

Wiener measure and Donsker's theorem

Jordan Bell

jordan.bell@gmail.com

Department of Mathematics, University of Toronto

September 4, 2015

1 Relatively compact sets of Borel probability measures on $C[0, 1]$

Let $E = C[0, 1]$, let \mathcal{B}_E be the Borel σ -algebra of E , and let \mathcal{P}_E be the collection of Borel probability measures on E . We assign \mathcal{P} the **narrow topology**, the coarsest topology on \mathcal{P}_E such that for each $F \in C_b(E)$ the map $\mu \mapsto \int_E F d\mu$ is continuous.

For $f \in E$ and $\delta > 0$ we define

$$\omega_f(\delta) = \sup_{s, t \in [0, 1], |s-t| \leq \delta} |f(s) - f(t)|.$$

For $f \in E$, $\omega_f(\delta) \downarrow 0$ as $\delta \downarrow 0$, and for $\delta > 0$, $f \mapsto \omega_f(\delta)$ is continuous. We shall use the following characterization of a relatively compact subset A of E , which is proved using the Arzelà-Ascoli theorem.

Lemma 1. *Let A be a subset of E . \overline{A} is compact if and only if*

$$\sup_{f \in A} |f(0)| < \infty$$

and

$$\sup_{f \in A} \omega_f(\delta) \downarrow 0, \quad \delta \downarrow 0.$$

We shall use **Prokhorov's theorem**:¹ for X a Polish space and for $\Gamma \subset \mathcal{P}_X$, $\overline{\Gamma}$ is compact if and only if for each $\epsilon > 0$ there is a compact subset K_ϵ of X such that $\mu(K_\epsilon) \geq 1 - \epsilon$ for all $\mu \in \Gamma$. Namely, a subset of \mathcal{P}_X is relatively compact if and only if it is **tight**. We use Prokhorov's theorem to prove a characterization of relatively compact subsets of \mathcal{P}_E , which we then use to prove the characterization in Theorem 3.²

¹K. R. Parthasarathy, *Probability Measures on Metric Spaces*, p. 47, Chapter II, Theorem 6.7.

²K. R. Parthasarathy, *Probability Measures on Metric Spaces*, p. 213, Chapter VII, Lemma 2.2.

Lemma 2. *Let Γ be a subset of \mathcal{P}_E . $\bar{\Gamma}$ is compact if and only if for each $\epsilon > 0$ there is some $M_\epsilon < \infty$ and a function $\delta \mapsto \omega_\epsilon(\delta)$ satisfying $\omega_\epsilon(\delta) \downarrow 0$ as $\delta \downarrow 0$ and such that for all $\mu \in \Gamma$,*

$$\mu(A_\epsilon) \geq 1 - \frac{\epsilon}{2}, \quad \mu(B_\epsilon) \geq 1 - \frac{\epsilon}{2},$$

where

$$A_\epsilon = \{f \in E : |f(0)| \leq M_\epsilon\}, \quad B_\epsilon = \{f \in E : \omega_f(\delta) \leq \omega_\epsilon(\delta) \text{ for all } \delta > 0\}.$$

Proof. Suppose that Γ satisfies the above conditions. Because $f \mapsto |f(0)|$ is continuous, A_ϵ is closed. For $\delta > 0$, suppose that f_n is a sequence in B_ϵ tending to some $f \in E$. Because $g \mapsto \omega_g(\delta)$ is continuous, $\omega_{f_n}(\delta) \rightarrow \omega_f(\delta)$, and because $\omega_{f_n}(\delta) \leq \omega_\epsilon(\delta)$ for each n , we get $\omega_f(\delta) \leq \omega_\epsilon(\delta)$ and hence $f \in B_\epsilon$, showing that B_ϵ is closed. Therefore $K_\epsilon = A_\epsilon \cap B_\epsilon$ is closed, i.e. $K_\epsilon = \overline{K_\epsilon}$. The set K_ϵ satisfies

$$\sup_{f \in K_\epsilon} |f(0)| \leq M_\epsilon$$

and

$$\limsup_{\delta \downarrow 0} \sup_{f \in K_\epsilon} \omega_f(\delta) \leq \limsup_{\delta \downarrow 0} \omega_\epsilon(\delta) = 0,$$

thus by Lemma 1, K_ϵ is compact. For $\mu \in \Gamma$,

$$\mu(K_\epsilon) \geq 1 - \frac{\epsilon}{2},$$

and because K_ϵ is compact, this means that Γ is tight, so by Prokhorov's theorem, Γ is relatively compact.

Now suppose that Γ is relatively compact and let $\epsilon > 0$. By Prokhorov's theorem, there is a compact set K_ϵ in E such that $\mu(K_\epsilon) \geq 1 - \frac{\epsilon}{2}$ for all $\mu \in \Gamma$. Define

$$M_\epsilon = \sup_{f \in K_\epsilon} |f(0)|, \quad \omega_\epsilon(\delta) = \sup_{f \in K_\epsilon} \omega_f(\delta), \quad \delta > 0.$$

Because K_ϵ is compact, by Lemma 1 we get that $M_\epsilon < \infty$ and $\omega_\epsilon(\delta) \downarrow 0$ as $\delta \downarrow 0$. For $\mu \in \Gamma$,

$$\mu(A_\epsilon) \geq \mu(K_\epsilon) \geq 1 - \frac{\epsilon}{2}, \quad \mu(B_\epsilon) \geq \mu(K_\epsilon) \geq 1 - \frac{\epsilon}{2},$$

showing that Γ satisfies the conditions of the theorem. \square

We now prove the characterization of relatively compact subsets of \mathcal{P}_E that we shall use in our proof of Donsker's theorem.³

Theorem 3 (Relatively compact sets in \mathcal{P}). *Let Γ be a subset of \mathcal{P}_E . $\bar{\Gamma}$ is compact if and only if the following conditions are satisfied:*

³K. R. Parthasarathy, *Probability Measures on Metric Spaces*, p. 214, Chapter VII, Theorem 2.2.

