Definition 1. Let $A : [0, a] \to \mathbb{R}$ be a RC function of finite variation. We define an finitely additive set function $\mu$ on $([0, a], \mathcal{F})$ by $\mu([0, t]) := A_t$ for any $0 \leq t \leq a$, where $\mathcal{F}$ 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 $\mu^*$ on $([0, a], \mathcal{B}([0, a]))$ so that $\mu^*(E) = \mu(E)$ for all $E \in \mathcal{F}$, where $\mu$ 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 $\mu^1, \mu^2$ are two finitely additive set functions on $([0, a], \mathcal{F})$. By [Theorem 1.2, Varadhan], for any $D_n \in \mathcal{F}$, $D_n \downarrow \emptyset$, we have $\mu^1(D_n) \downarrow 0$ and $\mu^2(D_n) \downarrow 0$. Therefore, by [Theorem 1.3, Varadhan], there exists two positive measures $(\mu^1)^*$ and $(\mu^2)^*$ on $([0, a], \mathcal{B}([0, a]))$, so that $(\mu^1)^*(E) = \mu^1(E)$ and $(\mu^2)^*(E) = \mu^2(E)$ for all $E \in \mathcal{F}$. We then define $\mu^* := (\mu^1)^* - (\mu^2)^*$, and the proof of existence is complete.

The uniqueness of $\mu^*$ follows from the monotone class theorem.

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

Lebesgue-Stieltjes integral

Since the Lebesgue-Stieltjes measure is a measure defined on $([0, a], \mathcal{B}([0, a]))$, any $f : [0, a] \to \mathbb{R}$ is a candidate for $\int_{[0, a]} f(s) \, dA_s$ to make sense.
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. 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.

For $A$ (RC) of finite variation 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.

Note that we may also define the Lebesgue-Stieltjes measure on $([a, b], B([a, b]))$. It follows the same construction procedure.

**Decomposition of $dA_s$**

By [Section 1.3, Chung], 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 $dA_s$ can be decomposed into three parts. Usually, it is the most 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 4.** Let $A$ be absolutely continuous, and let $f$ be a bounded Borel measurable function on $[0, a]$. Then $\int_{[0,a]} f(s) \, dA_s = \int_{[0,a]} f(s) A'_s \, ds$.

*Proof.* Monotone class theorem. \quad $\square$

**Total variation $|dA_s|$ of $dA_s$**

Let us recall the definition of signed measure, and its decomposition theorems. Our main reference is [Chap 2.10, James Yeh].

**Definition 5.** [Definition 2.10.1, James Yeh] Given a measurable space $(X, \mathcal{F})$. A set function $\lambda$ on $\mathcal{F}$ is called a signed measure on $\mathcal{F}$ if it satisfies the following conditions:
(1) \( \lambda(E) \in (-\infty, \infty] \) for every \( E \in \mathcal{F} \) or \( \lambda(E) \in [-\infty, \infty) \) for every \( E \in \mathcal{F} \).

(2) \( \lambda(\emptyset) = 0 \).

(3) countable additivity: for every disjoint sequence \( \{E_n : n \in \mathbb{N}\} \) in \( \mathcal{F} \),
\[
\sum_{n=1}^{\infty} \lambda(E_n) \text{ exists in } \mathbb{R} \text{ and } \sum_{n=0}^{\infty} \lambda(E_n) = \lambda(\bigcup_{n=1}^{\infty} E_n).
\]

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

**Theorem 6.** (Hahn Decomposition of Signed Measure Spaces) [Theorem 2.10.14, James Yeh] For an arbitrary signed measure space \( (X, \mathcal{F}, \lambda) \), a Hahn decomposition exists and is unique up to null sets of \( \lambda \), that is, there exist a positive set \( P \) and a negative set \( N \) for \( \lambda \) such that \( P \cap N = \emptyset \) 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 \( \lambda \).

**Theorem 7.** (Jordan Decomposition of Signed Measures) [Theorem 2.10.21, James Yeh] Given a signed measure space \( (X, \mathcal{F}, \lambda) \), a Jordan decomposition for \( (X, \mathcal{F}, \lambda) \) exists and is unique, that is, there exists a unique pair \( \{\mu, \nu\} \) of positive measures on \( (X, \mathcal{F}) \), 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 \( (X, \mathcal{F}, \lambda) \), if we define two set functions \( \mu \) and \( \nu \) on \( \mathcal{F} \) by setting \( \mu(E) = \lambda(E \cap P) \) and \( \nu(E) = -\lambda(E \cap N) \) for all \( E \in \mathcal{F} \), then \( \{\mu, \nu\} \) is a Jordan decomposition for \( (X, \mathcal{F}, \lambda) \).

