From Feynman Heuristics to Brownian Motion

In the physics literature one often finds the following informal expression,

$$d\mu_T(\omega) = \text{“}\frac{1}{Z(T)} e^{-\frac{1}{2}\int_0^T |\omega'(\tau)|^2 d\tau} \mathcal{D}_T\omega\text{”} \text{ for } \omega \in W_T, \tag{1.1}$$

where W_T is the set of continuous paths, $\omega : [0,T] \to \mathbb{R}$ (or $\mathbb{R}^d$), such that $\omega(0) = 0$,

$$\mathcal{D}_T\omega = \text{“}\prod_{0<t\leq T} m(d\omega(t))\text{”} \ (m \text{ is Lebesgue measure here})$$

and $Z(T)$ is a normalization constant such that $\mu_T(W_T) = 1$.

We begin by giving meaning to this expression. For $0 \leq s \leq t \leq T$, let

$$E_{[s,t]}(\omega) := \int_s^t |\omega'(\tau)|^2 d\tau.$$

If we decompose $\omega(\tau)$ as $\sigma(\tau) + \gamma(\tau)$ where

$$\sigma(\tau) := \omega(s) + \frac{\tau - s}{t-s}(\omega(t) - \omega(s)) \text{ and } \gamma(\tau) := \omega(\tau) - \sigma(\tau),$$

then we have, $\sigma\prime(\tau) = \frac{\omega(t)-\omega(s)}{t-s}$, $\gamma(s) = \gamma(t) = 0$, and hence

$$\begin{aligned}\int_s^t \sigma'(\tau)\cdot\gamma'(\tau)\,d\tau &= \int_s^t \sigma'(\tau)\cdot\gamma'(\tau)\,d\tau \\ &= \frac{\omega(t)-\omega(s)}{t-s}\cdot(\gamma(t)-\gamma(s)) = 0.\end{aligned}$$

Thus it follows that

$$\begin{aligned}E_{[s,t]}(\omega) &= E_{[s,t]}(\sigma) + E_{[s,t]}(\gamma) = \left|\frac{\omega(t)-\omega(s)}{t-s}\right|^2 (t-s) + E_{[s,t]}(\gamma) \\ &= \frac{|\omega(t)-\omega(s)|^2}{t-s} + E_{[s,t]}(\gamma).\end{aligned} \tag{1.2}$$

Thus if $f(\omega) = F\left(\omega|_{[0,s]}, \omega(t)\right)$, we will have,

$$\begin{aligned}&\frac{1}{Z_t}\int_{W_t} F\left(\omega|_{[0,s]}, \omega(t)\right) e^{-\frac{1}{2}E_t(\omega)}\mathcal{D}_t\omega \\ &\quad = \frac{1}{Z_t}\int_{W_t} F\left(\omega|_{[0,s]}, \omega(t)\right) e^{-\frac{1}{2}\left[E_s(\omega)+E_{[s,t]}(\omega)\right]}\mathcal{D}_t\omega\end{aligned}$$

and now fixing $\omega|_{[0,s]}$ and $\omega(t)$ and then doing the integral over $\omega|_{(s,t)}$ implies,

$$\begin{aligned}&\int F\left(\omega|_{[0,s]}, \omega(t)\right) e^{-\frac{1}{2}\left[E_s(\omega)+E_{[s,t]}(\omega)\right]}\mathcal{D}_{(s,t)}\omega \\ &= \int F\left(\omega|_{[0,s]}, \omega(t)\right) e^{-\frac{1}{2}\left[E_s(\omega)+\frac{|\omega(t)-\omega(s)|^2}{t-s}+E_{[s,t]}(\gamma)\right]}\mathcal{D}_{(s,t)}\gamma \\ &= C(s,t)\int F\left(\omega|_{[0,s]}, \omega(t)\right)\frac{e^{-\frac{1}{2}E_s(\omega)}}{Z(s)} e^{-\frac{1}{2}\frac{|\omega(t)-\omega(s)|^2}{t-s}}.\end{aligned}$$

Multiplying this equation by $\frac{1}{Z_t}\mathcal{D}\omega_{[0,s]}\cdot d\omega(t)$ and integrating the result then implies,

$$\begin{aligned}&\int_{W_t} F\left(\omega|_{[0,s]}, \omega(t)\right) d\mu_t(\omega) \\ &\quad = \frac{C(s,t)}{Z_t}\int\left[\int_{\mathbb{R}^d} F\left(\omega|_{[0,s]}, y\right) e^{-\frac{1}{2}\frac{|y-\omega(s)|^2}{t-s}}dy\right]\frac{e^{-\frac{1}{2}E_s(\omega)}}{Z(s)}\mathcal{D}\omega_{[0,s]} \\ &\quad = \frac{C(s,t)}{Z_t}\int_{W_s}\left[\int_{\mathbb{R}^d} F(\omega, y)\, e^{-\frac{1}{2}\frac{|y-\omega(s)|^2}{t-s}}dy\right] d\mu_s(\omega).\end{aligned}$$

Taking $F \equiv 1$ in this equation then implies,

$$\begin{aligned}1 &= \frac{C(s,t)}{Z_t}\int_{W_s}\left[\int_{\mathbb{R}^d} e^{-\frac{1}{2}\frac{|y-\omega(s)|^2}{t-s}}dy\right] d\mu_s(\omega) \\ &= \frac{C(s,t)}{Z_t}\int_{W_s}\left[(2\pi(t-s))^{d/2}\right] d\mu_s(\omega) = \frac{C(s,t)}{Z_t}(2\pi(t-s))^{d/2}.\end{aligned}$$

Thus the heuristic expression in Eq. (1.1) leads to the following **Markov property** for μ_t, namely.

Proposition 1.1 (Heuristic). *Suppose that $F : W_s \times \mathbb{R}^d \to \mathbb{R}$ is a reasonable function, then for any $t \geq s$ we have*

$$\int_{W_t} F\left(\omega|_{[0,s]}, \omega(t)\right) d\mu_t(\omega) = \int_{W_s} \left[\int_{\mathbb{R}^d} F(\omega, y)\, p_{t-s}(\omega(s), y)\, dy\right] d\mu_s(\omega),$$

where

$$p_s(x,y) := \left(\frac{1}{2\pi(t-s)}\right)^{d/2} e^{-\frac{1}{2}\frac{|y-x|^2}{t-s}}. \tag{1.3}$$

Corollary 1.2 (Heuristic). *If $0 = s_0 < s_1 < s_2 < \cdots < s_n = T$ and $f : \left(\mathbb{R}^d\right)^n \to \mathbb{R}$ is a reasonable function, then*

$$\int_{W_T} f(\omega(s_1), \dots, \omega(s_n))\, d\mu_T(\omega) = \int_{(\mathbb{R}^d)^n} f(y_1, \dots, y_n) \prod_{i=1}^n \left(p_{s_i - s_{i-1}}(y_{i-1}, y_i)\, dy_i\right) \tag{1.4}$$

where by convention, $y_0 = 0$.

Theorem 1.3 (Wiener 1923). *For all $t > 0$ there exists a unique probability measure, μ_t, on W_t, such that Eq. (1.4) holds for all n and all bounded measurable $f : \left(\mathbb{R}^d\right)^n \to \mathbb{R}$.*

Definition 1.4. *Let $B_t(\omega) := \omega(t)$. Then $\{B_t\}_{0 \le t \le T}$ as a process on (W_T, μ_T) is called **Brownian motion.** We further write $\mathbb{E}f$ for $\int_{W_T} f(\omega)\, d\mu_T(\omega)$.*

The following lemma is useful for computational purposes involving Brownian motion and follows readily form from Eq. (1.4).

Lemma 1.5. *Suppose that $0 = s_0 < s_1 < s_2 < \cdots < s_n = t$ and $f_i : \mathbb{R}^d \to \mathbb{R}$ are reasonable functions, then*

$$\mathbb{E}\left[\prod_{i=1}^n f_i\left(B_{s_i} - B_{s_{i-1}}\right)\right] = \prod_{i=1}^n \mathbb{E}\left[f_i\left(B_{s_i} - B_{s_{i-1}}\right)\right], \tag{1.5}$$

$$\mathbb{E}\left[f(B_t - B_s)\right] = \mathbb{E}\left[f(B_{t-s})\right], \tag{1.6}$$

and

$$\mathbb{E}\left[f(B_t)\right] = \mathbb{E}f\left(\sqrt{t}B_1\right). \tag{1.7}$$

As an example let us observe that

$$\mathbb{E}B_t = \int y p_t(y)\, dy = 0,$$

$$\mathbb{E}B_t^2 = t\mathbb{E}B_1^2 = t\int y^2 p_1(y)\, dy = t \cdot 1,$$

and for $s < t$,

$$\mathbb{E}[B_t B_s] = \mathbb{E}[(B_t - B_s) B_s] + \mathbb{E}B_s^2 = \mathbb{E}(B_t - B_s) \cdot \mathbb{E}B_s + s = s$$

and

$$\mathbb{E}[|B_t - B_s|^p] = |t-s|^{p/2}\, \mathbb{E}[|B_1|^p] = C_p |t-s|^{p/2}. \tag{1.8}$$

1.1 Construction and basic properties of Brownian motion

In this section we sketch one method of constructing Wiener measure ore equivalently Brownian motion. We begin with the existence of a measure ν_T on the $\tilde{W}_T := \prod_{0 \le s \le T} \bar{\mathbb{R}}$ which satisfies Eq. (1.4) where $\bar{\mathbb{R}}$ is a compactification of $\mathbb{R}$ – for example either one point compactificatoin so that $\bar{\mathbb{R}} \cong S^1$.

Theorem 1.6 (Kolmogorov's Existence Theorem). *There exists a probability measure, ν_T, on $\tilde{W}_T$ such that Eq. (1.4) holds.*

Proof. For a function $F(\omega) := f(\omega(s_1), \dots, \omega(s_n))$ where $f \in C\left(\bar{\mathbb{R}}^n, \mathbb{R}\right)$, define

$$I(F) := \int_{\mathbb{R}^n} f(y_1, \dots, y_n) \prod_{i=1}^n \left(p_{s_i - s_{i-1}}(y_{i-1}, y_i)\, dy_i\right).$$

Using the semi-group property;

$$\int_{\mathbb{R}^d} p_t(x,y)\, p_s(y,z)\, dy = p_{s+t}(x,z)$$

along with the fact that $\int_{\mathbb{R}^d} p_t(x,y)\, dy = 1$ for all $t > 0$, one shows that $I(F)$ is well defined independently of how we represent F as a "finitely based" continuous function.

By Tychonoff's Theorem $\tilde{W}_T$ is a compact Hausdorff space. By the Stone Weierstrass Theorem, the finitely based continuous functions are dense inside of $C\left(\tilde{W}_T\right)$. Since $|I(F)| \le \|F\|_\infty$ for all finitely based continuous functions, we may extend I uniquely to a positive continuous linear functional on $C\left(\tilde{W}_T\right)$. An application of the Riesz Markov theorem now gives the existence of the desired measure, ν_T. ■

Theorem 1.7 (Kolmogorov's Continuity Criteria). *Suppose that $(\Omega, \mathcal{F}, P)$ is a probability space and $\tilde{X}_t : \Omega \to S$ is a process for $t \in [0,T]$ where (S, ρ) is a complete metric space. Assume there exists positive constants, ε, β, and C, such that*

$$\mathbb{E}[\rho(\tilde{X}_t, \tilde{X}_s)^\varepsilon] \le C|t-s|^{1+\beta} \tag{1.9}$$

for all $s, t \in [0, T]$. *Then for any* $\alpha \in (0, \beta/\varepsilon)$ *there is a modification,* X, *of* $\tilde{X}$ *(i.e.* $P\left(X_t = \tilde{X}_t\right) = 1$ *for all* t*) which is* α*–Hölder continuous. Moreover, there is a random variable* K_α *such that,*

$$\rho(X_t, X_s) \leq K_\alpha |t - s|^\alpha \text{ for all } s, t \in [0, T] \tag{1.10}$$

and $\mathbb{E}K_\alpha^p < \infty$ *for all* $p < \frac{\beta - \alpha\varepsilon}{1 - \alpha}$. *(For the proof of this theorem see Section* **??** *below.)*

Corollary 1.8. *Let* $\tilde{B}_t : \tilde{W}_T \to \mathbb{R}$ *be the projection map,* $\tilde{B}_t(\omega) = \omega(t)$. *Then there is a modifications,* $\{B_t\}$ *of* $\left\{\tilde{B}_t\right\}$ *for which* $t \to B_t$ *is* α *– Hölder continuous* ν_T *– almost surely for any* $\alpha \in (0, 1/2)$.

Proof. Applying Theorem 1.7 with $\varepsilon := p$ and $\beta := p/2 - 1$ for any $p \in (2, \infty)$ shows there is a modification $\{B_t\}_{t \geq 0}$ of $\left\{\tilde{B}_t\right\}$ which is almost surely α – Hölder continuous for any

$$\alpha \in (0, \beta/\varepsilon) = \left(0, \frac{p/2 - 1}{p}\right) = (0, 1/2 - 1/p).$$

Letting $p \to \infty$ shows that $\{B_t\}_{t \geq 0}$ is almost surely α – Hölder continuous for all $\alpha < 1/2$. ■

We will see shortly that these Brownian paths are very rough. Before we do this we will pause to develope a quatitative measurement of roughness of a continuous path.

p – Variations and Controls

Let (E,d) be a metric space which will usually be assumed to be complete.

Definition 2.1. *Let $0 \le a < b < \infty$. Given a **partition** $\Pi := \{a = t_0 < t_1 < \dots < t_n = b\}$ of $[a,b]$ and a function $Z \in C([a,b],E)$, let $(t_i)_- := t_{i-1}$, $(t_i)_+ := t_{i+1}$, with the convention that $t_{-1} := t_0 = a$ and $t_{n+1} := t_n = T$. Furthermore for $1 \le p < \infty$ let*

$$V_p(Z:\Pi) := \left(\sum_{j=1}^n d^p\left(Z_{t_j}, Z_{t_{j-1}}\right)\right)^{1/p} = \left(\sum_{t\in\Pi} d^p\left(Z_t, Z_{t_-}\right)\right)^{1/p}. \tag{2.1}$$

*Furthermore, let $\mathcal{P}(a,b)$ denote the collection of partitions of $[a,b]$. Also let $\operatorname{mesh}(\Pi) := \max_{t\in\Pi}|t-t_-|$ be the **mesh** of the partition, Π.*

Definition 2.2. *and $Z \in C([a,b],E)$. For $1 \le p < \infty$, the **p - variation** of Z is;*

$$V_p(Z) := \sup_{\Pi\in\mathcal{P}(a,b)} V_p(Z:\Pi) = \sup_{\Pi\in\mathcal{P}(a,b)} \left(\sum_{j=1}^n d^p\left(Z_{t_j}, Z_{t_{j-1}}\right)\right)^{1/p}. \tag{2.2}$$

Moreover if $Z \in C([0,T],E)$ and $0 \le a \le b \le T$, we let

$$\omega_{Z,p}(a,b) := \left[\nu_p\left(Z|_{[a,b]}\right)\right]^p = \sup_{\Pi\in\mathcal{P}(a,b)} \sum_{j=1}^n d^p\left(Z_{t_j}, Z_{t_{j-1}}\right). \tag{2.3}$$

Remark 2.3. We can define $V_p(Z)$ for $p \in (0,1)$ as well but this is not so interesting. Indeed if $0 \le s \le T$ and $\Pi \in \mathcal{P}(0,T)$ is a partition such that $s \in \Pi$, then

$$\begin{aligned} d(Z(s), Z(0)) &\le \sum_{t\in\Pi} d(Z(t), Z(t_-)) = \sum_{t\in\Pi} d^{1-p}(Z(t), Z(t_-))\, d^p(Z(t), Z(t_-)) \\ &\le \max_{t\in\Pi} d^{1-p}(Z(t), Z(t_-)) \cdot V_p^p(Z:\Pi) \\ &\le \max_{t\in\Pi} d^{1-p}(Z(t), Z(t_-)) \cdot V_p^p(Z). \end{aligned}$$

Using the uniform continuity of Z (or $d(Z(s), Z(t))$ if you wish) we know that $\lim_{|\Pi|\to 0}\max_{t\in\Pi} d^{1-p}(Z(t), Z(t_-)) = 0$ and hence that,

$$d(Z(s), Z(0)) \le \lim_{|\Pi|\to 0}\max_{t\in\Pi} d^{1-p}(Z(t), Z(t_-)) \cdot V_p^p(Z) = 0.$$

Thus we may conclude $Z(s) = Z(0)$, i.e. Z must be constant.

Lemma 2.4. *Let $\{a_i > 0\}_{i=1}^n$, then*

$$\left(\sum_{i=1}^n a_i^p\right)^{1/p} \text{ is decreasing in } p \text{ and}$$

$$\varphi(p) := \ln\left(\sum_{i=1}^n a_i^p\right) \text{ is convex in } p.$$

Proof. Let $f(i) = a_i$ and $\mu(\{i\}) = 1$ be counting measure so that

$$\sum_{i=1}^n a_i^p = \mu(f^p) \text{ and } \varphi(p) = \ln\mu(f^p).$$

Using $\frac{d}{dp}f^p = f^p \ln f$, it follows that and

$$\varphi'(p) = \frac{\mu(f^p \ln f)}{\mu(f^p)} \text{ and}$$

$$\varphi''(p) = \frac{\mu\left(f^p \ln^2 f\right)}{\mu(f^p)} - \left[\frac{\mu(f^p \ln f)}{\mu(f^p)}\right]^2.$$

Thus if we let $\mathbb{E}X := \mu(f^p X)/\mu(f^p)$, we have shown, $\varphi'(p) = \mathbb{E}[\ln f]$ and

$$\varphi''(p) = \mathbb{E}\left[\ln^2 f\right] - (\mathbb{E}[\ln f])^2 = \operatorname{Var}(\ln f) \ge 0$$

which shows that φ is convex in p.

