A BRIEF INTRODUCTION TO LEBESGUE–STIELTJES INTEGRAL

SHIU-TANG LI

Abstract. In this article, we first show the the reader how to construct Lebesgue–Stieltjes measure, which is used to construct Lebesgue–Stieltjes integral. We then talk about some properties of Lebesgue–Stieltjes integral.

Contents

1. Introduction

In this article we will show how to construct Lebesgue–Stieltjes measure. To define an integral, we first have to specify what measure it is using; the ideas can be found in a lot of real analysis books. We will then talk about the properties of Lebesgue–Stieltjes integral.

We assume the reader is familiar with Lebesgue integral, knowing what Borel sets are, in order to follow the materials.

2. Abbreviations and notations used in this article

RC: right-continuous.
BV: bounded variation, finite variation.
$A_t = A(t)$: a real-valued function A with a single variable t.
$A_{t-} = \lim_{s\to t-} A(s)$: the left limit of A at t.
dA: the Lebesgue–Stieltjes measure associated with A.

Date: February 18, 2017.

$\mathscr{B}([0,a])$: the Borel sets on the space $[0,a]$.
$\otimes$: product measure.

3. Construction of Lebesgue–Stieltjes measure

Definition 1. *Let $A : [0,a] \to \mathbb{R}$ be an RC function of finite variation. We define an finitely additive set function μ on $([0,a], \mathscr{F})$ by $\mu([0,t]) := A_t$ for any $0 \leq t \leq a$, where $\mathscr{F}$ is the collection of any finite union of sets of form $\{0\}$ or $(c,d]$.*

Theorem 2. (Caratheodory Extension Theorem) *There exists a unique signed measure μ^* on $([0,a], \mathscr{B}([0,a]))$ so that $\mu^*(E) = \mu(E)$ for all $E \in \mathscr{F}$, where μ is defined in Definition 1.*

Proof. Decompose $A = A^1 - A^2$, where A^1, A^2 are both RC increasing functions. Write $\mu = \mu^1 - \mu^2$, where $\mu^1([0,t]) = A^1_t$ and $\mu^2([0,t]) = A^2_t$, and it's easily seen that μ^1, μ^2 are two finitely additive set functions on $([0,a], \mathscr{F})$. By the proof of Theorem 1.2 in [3], for any $D_n \in \mathscr{F}$, $D_n \downarrow \varnothing$, we have $\mu^1(D_n) \downarrow 0$ and $\mu^2(D_n) \downarrow 0$. Therefore, by Theorem 1.1 in [3], there exists two positive measures $(\mu^1)^*$ and $(\mu^2)^*$ on $([0,a], \mathscr{B}([0,a]))$, so that $(\mu^1)^*(E) = \mu^1(E)$ and $(\mu^2)^*(E) = \mu^2(E)$ for all $E \in \mathscr{F}$. We then define $\mu^* := (\mu^1)^* - (\mu^2)^*$, and the proof of existence is complete.

To prove the uniqueness of μ^*, we note that for any other measure ν^* which is another extension of μ on $\mathscr{F}$, $\{E : \nu^*(E) = \mu^*(E)\}$ is a λ-system, and $\mathscr{F}$ is a π-system, by Dynkin's $\pi - \lambda$ theorem, $\nu^*(E) = \mu^*(E)$ for all $E \in \mathscr{B}([0,a])$.

$\square$

Definition 3. *The measure μ^* constructed in Theorem 2 is called the Lebesgue-Stieltjes measure associated with A.*

4. Definition of Lebesgue–Stieltjes integral

We first consider the case that A is increasing on $[0,a]$. It is a routine job to define $\int_{[0,a]} f(s)\,dA_s$ for f simple, f bounded Borel measurable, and then f positive Borel measurable as is done in other measure theory textbooks. For arbitrary Borel measurable function f, we decompose it into $f^+ - f^-$, and define $\int_{[0,a]} f(s)\,dA_s := \int_{[0,a]} f^+(s)\,dA_s - \int_{[0,a]} f^-(s)\,dA_s$ if at least one of these two integrals is finite.

Given any RC, BV function A on $[0,a]$, we write $A = A^1 - A^2$, where A^1, A^2 are both RC increasing functions on $[0,a]$. The integral $\int_{[0,a]} f(s)\,dA_s := \int_{[0,a]} f(s)\,dA^1_s - \int_{[0,a]} f(s)\,dA^2_s$, where we require at least one of these two integrals to be finite.