**Definition 8.** [Definition 2.10.22, James Yeh] Given a signed measure space \( (X, \mathcal{F}, \lambda) \). Let \( \mu \) and \( \nu \) be the unique positive measures on \( \mathcal{F} \), at least one of which is finite, such that \( \mu \perp \nu \) and \( \lambda = \mu - \nu \). Let us call \( \mu \) and \( \nu \) the positive and negative parts of \( \lambda \) and write \( \lambda^+ \) for \( \mu \) and \( \lambda^- \) for \( \nu \). The total variation of \( X \) is a positive measure \( |\lambda| \) on \( \mathcal{F} \) defined by \( |\lambda|(E) = \lambda^+(E) + \lambda^-(E) \) for \( E \in \mathcal{F} \).

Since \( dA \) on \( ([0, a], \mathcal{B}([0, a])) \) is a signed measure, we may find \( P, N \in \mathcal{B}([0, a]) \), \( P \cup N = [0, a] \), \( P \cap N = \emptyset \), so that \( dA = \mu - \nu \), where \( \mu(E) = dA(E \cap P) \geq 0 \) and \( \nu(E) = -dA(E \cap N) \leq 0 \) for all \( E \in \mathcal{B}([0, a]) \). Therefore, we may define the total variation \( |dA| \) of \( dA \) by \( |dA| := \mu + \nu \).

It’s easily seen that \( \int_{[0,a]} f \cdot (1_P - 1_N) \, dA = \int_{[0,a]} f \cdot |dA| \) for \( f \in \mathcal{B}([0, a]) \).

Next, we show the connection between the variation of signed measures and the variation of a BV function. (Thanks for the assistance by Wei-Da}
We also assume that \( A_0 = 0 \). We claim that \( dV = |dA| \), as a measure on \([0, a], \mathcal{B}([0, a])\).

**Proof.** First, it is easily seen that the identity holds for \( |dA(s, t)| = |A(t) - A(s)| \leq V[s, t] = V(t) - V(s) = dV(s, t) \) for all intervals \( (s, t) \subseteq [0, a] \). By the monotone class theorem, \( |dA(B)| \leq dV(B) \) for any \( B \in \mathcal{B}([0, a]) \). Now, let \( \{P, N\} \) be the Hahn decomposition of \((0, a], \mathcal{B}([0, a]), A)\). For any \( E \in \mathcal{B}([0, a]) \), we have \( |dA|(B) = dA(B \cap P) - dA(B \cap N) \leq dV(B \cap P) + dV(B \cap N) = dV(B) \).

\[
dV \leq |dA|: \text{ since } dV(s, t] = V(t) - V(s) = \sup \Delta \sum |A_{t_i} - A_{t_{i-1}}| \leq \sup \Delta \sum |dA|(t_i - t_{i-1}) = |dA|(s, t]. \text{ The rest of the proof follows from our old friend, the monotone class theorem.} \]

**Properties of Lebesgue-Stieltjes integral**

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

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

**Proof.** Dominated convergence theorem. \( \square \)

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

**Proof.** \( g(t) \) is the difference of two increasing functions on \([0, a]\). \( \square \)

**Theorem 12.** (Associativity) Let \( f, g \) be as above. Let \( h \in \mathcal{B}([0, a]) \) 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 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 \( h \in \mathcal{B}([0, a]) \) so that \( \int_0^a |h(s)| \, dg_s < \infty \) or \( \int_0^a |h(s)f(s)| \, dA_s < \infty \), we write \( h = h^+ - h^- = h^+1_{P^+} + h^+1_{P^r} - h^-1_{N^-} - h^-1_{N^r} \). We then approximate each component with bounded Borel measurable functions, using the monotone convergence theorem.

**Remarks 13.** (1) We say \( f \) satisfies property (*) if \( f \in \mathcal{B}([0, a]) \) and \( \int_0^a |f(s)| \, dA_s < \infty \). When \( f \) is continuous on \([0, a]\), \( f \) of \( BV \) on \([0, a]\), or \( f \) bounded Borel measurable, then \( f \) satisfies property (*). (2) If \( A \) is of \( BV \) on \([0, a]\), then \( A_+ \) is of \( BV \) on \([0, a]\), thanks to the Jordan decomposition for \( BV \) functions.

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

**Proof.** Each term is equal to \( dA \otimes dB[0, t]^2 \).

**Theorem 15.** [Revuz, Yor] If \( F \) is a \( C^1 \)-function and \( A \) is of finite variation, then \( F(A) \) is of finite variation and

\[
F(A_t) = F(A_0) + \int_0^t F'(A_s-) \, dA_s + \sum_{0 < s \leq t} (F(A_s) - F(A_s-)) - F'(A_s-) (A_s - A_s-).
\]

**Proof.** 1. The first assertion follows from \( \sum_{\triangle} |F(A_t) - F(A_{t-1})| \leq \sum_{\triangle} |F'(\xi)| \cdot |A_t - A_{t-1}| \).

