Lemmas for the Skorohod Space

Jason Swanson
May 17, 2006

The following lemmas were proved while working on a recent paper, but as of today, they do not appear in the final draft. Rather than having my hard work go to waste, I have decided to present them here in this small article. The context in which these lemmas occurred is described at the end of this article.

1 The Lemmas

A function is cadlag if it is right continuous and has left limits. If \((E, r)\) is a metric space, then the Skorohod space, \(D = D_E[0, \infty)\), is the space of cadlag functions from \([0, \infty)\) to \(E\). A metric on \(D\) is given by

\[
d(x, y) = \inf_{\lambda \in \Lambda} \left[ \| \log \lambda' \|_\infty \vee \int_0^\infty e^{-u} \sup_{t \geq 0} \{ r(x(t \wedge u), y(\lambda(t) \wedge u)) \wedge 1 \} \, du \right],
\]

(1)

where \(\Lambda\) is the collection of all strictly increasing, surjective, Lipschitz continuous functions \(\lambda : [0, \infty) \to [0, \infty)\) such that \(\| \log \lambda' \|_\infty < \infty\). If \((E, r)\) is complete and separable, then \((D, d)\) is complete and separable. This metric generates the Skorohod topology on \(D\). See Chapter 3 of [1] for details.

Note that \(D_E \times D_E\) is not the same space as \(D_{E \times E}\). In particular, the map \((x, y) \mapsto x + y\) is not continuous when viewed as a map from \(D_{\mathbb{R}^d} \times D_{\mathbb{R}^d}\) to \(D_{\mathbb{R}^d}\), but it is continuous as a map from \(D_{\mathbb{R}^{2d}}\) to \(D_{\mathbb{R}^d}\).

**Lemma 1.1** Suppose \(x_n \to x\) in \(D_{\mathbb{R}^d}[0, \infty)\) and \(y_n \to y\) in \(D_{\mathbb{R}^d}[0, \infty)\). If \(\Delta x(t) \Delta y(t) = 0\) for all \(t \geq 0\), then \(x_n + y_n \to x + y\) in \(D_{\mathbb{R}^d}[0, \infty)\).

**Proof.** By Lemma 6.2 in [2], \(v_n \to v\) in \(D_{\mathbb{R}^d}[0, \infty)\) if and only if the following conditions hold.

(i) If \(t_n \to t\), then \(|v_n(t_n) - v(t)| \wedge |v_n(t_n) - v(t^-)| \to 0\).

(ii) If \(s_n \geq t_n, s_n, t_n \to t\), and \(v_n(t_n) \to v(t)\), then \(v_n(s_n) \to v(t)\).

Let \(z_n = x_n + y_n\) and \(z = x + y\). Suppose \(t_n \to t\). Since \(\Delta x(t) \Delta y(t) = 0\), either \(t\) is a continuity point of \(x\) or it is a continuity point of \(y\). By symmetry, suppose it is a continuity point of \(y\).
point of x. In this case, choose strictly increasing, surjective \( \lambda_n : [0, \infty) \to [0, \infty) \) such that \( \lambda_n(t) \to t \) and \( x_n(t) - x(\lambda_n(t)) \to 0 \) uniformly on compacts. Then

\[
|x_n(t_n) - x(t)| \leq |x_n(t_n) - x(\lambda_n(t_n))| + |x(\lambda_n(t_n)) - x(t)|.
\]

Since \( \lambda_n(t_n) \to t \) and \( t \) is a continuity point of \( x \), it follows that \( x_n(t_n) \to x(t) \). Hence,

\[
|z_n(t_n) - z(t)| \land |z_n(t_n) - z(t^-)| \leq (|x_n(t_n) - x(t)| + |y_n(t_n) - y(t)|) \\
\land (|x_n(t_n) - x(t^-)| + |y_n(t_n) - y(t^-)|) \\
= |x_n(t_n) - x(t)| + (|y_n(t_n) - y(t)| \land |y_n(t_n) - y(t^-)|).
\]

Since (i) holds for \( \{y_n\} \), this goes to zero, which verifies (i) for \( \{z_n\} \).