5. Total variation of Lebesgue–Stieltjes measure

5.1. **A quick review on decomposition of measures.** This subsection recalls the reader the definition of signed measure, and its decomposition theorems. The main reference here is Chap 2.10 in [4].

Definition 4. (Definition 2.10.1, [4]) *Given a measurable space $(X, \mathscr{F})$. A set function λ on $\mathscr{F}$ is called a signed measure on $\mathscr{F}$ if it satisfies the following conditions:*
(1) $\lambda(E) \in (-\infty, \infty]$ for every $E \in \mathscr{F}$ or $\lambda(E) \in [-\infty, \infty)$ for every $E \in \mathscr{F}$.
(2) $\lambda(\varnothing) = 0$.
(3) countable additivity: for every disjoint sequence $\{E_n : n \in \mathbb{N}\}$ in $\mathscr{F}$, $\sum_{n=1}^{\infty} \lambda(E_n)$ exists in $\overline{\mathbb{R}}$ and $\sum_{n=0}^{\infty} \lambda(E_n) = \lambda(\bigcup_{n=1}^{\infty} E_n)$.

If λ is a signed measure on $\mathscr{F}$, the triple $(X, \mathscr{F}, \lambda)$ is called a signed measure space.

Theorem 5. (Hahn Decomposition of Signed Measure Spaces) (Theorem 2.10.14, [4]) *For an arbitrary signed measure space $(X, \mathscr{F}, \lambda)$, a Hahn decomposition exists and is unique up to null sets of λ, that is, there exist a positive set P and a negative set N for λ such that $P \cap N = \varnothing$ and $P \cup N = X$, and moreover if P' and N' are another such pair, then $P \triangle P'$ and $N \triangle N'$ are null sets for λ.*

Theorem 6. (Jordan Decomposition of Signed Measures) (Theorem 2.10.21, [4]) *Given a signed measure space $(X, \mathscr{F}, \lambda)$. A Jordan decomposition for $(X, \mathscr{F}, \lambda)$ exists and is unique, that is, there exists a unique pair $\{\mu, \nu\}$ of positive measures on $(X, \mathscr{F})$, at least one of which is finite, such that $\mu \perp \nu$ and $\lambda = \mu - \nu$. Moreover with an arbitrary Hahn decomposition $\{P, N\}$ of $(X, \mathscr{F}, \lambda)$, if we define two set functions μ and ν on $\mathscr{F}$ by setting $\mu(E) = \lambda(E \cap P)$ and $\nu(E) = -\lambda(E \cap N)$ for all $E \in \mathscr{F}$, then $\{\mu, \nu\}$ is a Jordan decomposition for $(X, \mathscr{F}, \lambda)$.*

Definition 7. (Definition 2.10.22, [4]) *Given a signed measure space $(X, \mathscr{F}, \lambda)$. Let μ and ν be the unique positive measures on $\mathscr{F}$, at least one of which is finite, such that $\mu \perp \nu$ and $\lambda = \mu - \nu$. Let us call μ and ν the positive and negative parts of λ and write λ^+ for μ and λ^- for ν. The total variation of X is a positive measure $|\lambda|$ on $\mathscr{F}$ defined by $|\lambda|(E) = \lambda^+(E) + \lambda^-(E)$ for $E \in \mathscr{F}$.*

5.2. **Characterization of the total variation of Lebesgue–Stieltjes measure.** The following theorem shows the connection between the variation of signed measures and the variation of a BV function.

Theorem 8. *Let $V : [0, a] \to \mathbb{R}$, $V(t)$ be the variation of A on $[0, t]$. We also assume that $A_0 = 0$. Let dV be the Lebesgue–Stieltjes measure associated with V, and $|dA|$ be the total variation of measure dA, where dA is the Lebesgue–Stieltjes measure associated with A. We claim that $dV = |dA|$, as a measure on $([0, a], \mathscr{B}([0, a]))$.*

Proof. $|dA| \leq dV$: we first notice that $|dA(s, t]| = |A(t) - A(s)| \leq V[s, t] = V(t) - V(s) = dV(s, t]$ for all intervals $(s, t] \subset [0, a]$. By Dynkin's $\pi - \lambda$ theorem, $|dA(B)| \leq dV(B)$ for any $B \in \mathscr{B}([0, a])$. Now, let $\{P, N\}$ be the Hahn decomposition of $([0, a], \mathscr{B}([0, a]), A)$. For any $E \in \mathscr{B}([0, a])$, we have $|dA|(B) = dA(B \cap P) - dA(B \cap N) \leq dV(B \cap P) + dV(B \cap N) = dV(B)$.

