# BOCHNER INTEGRATION AND AN APPLICATION IN STOCHASTIC PROCESSES

CHUHUAN HUANG

## Contents



## 1. Introduction

In this note, we discuss the Bochner integration and an application in showing the distributions of $\int_0^T X_t dt$ and $\int_0^T Y_t dt$ coincide, given two progressively measurable stochastic processes $X, Y$.

*Date*: April, 2023.

## 2. Bochner Integration

Given a measure space $(X, \Sigma, \mu)$ and a real Banach Space $(V, |\cdot|_V)$, we say a $(\Sigma, \mathscr{B}_V)$ measurable function $s : X \to V$ is ***simple*** if

$$s(x) = \sum_{i=1}^{n} \mathbb{1}_{E_i}(x) v_i,$$

where $\{E_i\} \subseteq \Sigma$ are disjoint, and $v_i$ are distinct elements in $V$.

We say such $s$ is ***integrable***, and its integral is defined as

$$\int_X s d\mu \coloneqq \sum_{i=1}^{n} \mu(E_i) v_i,$$

if $\mu(E_i)$ is finite whenever $v_i \neq 0$.

A $(\Sigma, \mathscr{B}_V)$ measurable function $f : X \to V$ is said to be ***Bochner integrable***, if there exists a sequence of integrable simple functions $\{s_n\}_{n=1}^{\infty}$ s.t.

$$\lim_{n\to\infty} \int_X |f - s_n|_V d\mu = 0,$$

and its integral is defined as

$$\int_X f d\mu \coloneqq \lim_{n\to\infty} \int_X s_n d\mu,$$

where the first integral is in the sense of Lebesgue integral and the second convergence is in the sense of $|\cdot|_V$ norm topology.

**Remark 2.1.** we make several natural remarks here:

(1) It's easy to show that $\{\int_X s_n d\mu\}_{n=1}^{\infty}$ is a Cauchy sequence in the Banach space $V$, and hence by its completeness the limit exists. If $f$ takes value in a general normed space, we may have trouble that this Cauchy sequence does not necessary to converge.
(2) It's also easy to show that if there is another such sequence of simple functions, the integral of $f$ does not change. That is, the Bochner integral is well-defined (independent of choices).
(3) Moreover, $f$ is Bochner integrable if and only if $\int_X |f|_V d\mu < \infty$, and we may denote $\mathcal{L}^1(X; V)$ to be the set of all such functions. It is easy to see $\mathcal{L}^1(X; V)$ is a vector space (linearity is here!) and $\int_X |\cdot|_V d\mu$ is a semi-norm on it, and hence by using the standard trick: taking the quotient space w.r.t the equivalence relation $f \sim g$ if $f = g$ $\mu$-a.e.,

$$L^1(X; V) \coloneqq \mathcal{L}^1(X; V)/\sim$$

is a Banach space equiped with the norm $|\cdot|_{L^1(X;V)} \coloneqq \int_X |\cdot|_V d\mu$.
(4) From our definition, simple functions are naturally dense in $L^1(X; V)$, of course w.r.t. the $|\cdot|_{L^1(X;V)}$ norm topology.

(5) If $f$ is Bochner integrable, then the following usual inequlity holds:

$$|\int_E f d\mu|_V \leq \int_E |f|_V d\mu,$$

for all $E \in \Sigma$, where $\int_E f d\mu := \int_X f \mathbb{1}_E d\mu$ as usual.

(6) Typically, Suppose $T : V \to V'$ is a continous linear operator between Banach space $V$ and $V'$, and $f$ is Bochner integrable. Then,
- $Tf : X \to V'$ is also Bochner integrable;
- the Bochner integral commutes with $T$:

$$\int_X Tf d\mu = T \int_X f d\mu$$

(7) Lastly, given a Bochner integrable function $f$, for all $E \in \Sigma$, $\nu : E \mapsto \int_E f d\mu$ is a countably-additive $V$-valued measure on $X$, and $\nu \ll \mu$.

## 2.1. **Dominated Convergence Theorem for Banach-valued functions.**

More importantly, we still have our beloved *Dominated Convergence Theorem*, perhaps after a redefinition w.r.t to a completion of the measure $\mu$.