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.** I’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

\[
(A_t)^n = (A_0)^n + \int_0^t (A_s-)^{n-1} \, dA_s + \int_0^t A_s \, d(A_s)^{n-1}
+ \sum_{0 < s \leq t} (A_s - A_s-) \cdot ((A_s)^{n-1} - (A_s-)^{n-1}). \tag{1}
\]

By induction hypothesis, we have

\[
(A_t)^{n-1} = (A_0)^{n-1} + \int_0^t (n-1)(A_s-)^{n-2} \, dA_s
+ \sum_{0 < s \leq t} (A_s)^{n-1} - (A_s-)^{n-1} - (n-1)(A_s-)^{n-2}(A_s - A_s-). \tag{2}
\]
Apply monotone class theorem to (2) above, for all bounded Borel measurable $f$ we have
\[
\int_0^t f(s) d(A_s)^{n-1} = (n - 1) \int_0^t f(s)(A_s)_{n-2} dA_s \\
+ \sum_{0<s\leq t} f(s)(A_s)^{n-1} - f(s)(A_s_{n-1}) - (n - 1)f(s)(A_s)_{n-2}(A_s - A_s_{-}). \tag{3}
\]
Let $f(s) = A_s -$ in (3) and substitute (3) back to (1), we have
\[
(A_t^n) = (A_0^n) + \int_0^t (A_s)_{n-1} dA_s + (n - 1) \int_0^t (A_s_{n-1}) dA_s \\
+ \sum_{0<s\leq t} (A_s)_{n-1} - (A_s)_{n} - (n - 1)(A_s - A_s_{-}) \\
+ \sum_{0<s\leq t} (A_s - A_s_{-}) \cdot ((A_s)_{n-1} - (A_s_{-})_{n-1}) \\
= (A_0^n) + \int_0^t n(A_s_{n-1}) dA_s + \sum_{0<s\leq t} (A_s)^{n} - (A_s_{-})^{n} - n(A_s_{-})_{n-1}(A_s - A_s_{-}),
\]
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
\[
\sum_{0<s\leq t} |A_s - A_s_{-}| \leq \sum_{0<s\leq t} |B_s - B_s_{-}| + \sum_{0<s\leq t} |C_s - C_s_{-}| \\
\leq B_t - B_0 + C_t - C_0 < \infty.
\]

**Step 4.** Let $\Delta A_s := A_s - A_s_{-}$.
\[
\left| \sum_{0<s\leq t} F(A_s) - P_n(A_s) - F(A_s_{-}) + P_n(A_s_{-}) - \left( F'(A_s_{-}) - p_n(A_s_{-}) \right) \cdot \Delta A_s \right| \\
\leq \sum_{0<s\leq t} |F'(A_s^n) - p_n(A_s^n)| \cdot |\Delta A_s| + |F'(A_s_{-}) - p_n(A_s_{-})| \cdot |\Delta A_s| \\
\to 0 \text{ as } n \to \infty
\]

**Step 5.** Approximate $F$ using $P_n$. The proof is now complete. \qed

6
Corollary 16. Settings as in Theorem 15. Let $f$ be a bounded Borel measurable function, then we have

\[
\int_0^t f(s) dF(A_s) = \int_0^t f(s) F'(A_{s-}) dA_s \\
+ \sum_{0<s\leq t} f(s) F(A_s) - f(s) F(A_{s-}) - f(s) F'(A_{s-}) \cdot \Delta A_s.
\]

Before proceeding to the next theorem, we would like to show that when $A$ is of bounded variation on $[0, a]$, $A_0 := 0$, we may define $Y_t := \prod_{j=1}^\infty (1 + \Delta A_{a_j})$, where $\{a_1, a_2, \cdots\}$ is some enumeration of $\{0 \leq s \leq t : \Delta A_s \neq 0\}$. Note that $Y_0 = 1 + A_0$, and we define $Y_0 := 0$.

Here are a few remarks.

(1) We’d like to show the limit of $\prod_{j=1}^n (1 + \Delta A_{a_j})$ exists as $n$ goes to infinity. For $n > m$, $|\prod_{j=1}^n (1 + \Delta A_{a_j}) - \prod_{j=1}^m (1 + \Delta A_{a_j})| \leq e^{\sum_{s\in[0,a]} |\Delta A_s|} \cdot \sum_{m+1 \leq j \leq n} |\Delta A_{a_j}|$, showing that $\prod_{j=1}^n (1 + \Delta A_{a_j})$ is Cauchy.