1. For each $\epsilon > 0$ there is some $M_\epsilon < \infty$ such that

$$\mu(f : |f(0)| \leq M_\epsilon) \geq 1 - \frac{\epsilon}{2}, \quad \mu \in \Gamma.$$

2. For each $\epsilon > 0$ and $\delta > 0$ there is some $\eta = \eta(\epsilon, \delta) > 0$ such that

$$\mu(f : \omega_f(\eta) \leq \delta) \geq 1 - \frac{\epsilon}{2}, \quad \mu \in \Gamma.$$

Proof. Suppose that $\bar{\Gamma}$ is compact and let $\epsilon > 0$. By Lemma 2, there is some $M_\epsilon < \infty$ and a function $\eta \mapsto \omega_\epsilon(\eta)$ satisfying $\omega_\epsilon(\eta) \downarrow 0$ as $\eta \downarrow 0$ and

$$\mu(A_\epsilon) \geq 1 - \frac{\epsilon}{2}, \quad \mu(B_\epsilon) \geq 1 - \frac{\epsilon}{2}, \quad \mu \in \Gamma.$$

For $\delta > 0$, there is some $\eta = \eta(\epsilon, \delta)$ with $\omega_\epsilon(\eta) \leq \delta$. Then for $\mu \in \Gamma$,

$$\mu(f : \omega_f(\eta) \leq \delta) \geq \mu(f : \omega_f(\eta) \leq \omega_\epsilon(\eta)) \geq \mu(B_\epsilon) \geq 1 - \frac{\epsilon}{2}.$$

Now suppose that the conditions of the theorem hold. For each $\epsilon > 0$ and $n \geq 1$ there is some $\eta_{\epsilon,n} > 0$ such that

$$\mu(F_{\epsilon,n}) \geq 1 - \frac{\epsilon}{2^{n+1}}, \quad \mu \in \Gamma,$$

where

$$F_{\epsilon,n} = \left\{ f : \omega_f(\eta_{\epsilon,n}) \leq \frac{1}{n} \right\}.$$

Let

$$K_\epsilon = \{f : |f(0)| \leq M_\epsilon\} \cap \bigcap_{n=1}^{\infty} F_{\epsilon,n},$$

for which

$$\mu(K_\epsilon) \geq \mu(f : |f(0)| \leq M_\epsilon) \geq 1 - \frac{\epsilon}{2}, \quad \mu \in \Gamma.$$

For $f \in K_\epsilon$, then for each $n \geq 1$ we have $f \in F_{\epsilon,n}$, which means that $\omega_f(\eta_{\epsilon,n}) \leq \frac{1}{n}$, and therefore

$$\sup_{f \in K_\epsilon} \omega_f(\eta_{\epsilon,n}) \leq \frac{1}{n}.$$

Thus for $n \geq 1$, if $0 < \eta \leq \eta_{\epsilon,n}$ then

$$\sup_{f \in K_\epsilon} \omega_f(\eta) \leq \frac{1}{n},$$

which shows $\sup_{f \in K_\epsilon} \omega_f(\eta) \downarrow 0$ as $\eta \downarrow 0$. Then because

$$\sup_{f \in K_\epsilon} |f(0)| \leq M_\epsilon,$$

applying Lemma 1 we get that $\overline{K_\epsilon}$ is compact. The map $f \mapsto \omega_f(\eta_{\epsilon,n})$ is continuous, so the set $F_{\epsilon,n}$ is closed, and therefore the set K_ϵ is closed. Because K_ϵ is compact and $\mu(K_\epsilon) \geq 1 - \frac{\epsilon}{2}$ for all $\mu \in \Gamma$, it follows from by Prokhorov's theorem that Γ is relatively compact. \square

2 Wiener measure

For $t_1, \dots, t_d \in [0, 1]$, $t_1 < \dots < t_d$, define $\pi_{t_1, \dots, t_d} : E \rightarrow \mathbb{R}^d$ by

$$\pi_{t_1, \dots, t_d}(f) = (f(t_1), \dots, f(t_d)), \quad f \in E,$$

which is continuous. We state the following results, which we will use later.⁴

Theorem 4 (The Borel σ -algebra of E). *\mathcal{B}_E is equal to the σ -algebra generated by $\{\pi_t : t \in [0, 1]\}$.*

Two elements μ and ν of \mathcal{P}_E are equal if and only if for any d and any $t_1 < \dots < t_d$, the pushforward measures

$$\mu_{t_1, \dots, t_d} = (\pi_{t_1, \dots, t_d})_* \mu, \quad \nu_{t_1, \dots, t_d} = (\pi_{t_1, \dots, t_d})_* \nu$$

are equal.

Let $(\xi_t)_{t \in [0, 1]}$ be a stochastic process with state space \mathbb{R} and sample space (Ω, \mathcal{F}, P) . For $t_1 < \dots < t_d$, let $\xi_{t_1, \dots, t_d} = \xi_{t_1} \otimes \dots \otimes \xi_{t_d}$ and let $P_{t_1, \dots, t_d} = (\xi_{t_1, \dots, t_d})_* P$: for $B \in \mathcal{B}_{\mathbb{R}^d}$,

$$P_{t_1, \dots, t_d}(B) = ((\xi_{t_1, \dots, t_d})_* P)(B) = P(\xi_{t_1, \dots, t_d}^{-1}(B)) = P((\xi_{t_1}, \dots, \xi_{t_d}) \in B).$$

P_{t_1, \dots, t_d} is a Borel probability measure on \mathbb{R}^d and is called a **finite-dimensional distribution of the stochastic process**.