Now let us shows that $\|f\|_p$ is decreasing in in p. To this end we compute,

$$\begin{aligned}\frac{d}{dp}\left[\ln\|f\|_p\right] &= \frac{d}{dp}\left[\frac{1}{p}\varphi(p)\right] = \frac{1}{p}\varphi'(p) - \frac{1}{p^2}\varphi(p)\\ &= \frac{1}{p^2\mu(f^p)}\left[p\mu(f^p\ln f) - \mu(f^p)\ln\mu(f^p)\right]\\ &= \frac{1}{p^2\mu(f^p)}\left[\mu(f^p\ln f^p) - \mu(f^p)\ln\mu(f^p)\right]\\ &= \frac{1}{p^2\mu(f^p)}\left[\mu\left(f^p\ln\frac{f^p}{\mu(f^p)}\right)\right].\end{aligned}$$

Up to now our computation has been fairly general. The point where μ being counting measure comes in is that in this case $\mu(f^p) \geq f^p$ everywhere and therefore $\ln\frac{f^p}{\mu(f^p)} \leq 0$ and therefore, $\frac{d}{dp}\left[\ln\|f\|_p\right] \leq 0$ as desired.

Alternative proof that $\|f\|_p$ is decreasing in p. If we let $q = p + r$, then

$$\|a\|_q^q = \sum_{j=1}^n a_j^{p+r} \leq \left(\max_j a_j\right)^r \cdot \sum_{j=1}^n a_j^p \leq \|a\|_p^r \cdot \|a\|_p^p = \|a\|_p^q,$$

wherein we have used,

$$\max_j a_j = \left(\max_j a_j^p\right)^{1/p} \leq \left(\sum_{j=1}^n a_j^p\right)^{1/p} = \|a\|_p.$$

■

Remark 2.5. It is not too hard to see that the convexity of φ is equivalent to the interpolation inequality,

$$\|f\|_{p_s} \leq \|f\|_{p_0}^{1-s} \cdot \|f\|_{p_1}^s,$$

where $0 \leq s \leq 1$, $1 \leq p_0, p_1$, and

$$\frac{1}{p_s} := (1-s)\frac{1}{p_0} + s\frac{1}{p_1}.$$

This interpolation inequality may be proved via Hölder's inequality.

Corollary 2.6. *The function $V_p(Z)$ is a decreasing function of p and $\ln V_p(Z)^p$ is a convex function of p where they are finite. Moreover, for all $p_0 > 1$,*

$$\lim_{p\downarrow p_0} V_p(Z) = V_{p_0}(Z). \tag{2.4}$$

and $p \to V_p(Z)$ is continuous on the set of p's where $V_p(Z)$ is finite.

Proof. Given Lemma 2.4, it suffices to prove Eq. (2.4) and the continuity assertion on $p \to V_p(Z)$. Since $p \to V_p(Z)$ is a decreasing function, we know that $\lim_{p\uparrow p_0} V_p(Z)$ and $\lim_{p\downarrow p_0} V_p(Z)$ always exists and also that $\lim_{p\downarrow p_0} V_p(Z) = \sup_{p>p_0}\sup_{\Pi} V_p(Z:\Pi)$. Therefore,

$$\lim_{p\downarrow p_0} V_p(Z) = \sup_{p>p_0}\sup_{\Pi} V_p(Z:\Pi) = \sup_{\Pi}\sup_{p>p_0} V_p(Z:\Pi) = \sup_{\Pi} V_{p_0}(Z:\Pi) = V_{p_0}(Z)$$

which proves Eq. (2.4). The continuity of $V_p(Z) = \exp\left(\frac{1}{p}\ln V_p(Z)^p\right)$ follows directly from the fact that $\ln V_p(Z)^p$ is convex in p and that convex functions are continuous (where finite).

Here is a proof for this case. Let $\varphi(p) := \ln V_p(Z)^p$, $1 \leq p_0 < p_1$ such that $V_{p_0}(Z) < \infty$, and $p_s := (1-s)p_0 + sp_1$, then

$$\varphi(p_s) \leq (1-s)\varphi(p_0) + s\varphi(p_1).$$

Letting $s \uparrow 1$ then implies $p_s \uparrow p_1$ and $\varphi(p_1-) \leq \varphi(p_1)$, i.e. $V_{p_1-} \leq V_{p_1} \leq V_{p_1-}$. Therefore $V_{p_1-} = V_{p_1}$ and along with Eq. (2.4) proves the continuity of $p \to V_p(Z)$. ■

2.1 Computing $V_p(x)$

How do we actually compute $V_p(x) := V_p(x;0,T)$ for a given path $x \in C([0,T],\mathbb{R})$, even a very simple one? Suppose x is piecewise linear, with corners at the points $0 = s_0, s_1, \ldots, s_m = T$. Intuitively it would seem that the p-variation should be given by choosing the corners to be the partition points. That is, if $S = \{s_0, \ldots, s_n\}$ is the partition of corner points, we might think that $V_p(x) = V_p(x;S)$. Well, first we would have to leave out any corner which is not a local extremum (because of Lemma 2.8 below). But even then, this is not generally true as is seen in Example 2.9 below.

Lemma 2.7. *For all $a, b \geq 0$ and $p \geq 1$,*

$$(a+b)^p \geq a^p + b^p \tag{2.5}$$

and the inequality is strict if $a, b > 0$ and $p > 1$.

Proof. Observe that $(a+b)^p \geq a^p + b^p$ happens iff

$$1 \geq \left(\frac{a}{a+b}\right)^p + \left(\frac{b}{a+b}\right)^p$$

which obviously holds since

$$\left(\frac{a}{a+b}\right)^p + \left(\frac{b}{a+b}\right)^p \leq \frac{a}{a+b} + \frac{b}{a+b} = 1.$$

Moreover the latter inequality is strict if if $a, b > 0$ and $p > 1$. ■

Lemma 2.8. *Let x be a path, and $D = \{t_0, \ldots, t_n\}$ be a partition. Suppose x is monotone increasing (decreasing) on $[t_{i-1}, t_{i+1}]$. Then if $D' = D \backslash \{t_i\}$, $V_p(x : D\prime) \geq V_p(x : D)$. If x is strictly increasing and $p > 1$, the inequality is strict.*

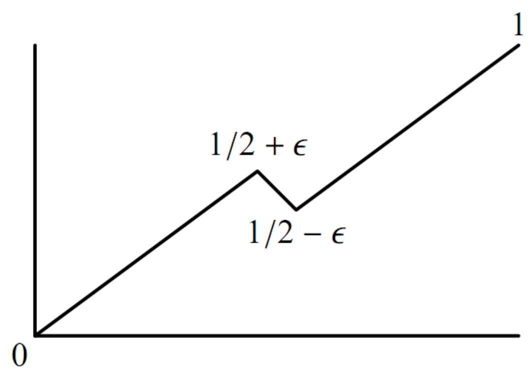
Proof. From Eq. (2.5) it follows

$$V_p(x : D\prime)^p - V_p(x : D)^p = (x(t_{i+1}) - x(t_{i-1}))^p - (x(t_{i+1}) - x(t_i))^p - (x(t_i) - x(t_{i-1}))^p$$
$$= \left(\Delta_{t_i} x + \Delta_{t_{i+1}} x\right)^p - \left(\Delta_{t_i} x\right)^p - \left(\Delta_{t_{i+1}} x\right)^p \geq 0$$

and the inequality is strict if $\Delta_{t_i} x > 0$, $\Delta_{t_{i+1}} x > 0$ and $p > 1$. ∎

In other words, on any monotone increasing segment, we should not include any intermediate points, because they can only hurt us.

Example 2.9. Consider a path like the following: If we partition $[0, T]$ at the



corner points, then

$$V_p(x : S)^p = (\frac{1}{2} + \epsilon)^p + (2\epsilon)^p + (\frac{1}{2} + \epsilon)^p \approx 2(\frac{1}{2})^p < 1$$

by taking ϵ small. On the other hand, taking the trivial partition $D = \{0, T\}$, $V_p(x : D) = 1$, so $V_p(x : S) < 1 \leq V_p(x)$ and in this case using all of local minimum and maximum does not maximize the p − variation.

The clean proof of the following theorem is due to Thomas Laetsch.

Theorem 2.10. *If $x : [0, T] \to \mathbb{R}$ having only finitely many local extremum in $(0, T)$ located at $\{s_1 < \cdots < s_{n-1}\}$. Then*

$$V_p(x) = \sup\{V_p(x : D) : \{0, T\} \subset D \subset S\},$$

where $S = \{0 = s_0 < s_1 < \cdots < s_n = T\}$.

Proof. Let $D = \{0 = t_0 < t_1 < \cdots < t_r = T\} \in \mathcal{P}(0, T)$ be an arbitrary partition of $[0, T]$. We are going to prove by induction that there is a partition $\Pi \subset S$ such that $V_p(x : D) \leq V_p(x : \Pi)$. The proof will be by induction on $n := \#(D \setminus S)$. If $n = 0$ there is nothing to prove. So let us now suppose that the theorem holds at some level $n \geq 0$ and suppose that $\#(D \setminus S) = n + 1$. Let $1 \leq k < r$ be chosen so that $t_k \in D \setminus S$. If $x(t_k)$ is between $x(t_{k-1})$ and $x(t_{k+1})$ (i.e. $(x(t_{k-1}), x(t_k), x(t_{k+1}))$ is a monotonic triple), then according Lemma 2.8 we will have $V_p(x : D) \leq V_p(x : D \setminus \{t_k\})$ and since $\#[(D \setminus \{t_k\}) \setminus S] = n$, the induction hypothesis implies there exists a partition, $\Pi \subset S$ such that

$$V_p(x : D) \leq V_p(x : D \setminus \{t_k\}) \leq V_p(x : \Pi).$$

Hence we may now assume that either $x(t_k) < \min(x(t_{k-1}), x(t_{k+1}))$ or $x(t_k) > \max(x(t_{k-1}), x(t_{k+1}))$. In the first case we let $t_k^* \in (t_{k-1}, t_{k+1})$ be a point where $x|_{[t_{k-1}, t_{k+1}]}$ has a minimum and in the second let $t_k^* \in (t_{k-1}, t_{k+1})$ be a point where $x|_{[t_{k-1}, t_{k+1}]}$ has a maximum. In either case if $D^* := (D \setminus \{t_k\}) \cup \{t_k^*\}$ we will have $V_p(x : D) \leq V_p(x : D^*)$ and $\#(D^* \setminus S) = n$. So again the induction hypothesis implies there exists a partition $\Pi \subset S$ such that

$$V_p(x : D) \leq V_p(x : D^*) \leq V_p(x : \Pi).$$

From these considerations it follows that

$$V_p(x : D) \leq \sup\{V_p(x : \Pi) : \Pi \in \mathcal{P}(0, T) \text{ s.t. } \Pi \subset S\}$$

and therefore

$$V_p(x) = \sup\{V_p(x : D) : D \in \mathcal{P}(0, T)\}$$
$$\leq \sup\{V_p(x : \Pi) : \Pi \in \mathcal{P}(0, T) \text{ s.t. } \Pi \subset S\} \leq V_p(x).$$

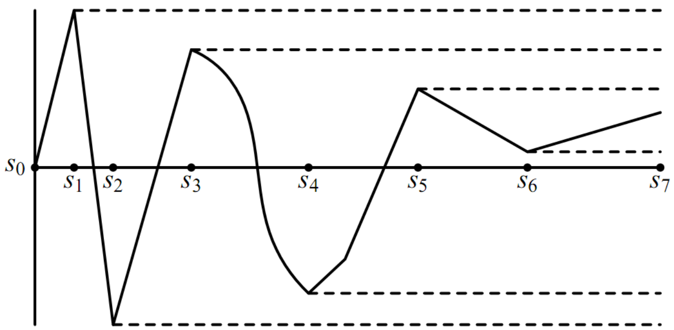
∎

Let us now suppose that x is (say) monotone increasing (not strictly) on $[s_0, s_1]$, monotone decreasing on $[s_1, s_2]$, and so on. Thus $s_0, s_2, \ldots$ are local minima, and $s_1, s_3, \ldots$ are local maxima. (If you want the reverse, just replace x with $-x$, which of course has the same p-variation.)

Definition 2.11. *Say that $s \in [0, T]$ is a **forward maximum** for x if $x(s) \geq x(t)$ for all $t \geq s$. Similarly, s is a **forward minimum** if $x(s) \leq x(t)$ for all $t \geq s$.*

Definition 2.12. *Suppose x is piecewise monotone, as above, with extrema $\{s_0, s_1, \ldots\}$. Suppose further that $s_2, s_4, \ldots$ are not only local minima but also forward minima, and that $s_1, s_3, \ldots$ are both local and forward maxima. Then we will say that x is **jog-free**.*

Note that $s_0 = 0$ does not *have to be a forward extremum. This is in order to admit a path with $x(0) = 0$ which can change signs.*



Here is an example.

Remark 2.13. Here is another way to state the jog-free condition. Let x be piecewise monotone with extrema $s_0, s_1, \dots$. Let $\xi_i = |x(s_{i+1}) - x(s_i)|$. Then x is jog-free iff $\xi_1 \geq \xi_2 \geq \dots$. The idea is that the oscillations are shrinking. (Notice that we don't need $\xi_0 \geq \xi_1$; this is because $s_0 = 0$ is not required to be a forward extremum.)

Remark 2.14. It is also okay if $s_1, s_2, \dots$ are *backwards* extrema; this corresponds to the oscillations getting larger. Just reverse time, replacing $x(t)$ by $x(T-t)$, which again doesn't change the p-variation. Note that if ξ_i are as above, this corresponds to having $\xi_0 \leq \xi_1 \leq \xi_2 \leq \dots$ (note that ξ_0 is included now, but ξ_{m-1} would not be). This case seems less useful, however.

Lemma 2.15. *Let x be jog-free with extrema $s_0, \dots, s_m$. Let $D = \{t_0, \dots, t_n\}$ be any partition not containing all the s_j. Then there is some $s_j \notin D$ such that if $D' = D \cup \{s_j\}$, $V_p(x : D\prime) \geq V_p(x : D)$.*

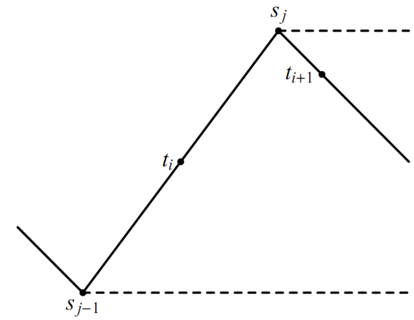
Proof. Let s_j be the first extremum not contained in D (note $s_0 = 0 \in D$ already, so j is at least 1 and s_j is also a forward extremum). Let t_i be the last element of D less than s_j. Note that $s_{j-1} \leq t_i < s_j < t_{i+1}$.

Now x is monotone on $[s_{j-1}, s_j]$; say WLOG it's monotone increasing, so that s_j is a local maximum and also a forward maximum. Since $t_i \in [s_{j-1}, s_j]$, where x is monotone increasing, $x(s_j) \geq x(t_i)$. And since s_j is a forward maximum, $x(s_j) \geq x(t_{i+1})$.

Therefore we have

$$x(s_j) - x(t_i) \geq x(t_{i+1}) - x(t_i)$$
$$x(s_j) - x(t_{i+1}) \geq x(t_i) - x(t_{i+1}).$$

One of the quantities on the right is equal to $|x(t_{i+1}) - x(t_i)|$, and so it follows that



$$|x(s_j) - x(t_i)|^p + |x(s_j) - x(t_{i+1})|^p \geq |x(t_{i+1}) - x(t_i)|^p$$

since one of the terms on the left is already $\geq$ the term on the right. This shows that $V_p(x : D\prime)^p \geq V_p(x : D)^p$. ∎

In other words, we should definitely include the extreme points, because they can only help.

Putting these together yields the desired result.

Proposition 2.16. *If x is jog-free with extrema $S = \{s_0, \dots, s_m\}$, then $V_p(x) = V_p(x : S) = (\sum \xi_i^p)^{1/p}$.*

Proof. Fix $\epsilon > 0$, and let D be a partition such that $V_p(x : D) \geq V_p(x) - \epsilon$. By repeatedly applying Lemma 2.15, we can add the points of S to D one by one (in some order), and only increase the p-variation. So $V_p(x : D \cup S) \geq V_p(x : D)$. Now, if $t \in D \backslash S$, it is inside some interval $[s_j, s_{j+1}]$ on which x is monotone, and so by Lemma 2.8 t can be removed from $D \cup S$ to increase the p-variation. Removing all such points one by one (in any order), we find that $V_p(x : S) \geq V_p(x : D \cup S)$. Thus we have $V_p(x : S) \geq V_p(x : D) \geq V_p(x) - \epsilon$; since ϵ was arbitrary we are done. ∎

Notice that we only considered the case of jog-free paths with only finitely many extrema. Of course, in order to get infinite p-variation for any p we would need infinitely many extrema. Let's just check that the analogous result holds there.

Proposition 2.17. *Suppose we have a sequence $s_0, s_1, \dots$ increasing to T, where x is alternately monotone increasing and decreasing on the intervals $[s_j, s_{j+1}]$. Suppose also that the s_j are forward extrema for x. Letting $\xi_j = |x(s_{j+1}) - x(s_j)|$ as before, we have*

$$V_p(x) = \left(\sum_{j=0}^{\infty} \xi_j^p \right)^{1/p}.$$

Actually, the extreme points s_j can converge to some earlier time than T, but x will have to be constant after that time.

Proof. For any m, we have $\sum_{j=0}^{m} \xi_j^p = V_p(x:D)^p$ for $D = \{s_0, \ldots, s_{m+1}\}$, so $V_p(x)^p \geq \sum_{j=0}^{m} \xi_j^p$. Passing to the limit, $V_p(x)^p \geq \sum_{j=0}^{\infty} \xi_j^p$.