Now suppose \( s_n \geq t_n, s_n, t_n \to t \), and \( z_n(t_n) \to z(t) \). Again, by symmetry, assume \( t \) is a continuity point of \( x \). We then have that \( y_n(t_n) = z_n(t_n) - x_n(t_n) \to z(t) - x(t) = y(t) \).

Hence, by (ii), we must have \( y_n(s_n) \to y(t) \). But this implies \( z_n(s_n) = x_n(s_n) + y_n(s_n) \to x(t) + y(t) = z(t) \) and this verifies (ii) for \( \{z_n\} \).

**Lemma 1.2** If \( 2 \leq d < \infty \), then \( \{X_n^1, \ldots, X_n^d\} \) is relatively compact in \( D_{\mathbb{R}^d}[0, \infty) \) if and only if \( \{X_n^k\} \) and \( \{X_n^k + X_n^k\} \) are relatively compact in \( D_{\mathbb{R}}[0, \infty) \).

**Proof.** Problem 3.22(c) in [1].

**Lemma 1.3** For each \( n \), let \( X_n \) and \( Y_n \) be independent random variables taking values in \( D_{\mathbb{R}^k}[0, \infty) \) and \( D_{\mathbb{R}^m}[0, \infty) \), respectively. Suppose that \( (X_n, Y_n) \Rightarrow (X, Y) \) in \( D_{\mathbb{R}^k}[0, \infty) \times D_{\mathbb{R}^m}[0, \infty) \). If

\[
P(\Delta X(t) \Delta Y(t) = 0 \text{ for all } t \geq 0) = 1,
\]

then \( (X_n, Y_n) \Rightarrow (X, Y) \) in \( D_{\mathbb{R}^k \times \mathbb{R}^m}[0, \infty) \).

**Proof.** By the Skorohod Representation Theorem, we can assume that \( X_n \to X \) and \( Y_n \to Y \) a.s. By Lemma 1.1, \( X_n + Y_n \to X + Y \) a.s. Hence, by Lemma 1.2, \( \{(X_n, Y_n)\} \) is relatively compact in \( D_{\mathbb{R}^k + \mathbb{R}^m}[0, \infty) \). If \((U, V)\) is a subsequential limit, then \( U \overset{d}{=} X, V \overset{d}{=} Y \), and \( U \) and \( V \) are independent. Hence, \((U, V) \overset{d}{=} (X, Y)\), so \((X_n, Y_n) \Rightarrow (X, Y)\).

**Lemma 1.4** Let \((E, r)\) be a complete and separable metric space. Let \( X_n \) be a sequence of \( E \)-valued random variables and suppose, for each \( k \), there exists a sequence \( \{X_{n,k}\}_{n=1}^{\infty} \) such that

\[
\limsup_{n \to \infty} E[r(X_n, X_{n,k})] \leq \delta_k,
\]

where \( \delta_k \to 0 \) as \( k \to \infty \). Suppose also that for each \( k \), there exists \( Y_k \) such that \( X_{n,k} \Rightarrow Y_k \) as \( n \to \infty \). Then there exists \( X \) such that \( X_n \Rightarrow X \) and \( Y_k \Rightarrow X \).

**Proof.** Let \( \mathcal{P}(E) \) be the family of all probability measures on \( E \), endowed with the Prohorov metric,

\[
\rho(P, Q) = \inf\{\varepsilon > 0 : P(F) \leq Q(F^\varepsilon) + \varepsilon \text{ for all } F \in \mathcal{C}\},
\]

where \( \mathcal{C} \) is the collection of closed sets in \( E \) and \( F^\varepsilon = \{x \in E : r(x, F) < \varepsilon\} \). Under this metric, \((\mathcal{P}(E), \rho)\) is complete and separable, and \( Z_n \Rightarrow Z \) if and only if \( \rho(PZ_n^{-1}, PZ^{-1}) \to 0 \).
Let $\varepsilon > 0$ be given and choose $k_0$ such that $\delta_k < \varepsilon^2$ whenever $k \geq k_0$. For each fixed $k \geq k_0$, choose $N(k)$ and $M(k)$ such that $E[r(X_n, X_{n,k})] < \varepsilon^2$ whenever $n \geq N(k)$ and $\rho(PX_{n,k}^{-1}, PY_k^{-1}) < \varepsilon$ whenever $n \geq M(k)$. Let $n \geq N(k)$ be arbitrary. Then for all $F \in C$,

$$P(X_n \in F) \leq P(X_n \in F, r(X_n, X_{n,k}) < \varepsilon) + P(r(X_n, X_{n,k}) \geq \varepsilon) \leq P(X_{n,k} \in F^c) + \varepsilon.$$ 

It follows then that $\rho(PX_n^{-1}, PX_{n,k}^{-1}) \leq \varepsilon$ whenever $n \geq N(k)$.