The **Kolmogorov continuity theorem**⁵ tells us that if there are $\alpha, \beta, K > 0$ such that for all $s, t \in [0, 1]$,

$$E|\xi_t - \xi_s|^\alpha \leq K|t - s|^{1+\beta},$$

then there is a unique $\mu \in \mathcal{P}_E$ such that for all k and for all $t_1 < \dots < t_d$,

$$\mu_{t_1, \dots, t_d} = P_{t_1, \dots, t_d}.$$

We now define and prove the existence of **Wiener measure**.⁶

Theorem 5 (Wiener measure). *There is a unique Borel probability measure W on E satisfying:*

1. $W(f \in E : f(0) = 0) = 1$.
2. For $0 \leq t_0 < t_1 < \dots < t_d \leq 1$ the random variables

$$\pi_{t_1} - \pi_{t_0}, \quad \pi_{t_2} - \pi_{t_1}, \quad \pi_{t_3} - \pi_{t_2}, \quad \pi_{t_d} - \pi_{t_{d-1}}$$

are independent $(E, \mathcal{B}_E, W) \rightarrow (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$.

⁴K. R. Parthasarathy, *Probability Measures on Metric Spaces*, p. 212, Chapter VII, Theorem 2.1.

⁵K. R. Parthasarathy, *Probability Measures on Metric Spaces*, p. 216, Chapter VII, Theorem 3.1.

⁶K. R. Parthasarathy, *Probability Measures on Metric Spaces*, p. 218, Chapter VII, Theorem 3.2.

3. If $0 \leq s < t \leq 1$, the random variable $\pi_t - \pi_s : (E, \mathcal{B}_E, W) \rightarrow (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ is normal with mean 0 and variance $t - s$.

Proof. There is a stochastic process $(\xi_t)_{t \in [0,1]}$ with state space \mathbb{R} and some sample space (Ω, \mathcal{F}, P) , such that (i) $P(\xi_0 = 0) = 1$, (ii) $(\xi_t)_{t \in [0,1]}$ has independent increments, and (iii) for $s < t$, $\xi_t - \xi_s$ is a normal random variable with mean 0 and variance $t - s$. (Namely, **Brownian motion with starting point 0**.) Because $\xi_t - \xi_s$ has mean 0 and variance $t - s$, we calculate (cf. Isserlis's theorem)

$$E|\xi_t - \xi_s|^4 = 3|t - s|^2.$$

Thus using the Kolmogorov continuity theorem with $\alpha = 4$, $\beta = 1$, $K = 3$, there is a unique $W \in \mathcal{P}_E$ such that for all $t_1 < \dots < t_d$,

$$W_{t_1, \dots, t_d} = P_{t_1, \dots, t_d},$$

i.e. for $B \in \mathcal{B}_{\mathbb{R}}^d$,

$$W(\pi_{t_1} \otimes \dots \otimes \pi_{t_d} \in B) = P(\xi_{t_1} \otimes \dots \otimes \xi_{t_d} \in B).$$

For $t_1 < \dots < t_d$ and $B \in \mathcal{B}_{\mathbb{R}}^d$, with $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$ defined by $T(x_1, \dots, x_d) = (x_1, x_2 - x_1, \dots, x_d - x_{d-1})$,

$$\begin{aligned} & W(\pi_{t_1} \otimes (\pi_{t_2} - \pi_{t_1}) \otimes \dots \otimes (\pi_{t_d} - \pi_{t_{d-1}}) \in B) \\ &= W(T \circ (\pi_{t_1} \otimes \pi_{t_2} \otimes \dots \otimes \pi_{t_d}) \in B) \\ &= W(\pi_{t_1} \otimes \pi_{t_2} \otimes \dots \otimes \pi_{t_d} \in T^{-1}(B)) \\ &= P(\xi_{t_1} \otimes \xi_{t_2} \otimes \dots \otimes \xi_{t_d} \in T^{-1}(B)) \\ &= P(T \circ (\xi_{t_1} \otimes \xi_{t_2} \otimes \dots \otimes \xi_{t_d}) \in B) \\ &= P(\xi_{t_1} \otimes (\xi_{t_2} - \xi_{t_1}) \otimes \dots \otimes (\xi_{t_d} - \xi_{t_{d-1}}) \in B). \end{aligned}$$

Hence, because $\xi_{t_1}, \xi_{t_2} - \xi_{t_1}, \dots, \xi_{t_d} - \xi_{t_{d-1}}$ are independent,

$$\begin{aligned} & (\pi_{t_1} \otimes (\pi_{t_2} - \pi_{t_1}) \otimes \dots \otimes (\pi_{t_d} - \pi_{t_{d-1}}))_* W \\ &= (\xi_{t_1} \otimes (\xi_{t_2} - \xi_{t_1}) \otimes \dots \otimes (\xi_{t_d} - \xi_{t_{d-1}}))_* P \\ &= (\xi_{t_1})_* P \otimes (\xi_{t_2} - \xi_{t_1})_* P \otimes \dots \otimes (\xi_{t_d} - \xi_{t_{d-1}})_* P \\ &= (\pi_{t_1})_* W \otimes (\pi_{t_2} - \pi_{t_1})_* W \otimes \dots \otimes (\pi_{t_d} - \pi_{t_{d-1}})_* W, \end{aligned}$$

which means that the random variables $\pi_{t_1}, \pi_{t_2} - \pi_{t_1}, \dots, \pi_{t_d} - \pi_{t_{d-1}}$ are independent.