For the reverse inequality, let $D = \{0 = t_0, t_1, \ldots, t_n = T\}$ be a partition with $V_p(x:D) \geq V_p(x) - \epsilon$. Choose m so large that $s_m > t_{n-1}$. Let $S = \{s_0, \ldots, s_m, T\}$, then by the same argument as in Proposition 2.16 we find that $V_p(x:S) \geq V_p(x:D)$. (Previously, the only way we used the assumption that S contained *all* extrema $s_j$was in order to have every $t_i \in D \backslash S$ contained in some monotone interval $[s_j, s_{j+1}]$. That is still the case here; we just take enough s_j's to ensure that we can surround each t_i. We do not need to surround $t_n = T$, since it is already in S.)

But $V_p(x:S)^p = \sum_{j=0}^{m-1} \xi_j^p \leq \sum_{j=0}^{\infty} \xi_j^p$, and so we have that

$$\left(\sum_{j=0}^{\infty} \xi_j^p\right)^{1/p} \geq V_p(x:D) \geq V_p(x) - \epsilon.$$

ϵ was arbitrary and we are done. ■

2.2 Brownian Motion in the Rough

Corollary 2.18. *For all $p > 2$ and $T < \infty$, $V_p\left(B|_{[0,T]}\right) < \infty$ a.s. (We will see later that $V_p\left(B|_{[0,T]}\right) = \infty$ a.s. for all $p < 2$.)*

Proof. By Corollary 1.8, there exists $K_p < \infty$ a.s. such that

$$|B_t - B_s| \leq K_p |t-s|^{1/p} \text{ for all } 0 \leq s, t \leq T. \tag{2.6}$$

Thus we have

$$\sum_i |\Delta_i B|^p \leq \sum_i \left(K_p |t_i - t_{i-1}|^{1/p}\right)^p \leq \sum_i K_p^p |t_i - t_{i-1}| = K_p^p T$$

and therefore, $V_p\left(B|_{[0,T]}\right) \leq K_p^p T < \infty$ a.s. ■

Proposition 2.19 (Quadratic Variation). *Let $\{\Pi_m\}_{m=1}^{\infty}$ be a sequence of partition of $[0,T]$ such that $\lim_{m\to\infty} |\Pi_m| = 0$ and define $Q_m := V_2^2(B:\Pi_m)$. Then*

$$\lim_{m\to\infty} \mathbb{E}\left[(Q_m - T)^2\right] = 0 \tag{2.7}$$

and if $\sum_{m=1}^{\infty} \operatorname{mesh}(\Pi_m) < \infty$ then $\lim_{m\to\infty} Q_m = T$ a.s. This result is often abbreviated by the writing, $dB_t^2 = dt$.

Proof. Let N be an $N(0,1)$ random variable, $\Delta t := t - t_-$, $\Delta_t B := B_t - B_{t_-}$ and observe that $\Delta_t B \sim \sqrt{\Delta t} N$. Thus we have,

$$\mathbb{E} Q_m = \sum_{t \in \Pi_m} \mathbb{E}\left(\Delta_t B\right)^2 = \sum_{t\in\Pi_m} \Delta t = T.$$

Let us define

$$\begin{aligned}\operatorname{Cov}(A,B) &:= \mathbb{E}[AB] - \mathbb{E}A \cdot \mathbb{E}B \text{ and}\\ \operatorname{Var}(A) &:= \operatorname{Cov}(A,A) = \mathbb{E}A^2 - (\mathbb{E}A)^2 = \mathbb{E}\left[(A - \mathbb{E}A)^2\right].\end{aligned}$$

and observe that

$$\operatorname{Var}\left(\sum_{i=1}^n A_i\right) = \sum_{i=1}^n \operatorname{Var}(A_i) + \sum_{i \neq j} \operatorname{Cov}(A_i, A_j).$$

As $\operatorname{Cov}(\Delta_t B, \Delta_s B) = 0$ if $s \neq t$, we may use the above computation to conclude,

$$\begin{aligned}\operatorname{Var}(Q_m) &= \sum_{t\in\Pi} \operatorname{Var}((\Delta_t B)^2) = \sum_{t\in\Pi} \operatorname{Var}(\Delta t \cdot N^2)\\ &= \operatorname{Var}(N^2) \sum_{t\in\Pi} (\Delta t)^2 \leq \operatorname{Var}(N^2)\, |\Pi_m| \sum_{t\in\Pi} \Delta t\\ &= T \cdot \operatorname{Var}(N^2)\, |\Pi_m| \to 0 \text{ as } m \to \infty.\end{aligned}$$

(By explicit Gaussian integral computations,

$$\operatorname{Var}(N^2) = \mathbb{E}N^4 - \left(\mathbb{E}N^2\right)^2 = 3 - 1 = 2 < \infty.)$$

Thus we have shown

$$\lim_{m\to\infty} \mathbb{E}\left[(Q_m - T)^2\right] = \lim_{m\to\infty} \mathbb{E}\left[(Q_m - \mathbb{E}Q)^2\right] = \lim_{m\to\infty} \operatorname{Var}(Q_m) = 0.$$

If $\sum_{m=1}^{\infty} |\Pi_m| < \infty$, then

$$\begin{aligned}\mathbb{E}\left[\sum_{m=1}^{\infty} (Q_m - T)^2\right] &= \sum_{m=1}^{\infty} \mathbb{E}\left(Q_m - T\right)^2 = \sum_{m=1}^{\infty} \operatorname{Var}(Q_m)\\ &\leq \operatorname{Var}(N^2) \cdot T \cdot \sum_{m=1}^{\infty} \operatorname{mesh}(\Pi_m) < \infty\end{aligned}$$

from which it follows that $\sum_{m=1}^{\infty} (Q_m - T)^2 < \infty$ a.s. In particular $(Q_m - T) \to 0$ almost surely. ■

Proposition 2.20. *If $p > q \geq 1$ and $V_q(Z) < \infty$, then $\lim_{|\Pi| \to 0} V_p(Z : \Pi) = 0$.*

Proof. Let $\Pi \in \mathcal{P}(0,T)$, then

$$\begin{aligned} V_p^p(Z:\Pi) &= \sum_{t \in \Pi} d^p(Z(t), Z(t_-)) = \sum_{t\in\Pi} d^{p-q}(Z(t), Z(t_-))\, d^q(Z(t), Z(t_-)) \\ &\leq \max_{t\in\Pi} d^{p-q}(Z(t), Z(t_-)) \cdot \sum_{t\in\Pi} d^q(Z(t), Z(t_-)) \\ &\leq \max_{t\in\Pi} d^{p-q}(Z(t), Z(t_-)) \cdot V_q^q(Z:\Pi) \\ &\leq \max_{t\in\Pi} d^{p-q}(Z(t), Z(t_-)) \cdot V_q^q(Z). \end{aligned}$$

Thus, by the uniform continuity of $Z|_{[0,T]}$ we have

$$\limsup_{|\Pi|\to 0} V_p(Z:\Pi) \leq \limsup_{|\Pi|\to 0} \max_{t\in\Pi} d^{p-q}(Z(t), Z(t_-)) \cdot V_q^q(Z) = 0.$$

■

Corollary 2.21. *If $p < 2$, then $V_p\left(B|_{[0,T]}\right) = \infty$ a.s.*

Proof. Choose partitions, $\{\Pi_m\}$, of $[0,T]$ such that $\lim_{m\to\infty} Q_m = T$ a.s. where $Q_m := V_2^2(B:\Pi_m)$ and let $\Omega_0 := \{\lim_{m\to\infty} Q_m = T\}$ so that $P(\Omega_0) = 1$. If $V_p\left(B|_{[0,T]}(\omega)\right) < \infty$ for then by Proposition 2.20,

$$\lim_{m\to\infty} Q_m(\omega) = \lim_{m\to\infty} V_2^2(B(\omega):\Pi_m) = 0$$

and hence $\omega \notin \Omega_0$, i.e. $\left\{V_p\left(B|_{[0,T]}(\cdot)\right) < \infty\right\} \subset \Omega_0^c$. Therefore $\Omega_0 \subset \left\{V_p\left(B|_{[0,T]}(\cdot)\right) = \infty\right\}$ and hence

$$P\left(\left\{V_p\left(B|_{[0,T]}(\cdot)\right) = \infty\right\}\right) \geq P(\Omega) = 1.$$

■

Fact 2.22 *If $\{B_t\}_{t\geq 0}$ is a Brownian motion, then*

$$P(V_p(B) < \infty) = \begin{cases} 1 \text{ if } p > 2 \\ 0 \text{ if } p \leq 2 \end{cases}.$$

See for example [7, Exercise 1.14 on p. 36].

Corollary 2.23 (Roughness of Brownian Paths). *A Brownian motion, $\{B_t\}_{t\geq 0}$, is* ***not*** *almost surely α – Hölder continuous for any $\alpha > 1/2$.*

Proof. According to Proposition 2.19 we may choose partition, Π_m, such that mesh$(\Pi_m) \to 0$ and $Q_m \to T$ a.s. If B were α – Hölder continuous for some $\alpha > 1/2$, then

$$\begin{aligned} Q_m &= \sum_{t\in\Pi_m} (\Delta_t B)^2 \leq C \sum_{t\in\Pi_m} (\Delta t)^{2\alpha} \leq C \max\left([\Delta t]^{2\alpha-1}\right) \sum_{t\in\Pi_m} \Delta t \\ &\leq C\,[|\Pi_m|]^{2\alpha-1}\, T \to 0 \text{ as } m \to \infty \end{aligned}$$

which contradicts the fact that $Q_m \to T$ as $m \to \infty$. ■

2.3 The Bounded Variation Obstruction

Proposition 2.24. *Suppose that $Z(t)$ is a real continuous function such that $Z_0 = 0$ for simplicity. Define*

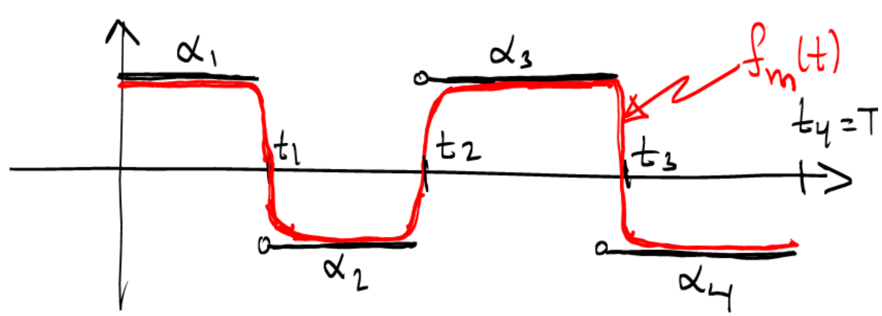
$$\int_0^T f(\tau)\, dZ(\tau) := -\int_0^T \dot{f}(\tau)\, Z(t)\, d\tau + f(t)\, Z(t)|_0^T$$

whenever f is a C^1 – function. If there exists, $C < \infty$ such that

$$\left|\int_0^T f(\tau)\, dZ(\tau)\right| \leq C \cdot \max_{0\leq\tau\leq T} |f(\tau)|, \tag{2.8}$$

then $V_1(Z) < \infty$ (See Definition 2.2 above) and the best possible choice for C in Eq. (2.8) is $V_1(Z)$.

Proof. Given a partition, $\Pi := \{0 = t_0 < t_1 < \cdots < t_n = T\}$ be a partition of $[0,T]$, $\{\alpha_k\}_{k=1}^n \subset \mathbb{R}$, and $f(t) := \alpha_1 1_{\{0\}} + \sum_{k=1}^n \alpha_k 1_{(t_{k-1},t_k]}$. Choose $f_m(t)$ in $C^1([0,T],\mathbb{R})$ "well approximating" $f(t)$ as in Figure 2.3. It then is fairly



easy to show,

$$\int_0^T \dot{f}_m(\tau) Z(t)\, d\tau \to \sum_{k=1}^{n-1} (\alpha_{k+1} - \alpha_k) Z(t_k)$$

and therefore,

$$\lim_{m\to\infty} \int_0^T f_m(t)\, dZ(t) = -\sum_{k=1}^{n-1} (\alpha_{k+1} - \alpha_k) Z(t_k) + \alpha_n Z(t_n) - \alpha_1 Z(t_0)$$
$$= \sum_{k=1}^{n} \alpha_k (Z(t_k) - Z(t_{k-1})).$$

Therefore we have,

$$\left| \sum_{k=1}^{n} \alpha_k (Z(t_k) - Z(t_{k-1})) \right| = \lim_{m\to\infty} \left| \int_0^T f_m(\tau)\, dZ(\tau) \right|$$
$$\leq C \cdot \limsup_{m\to\infty} \max_{0\leq\tau\leq T} |f_m(\tau)| = C \max_k |\alpha_k|.$$

Taking $\alpha_k = \operatorname{sgn}(Z(t_k) - Z(t_{k-1}))$ for each k, then shows $\sum_{k=1}^{n} |Z(t_k) - Z(t_{k-1})| \leq C$. Since this holds for any partition Π, it follows that $V_1(Z) \leq C$.

If $V_1(Z) < \infty$, then

$$\int_0^T \dot{f}(\tau) Z(t)\, d\tau = -\int_0^T f(t)\, d\lambda_Z(t) + f(t) Z(t)|_0^T$$

where λ_Z is the Lebesgue Stieltjes measure associated to Z. From this identity and integration by parts for such finite variation functions, it follows that

$$\int_0^T f(t)\, dZ(t) = \int_0^T f(t)\, d\lambda_Z(t)$$

and

$$\left| \int_0^T f(t)\, dZ(t) \right| = \left| \int_0^T f(t)\, d\lambda_Z(t) \right| \leq \int_0^T |f(t)|\, d\|\lambda_Z\|(t)$$
$$\leq \max_{0\leq\tau\leq T} |f(\tau)| \cdot \|\lambda_Z\|([0,T]) = V_1(Z) \cdot \max_{0\leq\tau\leq T} |f(\tau)|$$

Therefore C can be taken to be $V_1(Z)$ in Eq. (2.8) and hence $V_1(Z)$ is the best possible constant to use in this equation. ■

Combining Fact 2.22 with Proposition 2.24 explains why we are going to have trouble defining $\int_0^t f_s dB_s$ when B is a Brownian motion. However, one might hope to use Young's integral in this setting.

Theorem 2.25 (L. C. Young 1936). *Suppose that* $p, q > 0$ *with* $\frac{1}{p} + \frac{1}{q} =: \theta > 1$. *Then there exists a constant,* $C(\theta) < \infty$ *such that*

$$\left| \int_0^T f(t)\, dZ(t) \right| \leq C(\theta) \left(\|f\|_\infty + V_q(f) \right) \cdot V_p(Z)$$

for all $f \in C^1$. *Thus if* $V_p(Z) < \infty$ *the integral extends to those* $f \in C([0,T])$ *such that* $V_q(f) < \infty$.

Unfortunately, Young's integral is still not sufficiently general to allow us to solve the typical SDE that we would like to consider. For example, consider the "simple" SDE,

$$\dot{y}(t) = B(t) \dot{B}(t) \text{ with } y(0) = 0.$$

The solution to this equation should be,

$$y(T) = \int_0^T B(t)\, dB(t)$$

which still does not make sense as a Young's integral when B is a Brownian motion because for any $p > 2$, $\frac{1}{p} + \frac{1}{p} =: \theta < 1$. For more on this point view see the very interesting work of Terry Lyons on "rough path analysis," [4].

2.4 Controls

Notation 2.26 (Controls) *Let*

$$\Delta = \{(s,t) : 0 \leq s \leq t \leq T\}.$$

A ***control****, is a continuous function* $\omega : \Delta \to [0,\infty)$ *such that*

1. $\omega(t,t) = 0$ *for all* $t \in [0,T]$,
2. ω *is super-additive, i.e., for all* $s \leq t \leq v$ *we have*

$$\omega(s,t) + \omega(t,v) \leq \omega(s,v). \tag{2.9}$$

Remark 2.27. If ω is a control then and $\omega(s,t)$ is increasing in t and decreasing in s for $(s,t) \in \Delta$. For example if $s \leq \sigma \leq t$, then $\omega(s,\sigma) + \omega(\sigma,t) \leq \omega(s,t)$ and therefore, $\omega(\sigma,t) \leq \omega(s,t)$. Similarly if $s \leq t \leq \tau$, then $\omega(s,t) + \omega(t,\tau) \leq \omega(s,\tau)$ and therefore $\omega(s,t) \leq \omega(s,\tau)$.

Lemma 2.28. *If* ω *is a control and* $\varphi \in C([0,\infty) \to [0,\infty))$ *such that* $\varphi(0) = 0$ *and* φ *is convex and increasing*[1]*, then* $\varphi \circ \omega$ *is also a control.*

[1] The assumption that φ is increasing is redundant here since we are assuming $\varphi'' \geq 0$ and we may deduce that $\varphi'(0) \geq 0$, it follows that $\varphi'(x) \geq 0$ for all x. This assertion also follows from Eq. (2.11).

Proof. We must show $\varphi \circ \omega$ is still superaditive. and this boils down to showing if $0 \le a, b, c$ with $a + b \le c$, then

$$\varphi(a) + \varphi(b) \le \varphi(c).$$

As φ is increasing, it suffices to show,

$$\varphi(a) + \varphi(b) \le \varphi(a + b). \tag{2.10}$$

Making use of the convexity of φ, we have,

$$\begin{aligned}\varphi(b) &= \varphi\left(\frac{a}{a+b} \cdot 0 + \frac{b}{a+b}(a+b)\right) \\ &\le \frac{a}{a+b}\varphi(0) + \frac{b}{a+b}\varphi(a+b) = \frac{b}{a+b}\varphi(a+b)\end{aligned}$$

and interchanging the roles of a and b gives,

$$\varphi(a) \le \frac{a}{a+b}\varphi(a+b). \tag{2.11}$$

.

