

In these notes, we give the definition of the well-known **Wiener measure** defined on the separable Banach space  $(C[0, 1], \|\cdot\|_\infty)$ , and give the proof that it indeed exists. Here is the outline:

- The Wiener measure will be defined as any Borel probability measure on  $(C[0, 1], \|\cdot\|_\infty)$  with certain prescribed finite-dimensional distributions. Since the collection of finite-dimensional subsets of  $C[0, 1]$  is a separating class for its Borel  $\sigma$ -algebra (see Example 1.3, [3]), we see immediately that if a Wiener measure exists, it is unique.
- However, the above definition is axiomatic, rather than constructive; the existence of a probability measure on  $(C[0, 1], \|\cdot\|_\infty)$  satisfying the defining properties of a Wiener measure must therefore be proved.
- To present this proof, we will start by showing that the “scaling limits” of the finite-dimensional distributions of the sequence of probability measures induced on  $(C[0, 1], \|\cdot\|_\infty)$  by linearly interpolated random walks with I.I.D. steps possess exactly the defining properties (of the finite-dimensional distributions) of the Wiener measure.
- We then show that the sequence of probability measures induced on  $(C[0, 1], \|\cdot\|_\infty)$  by linearly interpolated random walks with I.I.D. steps is tight, and thus admits a weakly convergent subsequence.
- The Wiener measure is, by definition, the limit of this weakly convergent subsequence of probability measures on  $(C[0, 1], \|\cdot\|_\infty)$

## 1 Equivalence of $(C[0, 1], \|\cdot\|_\infty)$ -valued random variables and $\mathbb{R}$ -valued stochastic processes indexed by $[0, 1]$ with continuous sample paths

**Proposition 1.1** (The “one-dimensional subsets” of  $C[0, 1]$  generate its Borel  $\sigma$ -algebra)

Let  $(C[0, 1], \|\cdot\|_\infty)$  be the metric space of continuous  $\mathbb{R}$ -valued functions defined on the closed unit interval equipped with the supremum norm. For each  $t \in [0, 1]$ , let  $\text{ev}_t : C[0, 1] \rightarrow \mathbb{R} : x \mapsto x(t)$ . Define:

$$\mathcal{S} := \left\{ \text{ev}_t^{-1}(H) \subset C[0, 1] \mid \begin{array}{l} t \in [0, 1] \\ H \in \mathcal{O} \end{array} \right\} \subset \mathcal{P}(C[0, 1]).$$

Then,  $\mathcal{S}$  generates the Borel  $\sigma$ -algebra  $\mathcal{B} := \mathcal{B}(C[0, 1], \|\cdot\|_\infty)$  of the metric space  $(C[0, 1], \|\cdot\|_\infty)$ ; in other words,

$$\sigma(\mathcal{S}) = \mathcal{B}.$$

**PROOF** First, note that  $\sigma(\mathcal{S}) \subset \mathcal{B}$ . Indeed, recall that, for each  $t \in [0, 1]$ ,  $\text{ev}_t : C[0, 1] \rightarrow \mathbb{R}$  is continuous, hence  $(\mathcal{B}, \mathcal{O})$ -measurable, by Corollary B.4. In particular,  $\text{ev}_t^{-1}(H) \in \mathcal{B}$ , for each  $t \in [0, 1]$  and  $H \in \mathcal{O}$ . Thus,  $\mathcal{S} \subset \mathcal{B}$ ; hence,  $\sigma(\mathcal{S}) \subset \mathcal{B}$ .

It remains to establish the reverse inclusion. To this end, first observe that, for each  $x \in C[0, 1]$  and each  $\varepsilon > 0$ , we have

$$\overline{B(x, \varepsilon)} = \bigcap_{r \in \mathbb{Q} \cap [0, 1]} \left\{ y \in C[0, 1] \mid |y(r) - x(r)| \leq \varepsilon \right\} = \bigcap_{r \in \mathbb{Q} \cap [0, 1]} \text{ev}_r^{-1}([x(r) - \varepsilon, x(r) + \varepsilon]),$$

which shows that  $\sigma(\mathcal{S})$  contains all the closed balls in  $C[0, 1]$ . On the other hand, recall that, in any metric space, every open ball can be expressed as a countable union of closed balls; indeed, for any  $y$  in the given metric space, and any  $\delta > 0$ , we have:

$$B(y, \delta) = \bigcup_{n \in \mathbb{N}} \overline{B\left(y, \delta - \frac{1}{n}\right)}.$$

We thus see that  $\sigma(\mathcal{S})$  contains all the open balls in  $C[0, 1]$ . By the separability of  $C[0, 1]$  and Theorem C.1, we see that every open subset of  $C[0, 1]$  can be expressed as a countable union of open balls. Hence,  $\sigma(\mathcal{S})$  in fact contains all the open subsets of  $C[0, 1]$ , which immediately yields  $\mathcal{B} \subset \sigma(\mathcal{S})$ . This proves  $\sigma(\mathcal{S}) = \mathcal{B}$ .  $\square$

**Theorem 1.2**

*Suppose:*

- $(\Omega, \mathcal{A})$  is a measurable space, and  $\mathcal{O}$  is the Borel  $\sigma$ -algebra of  $\mathbb{R}$  (equipped with usual Euclidean metric).
- $(C[0, 1], \|\cdot\|_\infty)$  is the metric space of continuous  $\mathbb{R}$ -valued functions defined on the compact unit interval equipped with the supremum norm, and  $\mathcal{B} := \mathcal{B}(C[0, 1], \|\cdot\|_\infty)$  is its Borel  $\sigma$ -algebra.
- $X : \Omega \longrightarrow C[0, 1]$  is a function with domain  $\Omega$  and codomain  $C[0, 1]$ , but otherwise arbitrary.
- For each  $t \in [0, 1]$ , let  $\text{ev}_t : C[0, 1] \longrightarrow \mathbb{R} : x \longmapsto x(t)$ .
- For each  $t \in [0, 1]$ , let  $X_t := \text{ev}_t \circ X$ . In other words,  $X_t : \Omega \longrightarrow \mathbb{R} : \omega \longmapsto \text{ev}_t(X(\omega)) = X(\omega)(t)$ .

Then,  $X$  is  $(\mathcal{A}, \mathcal{B})$ -measurable if and only if, for each  $t \in [0, 1]$ ,  $X_t$  is  $(\mathcal{A}, \mathcal{O})$ -measurable.

PROOF

$(\implies)$

It is trivial to see that, for each  $t \in [0, 1]$ ,  $\text{ev}_t : (C[0, 1], \|\cdot\|_\infty) \longrightarrow (\mathbb{R}, |\cdot|) : x \longmapsto x(t)$  is continuous. Recall that continuous maps are necessarily Borel measurable; see Corollary B.4. Hence,  $\text{ev}_t : (C[0, 1], \|\cdot\|_\infty) \longrightarrow (\mathbb{R}, |\cdot|)$  is  $(\mathcal{B}, \mathcal{O})$ -measurable, for each  $t \in [0, 1]$ . Now, suppose  $X : \Omega \longrightarrow C[0, 1]$  is  $(\mathcal{A}, \mathcal{B})$ -measurable. Then, for each  $t \in [0, 1]$ , the composition  $X_t := \text{ev}_t \circ X$  is  $(\mathcal{A}, \mathcal{O})$ -measurable, as required.

$(\impliedby)$

Suppose that, for each  $t \in [0, 1]$ ,  $X_t := \text{ev}_t \circ X$  is  $(\mathcal{A}, \mathcal{O})$ -measurable. We seek to establish that  $X : (\Omega, \mathcal{A}) \longrightarrow (C[0, 1], \mathcal{B})$  is  $(\mathcal{A}, \mathcal{B})$ -measurable. To this end, let

$$\mathcal{S} := \left\{ \text{ev}_t^{-1}(H) \subset C[0, 1] \mid \begin{array}{l} t \in [0, 1] \\ H \in \mathcal{O} \end{array} \right\} \subset \mathcal{P}(C[0, 1]).$$

Then, note that the  $(\mathcal{A}, \mathcal{B})$ -measurability of  $X$  follows immediately from Theorem B.3, Proposition 1.1, and the following

**Claim:**  $X^{-1}(\mathcal{S}) \subset \mathcal{A}$ .

Proof of Claim: Every set in  $\mathcal{S}$  has the form  $\text{ev}_t^{-1}(H)$ , for some  $t \in [0, 1]$  and some  $H \in \mathcal{O}$ . Note that

$$X^{-1}(\text{ev}_t^{-1}(H)) = (\text{ev}_t \circ X)^{-1}(H) = X_t^{-1}(H) \in \mathcal{A},$$

where the last containment follows immediately from the  $(\mathcal{A}, \mathcal{O})$ -measurability hypothesis on  $X_t$ , for each  $t \in [0, 1]$ . This shows that  $X^{-1}(\mathcal{S}) \subset \mathcal{A}$  and completes the proof of the Claim.

The proof of the Theorem is now complete. □

**Theorem 1.3**

*Suppose:*

- $(\Omega, \mathcal{A})$  is a measurable space, and  $\mathcal{O}$  is the Borel  $\sigma$ -algebra of  $\mathbb{R}$  (equipped with usual Euclidean metric).
- $(C[0, 1], \|\cdot\|_\infty)$  is the metric space of continuous  $\mathbb{R}$ -valued functions defined on the closed unit interval equipped with the supremum norm, and  $\mathcal{B} := \mathcal{B}(C[0, 1], \|\cdot\|_\infty)$  is its Borel  $\sigma$ -algebra.
- $X : \Omega \longrightarrow C[0, 1]$  is a function with domain  $\Omega$  and codomain  $C[0, 1]$ , but otherwise arbitrary.

- For each  $t \in [0, 1]$ , let  $\text{ev}_t : C[0, 1] \rightarrow \mathbb{R} : x \mapsto x(t)$ .
- For each  $t \in [0, 1]$ , let  $X_t := \text{ev}_t \circ X$ . In other words,  $X_t : \Omega \rightarrow \mathbb{R} : \omega \mapsto \text{ev}_t(X(\omega)) = X(\omega)(t)$ .

Then, the following are equivalent:

- (i)  $X$  is a  $(C[0, 1], \|\cdot\|_\infty)$ -valued random variable (in other words,  $X$  is  $(\mathcal{A}, \mathcal{B})$ -measurable).
- (ii) For each  $t \in [0, 1]$ ,  $X_t$  is an  $\mathbb{R}$ -valued random variable (in other words, each  $X_t$  is  $(\mathcal{A}, \mathcal{O})$ -measurable).
- (iii)  $\{X_t : \Omega \rightarrow \mathbb{R}\}_{t \in [0, 1]}$  is a stochastic process indexed by the closed unit interval defined on the probability space  $(\Omega, \mathcal{A}, \mu)$  with state space  $\mathbb{R}$  and continuous sample paths.

PROOF The equivalence of (i) and (ii) is immediate by the preceding Theorem. The equivalence of (ii) and (iii) is immediate by the definition of stochastic processes.  $\square$

## 2 Scaling limits of finite-dimensional distributions of linearly interpolated random walks are multivariate Gaussian

### Definition 2.1

- A **random walk on  $\mathbb{R}$**  is a sequence  $\{S_n : \Omega \rightarrow \mathbb{R}\}_{n \in \mathbb{N}}$  of  $\mathbb{R}$ -valued random variables defined on a probability space  $\Omega$  of the following form: There exists a sequence  $\xi_1, \xi_2, \dots : \Omega \rightarrow \mathbb{R}$  of independent and identically distributed  $\mathbb{R}$ -valued random variables such that

$$S_n = \sum_{i=1}^n \xi_i, \quad \text{for each } n \in \mathbb{N}.$$

- Let  $\left\{ S_n = \sum_{i=1}^n \xi_i : \Omega \rightarrow \mathbb{R} \right\}_{n \in \mathbb{N}}$  be a random walk on  $\mathbb{R}$ . Let  $S_0 : \Omega \rightarrow \mathbb{R} : \omega \mapsto 0$  be the (constant) random variable defined on  $\Omega$  with constant value  $0 \in \mathbb{R}$ .

For each  $n \in \mathbb{N}$ , the **piecewise linear equi-spaced lattice interpolation of  $\{S_0, S_1, S_2, \dots, S_n\}$  over the unit interval  $[0, 1]$**  is the function  $\bar{S}^{(n)} : \Omega \rightarrow C[0, 1]$  defined as follows:

$$\bar{S}^{(n)}(\omega)(t) := \frac{1}{\sigma \cdot \sqrt{n}} \left\{ S_{i-1}(\omega) + n \left( t - \frac{i-1}{n} \right) \xi_i(\omega) \right\}, \quad \text{for each } \omega \in \Omega, \quad t \in \left[ \frac{i-1}{n}, \frac{i}{n} \right], \quad i = 1, 2, 3, \dots, n.$$

- For each  $n \in \mathbb{N}$  and each  $t \in [0, 1]$ , define  $\bar{S}_t^{(n)} : \Omega \rightarrow \mathbb{R}$  as follows:

$$\bar{S}_t^{(n)}(\omega) := \bar{S}^{(n)}(\omega)(t), \quad \text{for each } \omega \in \Omega.$$

### Theorem 2.2

Let  $\left\{ S_n = \sum_{i=1}^n \xi_i : \Omega \rightarrow \mathbb{R} \right\}_{n \in \mathbb{N}}$  be a random walk on  $\mathbb{R}$ , and let  $\bar{S}^{(n)} : \Omega \rightarrow C[0, 1]$  and  $\bar{S}_t^{(n)} : \Omega \rightarrow \mathbb{R}$  be the associated functions as defined in Definition 2.1. Let  $S_0 : \Omega \rightarrow \mathbb{R} : \omega \mapsto 0$  be the (constant)  $\mathbb{R}$ -valued random variable defined on  $\Omega$  with constant value  $0 \in \mathbb{R}$ . Then, the following statements are true:

- (i) For each  $\omega \in \Omega$  and each  $n \in \mathbb{N}$ ,

$$\bar{S}^{(n)}(\omega) \left( \frac{i}{n} \right) = \frac{1}{\sigma \cdot \sqrt{n}} \cdot S_i(\omega), \quad \text{for } i = 0, 1, 2, \dots, n.$$

(ii) For each  $\omega \in \Omega$  and each  $n \in \mathbb{N}$ ,

$$\bar{S}^{(n)}(\omega)(t) \text{ is the linear interpolation from } \frac{1}{\sigma \cdot \sqrt{n}} S_{i-1}(\omega) \text{ to } \frac{1}{\sigma \cdot \sqrt{n}} S_i(\omega) \text{ over } t \in \left[\frac{i-1}{n}, \frac{i}{n}\right],$$

where  $i = 1, 2, \dots, n$ . In particular, for each  $\omega \in \Omega$ ,  $\bar{S}^{(n)}(\omega)$  is a continuous  $\mathbb{R}$ -valued function defined on  $[0, 1]$ .

(iii) For each  $n \in \mathbb{N}$  and each  $t \in [0, 1]$ , the function  $\bar{S}_t^{(n)} : \Omega \rightarrow \mathbb{R}$  is an  $\mathbb{R}$ -valued random variable defined on the probability space  $\Omega$ . Hence, for each  $n \in \mathbb{N}$ ,  $\left\{ \bar{S}_t^{(n)} \right\}_{t \in [0, 1]}$  is a stochastic process indexed by  $[0, 1]$  with state space  $\mathbb{R}$  and continuous sample paths.

(iv) For each  $n \in \mathbb{N}$ , the function  $\bar{S}^{(n)} : \Omega \rightarrow C[0, 1]$  is a  $C[0, 1]$ -valued random variable defined on the probability space  $\Omega$ .

PROOF

(i) Obvious.

(ii) Obvious.

(iii) For each  $n \in \mathbb{N}$  and each  $t \in [0, 1]$ , the function  $\bar{S}_t^{(n)}$  is an  $\mathbb{R}$ -valued random variable defined on  $\Omega$  because it is a linear combination, with coefficients in  $\mathbb{R}$ , of the  $\mathbb{R}$ -valued random variables  $\xi_1, \xi_2, \dots, \xi_n$ . Thus,  $\left\{ \bar{S}_t^{(n)} \right\}_{t \in [0, 1]}$  is a stochastic process indexed by  $[0, 1]$  with state space  $\mathbb{R}$ . By (ii), this stochastic process has continuous sample paths.

(iv) Immediate by (iii) and Theorem 1.3.

□

## Theorem 2.3

- Let  $\xi_1, \xi_2, \dots : \Omega \rightarrow \mathbb{R}$  be a sequence of independent and identically distributed  $\mathbb{R}$ -valued random variables defined on the probability space  $(\Omega, \mathcal{A}, \mu)$ , with expectation value zero and common finite variance  $\sigma^2 > 0$ .
- Define the random variables:

$$\begin{cases} S_0 : \Omega \rightarrow \mathbb{R} : \omega \mapsto 0, & \text{and} \\ S_n : \Omega \rightarrow \mathbb{R} : \omega \mapsto \sum_{i=1}^n \xi_i(\omega), & \text{for each } n \in \mathbb{N}. \end{cases}$$

- For each  $n \in \mathbb{N}$ , define  $X^{(n)} : \Omega \rightarrow C[0, 1]$  as follows:

$$X^{(n)}(\omega)(t) := \frac{1}{\sigma \cdot \sqrt{n}} \left\{ S_{i-1}(\omega) + n \left( t - \frac{i-1}{n} \right) \xi_i(\omega) \right\}, \text{ for each } \omega \in \Omega, t \in \left[ \frac{i-1}{n}, \frac{i}{n} \right], i = 1, 2, 3, \dots, n.$$

- For each  $n \in \mathbb{N}$  and each  $t \in [0, 1]$ , define  $X_t^{(n)} : \Omega \rightarrow \mathbb{R}$  as follows:

$$X_t^{(n)}(\omega) := X^{(n)}(\omega)(t), \text{ for each } \omega \in \Omega.$$

Then, the following statements are true:

(i) For any  $0 \leq t_0 < t_1 < t_2 < \cdots < t_k \leq 1$ ,

$$\left( X_{t_1}^{(n)} - X_{t_0}^{(n)}, \dots, X_{t_k}^{(n)} - X_{t_{k-1}}^{(n)} \right) \xrightarrow{d} N\left(\mu = \mathbf{0}, \Sigma = \text{diag}(t_1 - t_0, \dots, t_k - t_{k-1})\right), \text{ as } n \rightarrow \infty.$$

(ii) For any pairwise distinct  $0 \leq t_1, t_2, \dots, t_k \leq 1$ ,

$$\left( X_{t_1}^{(n)}, X_{t_2}^{(n)}, \dots, X_{t_k}^{(n)} \right) \xrightarrow{d} N\left(\mu = \mathbf{0}, \Sigma = \left[ \min\{t_i, t_j\} \right]_{1 \leq i, j \leq k}\right), \text{ as } n \rightarrow \infty.$$

**PROOF** First, note that, by Theorem 2.2, for each  $n \in \mathbb{N}$ ,  $X^{(n)} : \Omega \rightarrow C[0, 1]$  is a  $C[0, 1]$ -valued random variable defined on  $\Omega$ .