**Theorem 2.2** (Dominated Convergence Theorem for Banach-valued functions)**.** *Suppose $\{f_n\}_{n=1}^{\infty} \subseteq L^1(X;V)$, $f_n \to f$ $\mu$-a.e., and there $\exists g \in L^1(X,\Sigma,\mu)$ s.t. $|f_n(x)|_V \leq g(x)$ for $\mu$-a.e. $x \in X$. Then, $f \in L^1(X;V)$,*

$$f_n \to f \text{ in } L^1(X;V),$$

*and hence*

$$\int_X f d\mu = \lim_{n\to\infty} \int_X f_n d\mu,$$

*where the second convergence is in the sense of $|\cdot|_V$ norm topology.*

The proof is clear by applying the usual Dominated Convergence Theorem [Folland Theorem 2.24] to $|f_n - f|_V \leq 2g$ and observe

$$|\int_X f d\mu - \int_X f_n d\mu|_V \leq \int_X |f_n - f|_V d\mu \to 0.$$

## 2.2. **Luzin's Theorem for Banach-valued functions.**

Now we have sufficient ingredients, it is time to present the main result: *Luzin's Theorem* for Banach-valued functions.

**Theorem 2.3** (Luzin's Theorem for Banach-valued functions)**.** *If $f : [a,b] \to V$ is $(\mathcal{L}_{[a,b]}, \mathscr{B}_V)$-measurable ($\mathcal{L}_{[a,b]}$ is the class of all Lebesgue measurable sets included in $[a,b]$) and $\varepsilon > 0$, then there is a compact set $K_\varepsilon \subseteq [a,b]$ s.t. the Lebesgue measure $m(K_\varepsilon^c) < \varepsilon$ and $f|_K$ is continuous.*

To prove this version of Luzin's, we give several natural lemmas[1].

**Lemma 2.4.** *If $f_n \to f$ in $L^1(X;V)$, there is a subsequence $\{f_{n_j}\}$ s.t. $f_{n_j} \to f$ a.e.*

**Lemma 2.5** (Egorov's for Banach-valued functions)**.** *Suppose $\mu(X) < \infty$, and $\{f_n\}, f$ are $(\Sigma, \mathscr{B}_V)$-measurable functions on $X$ s.t. $f_n \to f$ a.e. Then, $\forall \varepsilon > 0, \exists E_\varepsilon \subseteq X$ s.t. $\mu(E) < \varepsilon$ and $f_n \to f$ uniformly on $E^c$.*

**Lemma 2.6.** *If $\mu$ is a Lebesgue-Stieltjes measure on $\mathbb{R}$, then compactly supported continuous functions $f : \mathbb{R} \to V$ are dense in $L^1(\mathbb{R}; V)$, with respect to the $|\cdot|_{L^1(\mathbb{R};V)}$ norm topology.*

Now we could prove the Luzin's in the usual way

*Proof.* We first assume $f$ is bounded, i.e. $\sup_{x\in[a,b]} |f(x)|_V < \infty$, and then we combine lemma 2.4 and lemma 2.6 that there is a sequence of compactly supported continuous functions $\{f_n\}$ s.t. $f_n \to f$ a.e. Then, as $m([a,b]) = b - a < \infty$, we could apply Egorov's (lemma 2.5), for a given $\varepsilon > 0$, $\exists E_\varepsilon \subseteq [a,b]$ s.t. $m(E_\varepsilon) < \varepsilon$ and $f_n \to f$ uniformly on $E_\varepsilon^c$. That is, as the uniform limit of continuous functions on $E_\varepsilon^c$, $f$ is continuous on $E_\varepsilon^c$. Finally, recall Lebesgue measurable sets could be approached from inside by compact sets. That is, as $m(E_\varepsilon^c) > b-a-\varepsilon$, $b-a-\varepsilon$ is not the sup and hence $\exists$ compact $K_\varepsilon \subseteq E_\varepsilon^c$ s.t. $m(K_\varepsilon) > b-a-\varepsilon$. $K_\varepsilon$ is what we are looking for.