Adding these last two inequalities then proves Eq. (2.10). ∎

Example 2.29. Suppose that $u(t)$ is any increasing continuous function of t, then $\omega(s,t) := u(t) - u(s)$ is a control which is in fact additive, i.e.

$$\omega(s,t) + \omega(t,v) = \omega(s,v) \text{ for all } s \le t \le v.$$

So for example $\omega(s,t) = t - s$ is an additive control and for any $p > 1$, $\omega(s,t) = (t-s)^p$ or more generally, $\omega(s,t) = (u(t) - u(s))^p$ is a control.

Lemma 2.30. *Suppose that ω is a control, $p \in [1,\infty)$, and $Z \in C([0,T], E)$ is a function satisfying,*

$$d(Z_s, Z_t) \le \omega(s,t)^{1/p} \text{ for all } (s,t) \in \Delta,$$

then $V_p^p(Z) \le \omega(0,T) < \infty$. More generally,

$$\omega_{p,Z}(s,t) := V_p^p\left(Z|_{[s,t]}\right) \le \omega(s,t) \text{ for all } (s,t) \in \Delta.$$

Proof. Let $(s,t) \in \Delta$ and $\Pi \in \mathcal{P}([s,t])$, then using the superadditivity of ω we find

$$V_p^p\left(Z|_{[s,t]} : \Pi\right) = \sum_{t \in \Pi} d^p\left(Z_t, Z_{t_-}\right) \le \sum_{t \in \Pi} \omega\left(Z_t, Z_{t_-}\right) \le \omega(s,t).$$

Therefore,

$$\omega_{p,Z}(s,t) := V_p^p\left(Z|_{[s,t]}\right) = \sup_{\Pi \in \mathcal{P}([s,t])} V_p^p\left(Z|_{[s,t]} : \Pi\right) \le \omega(s,t).$$

∎

Notation 2.31 *Given $o \in E$ and $p \in [1,\infty)$, let*

$$C_p([0,T], E) := \{Z \in C([0,T], E) : V_p(Z) < \infty\} \text{ and}$$

$$C_{0,p}([0,T], E) := \{Z \in C_p([0,T], E) : Z(0) = o\}.$$

Theorem 2.32. *Let $\rho : \Delta \to [0,\infty)$ be a function and define,*

$$\omega(s,t) := \omega_\rho(s,t) := \sup_{\Pi \in \mathcal{P}(s,t)} V_1(\rho : \Pi), \tag{2.12}$$

where for any $\Pi \in \mathcal{P}(s,t)$,

$$V_1(\rho : \Pi) = \sum_{t \in \Pi} \rho(t_-, t). \tag{2.13}$$

We now assume:

1. *ρ is continuous,*
2. *$\rho(t,t) = 0$ for all $t \in [0,T]$ (This condition is redundant since next condition would fail if it were violated.), and*
3. *$V_1(\rho) := \omega(0,T) := \sup_{\Pi \in \mathcal{P}(0,T)} V_1(\rho : \Pi) < \infty$.*

Under these assumptions, $\omega : \Delta \to [0,\infty)$ is a control.

We will give the proof of Theorem 2.32 after a corollary and a few preparatory lemmas.

Corollary 2.33 (The variation control). *Let $p \in [1,\infty)$ and suppose that $Z \in C_p([0,T], E)$. Then $\omega_{Z,p} : \Delta \to [0,\infty)$ defined in Eq. (2.3) is a control satisfying, $d(Z(s), Z(t)) \le \omega_{Z,p}(s,t)^{1/p}$ for all $(s,t) \in \Delta$.*

Proof. Apply Theorem 2.32 with $\rho(s,t) := d^p(Z(s), Z(t))$ and observe that with this definition, $\omega_{Z,p} = \omega_\rho$. ∎

Lemma 2.34. *Let $\rho : \Delta \to [0,\infty)$ satisfy the hypothesis in Theorem 2.32, then $\omega = \omega_\rho$ (defined in Eq. (2.12)) is superadditive.*

Proof. If $0 \le u \le s \le v \le T$ and $\Pi_1 \in \mathcal{P}(u,s)$, $\Pi_2 \in \mathcal{P}(s,v)$, then $\Pi_1 \cup \Pi_2 \in \mathcal{P}(u,v)$. Thus we have,

$$V_1(\rho : \Pi_1) + V_1(\rho : \Pi_2) = V_1(\rho : \Pi_1 \cup \Pi_2) \le \omega(u,v).$$

Taking the supremum over all Π_1 and Π_2 then implies,

$$\omega(u,s) + \omega(s,v) \le \omega(u,v) \text{ for all } u \le s \le v,$$

i.e. ω is superaditive. ∎

Lemma 2.35. *Let* $Z \in C_p([0,T], E)$ *for some* $p \in [1,\infty)$ *and let* $\omega := \omega_{Z,p} : \Delta \to [0,\infty)$ *defined in Eq. (2.3). Then* ω *is superadditive. Furthermore if* $p = 1$, ω *is additive, i.e. Equality holds in Eq. (2.9).*

Proof. The superadditivity of $\omega_{Z,p}$ follows from Lemma 2.34 and since $\omega_{Z,p}(s,t) = \omega_\rho(s,t)$ where $\rho(s,t) := d^p(Z(s), Z(t))$. In the case $p = 1$, it is easily seen using the triangle inequality that if $\Pi_1, \Pi_2 \in \mathcal{P}(s,t)$ and $\Pi_1 \subset \Pi_2$, then $V_1(X : \Pi_1) \le V_1(X, \Pi_2)$. Thus in computing the supremum of $V_1(X : \Pi)$ over all partition in $\mathcal{P}(s,t)$ it never hurts to add more points to a partition. Using this remark it is easy to show,

$$\begin{aligned}\omega(u,s) + \omega(s,v) &= \sup_{\Pi_1 \in \mathcal{P}(u,s), \Pi_2 \in \mathcal{P}(s,v)} [V_1(X : \Pi_1) + V_1(X : \Pi_2)] \\ &= \sup_{\Pi_1 \in \mathcal{P}(u,s), \Pi_2 \in \mathcal{P}(s,v)} V_1(X : \Pi_1 \cup \Pi_2) \\ &= \sup_{\Pi \in \mathcal{P}(u,v)} V_1(X : \Pi) = \omega(u,v)\end{aligned}$$

as desired. ■

Lemma 2.36. *Let* $\rho : \Delta \to [0,\infty)$ *and* $\omega = \omega_\rho$ *be as in Theorem 2.32. Further suppose* $(a,b) \in \Delta$, $\Pi \in \mathcal{P}(a,b)$, *and let*

$$\varepsilon := \omega(a,b) - V_1(\rho : \Pi) \ge 0.$$

Then for any $\Pi' \in \mathcal{P}(a,b)$ *with* $\Pi' \subset \Pi$, *we have*

$$\sum_{t \in \Pi'} [\omega(t_-, t) - V_1(\rho : \Pi \cap [t_-, t])] \le \varepsilon. \tag{2.14}$$

In particular, if $(\alpha,\beta) \in \Delta \cap \Pi^2$ *then*

$$\omega(\alpha,\beta) \le V_1(\rho : \Pi \cap [\alpha,\beta]) + \varepsilon. \tag{2.15}$$

Proof. Equation (**??**) is a simple consequence of the superadditivity of ω (Lemma 2.34) and the identity,

$$\sum_{t \in \Pi'} V_1(\rho : \Pi \cap [t_-, t]) = V_1(\rho : \Pi)$$

where $t_- := t_-(\Pi')$. Indeed, using these properties we find,

$$\begin{aligned}\sum_{t \in \Pi'} [\omega(t_-, t) - V_1(\rho : \Pi \cap [t_-, t])] &= \sum_{t \in \Pi'} \omega(t_-, t) - V_1(\rho : \Pi) \\ &\le \omega(a,b) - V_1(\rho : \Pi) = \varepsilon.\end{aligned}$$

■

Lemma 2.37. *Suppose that* $\rho : \Delta \to [0,\infty)$ *is a continuous function such that* $\rho(t,t) = 0$ *for all* $t \in [0,T]$ *and* $\varepsilon > 0$ *is given. Then there exists* $\delta > 0$ *such that, for every* $\Pi \subset\subset [0,T]$ *and* $u \in [0,T]$ *such that* $\operatorname{dist}(u,\Pi) < \delta$ *we have,*

$$|V_1(\rho : \Pi) - V_1(\rho : \Pi \cup \{u\})| < \varepsilon.$$

Proof. By the uniform continuity of ρ, there exists $\delta > 0$ such that $|\rho(s,t) - \rho(u,v)| < \varepsilon/2$ provided $|(s,t) - (u,v)| < \delta$. Suppose that $\Pi = \{t_0 < t_1 < \dots < t_n\} \subset [0,T]$ and $u \in [0,T]$ such that $\operatorname{dist}(u,\Pi) < \delta$. There are now three case to consider, $u \in (t_0, t_n)$, $u < t_0$ and $u > t_1$. In the first case, suppose that $t_{i-1} < u < t_i$ and that (for the sake of definiteness) that $|t_i - u| < \delta$, then

$$\begin{aligned}|V_1(\rho : \Pi) - V_1(\rho : \Pi \cup \{u\})| &= |\rho(t_{i-1}, t_i) - \rho(t_{i-1}, u) - \rho(u, t_i)| \\ &\le |\rho(t_{i-1}, t_i) - \rho(t_{i-1}, u)| + |\rho(u, t_i) - \rho(t_i, t_i)| < \varepsilon.\end{aligned}$$

The second and third case are similar. For example if $u < t_0$, we will have,

$$|V_1(\rho : \Pi \cup \{u\}) - V_1(\rho : \Pi)| = \rho(u, t_0) = \rho(u, t_0) - \rho(t_0, t_0) < \varepsilon/2.$$

■

With these lemmas as preparation we are now ready to complete the proof of Theorem 2.32.

Proof. Proof of **Theorem 2.32.** Let $\omega(s,t) := \omega_\rho(s,t)$ be as in Theorem 2.32. It is clear by the definition of ω, the $\omega(t,t) = 0$ for all t and we have already seen in Lemma 2.34 that ω is superadditive. So to finish the proof we must show ω is continuous.

Using Remark 2.27, we know that $\omega(s,t)$ is increasing in t and decreasing in s and therefore $\omega(u+, v-) = \lim_{s \downarrow u, t \uparrow v} \omega(s,t)$ and $\omega(u-, v+) = \lim_{s \uparrow u,, t \downarrow v} \omega(s,t)$ exists and satisfies,

$$\omega(u+, v-) \le \omega(u,v) \le \omega(u-, v+). \tag{2.16}$$

The main crux of the continuity proof is to show that the inequalities in Eq. (2.16) are all equalities.

1. Suppose that $\varepsilon > 0$ is given and $\delta > 0$ is chosen as in Lemma 2.37 and suppose that $u < s < t < v$ with $|s - u| < \delta$ and $|v - t| < \delta$. Further let $\Pi \in \mathcal{P}(u,v)$ be a partition of $[u,v]$, then according to Lemma 2.37,

$$\begin{aligned}V_1(\rho : \Pi) &\le V_1(\rho : \Pi \cup \{s,t\}) + 2\varepsilon \\ &= \rho(u,s) + \rho(t,v) + V_1(\rho : \Pi \cap [s,t] \cup \{s,t\}) + 2\varepsilon \\ &\le \rho(u,s) + \rho(t,v) + \omega(s,t) + 2\varepsilon.\end{aligned}$$

Letting $s \downarrow u$ and $t \uparrow v$ in this inequality shows,

$$V_1(\rho : \Pi) \le \omega(u+, v-) + 2\varepsilon$$

and then taking the supremum over $\Pi \in \mathcal{P}(u,v)$ and then letting $\varepsilon \downarrow 0$ shows $\omega(u,v) \le \omega(u+,v-)$. Combined this with the first inequality in Eq. (2.16) shows, $\omega(u+,v-) = \omega(u,v)$.

2. We will now show $\omega(u,v) = \omega(u-,v+)$ by showing $\omega(u-,v+) \le \omega(u,v)$. Let $\varepsilon > 0$ and $\delta > 0$ be as in Lemma 2.37 and suppose that $s < u$ and $t > v$ with $|u-s| < \delta$ and $|t-v| < \delta$. Let us now choose a partition $\Pi \in \mathcal{P}(s,t)$ such that

$$\omega(s,t) \le V_1(\rho : \Pi) + \varepsilon.$$

Then applying Lemma 2.37 gives,

$$\omega(s,t) \le V_1(\rho : \Pi_1) + 3\varepsilon$$

where $\Pi_1 = \Pi \cup \{u,v\}$. As above, let u_- and v_+ be the elements in Π_1 just before u and just after v respectively. An application of Lemma 2.36 then shows,

$$\begin{aligned}\omega(u-,v+) &\le \omega(u_-,v_+) \le V_1(\rho : \Pi_1 \cap [u_-,v_+]) + 3\varepsilon \\ &= V_1(\rho : \Pi_1 \cap [u,v]) + \rho(u_-,u) + \rho(v,v_+) + 3\varepsilon \\ &\le \omega(u,v) + 5\varepsilon.\end{aligned}$$

As $\varepsilon > 0$ was arbitrary we may conclude $\omega(u-,v+) \le \omega(u,v)$ which completes the proof that $\omega(u-,v+) = \omega(u,v)$.

I now claim all the other limiting directions follow easily from what we have proved. For example,

$$\begin{aligned}\omega(u,v) \le \omega(u,v+) \le \omega(u-,v+) = \omega(u,v) &\implies \omega(u,v+) = \omega(u,v), \\ \omega(u,v) = \omega(u+,v-) \le \omega(u,v-) \le \omega(u,v) &\implies \omega(u,v-) = \omega(u,v),\end{aligned}$$

and similarly, $\omega(u\pm,v) = \omega(u,v)$. We also have,

$$\omega(u,v) = \omega(u+,v-) \le \liminf_{s\downarrow u,\, t\downarrow v} \omega(s,t) \le \limsup_{s\downarrow u,\, t\downarrow v} \omega(s,t) \le \omega(u-,v+) = \omega(u,v)$$

which shows $\omega(u+,v+) = \omega(u,v)$ and

$$\omega(u,v) = \omega(u+,v-) \le \liminf_{s\uparrow u,\, t\uparrow v} \omega(s,t) \le \liminf_{s\uparrow u,\, t\uparrow v} \omega(s,t) \le \omega(u-,v+) = \omega(u,v)$$

so that $\omega(u-,v-) = \omega(u,v)$. ■

Proposition 2.38 (See [2, Proposition 5.15 from p. 83.]). *Let (E,d) be a metric space, and let $x : [0,T] \to E$ be a continuous path. Then x is of finite p-variation if and only if there exists a continuous increasing (i.e. non – decreasing) function $h : [0,T] \to [0, V_p^p(Z)]$ and a $1/p$ – Hölder path $g : [0, V_p^p(Z)] \to E$ such that such that $x = g \circ h$. More explicitly we have,*

$$d(g(v), g(u)) \le |v-u|^{1/p} \text{ for all } u,v \in [0, V_p^p(Z)]. \tag{2.17}$$

Proof. Let $\omega(s,t) := \omega_{p,x}(s,t) = V_p^p(x|_{[s,t]})$ be the control associated to x and define $h(t) := \omega(0,t)$. Observe that h is increasing and for $0 \le s \le t \le T$ that $h(s) + \omega(s,t) \le h(t)$, i.e.

$$\omega(s,t) \le h(t) - h(s) \text{ for all } 0 \le s \le t \le T.$$

Let $g : [0, h(T)] \to E$ be defined by $g(h(t)) := x(t)$. This is well defined since if $s \le t$ and $h(s) = h(t)$, then $\omega(s,t) = 0$ and hence $x|_{[s,t]}$ is constant and in particular $x(s) = x(t)$. Moreover it now follows for $s < t$ such that $u := h(s) < h(t) =: v$, that

$$\begin{aligned}d^p(g(v), g(u)) &= d^p(g(h(t)), g(h(s))) = d^p(x(t), x(s)) \\ &\le \omega(s,t) \le h(t) - h(s) = v - u\end{aligned}$$

from which Eq. (2.17) easily follows. ■

2.5 Banach Space Structures

This section needs more work and may be moved later.

To put a metric on Hölder spaces seems to require some extra structure on the metric space, E. What is of interest here is the case $E = G$ is a group with a left (right) invariant metric, d. In this case suppose that we consider p - variation paths, x and y starting at $e \in G$ in which case we define,

$$d_{p-var}(x,y) := \sup_{\Pi \in \mathcal{P}(0,T)} \left(\sum_{t \in \Pi} d^p(\Delta_t x, \Delta_t y) \right)^{1/p}$$

where $\Delta_t x := x_{t_-}^{-1} x_t$ for all $t \in \Pi$. The claim is that this should now be a complete metric space.