$dV \leq |dA|$: since $dV(s, t] = V(t) - V(s) = V[s, t] = \sup_{\triangle} \sum |A_{t_i} - A_{t_{i-1}}| \leq \sup_{\triangle} \sum |dA|(t_i - t_{i-1}] = |dA|(s, t]$. The rest of the proof follows from Dynkin's $\pi - \lambda$ theorem. $\square$

6. Properties of Lebesgue–Stieltjes integral

Throughout this section we, $\int_0^t f(s)\,dA_s := \int_{(0,t]} f(s)\,dA_s$.

6.1. **Conversion between Lebesgue–Stieltjes integral and Lebesgue integral.**

As the reader can find in Section 1.3 in [1], when A is RC increasing on $[0,a]$, we may decompose it into a convex combination of three different increasing functions: a RC discrete increasing function, a singular continuous increasing function (not identically zero but with zero derivatives a.e.), and an absolutely continuous increasing function.

This implies the Lebesgue-Stieltjes measure associated with RC function A, dA_s, can be decomposed into three parts. It is quite difficult to compute the Lebesgue-Stieltjes integral when A is singular continuous. On the other hand, when A is absolutely continuous, we have the following result:

Theorem 9. *Let A be absolutely continuous, and let f be a bounded Borel measurable function on $[0,a]$. Then $\int_0^a f(s)\,dA_s = \int_0^a f(s)A'_s\,ds$, where A'_t is the a.e. derivative of A_t.*

Proof. Show the identity holds for simple functions first, then use the functional monotone class theorem to show it holds for Borel measurable functions as well. □

6.2. **Other properties of Lebesgue–Stieltjes integral.**

Theorem 10. (Right continuity) *Let $\int_0^a |f(s)|\,|dA_s| < \infty$, where $f \in \mathscr{B}([0,a])$. Then $g(t) := \int_0^t f(s)\,dA_s$ is RC on $(0,a]$.*

Proof. The right-continuity of g follows from the dominated convergence theorem. □

Theorem 11. *Let $\int_0^a |f(s)|\,|dA_s| < \infty$, where $f \in \mathscr{B}([0,a])$. Then $g(t) := \int_0^t f(s)\,dA_s$ is of BV on $[0,a]$.*

Proof. One can decompose $f = f^+ - f^-$, $f^+, f^- \geq 0$, and $dA = \mu + \nu$, where ν is a negative measure. $\int_0^t f(s)\,dA_s = \Big(\int_0^t f^+(s)\mu(s) - \int_0^t f^-(s)\nu(s)\Big) - \Big(-\int_0^t f^+(s)\nu(s) + \int_0^t f^-(s)\mu(s)\Big)$. So $g(t)$ is the difference of two increasing functions on $[0,a]$. □

Theorem 12. (Associativity) *Let f, g be as above. Let $h \in \mathscr{B}([0,a])$ so that $\int_0^a |h(s)|\,|dg_s| < \infty$ or $\int_0^a |h(s)f(s)|\,|dA_s| < \infty$. Then $\int_0^a h(s)\,dg_s = \int_0^a h(s)f(s)\,dA_s$.*

Proof. First, it is easily seen that the identity holds for $h(s) = 1_{(a,b]}(s)$. By the functional monotone class theorem, the identity holds for all bounded Borel measurable h's.

Let $\{P, N\}$ be the Hahn decomposition of dA. This implies $\{P', N'\}$ is the Hahn decomposition of dg, where $P' = (\{f \geq 0\} \cap P) \cup (\{f < 0\} \cap N)$, $N' = (\{f \geq 0\} \cap N) \cup (\{f < 0\} \cap P)$. The decomposition of dg follows from taking $h = 1_{A \cap P'}$ and $h = 1_{A \cap N'}$.