If $s < t$ and $B_1, B_2 \in \mathcal{B}_{\mathbb{R}}$, and for $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by $T(x, y) = (x, y - x)$,

$$\begin{aligned} W((\pi_s, \pi_t - \pi_s) \in (B_1, B_2)) &= W(T \circ (\pi_s, \pi_t) \in (B_1, B_2)) \\ &= P((\xi_s, \xi_t) \in T^{-1}(B_1, B_2)) \\ &= P((\xi_s, \xi_t - \xi_s) \in (B_1, B_2)), \end{aligned}$$

which implies that $(\pi_t - \pi_s)_*W = (\xi_t - \xi_s)_*P$, and because $\xi_t - \xi_s$ is a normal random variable with mean 0 and variance $t - s$, so is $\pi_t - \pi_s$.

Finally,

$$W(f : f(0) = 0) = W(\pi_0 = 0) = P(\xi_0 = 0) = 1.$$

□

(E, \mathcal{B}_E, W) is a probability space, and the stochastic process $(\pi_t)_{t \in [0,1]}$ is a Brownian motion.

3 Interpolation and continuous stochastic processes

Let $(\xi_t)_{t \in [0,1]}$ be a **continuous stochastic process** with state space \mathbb{R} and sample space (Ω, \mathcal{F}, P) . To say that the stochastic process is continuous means that for each $\omega \in \Omega$ the map $t \mapsto \xi_t(\omega)$ is continuous $[0, 1] \rightarrow \mathbb{R}$. Define $\xi : \Omega \rightarrow E$ by

$$\xi(\omega) = (t \mapsto \xi_t(\omega)), \quad \omega \in \Omega.$$

For $t \in [0, 1]$ and B a Borel set in \mathbb{R} ,

$$\xi^{-1}\pi_t^{-1}B = \{\omega \in \Omega : \xi_t(\omega) \in B\} = \xi_t^{-1}B,$$

and because $\xi_t : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ is measurable this belongs to \mathcal{F} . But by Theorem 4, \mathcal{B}_E is generated by the collection $\{\pi_t^{-1}B : t \in [0, 1], B \in \mathcal{B}_{\mathbb{R}}\}$. Now, for $f : X \rightarrow Y$ and for a nonempty collection \mathcal{F} of subsets of Y ,⁷

$$\sigma(f^{-1}(\mathcal{F})) = f^{-1}(\sigma(\mathcal{F})).$$

Therefore $\xi^{-1}(\mathcal{B}_E) \subset \mathcal{F}$, which means that $\xi : (\Omega, \mathcal{F}) \rightarrow (E, \mathcal{B}_E)$ is measurable. This means that a continuous stochastic process with index set $[0, 1]$ induces a random variable with state space E . Then the pushforward measure of P by ξ is a Borel probability measure on E . We shall end up constructing a sequence of pushforward measures from a sequence of continuous stochastic processes, that converge in \mathcal{P}_E to Wiener measure W .

Let $(X_n)_{n \geq 1}$ be a sequence of independent identically distributed random variables on a sample space (Ω, \mathcal{F}, P) with $E(X_n) = 0$ and $V(X_n) = 1$, and let $S_0 = 0$ and

$$S_k = \sum_{i=1}^k X_i.$$

Then $E(S_k) = 0$ and $V(S_k) = k$. For $t \geq 0$ let

$$Y_t = S_{[t]} + (t - [t])X_{[t]+1}.$$

⁷Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 140, Lemma 4.23.

Thus, for $k \geq 0$ and $k \leq t \leq k+1$,

$$\begin{aligned} Y_t &= S_k + (t-k)X_{k+1} \\ &= S_k + (t-k)(S_{k+1} - S_k) \\ &= (1-t+k)S_k + (t-k)S_{k+1}. \end{aligned}$$

For each $\omega \in \Omega$, the map $t \mapsto Y_t(\omega)$ is piecewise linear, equal to $S_k(\omega)$ when $t = k$, and in particular it is continuous. For $n \geq 1$, define

$$X_t^{(n)} = n^{-1/2}Y_{nt} = n^{-1/2}S_{[nt]} + n^{-1/2}(nt - [nt])X_{[nt]+1}, \quad t \in [0, 1]. \quad (1)$$

For $0 \leq k \leq n$,

$$X_{k/n}^{(n)} = n^{-1/2}S_k.$$

For each $n \geq 1$, $(X_t^{(n)})_{t \in [0,1]}$ is a continuous stochastic process on the sample space (Ω, \mathcal{F}, P) , and we denote by $P_n \in \mathcal{P}_E$ the pushforward measure of P by $X^{(n)}$.

4 Donsker's theorem

Lemma 6. *If Z_n and U_n are random variables with state space \mathbb{R}^d such that $Z_n \rightarrow Z$ in distribution and $U_n \rightarrow 0$ in distribution, then $Z_n + U_n \rightarrow Z$ in distribution.*

If Z_n are random variables with state space \mathbb{R} that converge in distribution to some random variable Z and c_n are real numbers that converge to some real number c , then $c_n Z_n \rightarrow cZ$ in distribution.

For $\sigma \geq 0$, let ν_{σ^2} be the Gaussian measure on \mathbb{R} with mean 0 and variance σ^2 . The **characteristic function** of ν_{σ^2} is, for $\sigma > 0$,

$$\tilde{\nu}_{\sigma^2}(\xi) = \int_{\mathbb{R}} e^{i\xi x} d\nu_{\sigma^2}(x) = \int_{\mathbb{R}} e^{i\xi x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx = e^{-\frac{1}{2}\sigma^2\xi^2},$$

and $\tilde{\nu}_0(\xi) = 1$. One checks that $c_*\nu_1 = \nu_{c^2}$ for $c \geq 0$.