Lemma 2.39. *$(C_{0,p}(\Delta, T^{(n)}(V)), d_p)$ is a metric space.*

Proof. For each fixed partition D and each $1 \le i \le \lfloor p \rfloor$, we have

$$v_D^i(X) = \left(\sum_{\ell=1}^{r} \left| X^i_{t_{\ell-1} t_\ell} \right|^{p/i} \right)^{i/p}$$

is a semi-norm on $C_{0,p}(\Delta, T^{(n)}(V))$ and in particular satisfies the triangle inequality. Moreover,

$$v_D^i(X+Y) \le \sup_{D'} [v_D^i(X) + v_{D'}^i(Y)] \le \sup_{D'} v_{D'}^i(X) + \sup_{D'} v_{D'}^i(Y)$$

and therefore

$$\sup_D v_D^i(X+Y) \le \sup_D v_D^i(X) + \sup_D v_D^i(Y)$$

which shows $\sup_D v_D^i(X)$ still satisfies the triangle inequality. (i.e., the supremum of a family of semi-norms is a semi-norm). Thus we have

$$d_p(X) := \max_{1\le i\le \lfloor p \rfloor} \sup_D v_D^i(X)$$

is also a semi-norm on $C_{0,p}(\Delta, T^{(n)}(V))$. Thus $d_p(X,Y) = d_p(X-Y)$ satisfies the triangle inequality. Moreover we have $d_p(X,Y) = 0$ implies that

$$\left|X_{st}^i - Y_{st}^i\right|^{p/i} = 0 \quad \forall \;\; 1 \le i \le \lfloor p \rfloor$$

and $(s,t) \in \Delta$, i.e., $X^i = Y^i$ for all $1 \le i \le \lfloor p \rfloor$ and we have verified $d_p(X,Y)$ is a metric. ∎

The Bounded Variation Theory

3.1 Integration Theory

Let $T \in (0, \infty)$ be fixed,

$$\mathcal{S} := \{(a, b] : 0 \leq a \leq b \leq T\} \cup \{[0, b] \cap \mathbb{R} : 0 \leq b \leq T\}. \tag{3.1}$$

Further let $\mathcal{A}$ be the algebra generated by $\mathcal{S}$. Since $\mathcal{S}$ is an elementary set, $\mathcal{A}$ may be described as the collection of sets which are finite disjoint unions of subsets from $\mathcal{S}$. Given any function, $Z : [0, T] \to V$ with V being a vector define $\mu_Z : \mathcal{S} \to V$ via,

$$\mu_Z((a, b]) := Z_b - Z_a \text{ and } \mu_Z([0, b]) = Z_b - Z_0 \ \forall\ 0 \leq a \leq b \leq T.$$

With this definition we are asserting that $\mu_Z(\{0\}) = 0$. Another common choice is to take $\mu_Z(\{0\}) = Z_0$ which would be implemented by taking $\mu_Z([0, b]) = Z_b$ instead of $Z_b - Z_0$.

Lemma 3.1. *μ_Z is finitely additive on $\mathcal{S}$ and hence extends to a finitely additive measure on $\mathcal{A}$.*

Proof. See Chapter **??** and in particular make the minor necessary modifications to Examples **??**, **??**, and Proposition **??**. ∎

Let W be another vector space and $f : [0, T] \to \mathrm{End}(V, W)$ be an $\mathcal{A}$ – **simple function**, i.e. $f([0, T])$ is a finite set and $f^{-1}(\lambda) \in \mathcal{A}$ for all $\lambda \in \mathrm{End}(V, W)$. For such functions we define,

$$\int_{[0,T]} f(t)\, dZ(t) := \int_{[0,T]} f d\mu_Z = \sum_{\lambda \in \mathrm{End}(V,W)^\times} \lambda \mu_Z(f = \lambda) \in W. \tag{3.2}$$

The basic linearity properties of this integral are explained in Proposition **??**. For later purposes, it will be useful to have the following substitution formula at our disposal.

Theorem 3.2 (Substitution formula). *Suppose that f and Z are as above and $Y_t = \int_{[0,t]} f d\mu_Z \in W$. Further suppose that $g : \mathbb{R}_+ \to \mathrm{End}(W, U)$ is another $\mathcal{A}$ – simple function with finite support. Then*

$$\int_{[0,T]} g d\mu_Y = \int_{[0,T]} g f d\mu_Z.$$

Proof. By definition of these finitely additive integrals,

$$\begin{aligned}\mu_Y((a, b]) &= Y_b - Y_a = \int_{[0,b]} f d\mu_Z - \int_{[0,a]} f d\mu_Z \\ &= \int_{[0,T]} \left(1_{[0,b]} - 1_{[0,a]}\right) f d\mu_Z = \int_{[0,T]} 1_{(a,b]} f d\mu_Z.\end{aligned}$$

Therefore, it follows by the finite additivity of μ_Y and linearity $\int_{[0,T]} (\cdot)\, d\mu_Z$, that

$$\mu_Y(A) = \int_A f d\mu_Z = \int_{[0,T]} 1_A f d\mu_Z \text{ for all } A \in \mathcal{A}.$$

Therefore,

$$\begin{aligned}\int_{[0,T]} g d\mu_Y &= \sum_{\lambda \in \mathrm{End}(W,U)^\times} \lambda \mu_Y(g = \lambda) = \sum_{\lambda \in \mathrm{End}(W,U)^\times} \lambda \int_{[0,T]} 1_{\{g=\lambda\}} f d\mu_Z \\ &= \int_{[0,T]} \sum_{\lambda \in \mathrm{End}(W,U)^\times} 1_{\{g=\lambda\}} \lambda f d\mu_Z = \int_{[0,T]} g f d\mu_Z\end{aligned}$$

as desired. ∎

Let us observe that

$$\left\| \int_{[0,T]} f(t)\, dZ_t \right\| \leq \sum_{\lambda \in \mathrm{End}(V,W)} \|\lambda\| \, \|\mu_Z(f = \lambda)\|.$$

Let us now define,

$$\begin{aligned}\|\mu_Z\|((a, b]) &:= V_1\left(Z|_{[a,b]}\right) \\ &= \sup\left\{\sum_{j=1}^n \left\|Z_{t_j} - Z_{t_{j-1}}\right\| : a = t_0 < t_1 < \cdots < t_n = b \text{ and } n \in \mathbb{N}\right\}\end{aligned}$$

be the variation measure associated to μ_Z.

Lemma 3.3. *If $\|\mu_Z\|((0, T]) < \infty$, then $\|\mu_Z\|$ is a finitely additive measure on $\mathcal{S}$ and hence extends to a finitely additive measure on $\mathcal{A}$. Moreover for all $A \in \mathcal{A}$ we have,*

$$\|\mu_Z(A)\| \leq \|\mu_Z\|(A). \tag{3.3}$$

Proof. The additivity on $\mathcal{S}$ was already verified in Lemma 2.34. Here is the proof again for sake of convenience.

Suppose that $\Pi = \{a = t_0 < t_1 < \cdots < t_n = b\}$, $s \in (t_{l-1}, t_l)$ for some l, and $\Pi' := \Pi \cup \{s\}$. Then

$$\begin{aligned}\|\mu_Z\|^{\Pi}((a,b]) &:= \sum_{j=1}^{n} \left\|Z_{t_j} - Z_{t_{j-1}}\right\| \\ &= \sum_{j=1: j\neq l}^{n} \left\|Z_{t_j} - Z_{t_{j-1}}\right\| + \left\|Z_{t_l} - Z_s + Z_s - Z_{t_{l-1}}\right\| \\ &\leq \sum_{j=1: j\neq l}^{n} \left\|Z_{t_j} - Z_{t_{j-1}}\right\| + \left\|Z_{t_l} - Z_s\right\| + \left\|Z_s - Z_{t_{l-1}}\right\| \\ &= \|\mu_Z\|^{\Pi'}((a,b]) \leq \|\mu_Z\|((a,s]) + \|\mu_Z\|((s,b]).\end{aligned}$$

Hence it follows that

$$\|\mu_Z\|((a,b]) = \sup_{\Pi} \|\mu_Z\|^{\Pi}((a,b]) \leq \|\mu_Z\|((a,s]) + \|\mu_Z\|((s,b]).$$

Conversely if Π_1 is a partition of $(a,s]$ and Π_2 is a partition of $(s,b]$, then $\Pi := \Pi_1 \cup \Pi_2$ is a partition of $(a,b]$. Therefore,

$$\|\mu_Z\|^{\Pi_1}((a,s]) + \|\mu_Z\|^{\Pi_2}((s,b]) = \|\mu_Z\|^{\Pi}((a,b]) \leq \|\mu_Z\|((a,b])$$

and therefore,

$$\|\mu_Z\|((a,s]) + \|\mu_Z\|((s,b]) \leq \|\mu_Z\|((a,b]).$$

Lastly if $A \in \mathcal{A}$, then A is the disjoint union of intervals, J_i from $\mathcal{S}$ and we have,

$$\|\mu_Z(A)\| = \left\|\sum_i \mu_Z(J_i)\right\| \leq \sum_i \|\mu_Z(J_i)\| \leq \sum_i \|\mu_Z\|(J_i) = \|\mu_Z\|(A).$$

■

Corollary 3.4. *If Z has finite variation on $[0,T]$, then we have*

$$\left\|\int_{[0,T]} f(t)\, dZ_t\right\| \leq \int_{[0,T]} \|f(\lambda)\| \, \|\mu_Z\|(d\lambda) \leq \|f\|_\infty \cdot \|\mu_Z\|([0,T]).$$

Proof. Simply observe that $\|\mu_Z(A)\| \leq \|\mu_Z\|(A)$ for all $A \in \mathcal{A}_T$ and hence from Eq. (3.2) and the bound in Eq. (3.3) we have

$$\begin{aligned}\left\|\int_{[0,T]} f(t)\, dZ_t\right\| &\leq \sum_{\lambda \in \mathrm{End}(V,W)} \|\lambda\| \, \|\mu_Z(f = \lambda)\| \\ &\leq \sum_{\lambda \in \mathrm{End}(V,W)} \|\lambda\| \, \|\mu_Z\|(f = \lambda) = \int_{[0,T]} \|f(\lambda)\| \, \|\mu_Z\|(d\lambda) \\ &\leq \|f\|_\infty \cdot \|\mu_Z\|([0,T]).\end{aligned}$$

■

Notation 3.5 *In the future we will often write $\|dZ\|$ for $d\|\mu_Z\|$.*

Theorem 3.6. *If V and W are Banach spaces and $V_1(Z) = \|\mu_Z\|([0,T]) < \infty$, we may extend the integral, $\int_{[0,T]} f(t)\, dZ_t$, by continuity to all functions which are in the uniform closure of the $\mathcal{A}$ – simple functions. In fact we may extend the integral to $L^1(\|\mu_Z\|)$ – closure of the $\mathcal{A}$ – simple functions. In particular, if $f : [0,T] \to \mathrm{Hom}(V,W)$ is a continuous function,*

$$\int_{[0,T]} f(t)\, dZ(t) = \lim_{|\Pi| \to 0} \sum_{t \in \Pi} f(t_-)(Z(t) - Z(t_-)). \tag{3.4}$$

Proof. These results are elemetary soft analysis except possibly for the last assertion for the statement in Eq. (3.4). To prove this, to any partition, $\Pi \in \mathcal{P}(0,T)$, let

$$f_\Pi := \sum_{\tau \in \Pi} f(t_-) 1_{(t_-, t]} + f(0) 1_{\{0\}}$$

in which case,

$$\sum_{t \in \Pi} f(t_-)(Z(t) - Z(t_-)) = \int_{[0,T]} f_\Pi(t)\, dZ(t).$$

This completes the proof since $f_\Pi \to f$ uniformly on $[0,T]$ as $|\Pi| \to 0$ by the uninform continuity of f. ■

Theorem 3.7 (Substitution formula II). *Let $Z : [0,T] \to V$ be a finite variation process, $f : [0,T] \to \mathrm{End}(V,W)$ and $g : [0,T] \to \mathrm{End}(W,U)$ be continuous maps and define,*

$$Y_t = \int_{[0,t]} f dZ \in W.$$

Then Y is a continuous finite variation process and the following substitution formula holds,

$$\int_{[0,T]} g dY = \int_{[0,T]} g f dZ. \tag{3.5}$$

In short, $dY = f dZ$.

Proof. First off observe that

$$\|Y_t - Y_s\| \leq \int_s^t \|f\| \, \|dZ\| =: \omega(s,t)$$

where the right side is a continuous control. This follows from the fact that $\|dZ\|$ is a continuous measure. Therefore

$$V_1(Y) \leq \int_0^T \|f\| \, \|dZ\| < \infty.$$

If $g = \lambda 1_{(a,b]}$ with $\lambda \in \mathrm{End}(W,U)$, then

$$\int_{[0,T]} g dY = \lambda (Y_b - Y_a) = \lambda \int_a^b f dZ = \int_a^b \lambda f dZ = \int_0^T g f dZ.$$

Thus Eq. (3.5) holds for all $\mathcal{A}$ – simple functions and hence also for all uniform limits of simple functions. In particular this includes all continuous $g : [0,T] \to \mathrm{End}(W,U)$. ∎

Remark 3.8. If we keep the same hypothesis as in Theorem 3.7 but now take $Y_t := \int_t^T f dZ$ instead. In this case we have,

$$\int_{[0,T]} g dY = -\int_{[0,T]} g f dZ.$$

To prove this just observe that $Y_t = W_T - W_t$ where $W_t := \int_0^t f dZ$. It is now easy to see that

$$dY_t = d(-W_t) = -dW_t = -f dZ$$

and the claim follows.

3.2 The Fundamental Theorem of Calculus

As above, let V and W be Banach spaces and $0 \leq a < b \leq T$.

Proposition 3.9. *Suppose that $f : [a,b] \to V$ is a continuous function such that $\dot{f}(t)$ exists and is equal to zero for $t \in (a,b)$. Then f is constant.*

Proof. First Proof. For $\ell \in V^*$, we have $f_\ell := \ell \circ f : [a,b] \to \mathbb{R}$ with $\dot{f}_\ell(t) = 0$ for all $t \in (a,b)$. Therefore by the mean value theory, it follows that $f_\ell(t)$ is constant, i.e. $\ell(f(t) - f(a)) = 0$ for all $t \in [a,b]$. Since $\ell \in V^*$ is arbitrary, it follows from the Hahn – Banach theorem that $f(t) - f(a) = 0$, i.e. $f(t) = f(a)$ independent of t.

Second Proof (with out Hahn − Banach). Let $\varepsilon > 0$ and $\alpha \in (a,b)$ be given. (We will later let $\varepsilon \downarrow 0$.) By the definition of the derivative, for all $\tau \in (a,b)$ there exists $\delta_\tau > 0$ such that

$$\|f(t) - f(\tau)\| = \left\| f(t) - f(\tau) - \dot{f}(\tau)(t-\tau) \right\| \leq \varepsilon |t-\tau| \text{ if } |t-\tau| < \delta_\tau. \tag{3.6}$$

Let

$$A = \{t \in [\alpha, b] : \|f(t) - f(\alpha)\| \leq \varepsilon(t-\alpha)\} \tag{3.7}$$

and t_0 be the least upper bound for A. We will now use a standard argument which is sometimes referred to as **continuous induction** to show $t_0 = b$. Eq. (3.6) with $\tau = \alpha$ shows $t_0 > \alpha$ and a simple continuity argument shows $t_0 \in A$, i.e.

$$\|f(t_0) - f(\alpha)\| \leq \varepsilon(t_0 - \alpha). \tag{3.8}$$

For the sake of contradiction, suppose that $t_0 < b$. By Eqs. (3.6) and (3.8),

$$\begin{aligned}\|f(t) - f(\alpha)\| &\leq \|f(t) - f(t_0)\| + \|f(t_0) - f(\alpha)\| \\ &\leq \varepsilon(t_0 - \alpha) + \varepsilon(t - t_0) = \varepsilon(t - \alpha)\end{aligned}$$

for $0 \leq t - t_0 < \delta_{t_0}$ which violates the definition of t_0 being an upper bound. Thus we have shown $b \in A$ and hence

$$\|f(b) - f(\alpha)\| \leq \varepsilon(b - \alpha).$$

Since $\varepsilon > 0$ was arbitrary we may let $\varepsilon \downarrow 0$ in the last equation to conclude $f(b) = f(\alpha)$. Since $\alpha \in (a,b)$ was arbitrary it follows that $f(b) = f(\alpha)$ for all $\alpha \in (a,b]$ and then by continuity for all $\alpha \in [a,b]$, i.e. f is constant. ∎

Theorem 3.10 (Fundamental Theorem of Calculus). *Suppose that $f \in C([a,b],V)$, Then*

1. *$\frac{d}{dt}\int_a^t f(\tau)\, d\tau = f(t)$ for all $t \in (a,b)$.*
2. *Now assume that $F \in C([a,b],V)$, F is continuously differentiable on (a,b) (i.e. $\dot{F}(t)$ exists and is continuous for $t \in (a,b)$) and $\dot{F}$ extends to a continuous function on $[a,b]$ which is still denoted by $\dot{F}$. Then*

$$\int_a^b \dot{F}(t)\, dt = F(b) - F(a). \tag{3.9}$$

Proof. Let $h > 0$ be a small number and consider

$$\begin{aligned}\left\| \int_a^{t+h} f(\tau) d\tau - \int_a^t f(\tau) d\tau - f(t) h \right\| &= \left\| \int_t^{t+h} (f(\tau) - f(t))\, d\tau \right\| \\ &\leq \int_t^{t+h} \|(f(\tau) - f(t))\|\, d\tau \leq h\varepsilon(h),\end{aligned}$$

where $\varepsilon(h) := \max_{\tau\in[t,t+h]} \|(f(\tau) - f(t))\|$. Combining this with a similar computation when $h < 0$ shows, for all $h \in \mathbb{R}$ sufficiently small, that

$$\left\| \int_a^{t+h} f(\tau) d\tau - \int_a^t f(\tau) d\tau - f(t)h \right\| \le |h|\varepsilon(h),$$

where now $\varepsilon(h) := \max_{\tau\in[t-|h|,t+|h|]} \|(f(\tau) - f(t))\|$. By continuity of f at t, $\varepsilon(h) \to 0$ and hence $\frac{d}{dt}\int_a^t f(\tau)\, d\tau$ exists and is equal to $f(t)$.

For the second item, set $G(t) := \int_a^t \dot{F}(\tau)\, d\tau - F(t)$. Then G is continuous by Lemma **??** and $\dot{G}(t) = 0$ for all $t \in (a, b)$ by item 1. An application of Proposition 3.9 shows G is a constant and in particular $G(b) = G(a)$, i.e. $\int_a^b \dot{F}(\tau)\, d\tau - F(b) = -F(a)$.