(2) The limit is independent of our choice of $\{a_j\}$. Let $\{b_1, b_2, \cdots\}$ be another enumeration of $\{0 \leq s \leq t : \Delta A_s \neq 0\}$. For each $n \in \mathbb{N}$, we may find $k = k(n) \in \mathbb{N}$ so that $\{b_1, b_2, \cdots, b_n\} \subset \{a_1, a_2, \cdots, a_k\}$. We may find $n$ large so that $|\prod_{j=1}^n (1 + \Delta A_{b_j}) - \prod_{j=1}^k (1 + \Delta A_{a_j})|$ is small enough.

By (2), we may write $Y_t = \prod_{0 \leq s \leq t} (1 + \Delta A_s)$.

(3) Now we extend the definition of $Y_t = \prod_{0 \leq s \leq t} (1 + \Delta A_s) = \prod_{s \in I} (1 + \Delta A_s)$, where $I = [0, a]$, to arbitrary finite subinterval of $[0, a]$ or a single point in $[0, a]$, with the same approach of construction. The new definition is denoted by $Y_I$. Here are some properties of $Y_I$:

(3-1) For $I_1, I_2$ disjoint, $I_1 \cup I_2 = I$, $Y_I = Y_{I_1} \cdot Y_{I_2}$.

(3-2) For $I_1, I_2$ disjoint, $I_1 \cup I_2 = I$, $|Y_I - Y_{I_1}| \leq e^{\sum_{s \in I} |\Delta A_s|} \cdot \sum_{s \in I_2} |\Delta A_s|$.

The following ones are proved using (3-2):

(3-3) $Y_I$ is of BV on $[0, a]$.

(3-4) If $I_n \uparrow I$, then $Y_{I_n} \to Y_I$. 

7
\[(3-5) \quad Y_{t-} = \Pi_{0 \leq s < t}(1 + \Delta A_s) \text{; } \Delta Y_t = \Pi_{0 \leq s < t}(1 + \Delta A_s) \cdot \Delta A_s. \]

(4) We’d like to show \( Y_t = \sum_{0 \leq s \leq t} \Delta Y_s. \) Let \( s_{n,1} < s_{n,2} < \cdots < s_{n,n} \) be a reordering of \( \{a_1, \cdots, a_n\} \), we have

\[
Y_t = \lim_{n \to \infty} \Pi_{j=1}^n (1 + \Delta A_{s_{n,j}})
= \lim_{n \to \infty} \Pi_{j=1}^n (1 + \Delta A_{s_{n,1}}) + \sum_{m=2}^n (\Pi_{j=1}^m (1 + \Delta A_{s_{n,j}}) - \Pi_{j=1}^{m-1} (1 + \Delta A_{s_{n,j}}))
= \lim_{n \to \infty} \left( (1 + \Delta A_{s_{n,1}}) + \sum_{m=2}^n \Delta A_{s_{n,m}} \cdot \Pi_{j=1}^{m-1} (1 + \Delta A_{s_{n,j}}) \right)
= 1 + \Delta A_0 + \sum_{0 < s \leq t} \Delta A_s \cdot \Pi_{0 \leq z < s} (1 + \Delta A_z)
= \sum_{0 \leq s \leq t} \Delta Y_s.
\]

We have used the bounded convergence theorem in the second last equality. It is bounded by \( e^{\sum_{s \in [0, a]} |\Delta A_s|} \), with counting measure \( \Delta A_s \).

**Theorem 17. [Revuz. Yor]** If \( A \) is a RC function of finite variation, \( A_0 = A_{0-} = 0 \), then \( Y_t := Y_0 \Pi_{0 \leq s \leq t} (1 + \Delta A_s) e^{A^c_t} \) is the only locally bounded solution of the equation \( Y_t = Y_0 + \int_0^t Y_s \cdot dA_s \), where \( A^c_t := A_t - \sum_{0 \leq s \leq t} \Delta A_s \).

**Proof.** Apply Theorem 14. to \( C_t := Y_0 \Pi_{0 \leq s \leq t} (1 + \Delta A_s) \) and \( D_t := e^{A^c_t} \), we have

\[
Y_0 \Pi_{0 \leq s \leq t} (1 + \Delta A_s) e^{A^c_t} = Y_0 + \int_0^t e^{A^c_s} dC_s + \int_0^t Y_0 \Pi_{0 \leq z < s} (1 + \Delta A_z) dD_s
= Y_0 + \int_0^t Y_0 e^{A^c_s} \cdot \Pi_{0 \leq z < s} (1 + \Delta A_z) d \sum_{0 \leq z \leq s} \Delta A_s
+ \int_0^t Y_0 \Pi_{0 \leq z < s} (1 + \Delta A_z) e^{A^c_s} dA^c_s
= Y_0 + \int_0^t Y_0 e^{A^c_s} \cdot \Pi_{0 \leq z < s} (1 + \Delta A_z) dA_s.
\]

Assume there are two locally bounded solutions with the same initial value \( Y_0 \), and we denote their difference by \( Z_t \). See [Revuz. Yor] for details. \( \square \)