In following theorem and in what follows, $X^{(n)}$ is the piecewise linear stochastic process defined in (1). We prove that a sequence of finite-dimensional distributions converge to a Gaussian measure.⁸

Theorem 7. *For $0 \leq t_0 < t_1 < t_1 < \dots < t_d \leq 1$, the random vectors*

$$(X_{t_1}^{(n)} - X_{t_0}^{(n)}, \dots, X_{t_d}^{(n)} - X_{t_{d-1}}^{(n)}), \quad (\Omega, \mathcal{F}, P) \rightarrow (\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d}),$$

converge in distribution to $\nu_{t_1-t_0} \otimes \dots \otimes \nu_{t_d-t_{d-1}}$ as $n \rightarrow \infty$.

⁸Bert Fristedt and Lawrence Gray, *A Modern Approach to Probability Theory*, p. 368, §19.1, Lemma 1.

Proof. For $0 < j \leq d$ and $n \geq 1$ let

$$r_{j,n} = \frac{[nt_j]}{n}, \quad U_{j,n} = X_{t_j}^{(n)} - X_{r_{j,n}}^{(n)},$$

and for $0 \leq j < d$ and $n \geq 1$ let

$$s_{j,n} = \frac{[nt_j]}{n}, \quad V_{j,n} = X_{s_{j,n}}^{(n)} - X_{t_j}^{(n)},$$

with which

$$\begin{aligned} (X_{t_1}^{(n)} - X_{t_0}^{(n)}, \dots, X_{t_d}^{(n)} - X_{t_{d-1}}^{(n)}) &= (X_{r_{1,n}}^{(n)} - X_{s_{0,n}}^{(n)}, \dots, X_{r_{d,n}}^{(n)} - X_{s_{d-1,n}}^{(n)}) \\ &\quad + (U_{1,n}, \dots, U_{d,n}) + (V_{0,n}, \dots, V_{d-1,n}). \end{aligned}$$

Because $E(X_t^{(n)}) = 0$,

$$E(U_{j,n}) = 0, \quad E(V_{j,n}) = 0.$$

Furthermore,

$$\begin{aligned} &V(U_{j,n}) \\ &= V(X_{t_j}^{(n)} - X_{r_{j,n}}^{(n)}) \\ &= n^{-1}V(S_{[nt_j]} + (nt_j - [nt_j])X_{[nt_j]+1} - S_{[nr_{j,n}]} - (nr_{j,n} - [nr_{j,n}])X_{[nr_{j,n}]+1}) \\ &= n^{-1}V(S_{[nt_j]} + (nt_j - [nt_j])X_{[nt_j]+1} - S_{[nt_j]} - ([nt_j] - [nt_j])X_{[nr_{j,n}]+1}) \\ &= n^{-1}(nt_j - [nt_j])^2V(X_{[nt_j]+1}) \\ &= n^{-1}(nt_j - [nt_j])^2, \end{aligned}$$

and because $0 \leq nt_j - [nt_j] < 1$ this tends to 0 as $n \rightarrow \infty$. Likewise, $V(V_{j,n}) \rightarrow 0$ as $n \rightarrow \infty$.

For $1 \leq j \leq d$,

$$\begin{aligned} X_{r_{j,n}}^{(n)} - X_{s_{j-1,n}}^{(n)} &= n^{-1/2}S_{[nr_{j,n}]} + n^{-1/2}(nr_{j,n} - [nr_{j,n}])X_{[nr_{j,n}]+1} \\ &\quad - n^{-1/2}S_{[ns_{j-1,n}]} - n^{-1/2}(ns_{j-1,n} - [ns_{j-1,n}])X_{[ns_{j-1,n}]+1} \\ &= n^{-1/2}S_{[nt_j]} - n^{-1/2}S_{[nt_{j-1}]} \\ &= n^{-1/2} \frac{([nt_j] - [nt_{j-1}] - 1)^{1/2}}{([nt_j] - [nt_{j-1}] - 1)^{1/2}} \sum_{i=[nt_{j-1}]+1}^{[nt_j]} X_i. \end{aligned}$$

By the central limit theorem,

$$([nt_j] - [nt_{j-1}] - 1)^{1/2} \sum_{i=[nt_{j-1}]+1}^{[nt_j]} X_i \rightarrow \nu_1$$

in distribution as $n \rightarrow \infty$. But

$$n^{-1/2}([nt_j] - [nt_{j-1}] - 1)^{1/2} \rightarrow (t_j - t_{j-1})^{1/2}$$

as $n \rightarrow \infty$, and $(t_j - t_{j-1})_*^{1/2} \nu_1 = \nu_{t_j - t_{j-1}}$, so by Lemma 6,

$$X_{r_j, n}^{(n)} - X_{s_{j-1}, n}^{(n)} \rightarrow \nu_{t_j - t_{j-1}}$$

in distribution as $n \rightarrow \infty$.

For sufficiently large n , depending on t_0, \dots, t_d ,

$$t_0 \leq s_{0,n} < r_{1,n} \leq t_1 \leq s_{1,n} < r_{2,n} \leq \dots \leq t_{d-1} \leq s_{d-1,n} < r_{d,n} \leq t_d.$$

Check that $(U_{1,n}, \dots, U_{d,n}) \rightarrow 0$ in probability and that $(V_{0,n}, \dots, V_{d-1,n}) \rightarrow 0$ in probability, and hence these random vectors converge to 0 in distribution as $n \rightarrow \infty$. The random variables $X_{r_{1,n}}^{(n)} - X_{s_{0,n}}^{(n)}, \dots, X_{r_{d,n}}^{(n)} - X_{s_{d-1,n}}^{(n)}$ are independent, and therefore their joint distribution is equal to the product of their distributions. Now, if $\mu_n = \mu_n^1 \otimes \dots \otimes \mu_n^d$ and $\mu_n^j \rightarrow \mu^j$ as $n \rightarrow \infty$, $1 \leq j \leq d$, then for $\xi \in \mathbb{R}^d$,