Alternative proof of Eq. (3.9). It is easy to show

$$\ell\left(\int_a^b \dot{F}(t)\, dt\right) = \int_a^b \ell\circ\dot{F}(t)\, dt = \int_a^b \frac{d}{dt}(\ell\circ F)(t)\, dt \text{ for all } \ell\in V^*.$$

Moreover by the real variable fundamental theorem of calculus we have,

$$\int_a^b \frac{d}{dt}(\ell\circ F)(t)\, dt = \ell\circ F(b) - \ell\circ F(a) \text{ for all } \ell\in V^*.$$

Combining the last two equations implies,

$$\ell\left(\int_a^b \dot{F}(t)\, dt - F(b) + F(a)\right) = 0 \text{ for all } \ell\in V^*.$$

Equation (3.9) now follows from these identities after an application of the Hahn – Banach theorem. ∎

Corollary 3.11 (Mean Value Inequality). *Suppose that $f : [a, b] \to V$ is a continuous function such that $\dot{f}(t)$ exists for $t \in (a, b)$ and $\dot{f}$ extends to a continuous function on $[a, b]$. Then*

$$\|f(b) - f(a)\| \le \int_a^b \|\dot{f}(t)\|dt \le (b-a)\cdot\left\|\dot{f}\right\|_\infty. \tag{3.10}$$

Proof. By the fundamental theorem of calculus, $f(b) - f(a) = \int_a^b \dot{f}(t)dt$ and then (by the triangle inequality for integrals)

$$\begin{aligned}\|f(b) - f(a)\| &= \left\|\int_a^b \dot{f}(t)dt\right\| \le \int_a^b \|\dot{f}(t)\|dt \\ &\le \int_a^b \left\|\dot{f}\right\|_\infty dt = (b-a)\cdot\left\|\dot{f}\right\|_\infty.\end{aligned}$$

∎

Corollary 3.12 (Change of Variable Formula). *Suppose that $f \in C([a, b], V)$ and $T : [c, d] \to (a, b)$ is a continuous function such that $T(s)$ is continuously differentiable for $s \in (c, d)$ and $T'(s)$ extends to a continuous function on $[c, d]$. Then*

$$\int_c^d f(T(s))\, T'(s)\, ds = \int_{T(c)}^{T(d)} f(t)\, dt.$$

Proof. For $t \in (a, b)$ define $F(t) := \int_{T(c)}^t f(\tau)\, d\tau$. Then $F \in C^1((a, b), V)$ and by the fundamental theorem of calculus and the chain rule,

$$\frac{d}{ds} F(T(s)) = F'(T(s))\, T'(s) = f(T(s))\, T'(s).$$

Integrating this equation on $s \in [c, d]$ and using the chain rule again gives

$$\int_c^d f(T(s))\, T'(s)\, ds = F(T(d)) - F(T(c)) = \int_{T(c)}^{T(d)} f(t)\, dt.$$

∎

Exercise 3.1 (Fundamental Theorem of Calculus II). Prove the fundamental theorem of calculus in this context. That is; if $f : V \to W$ be a C^1 – function and $\{Z_t\}_{t\ge0}$ is a V – valued function of locally bounded variation, then for all $0 \le a < b \le T$,

$$f(Z_b) - f(Z_a) = \int_a^b f'(Z_\tau)\, dZ_\tau := \int_{[a,b]} f'(Z_\tau)\, dZ_\tau,$$

where $f'(z) \in \mathrm{End}(V, W)$ is defined by, $f'(z)\, v := \frac{d}{dt}|_0 f(z + tv)$. In particular it follows that $f(Z(t))$ has finite variation and

$$df(Z(t)) = f'(Z(t))\, dZ(t).$$

Solution to Exercise (3.1). Let $\Pi \in \mathcal{P}(0, T)$. Then by a telescoping series argument,

$$f(Z_b) - f(Z_a) = \sum_{t\in\Pi} \Delta_t f(Z_\cdot)$$

where

$$\begin{aligned}\Delta_t f(Z_\cdot) &= f(Z_t) - f\left(Z_{t_-}\right) = f\left(Z_{t_-} + \Delta_t Z\right) - f\left(Z_{t_-}\right) \\ &= \int_0^1 f'\left(Z_{t_-} + s\Delta_t Z\right)\Delta_t Z\, ds = f'\left(Z_{t_-}\right)\Delta_t Z + \varepsilon_t^\Pi \Delta_t Z\end{aligned}$$

and

$$\varepsilon_t^{\Pi} := \int_0^1 \left[f'\left(Z_{t_-} + s\Delta_t Z\right) - f'\left(Z_{t_-}\right)\right] \, ds.$$

Thus we have,

$$f(Z_b) - f(Z_a) = \sum_{t\in\Pi} f'\left(Z_{t_-}\right)\Delta_t Z + \delta_\Pi = \int_{[a,b]} f'\left(Z_{t_-}\right) dZ(t) + \delta_\Pi \quad (3.11)$$

where $\delta_\Pi := \sum_{t\in\Pi} \varepsilon_t^\Pi \Delta_t Z$. Since,

$$\begin{aligned}\|\delta_\Pi\| &\le \sum_{t\in\Pi}\left\|\varepsilon_t^\Pi \Delta_t Z\right\| \le \sum_{t\in\Pi}\left\|\varepsilon_t^\Pi\right\| \|\Delta_t Z\| \le \max_{t\in\Pi}\left\|\varepsilon_t^\Pi\right\| \cdot \sum_{t\in\Pi}\|\Delta_t Z\| \\ &\le \max_{t\in\Pi}\left\|\varepsilon_t^\Pi\right\| \cdot V_1(Z),\end{aligned}$$

and

$$\left\|\varepsilon_t^\Pi\right\| := \int_0^1 \left\|\left[f'\left(Z_{t_-} + s\Delta_t Z\right) - f'\left(Z_{t_-}\right)\right]\right\| ds.$$

Since $g(s,\tau,t) := \|[f'(Z_\tau + s(Z_t - Z_\tau)) - f'(Z_\tau)]\|$ is a continuous function in $s \in [0,1]$ and $\tau, t \in [0,T]$ with $g(s,t,t) = 0$ for all s and t, it follows by uniform continuity arguments that $g(s,\tau,t)$ is small whenever $|t-\tau|$ is small. Therefore, $\lim_{|\Pi|\to 0}\left\|\varepsilon_t^\Pi\right\| = 0$. Moreover, again by a uniform continuity argument, $f'\left(Z_{t_-}\right) \to f'(Z_t)$ uniformly as $|\Pi| \to 0$. Thus we may pass to the limit as $|\Pi| \to 0$ in Eq. (3.11) to complete the proof.

3.3 Calculus Bounds

For the exercises to follow we suppose that μ is a positive σ − finite measure on $\left([0,\infty), \mathcal{B}_{[0,\infty)}\right)$ such that μ is continuous, i.e. $\mu(\{s\}) = 0$ for all $s \in [0,\infty)$. We will further write,

$$\int_0^t f(s)\, d\mu(s) := \int_{[0,t]} f(s)\, d\mu(s) = \int_{(0,t)} f(s)\, d\mu(s),$$

wherein the second equality holds since μ is continuous. Although it is not necessary, you may use Exercise 3.1 with $Z_t := \mu([0,t])$ to solve the following problems.

Exercise 3.2. Show for all $0 \le a < b < \infty$ and $n \in \mathbb{N}$ that

$$h_n(b) := \int_{a\le s_1\le s_2\le\dots\le s_n\le b} d\mu(s_1)\dots d\mu(s_n) = \frac{\mu([a,b])^n}{n!}. \quad (3.12)$$

Solution to Exercise (3.2). First solution. Let us observe that $h(t) := h_1(t) = \mu([a,t])$ and $h_n(t)$ satisfies the recursive relation,

$$h_{n+1}(t) := \int_a^t h_n(s)\, d\mu(s) = \int_a^t h_n(s)\, dh(s) \text{ for all } t \ge a.$$

Now let $H_n(t) := \frac{1}{n!} h^n(t)$, by an application of Exercise 3.1 with $f(x) = x^{n+1}/(n+1)!$ implies,

$$H_{n+1}(t) = H_{n+1}(t) - H_{n+1}(a) = \int_a^t f'(h(\tau))\, dh(\tau) = \int_a^t H_n(\tau)\, dh(\tau)$$

and therefore it follows that $H_n(t) = h_n(t)$ for all $t \ge a$ and $n \in \mathbb{N}$.

Second solution. If $i \ne j$, it follows by Fubini's theorem that

$$\begin{aligned}&\mu^{\otimes n}\left(\{(s_1,\dots,s_n) \in [a,b]^n : s_i = s_j\}\right) \\ &= \mu([a,b])^{n-2} \cdot \int_{[a,b]^2} 1_{s_i = s_j} d\mu(s_i)\, d\mu(s_j) \\ &= \mu([a,b])^{n-2} \cdot \int_{[a,b]} \mu(\{s_j\})\, d\mu(s_j) = 0.\end{aligned}$$

From this observation it follows that

$$1_{[a,b]^n}(s_1,\dots,s_n) = \sum_{\sigma\in\mathcal{S}_n} 1_{a\le s_{\sigma 1}\le s_{\sigma 2}\le\dots\le s_{\sigma n}\le b} - \mu^{\otimes n} - \text{a.e.},$$

where σ ranges over the permutations, $\mathcal{S}_n$, of $\{1,2,\dots,n\}$. Integrating this equation relative with respect to $\mu^{\otimes n}$ and then using Fubini's theorem gives,

$$\begin{aligned}\mu([a,b])^n = \mu^{\otimes n}([a,b]^n) &= \sum_{\sigma\in\mathcal{S}_n} \int 1_{a\le s_{\sigma 1}\le s_{\sigma 2}\le\dots\le s_{\sigma n}\le b} d\mu^{\otimes n}(\mathbf{s}) \\ &= \sum_{\sigma\in\mathcal{S}_n} \int 1_{a\le s_{\sigma 1}\le s_{\sigma 2}\le\dots\le s_{\sigma n}\le b} d\mu(s_1)\dots d\mu(s_n) \\ &= \sum_{\sigma\in\mathcal{S}_n} \int_{a\le s_1\le s_2\le\dots\le s_n\le b} d\mu(s_1)\dots d\mu(s_n) \\ &= n! \int_{a\le s_1\le s_2\le\dots\le s_n\le b} d\mu(s_1)\dots d\mu(s_n).\end{aligned}$$

Exercise 3.3 (Gronwall's Lemma). If $\varepsilon(t)$ and $f(t)$ are continuous non-negative functions such that

$$f(t) \le \varepsilon(t) + \int_0^t f(\tau)\, d\mu(\tau), \quad (3.13)$$

then

$$f(t) \le \varepsilon(t) + \int_0^t e^{\mu([\tau,t])}\varepsilon(\tau)\,d\mu(\tau). \tag{3.14}$$

If we further assume that ε is increasing, then

$$f(t) \le \varepsilon(t)\,e^{\mu([0,t])}. \tag{3.15}$$

Solution to Exercise (3.3). Feeding Eq. (3.13) back into itself implies

$$\begin{aligned} f(t) &\le \varepsilon(t) + \int_0^t \left[\varepsilon(\tau) + \int_0^\tau f(s)\,d\mu(s)\right] d\mu(\tau) \\ &= \varepsilon(t) + \int_0^t \varepsilon(s_1)\,d\mu(s_1) + \int_{0\le s_2\le s_1\le t} f(s_2)\,d\mu(s_1)\,d\mu(s_2) \\ &\le \varepsilon(t) + \int_0^t \varepsilon(s_1)\,d\mu(s_1) + \int_{0\le s_2\le s_1\le t}\left[\varepsilon(s_2) + \int_0^{s_2} f(s_3)\,d\mu(s_3)\right] d\mu(s_1)\,d\mu(s_2) \\ &= \varepsilon(t) + \int_0^t \varepsilon(s_1)\,d\mu(s_1) + \int_{0\le s_2\le s_1\le t} \varepsilon(s_2)\,d\mu(s_1)\,d\mu(s_2) \\ &\quad + \int_{0\le s_3\le s_2\le s_1\le t} f(s_3)\,d\mu(s_1)\,d\mu(s_2)\,d\mu(s_3). \end{aligned}$$

Continuing in this manner inductively shows,

$$f(t) \le \varepsilon(t) + \sum_{k=1}^N \int_{0\le s_k\le\dots\le s_2\le s_1\le t} \varepsilon(s_k)\,d\mu(s_1)\dots d\mu(s_k) + R_N(t) \tag{3.16}$$

where, using Exercise 3.2,

$$\begin{aligned} R_N(t) &:= \int_{0\le s_{k+1}\le\dots\le s_2\le s_1\le t} f(s_{k+1})\,d\mu(s_1)\dots d\mu(s_k)\,d\mu(s_{k+1}) \\ &\le \max_{0\le s\le t} f(t)\cdot\frac{\mu([0,t])^{N+1}}{(N+1)!} \to 0 \text{ as } N\to\infty. \end{aligned}$$

So passing to the limit in Eq. (3.16) and again making use of Exercise 3.2 shows,

$$\begin{aligned} f(t) &\le \varepsilon(t) + \sum_{k=1}^\infty \int_{0\le s_k\le\dots\le s_2\le s_1\le t} \varepsilon(s_k)\,d\mu(s_1)\dots d\mu(s_k) \qquad (3.17) \\ &= \varepsilon(t) + \sum_{k=1}^\infty \int_0^t \varepsilon(s_k)\frac{\mu([s_k,t])^{k-1}}{(k-1)!}d\mu(s_k) \\ &= \varepsilon(t) + \int_0^t \varepsilon(\tau)\cdot\sum_{k=1}^\infty \frac{\mu([\tau,t])^{k-1}}{(k-1)!}d\mu(\tau) \\ &= \varepsilon(t) + \int_0^t e^{\mu([\tau,t])}\varepsilon(\tau)\,d\mu(\tau). \end{aligned}$$

If we further assume that ε is increasing, then from Eq. (3.17) and Exercise 3.2 we have

$$\begin{aligned} f(t) &\le \varepsilon(t) + \varepsilon(t)\sum_{k=1}^\infty \int_{0\le s_k\le\dots\le s_2\le s_1\le t} d\mu(s_1)\dots d\mu(s_k) \\ &= \varepsilon(t) + \varepsilon(t)\sum_{k=1}^\infty \frac{\mu([0,t])^k}{k!} = \varepsilon(t)\,e^{\mu([0,t])}. \end{aligned}$$

Alternatively if we let $Z_t := \mu([0,t])$, then

$$\begin{aligned} \int_0^t e^{\mu([\tau,t])}d\mu(\tau) &= \int_0^t e^{Z_t - Z_\tau}dZ_\tau = \int_0^t d_\tau\left(-e^{Z_t - Z_\tau}\right) \\ &= \left(-e^{Z_t - Z_\tau}\right)_0^t = e^{Z_t} - 1. \end{aligned}$$

Therefore,

$$f(t) \le \varepsilon(t) + \varepsilon(t)\left(e^{Z_t} - 1\right) = \varepsilon(t)\,e^{Z_t}.$$

Exercise 3.4. Suppose that $\{\varepsilon_n(t)\}_{n=0}^\infty$ is a sequence of non-negative continuous functions such that

$$\varepsilon_{n+1}(t) \le \int_0^t \varepsilon_n(\tau)\,d\mu(\tau) \text{ for all } n\ge 0 \tag{3.18}$$

and $\delta(t) = \max_{0\le\tau\le t}\varepsilon_0(\tau)$. Show

$$\varepsilon_n(t) \le \delta(t)\frac{\mu([0,t])^n}{n!} \text{ for all } n\ge 0.$$

Solution to Exercise (3.4). By iteration of Eq. (3.18) we find,

$$\begin{aligned} \varepsilon_1(t) &\le \int_0^t \varepsilon_0(\tau)\,d\mu(\tau) \le \delta(t)\int_{0\le s_1\le t} d\mu(s_1), \\ \varepsilon_2(t) &\le \int_0^t \varepsilon_1(s_2)\,d\mu(s_2) \le \delta(t)\int_0^t\left[\int_{0\le s_1\le t} d\mu(s_1)\right] d\mu(s_2) \\ &= \delta(t)\int_{0\le s_2\le s_1\le t} d\mu(s_1)\,d\mu(s_2), \\ &\vdots \\ \varepsilon_n(t) &\le \delta(t)\int_{0\le s_n\le\dots\le s_1\le t} d\mu(s_1)\dots d\mu(s_n). \end{aligned}$$

The result now follows directly from Exercise 3.2.