(i) For each  $\omega \in \Omega$ ,  $n \in \mathbb{N}$ , and  $t \in [0, 1]$ , we have

$$X_t^{(n)}(\omega) = \frac{1}{\sigma \cdot \sqrt{n}} \left\{ S_{[nt]}(\omega) + (nt - [nt]) \cdot \xi_{[nt]+1}(\omega) \right\},$$

where  $\lfloor \cdot \rfloor : \mathbb{R} \rightarrow \mathbb{Z}$ , defined by

$$\lfloor x \rfloor := \max \left\{ k \in \mathbb{Z} \mid k \leq x \right\}, \text{ for each } x \in \mathbb{R},$$

is the round-down function. We next state three Claims, whose proofs will be given below. We note that the desired conclusion follows readily from Claim 3 and the Cramér-Wold Theorem (Theorem 1.9(iii), p.56, [5]); hence the present proof is complete once we establish the three Claims below.

**Claim 1:** If  $\{a_n\}_{n \in \mathbb{N}}$  is a sequence of non-negative integers and  $\{b_n\}_{n \in \mathbb{N}} \subset \mathbb{N}$  a sequence of positive integers satisfying:

$$a_n < b_n, \text{ for sufficiently large } n \in \mathbb{N}, \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{b_n - a_n}{n} = c > 0,$$

then

$$\frac{1}{\sigma \cdot \sqrt{n}} \cdot \sum_{i=1+a_n}^{b_n} \xi_i \xrightarrow{d} \sqrt{c} \cdot Z, \text{ where } Z \sim N(0, 1).$$

**Claim 2:** For each fixed  $t \in [0, 1]$ ,

$$W(t)_n := \frac{1}{\sigma \cdot \sqrt{n}} \cdot (nt - [nt]) \cdot \xi_{[nt]+1} \xrightarrow{d} 0.$$

**Claim 3:** For  $0 \leq t_0 < t_1 < t_2 < \cdots < t_k \leq 1$ , and arbitrary  $c_1, c_2, \dots, c_k \in \mathbb{R}$ ,

$$\sum_{i=1}^k c_i \left( X_{t_i}^{(n)} - X_{t_{i-1}}^{(n)} \right) \xrightarrow{d} N\left(0, \sum_{i=1}^k c_i^2 (t_i - t_{i-1})\right), \text{ as } n \rightarrow \infty.$$

Proof of Claim 1: Note that, for sufficiently large  $n \in \mathbb{N}$ , we may write

$$\frac{1}{\sigma \cdot \sqrt{n}} \cdot \sum_{i=1+a_n}^{b_n} \xi_i = \frac{\sqrt{b_n - a_n}}{\sqrt{n}} \cdot \left( \frac{1}{\sigma \cdot \sqrt{b_n - a_n}} \cdot \sum_{i=1+a_n}^{b_n} \xi_i \right).$$

Since, by hypothesis, that

$$\lim_{n \rightarrow \infty} \frac{b_n - a_n}{n} = c > 0,$$

Claim 1 follows by Slutsky's Theorem (Example 6, p.40, [4]), once we establish the following:

$$\frac{1}{\sigma \cdot \sqrt{b_n - a_n}} \cdot \sum_{i=1+a_n}^{b_n} \xi_i \xrightarrow{d} N(0, 1), \quad \text{as } n \rightarrow \infty.$$

We establish the above convergence by invoking the Lindeberg Central Limit Theorem (Theorem 1.15, §1.5.5, p.67, [5]). In the present context, the Lindeberg Condition is the following:

$$\lim_{n \rightarrow \infty} \frac{1}{B_n^2} \cdot E \left[ \sum_{i=1+a_n}^{b_n} \xi_i^2 \cdot I_{\{|\xi_i| \geq \varepsilon S_n\}} \right] = 0, \quad \text{for each } \varepsilon > 0,$$

where

$$B_n^2 := \text{Var} \left[ \sum_{i=1+a_n}^{b_n} \xi_i \right] = (b_n - a_n) \sigma^2 > 0.$$

The last equality used the hypothesis that  $\xi_1, \xi_2, \dots$  are independent and identically distributed with common finite variance  $0 < \sigma^2 < \infty$ . Hence, for each  $\varepsilon > 0$ ,

$$\begin{aligned} \frac{1}{B_n^2} \cdot E \left[ \sum_{i=1+a_n}^{b_n} \xi_i^2 \cdot I_{\{|\xi_i| \geq \varepsilon B_n\}} \right] &= \frac{1}{(b_n - a_n) \sigma^2} \cdot (b_n - a_n) \cdot E \left[ \xi_1^2 \cdot I_{\{|\xi_1| \geq \varepsilon \sigma \sqrt{b_n - a_n}\}} \right] \\ &= \frac{1}{\sigma^2} \cdot E \left[ \xi_1^2 \cdot I_{\{|\xi_1| / \varepsilon \sigma \geq \sqrt{b_n - a_n}\}} \right] \rightarrow 0, \quad \text{as } n \rightarrow \infty, \end{aligned}$$

since  $\lim_{n \rightarrow \infty} \sqrt{b_n - a_n} = \infty$  and  $\sigma^2 = E[\xi_1^2]$  is finite. This verifies that the Lindeberg Condition indeed holds in the present context, and completes the proof of Claim 1.

Proof of Claim 2: First, note that  $E[W(t)_n] = 0$ . We now argue that  $W(t)_n \xrightarrow{p} 0$ . To this end, let  $\varepsilon > 0$  be given. Then,

$$\begin{aligned} \varepsilon^2 \cdot P(|W(t)_n| \geq \varepsilon) &\leq E[W(t)_n^2 \cdot I_{\{|W(t)_n| \geq \varepsilon\}}] \\ &\leq E[W(t)_n^2] = \text{Var}(W(t)_n) = \text{Var} \left[ \frac{1}{\sigma \cdot \sqrt{n}} \cdot (nt - \lfloor nt \rfloor) \cdot \xi_{\lfloor nt \rfloor + 1} \right] \\ &= \frac{1}{n \cdot \sigma^2} \cdot (nt - \lfloor nt \rfloor)^2 \cdot \text{Var}(\xi_{\lfloor nt \rfloor + 1}) = \frac{1}{n \cdot \sigma^2} \cdot (nt - \lfloor nt \rfloor)^2 \cdot \sigma^2 \\ &\leq \frac{1}{n}, \end{aligned}$$

which implies

$$\lim_{n \rightarrow \infty} P(|W(t)_n| \geq \varepsilon) = 0, \quad \text{for each } \varepsilon > 0,$$

i.e.  $W(t)_n \xrightarrow{p} 0$ , as  $n \rightarrow \infty$  (Definition 2, Chapter 1, [4]), which is equivalent to  $W(t)_n \xrightarrow{d} 0$ , as  $n \rightarrow \infty$  (by Theorem 1, Chapter 1 and Theorem 2, Chapter 2, [4]). This proves Claim 2.

Proof of Claim 3: Let  $0 \leq t_0 < t_1 < t_2 < \dots < t_k \leq 1$ , and  $c_1, c_2, \dots, c_k \in \mathbb{R}$  be arbitrary. Observe that:

$$\begin{aligned}
 & \sum_{i=1}^k c_i \left( X_{t_i}^{(n)} - X_{t_{i-1}}^{(n)} \right) \\
 &= \sum_{i=1}^k \frac{c_i}{\sigma \cdot \sqrt{n}} \left\{ S_{\lfloor nt_i \rfloor} - S_{\lfloor nt_{i-1} \rfloor} \right\} + \sum_{i=1}^k \frac{c_i}{\sigma \cdot \sqrt{n}} \left\{ \left( nt_i - \lfloor nt_i \rfloor \right) \cdot \xi_{\lfloor nt_i \rfloor + 1} - \left( nt_{i-1} - \lfloor nt_{i-1} \rfloor \right) \cdot \xi_{\lfloor nt_{i-1} \rfloor + 1} \right\} \\
 &= \sum_{i=1}^k \frac{c_i}{\sigma \cdot \sqrt{n}} \left\{ \sum_{j=1+\lfloor nt_{i-1} \rfloor}^{\lfloor nt_i \rfloor} \xi_j \right\} + \sum_{i=1}^k c_i \left\{ W(t_i)_n - W(t_{i-1})_n \right\} \\
 &= \sum_{i=1}^k c_i Y_i^{(n)} + \sum_{i=1}^k c_i \left\{ W(t_i)_n - W(t_{i-1})_n \right\}
 \end{aligned}$$

By Claim 2 and Slutsky's Theorem (Corollary, p.40, [4]),

$$\sum_{i=1}^k c_i \left\{ W(t_i)_n - W(t_{i-1})_n \right\} \xrightarrow{d} 0, \text{ as } n \rightarrow \infty. \quad (2.1)$$

Next, note that since  $\xi_1, \xi_2, \xi_3, \dots$  are independent, we see that, for each fixed  $n \in \mathbb{N}$ ,

$$Y_i^{(n)} := \frac{1}{\sigma \cdot \sqrt{n}} \cdot \sum_{j=1+\lfloor nt_{i-1} \rfloor}^{\lfloor nt_i \rfloor} \xi_j, \quad i = 1, 2, 3, \dots, k,$$

are independent. Now, since  $0 \leq t_{i-1} < t_i \leq 1$ , it follows that  $\lfloor nt_{i-1} \rfloor < \lfloor nt_i \rfloor$  for sufficiently large  $n \in \mathbb{N}$ . In addition,

$$\begin{aligned}
 \frac{\lfloor nt_i \rfloor - \lfloor nt_{i-1} \rfloor}{n} &= \frac{\lfloor nt_i \rfloor}{n} - \frac{\lfloor nt_{i-1} \rfloor}{n} = \left( \frac{nt_i}{n} + \frac{\lfloor nt_i \rfloor - nt_i}{n} \right) - \left( \frac{nt_{i-1}}{n} + \frac{\lfloor nt_{i-1} \rfloor - nt_{i-1}}{n} \right) \\
 &= t_i - t_{i-1} + \frac{\lfloor nt_i \rfloor - nt_i}{n} - \frac{\lfloor nt_{i-1} \rfloor - nt_{i-1}}{n},
 \end{aligned}$$

which implies

$$\left| \frac{\lfloor nt_i \rfloor - \lfloor nt_{i-1} \rfloor}{n} - (t_i - t_{i-1}) \right| = \left| \frac{\lfloor nt_i \rfloor - nt_i}{n} - \frac{\lfloor nt_{i-1} \rfloor - nt_{i-1}}{n} \right| \leq \frac{2}{n} \rightarrow 0, \text{ as } n \rightarrow \infty.$$

Thus,

$$\lim_{n \rightarrow \infty} \frac{\lfloor nt_i \rfloor - \lfloor nt_{i-1} \rfloor}{n} = t_i - t_{i-1} > 0.$$

Thus, by Claim 1, we see that, for each  $i = 1, 2, \dots, k$ ,

$$Y_i^{(n)} := \frac{1}{\sigma \cdot \sqrt{n}} \cdot \sum_{j=1+\lfloor nt_{i-1} \rfloor}^{\lfloor nt_i \rfloor} \xi_j \xrightarrow{d} \sqrt{t_i - t_{i-1}} \cdot N(0, 1) = N\left(0, t_i - t_{i-1}\right), \text{ as } n \rightarrow \infty. \quad (2.2)$$

By (2.1), (2.2), Proposition A.1, and Slutsky's Theorem (Corollary, p.40, [4]), we now see that

$$\sum_{i=1}^k c_i \left( X_{t_i}^{(n)} - X_{t_{i-1}}^{(n)} \right) = \sum_{i=1}^k c_i Y_i^{(n)} + \sum_{i=1}^k c_i \left\{ W(t_i)_n - W(t_{i-1})_n \right\} \xrightarrow{d} N\left(0, \sum_{i=1}^k c_i^2 (t_i - t_{i-1})\right).$$

This completes the proof of Claim 3.

- (ii) Let  $t_0 := 0$ , hence,  $X_{t_0}^{(n)} \equiv 0$  for each  $n \in \mathbb{N}$ . Without loss of generality, we may re-label  $t_1, t_2, \dots, t_k$  if necessary such that  $t_1 < t_2 < \dots < t_k$ . We thus have, for each  $n \in \mathbb{N}$ ,

$$\begin{pmatrix} X_{t_1}^{(n)} \\ X_{t_2}^{(n)} \\ \vdots \\ X_{t_k}^{(n)} \end{pmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ 1 & 1 & 0 & \cdots & \cdots & 0 & 0 \\ 1 & 1 & 1 & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 & 0 \\ 1 & 1 & 1 & \cdots & \cdots & 1 & 0 \\ 1 & 1 & 1 & \cdots & \cdots & 1 & 1 \end{bmatrix}}_T \cdot \begin{pmatrix} X_{t_1}^{(n)} \\ X_{t_2}^{(n)} - X_{t_1}^{(n)} \\ \vdots \\ X_{t_k}^{(n)} - X_{t_{k-1}}^{(n)} \end{pmatrix}.$$

By (i), we know that

$$\begin{pmatrix} X_{t_1}^{(n)} \\ X_{t_2}^{(n)} - X_{t_1}^{(n)} \\ \vdots \\ X_{t_k}^{(n)} - X_{t_{k-1}}^{(n)} \end{pmatrix} \xrightarrow{d} \begin{pmatrix} Z_{t_1} \\ Z_{t_2-t_1} \\ \vdots \\ Z_{t_k-t_{k-1}} \end{pmatrix} \sim N\left(\mu = \mathbf{0}, \Sigma = \text{diag}(t_1, t_2 - t_1, \dots, t_k - t_{k-1})\right), \text{ as } n \rightarrow \infty.$$

Since the map  $\mathbb{R}^k \rightarrow \mathbb{R}^k : x \mapsto T \cdot x$  is continuous, we see immediately by Slutsky's Theorem (Theorem 6(a), p.39, [4]) that

$$\begin{pmatrix} X_{t_1}^{(n)} \\ X_{t_2}^{(n)} \\ \vdots \\ X_{t_k}^{(n)} \end{pmatrix} \xrightarrow{d} T \cdot \begin{pmatrix} Z_{t_1} \\ Z_{t_2-t_1} \\ \vdots \\ Z_{t_k-t_{k-1}} \end{pmatrix}, \text{ as } n \rightarrow \infty.$$

Since the map  $\mathbb{R}^k \rightarrow \mathbb{R}^k : x \mapsto T \cdot x$  is an invertible linear automorphism on  $\mathbb{R}^k$ , we see that

$$L = \begin{pmatrix} L_{t_1} \\ L_{t_2} \\ \vdots \\ L_{t_k} \end{pmatrix} := T \cdot \begin{pmatrix} Z_{t_1} \\ Z_{t_2-t_1} \\ \vdots \\ Z_{t_k-t_{k-1}} \end{pmatrix}$$

is still an  $\mathbb{R}^k$ -valued Gaussian random variable, and it clearly has expectation value  $\mathbf{0} \in \mathbb{R}^k$ , since each of  $Z_{t_1}, Z_{t_2-t_1}, \dots, Z_{t_k-t_{k-1}}$  has expectation value  $0 \in \mathbb{R}$ . It remains only to compute the covariance matrix of the  $\mathbb{R}^k$ -valued Gaussian random variable  $L$ . To this end, consider  $1 \leq i \leq j \leq k$ , i.e.  $t_i \leq t_j$ . Then, using the



alternative notation  $Z_{t_1-t_0} := Z_{t_1}$ , we have

$$\begin{aligned}
 \text{Cov}(L_{t_i}, L_{t_j}) &= \text{Cov}(Z_{t_1} + Z_{t_2-t_1} + \cdots + Z_{t_i-t_{i-1}}, Z_{t_1} + Z_{t_2-t_1} + \cdots + Z_{t_j-t_{j-1}}) \\
 &= \text{Cov}\left(\sum_{a=1}^i Z_{t_a-t_{a-1}}, \sum_{b=1}^j Z_{t_b-t_{b-1}}\right) = \sum_{a=1}^i \sum_{b=1}^j \text{Cov}(Z_{t_a-t_{a-1}}, Z_{t_b-t_{b-1}}) \\
 &= \sum_{a=1}^i \text{Cov}(Z_{t_a-t_{a-1}}, Z_{t_a-t_{a-1}}) = \sum_{a=1}^i \text{Var}(Z_{t_a-t_{a-1}}) = \sum_{a=1}^i (t_a - t_{a-1}) \\
 &= (t_1 - t_0) + (t_2 - t_1) + \cdots + (t_{i-1} - t_{i-2}) + (t_i - t_{i-1}) \\
 &= t_i = \min\{t_i, t_j\},
 \end{aligned}$$

as required. □

### 3 A sufficient condition for the tightness of a sequence of linearly interpolated random walks

**Lemma 3.1** (Lemma, p.88, [3])

- Let  $\xi_1, \xi_2, \dots : \Omega \rightarrow \mathbb{R}$  be a sequence of  $\mathbb{R}$ -valued random variables defined on the probability space  $(\Omega, \mathcal{A}, \mu)$ , with expectation value zero and common finite variance  $\sigma^2 > 0$ .
- Define the random variables:

$$\begin{cases} S_0 & : \Omega \rightarrow \mathbb{R} : \omega \mapsto 0, & \text{and} \\ S_n & : \Omega \rightarrow \mathbb{R} : \omega \mapsto \sum_{i=1}^n \xi_i(\omega), & \text{for each } n \in \mathbb{N}. \end{cases}$$

- For each  $n \in \mathbb{N}$ , define  $X^{(n)} : \Omega \rightarrow C[0, 1]$  as follows:

$$X^{(n)}(\omega)(t) := \frac{1}{\sigma \cdot \sqrt{n}} \left\{ S_{i-1}(\omega) + n \left( t - \frac{i-1}{n} \right) \xi_i(\omega) \right\}, \text{ for each } \omega \in \Omega, t \in \left[ \frac{i-1}{n}, \frac{i}{n} \right], i = 1, 2, 3, \dots, n.$$

If

- (i) the sequence  $\{\xi_n\}_{n \in \mathbb{N}}$  is stationary  
(i.e. for each fixed  $j = 0, 1, 2, \dots$ , the distribution of  $(\xi_k, \xi_{k+1}, \dots, \xi_{k+j})$  is the same of all  $k \in \mathbb{N}$ ), and
- (ii)

$$\lim_{\lambda \rightarrow \infty} \limsup_{n \rightarrow \infty} \lambda^2 \cdot P\left(\max_{1 \leq k \leq n} |S_k| \geq \lambda \sigma \sqrt{n}\right) = 0,$$

then  $\{X^{(n)}\}_{n \in \mathbb{N}}$  is tight.