$$\begin{aligned} \tilde{\mu}_n(\xi) &= \tilde{\mu}_n^1(\xi_1) \cdots \tilde{\mu}_n^d(\xi_d) \\ &\rightarrow \tilde{\mu}^1(\xi_1) \cdots \tilde{\mu}^d(\xi_d) \\ &= (\mu^1 \otimes \dots \otimes \mu^d)^\sim(\xi) \end{aligned}$$

as $n \rightarrow \infty$, and therefore by **Lévy's continuity theorem**, $\mu_n \rightarrow \mu^1 \otimes \dots \otimes \mu^d$ as $n \rightarrow \infty$. This means that the joint distribution of $X_{r_{1,n}}^{(n)} - X_{s_{0,n}}^{(n)}, \dots, X_{r_{d,n}}^{(n)} - X_{s_{d-1,n}}^{(n)}$ converges to

$$\nu_{t_1 - t_0} \otimes \dots \otimes \nu_{t_d - t_{d-1}}$$

as $n \rightarrow \infty$. Because $(U_{1,n}, \dots, U_{d,n}) \rightarrow 0$ in distribution as $n \rightarrow \infty$ and $(V_{0,n}, \dots, V_{d-1,n}) \rightarrow 0$ in distribution as $n \rightarrow \infty$, applying Lemma 6 we get that

$$(X_{t_1}^{(n)} - X_{t_0}^{(n)}, \dots, X_{t_d}^{(n)} - X_{t_{d-1}}^{(n)}) \rightarrow \nu_{t_1 - t_0} \otimes \dots \otimes \nu_{t_d - t_{d-1}}$$

in distribution as $n \rightarrow \infty$, completing the proof. \square

Let $t_0 = 0$ and let $0 < t_1 < \dots < t_d \leq 1$. As $X_0^{(n)} = 0$, the above lemma tells us that

$$(X_{t_1}^{(n)}, X_{t_2}^{(n)} - X_{t_1}^{(n)}, \dots, X_{t_d}^{(n)} - X_{t_{d-1}}^{(n)}) \rightarrow \nu_{t_1} \otimes \nu_{t_2 - t_1} \otimes \dots \otimes \nu_{t_d - t_{d-1}}$$

in distribution as $n \rightarrow \infty$. Define $g : \mathbb{R}^d \rightarrow \mathbb{R}^d$ by

$$g(x_1, x_2, \dots, x_d) = (x_1, x_1 + x_2, \dots, x_1 + x_2 + \dots + x_d).$$

The function g is continuous and satisfies

$$g \circ (X_{t_1}^{(n)} - X_{t_0}^{(n)}, \dots, X_{t_d}^{(n)} - X_{t_{d-1}}^{(n)}) = (X_{t_1}^{(n)}, X_{t_2}^{(n)}, \dots, X_{t_d}^{(n)}).$$

Then by the **continuous mapping theorem**,

$$(X_{t_1}^{(n)}, X_{t_2}^{(n)} - X_{t_1}^{(n)}, \dots, X_{t_d}^{(n)} - X_{t_{d-1}}^{(n)}) \rightarrow g_*(\nu_{t_1} \otimes \nu_{t_2 - t_1} \otimes \dots \otimes \nu_{t_d - t_{d-1}}) \quad (2)$$

in distribution as $n \rightarrow \infty$.⁹

We prove a result that we use to prove the next lemma, and that lemma is used in the proof of Donsker's theorem.¹⁰

Lemma 8. For $\epsilon > 0$,

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{\delta} P \left(\max_{1 \leq j \leq [n\delta]+1} |S_j| > \epsilon n^{1/2} \right) = 0.$$

Proof. For each $\delta > 0$, by the central limit theorem,

$$([n\delta] + 1)^{-1/2} S_{[n\delta]+1} \rightarrow Z$$

in distribution as $n \rightarrow \infty$, where $Z_* P = \nu_1$. Because $\frac{([n\delta]+1)^{1/2}}{(n\delta)^{1/2}} \rightarrow 1$ as $n \rightarrow \infty$, by Lemma 6 we then get that

$$(n\delta)^{-1/2} S_{[n\delta]+1} \rightarrow Z$$

in distribution as $n \rightarrow \infty$. Now let $\lambda > 0$, and there is a sequence ϕ_k in $C_b(\mathbb{R})$ such that $\phi_k \downarrow 1_{(-\infty, -\lambda] \cup [\lambda, \infty)} = \chi_\lambda$ pointwise as $k \rightarrow \infty$. For each k , writing $X = S_{[n\delta]+1}$, using the change of variables formula,

$$\begin{aligned} P(|X| \geq \lambda(n\delta)^{1/2}) &= \int_{\Omega} \chi_{\lambda(n\delta)^{1/2}}(X(\omega)) dP(\omega) \\ &= \int_{\Omega} \chi_{\lambda}((n\delta)^{-1/2} X(\omega)) dP(\omega) \\ &\leq \int_{\Omega} \phi_k((n\delta)^{-1/2} X(\omega)) dP(\omega) \\ &= E(\phi_k((n\delta)^{-1/2} X)). \end{aligned}$$

Therefore, by the continuous mapping theorem,

$$\begin{aligned} \limsup_{n \rightarrow \infty} P(|S_{[n\delta]+1}| \geq \lambda(n\delta)^{1/2}) &\leq \lim_{n \rightarrow \infty} E(\phi_k((n\delta)^{-1/2} S_{[n\delta]+1})) \\ &= E(\phi_k \circ Z). \end{aligned}$$

Because $\phi_k \downarrow \chi_\lambda$ pointwise as $k \rightarrow \infty$, using the monotone convergence theorem and then using Chebyshev's inequality,

$$E(\phi_k \circ Z) \rightarrow E(\chi_\lambda \circ Z) = P(|Z| \geq \lambda) \leq \lambda^{-3} E|Z|^3.$$

We have established that for each $\lambda > 0$,

$$\limsup_{n \rightarrow \infty} P(|S_{[n\delta]+1}| \geq \lambda(n\delta)^{1/2}) \leq \lambda^{-3} E|Z|^3. \quad (3)$$

⁹Allan Gut, *Probability: A Graduate Course*, second ed., p. 245, Chapter 5, Theorem 10.4.