If $f$ is unbounded, we can consider the sets $E_n = [a,b] \cap \{x : |f(x)|_V \le n\}$ for $n \in \mathbb{N}$, which is measurable and have finite measure since $[a,b]$ has finite measure. Then, we can apply the previous arguments to each $f|_{E_n}$, which is a bounded function, to obtain a sequence of compact $K_n \subseteq E_n$ s.t.

$$m(K_n^c) < \frac{\varepsilon}{2^n},$$

and $f|_{K_n}$ is continuous. Then, typically, $f|_{\cap_{n\in\mathbb{N}} K_n}$ is continuous and $\cap_{n\in\mathbb{N}} K_n$ is still compact, with

$$m(\cup_{n\in\mathbb{N}} K_n^c) \le \sum_{n\in\mathbb{N}} m(K_n^c) < \varepsilon.$$

□

## 3. An application

### 3.1. Stochastic Analysis Preliminaries.

Recall that, given a filtered probability space $(\Omega, \mathscr{F}, \mathbb{F} := \{\mathscr{F}_t\}_{0\le t\le T}, \mathbb{P})$

- A *stochastic process* is a real-valued[2] function $X : [0,T] \times \Omega \to \mathbb{R}$ that is $(\mathscr{B}_{[0,T]} \times \mathscr{F}, \mathscr{B}_\mathbb{R})$ measurable.

[1]We attached the proofs in the appendix.

[2]We use the real-valued functions for simplicity, but similar notion could be developed for complex-valued functions.

- For given fixed $t \in [0,T]$, we view $X_t := X(t,\cdot) : \Omega \to \mathbb{R}$ as a *random variable* (simply a $(\mathscr{F}, \mathscr{B}_{\mathbb{R}})$-measurable function). Equivalently, we could view $X$ as $X : [0,T] \to L^0(\Omega)$ with $X(t) := X_t$, where $L^0(\Omega)$ is the space of all random variables $Z : \Omega \to \mathbb{R}$. For $p \in [0,\infty]$, $L^p(\Omega)$ are as usual.
- A stochastic process $X$ is *progressively measurable* if its restriction $X|_{[0,t]} : [0,t] \times \Omega \to \mathbb{R}$ is $(\mathscr{B}_{[0,t]} \times \mathscr{F}_t, \mathscr{B}_{\mathbb{R}})$ measurable for all $t \in [0,T]$, denoted as $X \in \mathbb{L}^0(\mathbb{F})$. For $p \in [1,\infty]$, $\mathbb{L}^p(\mathbb{F}) := \mathbb{L}^0(\mathbb{F}) \cap L^p([0,T] \times \Omega)$ are as usual.
- The *distribution* of a random variable $Z$ is a Borel measure, denoted as $\mu_Z : \mathscr{B}_{\mathbb{R}} \to [0,1]$, is defined as
$$\mu_Z(A) := \mathbb{P}(Z \in A).$$
We say such two random variables $Z_1, Z_2$ have the same distribution, denoted as $Z_1 \overset{(d)}{=} Z_2$, if $\mu_{Z_1} = \mu_{Z_2}$.
- The *finite distribution* of a stochastic process $X$ is a collection of Borel measures
$$\{\mu_{X_{t_1},...,X_{t_n}}\}_{\text{all } 0 \le t_1 < t_2 < ... < t_n \le T},$$
where
$$\mu_{X_{t_1},...,X_{t_n}}(A) := \mathbb{P}\Big((X_{t_1},...,X_{t_n}) \in A\Big)$$
for all $A \in \mathscr{B}_{\mathbb{R}^n}$.
- Moreover, we say such two stochastic processes $X, Y$ have the same finite distribution, denoted as $X \overset{(d)}{=} Y$, if $\mu_{X_{t_1},...,X_{t_n}} = \mu_{Y_{t_1},...,Y_{t_n}}$, for any $0 \le t_1 < t_2 < ... < t_n \le T$.

3.2. **Main Proposition.**

We want to show that

**Proposition 3.1.** Given $(\Omega, \mathscr{F}, \mathbb{F} := \{\mathscr{F}_t\}_{0 \le t \le T}, \mathbb{P})$ and $X, Y \in \mathbb{L}^0(\mathbb{F})$, if $X \overset{(d)}{=} Y$, then $\int_0^T X_t dt \overset{(d)}{=} \int_0^T Y_t dt$.

We first give out a lemma.