For arbitrary $h \in \mathscr{B}([0,a])$ so that $\int_0^a |h(s)|\,|dg_s| < \infty$ or $\int_0^a |h(s)f(s)|\,|dA_s| < \infty$, we write $h = h^+ - h^- = h^+ 1_{P'} + h^+ 1_{P'} - h^- 1_{N'} - h^- 1_{N'}$. We then approximate each component with bounded Borel measurable functions, using the functional monotone class theorem. □

Remarks 13. *(1) When f is continuous on $[0,a]$, f of BV on $[0,a]$, or f bounded Borel measurable, then $f \in \mathscr{B}([0,a])$ and $\int_0^a |f(s)|\,|dA_s| < \infty$. (2) If A is of BV on $[0,a]$, then A_-, the function obtained by replacing A_t with A_{t-} for every $t \in [0,a]$, is of BV on $[0,a]$, thanks to the Jordan decomposition for BV functions.*

The next two theorems are taken from Chap 0 in [2]. More detailed proofs are given here.

Theorem 14. (Integration by parts) *Let A, B be two functions of finite variation. Then for any $t > 0$, $A_tB_t = A_0B_0 + \int_0^t A_{s-}\,dB_s + \int_0^t B_s\,dA_s = A_0B_0 + \int_0^t A_{s-}\,dB_s + \int_0^t B_{s-}\,dA_s + \sum_{0<s\le t}(A_s - A_{s-})\cdot(B_s - B_{s-})$.*

Proof.

$$
\begin{aligned}
A_tB_t &= \left(\int_{[0,t]} dA_s\right)\cdot\left(\int_{[0,t]} dB_s\right)\\
&= \int_{[0,t]\times[0,t]} dA_x \otimes dB_y\\
&= \int_{\{0\}\times\{0\}} dA_x \otimes dB_y + \int_{A_1} dA_x \otimes dB_y + \int_{A_2} dA_x \otimes dB_y + \int_{A_3} dA_x \otimes dB_y,
\end{aligned}
$$

where $A_1 = \{(x,y) : 0 \le x,y \le t, x < y\}$, $A_2 = \{(x,y) : 0 \le x,y \le t, x > y\}$, $A_3 = \{(x,y) : 0 \le x,y \le t, x = y, (x,y) \ne (0,0)\}$.

It's easily seen that $\int_{\{0\}\times\{0\}} dA_x \otimes dB_y = A_0B_0$, by Fubini's theorem. By Fubini's theorem again, we have

$$
\begin{aligned}
\int_{A_1} dA_x \otimes dB_y &= \int_{(0,t]} \left(\int_{[0,y)} dA_x\right) dB_y\\
&= \int_{(0,t]} A_{y-}\,dB_y.
\end{aligned}
$$

We have used the fact that $A_t = dA([0,t])$, $A_{t-} = \lim_{s\to t-} dA([0,s]) = dA([0,t))$. Similar calculations for $\int_{A_2} dA_x \otimes dB_y$. For the last term, we have

$$
\begin{aligned}
\int_{A_3} dA_x \otimes dB_y &= \int_{(0,t]} \left(\int_{\{y\}} dA_x\right) dB_y\\
&= \int_{(0,t]} A_y - A_{y-}\,dB_y\\
&= \sum_{0<y\le t} (A_y - A_{y-})\cdot(B_y - B_{y-}).
\end{aligned}
$$

The last line is due to the fact that $A_y - A_{y-}$ is nonzero for only countably many y's. In this case we may rewrite the integral as summation. $\square$

Theorem 15. *If F is a C^1-function and A is of finite variation, then $F(A)$ is of finite variation and*

$F(A_t) = F(A_0) + \int_0^t F'(A_{s-})dA_s + \sum_{0<s\le t}\big(F(A_s) - F(A_{s-}) - F'(A_{s-})(A_s - A_{s-})\big)$.

Proof. 1. The first assertion follows from $\sum_\triangle |F(A_{t_i}) - F(A_{t_{i-1}})| \le \sum_\triangle |F'(\xi_i)|\cdot|A_{t_i} - A_{t_{i-1}}| \le M\cdot\sum_\triangle |A_{t_i} - A_{t_{i-1}}|$.

2. It's easily seen that the identity holds for $F \equiv c$, and if F_1, F_2 both satisfy the identity, so does $F_1 + cF_2$.