¹⁰Ioannis Karatzas and Steven E. Shreve, *Brownian Motion and Stochastic Calculus*, second ed., p. 68, Lemma 4.18.

Define

$$\tau = \min\{j \geq 1 : |S_j| > n^{1/2}\epsilon\}.$$

For $0 < \delta < \epsilon^2/2$, it is a fact that

$$\begin{aligned} & P\left(\max_{0 \leq j \leq [n\delta]+1} |S_j| > n^{1/2}\epsilon\right) \\ & \leq P(|S_{[n\delta]+1}| \geq n^{1/2}(\epsilon - (2\delta)^{1/2})) \\ & \quad + \sum_{j=1}^{[n\delta]} P(|S_{[n\delta]+1}| < n^{1/2}(\epsilon - (2\delta)^{1/2}) | \tau = j) P(\tau = j). \end{aligned}$$

If $\tau(\omega) = j$ and $|S_{[n\delta]+1}(\omega)| < n^{1/2}(\epsilon - (2\delta)^{1/2})$ then

$$|S_j(\omega) - S_{[n\delta]+1}(\omega)| \geq |S_j(\omega)| - |S_{[n\delta]+1}(\omega)| > n^{1/2}\epsilon - n^{1/2}(\epsilon - (2\delta)^{1/2}) = (2n\delta)^{1/2}.$$

But by Chebyshev's inequality and the fact that the random variables X_1, X_2, \dots are independent with mean 0 and variance 1,

$$P(|S_j - S_{[n\delta]+1}| > (2n\delta)^{1/2}) \leq \frac{1}{2n\delta} E((S_j - S_{[n\delta]+1})^2) = \frac{1}{2n\delta} ([n\delta] - j) \leq \frac{1}{2},$$

so

$$P(|S_{[n\delta]+1}(\omega)| < n^{1/2}(\epsilon - (2\delta)^{1/2}) | \tau = j) \leq \frac{1}{2}.$$

Therefore,

$$\begin{aligned} & P\left(\max_{0 \leq j \leq [n\delta]+1} |S_j| > n^{1/2}\epsilon\right) \\ & \leq P(|S_{[n\delta]+1}| \geq n^{1/2}(\epsilon - (2\delta)^{1/2})) + \sum_{j=1}^{[n\delta]} \frac{1}{2} \cdot P(\tau = j) \\ & = P(|S_{[n\delta]+1}| \geq n^{1/2}(\epsilon - (2\delta)^{1/2})) + \frac{1}{2} P(\tau \leq [n\delta]) \\ & = P(|S_{[n\delta]+1}| \geq n^{1/2}(\epsilon - (2\delta)^{1/2})) + \frac{1}{2} P\left(\max_{0 \leq j \leq [n\delta]+1} |S_j| > n^{1/2}\epsilon\right), \end{aligned}$$

so

$$P\left(\max_{0 \leq j \leq [n\delta]+1} |S_j| > n^{1/2}\epsilon\right) \leq 2P(|S_{[n\delta]+1}| \geq n^{1/2}(\epsilon - (2\delta)^{1/2})).$$

Now using (3) with $\lambda = (\epsilon - (2\delta)^{1/2})\delta^{-1/2}$,

$$\limsup_{n \rightarrow \infty} P(|S_{[n\delta]+1}| \geq (\epsilon - (2\delta)^{1/2})\delta^{-1/2}(n\delta)^{1/2}) \leq (\epsilon - (2\delta)^{1/2})^{-3} \delta^{3/2} E|Z|^3,$$

hence

$$\limsup_{n \rightarrow \infty} P\left(\max_{0 \leq j \leq [n\delta]+1} |S_j| > n^{1/2}\epsilon\right) \leq 2(\epsilon - (2\delta)^{1/2})^{-3} \delta^{3/2} E|Z|^3.$$

Dividing both sides by δ and then taking $\delta \downarrow 0$ we obtain the claim. \square

We prove one more result that we use to prove Donsker's theorem.¹¹

Lemma 9. For $T > 0$ and $\epsilon > 0$,

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} P \left(\max_{0 \leq k \leq [nT] + 1} \max_{1 \leq j \leq [n\delta] + 1} |S_{j+k} - S_k| > n^{1/2} \epsilon \right) = 0.$$

Proof. For $0 < \delta \leq T$, let $m = \lceil T/\delta \rceil$, so $T/m < \delta \leq T/(m-1)$. Then

$$\lim_{n \rightarrow \infty} \frac{[nT] + 1}{[n\delta] + 1} = \frac{T}{\delta} < m,$$

so for all $n \geq n_\delta$ it is the case that $[nT] + 1 < ([n\delta] + 1)m$. Suppose that $\omega \in \Omega$ is such that there are $1 \leq j \leq [n\delta] + 1$ and $0 \leq k \leq [nT] + 1$ satisfying

$$|S_{j+k}(\omega) - S_k(\omega)| > n^{1/2} \epsilon,$$

and then let $p = \lfloor k/([n\delta] + 1) \rfloor$, which satisfies $0 \leq p \leq m-1$ and

$$([n\delta] + 1)p \leq k < ([n\delta] + 1)(p+1).$$

Because $1 \leq j \leq [n\delta] + 1$, either

$$([n\delta] + 1)p < k + j \leq ([n\delta] + 1)(p+1)$$

or

$$([n\delta] + 1)(p+1) < k + j < ([n\delta] + 1)(p+2).$$

We separate the first case into the cases

$$|S_k(\omega) - S_{([n\delta] + 1)p}(\omega)| > \frac{1}{2} n^{1/2} \epsilon$$

and

$$|S_{j+k}(\omega) - S_{([n\delta] + 1)p}(\omega)| > \frac{1}{2} n^{1/2} \epsilon,$$