**PROOF** We apply the necessary and sufficient condition for tightness in Theorem F.1(iv). Thus, it suffices to prove the following two claims:

**Claim 1:** For each  $\eta > 0$ , there exist  $a > 0$  and  $n_0 \in \mathbb{N}$  such that

$$P_{X^{(n)}}\left(\left\{f \in C[0, 1] \mid |f(0)| \geq a\right\}\right) = P\left(\left|X_0^{(n)}\right| \geq a\right) \leq \eta, \text{ for each } n \geq n_0,$$

**Claim 2:** For each  $\varepsilon > 0$ ,

$$\lim_{\delta \rightarrow 0^+} \limsup_{n \rightarrow \infty} P_{X^{(n)}}\left(\left\{f \in C[0, 1] \mid w(f, \delta) \geq \varepsilon\right\}\right) = \lim_{\delta \rightarrow 0^+} \limsup_{n \rightarrow \infty} P\left(w(X^{(n)}, \delta) \geq \varepsilon\right) = 0.$$

Proof of Claim 1: Since, for each  $n \in \mathbb{N}$ ,  $X_0^{(n)}$  is identically zero, we may choose  $a = 1$  and  $n_0 = 1$ . We then have, for any  $\eta > 0$ ,

$$P_{X^{(n)}}\left(\left\{f \in C[0, 1] \mid |f(0)| \geq a\right\}\right) = P\left(\left|X_0^{(n)}\right| \geq a\right) = P\left(\left|X_0^{(n)}\right| \geq 1\right) = 0 \leq \eta, \text{ for each } n \geq n_0 := 1.$$

This proves Claim 1.

In order to simplify the proof of Claim 2, before presenting it, we first prove instead the following:

**Claim 3:** For each  $\varepsilon > 0$  and  $\delta \in (0, 1)$ ,

$$P\left(w(X^{(n)}, \delta) \geq 3\varepsilon\right) \leq \frac{2}{\delta} \cdot P\left(\max_{1 \leq k \leq \lceil n\delta \rceil} |S_k| \geq \frac{\varepsilon}{\sqrt{2\delta}} \sigma \sqrt{\lceil n\delta \rceil}\right), \text{ for all sufficiently large } n \in \mathbb{N}.$$

Proof of Claim 3: For each  $\delta \in (0, 1)$  and  $n \in \mathbb{N}$ , let  $m := \lceil n\delta \rceil$  be the round-up (“smallest integer greater than or equal to”) of  $n\delta$  and  $v = \lceil n/m \rceil$  be that of  $n/m$ . Let

$$t_0 = 0, \quad t_1 = \frac{m}{n}, \quad t_2 = \frac{2 \cdot m}{n}, \quad \dots, \quad t_{v-1} = \frac{(v-1) \cdot m}{n}, \quad t_v = 1.$$

Then,

$$t_i - t_{i-1} = \frac{i \cdot m}{n} - \frac{(i-1) \cdot m}{n} = \frac{m}{n} = \frac{\lceil n\delta \rceil}{n} \geq \frac{n\delta}{n} = \delta, \text{ for } i = 1, 2, \dots, v-1,$$

which implies, by Proposition D.5(i), that for any  $\varepsilon > 0$  and any  $f \in C[0, 1]$ , we have

$$w(f, \delta) \leq 3 \cdot \max_{1 \leq i \leq v} \left\{ \sup_{s \in [t_{i-1}, t_i]} |f(s) - f(t_{i-1})| \right\}.$$

Thus, in particular, we have

$$w(X^{(n)}(\omega), \delta) \leq 3 \cdot \max_{1 \leq i \leq v} \left\{ \sup_{s \in [t_{i-1}, t_i]} |X^{(n)}(\omega)(s) - X^{(n)}(\omega)(t_{i-1})| \right\}, \text{ for each } \omega \in \Omega.$$

Next, recall that  $X^{(n)}$  is continuous and piecewise linear with its set of discontinuity points contained in  $\{1/n, 2/n, \dots, (n-1)/n, 1\}$ , and

$$X^{(n)}(\omega) \left( \frac{i}{n} \right) = \frac{1}{\sigma \cdot \sqrt{n}} \cdot S_i(\omega), \text{ for } i = 1, 2, \dots, n.$$

Hence, writing  $m_i := i \cdot m$ ,  $i = 0, 1, 2, \dots, v-1$ , and  $m_v := n$  (hence  $t_i = m_i/n$ ,  $i = 0, 1, 2, \dots, v$ ), the preceding inequality becomes:

$$w(X^{(n)}(\omega), \delta) \leq 3 \cdot \max_{1 \leq i \leq v} \left\{ \max_{m_{i-1} \leq k \leq m_i} \frac{|S_k(\omega) - S_{m_{i-1}}(\omega)|}{\sigma \sqrt{n}} \right\}, \text{ for each } \omega \in \Omega,$$

which implies

$$\begin{aligned}
 P\left(\left\{\omega \in \Omega \mid w(X^{(n)}(\omega), \delta) \geq 3\varepsilon\right\}\right) &\leq P\left(\left\{\omega \in \Omega \mid \max_{1 \leq i \leq v} \left\{ \max_{m_{i-1} \leq k \leq m_i} \frac{|S_k(\omega) - S_{m_{i-1}}(\omega)|}{\sigma\sqrt{n}} \right\} \geq \varepsilon\right\}\right) \\
 &= P\left(\bigcup_{i=1}^v \left\{\omega \in \Omega \mid \max_{m_{i-1} \leq k \leq m_i} \frac{|S_k(\omega) - S_{m_{i-1}}(\omega)|}{\sigma\sqrt{n}} \geq \varepsilon\right\}\right) \\
 &\leq \sum_{i=1}^v P\left(\left\{\omega \in \Omega \mid \max_{m_{i-1} \leq k \leq m_i} |S_k(\omega) - S_{m_{i-1}}(\omega)| \geq \varepsilon\sigma\sqrt{n}\right\}\right) \\
 &= \sum_{i=1}^v P\left(\left\{\omega \in \Omega \mid \max_{1 \leq k \leq m_i - m_{i-1}} |S_k(\omega) - S_0(\omega)| \geq \varepsilon\sigma\sqrt{n}\right\}\right)
 \end{aligned}$$

Next, note that

$$\begin{aligned}
 n\delta \leq m := \lceil n\delta \rceil \leq n\delta + 1 &\implies \frac{1}{n\delta} \geq \frac{1}{m} \geq \frac{1}{n\delta + 1} \\
 &\implies \frac{1}{\delta} = \frac{n}{n\delta} \geq \frac{n}{m} \geq \frac{n}{n\delta + 1} = \frac{1}{\delta + 1/n},
 \end{aligned}$$

which implies  $\lim_{n \rightarrow \infty} \frac{n}{m} = \frac{1}{\delta}$ . In particular,

$$\frac{n}{m} > \frac{1}{2\delta}, \quad \text{for sufficiently large } n.$$

Consequently,

$$\varepsilon\sigma\sqrt{n} = \varepsilon\sigma\sqrt{\frac{n}{m}}\sqrt{m} \geq \frac{\varepsilon\sigma}{\sqrt{2\delta}}\sqrt{m}, \quad \text{for sufficiently large } n.$$

Furthermore, we have  $m_i - m_{i-1} = im - (i-1)m = m$ , for  $i = 1, 2, \dots, v-1$ , and

$$v := \lceil n/m \rceil \implies (v-1)m < n \leq vm \implies \begin{cases} m_v - m_{v-1} = n - (v-1)m \leq vm - (v-1)m \leq m \\ v < \frac{n}{m} + 1 < \frac{n}{m} + \frac{1}{\delta} \leq \frac{1}{\delta} + \frac{1}{\delta} = \frac{2}{\delta} \end{cases}$$

Thus, we see that, for sufficiently large  $n$ ,

$$\begin{aligned}
 P\left(\left\{\omega \in \Omega \mid w(X^{(n)}(\omega), \delta) \geq 3\varepsilon\right\}\right) &\leq \sum_{i=1}^v P\left(\left\{\omega \in \Omega \mid \max_{1 \leq k \leq m_i - m_{i-1}} |S_k(\omega) - S_0(\omega)| \geq \varepsilon\sigma\sqrt{n}\right\}\right) \\
 &\leq \sum_{i=1}^v P\left(\left\{\omega \in \Omega \mid \max_{1 \leq k \leq m} |S_k(\omega) - S_0(\omega)| \geq \frac{\varepsilon\sigma}{\sqrt{2\delta}}\sqrt{m}\right\}\right) \\
 &\leq v \cdot P\left(\left\{\omega \in \Omega \mid \max_{1 \leq k \leq m} |S_k(\omega)| \geq \frac{\varepsilon}{\sqrt{2\delta}}\sigma\sqrt{m}\right\}\right) \\
 &\leq \frac{2}{\delta} \cdot P\left(\left\{\omega \in \Omega \mid \max_{1 \leq k \leq m} |S_k(\omega)| \geq \frac{\varepsilon}{\sqrt{2\delta}}\sigma\sqrt{m}\right\}\right),
 \end{aligned}$$

where the second inequality uses the stationarity hypothesis (i). This completes the proof of Claim 3.

Proof of Claim 2: By Claim 3, we have, for each  $\varepsilon > 0$ , each  $\delta \in (0, 1)$  and each sufficiently large  $n \in \mathbb{N}$ ,

$$P\left(\left\{\omega \in \Omega \mid w(X^{(n)}(\omega), \delta) \geq 3\varepsilon\right\}\right) \leq \frac{2}{\delta} \cdot P\left(\left\{\omega \in \Omega \mid \max_{1 \leq k \leq \lceil n\delta \rceil} |S_k(\omega)| \geq \frac{\varepsilon}{\sqrt{2\delta}}\sigma\sqrt{\lceil n\delta \rceil}\right\}\right).$$

This implies:

$$\begin{aligned}
 & \limsup_{n \rightarrow \infty} P\left(\left\{\omega \in \Omega \mid w(X^{(n)}(\omega), \delta) \geq 3\varepsilon\right\}\right) \\
 & \leq \limsup_{n \rightarrow \infty} \frac{2}{\delta} \cdot P\left(\left\{\omega \in \Omega \mid \max_{1 \leq k \leq \lceil n\delta \rceil} |S_k(\omega)| \geq \frac{\varepsilon}{\sqrt{2\delta}} \sigma \sqrt{\lceil n\delta \rceil}\right\}\right) \\
 & \leq \limsup_{n \rightarrow \infty} \frac{2}{\delta} \cdot P\left(\left\{\omega \in \Omega \mid \max_{1 \leq k \leq n} |S_k(\omega)| \geq \frac{\varepsilon}{\sqrt{2\delta}} \sigma \sqrt{n}\right\}\right) \\
 & = \frac{4}{\varepsilon^2} \cdot \limsup_{n \rightarrow \infty} \left(\frac{\varepsilon}{\sqrt{2\delta}}\right)^2 \cdot P\left(\left\{\omega \in \Omega \mid \max_{1 \leq k \leq n} |S_k(\omega)| \geq \frac{\varepsilon}{\sqrt{2\delta}} \sigma \sqrt{n}\right\}\right),
 \end{aligned}$$

where the second inequality follows from the fact that, for each fixed  $\delta \in (0, 1)$ ,  $\{\lceil n\delta \rceil\}_{n \in \mathbb{N}}$  is a sequence of non-decreasing natural numbers but which may contain repeated entries. The preceding inequality and hypothesis (ii) together imply:

$$\begin{aligned}
 & \lim_{\delta \rightarrow 0^+} \limsup_{n \rightarrow \infty} P\left(\left\{\omega \in \Omega \mid w(X^{(n)}(\omega), \delta) \geq 3\varepsilon\right\}\right) \\
 & \leq \frac{4}{\varepsilon^2} \cdot \lim_{\delta \rightarrow 0^+} \limsup_{n \rightarrow \infty} \left(\frac{\varepsilon}{\sqrt{2\delta}}\right)^2 \cdot P\left(\left\{\omega \in \Omega \mid \max_{1 \leq k \leq n} |S_k(\omega)| \geq \frac{\varepsilon}{\sqrt{2\delta}} \sigma \sqrt{n}\right\}\right) \\
 & = \frac{4}{\varepsilon^2} \cdot \lim_{\lambda \rightarrow \infty} \limsup_{n \rightarrow \infty} (\lambda)^2 \cdot P\left(\left\{\omega \in \Omega \mid \max_{1 \leq k \leq n} |S_k(\omega)| \geq \lambda \sigma \sqrt{n}\right\}\right) \\
 & = \frac{4}{\varepsilon^2} \cdot 0 = 0.
 \end{aligned}$$

This completes the proof of Claim 2, as well as that of the present Lemma. □

## 4 The Wiener measure on $(C[0, 1], \|\cdot\|_\infty)$

**Definition 4.1 (Wiener measure on  $(C[0, 1], \|\cdot\|_\infty)$ )**

A Borel probability measure  $W$  on  $(C[0, 1], \|\cdot\|_\infty)$  is called a **Wiener measure** if it satisfies the following two conditions:

- (i) Its induced measure  $W \circ \text{ev}_0^{-1}$  on  $\mathbb{R}$  via the evaluation map  $\text{ev}_0 : C[0, 1] \rightarrow \mathbb{R} : f \mapsto f(0)$  is the point-mass measure on  $\mathbb{R}$  concentrated at  $0 \in \mathbb{R}$ , i.e.

$$W(\text{ev}_0^{-1}(\{0\})) = W\left(\left\{f \in C[0, 1] \mid f(0) = 0\right\}\right) = 1.$$

- (ii) For any  $0 \leq t_0 < t_1 < t_2 < \dots < t_k \leq 1$ ,

$$W \circ \Pi_{t_0 t_1 \dots t_k}^{-1} = N\left(\mu = \mathbf{0}, \Sigma = \text{diag}(t_1 - t_0, \dots, t_k - t_{k-1})\right),$$

where the map  $\Pi_{t_0 t_1 \dots t_k} : C[0, 1] \rightarrow \mathbb{R}^k$  is defined as follows:

$$\Pi_{t_0 t_1 \dots t_k} : C[0, 1] \rightarrow \mathbb{R}^k : f \mapsto (f(t_1) - f(t_0), f(t_2) - f(t_1), \dots, f(t_k) - f(t_{k-1})).$$

**Theorem 4.2 (Finite-dimensional distributions of the Wiener measure)**

Suppose  $W$  is a Wiener measure defined on  $(C[0, 1], \|\cdot\|_\infty)$ . Then, for any pairwise distinct  $0 \leq t_1, t_2, \dots, t_k \leq 1$ ,

$$W \circ \pi_{t_1 \dots t_k}^{-1} = N\left(\mu = \mathbf{0}, \Sigma = \left[\min\{t_i, t_j\}\right]_{1 \leq i, j \leq k}\right),$$

where the map  $\pi_{t_1 \dots t_k} : C[0, 1] \longrightarrow \mathbb{R}^k$  is defined as follows:

$$\pi_{t_1 \dots t_k} : C[0, 1] \longrightarrow \mathbb{R}^k : f \longmapsto (f(t_1), f(t_2), \dots, f(t_k)).$$

PROOF Let  $t_0 := 0$ . Re-labeling the  $t_i$ 's if necessary, without loss of generality, we may assume that  $t_1 < t_2 < \dots < t_k$ . Let  $\mathcal{B}$  denote the Borel  $\sigma$ -algebra of  $(C[0, 1], \|\cdot\|_\infty)$ . Then,  $(C[0, 1], \mathcal{B}, W)$  is the probability space obtained by equipping the measurable space  $(C[0, 1], \mathcal{B})$  with the Wiener measure  $W$ .