Step 1. We'd like to show the identity holds for $F(x) = x^n$, using induction. First we perform integration by parts formula on $(A_t)^{n-1}$ and A_t, and we have

$$\begin{aligned}(A_t)^n =&(A_0)^n + \int_0^t (A_{s-})^{n-1}\,dA_s + \int_0^t A_{s-}\,d(A_s)^{n-1}\\ &+ \sum_{0<s\le t}(A_s - A_{s-})\cdot((A_s)^{n-1} - (A_{s-})^{n-1}).\end{aligned} \tag{6.1}$$

By induction hypothesis, we have

$$\begin{aligned}(A_t)^{n-1} =&(A_0)^{n-1} + \int_0^t (n-1)(A_{s-})^{n-2}\,dA_s\\ &+ \sum_{0<s\le t}(A_s)^{n-1} - (A_{s-})^{n-1} - (n-1)(A_{s-})^{n-2}(A_s - A_{s-}).\end{aligned} \tag{6.2}$$

Apply the functional monotone class theorem to (6.2) above, for all bounded Borel measurable f we have

$$\begin{aligned}&\int_0^t f(s)\,d(A_s)^{n-1} = (n-1)\int_0^t f(s)(A_{s-})^{n-2}\,dA_s\\ &+ \sum_{0<s\le t} f(s)(A_s)^{n-1} - f(s)(A_{s-})^{n-1} - (n-1)f(s)(A_{s-})^{n-2}(A_s - A_{s-}).\end{aligned} \tag{6.3}$$

Let $f(s) = A_{s-}$ in (6.3) and substitute (6.3) back to (6.1), we have

$$\begin{aligned}(A_t)^n =&(A_0)^n + \int_0^t (A_{s-})^{n-1}\,dA_s + (n-1)\int_0^t (A_{s-})^{n-1}\,dA_s\\ &+ \sum_{0<s\le t}(A_{s-})(A_s)^{n-1} - (A_{s-})^n - (n-1)(A_{s-})^{n-1}(A_s - A_{s-})\\ &+ \sum_{0<s\le t}(A_s - A_{s-})\cdot((A_s)^{n-1} - (A_{s-})^{n-1})\\ =&(A_0)^n + \int_0^t n(A_{s-})^{n-1}\,dA_s + \sum_{0<s\le t}(A_s)^n - (A_{s-})^n - n(A_{s-})^{n-1}(A_s - A_{s-}),\end{aligned}$$

which proves the claim made in Step 1.

Step 2. Fix K large so that $[-K, K]$ contains the image of $[0, t]$ under A. By Weierstrass approximation theorem, we may find a sequence of polynomials $\{p_n\}$ so that $p_n \to F'$

uniformly in $[-K, K]$, and $p_n(-K) = F'(-K)$ for all $n \in \mathbb{N}$. Now we let $P_n(x) := \int_{-K}^{x} p_n(y)\,dy + F(-K)$ for all $n \in \mathbb{N}$, and it follows that $P_n \to F$ uniformly in $[-K, K]$.

Step 3. Since A is of BV on $[0, t]$, we may decompose $A_s = B_s - C_s$, where B, C are increasing functions on $[0, t]$. We have

$$\sum_{0<s\le t} |A_s - A_{s-}| \le \sum_{0<s\le t} |B_s - B_{s-}| + \sum_{0<s\le t} |C_s - C_{s-}|$$
$$\le B_t - B_0 + C_t - C_0 < \infty.$$

Step 4. Let $\triangle A_s := A_s - A_{s-}$.

$$\Big|\sum_{0<s\le t} F(A_s) - P_n(A_s) - F(A_{s-}) + P_n(A_{s-}) - \Big(F'(A_{s-}) - p_n(A_{s-})\Big) \cdot \triangle A_s\Big|$$
$$\le \sum_{0<s\le t} |F'(A_s^{(n)}) - p_n(A_s^{(n)})| \cdot |\triangle A_s| + |F'(A_{s-}) - p_n(A_{s-})| \cdot |\triangle A_s|$$
$$\to 0 \text{ as } n \to \infty$$

Step 5. Approximate F using P_n. The proof is now complete. □

References

[1] K. L. Chung. *A course in probability theory.* Academic Press, 2001.
[2] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*, volume 293. Springer-Verlag, Berlin, 2004.
[3] S. R. S. Varadhan. *Probability theory.* Courant Lecture Notes, AMS, 2001.
[4] J. Yeh. *Real analysis, Theory of measure and integration.* World Scientific, 2006.

Shiu-Tang Li, University of Utah, Mathematics Department