and we separate the second case into the cases

$$|S_k - S_{([n\delta] + 1)p}(\omega)| > \frac{1}{3} n^{1/2} \epsilon,$$

and

$$|S_{([n\delta] + 1)p}(\omega) - S_{([n\delta] + 1)(p+1)}(\omega)| > \frac{1}{3} n^{1/2} \epsilon,$$

and

$$|S_{([n\delta] + 1)(p+1)}(\omega) - S_{([n\delta] + 1)(p+2)}(\omega)| > \frac{1}{3} n^{1/2} \epsilon.$$

¹¹Ioannis Karatzas and Steven E. Shreve, *Brownian Motion and Stochastic Calculus*, second ed., p. 69, Lemma 4.19.

It follows that¹²

$$\begin{aligned} & \left\{ \max_{1 \leq j \leq [n\delta]+1} \max_{0 \leq k \leq [nT]+1} |S_{j+k} - S_k| > n^{1/2}\epsilon \right\} \\ & \subset \bigcup_{p=0}^{m-1} \left\{ \max_{1 \leq j \leq [n\delta]+1} |S_{j+([n\delta]+1)p} - S_{([n\delta]+1)p}| > \frac{1}{3}n^{1/2}\epsilon \right\}. \end{aligned}$$

For $0 \leq p \leq m-1$,

$$\begin{aligned} & P \left(\max_{1 \leq j \leq [n\delta]+1} |S_{j+([n\delta]+1)p} - S_{([n\delta]+1)p}| > \frac{1}{3}n^{1/2}\epsilon \right) \\ & \leq P \left(\max_{1 \leq j \leq [n\delta]+1} |S_j| > \frac{1}{3}n^{1/2}\epsilon \right), \end{aligned}$$

so

$$\begin{aligned} & P \left\{ \max_{1 \leq j \leq [n\delta]+1} \max_{0 \leq k \leq [nT]+1} |S_{j+k} - S_k| > n^{1/2}\epsilon \right\} \\ & \leq \sum_{p=0}^{m-1} P \left(\max_{1 \leq j \leq [n\delta]+1} |S_j| > \frac{1}{3}n^{1/2}\epsilon \right) \\ & = mP \left(\max_{1 \leq j \leq [n\delta]+1} |S_j| > \frac{1}{3}n^{1/2}\epsilon \right). \end{aligned}$$

Lemma 8 tells us

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{\delta} P \left(\max_{1 \leq j \leq [n\delta]+1} |S_j| > \frac{1}{3}n^{1/2}\epsilon \right) = 0,$$

and because $m \leq \frac{T}{\delta} + 1 = \frac{T+\delta}{\delta}$,

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} P \left\{ \max_{1 \leq j \leq [n\delta]+1} \max_{0 \leq k \leq [nT]+1} |S_{j+k} - S_k| > n^{1/2}\epsilon \right\} = 0,$$

proving the claim. \square

In the following, $P_n \in \mathcal{P}_E$ denotes the pushforward measure of P by $X^{(n)}$, for $X^{(n)}$ defined in (1). We now prove **Donsker's theorem**.¹³

Theorem 10 (Donsker's theorem). $P_n \rightarrow W$.

Proof. We shall use Theorem 3 to prove that $\Gamma = \{P_n : n \geq 1\}$ is relatively compact in \mathcal{P}_E . For $n \geq 1$,

$$P_n(f \in E : |f(0)| = 0) = P(\omega \in \Omega : |X_0^{(n)}(\omega)| = 0) = 1,$$

¹²This should be worked out more carefully. In Karatzas and Shreve, there is $m+1$ where I have m .

¹³Ioannis Karatzas and Steven E. Shreve, *Brownian Motion and Stochastic Calculus*, second ed., p. 70, Theorem 4.20.

thus the first condition of Theorem 3 is satisfied with $M_\epsilon = 0$. For the second condition of Theorem 3 to be satisfied it suffices that for each $\epsilon > 0$,

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} P \left(\sup_{0 \leq s, t \leq 1, |s-t| \leq \delta} |X^{(n)}(s) - X^{(n)}(t)| > \epsilon \right) = 0.$$

Now,

$$P \left(\sup_{0 \leq s, t \leq 1, |s-t| \leq \delta} |X_s^{(n)} - X_t^{(n)}| > \epsilon \right) = P \left(\sup_{0 \leq s, t \leq n, |s-t| \leq n\delta} |Y_s - Y_t| > n^{1/2}\epsilon \right).$$

Also,

$$\begin{aligned} \sup_{0 \leq s, t \leq n, |s-t| \leq n\delta} |Y_s - Y_t| &\leq \sup_{0 \leq s, t \leq n, |s-t| \leq n\delta} |Y - s - Y_t| \\ &\leq \max_{1 \leq j \leq [n\delta]+1} \max_{0 \leq k \leq n+1} |S_{j+k} - S_k|, \end{aligned}$$

so applying Lemma 9,

$$\begin{aligned} &\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} P \left(\sup_{0 \leq s, t \leq 1, |s-t| \leq \delta} |X_s^{(n)} - X_t^{(n)}| > \epsilon \right) \\ &\leq \lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} P \left(\max_{1 \leq j \leq [n\delta]+1} \max_{0 \leq k \leq n+1} |S_{j+k} - S_k| > n^{1/2}\epsilon \right) \\ &\rightarrow 0, \end{aligned}$$

from which we get that Γ is tight in \mathcal{P}_E . □