Now let $n, m \geq N(k_0) \lor M(k_0)$. Then

$$\rho(PX_n^{-1}, PX_m^{-1}) \leq \rho(PX_n^{-1}, PX_{n,k_0}^{-1}) + \rho(PX_{n,k_0}^{-1}, PY_k^{-1}) + \rho(PY_k^{-1}, PX_{m,k_0}^{-1}) + \rho(PX_{m,k_0}^{-1}, PX_m^{-1}) < 4\varepsilon.$$ 

Hence, $\{PX_n^{-1}\}$ is Cauchy in $P(E)$, so there exists $X$ such that $X_n \Rightarrow X$.

Now let $k \geq k_0$ and choose $n \geq N(k) \lor M(k)$ such that $\rho(PX_n^{-1}, PX_k^{-1}) < \varepsilon$. Then

$$\rho(PY_k^{-1}, PX_k^{-1}) \leq \rho(PY_k^{-1}, PX_{n,k}^{-1}) + \rho(PX_{n,k}^{-1}, PX_n^{-1}) + \rho(PX_n^{-1}, PX_k^{-1}) < 3\varepsilon.$$ 

Hence, $Y_k \Rightarrow X$. \hfill \Box

**Lemma 1.5** Suppose $x, y \in D$ and $x(t) = y(t)$ for all $t < T$. Then $d(x, y) \leq e^{-T}$.

**Proof.** Taking $\lambda(t) = t$ in ([1]) gives

$$d(x, y) \leq \int_0^\infty e^{-u} \sup_{t \in [0,u]} \{r(x(t), y(t)) \land 1\} \, du.$$ 

If $x(t) = y(t)$ for all $t < T$, then $d(x, y) \leq \int_T^\infty e^{-u} \, du = e^{-T}$. \hfill \Box

**Lemma 1.6** For $x \in D = D_{\mathbb{R}^d}(0, \infty)$ and $\varepsilon > 0$, let

$$h_\varepsilon(x) = \inf\{t \geq 0 : |x(t)| \land |x(t-)| \leq \varepsilon\}. \quad (2)$$

If $(x_n, h_\varepsilon(x_n)) \rightarrow (x, T) \in D \times [0, \infty)$, then $h_\varepsilon(x) \leq T$.

**Proof.** Let $s < t < h_\varepsilon(x)$, so that $\inf_{[0,t]} |x(r)| > \varepsilon$. Since $x(\cdot) \mapsto \inf_{[0,\cdot]} |x(r)|$ is continuous in the Skorohod topology, for sufficiently large $n$, $\inf_{[0,s]} |x_n(r)| > \varepsilon$, which implies that $s \leq h_\varepsilon(x_n)$. Letting $n \rightarrow \infty$ gives $s \leq T$. Letting $s \uparrow h_\varepsilon(x)$ gives $h_\varepsilon(x) \leq T$. \hfill \Box

## 2 The Context

These lemmas were proved while working on a paper in which we applied the theorems in [2]. We did not need the full power of these theorems. Rather, we simply used the following “watered down” versions.

This first theorem is a special case of Theorem 2.2 in [2].
Theorem 2.1 For each \( n \), let \( Y_n \) be a cadlag, \( \mathbb{R}^m \)-valued semimartingale with respect to a filtration \( \{ \mathcal{F}_t^n \} \). Suppose that \( Y_n = M_n + A_n \), where \( M_n \) is an \( \{ \mathcal{F}_t^n \} \)-local martingale and \( A_n \) is a finite variation process, and that

\[
\sup_n E[[M_n]_t + V_t(A_n)] < \infty
\]

for each \( t \geq 0 \), where \( V_t(A_n) \) is the total variation of \( A_n \) on \([0, t]\). Let \( X_n \) be a cadlag, \( \{ \mathcal{F}_t^n \} \)-adapted, \( \mathbb{R}^{k \times m} \)-valued process and define

\[
Z_n(t) = \int_0^t X_n(s-) dY_n(s).
\]