**Lemma 3.2.** *Suppose $X, Y \in \mathbb{L}^0(\mathbb{F})$ and $X \overset{(d)}{=} Y$. If $X$ is continuous in $t$ for a.e. $\omega$, then $Y$ is also continuous in $t$ a.e. $\omega$.*

*Proof of the Proposition 3.1.*

Assume $X, Y$ bounded (i.e. $X, Y \in \mathbb{L}^\infty(\mathbb{F})$ and equivalently $X, Y : [0,T] \to L^\infty(\Omega)$.) and $X$ continuous. By lemma 3.2, $Y$ is also continuous. Then, consider partition of $\pi_n : \{t_i := \frac{i}{2^n}T\}$, $i = 0, ..., 2^n$ and
$$X_t^{\pi_n} := \sum_{i=0}^{n-1} X_{t_i} \mathbb{1}_{\{t_i \le t < t_{i+1}\}}$$
and
$$Y_t^{\pi_n} := \sum_{i=0}^{n-1} Y_{t_i} \mathbb{1}_{\{t_i \le t < t_{i+1}\}}.$$

It is clear, as $X, Y$ has same finite distribution, that

$$\int_0^T X_t^{\pi_n} dt = \sum_{i=0}^{n-1} X_{t_i} \Delta t_{i+1} \overset{(d)}{=} \sum_{i=0}^{n-1} Y_{t_i} \Delta t_{i+1} = \int_0^T Y_t^{\pi_n} dt.$$

Notice $X^{\pi_n} \to X$ pointwisely in $t$, $Y^{\pi_n} \to Y$ in $t$ pointwisely, and boundedness of $X, Y$, by dominated convergence theorem (Theorem 2.2), we know

$$\int_0^T X_t dt = \lim_{n\to\infty} \int_0^T X_t^{\pi_n} dt \overset{(d)}{=} \lim_{n\to\infty} \int_0^T Y_t^{\pi_n} dt = \int_0^T Y_t dt.$$

To see the $\overset{(d)}{=}$ above, denote $Z_n \coloneqq \int_0^T X_t^{\pi_n} dt$, $\tilde{Z}_n \coloneqq \int_0^T Y_t^{\pi_n} dt$, $Z \coloneqq \int_0^T X_t dt$ and $\tilde{Z} \coloneqq \int_0^T Y_t dt$. By Folland 6.8(c), after taking a union of $\mathbb{P}$-null sets, we have $Z_n \to Z, \tilde{Z}_n \to \tilde{Z}$ uniformly on $\Omega \backslash E$ with $\mathbb{P}(E) = 0$. For a given closed $A \subseteq \mathbb{R}$, we denote a *$\varepsilon$-perturbation* of $A$, $A_\varepsilon \coloneqq \{x \in \mathbb{R} : |x - y| < \varepsilon \text{ for some } y \in A\}$. we **claim**:

$$\bigcap_{\varepsilon>0} \bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} \{\omega \notin E : Z_n(\omega) \in A_\varepsilon\} = \{\omega \notin E : Z(\omega) \in A\}$$

and similarly $\bigcap_{\varepsilon>0} \bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} \{\omega \notin E : \tilde{Z}_n(\omega) \in A_\varepsilon\} = \{\omega \notin E : \tilde{Z}(\omega) \in A\}$. Then, as $Z_n \overset{(d)}{=} \tilde{Z}_n$,

$$\mathbb{P}\Big(\{\omega \notin E : Z(\omega) \in A\}\Big) = \mathbb{P}\Big(\{\omega \notin E : \tilde{Z}(\omega) \in A\}\Big),$$

and we are done as $\mathscr{B}_{\mathbb{R}}$ is generated by closed sets.

To see the claim, $\supseteq$ direction is clear by the uniform convergence on $E^c$. For the other direction, let $\omega \in \bigcap_{\varepsilon>0} \bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} \{\omega \notin E : Z_n(\omega) \in A_\varepsilon\}$. For a given $\varepsilon/2$, $\exists N_1$ s.t. $Z_n(\omega) \in A_{\varepsilon/2}$ for all $n \geq N_1$. Pick a $y_\varepsilon$ s.t.

$$|Z_n(\omega) - y_\varepsilon| < \frac{\varepsilon}{2}.$$

Also, the uniform convergence on $E^c$ allows us to choose $N_2$ s.t.

$$|Z(\omega) - Z_n(\omega)| < \frac{\varepsilon}{2}$$

for all $n \geq N_2$. We choose $N = N_1 \vee N_2$, and hence we have for all $\varepsilon, \exists y_\varepsilon \in A$

$$|Z(\omega) - y_\varepsilon| \leq |Z(\omega) - Z_n(\omega)| + |Z_n(\omega) - y_\varepsilon| < \varepsilon,$$

i.e. $Z(\omega)$ is a limit point of $A$ and hence $Z(\omega) \in A$ as $A$ is closed.