Define the following  $\mathbb{R}$ -valued random variables on the probability space  $(C[0, 1], \mathcal{B}, W)$ :

$$\begin{aligned} Z_{t_1-t_0} &: C[0, 1] \longrightarrow \mathbb{R}, & Z_{t_1-t_0}(f) &:= f(t_1) \\ Z_{t_i-t_{i-1}} &: C[0, 1] \longrightarrow \mathbb{R}, & Z_{t_i-t_{i-1}}(f) &:= f(t_i) - f(t_{i-1}), \quad \text{for } i = 2, \dots, k \\ L_{t_i} &: C[0, 1] \longrightarrow \mathbb{R}, & L_{t_i}(f) &:= f(t_i), \quad \text{for } i = 1, \dots, k \end{aligned}$$

Then, we have

$$\begin{pmatrix} L_{t_1}(f) \\ L_{t_2}(f) \\ \vdots \\ L_{t_k}(f) \end{pmatrix} = \begin{pmatrix} f(t_1) \\ f(t_2) \\ \vdots \\ f(t_k) \end{pmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ 1 & 1 & 0 & \cdots & \cdots & 0 & 0 \\ 1 & 1 & 1 & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 & 0 \\ 1 & 1 & 1 & \cdots & \cdots & 1 & 0 \\ 1 & 1 & 1 & \cdots & \cdots & 1 & 1 \end{bmatrix}}_T \cdot \begin{pmatrix} f(t_1) \\ f(t_2) - f(t_1) \\ \vdots \\ f(t_k) - f(t_{k-1}) \end{pmatrix} = T \cdot \begin{pmatrix} Z_{t_1-t_0}(f) \\ Z_{t_2-t_1}(f) \\ \vdots \\ Z_{t_k-t_{k-1}}(f) \end{pmatrix}$$

The present Theorem will be proved once we establish that the  $\mathbb{R}^k$ -valued random variable  $\pi_{t_1 \dots t_k} = (L_{t_1}, L_{t_2}, \dots, L_{t_k}) : C[0, 1] \longrightarrow \mathbb{R}^k$  defined on  $(C[0, 1], \mathcal{B}, W)$  has the following multivariate Gaussian distribution:

$$\pi_{t_1 \dots t_k} = \begin{pmatrix} L_{t_1} \\ L_{t_2} \\ \vdots \\ L_{t_k} \end{pmatrix} \sim N\left(\mu = \mathbf{0}, \Sigma = \left[\min\{t_i, t_j\}\right]_{1 \leq i, j \leq k}\right).$$

To this end, note that since  $t_0 := 0$ , by the definition of a Wiener measure, the  $\mathbb{R}^k$ -valued random variable  $(Z_{t_1-t_0}, Z_{t_2-t_1}, \dots, Z_{t_k-t_{k-1}}) : C[0, 1] \longrightarrow \mathbb{R}^k$  equals  $\Pi_{t_0 t_1 \dots t_k} W$ -almost-surely (since  $W(f(0) = 0) = 1$ ). These two  $\mathbb{R}^k$ -valued random variables therefore have the same probability distribution. By the definition of a Wiener measure again, this common distribution is the following multivariate Gaussian distribution:

$$\begin{pmatrix} Z_{t_1-t_0} \\ Z_{t_2-t_1} \\ \vdots \\ Z_{t_k-t_{k-1}} \end{pmatrix} \sim N\left(\mu = \mathbf{0}, \Sigma = \text{diag}(t_1 - t_0, \dots, t_k - t_{k-1})\right).$$

Since the  $\mathbb{R}^k$ -valued random variable  $(L_{t_1}, L_{t_2}, \dots, L_{t_k}) : C[0, 1] \rightarrow \mathbb{R}^k$  can be obtained from  $(Z_{t_1-t_0}, Z_{t_2-t_1}, \dots, Z_{t_k-t_{k-1}}) : C[0, 1] \rightarrow \mathbb{R}^k$  via the non-singular linear transformation  $T : \mathbb{R}^k \rightarrow \mathbb{R}^k$ , it follows that  $(L_{t_1}, L_{t_2}, \dots, L_{t_k})$  also has a multivariate Gaussian distribution. Since  $(Z_{t_1-t_0}, Z_{t_2-t_1}, \dots, Z_{t_k-t_{k-1}})$  has mean vector zero, so does  $(L_{t_1}, L_{t_2}, \dots, L_{t_k})$ . It remains only to compute the covariance matrix of  $(L_{t_1}, L_{t_2}, \dots, L_{t_k})$ , which we now do: For  $t_i < t_j$ , we have

$$\begin{aligned} \text{Cov}(L_{t_i}, L_{t_j}) &= \text{Cov}(Z_{t_1-t_0} + Z_{t_2-t_1} + \dots + Z_{t_i-t_{i-1}}, Z_{t_1-t_0} + Z_{t_2-t_1} + \dots + Z_{t_j-t_{j-1}}) \\ &= \text{Cov}\left(\sum_{a=1}^i Z_{t_a-t_{a-1}}, \sum_{b=1}^j Z_{t_b-t_{b-1}}\right) = \sum_{a=1}^i \sum_{b=1}^j \text{Cov}(Z_{t_a-t_{a-1}}, Z_{t_b-t_{b-1}}) \\ &= \sum_{a=1}^i \text{Cov}(Z_{t_a-t_{a-1}}, Z_{t_a-t_{a-1}}) = \sum_{a=1}^i \text{Var}(Z_{t_a-t_{a-1}}) = \sum_{a=1}^i (t_a - t_{a-1}) \\ &= (t_1 - t_0) + (t_2 - t_1) + \dots + (t_{i-1} - t_{i-2}) + (t_i - t_{i-1}) \\ &= t_i = \min\{t_i, t_j\}, \end{aligned}$$

as required. This completes the proof of the Theorem. □

### Theorem 4.3 (Uniqueness of the Wiener measure)

Any two Wiener measures on  $(C[0, 1], \|\cdot\|_\infty)$  must in fact be equal.

**PROOF** By Theorem 4.2, the collection of finite-dimensional distributions of a Wiener measure on  $(C[0, 1], \|\cdot\|_\infty)$  is completely determined. Thus, any two Wiener measures  $W_1$  and  $W_2$  must have the exactly the same finite-dimensional distributions, which in turn implies that  $W_1$  and  $W_2$  must agree on the entire collection of finite-dimensional subsets of  $C[0, 1]$ . Lastly, recall that the finite-dimensional subsets of  $C[0, 1]$  form a separating class of the Borel  $\sigma$ -algebra of  $(C[0, 1], w\|\cdot\|_\infty)$  (Example 1.3, p.11, [3]). We may now conclude that  $W_1 = W_2$ , as Borel probability measures on  $(C[0, 1], w\|\cdot\|_\infty)$ . □

### Theorem 4.4 (Existence of the Wiener measure)

There exists a Wiener measure on  $(C[0, 1], \|\cdot\|_\infty)$ .

**PROOF** Let  $\xi_1, \xi_2, \dots : \Omega \rightarrow \mathbb{R}$  be a sequence of independent and identically distributed standard Gaussian  $\mathbb{R}$ -valued random variables defined on the probability space  $(\Omega, \mathcal{A}, \mu)$  (i.e. with expectation value zero and common finite variance  $\sigma^2 = 1$ ). Define the random variables:

$$\begin{cases} S_0 & : \Omega \rightarrow \mathbb{R} : \omega \mapsto 0, & \text{and} \\ S_n & : \Omega \rightarrow \mathbb{R} : \omega \mapsto \sum_{i=1}^n \xi_i(\omega), & \text{for each } n \in \mathbb{N}. \end{cases}$$

For each  $n \in \mathbb{N}$ , define  $X^{(n)} : \Omega \rightarrow C[0, 1]$  as follows:

$$X^{(n)}(\omega)(t) := \frac{1}{\sqrt{n}} \left\{ S_{i-1}(\omega) + n \left( t - \frac{i-1}{n} \right) \xi_i(\omega) \right\}, \text{ for each } \omega \in \Omega, t \in \left[ \frac{i-1}{n}, \frac{i}{n} \right], i = 1, 2, 3, \dots, n.$$

#### Claim 1:

The sequence  $\{P_n := P_{X^{(n)}}\}_{n \in \mathbb{N}}$  of Borel probability measures on  $(C[0, 1], \|\cdot\|_\infty)$  induced by the  $X^{(n)}$ 's is tight.

We defer the proof of Claim 1 below. Granting the validity of Claim 1, we see that  $\{P_n := P_{X^{(n)}}\}_{n \in \mathbb{N}}$  admits a weakly convergent subsequence  $\{P_{n(i)}\}_{i \in \mathbb{N}}$ , say  $P_{n(i)} \xrightarrow{d} W$ , as  $i \rightarrow \infty$ , where  $W$  is some Borel probability measure on  $(C[0, 1], \|\cdot\|_\infty)$ .

**Claim 2:**  $W$  is a Wiener measure on  $(C[0, 1], \|\cdot\|_\infty)$ .

Of course, Claim 2 is the desired conclusion of the present Theorem. Thus, the proof of the Theorem will be complete once we prove Claim 1 and Claim 2, which we now do.

Proof of Claim 1: We will use Lemma 3.1. Since  $\xi_1, \xi_2, \dots$  are independent and identically distributed, they satisfy trivially the stationarity hypothesis in Lemma 3.1. Thus, Claim 1 will follow once we establish the following:

$$\lim_{\lambda \rightarrow \infty} \limsup_{n \rightarrow \infty} \lambda^2 \cdot P\left(\max_{1 \leq k \leq n} |S_k| \geq \lambda \sqrt{n}\right) = 0.$$

Since, for each  $k \in \mathbb{N}$ ,  $S_k = \xi_1 + \xi_2 + \dots + \xi_k$  is a sum of  $k$  independent standard Gaussian  $\mathbb{R}$ -valued random variables,  $S_k/\sqrt{k}$  itself is a standard Gaussian  $\mathbb{R}$ -valued random variable. Now, observe that:

$$\begin{aligned} P\left(\max_{1 \leq k \leq n} |S_k| \geq \lambda \sqrt{n}\right) &\leq 3 \cdot \max_{1 \leq k \leq n} P\left(|S_k| \geq \frac{\lambda \sqrt{n}}{3}\right), \quad \text{by Etemadi's inequality (Theorem H.1)} \\ &= 3 \cdot \max_{1 \leq k \leq n} P\left(\left|\frac{S_k}{\sqrt{k}}\right| \geq \frac{\lambda \sqrt{n}}{3\sqrt{k}}\right) \\ &= 3 \cdot \max_{1 \leq k \leq n} P\left(|Z| \geq \frac{\lambda \sqrt{n}}{3\sqrt{k}}\right), \quad \text{where } Z \sim N(0, 1) \\ &\leq 3 \cdot \max_{1 \leq k \leq n} \left(\frac{3\sqrt{k}}{\lambda \sqrt{n}}\right)^4 \cdot E(|Z|^4), \quad \text{by Chebyshev's inequality} \\ &= 3 \cdot \frac{3^4}{\lambda^4} \cdot 3, \quad \text{since } E(|Z|^4) = 3 \\ &\leq \frac{3^6}{\lambda^4}, \end{aligned}$$

which implies

$$\lim_{\lambda \rightarrow \infty} \limsup_{n \rightarrow \infty} \lambda^2 \cdot P\left(\max_{1 \leq k \leq n} |S_k| \geq \lambda \sqrt{n}\right) \leq \lim_{\lambda \rightarrow \infty} \limsup_{n \rightarrow \infty} \lambda^2 \cdot \frac{3^6}{\lambda^4} = \lim_{\lambda \rightarrow \infty} \frac{3^6}{\lambda^2} = 0.$$

This proves Claim 1.

Proof of Claim 2: First of all,  $W$  above was chosen to be the weak limit of the subsequence  $P_{n(i)}$ , so we have  $P_{n(i)} \xrightarrow{d} W$ . Now, recall that  $\text{ev}_0 : C[0, 1] \rightarrow \mathbb{R} : f \mapsto f(0)$  is continuous. Hence,

$$\begin{aligned} &P_{n(i)} \xrightarrow{d} W \\ \implies &P_{n(i)} \circ \text{ev}_0^{-1} \xrightarrow{d} W \circ \text{ev}_0^{-1}, \quad \text{by the Continuous Mapping Theorem (Theorem 2.7, p.21, [3])} \\ \implies &\limsup_{i \rightarrow \infty} P_{n(i)} \circ \text{ev}_0^{-1}(\{0\}) \leq W \circ \text{ev}_0^{-1}(\{0\}), \quad \text{by the Portmanteau Theorem (Theorem 2.1, p.16, [3])} \\ \implies &\limsup_{i \rightarrow \infty} P(X^{(n(i))}(0) = 0) \leq W \circ \text{ev}_0^{-1}(\{0\}) \\ \implies &1 = \limsup_{i \rightarrow \infty} P(S_0 = 0) \leq W \circ \text{ev}_0^{-1}(\{0\}), \end{aligned}$$

which implies

$$W \circ \text{ev}_0^{-1}(\{0\}) = 1.$$

Secondly, since, for each  $0 \leq t_0 < t_1 < t_2 < \dots < t_k \leq 1$ , the map

$$\Pi_{t_0 t_1 \dots t_k} : C[0, 1] \longrightarrow \mathbb{R}^k : f \longmapsto \left( f(t_1) - f(t_0), f(t_2) - f(t_1), \dots, f(t_k) - f(t_{k-1}) \right)$$

is continuous (proof is routine, hence omitted), by the Continuous Mapping Theorem (Theorem 2.7, p.21, [3]), we have

$P_{n(i)} \circ \Pi_{t_0 t_1 \dots t_k}^{-1} \xrightarrow{d} W \circ \Pi_{t_0 t_1 \dots t_k}^{-1}$ . By Theorem 2.3, we have also

$$P_n \circ \Pi_{t_0 t_1 \dots t_k}^{-1} \xrightarrow{d} N\left(\mu = \mathbf{0}, \Sigma = \text{diag}(t_1 - t_0, \dots, t_k - t_{k-1})\right), \text{ as } n \longrightarrow \infty,$$

in particular,

$$P_{n(i)} \circ \Pi_{t_0 t_1 \dots t_k}^{-1} \xrightarrow{d} N\left(\mu = \mathbf{0}, \Sigma = \text{diag}(t_1 - t_0, \dots, t_k - t_{k-1})\right), \text{ as } i \longrightarrow \infty.$$

Thus, both  $W \circ \Pi_{t_0 t_1 \dots t_k}^{-1}$  and  $N\left(\mu = \mathbf{0}, \Sigma = \text{diag}(t_1 - t_0, \dots, t_k - t_{k-1})\right)$  are weak limits of  $P_{n(i)} \circ \Pi_{t_0 t_1 \dots t_k}^{-1}$ , as  $i \longrightarrow \infty$ . By Theorem A.3, it follows that

$$W \circ \Pi_{t_0 t_1 \dots t_k}^{-1} = N\left(\mu = \mathbf{0}, \Sigma = \text{diag}(t_1 - t_0, \dots, t_k - t_{k-1})\right).$$

This proves that  $W$  is indeed a Wiener measure on  $(C[0, 1], \|\cdot\|_\infty)$ . This completes the proof of Claim 2, as well as that of the present Theorem.  $\square$



## A Technical Lemmas

Note that  $X_n \xrightarrow{d} X$  and  $Y_n \xrightarrow{d} Y$  does NOT in general imply  $X_n + Y_n \xrightarrow{d} X + Y$ . But the implication does hold if  $X_n$  and  $Y_n$  are independent for each  $n \in \mathbb{N}$ , and both  $X$  and  $Y$  are Gaussian random variables, as the following Proposition shows.

**Proposition A.1** *Let  $k \in \mathbb{N}$  be fixed. Suppose:*

- For each  $n \in \mathbb{N}$ ,

$$Y_1^{(n)}, Y_2^{(n)}, \dots, Y_k^{(n)} : \Omega^{(n)} \longrightarrow \mathbb{R}$$

*are independent  $\mathbb{R}$ -valued random variables defined on the probability space  $\Omega^{(n)}$ .*

- For each  $i = 1, 2, \dots, k$ ,

$$Y_i^{(n)} \xrightarrow{d} N(\mu_i, \sigma_i^2), \quad \text{as } n \longrightarrow \infty.$$

Then, for any  $c_1, c_2, \dots, c_k \in \mathbb{R}$ ,

$$\sum_{i=1}^k c_i Y_i^{(n)} \xrightarrow{d} N\left(\sum_{i=1}^k c_i \mu_i, \sum_{i=1}^k c_i^2 \sigma_i^2\right), \quad \text{as } n \longrightarrow \infty.$$

PROOF Let  $Y^{(n)} := \sum_{i=1}^k c_i Y_i^{(n)}$ . Let  $\varphi_X$  denote the characteristic function of a  $\mathbb{R}$ -valued random variable  $X$ . Then,

$$\begin{aligned} \varphi_{Y^{(n)}}(t) &= \varphi_{\sum_{i=1}^k c_i Y_i^{(n)}}(t) \\ &= \prod_{i=1}^k \varphi_{c_i Y_i^{(n)}}(t), \quad \text{since } Y_1^{(n)}, \dots, Y_k^{(n)} \text{ are independent} \\ &= \prod_{i=1}^k \varphi_{Y_i^{(n)}}(c_i t) \\ &\longrightarrow \prod_{i=1}^k \exp\left\{\sqrt{-1} \mu_i (c_i t) - \frac{1}{2} \sigma_i^2 (c_i t)^2\right\} \\ &= \exp\left\{\sqrt{-1} \left(\sum_{i=1}^k c_i \mu_i\right) t - \frac{1}{2} \left(\sum_{i=1}^k c_i^2 \sigma_i^2\right) t^2\right\}, \quad \text{as } n \longrightarrow \infty, \end{aligned}$$

where the second and third equalities follow from the properties of characteristic functions of random variables (see p.21, [4]), while the expression of the limit follows from the fact that the characteristic function  $\varphi_Z$  of a random variable  $Z$  with distribution  $N(\mu, \sigma^2)$  is

$$\varphi_Z = \exp\left\{\sqrt{-1} \mu t - \frac{1}{2} \sigma^2 t^2\right\}.$$

The Proposition now follows immediately from the Lévy-Cramér Continuity Theorem (Theorem 1.9(ii), p.56, [5]).  $\square$

### Proposition A.2

Suppose:

- $k \in \mathbb{N}$  and  $Q_1, Q_2, \dots, Q_k$  are Borel probability measures on  $(C[0, 1], \|\cdot\|_\infty)$ .
- $\{P_n\}_{n \in \mathbb{N}}$  is a sequence of Borel probability measures on  $(C[0, 1], \|\cdot\|_\infty)$ .

Then, the following are equivalent:

- (i)  $\{P_n\}_{n \in \mathbb{N}}$  is tight.
- (ii)  $\{Q_1, \dots, Q_k\} \cup \{P_n\}_{n \in \mathbb{N}}$  is tight.

PROOF Note that once we prove the Proposition for  $k = 1$ , the cases  $k \geq 2$  will follow immediately by finite induction. We now proceed to prove the Proposition for  $k = 1$ , and we write  $Q$  for  $Q_1$ .

(ii)  $\implies$  (i) This is trivial. Indeed,

$$\begin{aligned}
 & \{Q\} \cup \{P_n\}_{n \in \mathbb{N}} \text{ is tight} \\
 \iff & \text{for each } \varepsilon > 0, \exists \text{ compact } K \subset C[0, 1] \text{ such that } 1 - \varepsilon < Q(K) \leq 1 \text{ and } 1 - \varepsilon < P_n(K) \leq 1, \text{ for all } n \in \mathbb{N} \\
 \implies & \text{for each } \varepsilon > 0, \exists \text{ compact } K \subset C[0, 1] \text{ such that } 1 - \varepsilon < P_n(K) \leq 1, \text{ for all } n \in \mathbb{N} \\
 \iff & \{P_n\}_{n \in \mathbb{N}} \text{ is tight}
 \end{aligned}$$

This proves that (ii)  $\implies$  (i).