3.4 Bounded Variation Ordinary Differential Equations

In this section we begin by reviewing some of the basic theory of ordinary differential equations – O.D.E.s for short. Throughout this chapter we will let X and Y be Banach spaces, $U \subset_o Y$ an open subset of Y, and $y_0 \in U$, $x : [0,T] \to X$ is a continuous process of bounded variation, and $F : [0,T] \times U \to \mathrm{End}(X,Y)$ is a continuous function. (We will make further assumptions on F as we need them.) Our goal here is to investigate the "ordinary differential equation,"

$$\dot{y}(t) = F(t, y(t))\, \dot{x}(t) \text{ with } y(0) = y_0 \in U. \tag{3.19}$$

Since x is only of bounded variation, to make sense of this equation we will interpret it in its integrated form,

$$y(t) = y_0 + \int_0^t F(\tau, y(\tau))\, dx(\tau). \tag{3.20}$$

Proposition 3.13 (Continuous dependence on the data). *Suppose that* $G : [0,T] \times U \to \mathrm{End}(X,Y)$ *is another continuous function,* $z : [0,T] \to X$ *is another continuous function with bounded variation, and* $w : [0,T] \to U$ *satisfies the differential equation,*

$$w(t) = w_0 + \int_0^t G(\tau, w(\tau))\, dz(\tau) \tag{3.21}$$

for some $w_0 \in U$. *Further assume there exists a continuous function,* $K(t) \geq 0$ *such that* F *satisfies the* ***Lipschitz condition,***

$$\|F(t,y) - F(t,w)\| \leq K(t)\, \|y - w\| \text{ for all } 0 \leq t \leq T \text{ and } y, w \in U. \tag{3.22}$$

Then

$$\|y(t) - w(t)\| \leq \varepsilon(t) \exp\left(\int_0^t K(\tau)\, \|dx(\tau)\|\right). \tag{3.23}$$

where

$$\begin{aligned}\varepsilon(t) := \|y_0 - w_0\| &+ \int_0^t \|F(\tau, w(\tau)) - G(\tau, w(\tau))\|\, \|dx(\tau)\| \\ &+ \int_0^t \|G(\tau, w(\tau))\|\, \|d(x - z)(\tau)\| \end{aligned} \tag{3.24}$$

Proof. Let $\delta(t) := y(t) - w(t)$, so that $y = w + \delta$. We then have,

$$\begin{aligned}\delta(t) &= y_0 - w_0 + \int_0^t F(\tau, y(\tau))\, dx(\tau) - \int_0^t G(\tau, w(\tau))\, dz(\tau) \\ &= y_0 - w_0 + \int_0^t F(\tau, w(\tau) + \delta(\tau))\, dx(\tau) - \int_0^t G(\tau, w(\tau))\, dz(\tau) \\ &= y_0 - w_0 + \int_0^t [F(\tau, w(\tau)) - G(\tau, w(\tau))]\, dx(\tau) + \int_0^t G(\tau, w(\tau))\, d(x - z)(\tau) \\ &+ \int_0^t [F(\tau, w(\tau) + \delta(\tau)) - F(\tau, w(\tau))]\, dx(\tau).\end{aligned}$$

Crashing through this identity with norms shows,

$$\|\delta(t)\| \leq \varepsilon(t) + \int_0^t K(\tau)\, \|\delta(\tau)\|\, \|dx(\tau)\|$$

where $\varepsilon(t)$ is given in Eq. (3.24). The estimate in Eq. (3.23) is now a consequence of this inequality and Exercise 3.3 with $d\mu(\tau) := K(\tau)\, \|dx(\tau)\|$. ■

Corollary 3.14 (Uniquness of solutions). *If* F *satisfies the Lipschitz hypothesis in Eq. (3.22), then there is at most one solution to the ODE in Eq. (3.20).*

Proof. Simply apply Proposition 3.13 with $F = G$, $y_0 = w_0$, and $x = z$. In this case $\varepsilon \equiv 0$ and the result follows. ■

Proposition 3.15 (An apriori growth bound). *Suppose that* $U = Y$, $T = \infty$, *and there are continuous functions,* $a(t) \geq 0$ *and* $b(t) \geq 0$ *such that*

$$\|F(t,y)\| \leq a(t) + b(t)\, \|y\| \text{ for all } t \geq 0 \text{ and } y \in Y.$$

Then

$$\|y(t)\| \leq \left(\|y_0\| + \int_0^t a(\tau)\, d\nu(\tau)\right) \exp\left(\int_0^t b(\tau)\, d\nu(\tau)\right), \text{ where} \tag{3.25}$$

$$\nu(t) := \omega_{x,1}(0,t) = \|s\|_{1\text{-}Var}(t). \tag{3.26}$$

Proof. From Eq. (3.20) we have,

$$\begin{aligned}\|y(t)\| &\leq \|y_0\| + \int_0^t \|F(\tau, y(\tau))\|\, d\nu(\tau) \\ &\leq \|y_0\| + \int_0^t (a(\tau) + b(\tau)\, \|y(\tau)\|)\, d\nu(\tau) \\ &= \varepsilon(t) + \int_0^t \|y(\tau)\|\, d\mu(\tau)\end{aligned}$$

where

$$\varepsilon(t) := \|y_0\| + \int_0^t a(\tau)\, d\nu(\tau) \text{ and } d\mu(\tau) := b(\tau)\, d\nu(\tau).$$

Hence we may apply Exercise 3.3 to learn $\|y(t)\| \le \varepsilon(t) e^{\mu([0,t])}$ which is the same as Eq. (3.25). ■

Theorem 3.16 (Global Existence). *Let us now suppose $U = X$ and F satisfies the Lipschitz hypothesis in Eq. (3.22). Then there is a unique solution, $y(t)$ to the ODE in Eq. (3.20).*

Proof. We will use the standard method of Picard iterates. Namely let $y_0(t) \in W$ be **any** continuous function and then define $y_n(t)$ inductively by,

$$y_{n+1}(t) := y_0 + \int_0^t F(\tau, y_n(\tau))\, dx(\tau). \tag{3.27}$$

Then from our assumptions and the definition of $y_n(t)$, we find for $n \ge 1$ that

$$\begin{aligned}\|y_{n+1}(t) - y_n(t)\| &= \left\|\int_0^t F(\tau, y_n(\tau))\, dx(\tau) - \int_0^t F(\tau, y_{n-1}(\tau))\, dx(\tau)\right\| \\ &\le \int_0^t \|F(\tau, y_n(\tau)) - F(\tau, y_{n-1}(\tau))\| \|dx(\tau)\| \\ &\le \int_0^t K(\tau) \|y_n(\tau) - y_{n-1}(\tau)\| \|dx(\tau)\|.\end{aligned}$$

Since,

$$\begin{aligned}\|y_1(t) - y_0(t)\| &= \left\|y_0 + \int_0^t F(\tau, y_0(\tau))\, dx(\tau) - y_0(t)\right\| \\ &\le \max_{0\le\tau\le t} \|y_0(\tau) - y_0\| + \int_0^t \|F(\tau, y_0)\| \|dx(\tau)\| =: \delta(t),\end{aligned}$$

it follows by an application of Exercise 3.4 with

$$\varepsilon_n(t) := \|y_{n+1}(t) - y_n(t)\|$$

that

$$\|y_{n+1}(t) - y_n(t)\| \le \delta(t) \cdot \left(\int_0^t K(\tau) \|dx(\tau)\|\right)^n / n!. \tag{3.28}$$

Since the right side of this equation is increasing in t, we may conclude by summing Eq. (3.28) that

$$\sum_{n=0}^{\infty} \sup_{0\le t\le T} \|y_{n+1}(t) - y_n(t)\| \le \delta(T) e^{\int_0^T K(\tau)\|dx(\tau)\|} < \infty.$$

Therefore, it follows that $y_n(t)$ is uniformly convergent on compact subsets of $[0,\infty)$ and therefore $y(t) := \lim_{n\to\infty} y_n(t)$ exists and is a continuous function. Moreover, we may now pass to the limit in Eq. (3.27) to learn this function y satisfies Eq. (3.20). Indeed,

$$\begin{aligned}&\left\|\int_0^t F(\tau, y_n(\tau))\, dx(\tau) - \int_0^t F(\tau, y(\tau))\, dx(\tau)\right\| \\ &\quad\le \int_0^t \|F(\tau, y_n(\tau)) - F(\tau, y(\tau))\| \|dx(\tau)\| \\ &\quad\le \int_0^t K(\tau) \|y_n(\tau) - y(\tau)\| \|dx(\tau)\| \\ &\quad\le \sup_{0\le\tau\le t} \|y_n(\tau) - y(\tau)\| \cdot \int_0^t K(\tau) \|dx(\tau)\| \to 0 \text{ as } n \to \infty.\end{aligned}$$

■

Remark 3.17 (Independence of initial guess). In the above proof, we were allowed to choose $y_0(t)$ as we pleased. In all cases we ended up with a solution to the ODE which we already knew to be unique if it existed. Therefore all initial guesses give rise to the same solution. This can also be see directly. Indeed, if $z_0(t)$ is another continuous path in W and $z_n(t)$ is defined inductively by,

$$z_{n+1}(t) := y_0 + \int_0^t F(\tau, z_n(\tau))\, dx(\tau) \text{ for } n \ge 0.$$

Then

$$z_{n+1}(t) - y_{n+1}(t) = \int_0^t [F(\tau, z_n(\tau)) - F(\tau, y_n(\tau))]\, dx(\tau)$$

and therefore,

$$\begin{aligned}\|z_{n+1}(t) - y_{n+1}(t)\| &\le \int_0^t \|F(\tau, z_n(\tau)) - F(\tau, y_n(\tau))\| \|dx(\tau)\| \\ &\le \int_0^t K(\tau) \|z_n(\tau) - y_n(\tau)\| \|dx(\tau)\|.\end{aligned}$$

Thus it follows from Exercise 3.4 that

$$\|z_n(t) - y_n(t)\| \le \frac{1}{n!}\left(\int_0^t K(\tau) \|dx(\tau)\|\right)^n \cdot \max_{0\le\tau\le t} \|z_0(\tau) - y_0(\tau)\| \to 0 \text{ as } n \to \infty.$$

3.5 Some Linear ODE Results

In this section we wish to consider linear ODE of the form,

$$y(t) = \int_0^t dx(\tau)\, y(\tau) + f(t) \tag{3.29}$$

where $x(t) \in \mathrm{End}(W)$ and $f(t) \in W$ are finite variation paths. To put this in the form considered above, let $V := \mathrm{End}(W)$ and define, for $y \in W$, $F(y) : V \times W \to W$ by,

$$F(y)(x, f) := xy + f \text{ for all } (x, f) \in V \times W = \mathrm{End}(W) \times W.$$

Then the above equation may be written as,

$$y(t) = f(0) + \int_0^t F(y(\tau))\, d(x, f)(\tau).$$

Notice that

$$\|[F(y) - F(y')](x, f)\| = \|x(y - y')\| \le \|x\| \cdot \|y - y'\|$$

and therefore,

$$\|F(y) - F(y')\| \le \|y - y'\|$$

where we use any reasonable norm on $V \times W$, for example $\|(x, w)\| := \|x\| + \|w\|$ or $\|(x, w)\| := \max(\|x\|, \|w\|)$. Thus the theory we have developed above guarantees that Eq. (3.29) has a unique solution which we can construct via the method of Picard iterates.

Theorem 3.18. *The unique solution to Eq. (3.29) is given by*

$$y(t) = f(t) + \sum_{n=1}^{\infty} \int_{0 \le \tau_1 \le \dots \le \tau_n \le t} dx(\tau_n)\, dx(\tau_{n-1})\, dx(\tau_{n-2}) \dots dx(\tau_1)\, f(\tau_1).$$

More generally if $0 \le s \le t \le T$, then the unique solution to

$$y(t) = \int_s^t dx(\tau)\, y(\tau) + f(t) \text{ for } s \le t \le T \tag{3.30}$$

is given by

$$y(t) = f(t) + \sum_{n=1}^{\infty} \int_{s \le \tau_1 \le \dots \le \tau_n \le t} dx(\tau_n)\, dx(\tau_{n-1})\, dx(\tau_{n-2}) \dots dx(\tau_1)\, f(\tau_1). \tag{3.31}$$

Proof. Let us first find the formula for $y(t)$. To this end, let

$$(Ay)(t) := \int_s^t dx(\tau)\, y(\tau).$$

Then Eq. (3.29) may be written as $y - Ay = f$ or equivalently as,

$$(I - A)\, y = f.$$

Thus the solution to this equation should be given by,

$$y = (I - A)^{-1} f = \sum_{n=0}^{\infty} A^n f. \tag{3.32}$$

But

$$\begin{aligned}
(A^n f)(t) &= \int_s^t dx(\tau_n) \left(A^{n-1} f\right)(\tau_n) = \int_s^t dx(\tau_n) \int_s^{\tau_n} dx(\tau_{n-1}) \left(A^{n-2} f\right)(\tau_{n-1}) \\
&\vdots \\
&= \int_s^t dx(\tau_n) \int_s^{\tau_n} dx(\tau_{n-1}) \int_s^{\tau_{n-1}} dx(\tau_{n-2}) \dots \int_s^{\tau_1} dx(\tau_1)\, f(\tau_1) \\
&= \int_{s \le \tau_1 \le \dots \le \tau_n \le t} dx(\tau_n)\, dx(\tau_{n-1})\, dx(\tau_{n-2}) \dots dx(\tau_1)\, f(\tau_1) \qquad (3.33)
\end{aligned}$$

and therefore, Eq. (3.31) now follows from Eq. (3.32) and (3.33).

For those not happy with this argument one may use Picard iterates instead. So we begin by setting $y_0(t) = f(t)$ and then define $y_n(t)$ inductively by,

$$\begin{aligned}
y_{n+1}(t) &= f(s) + \int_s^t F(y_n(\tau))\, d(x, f)(\tau) \\
&= f(s) + \int_s^t [dx(\tau)\, y_n(\tau) + df(\tau)] \\
&= \int_s^t dx(\tau)\, y_n(\tau) + f(t).
\end{aligned}$$

Therefore,

$$\begin{aligned}
y_1(t) &= \int_s^t dx(\tau)\, f(\tau) + f(t) \\
y_2(t) &= \int_s^t dx(\tau_2)\, y_1(\tau_2) + f(t) \\
&= \int_s^t dx(\tau_2) \left[\int_s^{\tau_2} dx(\tau_1)\, f(\tau_1) + f(\tau_2)\right] + f(t) \\
&= \int_{s \le \tau_1 \le \tau_2 \le t} dx(\tau_2)\, dx(\tau_1)\, f(\tau_1) + \int_s^t dx(\tau_2)\, f(\tau_2) + f(t)
\end{aligned}$$

and likewise,

$$y_3(t) = \int_{s\le\tau_1\le\tau_2\le\tau_3\le t} dx(\tau_3)\, dx(\tau_2)\, dx(\tau_1)\, f(\tau_1) + \int_{s\le\tau_1\le\tau_2\le t} dx(\tau_2)\, dx(\tau_1)\, f(\tau_1) + \int_s^t dx(\tau_2)\, f(\tau_2) + f(t).$$

So by induction it follows that

$$y_n(t) = \sum_{k=1}^{n} \int_{s\le\tau_1\le\dots\le\tau_k\le t} dx(\tau_k)\dots dx(\tau_1)\, f(\tau_1) + f(t).$$

Letting $n\to\infty$ making use of the fact that

$$\left\| \int_{s\le\tau_1\le\dots\le\tau_n\le t} dx(\tau_n)\dots dx(\tau_1) \right\| \le \int_{s\le\tau_1\le\dots\le\tau_n\le t} \|dx(\tau_n)\|\dots\|dx(\tau_1)\| = \frac{1}{n!}\left(\int_s^t \|dx\|\right)^n, \tag{3.34}$$

we find as before, that

$$y(t) = \lim_{n\to\infty} y_n(t) = f(t) + \sum_{k=1}^{\infty} \int_{s\le\tau_1\le\dots\le\tau_k\le t} dx(\tau_k)\dots dx(\tau_1)\, f(\tau_1).$$

■

Definition 3.19. *For $0\le s\le t\le T$, let $T_0(t,s) := I$,*

$$T_n^x(t,s) = \int_{s\le\tau_1\le\dots\le\tau_n\le t} dx(\tau_n)\, dx(\tau_{n-1})\, dx(\tau_{n-2})\dots dx(\tau_1), \text{ and} \tag{3.35}$$

$$T^x(t,s) = \sum_{n=0}^{\infty} T_n(t,s) = I + \sum_{n=1}^{\infty} \int_{s\le\tau_1\le\dots\le\tau_n\le t} dx(\tau_n)\, dx(\tau_{n-1})\, dx(\tau_{n-2})\dots dx(\tau_1). \tag{3.36}$$

Example 3.20. Suppose that $x(t) = tA$ where $A\in \mathrm{End}(V)$, then

$$\int_{s\le\tau_1\le\dots\le\tau_n\le t} dx(\tau_n)\, dx(\tau_{n-1})\, dx(\tau_{n-2})\dots dx(\tau_1) = A^n \int_{s\le\tau_1\le\dots\le\tau_n\le t} d\tau_n d\tau_{n-1} d\tau_{n-2}\dots d\tau_1 = \frac{(t-s)^n}{n!} A^n$$

and therefore we may conclude in this case that $T^x(t,s) = e^{(t-s)A}$ where

$$e^{tA} := \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n.$$

Theorem 3.21 (Duhamel's principle I). *As a function of $t\in[s,T]$ or $s\in[0,t]$, $T(t,s)$ is of bounded variation and $T(t,s) := T^x(t,s)$ satisfies the ordinary differential equations,*

$$T(dt,s) = dx(t)\, T(t,s) \text{ with } T(s,s) = I \quad (\text{in } t\ge s), \tag{3.37}$$

and

$$T(t,s) = -T(t,s)\, dx(s) \text{ with } T(t,t) = I \ (\text{in } 0\le s\le t). \tag{3.38}$$

Moreover, T, obeys the semi-group property[1]*,*

$$T(t,s)\, T(s,u) = T(t,u) \text{ for all } 0\le u\le s\le t\le T,$$

and the solution to Eq. (3.30) is given by

$$y(t) = f(t) - \int_s^t T(t,d\tau)\, f(\tau). \tag{3.39}$$

In particular when $f(t) = y_0$ is a constant we have,

$$y(t) = y_0 - T(t,\tau)\, y_0|_{\tau=s}^{\tau=t} = T(t,s)\, y_0. \tag{3.40}$$

Proof. 1. One may directly conclude that $T(t,s)$ solves Eq. (3.37) by applying Theorem 3.18 with $y(t)$ and $f(t)$ now taking values in $\mathrm{End}(V)$ with $f(t)\equiv I$. Then Theorem 3.18 asserts the solution to $dy(t) = dx(t)\, y(t)$ with $y(0) = I$ is given by $T^x(t,s)$ with $T^x(t,s)$ as in Eq. (3.36). Alternatively it is possible to use the definition of $T^x(t,s)$ in Eq. (3.36) to give a direct proof the Eq. (3.37) holds. We will carry out this style of proof for Eq. (3.38) and leave the similar proof of Eq. (3.37) to the reader if they so desire to do it.