Suppose that \( (X_n, Y_n) \Rightarrow (X, Y) \) in \( D_{\mathbb{R}^{k \times m} \times \mathbb{R}^m}[0, \infty) \). Then \( Y \) is a semimartingale with respect to a filtration to which \( X \) and \( Y \) are adapted, and \( (X_n, Y_n, Z_n) \Rightarrow (X, Y, Z) \) in \( D_{\mathbb{R}^{k \times m} \times \mathbb{R}^m \times \mathbb{R}^k}[0, \infty) \), where

\[
Z(t) = \int_0^t X(s-) dY(s).
\]

If \( (X_n, Y_n) \rightarrow (X, Y) \) in probability, then \( Z_n \rightarrow Z \) in probability.

Remark 2.2 In the setting of Theorem 2.1, if \( \{ V_n \} \) is another sequence of cadlag adapted processes and \( (V_n, X_n, Y_n) \Rightarrow (V, X, Y) \), then \( (V_n, X_n, Y_n, Z_n) \Rightarrow (V, X, Y, Z) \). This can be seen by applying Theorem 2.1 to \( (X_n, Y_n) \), where \( \tilde{X}_n = (V_n, X_n) \) and \( \tilde{Y}_n = (0, Y_n)^T \).

This next theorem is a special case of Theorem 5.4 and Corollary 5.6 in [2].

Theorem 2.3 For each \( n \), let \( Y_n \) be a cadlag, \( \mathbb{R}^m \)-valued semimartingale with respect to a filtration \( \{ \mathcal{F}_t^n \} \). Suppose that \( \{ Y_n \} \) satisfies (3). Let \( U_n \) be a cadlag, \( \{ \mathcal{F}_t^n \} \)-adapted, \( \mathbb{R}^k \)-valued process and suppose that \( (U_n, Y_n) \Rightarrow (U, Y) \) in \( D_{\mathbb{R}^k \times \mathbb{R}^m}[0, \infty) \). Let \( G_n \) and \( G \) be continuous functions from \( \mathbb{R}^k \) to \( \mathbb{R}^{k \times m} \) such that \( G_n \rightarrow G \) uniformly on compacts, and suppose that \( X_n \) satisfies

\[
X_n(t) = U_n(t) + \int_0^t G_n(X_n(s-)) dY_n(s).
\]

Consider the integral equation

\[
X(t) = U(t) + \int_0^t G(X(s-)) dY(s).
\]

Suppose that for every version of \( (U, Y) \), this equation has a unique strong solution for all time \( t \geq 0 \). Then \( (U_n, X_n, Y_n) \Rightarrow (U, X, Y) \) in \( D_{\mathbb{R}^k \times \mathbb{R}^k \times \mathbb{R}^m}[0, \infty) \). If \( (U_n, Y_n) \rightarrow (U, Y) \) in probability, then \( X_n \rightarrow X \) in probability.

Remark 2.4 As in Remark 2.2, if \( \{ V_n \} \) is another sequence of cadlag adapted processes and \( (V_n, U_n, Y_n) \Rightarrow (V, U, Y) \), then \( (V_n, U_n, X_n, Y_n) \Rightarrow (V, U, X, Y) \). This can be seen by applying Theorem 2.3 to \( (U_n, Y_n) \) and \( \tilde{G}_n \), where \( \tilde{U}_n = (V_n, U_n) \) and \( \tilde{G}_n = (0, G_n) \).

The final theorem in this section is a generalization of these two, which follows from Remark 2.5 in [2].
Theorem 2.5 Suppose all of the hypotheses of Theorem 2.1 (or Theorem 2.3) hold, except for (3). If \( \{(\bar{X}_n, \bar{Y}_n)\} \) is relatively compact in \( D_{\mathbb{R}^m \times \mathbb{R}^\ell \times [0, \infty)} \), \( \{\bar{Y}_n\} \) satisfies (3), and

\[
Y_n(t) = \int_0^t \bar{X}_n(s-) \, d\bar{Y}_n(s),
\]

then the conclusions of Theorem 2.1 (or Theorem 2.3) hold.

References