(i)  $\implies$  (ii)

Let  $\varepsilon > 0$  be given. Since  $(C[0, 1], \|\cdot\|_\infty)$  is separable and complete, the single Borel probability measure  $Q$  on  $(C[0, 1], \|\cdot\|_\infty)$  is tight (Theorem 1.3, [3]). Thus, there exists a compact subset  $K_1 \subset C[0, 1]$  such that  $1 - \varepsilon < Q(K_1) \leq 1$ . On the other hand, the tightness hypothesis on  $\{P_n\}_{n \in \mathbb{N}}$  implies there exists a compact subset  $K_2 \subset C[0, 1]$  such that  $1 - \varepsilon < P_n(K_2) \leq 1$ , for each  $n \in \mathbb{N}$ . Let  $K := K_1 \cup K_2$ . Then,  $K$  is itself a compact subset of  $C[0, 1]$ , and

$$Q(K) = Q(K_1 \cup K_2) \geq Q(K_1) > 1 - \varepsilon, \quad \text{and} \quad P_n(K) = P_n(K_1 \cup K_2) \geq P_n(K_2) > 1 - \varepsilon.$$

This proves the tightness of  $\{Q\} \cup \{P_n\}_{n \in \mathbb{N}}$ , as required.  $\square$

### Theorem A.3 (Uniqueness of weak limit of a sequence of Borel probability measures on a metric space)

Suppose  $S$  is a metric space and  $\mathcal{B}(S)$  is its Borel  $\sigma$ -algebra. If  $P, Q, P_1, P_2, \dots$  are Borel probability measures on  $(S, \mathcal{B}(S))$  such that

$$P_n \xrightarrow{d} P \quad \text{and} \quad P_n \xrightarrow{d} Q, \quad \text{as } n \longrightarrow \infty,$$

then  $P = Q$ , as Borel probability measures on  $(S, \mathcal{B}(S))$ .

PROOF Observe that, by the definition of weak convergence of measures, for each bounded continuous  $\mathbb{R}$ -valued function  $f : S \longrightarrow \mathbb{R}$  defined on  $S$ , we have:

$$\int_S f \, dP = \lim_{n \rightarrow \infty} \int_S f \, dP_n = \int_S f \, dQ.$$

Thus,  $P = Q$ , as Borel probability measures on  $(S, \mathcal{B}(S))$ , by Theorem 1.2, p.8, [3].  $\square$

## B Continuous maps are Borel measurable

### Lemma B.1 (The pre-image of a $\sigma$ -algebra is itself a $\sigma$ -algebra.)

Suppose  $\Omega$  is a non-empty set,  $(X, \mathcal{X})$  is a measurable space, and  $f : \Omega \longrightarrow X$  is a map from  $\Omega$  into  $X$ . Then,

$$f^{-1}(\mathcal{X}) := \{f^{-1}(V) \subset \Omega \mid V \in \mathcal{X}\}$$

is a  $\sigma$ -algebra of subsets of  $\Omega$ .

PROOF

$$\underline{\Omega \in f^{-1}(\mathcal{X})} \quad f(\Omega) \subset X \implies \Omega = f^{-1}(X) \in f^{-1}(\mathcal{X}).$$

$f^{-1}(\mathcal{X})$  is closed under complementations Let  $V \in \mathcal{X}$ . Then,  $X \setminus V \in \mathcal{X}$ , and

$$\Omega \setminus f^{-1}(V) = \{ \omega \in \Omega \mid f(\omega) \notin V \} = \{ \omega \in \Omega \mid f(\omega) \in X \setminus V \} = f^{-1}(X \setminus V) \in f^{-1}(\mathcal{X}),$$

which shows that  $f^{-1}(\mathcal{X})$  is indeed closed under complementations.

$f^{-1}(\mathcal{X})$  is closed countable unions Let  $V_1, V_2, \dots \in \mathcal{X}$ . Then,  $\bigcup_{i=1}^{\infty} V_i \in \mathcal{X}$ , and

$$\bigcup_{i=1}^{\infty} f^{-1}(V_i) = \left\{ \omega \in \Omega \mid \begin{array}{l} f(\omega) \in V_i \\ \text{for some } i \in \mathbb{N} \end{array} \right\} = f^{-1}\left(\bigcup_{i=1}^{\infty} V_i\right) \in f^{-1}(\mathcal{X}),$$

which proves that  $f^{-1}(\mathcal{X})$  is indeed closed under countable unions.

This concludes the proof that that  $f^{-1}(\mathcal{X})$  is a  $\sigma$ -algebra of subsets of  $\Omega$ . □

**Lemma B.2 (The push-forward of a  $\sigma$ -algebra is itself a  $\sigma$ -algebra.)**

Suppose  $(\Omega, \mathcal{A})$  is a measurable space,  $X$  is a non-empty set, and  $f : \Omega \rightarrow X$  is a map from  $\Omega$  into  $X$ . Then,

$$\mathcal{F} := \{ V \subset X \mid f^{-1}(V) \in \mathcal{A} \}$$

is a  $\sigma$ -algebra of subsets of  $X$ .

PROOF

$$\underline{X \in \mathcal{F}} \quad f^{-1}(X) = \Omega \in \mathcal{A} \implies X \in \mathcal{F}.$$

$\mathcal{F}$  is closed under complementations  $V \in \mathcal{F} \implies f^{-1}(V) \in \mathcal{A} \implies f^{-1}(X \setminus V) = \Omega \setminus f^{-1}(V) \in \mathcal{A} \implies X \setminus V \in \mathcal{F}$ , which proves that  $\mathcal{F}$  is indeed closed under complementations.

$\mathcal{F}$  is closed under countable unions

$$\begin{aligned} V_1, V_2, \dots \in \mathcal{F} &\implies f^{-1}(V_1), f^{-1}(V_2), \dots \in \mathcal{A} \\ &\implies f^{-1}\left(\bigcup_{i=1}^{\infty} V_i\right) = \bigcup_{i=1}^{\infty} f^{-1}(V_i) \in \mathcal{A} \\ &\implies \bigcup_{i=1}^{\infty} V_i \in \mathcal{F}, \end{aligned}$$

which proves that  $\mathcal{F}$  is indeed closed under countable unions. □

**Theorem B.3**

Suppose  $(\Omega, \mathcal{A})$  and  $(X, \mathcal{X})$  are measurable spaces, and  $f : \Omega \rightarrow X$  is a map from  $\Omega$  into  $X$ . Then,  $f$  is  $(\mathcal{A}, \mathcal{X})$ -measurable if there exists  $\mathcal{S} \subset \mathcal{X}$  satisfying the following conditions:

- $\mathcal{S}$  generates  $\mathcal{X}$ , i.e.  $\sigma(\mathcal{S}) = \mathcal{X}$ , and
- $f^{-1}(\mathcal{S}) \subset \mathcal{A}$ .

PROOF By Lemma B.2,

$$\mathcal{F} := \{ V \subset X \mid f^{-1}(V) \in \mathcal{A} \}$$

is a  $\sigma$ -algebra of subsets of  $X$ . By hypothesis,  $\mathcal{S} \subset \mathcal{F}$ ; hence,  $\mathcal{X} = \sigma(\mathcal{S}) \subset \mathcal{F}$ . Thus,  $f^{-1}(\mathcal{X}) \subset \mathcal{A}$ ; equivalently,  $f$  is  $(\mathcal{A}, \mathcal{X})$ -measurable.  $\square$

**Corollary B.4 (Continuous maps are Borel measurable.)**

Suppose  $X_1, X_2$  are topological spaces, and  $\mathcal{B}_1, \mathcal{B}_2$  are their respective Borel  $\sigma$ -algebras. Then, every continuous map  $f : X_1 \rightarrow X_2$  is  $(\mathcal{B}_1, \mathcal{B}_2)$ -measurable.

## C Topology

**Theorem C.1 (Appendix M3, [3])**

Suppose  $S$  is a metric space. Then, the following conditions are equivalent:

- (i)  $S$  is separable.
- (ii) The topology of  $S$  has a countable basis.
- (iii) Every open cover of *each subset* of  $S$  has a countable subcover.

## D Modulus of continuity

**Definition D.1 (Modulus of continuity)**

Let  $\mathbb{R}^{[0,1]}$  denote the set of all arbitrary  $\mathbb{R}$ -valued functions defined on the closed unit interval  $[0, 1]$ . The **modulus of continuity** is, by definition, the following function:

$$w : \mathbb{R}^{[0,1]} \times (0, 1] \rightarrow [0, \infty] : (f, \delta) \mapsto \sup \left\{ |f(s) - f(t)| \mid \begin{array}{l} (s, t) \in [0, 1] \times [0, 1] \\ |s - t| \leq \delta \end{array} \right\}.$$

**Proposition D.2**

- (i) The restriction of  $w$  to  $C[0, 1] \times (0, 1]$  takes values in  $[0, \infty)$ .
- (ii) For each  $\delta \in (0, 1]$  and each  $f, g \in C[0, 1]$ , we have

$$|w(f, \delta) - w(g, \delta)| \leq 2 \|f - g\|_\infty.$$

- (iii) For each  $\delta \in (0, 1]$ , the map  $w(\cdot, \delta) : (C[0, 1], \|\cdot\|_\infty) \rightarrow \mathbb{R} : f \mapsto w(f, \delta)$  is continuous.

PROOF

For each  $\delta \in (0, 1]$ , the set

$$D(\delta) := \left\{ (s, t) \in [0, 1] \times [0, 1] \mid |s - t| \leq \delta \right\}$$

is a subset of the compact set  $[0, 1] \times [0, 1]$ .

- (i) For each  $f \in C[0, 1]$ , the map

$$[0, 1] \times [0, 1] \rightarrow \mathbb{R} : (s, t) \mapsto |f(s) - f(t)|$$

is continuous on the compact set  $[0, 1] \times [0, 1]$ ; hence, it is bounded and attains its supremum on  $[0, 1] \times [0, 1]$ ; in particular, its supremum on  $[0, 1] \times [0, 1]$  is a (finite non-negative) real number. Hence,

$$\begin{aligned} w(f, \delta) &:= \sup \left\{ |f(s) - f(t)| \mid \begin{array}{l} (s, t) \in [0, 1] \times [0, 1] \\ |s - t| \leq \delta \end{array} \right\} = \sup_{(s, t) \in D(\delta)} \left\{ |f(s) - f(t)| \right\} \\ &\leq \sup_{(s, t) \in [0, 1] \times [0, 1]} \left\{ |f(s) - f(t)| \right\} < \infty. \end{aligned}$$

This proves (i).

(ii) Recall that for any  $a, b \in \mathbb{R}$ , we have

$$|a| = |a - b + b| \leq |a - b| + |b|,$$

which in turn implies:

$$|a| - |b| \leq |a - b|.$$

Consequently, for each  $f, g \in C[0, 1]$  and each  $s, t \in [0, 1]$ , we have:

$$\begin{aligned} \left| |f(s) - f(t)| - |g(s) - g(t)| \right| &\leq \left| f(s) - f(t) - g(s) + g(t) \right| \\ &\leq |f(s) - g(s)| + |f(t) - g(t)| \leq 2 \|f - g\|_\infty. \end{aligned}$$

On the other hand, note that for any  $g \in C[0, 1]$  and any  $(s, t) \in [0, 1] \times [0, 1]$  with  $|s - t| \leq \delta$ , i.e.  $(s, t) \in D(\delta)$ , we have

$$\left| g(s) - g(t) \right| \leq \sup_{(\xi, \zeta) \in D(\delta)} \left\{ |g(\xi) - g(\zeta)| \right\} =: w(g, \delta),$$

hence,

$$-w(g, \delta) \leq -|g(s) - g(t)|, \quad \text{for any } g \in C[0, 1] \text{ and any } (s, t) \in D(\delta).$$

Thus, we see that, for any  $f, g \in C[0, 1]$  and any  $(s, t) \in D(\delta)$ , we have

$$\left| |f(s) - f(t)| - w(g, \delta) \right| \leq \left| |f(s) - f(t)| - |g(s) - g(t)| \right| \leq 2 \|f - g\|_\infty.$$

Taking supremum of the left-hand side of the preceding inequality over  $(s, t) \in D(\delta)$  now yields:

$$\begin{aligned} w(f, \delta) - w(g, \delta) &= \sup_{(\xi, \zeta) \in D(\delta)} \left\{ |f(\xi) - f(\zeta)| \right\} - w(g, \delta) \\ &= \sup_{(\xi, \zeta) \in D(\delta)} \left\{ |f(\xi) - f(\zeta)| - w(g, \delta) \right\} \\ &\leq 2 \|f - g\|_\infty. \end{aligned}$$

Interchanging  $f$  and  $g$  in the preceding inequality now yields:

$$\left| w(f, \delta) - w(g, \delta) \right| \leq 2 \|f - g\|_\infty.$$

This completes the proof of (ii).

(iii) This is an immediate consequence of (ii).

□

### Corollary D.3

For any  $\delta \in (0, 1]$ , the map  $w(\cdot, \delta) : C[0, 1] \rightarrow \mathbb{R}$  is an  $\mathbb{R}$ -valued random variable (i.e. an  $\mathbb{R}$ -valued Borel measurable function).

PROOF  $w(\cdot, \delta)$  is continuous by the preceding Theorem, and hence, Borel measurable, by Corollary B.4.  $\square$

### Proposition D.4

For each fixed  $f \in C[0, 1]$ , the map  $w(f, \cdot) : (0, 1] \rightarrow \mathbb{R}$  is a non-decreasing function on  $(0, 1]$ .

PROOF For  $\delta_1, \delta_2 \in (0, 1]$  with  $\delta_1 \leq \delta_2$ , we have

$$D(\delta_1) := \left\{ (s, t) \in [0, 1] \times [0, 1] \mid |s - t| \leq \delta_1 \right\} \subset \left\{ (s, t) \in [0, 1] \times [0, 1] \mid |s - t| \leq \delta_2 \right\} =: D(\delta_2).$$

Hence, for each  $f \in C[0, 1]$  and each  $\delta_1, \delta_2 \in (0, 1)$  with  $\delta_1 \leq \delta_2$ , we have

$$w(f, \delta_1) := \sup_{(s, t) \in D(\delta_1)} \left\{ |f(s) - f(t)| \right\} \leq \sup_{(s, t) \in D(\delta_2)} \left\{ |f(s) - f(t)| \right\} =: w(f, \delta_2).$$

$\square$

### Proposition D.5 (Theorem 7.4, [3])

Suppose:

- $\delta > 0$  and  $0 = t_0 < t_1 < \cdots < t_n = 1$  satisfy:

$$\min_{1 \leq i \leq n} \left\{ t_i - t_{i-1} \right\} \geq \delta.$$

- $C[0, 1]$  is the Banach space of continuous  $\mathbb{R}$ -valued functions defined on  $[0, 1]$  equipped with the supremum norm.

Then, the following statements are true:

- (i) For each  $f \in C[0, 1]$ , we have:

$$w(f, \delta) \leq 3 \cdot \max_{1 \leq i \leq n} \left\{ \sup_{s \in [t_{i-1}, t_i]} |f(s) - f(t_{i-1})| \right\}.$$

- (ii) For each  $\varepsilon > 0$  and each Borel probability measure  $P$  on the (separable) Banach space  $(C[0, 1], \|\cdot\|_\infty)$ , we have:

$$P\left(\left\{ f \in C[0, 1] \mid w(f, \delta) \geq 3\varepsilon \right\}\right) \leq \sum_{i=1}^n P\left(\left\{ f \in C[0, 1] \mid \sup_{s \in [t_{i-1}, t_i]} |f(s) - f(t_{i-1})| \geq \varepsilon \right\}\right).$$

PROOF

- (i) First, note that

$$\min_{1 \leq i \leq n} \left\{ \begin{array}{l} |s - t| \leq \delta \\ t_i - t_{i-1} \geq \delta \end{array} \right\} \implies \left\{ \begin{array}{ll} \text{either } s, t \in [t_{i-1}, t_i], & \text{for some } i \in \{1, 2, \dots, n\} \\ \text{or } s, t \in [t_{i-2}, t_{i-1}] \cup [t_{i-1}, t_i], & \text{for some } i \in \{2, \dots, n\} \end{array} \right.$$

For the case in which both  $s$  and  $t$  lie in the same subinterval  $[t_{i-1}, t_i]$ , for some  $i \in \{1, 2, \dots, n\}$ , we have

$$|f(s) - f(t)| \leq |f(s) - f(t_{i-1})| + |f(t_{i-1}) - f(t)| \leq 2 \cdot \max_{1 \leq i \leq n} \left\{ \sup_{s \in [t_{i-1}, t_i]} |f(s) - f(t_{i-1})| \right\}.$$

For the case in which  $s$  and  $t$  lie in adjacent subintervals, say  $s \in [t_{i-2}, t_{i-1}]$  and  $t \in [t_{i-1}, t_i]$ , for some  $i \in \{2, \dots, n\}$ , we have

$$\begin{aligned} |f(s) - f(t)| &\leq |f(s) - f(t_{i-2})| + |f(t_{i-2}) - f(t_{i-1})| + |f(t_{i-1}) - f(t)| \\ &\leq 3 \cdot \max_{1 \leq i \leq n} \left\{ \sup_{s \in [t_{i-1}, t_i]} |f(s) - f(t_{i-1})| \right\}. \end{aligned}$$

Thus we see that, for any  $f \in C[0, 1]$  and any  $(s, t) \in [0, 1] \times [0, 1]$  with  $|s - t| \leq \delta$ , we have

$$|f(s) - f(t)| \leq 3 \cdot \max_{1 \leq i \leq n} \left\{ \sup_{s \in [t_{i-1}, t_i]} |f(s) - f(t_{i-1})| \right\},$$

which implies, for each  $f \in C[0, 1]$ ,

$$w(f, \delta) := \sup \left\{ |f(s) - f(t)| \mid \begin{array}{c} (s, t) \in [0, 1] \times [0, 1] \\ |s - t| \leq \delta \end{array} \right\} \leq 3 \cdot \max_{1 \leq i \leq n} \left\{ \sup_{s \in [t_{i-1}, t_i]} |f(s) - f(t_{i-1})| \right\}.$$

This proves (i).