2. **Proof of the semi-group property.** Simply observe that both $t\to T(t,s)\, T(s,u)$ and $t\to T(t,u)$ solve the same differential equation, namely,

$$dy(t) = dx(t)\, y(t) \text{ with } y(s) = T(s,u)\in \mathrm{End}(V),$$

hence by our uniqueness results we know that $T(t,s)\, T(s,u) = T(t,u)$.

3. **Proof of Eq. (3.38).** Let $T_n(t,s) := T_n^x(t,s)$ and observe that

[1] This is a key algebraic idenitity that we must demand in the rough path theory to come later.

$$T_n(t,s) = \int_s^t T_{n-1}(t,\sigma)\, dx(\sigma) \ \text{ for } n \geq 1. \tag{3.41}$$

Thus if let

$$T^{(N)}(t,s) := \sum_{n=0}^{N} T_n(t,s) = I + \sum_{n=1}^{N} T_n(t,s),$$

it follows that

$$\begin{aligned} T^{(N)}(t,s) &= I + \sum_{n=1}^{N} \int_s^t T_{n-1}(t,\sigma)\, dx(\sigma) \\ &= I + \int_s^t \sum_{n=0}^{N-1} T_n(t,\sigma)\, dx(\sigma) = I + \int_s^t T^{(N-1)}(t,\sigma)\, dx(\sigma). \end{aligned} \tag{3.42}$$

We already now that $T^{(N)}(t,s) \to T(t,s)$ uniformly in (t,s) which also follows from the estimate in Eq. (3.34) as well. Passing to the limit in Eq. (3.42) as $N \to \infty$ then implies,

$$T(t,s) = I + \int_s^t T(t,\sigma)\, dx(\sigma)$$

which is the integrated form of Eq. (3.38) owing to the fundamental theorem of calculus which asserts that

$$d_s \int_s^t T(t,\sigma)\, dx(\sigma) = -T(t,s)\, dx(s).$$

4. **Proof of Eq. (3.39).** From Eq. (3.31) and Eq. (3.41) which reads in differential form as, $T_n(t,d\sigma) = -T_{n-1}(t,\sigma)\, dx(\sigma)$, we have,

$$\begin{aligned} y(t) &= f(t) - \sum_{n=1}^{\infty} \int_s^t T_n(t,d\sigma) f(\sigma) \\ &= f(t) - \sum_{n=1}^{\infty} \int_s^t T_{n-1}(t,\sigma)\, dx(\sigma) f(\sigma) \\ &= f(t) - \int_s^t \sum_{n=1}^{\infty} T_{n-1}(t,\sigma)\, dx(\sigma) f(\sigma) \\ &= f(t) - \int_s^t T(t,d\sigma) f(\sigma). \end{aligned}$$

■

Corollary 3.22 (Duhamel's principle II). *Equation (3.39) may also be expressed as,*

$$y(t) = T^x(t,s) f(s) + \int_s^t T^x(t,\tau)\, df(\tau) \tag{3.43}$$

which is one of the standard forms of Duhamel's principle. In words it says,

$$\begin{aligned} y(t) &= \begin{pmatrix} \text{solution to the homogeneous eq.} \\ dy(t) = dx(t)\, y(t) \ \text{ with } y(s) = f(s) \end{pmatrix} \\ &\quad + \int_s^t \begin{pmatrix} \text{solution to the homogeneous eq.} \\ dy(t) = dx(t)\, y(t) \ \text{ with } y(\tau) = df(\tau) \end{pmatrix}. \end{aligned}$$

Proof. This follows from Eq. (3.39) by integration by parts (you should modify Exercise 3.5 below as necessary);

$$\begin{aligned} y(t) &= f(t) - T^x(t,\tau) f(\tau)|_{\tau=s}^{\tau=t} + \int_s^t T^x(t,\tau)\, df(\tau) \\ &= T^x(t,s) f(s) + \int_s^t T^x(t,\tau)\, df(\tau). \end{aligned} \tag{3.44}$$

■

Exercise 3.5 (Product Rule). Suppose that V is a Banach space and $x : [0,T] \to \mathrm{End}(V)$ and $y : [0,T] \to \mathrm{End}(V)$ are continuous finite $1-$variation paths. Show for all $0 \leq s < t \leq T$ that,

$$x(t)\, y(t) - x(s)\, y(s) = \int_s^t dx(\tau)\, y(\tau) + \int_s^t x(\tau)\, dy(\tau). \tag{3.45}$$

Alternatively, one may interpret this an an integration by parts formula;

$$\int_s^t x(\tau)\, dy(\tau) = x(\tau)\, y(\tau)|_{\tau=s}^{\tau=t} - \int_s^t dx(\tau)\, y(\tau)$$

Solution to Exercise (3.5). For $\Pi \in \mathcal{P}(s,t)$ we have,

$$\begin{aligned} x(t)\, y(t) - x(s)\, y(s) &= \sum_{\tau \in \Pi} \Delta_\tau (x(\cdot)\, y(\cdot)) \\ &= \sum_{\tau \in \Pi} \left[(x(\tau_-) + \Delta_\tau x)(y(\tau_-) + \Delta_\tau y) - x(\tau_-)\, y(\tau_-) \right] \\ &= \sum_{\tau \in \Pi} \left[x(\tau_-)\, \Delta_\tau y + (\Delta_\tau x)\, y(\tau_-) + (\Delta_\tau x)\, \Delta_\tau y \right]. \end{aligned} \tag{3.46}$$

The last term is easy to estimate as,

$$\left\|\sum_{\tau\in\Pi}(\Delta_\tau x)\,\Delta_\tau y\right\| \le \sum_{\tau\in\Pi}\|\Delta_\tau x\|\,\|\Delta_\tau y\| \le \max_{\tau\in\Pi}\|\Delta_\tau x\|\cdot\sum_{\tau\in\Pi}\|\Delta_\tau y\|$$
$$\le \max_{\tau\in\Pi}\|\Delta_\tau x\|\cdot V_1(y)\to 0 \text{ as } |\Pi|\to 0.$$

So passing to the limit as $|\Pi|\to 0$ in Eq. (3.46) gives Eq. (3.45).

Exercise 3.6 (Inverses). Let V be a Banach space and $x:[0,T]\to \mathrm{End}(V)$ be a continuous finite 1 – variation paths. Further suppose that $S(t,s)\in \mathrm{End}(V)$ is the unique solution to,

$$S(dt,s) = -S(t,s)\,dx(t) \text{ with } S(s,s)=I\in\mathrm{End}(V).$$

Show

$$S(t,s)\,T^x(t,s) = I = T^x(t,s)\,S(t,s) \text{ for all } 0\le s\le t\le T,$$

that is to say $T^x(t,s)$ is invertible and $T^x(t,s)^{-1}$ may be described as the unique solution to the ODE,

$$T^x(dt,s)^{-1} = -T^x(dt,s)^{-1}\,dx(t) \text{ with } T^x(s,s)^{-1}=I. \tag{3.47}$$

Solution to Exercise (3.6). Using the product rule we find,

$$d_t\left[S(t,s)\,T^x(t,s)\right] = -S(t,s)\,dx(t)\,T^x(t,s)+S(t,s)\,dx(t)\,T^x(t,s)=0.$$

Since $S(s,s)\,T^x(s,s)=I$, it follows that $S(t,s)\,T^x(t,s)=I$ for all $0\le s\le t\le T$.

For opposite product, let $g(t):=T^x(t,s)\,S(t,s)\in\mathrm{End}(V)$ so that

$$d_t[g(t)] = d_t\left[T^x(t,s)\,S(t,s)\right] = -T^x(t,s)\,S(t,s)\,dx(t)+dx(t)\,T^x(t,s)\,S(t,s)$$
$$= dx(t)\,g(t)-g(t)\,dx(t) \text{ with } g(s)=I.$$

Observe that $g(t)\equiv I$ solves this ODE and therefore by uniqueness of solutions to such linear ODE we may conclude that $g(t)$ must be equal to I, i.e. $T^x(t,s)\,S(t,s)=I$.

As usual we say that $A,B\in\mathrm{End}(V)$ *commute* if

$$0=[A,B]:=AB-BA. \tag{3.48}$$

Exercise 3.7 (Commute). Suppose that V is a Banach space and $x:[0,T]\to\mathrm{End}(V)$ is a continuous finite 1 – variation paths and $f:[0,T]\to\mathrm{End}(V)$ is continuous. If $A\in\mathrm{End}(V)$ commutes with $\{x(t),f(t):0\le t\le T\}$, then A commutes with $\int_s^t f(\tau)\,dx(\tau)$ for all $0\le s\le t\le T$. Also show that $[A,T^x(s,t)]=0$ for all $0\le s\le t\le T$.

Exercise 3.8 (Abelian Case). Suppose that $[x(s),x(t)]=0$ for all $0\le s,t\le T$, show

$$T^x(t,s)=e^{(x(t)-x(s))}. \tag{3.49}$$

Solution to Exercise (3.8). By replacing $x(t)$ by $x(t)-x(s)$ if necessary, we may assume that $x(s)=0$. Then by the product rule and the assumed commutativity,

$$d\frac{x^n(t)}{n!} = dx(t)\,\frac{x(t)^{n-1}}{(n-1)!}$$

or in integral form,

$$\frac{x^n(t)}{n!}=\int_s^t dx(\tau)\,\frac{x(\tau)^{n-1}}{(n-1)!}$$

which shows that $\frac{1}{n!}x^n(t)$ satisfies the same recursion relations as $T_n^x(t,s)$. Thus we may conclude that $T_n^x(t,s)=(x(t)-x(s))^n/n!$ and thus,

$$T(t,s)=\sum_{n=0}^{\infty}\frac{1}{n!}(x(t)-x(s))^n = e^{(x(t)-x(s))}$$

as desired.

Exercise 3.9 (Abelian Factorization Property). Suppose that V is a Banach space and $x:[0,T]\to\mathrm{End}(V)$ and $y:[0,T]\to\mathrm{End}(V)$ are continuous finite 1 – variation paths such that $[x(s),y(t)]=0$ for all $0\le s,t\le T$, then

$$T^{x+y}(s,t)=T^x(s,t)\,T^y(s,t) \text{ for all } 0\le s\le t\le T. \tag{3.50}$$

Hint: show both sides satisfy the same ordinary differential equations – see the next problem.

Exercise 3.10 (General Factorization Property). Suppose that V is a Banach space and $x:[0,T]\to\mathrm{End}(V)$ and $y:[0,T]\to\mathrm{End}(V)$ are continuous finite 1 – variation paths. Show

$$T^{x+y}(s,t)=T^x(s,t)\,T^z(s,t),$$

where

$$z(t):=\int_s^t T^x(s,\tau)^{-1}\,dy(\tau)\,T^x(s,\tau)$$

Hint: see the hint for Exercise 3.9.

Solution to Exercise (3.10). Let $g(t):=T^x(s,t)^{-1}\,T^{x+y}(s,t)$. Then making use of Exercise 3.6 we have,

$$\begin{aligned}dg(t) &= T^x(s,t)^{-1}(dx(t)+dy(t))T^{x+y}(s,t) - T^x(s,t)^{-1}dx(t)T^{x+y}(s,t)\\ &= T^x(s,t)^{-1}dy(t)T^{x+y}(s,t)\\ &= \left(T^x(s,t)^{-1}dy(t)T^x(s,t)\right)T^x(s,t)^{-1}T^{x+y}(s,t)\\ &= dz(t)g(t) \text{ with } g(s)=I.\end{aligned}$$

Remark 3.23. If $g(t) \in \operatorname{End}(V)$ is a C^1 − path such that $g(t)^{-1}$ is invertible for all t, then $t \to g(t)^{-1}$ is invertible and

$$\frac{d}{dt}g(t)^{-1} = -g(t)^{-1}\dot{g}(t)g(t)^{-1}.$$

Exercise 3.11. Suppose that $g(t) \in \operatorname{Aut}(V)$ is a continuous finite variation path. Show $g(t)^{-1} \in \operatorname{Aut}(V)$ is again a continuous path with finite variation and that

$$dg(t)^{-1} = -g(t)^{-1}dg(t)g(t)^{-1}. \tag{3.51}$$

Hint: recall that the invertible elements, $\operatorname{Aut}(V) \subset \operatorname{End}(V)$, is an open set and that $\operatorname{Aut}(V) \ni g \to g^{-1} \in \operatorname{Aut}(V)$ is a smooth map.

Solution to Exercise (3.11). Let $V(t) := g(t)^{-1}$ which is again a finite variation path by the fundamental theorem of calculus and the fact that $\operatorname{Aut}(V) \ni g \to g^{-1} \in \operatorname{Aut}(V)$ is a smooth map. Moreover we know that $V(t)g(t) = I$ for all t and therefore by the product rule $(dV)g + Vdg = dI = 0$. Making use of the substitution formula we then find,

$$V(t) = V(0) + \int_0^t dV(\tau) = \int_0^t dV(\tau)g(\tau)g(\tau)^{-1} = -\int_0^t V(\tau)dg(\tau)g(\tau)^{-1}.$$

Replacing $V(t)$ by $g(t)^{-1}$ in this equation then shows,

$$g(t)^{-1} - g(0)^{-1} = -\int_0^t g(\tau)dg(\tau)g(\tau)^{-1}$$

which is the integrated form of Eq. (3.51).

Exercise 3.12. Suppose now that B is a Banach algebra and $x(t) \in B$ is a continuous finite variation path. Let

$$X(s,t) := X^x(s,t) := 1 + \sum_{n=1}^{\infty} X_n^x(s,t),$$

where

$$X_n^x(s,t) := \int_{s\le\tau_1\le\dots\le\tau_n\le t} dx(\tau_1)\dots dx(\tau_n)$$

Show $t \to X(s,t)$ is the unique solution to the ODE,

$$X(s,dt) = X(s,t)dx(t) \text{ with } X(s,s) = 1$$

and that

$$X(s,t)X(t,u) = X(s,u) \text{ for all } 0 \le s \le t \le u \le T.$$

Solution to Exercise (3.12). This can be deduced from what we have already done. In order to do this, let $y(t) := R_{x(t)} \in \operatorname{End}(B)$ so that for $a \in B$,

$$\begin{aligned}T^y(t,s)a &= a + \sum_{n=1}^{\infty}\int_{s\le\tau_1\le\dots\le\tau_n\le t} dy(\tau_n)\dots dy(\tau_1)a\\ &= a + \sum_{n=1}^{\infty}\int_{s\le\tau_1\le\dots\le\tau_n\le t} adx(\tau_1)\dots dx(\tau_n)\\ &= aX(s,t).\end{aligned}$$

Therefore, taking $a = 1$, we find,

$$X(s,dt) = T^y(dt,s)1 = dy(t)T^y(t,s)1 = dy(t)X(s,t) = X(s,t)dx(t)$$

with $X(s,s) = T(s,s)1 = 1$. Moreover we also have,

$$X(s,t)X(t,u) = T^y(u,t)X(s,t) = T^y(u,t)T^y(t,s)1 = T^y(u,s)1 = X(s,u).$$

Alternatively: one can just check all the statements as we did for $T(t,s)$. The main point is that if $g(t)$ solves $dg(t) = g(t)dx(t)$, then $ag(t)$ also solves the same equation.

Remark 3.24. Let $\lambda \in \mathbb{R}$ or $\mathbb{C}$ as the case may be and define,

$$X^\lambda(s,t) := X^{\lambda x}(s,t) \text{ and } X_n^\lambda(s,t) := X_n^{\lambda x}(s,t) = \lambda^n X_n(s,t). \tag{3.52}$$

Then the identity in Eq. (3.52) becomes,

$$\begin{aligned}\sum_{n=0}^{\infty}\lambda^n X_n(s,u) &= X^\lambda(s,u) = X^\lambda(s,t)X^\lambda(t,u)\\ &= \sum_{k,l=0}^{\infty}\lambda^k\lambda^l X_k(s,t)X_l(t,u)\\ &= \sum_{n=0}^{\infty}\lambda^n\sum_{k+l=n}X_k(s,t)X_l(t,u)\end{aligned}$$

from which we conclude,

$$X_n(s,u) = \sum_{k=0}^{n} X_k(s,t)\, X_{n-k}(t,u) \ \text{ for } n = 0, 1, 2, \ldots \tag{3.53}$$

Terry. Lyons refers the identities in Eq. (3.53) as Chen's identities. These identities may be also be deduced directly by looking at the multiple integral expressions defining $X_k(s,t)$.

References

1. A. M. Davie, *Differential equations driven by rough paths: an approach via discrete approximation*, Appl. Math. Res. Express. AMRX (2007), no. 2, Art. ID abm009, 40. MR MR2387018
2. Peter Friz and Nicolas Victoir, *Multidimensional stochastic processes as rough paths: Theory and applications*, Springer ??, 2008.
3. Terry Lyons and Zhongmin Qian, *System Control and Rough Paths*, Oxford University Press, 2002, Oxford Mathematical Monographs.
4. ______, *System control and rough paths*, Oxford Mathematical Monographs, Oxford University Press, Oxford, 2002, Oxford Science Publications. MR MR2036784 (2005f:93001)
5. Terry Lyons and Nicolas Victoir, *An extension theorem to rough paths*, Ann. Inst. H. Poincaré Anal. Non Linéaire **24** (2007), no. 5, 835–847. MR MR2348055 (2008h:60229)
6. Terry J. Lyons, Michael Caruana, and Thierry Lévy, *Differential equations driven by rough paths*, Lecture Notes in Mathematics, vol. 1908, Springer, Berlin, 2007, Lectures from the 34th Summer School on Probability Theory held in Saint-Flour, July 6–24, 2004, With an introduction concerning the Summer School by Jean Picard. MR MR2314753
7. Peter Mrters, Yuval Peres with appendix by Schramm, and Werner, *Brownian motion*, http://people.bath.ac.uk/maspm/book.pdf, 2006.