Now for general bounded $X, Y$, we apply the Luzin's (Theorem 2.3), and, after taking a intersection of two compact sets, we may assume that $\forall \delta > 0$, $\exists K_\delta \subseteq [0, T]$ s.t. $m(K_\delta^c) < \delta$, and $X|_{K_\delta}$, $Y|_{K_\delta}$ are continuous. Apply the same discretization and perturbation technique, we have

$$\int_{K_\delta} X_t dt = \lim_{n\to\infty} \int_{K_\delta} X_t^{\pi_n} dt \overset{(d)}{=} \lim_{n\to\infty} \int_{K_\delta} Y_t^{\pi_n} dt = \int_{K_\delta} Y_t dt.$$

Then, we choose $\delta_m \coloneqq \frac{1}{m|X|_\infty} \wedge \frac{1}{m|Y|_\infty}$, denote $K_m \coloneqq K_{\delta_m}$, and hence we have

$$\int_{K_m^c} |X_t|_\infty dt \leq |X|_\infty m(K_m^c) < |X|_\infty \delta_m \leq 1/m,$$

and observe that

$$|\int_0^T X_t dt - \int_{K_m} X_t dt|_\infty = |\int_{K_m^c} X_t dt|_\infty \leq \int_{K_m^c} |X_t|_\infty dt < \frac{1}{m},$$

i.e. $\int_{K_m} X_t dt \to \int_0^T X_t dt$ uniformly except on a $\mathbb{P}$-null set $E$. Similarly, after taking a union of $\mathbb{P}$-null sets, we may say $\int_{K_m} Y_t dt \to \int_0^T Y_t dt$ uniformly except on the same $\mathbb{P}$-null set $E$. Then, we further apply the perturbation technique on $\int_{K_m} X_t dt, \int_0^T X_t dt$ and we may conclude $\int_0^T X_t dt \overset{(d)}{=} \int_0^T Y_t dt$.

Finally, for general $X, Y$, we follow the standard space truncation $X^{(n)} := X \wedge n$. It gives an a.e. convergence and hence allows us to follow the same perturbation procedure. $\square$

**Remark 3.3.** In fact this is a trivial application, as all we need to apply the perturbation technique is the a.e. convergence, which could be achieved by using $\mathbb{P}$-a.e. defined $(\int_0^T X_t dt)(\omega) := \int_0^T X(t,\omega)dt$ from the Fubini's. Anyway, it is good to realize the integration for vector-valued functions.

## 4. Reference


K. Yosida, *Functional analysis*, Springer, Ch. 8, §1., 1980

G.B. Folland. *Real Analysis: Modern Techniques and Their Applications*, Second Edition, 1999.

W. Rudin. *Functional analysis*. New York: McGraw-Hill, 1973


## 5. Appendix

To prove Lemma 2.4, we add the following generalized notion of convergence in measure and lemma.

We say that a sequence of $(\Sigma, \mathscr{B}_V)$-measurable functions $\{f_n\}$ on $(X, \Sigma, \mu)$

- ***converges in measure*** to $f$ if, for every $\varepsilon > 0$, we have
$$\lim_{n\to\infty} \mu(\{x : |f_n(x) - f(x)|_V \geq \varepsilon\}) = 0.$$
- ***Cauchy in measure*** if, for every $\varepsilon > 0$, we have
$$\mu(\{x : |f_n(x) - f_m(x)|_V \geq \varepsilon\}) \to 0$$
as $n, m \to \infty$.