(ii) By (i), we see that, for any  $\varepsilon, \delta > 0$  and any  $f \in C[0, 1]$ , we have:

$$3\varepsilon \leq w(f, \delta) \implies \varepsilon \leq \max_{1 \leq i \leq n} \left\{ \sup_{s \in [t_{i-1}, t_i]} |f(s) - f(t_{i-1})| \right\}.$$

For any  $\varepsilon, \delta > 0$  and any  $C[0, 1]$ -valued random variable  $X : (\Omega, \mathcal{A}, \mu) \longrightarrow (C[0, 1], \|\cdot\|_\infty)$ , we have:

$$\begin{aligned} \left\{ f \in C[0, 1] \mid 3\varepsilon \leq w(f, \delta) \right\} &\subset \left\{ f \in C[0, 1] \mid \varepsilon \leq \max_{1 \leq i \leq n} \left\{ \sup_{s \in [t_{i-1}, t_i]} |f(s) - f(t_{i-1})| \right\} \right\} \\ &= \bigcup_{i=1}^n \left\{ f \in C[0, 1] \mid \varepsilon \leq \sup_{s \in [t_{i-1}, t_i]} |f(s) - f(t_{i-1})| \right\}, \end{aligned}$$

and (ii) now follows by sub-additivity of measures.

□

## Corollary D.6 (Corollary of Theorem 7.4, [3])

Suppose  $\{P_n\}_{n \in \mathbb{N}}$  is a sequence of Borel probability measures on  $(C[0, 1], \|\cdot\|_\infty)$ .

Then, (i) implies (ii):

(i) For each  $\varepsilon, \eta > 0$ , there exist  $\delta \in (0, 1)$  and  $n_0 \in \mathbb{N}$  such that

$$\frac{1}{\delta} \cdot P_n \left( \left\{ f \in C[0, 1] \mid \varepsilon \leq \sup_{s \in [t, \min\{1, t+\delta\}]} |f(s) - f(t)| \right\} \right) \leq \eta, \quad \text{for each } t \in [0, 1] \text{ and each } n \geq n_0.$$

(ii) For each  $\varepsilon, \eta > 0$ , there exist  $\delta \in (0, 1)$  and  $n_0 \in \mathbb{N}$  such that

$$P_n \left( \left\{ f \in C[0, 1] \mid \varepsilon \leq w(f, \delta) \right\} \right) \leq \eta, \quad \text{for each } n \geq n_0.$$

PROOF Suppose (i) holds and let  $\varepsilon, \eta > 0$  be given. By (i), there exists  $\delta \in (0, 1)$  and  $n_0 \in \mathbb{N}$  such that

$$\frac{1}{\delta} \cdot P_n \left( \left\{ f \in C[0, 1] \mid \frac{\varepsilon}{3} \leq \sup_{s \in [t, \min\{1, t+\delta\}]} |f(s) - f(t)| \right\} \right) \leq \eta, \quad \text{for each } t \in [0, 1] \text{ and each } n \geq n_0.$$

Now, let  $t_0 = 0$ , and  $t_i = i\delta$ , for  $i = 1, 2, 3, \dots, k := \lfloor 1/\delta \rfloor$ . Then, the preceding inequality and the preceding Theorem together imply:

$$\begin{aligned} P_n \left( \left\{ f \in C[0, 1] \mid \varepsilon \leq w(f, \delta) \right\} \right) &\leq \sum_{i=1}^k P_n \left( \left\{ f \in C[0, 1] \mid \frac{\varepsilon}{3} \leq \sup_{s \in [t_{i-1}, t_i]} |f(s) - f(t_{i-1})| \right\} \right) \\ &\leq k \cdot \delta \cdot \eta = \lfloor 1/\delta \rfloor \cdot \delta \cdot \eta \leq 1 \cdot \eta = \eta, \quad \text{for each } n \geq n_0. \end{aligned}$$

This completes the proof of the Corollary.  $\square$

## E The Arzelà-Ascoli Theorem: compactness of subsets of $C(X)$

Recall that the space  $C(X)$  of continuous  $\mathbb{R}$ -valued functions defined on a compact topological space  $X$  equipped with the supremum norm is a complete metric space (see Theorem 9.3, [1]). The Arzelà-Ascoli Theorem characterizes compactness of subsets of  $C(X)$ .

### Definition E.1 (Equicontinuity)

Let  $X$  be a topological space and  $(Y, d)$  a metric space. Let  $Y^X$  denote the set of arbitrary functions from  $X$  into  $Y$ .

- A subset  $S \subset Y^X$  is said to be **equicontinuous at**  $x_0 \in X$  if, for each  $\varepsilon > 0$ , there exists an open subset  $V \subset X$  satisfying:

$$x_0 \in V, \quad \text{and} \quad \sup_{(x, f) \in V \times S} \left\{ d(f(x), f(x_0)) \right\} \leq \varepsilon.$$

- A subset  $S \subset Y^X$  is said to be **equicontinuous** if it is equicontinuous at each  $x_0 \in X$ .

### Proposition E.2

Let  $X$  be a topological space and  $(Y, d)$  a metric space. Let  $Y^X$  denote the set of arbitrary functions from  $X$  into  $Y$ . Suppose  $x_0 \in X$  and  $S_1, S_2 \subset Y^X$ . Then,

$$\left. \begin{array}{l} S_1 \subset S_2, \text{ and} \\ S_2 \text{ is equicontinuous at } x_0 \end{array} \right\} \implies S_1 \text{ is equicontinuous at } x_0$$

PROOF Let  $\varepsilon > 0$  be given. By the equicontinuity of  $S_2$  at  $x_0$ , there exists open  $V \subset X$  such that

$$x_0 \in V \quad \text{and} \quad \sup_{(x, f) \in V \times S_2} \left\{ d(f(x), f(x_0)) \right\} \leq \varepsilon.$$

However, the hypothesis  $S_1 \subset S_2$  implies that

$$\sup_{(x, f) \in V \times S_1} \left\{ d(f(x), f(x_0)) \right\} \leq \sup_{(x, f) \in V \times S_2} \left\{ d(f(x), f(x_0)) \right\}$$

which immediately implies the following conditions hold:

$$x_0 \in V \quad \text{and} \quad \sup_{(x, f) \in V \times S_1} \left\{ d(f(x), f(x_0)) \right\} \leq \varepsilon.$$

This proves the equicontinuity of  $S_1$  at  $x_0$ , as required.  $\square$



## Definition E.3 (Uniform equicontinuity)

Let  $(X, \rho)$  and  $(Y, d)$  two metric spaces. Let  $Y^X$  denote the set of arbitrary functions from  $X$  into  $Y$ . A subset  $S \subset Y^X$  is said to be **uniformly equicontinuous** if, for each  $\varepsilon > 0$ , there exists  $\delta > 0$  such that:

$$\sup \left\{ d(f(x_1), f(x_2)) \mid \begin{array}{l} f \in S, x_1, x_2 \in X, \\ \rho(x_1, x_2) < \delta \end{array} \right\} \leq \varepsilon.$$

## Proposition E.4

Let  $(X, \rho)$  and  $(Y, d)$  two metric spaces. Let  $Y^X$  denote the set of arbitrary functions from  $X$  into  $Y$ . Then, the following are true:

- (i) Uniform equicontinuity of a subset  $S \subset Y^X$  implies equicontinuity of  $S$ .
- (ii) Suppose furthermore that  $(X, \rho)$  is compact. Then, equicontinuity of a subset  $S \subset Y^X$  implies uniform equicontinuity of  $S$ .

PROOF

- (i) Suppose  $S \subset Y^X$  is uniformly equicontinuous; we seek to prove that  $S$  is also equicontinuous. Let  $x_0 \in X$  and  $\varepsilon > 0$ . By uniform equicontinuity of  $S$ , there exists  $\delta > 0$  such that:

$$\sup \left\{ d(f(x_1), f(x_2)) \mid \begin{array}{l} f \in S, x_1, x_2 \in X, \\ \rho(x_1, x_2) < \delta \end{array} \right\} \leq \varepsilon.$$

Let  $V(x_0) := \{x \in X \mid \rho(x, x_0) < \delta\}$ . Then,  $V(x_0)$  is an open subset of  $X$ , with  $x_0 \in V(x_0)$ , and

$$\sup_{(x, f) \in V(x_0) \times S} \left\{ d(f(x), f(x_0)) \right\} \leq \sup \left\{ d(f(x_1), f(x_2)) \mid \begin{array}{l} f \in S, x_1, x_2 \in X, \\ \rho(x_1, x_2) < \delta \end{array} \right\} \leq \varepsilon.$$

This proves the equicontinuity of  $S$ .

- (ii) Suppose  $(X, \rho)$  is compact and  $S \subset Y^X$  is equicontinuous. Let  $\varepsilon > 0$  be given. By equicontinuity of  $S$ , for each  $x \in X$ , there exists an open ball  $B(x, \delta_x) \subset X$  such that

$$\sup_{(\xi, f) \in B(x, \delta_x) \times S} \left\{ d(f(\xi), f(x)) \right\} \leq \frac{\varepsilon}{2}.$$

Thus,  $X = \bigcup_{x \in X} B(x, \delta_x/2)$  is an open cover of  $X$ . By compactness of  $X$ , this open cover admits a finite subcover:

$$X = \bigcup_{i=1}^n B(x_i, \delta_{x_i}/2).$$

Define  $\delta := \min_{1 \leq i \leq n} \{\delta_{x_i}/2\} > 0$ . Now note the uniform equicontinuity of  $S$  will be established once we prove the validity of the following:

Claim: For any  $\xi_1, \xi_2 \in X$ , and any  $f \in S$ , we have:

$$\rho(\xi_1, \xi_2) < \delta \implies d(f(\xi_1), f(\xi_2)) \leq \varepsilon.$$

Proof of Claim: Suppose  $\rho(\xi_1, \xi_2) < \delta$ . Note that  $\xi_1 \in B(x_i, \delta_{x_i}/2)$ , for some  $i = 1, 2, \dots, n$ . Next, observe that

$$\rho(x_i, \xi_2) \leq \rho(x_i, \xi_1) + \rho(\xi_1, \xi_2) \leq \frac{\delta_{x_i}}{2} + \delta \leq \frac{\delta_{x_i}}{2} + \frac{\delta_{x_i}}{2} = \delta_{x_i}.$$

This shows that both  $\xi_1, \xi_2 \in B(x_i, \delta_{x_i})$ , which implies

$$d(f(\xi_1), f(x_i)) \leq \frac{\varepsilon}{2}, \quad \text{and} \quad d(f(\xi_2), f(x_i)) \leq \frac{\varepsilon}{2}, \quad \text{for each } f \in S,$$

which in turn implies:

$$d(f(\xi_1), f(\xi_2)) \leq d(f(\xi_1), f(x_i)) + d(f(x_i), f(\xi_2)) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \quad \text{for each } f \in S.$$

This completes the proof of the Claim and the uniform equicontinuity of  $S$ . □

## Theorem E.5 (Arzelà-Ascoli, Theorem 9.10, [1])

Suppose  $X$  is a compact topological space and  $C(X)$  is the space of continuous  $\mathbb{R}$ -valued functions defined on  $X$  equipped with the supremum norm. Then, for each  $S \subset C(X)$ , the following conditions are equivalent:

- (i)  $S$  is a compact subset of  $C(X)$ .
- (ii)  $S$  is closed, bounded, and equicontinuous subset of  $C(X)$ .

PROOF

(i)  $\implies$  (ii)

Recall that every compact subset in a metric space is closed and bounded. Thus, it remains only to show that  $S \subset C(X)$  is equicontinuous. To this end, let  $\varepsilon > 0$  be given. Recall that a metric space is compact if and only if it is complete and totally bounded (Theorem 7.8, [1]). Thus, the compactness hypothesis on  $S$  implies  $S$  is totally bounded; in particular, there exist  $f_1, \dots, f_n \in S$  such that

$$S \subset \bigcup_{i=1}^n B\left(f_i, \frac{\varepsilon}{3}\right).$$

Hence, for each  $x_0 \in X$ , we may define

$$V(x_0) := \bigcap_{i=1}^n f_i^{-1}\left(\left(f_i(x_0) - \frac{\varepsilon}{3}, f_i(x_0) + \frac{\varepsilon}{3}\right)\right).$$

Note that  $V(x_0)$  is open and  $x_0 \in V(x_0)$ . Now, let  $f \in S$  and  $x \in V(x_0)$  be given. We may choose  $i \in \{1, 2, \dots, n\}$  such that  $f \in B\left(f_i, \frac{\varepsilon}{3}\right)$ , i.e.  $\|f - f_i\|_\infty \leq \frac{\varepsilon}{3}$ . Hence,

$$\left|f(x) - f(x_0)\right| \leq \left|f(x) - f_i(x)\right| + \left|f_i(x) - f_i(x_0)\right| + \left|f_i(x_0) - f(x_0)\right| \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Thus,

$$\sup_{(x,f) \in V(x_0) \times S} \left\{ \left|f(x) - f(x_0)\right| \right\} \leq \varepsilon.$$

This shows equicontinuity of  $S$  at  $x_0 \in X$ . Since  $x_0 \in X$  is arbitrary, we may conclude that  $S$  is equicontinuous.

(ii)  $\implies$  (i)

Suppose  $S \subset C(X)$  is closed, bounded, and equicontinuous. We need to show that  $S$  is a compact subset of  $C(X)$ . Recall that every subset of a metric space is compact if and only if it is sequentially compact (Theorem 7.3, [1]). Thus, it suffices to show that every sequence  $\{f_n\}_{n \in \mathbb{N}} \subset S$  has a convergent subsequence with limit in  $S$ . We start by stating and proving the following:

Claim 1:

For each  $k \in \mathbb{N}$ , there exists a finite subset  $F_k \subset X$  and open neighbourhoods  $\{V_y\}_{y \in F_k}$  such that

$$X = \bigcup_{y \in F_k} V_y, \quad \text{and} \quad \sup \left\{ \left| f(x) - f(y) \right| \mid \begin{array}{l} x \in V_y, y \in F_k \\ f \in S \end{array} \right\} \leq \frac{1}{3k}.$$

Proof of Claim 1: By equicontinuity of  $S$ , for each  $y \in X$ , there exists an open neighbourhood  $V_y \subset X$  of  $y \in X$  such that

$$\sup \left\{ \left| f(x) - f(y) \right| \mid \begin{array}{l} x \in V_y \\ f \in S \end{array} \right\} \leq \frac{1}{3k}.$$

Thus,  $X = \bigcup_{y \in X} V_y$  is an open cover of  $X$ . Compactness of  $X$  now implies that this open cover of  $X$  admits a finite subcover, i.e.

$$X = \bigcup_{y \in F_k} V_y, \quad \text{for some finite subset } F_k \subset X.$$

Lastly, note that

$$\sup \left\{ \left| f(x) - f(y) \right| \mid \begin{array}{l} x \in V_y, y \in F_k \\ f \in S \end{array} \right\} = \sup_{y \in F_k} \left\{ \sup \left\{ \left| f(x) - f(y) \right| \mid \begin{array}{l} x \in V_y \\ f \in S \end{array} \right\} \right\} \leq \frac{1}{3k}.$$

This completes the proof of Claim 1.

Next, let  $F := \bigcup_{k=1}^{\infty} F_k$ . Note that  $F$  is a countably infinite set. Let  $F = \{x_1, x_2, \dots\}$  be an enumeration of  $F$ . Recall that we wish to prove that every sequence  $\{f_n\}_{n \in \mathbb{N}} \subset S$  contains a convergent subsequence with limit in  $S$ . Now, consider the array of real numbers:

$$\begin{array}{cccc} f_1(x_1) & f_2(x_1) & f_3(x_1) & \cdots \\ f_1(x_2) & f_2(x_2) & f_3(x_2) & \cdots \\ \vdots & \vdots & \vdots & \end{array}$$

Since  $S \subset C(X)$  is bounded (with respect to the  $\|\cdot\|_\infty$  norm on  $C(X)$ ), there exists  $M > 0$  such that  $\sup_{f \in S} \|f\|_\infty \leq M$ .

In particular, every row in the above array is bounded. By Theorem A.14, p.538, [2], there exists an increasing sequence of positive integers  $n(1), n(2), n(3), \dots$  such that the limit

$$\lim_{i \rightarrow \infty} f_{n(i)}(x_k) \text{ exists, for each } k = 1, 2, \dots$$

Claim 2:  $\{f_{n(i)}\}_{i \in \mathbb{N}}$  is a Cauchy sequence in  $(C(X), \|\cdot\|_\infty)$ .

Proof of Claim 2: For each  $k \in \mathbb{N}$ , the convergence of  $\{f_{n(i)}(x_k)\}_{i \in \mathbb{N}}$  in  $\mathbb{R}$  implies that each  $\{f_{n(i)}(x_k)\}_{i \in \mathbb{N}}$  is a Cauchy sequence in  $\mathbb{R}$ . Since the set  $F_k$  is finite, we see that there exists  $m_k \in \mathbb{N}$  such that

$$\left| f_{n(i)}(y) - f_{n(j)}(y) \right| < \frac{1}{3k}, \quad \text{for any } i, j \geq m_k, \text{ and each } y \in F_k.$$

Now, for each  $x \in X$ , there exists  $y \in F_k$  such that  $x \in V_y$ . Hence, for any  $i, j \geq m_k$  and any  $x \in X$ , we have

$$\begin{aligned} \left| f_{n(i)}(x) - f_{n(j)}(x) \right| &\leq \left| f_{n(i)}(x) - f_{n(i)}(y) \right| + \left| f_{n(i)}(y) - f_{n(j)}(y) \right| + \left| f_{n(j)}(y) - f_{n(j)}(x) \right| \\ &\leq \frac{1}{3k} + \frac{1}{3k} + \frac{1}{3k} = \frac{1}{k}. \end{aligned}$$

In other words,

$$\|f_{n(i)} - f_{n(j)}\|_\infty \leq \frac{1}{k}, \quad \text{for any } i, j \geq m_k.$$

This shows that  $\{f_{n(i)}\}_{i \in \mathbb{N}}$  is indeed a Cauchy sequence in  $(C(X), \|\cdot\|_\infty)$  and completes the proof of Claim 2.

