists $(z_2', t_2) \in \Gamma_{\eta x_2}$ such that

$$|z_1' - z_2'| \vee |t_1 - t_2| \leq (\|\eta\| \vee 1) m_s(x_1, x_2) + \epsilon \ .$$

However, for $(z_1', t_1) \in \Gamma_{\eta x_1}$, there exists $(z_1, t_1) \in \Gamma_{x_1}$ such that $\eta z_1 = z_1'$. Then choose $(z_2, t_2) \in \Gamma_{x_2}$ such that

$$\|z_1 - z_2\| \vee |t_1 - t_2| \leq m_s(x_1, x_2) + \epsilon$$

Let $(z_2', t_2) = (\eta z_2, t_2)$. Then

$$|z_1' - z_2'| = |\eta z_1 - \eta z_2| \leq \|\eta\| \|z_1 - z_2\| \ . \quad \blacksquare$$

We have an analog of Corollary 6.7.1 for the $M_2$ topology.

**Corollary 6.11.5.** ($SM_2$-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*

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

**Proof.** First apply Theorem 6.11.3 to get $m_s((x_n, y_n), (x, y)) \to 0$ in $D([0,T], \mathbb{R}^{k+l})$. Then apply Theorem 6.11.4. $\blacksquare$

**Theorem 6.11.5.** (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 $\eta \in \mathbb{R}^k$.*

**Proof.** One direction is covered by Theorem 6.11.4. Suppose that $x_n \not\to x$ as $n \to \infty$ in $SM_2$. Then apply part (v) of Theorem 6.11.1 to deduce that $\eta x_n \not\to \eta x$ as $\eta \to \infty$ for some $\eta$. Note that $\|a\| > 0$ for $a \in \mathbb{R}^k$ if and only if $|\eta a| > 0$ in $\mathbb{R}$ for some $\eta \in \mathbb{R}^k$. Also, $\|a - A\| > 0$ for $A \subseteq R^k$ if and only if $|\eta a - \eta A| > 0$ in $\mathbb{R}$ for some $\eta \in \mathbb{R}^k$, where $\eta A = \{\eta b : b \in A\}$. $\blacksquare$

Just as with the $M_1$ topology, we can get convergence of sums under more general conditions than in Corollary 6.11.5. 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 6.11.6.** (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*

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

**Proof.** Apply the characterization of $SM_2$ convergence in Theorem 6.11.1 (v). At points $t$ in $Disc(x)^c \cap Disc(y)^c$, use the local uniform convergence in Lemma 12.5.1 of the book and Corollary 6.11.1 here. For other $t$ not in $Disc(x) \cap Disc(y)$, use Theorem 6.11.3. For $t \in Disc(x) \cap Disc(y)$, exploit condition (7.2) to deduce that, for all $\epsilon > 0$, there exists $\delta$ and $n_0$ such that

$$\bar{w}_s(x_n + y_n, x + y, t, \delta) \le w_s(x_n, x, t, \delta) + w_s(y_n, y, t, \delta) + \epsilon \tag{11.25}$$

for all $n \ge n_0$. ■

We now apply Theorem 6.11.5 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 $0 \le t_1 < t_2 \le T$, let

$$M_{t_1,t_2}(x) \equiv \sup_{t_1 \le t \le t_2} x(t) \ . \tag{11.26}$$

The proof exploits the $SM_2$ analog of Corollary 6.9.1.

In preparation for the next result, we state a basic lemma about preservation of convergence under restriction maps. 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 $r_{t_1,t_2}(x)(s) = x(s)$, $t_1 \le s \le t_2$. We omit the proof.

**Lemma 6.11.1.** (continuity of restriction maps) *If $x_n \to x$ as $n \to \infty$ in $D([0,T], \mathbb{R}^k)$ with one of the $SM_1$, $WM_1$, $SM_2$ and $WM_2$ topologies and if $t_1, t_2 \in Disc(x)^c$, then*

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

*with the same topology.*

**Theorem 6.11.7.** (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*

$$M_{t_1,t_2}(\eta x_n) \to M_{t_1,t_2}(\eta x) \quad as \quad n \to \infty \tag{11.27}$$

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

**Proof.** By Theorem 6.11.5, it suffices to consider the case of real-valued functions. By considering $\eta = \pm 1$ in (11.27), we get both the minimum and the maximum over $[t_1, t_2]$. it is easy to see that (11.27) for $\eta = \pm 1$ implies characterization (v) in Theorem 6.11.1: For $x, t$ and $\epsilon$ given, choose $\gamma$ so that $v(x, [t-\gamma, t)) < \epsilon/2$, $v(x, [t, t+\gamma]) < \epsilon/2$ and $0 < t-\gamma < t+\gamma < T$. Then finding $n_0$ such that $|M_{t_1,t_2}(\eta x_n) - M_{t_1,t_2}(\eta x)| < \epsilon/2$ for $n \ge n_0$, $\eta = \pm 1$ and

$$t-\gamma < t_1 < t-\delta < t < t+\delta < t_2 < t+\gamma$$

implies that $\bar{w}_s(x_n, x, t, \delta) < \epsilon$ for $n \ge n_0$. On the other hand, if $x_n \to x$ in $D([0,T], \mathbb{R}^1, M_2)$, then the restrictions converge in $D([t_1,t_2], \mathbb{R}^1, M_2)$ for all $t_1, t_2 \in Disc(x)^c$ by Lemma 6.11.1. If $m_s(x_n, x) < \epsilon$ in $D([t_1,t_2], \mathbb{R}^1, M_2)$, then clearly $|M_{t_1,t_2}(x_n) - M_{t_1,t_2}(x)| < \epsilon$ and $|M_{t_1,t_2}(-x_n) - M_{t_1,t_2}(-x)| < \epsilon$, so characterization (ii) of Theorem 6.11.1 implies (11.27). ■

We can apply the characterization of $M_2$ convergence in Theorem 6.11.7 to show the preservation of convergence under bounding functions in the $M_2$ topology. See Corollary 12.11.6 in the book.

## 6.12. Compactness

We have nothing to add in this final section.