**Lemma 5.1.** *Suppose $\{f_n\}$ is Cauchy in measure, then there is $(\Sigma, \mathscr{B}_V)$-measurable function $f$ s.t. $f_n \to f$ in measure, and there is a subsequence $f_{n_j}$ that $f_{n_j} \to f$ a.e. Moreover, if $f_n \to g$ in measure, then $f = g$ a.e.*

**Corollary 5.2.** *If $f_n \to f$ in measure, then there is a subsequence $\{f_{n_j}\}$ s.t. $f_{n_j} \to f$ a.e.*

**Lemma 5.3.** *if $f_n \to f$ in $L^1(X;V)$, then $f_n \to f$ in measure.*

Together, lemma 2.4 is proved.

*Proof of Lemma 5.1.* Since $\forall \varepsilon > 0$, $\lim_{n,m\to\infty} \mu(\{x : |f_n(x) - f_m(x)|_V \geq \varepsilon\}) = 0$, we choose $n_1$ large enough to be the first index s.t. $\mu(\{x : |f_n(x) - f_{n+1}(x)|_V \geq 2^{-1}\}) < 2^{-1}$. Similarly, We could choose $\{g_j\} = \{f_{n_j}\}$ s.t. if $E_j \coloneqq \{x : |g_j(x) - g_{j+1}(x)|_V \geq 2^{-j}\}$, then $\mu(E_j) \leq 2^{-j}$. Denote $F_k \coloneqq \bigcup_{j\geq k} E_j$, and hence $\mu(F_k) \leq \sum_{j=k}^{\infty} 2^{-j} = 2^{1-k}$, and if $x \notin F_k$, i,e. $x \in \bigcap_{j\geq k} E_j^c$ for $i \geq j \geq k$, we have

$$|g_i(x) - g_j(x)|_V \leq \sum_{l=j}^{i-1} |g_l(x) - g_{l+1}(x)|_V \leq \sum_{l=j}^{i-1} 2^{-l} < 2^{1-j}.$$

Consider $x \in \bigcup_{k=1}^{\infty} F_k^c$, then $x \in F_k^c$ for some $k$, and in here we have $|g_i(x) - g_j(x)|_V < 2^{1-j}$ for all $i \geq j \geq k$. Pick $K \geq k$ s.t. $2^{1-K} < \varepsilon$, and then, whenever $i \geq j \geq K$

$$|g_i(x) - g_j(x)|_V < 2^{1-j} \leq 2^{1-K} < \varepsilon,$$

i.e. $\{g_j\}$ pointwise Cauchy in $\bigcup_{k=1}^{\infty} F_k^c$. Observe $\mu(\Big(\bigcup_{k=1}^{\infty} F_k^c\Big)^c) = \mu(\bigcap_{k=1}^{\infty} F_k) = \lim_{k\to\infty} \mu(F_k) = 0$, and hence we know $\{g_j\}$ pointwise Cauchy almost everywhere. Define

$$f(x) \coloneqq \begin{cases} \lim_{j\to\infty} g_j(x) & \text{if } x \in \bigcup_{k=1}^{\infty} F_k^c \\ \mathbf{0} & \text{otherwise} \end{cases}.$$

It is easy to see $f$ is $(\Sigma, \mathscr{B}_V)$-measurable and hence $f_{n_j} = g_j \to f$ a.e.

We left to show $f_n \to f$ in measure. To see this, observe by triangle inequality

$$\{x : |f_n(x) - f(x)|_V \geq \varepsilon\} \subseteq \{x : |f_n(x) - f_{n_j}(x)|_V \geq \frac{\varepsilon}{2}\} \cup \{x : |f_{n_j}(x) - f(x)|_V \geq \frac{\varepsilon}{2}\},$$

where the later two sets have small measure when $n, j \to \infty$, one given by convergence in measure and the other given by the a.e. pointwise convergence.

If $f_n \to g$ in measure, similarly, observe that

$$\{x : |f(x) - g(x)|_V \geq \varepsilon\} \subseteq \{x : |f(x) - f_n(x)|_V \geq \frac{\varepsilon}{2}\} \cup \{x : |f_n(x) - g(x)|_V \geq \frac{\varepsilon}{2}\},$$

i.e. $\mu(\{x : |f(x) - g(x)|_V \geq \varepsilon\}) = 0$ for all $\varepsilon > 0$. Note

$$\{x : |f(x) - g(x)|_V > 0\} = \bigcup_{m\in\mathbb{N}} \{x : |f(x) - g(x)|_V \geq \frac{1}{m}\},$$