Lastly, by Theorem 9.3, [1],  $(C(X), \|\cdot\|_\infty)$  is a complete metric space. Thus, the Cauchy sequence  $\{f_{n(i)}\}_{i \in \mathbb{N}} \subset C(X)$  converges to some element  $f_0 \in C(X)$ . Since  $S \subset C(X)$  is, by hypothesis, a closed subset of  $C(X)$ , we see furthermore that  $f_0 \in S$ . This proves the sequential compactness of  $S$  and completes the proof of the Arzelà-Ascoli Theorem.  $\square$

## Proposition E.6

Suppose  $X$  is a compact topological space and  $C(X)$  is the space of continuous  $\mathbb{R}$ -valued functions defined on  $X$  equipped with the supremum norm. Let  $S \subset C(X)$ .

- (i)  $S$  is equicontinuous at  $x_0 \in X$  if and only if its closure  $\bar{S}$  in  $C(X)$  is equicontinuous at  $x_0$ .
- (ii)  $S$  is equicontinuous if and only if its closure  $\bar{S}$  in  $C(X)$  is equicontinuous.

PROOF It is obvious that (ii) is an immediate consequence of (i). Thus, it suffices to establish (i). First, by Proposition E.2, we immediately see that the equicontinuity of  $\bar{S}$  at  $x_0$  implies the equicontinuity of  $S$  at  $x_0$ . It remains to prove the converse. So, suppose that  $S \subset C(X)$  is equicontinuous at  $x_0 \in X$ . Thus, for each  $\varepsilon > 0$ , there exists an open subset  $V \subset X$  satisfying:

$$x_0 \in V, \quad \text{and} \quad \sup_{(x,f) \in V \times S} \left\{ \left| f(x) - f(x_0) \right| \right\} \leq \varepsilon.$$

Observe that, in order to show the equicontinuity of  $\bar{S}$  at  $x_0$ , it suffices to show that the following inequality is also valid:

$$\sup_{(x,g) \in V \times \bar{S}} \left\{ \left| g(x) - g(x_0) \right| \right\} \leq \varepsilon.$$

To this end, let  $g \in \bar{S} \subset C(X)$ . Then, there exist a sequence  $f_1, f_2, \dots \in S$  such that

$$\lim_{n \rightarrow \infty} \|f_n - g\|_\infty = \lim_{n \rightarrow \infty} \sup_{x \in X} \left\{ \left| f_n(x) - g(x) \right| \right\} = 0.$$

Consequently, for any  $x \in V$  and  $n \in \mathbb{N}$ , we have:

$$\begin{aligned} \left| g(x) - g(x_0) \right| &\leq \left| g(x) - f_n(x) \right| + \left| f_n(x) - f_n(x_0) \right| + \left| f_n(x_0) - g(x_0) \right| \\ &\leq \|g - f_n\|_\infty + \varepsilon + \|f_n - g\|_\infty \longrightarrow 0 + \varepsilon + 0 = \varepsilon. \end{aligned}$$

This implies:

$$\sup_{(x,g) \in V \times \bar{S}} \left\{ \left| g(x) - g(x_0) \right| \right\} \leq \varepsilon,$$

as desired. This completes the proof of the Proposition.  $\square$

## Theorem E.7 (Theorem 7.2, p.81, [3])

Let  $C[0, 1]$  denote the space of continuous  $\mathbb{R}$ -valued functions defined on the closed unit interval  $[0, 1]$  equipped with the supremum norm. Then, for each subset  $S \subset C[0, 1]$ , the following are equivalent:

- (i)  $S$  is a relatively compact subset of  $C[0, 1]$ , i.e. the closure of  $S$  is a compact subset of  $C[0, 1]$ .
- (ii)  $\sup_{f \in S} \left\{ \|f\|_\infty \right\} < \infty$ , and  $S$  is uniformly equicontinuous.

(iii)  $\sup_{f \in S} \{ |f(0)| \} < \infty$ , and for each  $\varepsilon > 0$ , there exists  $\delta > 0$  such that

$$\sup_{f \in S} \{ w(f, \delta) \} = \sup \left\{ \left| f(t_1) - f(t_2) \right| \mid \begin{array}{l} f \in S, t_1, t_2 \in [0, 1], \\ |t_1 - t_2| < \delta \end{array} \right\} \leq \varepsilon.$$

(iv)  $\sup_{f \in S} \{ |f(0)| \} < \infty$ , and  $\lim_{\delta \rightarrow 0^+} \sup_{f \in S} \{ w(f, \delta) \} = 0$ .

PROOF

(i)  $\iff$  (ii)

$$\begin{array}{ll} \text{(i)} \iff \bar{S} \text{ is compact,} & \text{(by definition of relative compactness)} \\ \iff \bar{S} \text{ is bounded and equicontinuous,} & \text{(by the Arzelà-Ascoli Theorem)} \\ \iff S \text{ is bounded and equicontinuous,} & \text{(by Proposition E.6)} \\ \iff S \text{ is bounded and uniformly equicontinuous,} & \text{(by Proposition E.4)} \\ \iff \text{(ii).} & \end{array}$$

(ii)  $\iff$  (iii)

Noting that the second condition in (iii) is precisely uniform equicontinuity of  $S$ , we see immediately that (ii)  $\implies$  (iii). Conversely, we may conclude that (iii)  $\implies$  (ii) once we prove that (iii) implies  $\sup_{f \in S} \{ \|f\|_\infty \} < \infty$ . To this end, take  $\varepsilon = 1$ . Then, by the second condition in (iii) (uniform equicontinuity of  $S$ ), there exists  $\delta > 0$  such that

$$\sup \left\{ \left| f(t_1) - f(t_2) \right| \mid \begin{array}{l} f \in S, t_1, t_2 \in [0, 1], \\ |t_1 - t_2| < \delta \end{array} \right\} \leq \varepsilon := 1.$$

Next, choose  $k \in \mathbb{N}$  sufficiently large such that  $\frac{1}{k} < \delta$ . Hence, for any  $f \in S$  and any  $t \in [0, 1]$ , we have

$$\begin{aligned} |f(t)| &= \left| f(t) - f\left(\frac{k-1}{k} \cdot t\right) + f\left(\frac{k-1}{k} \cdot t\right) - \dots - f\left(\frac{1}{k} \cdot t\right) + f\left(\frac{1}{k} \cdot t\right) - f(0) + f(0) \right| \\ &\leq |f(0)| + \sum_{i=1}^k \left| f\left(\frac{i}{k} \cdot t\right) - f\left(\frac{i-1}{k} \cdot t\right) \right| \leq |f(0)| + k \cdot 1 \\ &\leq \sup_{f \in S} \{ |f(0)| \} + k, \end{aligned}$$

where the last inequality follows from the first condition in (iii). Consequently, we see that, for each  $f \in S$ ,

$$\|f\|_\infty := \sup_{t \in [0, 1]} \{ |f(t)| \} \leq \sup_{f \in S} \{ |f(0)| \} + k < \infty,$$

which in turn implies

$$\sup_{f \in S} \{ \|f\|_\infty \} \leq \sup_{f \in S} \{ |f(0)| \} + k < \infty.$$

This completes the proof that (ii)  $\iff$  (iii).

(iii)  $\iff$  (iv)

This follows trivially from the definition of the right-limit at zero of a  $\mathbb{R}$ -valued function defined on an interval  $[0, \delta_0)$ , for some  $\delta_0 > 0$ .

□

## F Tightness of sequences of Borel measures on $(C[0, 1], \|\cdot\|_\infty)$

### Theorem F.1 (Theorem 7.3, [3])

Suppose  $\{P_n\}_{n \in \mathbb{N}}$  is a sequence of Borel probability measures on  $(C[0, 1], \|\cdot\|_\infty)$ .

Then, the following are equivalent:

- (i)  $\{P_n\}_{n \in \mathbb{N}}$  is tight.
- (ii) For each  $\varepsilon, \eta > 0$ , there exist  $a > 0$  and  $\delta \in (0, 1)$  such that

$$P_n\left(\left\{f \in C[0, 1] \mid |f(0)| \geq a\right\}\right) \leq \eta, \text{ and } P_n\left(\left\{f \in C[0, 1] \mid w(f, \delta) \geq \varepsilon\right\}\right) \leq \eta, \text{ for each } n \in \mathbb{N}.$$

- (iii) For each  $\varepsilon, \eta > 0$ , there exist  $a > 0$ ,  $\delta \in (0, 1)$  and  $n_0 \in \mathbb{N}$  such that

$$P_n\left(\left\{f \in C[0, 1] \mid |f(0)| \geq a\right\}\right) \leq \eta, \text{ and } P_n\left(\left\{f \in C[0, 1] \mid w(f, \delta) \geq \varepsilon\right\}\right) \leq \eta, \text{ for each } n \geq n_0.$$

- (iv) For each  $\eta > 0$ , there exist  $a > 0$  and  $n_0 \in \mathbb{N}$  such that

$$P_n\left(\left\{f \in C[0, 1] \mid |f(0)| \geq a\right\}\right) \leq \eta, \text{ for each } n \geq n_0,$$

and, for each  $\varepsilon > 0$ ,

$$\lim_{\delta \rightarrow 0^+} \limsup_{n \rightarrow \infty} P_n\left(\left\{f \in C[0, 1] \mid w(f, \delta) \geq \varepsilon\right\}\right) = 0.$$

PROOF

(i)  $\implies$  (ii)

Suppose  $\{P_n\}_{n \in \mathbb{N}}$  is tight and let  $\varepsilon, \eta > 0$  be given. By tightness of  $\{P_n\}_{n \in \mathbb{N}}$ , there exists a compact subset  $K \subset C[0, 1]$  such that  $1 - \eta \leq P_n(K) \leq 1$ , for every  $n \in \mathbb{N}$ . By Theorem E.7(iii),

compactness of  $K \subset C[0, 1]$

$$\implies \exists a > 0 \text{ such that } \sup_{f \in K} \left\{|f(0)|\right\} < a < \infty, \text{ and } \exists \delta > 0 \text{ such that } \sup_{f \in K} \left\{w(f, \delta)\right\} < \varepsilon$$

$$\implies \exists a > 0 \text{ such that } K \subset \left\{f \in C[0, 1] \mid |f(0)| < a\right\}, \text{ and } \exists \delta > 0 \text{ such that } K \subset \left\{f \in C[0, 1] \mid w(f, \delta) < \varepsilon\right\}$$

Hence, with  $a > 0$  and  $\delta \in (0, 1)$  as chosen above, we see that, for each  $n \in \mathbb{N}$ , we have:

$$P_n\left(\left\{f \in C[0, 1] \mid |f(0)| \geq a\right\}\right) \leq P_n(C[0, 1] \setminus K) = 1 - P_n(K) \leq \eta,$$

and

$$P_n\left(\left\{f \in C[0, 1] \mid w(f, \delta) \geq \varepsilon\right\}\right) \leq P_n(C[0, 1] \setminus K) = 1 - P_n(K) \leq \eta.$$

This completes the proof that (i)  $\implies$  (ii).

(ii)  $\implies$  (i)

Let  $\eta > 0$  be given. We need to show that  $\{P_n\}_{n \in \mathbb{N}}$  is tight; equivalently, we need to find a compact subset  $K \subset C[0, 1]$  such that  $1 - \eta < P_n(K) \leq 1$ , for each  $n \in \mathbb{N}$ .

By (ii), we may choose  $a > 0$  such that

$$P_n\left(\left\{f \in C[0, 1] \mid |f(0)| \geq a\right\}\right) \leq \frac{\eta}{2}, \text{ for each } n \in \mathbb{N}.$$

Let  $B_0 := \left\{ f \in C[0, 1] \mid |f(0)| < a \right\}$ . Note that  $P(B_0^c) \leq \frac{\eta}{2}$ .

Next, applying (ii) again, for each  $k \in \mathbb{N}$ , we may choose  $\delta_k > 0$  such that

$$P_n \left( \left\{ f \in C[0, 1] \mid w(f, \delta_k) \geq \frac{1}{k} \right\} \right) \leq \frac{\eta}{2^{k+1}}, \quad \text{for each } n \in \mathbb{N}.$$

Let  $B_k := \left\{ f \in C[0, 1] \mid w(f, \delta_k) < \frac{1}{k} \right\}$ , for  $k \in \mathbb{N}$ . Note that  $P(B_k^c) \leq \frac{\eta}{2^{k+1}}$ , for each  $k \in \mathbb{N}$ .

Now, let  $A := \bigcap_{k=0}^{\infty} B_k \subset C[0, 1]$ , and  $K := \overline{A} \subset C[0, 1]$ . Note that the desired implication (ii)  $\implies$  (i) follows from the following two Claims:

**Claim 1:**  $1 - \eta \leq P_n(K) \leq 1$ , for every  $n \in \mathbb{N}$ . In particular,  $K$  is non-empty.

**Claim 2:**  $A$  is a relatively compact subset of  $C[0, 1]$ , hence its closure  $K$  is a compact subset of  $C[0, 1]$ .

Proof of Claim 1: Note that

$$A^c = \left( \bigcap_{k=0}^{\infty} B_k \right)^c = \bigcup_{k=0}^{\infty} B_k^c,$$

which implies, for each  $n \in \mathbb{N}$ ,

$$\begin{aligned} P_n(A^c) &= P_n \left( \bigcup_{k=0}^{\infty} B_k^c \right) \leq \sum_{k=0}^{\infty} P_n(B_k^c) = P_n(B_0^c) + \sum_{k=1}^{\infty} P_n(B_k^c) \\ &\leq \frac{\eta}{2} + \sum_{k=1}^{\infty} \frac{\eta}{2^{k+1}} = \dots = \eta, \end{aligned}$$

which in turn implies

$$1 - \eta \leq 1 - P_n(A^c) = P_n(A) \leq P_n(\overline{A}) = P_n(K), \quad \text{for each } n \in \mathbb{N}.$$

This proves Claim 1.

Proof of Claim 2: Note that  $f \in A \implies f \in B_0 \implies |f(0)| < a < \infty$ . Hence,  $\sup_{f \in A} \{|f(0)|\} \leq a < \infty$ . Secondly, let  $\nu > 0$  be given. Choose  $k \in \mathbb{N}$  such that  $\frac{1}{k} < \nu$ . Then,  $f \in A \implies f \in B_k \implies w(f, \delta_k) < \frac{1}{k}$ , which in turn implies:

$$\sup_{f \in A} \{w(f, \delta_k)\} \leq \frac{1}{k} < \nu.$$

By Theorem E.7(iii) (or, essentially, the Arzelà-Ascoli Theorem),  $A$  is a relatively compact subset of  $C[0, 1]$ . Its closure  $K := \overline{A}$  is therefore a compact subset of  $C[0, 1]$ . This proves Claim 2, and completes the proof of (ii)  $\implies$  (i).

(ii)  $\iff$  (iii)

It is obvious that (ii)  $\implies$  (iii). It remains to prove the reverse implication (iii)  $\implies$  (ii). So, suppose (iii) holds. Let  $\varepsilon, \eta > 0$  be given. By (iii), there exists  $a' > 0$ ,  $\delta' \in (0, 1)$ , and  $n_0 \in \mathbb{N}$  such that

$$P_n \left( \left\{ f \in C[0, 1] \mid |f(0)| \geq a' \right\} \right) \leq \eta, \quad \text{and} \quad P_n \left( \left\{ f \in C[0, 1] \mid w(f, \delta') \geq \varepsilon \right\} \right) \leq \eta, \quad \text{for each } n \geq n_0.$$

On the other hand, since  $(C[0, 1], \|\cdot\|_\infty)$  is separable and complete, every Borel probability measure on  $(C[0, 1], \|\cdot\|_\infty)$  is tight, by Theorem 1.3, p.8, [3]. Thus, by the tightness of each  $P_1, P_2, \dots, P_{n_0-1}$ , and the equivalence (i)  $\iff$  (ii), which we have already established, there exist  $a_1, a_2, \dots, a_{n_0-1} > 0$  and  $\delta_1, \delta_2, \dots, \delta_{n_0-1} \in (0, 1)$  such that

$$P_i\left(\left\{f \in C[0, 1] \mid |f(0)| \geq a_i\right\}\right) \leq \eta, \text{ and } P_i\left(\left\{f \in C[0, 1] \mid w(f, \delta_i) \geq \varepsilon\right\}\right) \leq \eta, \text{ for } i = 1, 2, \dots, n_0 - 1.$$

Now, let  $a := \max\{a', a_1, a_2, \dots, a_{n_0-1}\}$  and  $\delta := \min\{\delta', \delta_1, \delta_2, \dots, \delta_{n_0-1}\}$ . Then,

$$\begin{aligned} P_n\left(\left\{f \in C[0, 1] \mid |f(0)| \geq a\right\}\right) &\leq P_n\left(\left\{f \in C[0, 1] \mid |f(0)| \geq a'\right\}\right) \leq \eta, \text{ for } n \geq n_0, \text{ and} \\ P_n\left(\left\{f \in C[0, 1] \mid |f(0)| \geq a\right\}\right) &\leq P_n\left(\left\{f \in C[0, 1] \mid |f(0)| \geq a_n\right\}\right) \leq \eta, \text{ for } n = 1, 2, \dots, n_0 - 1, \end{aligned}$$

which together imply

$$P_n\left(\left\{f \in C[0, 1] \mid |f(0)| \geq a\right\}\right) \leq \eta, \text{ for each } n \in \mathbb{N}.$$

Next, recall that for each fixed  $f \in C[0, 1]$ , the map  $\delta \mapsto w(f, \delta)$  is non-decreasing in  $\delta$ , by Proposition D.4. We thus see that:

$$\begin{aligned} P_n\left(\left\{f \in C[0, 1] \mid w(f, \delta) \geq \varepsilon\right\}\right) &\leq P_n\left(\left\{f \in C[0, 1] \mid w(f, \delta') \geq \varepsilon\right\}\right) \leq \eta, \text{ for } n \geq n_0, \text{ and} \\ P_n\left(\left\{f \in C[0, 1] \mid w(f, \delta) \geq \varepsilon\right\}\right) &\leq P_n\left(\left\{f \in C[0, 1] \mid w(f, \delta_n) \geq \varepsilon\right\}\right) \leq \eta, \text{ for } n = 1, 2, \dots, n_0 - 1, \end{aligned}$$

which together imply

$$P_n\left(\left\{f \in C[0, 1] \mid w(f, \delta) \geq \varepsilon\right\}\right) \leq \eta, \text{ for each } n \in \mathbb{N}.$$

This proves (iii)  $\implies$  (ii), and completes the proof of the equivalence (ii)  $\iff$  (iii).

(iii)  $\implies$  (iv)

(iii)  $\iff$  for each  $\varepsilon, \eta > 0$ , there exist  $a > 0, \delta \in (0, 1)$  and  $n_0 \in \mathbb{N}$  such that

$$P_n\left(\left\{f \in C[0, 1] \mid |f(0)| \geq a\right\}\right) \leq \eta, \text{ and } P_n\left(\left\{f \in C[0, 1] \mid w(f, \delta) \geq \varepsilon\right\}\right) \leq \eta, \text{ for each } n \geq n_0$$

$$\iff \begin{aligned} &\text{for each } \eta > 0, \exists a > 0 \text{ and } n_0 \in \mathbb{N} \text{ such that } P_n\left(\left\{f \in C[0, 1] \mid |f(0)| \geq a\right\}\right) \leq \eta, \text{ for each } n \geq n_0, \text{ and} \\ &\text{for each } \varepsilon, \eta > 0, \exists \delta \in (0, 1) \text{ such that } \limsup_{n \rightarrow \infty} P_n\left(\left\{f \in C[0, 1] \mid w(f, \delta) \geq \varepsilon\right\}\right) \leq \eta \end{aligned}$$

$$\iff \begin{aligned} &\text{for each } \eta > 0, \exists a > 0 \text{ and } n_0 \in \mathbb{N} \text{ such that } P_n\left(\left\{f \in C[0, 1] \mid |f(0)| \geq a\right\}\right) \leq \eta, \text{ for each } n \geq n_0, \text{ and} \\ &\text{for each } \varepsilon > 0, \lim_{\delta \rightarrow 0^+} \limsup_{n \rightarrow \infty} P_n\left(\left\{f \in C[0, 1] \mid w(f, \delta) \geq \varepsilon\right\}\right) = 0 \end{aligned}$$

$$\iff \text{(iv)}$$

The proof of the Theorem is now complete. □

## G Weak convergence of Borel probability measures on $(C[0, 1], \|\cdot\|_\infty)$

### Theorem G.1



Suppose  $P, P_1, P_2, \dots$  are Borel probability measures on  $(C[0, 1], \|\cdot\|_\infty)$ . If

$$P_n \circ \pi_{t_1 t_2 \dots t_k}^{-1} \xrightarrow{d} P \circ \pi_{t_1 t_2 \dots t_k}^{-1}, \quad \text{as } n \rightarrow \infty, \quad \text{for any } t_1, t_2, \dots, t_k \in [0, 1],$$

then the limit of every weakly convergent subsequence of  $\{P_n\}$  must be  $P$ .

PROOF Suppose  $P_{n(i)} \xrightarrow{d} Q$ , as  $i \rightarrow \infty$ , where  $Q$  is some Borel probability measure on  $(C[0, 1], \|\cdot\|_\infty)$ . We need to show that  $Q$  in fact must be  $P$ . Since, for any  $t_1, t_2, \dots, t_k \in [0, 1]$ , the map  $\pi_{t_1 t_2 \dots t_k} : C[0, 1] \rightarrow \mathbb{R}^k : f \mapsto (f(t_1), \dots, f(t_k))$  is continuous, it follows that  $P_{n(i)} \circ \pi_{t_1 t_2 \dots t_k}^{-1} \xrightarrow{d} Q \circ \pi_{t_1 t_2 \dots t_k}^{-1}$ , as  $i \rightarrow \infty$ , for any  $t_1, t_2, \dots, t_k \in [0, 1]$ , by the Continuous Mapping Theorem (Theorem 2.7, [3], or simply see remark on p.20, [3]). Then, for each bounded continuous function  $\varphi : \mathbb{R}^k \rightarrow \mathbb{R}$ , we have

$$\int_{\mathbb{R}^k} \varphi d(Q \circ \pi_{t_1 \dots t_k}^{-1}) = \lim_{i \rightarrow \infty} \int_{\mathbb{R}^k} \varphi d(P_{n(i)} \circ \pi_{t_1 \dots t_k}^{-1}) = \lim_{n \rightarrow \infty} \int_{\mathbb{R}^k} \varphi d(P_n \circ \pi_{t_1 \dots t_k}^{-1}) = \int_{\mathbb{R}^k} \varphi d(P \circ \pi_{t_1 \dots t_k}^{-1}),$$

where the first and third equalities follow directly from the definition of weak convergence of probability measures, while the second equality follows from the elementary fact that every subsequence of a convergent sequence of real numbers converges to the same limit as the full sequence. By Theorem 1.2, p.8, [3], we see that  $Q \circ \pi_{t_1 \dots t_k}^{-1} = P \circ \pi_{t_1 \dots t_k}^{-1}$ , as Borel measures on  $\mathbb{R}^k$ , for any  $t_1, t_2, \dots, t_k \in [0, 1]$ . This in turn implies that  $Q \circ \pi_{t_1 \dots t_k}^{-1}(B) = P \circ \pi_{t_1 \dots t_k}^{-1}(B)$ , for every Borel subset  $B \subset \mathbb{R}^k$ . In other words,  $Q$  and  $P$  agree on the collection of finite-dimensional subsets of  $C[0, 1]$ . Since the finite-dimensional subsets of  $C[0, 1]$  form a separating class (Example 1.3, p.11, [3]), we may now conclude that  $Q = P$ . This completes the proof of the Theorem.  $\square$

## Theorem G.2 (Example 5.1, p.57, [3])

Suppose  $P, P_1, P_2, \dots$  are Borel probability measures on  $(C[0, 1], \|\cdot\|_\infty)$ . If

- (i) the sequence  $\{P_n\}_{n \in \mathbb{N}}$  of Borel probability measures on  $(C[0, 1], \|\cdot\|_\infty)$  is relatively compact, and
- (ii) for any  $t_1, t_2, \dots, t_k \in [0, 1]$ , we have

$$P_n \circ \pi_{t_1 t_2 \dots t_k}^{-1} \xrightarrow{d} P \circ \pi_{t_1 t_2 \dots t_k}^{-1}, \quad \text{as } n \rightarrow \infty,$$

then  $P_n \xrightarrow{d} P$ , as  $n \rightarrow \infty$ .

PROOF Recall that, by Theorem 2.6, p.20, [3],  $P_n \xrightarrow{d} P$  if and only if every subsequence of  $\{P_n\}$  contains a further subsequence that weakly converges to  $P$ . So, let  $\{P_{n(i)}\}_{i \in \mathbb{N}}$  be a subsequence of  $\{P_n\}_{n \in \mathbb{N}}$ . By hypothesis (i) (i.e. relative compactness of  $\{P_n\}_{n \in \mathbb{N}}$ ), the subsequence  $\{P_{n(i)}\}_{i \in \mathbb{N}}$  contains a weakly convergent further subsequence  $\{P_{n(i_m)}\}_{m \in \mathbb{N}}$ , say  $P_{n(i_m)} \xrightarrow{d} Q$ , as  $m \rightarrow \infty$ , where  $Q$  is some Borel probability measure on  $(C[0, 1], \|\cdot\|_\infty)$ . By hypothesis (ii) and Theorem G.1, we see that in fact  $Q = P$ . Thus, we have shown that every subsequence  $\{P_{n(i)}\}_{i \in \mathbb{N}}$  of  $\{P_n\}_{n \in \mathbb{N}}$  contains a further subsequence  $P_{n(i_m)}$  which weakly converges to  $P$ . By Theorem 2.6, p.20, [3], we may now conclude that the full original sequence  $P_n$  converges weakly to  $P$ . This completes the proof of the Theorem.  $\square$

## Theorem G.3 (Theorem 7.5, p.84, [3])

Suppose:

- $(\Omega, \mathcal{A}, \mu), (\Omega^{(1)}, \mathcal{A}^{(1)}, \mu^{(1)}), (\Omega^{(2)}, \mathcal{A}^{(2)}, \mu^{(2)}), \dots$  are probability spaces.
- $X : \Omega \rightarrow C[0, 1]$  is a  $C[0, 1]$ -valued random variable, and  $\left\{X^{(n)} : \Omega^{(n)} \rightarrow C[0, 1]\right\}_{n \in \mathbb{N}}$  is a sequence of  $C[0, 1]$ -valued random variables.

If

(i) for any  $t_1, t_2, \dots, t_k \in [0, 1]$ ,

$$\left( X_{t_1}^{(n)}, X_{t_2}^{(n)}, \dots, X_{t_k}^{(n)} \right) \xrightarrow{d} \left( X_{t_1}, X_{t_2}, \dots, X_{t_k} \right), \text{ as } n \rightarrow \infty, \text{ and}$$

(ii) for each  $\varepsilon > 0$ ,

$$\lim_{\delta \rightarrow 0^+} \limsup_{n \rightarrow \infty} P\left(w(X^{(n)}, \delta) \geq \varepsilon\right) = \lim_{\delta \rightarrow 0^+} \limsup_{n \rightarrow \infty} \mu^{(n)}\left(\left\{\omega \in \Omega^{(n)} \mid w(X^{(n)}(\omega), \delta) \geq \varepsilon\right\}\right) = 0,$$

then

$$X^{(n)} \xrightarrow{d} X, \text{ as } n \rightarrow \infty.$$

**PROOF** Note that, in order to prove the Theorem, it suffices to establish tightness of  $\{X^{(n)}\}_{n \in \mathbb{N}}$ . This is because tightness implies relative compactness (Prokhorov's Theorem, Theorem 5.1, p.59, [3]), which, together with hypothesis (i) and Theorem G.2, will imply the desired weak convergence  $X^{(n)} \xrightarrow{d} X$ . We use Theorem F.1(iv) to establish the tightness of  $\{X^{(n)}\}_{n \in \mathbb{N}}$ . Thus, the proof the present Theorem will be complete once we prove the following two claims:

**Claim 1:**  $\{X^{(n)}\}_{n \in \mathbb{N}}$  satisfies the first condition in Theorem F.1(iv).

In other words, for each  $\eta > 0$ , there exist  $a > 0$  and  $n_0 \in \mathbb{N}$  such that

$$P_{X^{(n)}}\left(\left\{f \in C[0, 1] \mid |f(0)| \geq a\right\}\right) \leq \eta, \text{ for each } n \geq n_0.$$

**Claim 2:**  $\{X^{(n)}\}_{n \in \mathbb{N}}$  satisfies the second condition in Theorem F.1(iv).

In other words, for each  $\varepsilon > 0$ ,

$$\lim_{\delta \rightarrow 0^+} \limsup_{n \rightarrow \infty} P_{X^{(n)}}\left(\left\{f \in C[0, 1] \mid w(f, \delta) \geq \varepsilon\right\}\right) = 0.$$

**Proof of Claim 1:** By hypothesis (i),  $X_0^{(n)} \xrightarrow{d} X_0$ , which is equivalent to the statement that every subsequence of  $\{X_0^{(n)}\}_{n \in \mathbb{N}}$  admits a further subsequence which weakly converges to  $X_0$  (Theorem 2.6, p.20, [3]). In particular, weak convergence of  $\{X_0^{(n)}\}_{n \in \mathbb{N}}$  implies relative compactness of  $\{X_0^{(n)}\}_{n \in \mathbb{N}}$ . Since each  $X_0^{(n)}$  induces a Borel probability measure on  $\mathbb{R}$ , and  $\mathbb{R}$  is separable and complete, Prokhorov's Theorem (Theorem 5.2, p.60, [3]) implies that  $\{X_0^{(n)}\}_{n \in \mathbb{N}}$  is tight (as a sequence of Borel probability measures on  $\mathbb{R}$ ). Thus, for any  $\eta > 0$ , there exists a compact set  $K \subset \mathbb{R}$  such that

$$1 - \eta < P_{X_0^{(n)}}(K) = P(X_0^{(n)} \in K) \leq 1, \text{ for each } n \in \mathbb{N}.$$

Since  $K \subset \mathbb{R}$  is compact, it is bounded, and thus there exists  $a > 0$  such that  $K \subset (-a, a)$ . Thus, we see that:

$$1 - \eta \leq P(X_0^{(n)} \in K) \leq P(|X_0^{(n)}| < a) = 1 - P(|X_0^{(n)}| \geq a), \text{ for each } n \in \mathbb{N},$$

which implies that

$$P_{X^{(n)}}\left(\left\{f \in C[0, 1] \mid |f(0)| \geq a\right\}\right) = P(|X_0^{(n)}| \geq a) \leq \eta, \text{ for each } n \in \mathbb{N}.$$

This completes the proof of Claim 1.

**Proof of Claim 2:** Claim 2 holds precisely by hypothesis (ii), since

$$P_{X^{(n)}}\left(\left\{f \in C[0, 1] \mid w(f, \delta) \geq \varepsilon\right\}\right) = P(w(X^{(n)}, \delta) \geq \varepsilon).$$

This completes the proof of Claim 2, as well as that of the present Theorem. □

## H Etemadi's inequality

### Theorem H.1 (Appendix M19, [3])

Suppose  $X_1, \dots, X_n : \Omega \longrightarrow \mathbb{R}$  are independent  $\mathbb{R}$ -valued random variables defined on the common probability space  $\Omega$ .

Let  $S_k := \sum_{i=1}^k X_i$ , for  $k = 1, 2, \dots, n$ . Then, for any  $x > 0$ ,

$$P\left(\left\{\omega \in \Omega \mid \max_{1 \leq k \leq n} |S_k(\omega)| \geq 3x\right\}\right) \leq 3 \cdot \max_{1 \leq k \leq n} P\left(\left\{\omega \in \Omega \mid |S_k(\omega)| \geq x\right\}\right).$$

PROOF Define:

$$\begin{aligned} A &:= \left\{\omega \in \Omega \mid \max_{1 \leq k \leq n} |S_k(\omega)| \geq 3x\right\}, \\ A_1 &:= \left\{\omega \in \Omega \mid |S_1(\omega)| \geq 3x\right\}, \\ A_k &:= \left\{\omega \in \Omega \mid \max_{1 \leq i \leq k-1} |S_i(\omega)| < 3x \text{ and } |S_k(\omega)| \geq 3x\right\}, \text{ for } k = 2, 3, \dots, n. \end{aligned}$$

Note that  $A$  can be expressed as the following disjoint union:  $A = \bigsqcup_{k=1}^n A_k$ .

**Claim 1:**  $A_k \cap \left\{|S_n| < x\right\} \subset A_k \cap \left\{|S_n - S_k| > 2x\right\}$ , for each  $k = 1, 2, \dots, n$ .

**Claim 2:**  $\left\{|S_n - S_k| > 2x\right\} \subset \left\{|S_n| > x\right\} \cup \left\{|S_k| > x\right\}$ , for each  $k = 1, 2, \dots, n$ .

Proof of Claim 1:

$$\begin{aligned} &\omega \in A_k \cap \left\{|S_n| < x\right\} \\ \iff &\max_{1 \leq i \leq k-1} |S_i(\omega)| < 3x \text{ and } |S_k(\omega)| \geq 3x \text{ and } |S_n(\omega)| < x \\ \implies &\max_{1 \leq i \leq k-1} |S_i(\omega)| < 3x \text{ and } |S_k(\omega)| \geq 3x \text{ and } |S_k(\omega) - S_n(\omega)| \geq |S_k(\omega)| - |S_n(\omega)| > 3x - x = 2x \\ \implies &\omega \in A_k \cap \left\{|S_n - S_k| > 2x\right\} \end{aligned}$$

This proves Claim 1.

Proof of Claim 2: Simply note that for any two real numbers  $a, b \in \mathbb{R}$ , we have:

$$|a| \leq x \text{ and } |b| \leq x \implies |a - b| \leq |a| + |b| \leq x + x = 2x,$$

whose contrapositive is:

$$|a - b| > 2x, \implies |a| > x \text{ or } |b| > x.$$

This proves Claim 2.

Using the above two Claims, we now see that:

$$\begin{aligned}
 P(A) &= P\left(A \cap \{|S_n| \geq x\}\right) + P\left(A \cap \{|S_n| < x\}\right) \\
 &\leq P\left(|S_n| \geq x\right) + P\left(A \cap \{|S_n| < x\}\right) \\
 &= P\left(|S_n| \geq x\right) + \sum_{k=1}^n P\left(A_k \cap \{|S_n| < x\}\right) \\
 &\leq P\left(|S_n| \geq x\right) + \sum_{k=1}^n P\left(A_k \cap \{|S_n - S_k| > 2x\}\right), \quad \text{by Claim 1} \\
 &= P\left(|S_n| \geq x\right) + \sum_{k=1}^n P(A_k) \cdot P\left(|S_n - S_k| > 2x\right), \quad \text{by independence of summands} \\
 &\leq P\left(|S_n| \geq x\right) + \left(\max_{1 \leq k \leq n} P\left(|S_n - S_k| > 2x\right)\right) \cdot \sum_{k=1}^n P(A_k) \\
 &= P\left(|S_n| \geq x\right) + P(A) \cdot \max_{1 \leq k \leq n} P\left(|S_n - S_k| > 2x\right) \\
 &= P\left(|S_n| \geq x\right) + \max_{1 \leq k \leq n} P\left(|S_n - S_k| > 2x\right) \\
 &\leq P\left(|S_n| \geq x\right) + \max_{1 \leq k \leq n} \left\{ P\left(|S_n| > x\right) + P\left(|S_k| > x\right) \right\}, \quad \text{by Claim 2} \\
 &\leq 3 \cdot \max_{1 \leq k \leq n} \left\{ P\left(|S_k| \geq x\right) \right\}.
 \end{aligned}$$

This completes the proof of Etemadi's inequality. □

## References

- [1] ALIPRANTIS, C. D., AND BURKINSHAW, O. *Principles of Real Analysis*, third ed. Academic Press, 1998.
- [2] BILLINGSLEY, P. *Probability and Measure*, third ed. John Wiley & Sons, 1995.
- [3] BILLINGSLEY, P. *Convergence of Probability Measures*, second ed. John Wiley & Sons, 1999.
- [4] FERGUSON, T. S. *A Course in Large Sample Theory*, first ed. Texts in Statistical Science. CRC Press, 1996.
- [5] SHAO, J. *Mathematical Statistics*, second ed. Springer, 2003.