we conclude $f = g$ a.e. by the continuity from below. $\square$

*Proof of Lemma 5.2.*

Let $E_{n,\varepsilon} \coloneqq \{x : |f_n(x) - f(x)|_V \geq \varepsilon\}$, then

$$\varepsilon\mu(E_{n,\varepsilon}) \leq \int_{E_{n,\varepsilon}} |f_n - f| d\mu \leq \int_X |f_n - f| d\mu \to 0$$

as $n \to \infty$. $\square$

*Proof of Lemma 2.5.*

Let $f_n \to f$ pointwise on $K$ with $\mu(K^c) = 0$. Note

$$K = \bigcap_{k=1}^{\infty} \bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} \{x : |f_n(x) - f(x)|_V < \frac{1}{k}\}.$$

That is, for any fixed $k$, $K(k) \coloneqq \bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} \{x : |f_n(x) - f(x)|_V < \frac{1}{k}\}$ has full measure. For $N \in \mathbb{N}$, denote

$$E_N(k) \coloneqq \bigcup_{n\geq N} \{x : |f_n(x) - f(x)|_V \geq \frac{1}{k}\}.$$

Observe $E_N(k)$ decreases as $N$ increases, and

$$K(k) = \bigcup_{N\in\mathbb{N}} E_N(k)^c = \Big(\bigcap_{N\in\mathbb{N}} E_N(k)\Big)^c.$$

Thus, $F(k) \coloneqq \bigcap_{N\in\mathbb{N}} E_N(k)$ is a null set. By continuity from above, we may choose $N_k$ s.t. $\mu(E_{N_k}(k)) < \varepsilon 2^{-k}$. Denote $E \coloneqq \bigcup_{k=1}^{\infty} E_{N_k}(k)$, so we have $\mu(E) \leq \sum_{k=1}^{\infty} \mu(E_{N_k}(k)) < \varepsilon$.

On $E^c$, observe $E^c \subseteq E_{N_k}(k)^c$ for any $k$.Therefore, For any $\varepsilon' > 0$, there is a $k$ s.t. $\frac{1}{k} < \varepsilon$, and for $\forall x \in E^c \subseteq E_{N_k}(k)^c$, $|f_n(x) - f(x)|_V < \frac{1}{k} < \varepsilon'$ whenever $n \geq N_k$, i.e. $f_n \to f$ uniformly on $E^c$. $\square$

*Proof of Lemma 2.6.*

By linearity and denseness of simple functions in $L^1(\mathbb{R}; V)$, it suffice to show any simple function of form $s(x) \coloneqq \mathbb{1}_E(x)\mathbf{1}$ (we abuse the term to say it is a *indicator*) could approximated arbitrarily close in $|\cdot|_{L^1(X;V)}$ by a compactly supported continuous function, where $E \in \mathcal{L}_{\mathbb{R}}$ Lebesgue-Stieltjes measurable set, and $\mathbf{1}$ is a unit vector in $V$.

Recall that $E$ could be approximated by a finite union of open intervals $\{I_i\}_{i=1}^n$ s.t. $\mu(E \triangle \cup_{i=1}^n I_i) < \varepsilon$. Note

$$\int_X |\mathbb{1}_E \mathbf{1} - \mathbb{1}_{\cup_{i=1}^n I_i} \mathbf{1}|_V d\mu = \int_X |\mathbb{1}_E - \mathbb{1}_{\cup_{i=1}^n I_i}| d\mu = \mu(E \triangle \cup_{i=1}^n I_i) < \varepsilon. \tag{1}$$

So we know the indicator of any $E$ could be approximated arbitrarily by a indicator of finite union of open intervals.

Meanwhile, for the indicator of an open interval $\mathbb{1}_{(a,b)}\mathbf{1}$, we could approximate by constructing

$$f_\varepsilon(x) \coloneqq \begin{cases} \mathbf{1} & \text{if } x \in [a+\frac{\varepsilon}{2}, b-\frac{\varepsilon}{2}] \\ Linear & \text{if } x \in [a-\frac{\varepsilon}{2}, a+\frac{\varepsilon}{2}] \cup [b-\frac{\varepsilon}{2}, b+\frac{\varepsilon}{2}] \\ \mathbf{0} & \text{otherwise} \end{cases}.$$

□