Rev. Mat. Iberoam. **33** (2017), no. 3, 885–950
DOI 10.4171/RMI/959

Controlled rough paths on manifolds I

Bruce K. Driver and Jeremy S. Semko

Abstract. In this paper, we build the foundation for a theory of controlled rough paths on manifolds. A number of natural candidates for the definition of manifold valued controlled rough paths are developed and shown to be equivalent. The theory of controlled rough one-forms along such a controlled path and their resulting integrals are then defined. This general integration theory does require the introduction of an additional geometric structure on the manifold which we refer to as a "parallelism". A choice of parallelism allows us to compare nearby tangent spaces on the manifold which is necessary to fully discuss controlled rough one-forms. The transformation properties of the theory under change of parallelisms is explored. Although the integration of a general controlled one-form along a rough path does depend on the choice of parallelism, we show for a special class of controlled one-forms – those which are the restriction of smooth one-forms on the manifold – the resulting path integral is in fact independent of any choice of parallelism. We present a theory of push-forwards and show how it is compatible with our integration theory. Lastly, we give a number of characterizations for solving a rough differential equation when the solution is interpreted as a controlled rough path on a manifold and then show such solutions exist and are unique.

1. Introduction

In a series of papers [24], [25], [26], Terry Lyons introduced and developed the far reaching theory of rough path analysis. This theory allows one to solve (deterministically) differential equations driven by rough signals at the expense of "enhancing" the rough signal with some additional information. Lyons' theory has found numerous applications to stochastic calculus and stochastic differential equations, for example see [4], [5], [6], [8], and the references therein. For some more recent applications, see [1], [21], [20], [9] , and [2].

The rough path theory mentioned above has been almost exclusively developed in the context of state spaces being either finite or infinite dimensional Banach

Mathematics Subject Classification (2010): Primary 58J65; Secondary 60H30.
Keywords: Dynamical systems, rough paths.

spaces with the two exceptions of [7] and [3]. In [7], a version of manifold valued rough paths is developed in the context of "currents". Although the definition in [7] is global and intrinsic, it does rely on the non-trivial restriction that the underlying manifold is a "Lip-γ manifold" in order to obtain uniform estimates. In [3], the authors avoid the Lip-γ restriction at the expense of considering embedded submanifolds and introducing the notion of weakly geometric rough paths "constrained" to lie in the submanifold. It is eventually shown in [3] that the theory is independent of a choice of embedding. This embedded theory follows very closely Terry Lyons' original development of the theory.

The purpose of this paper is to define and develop a third interpretation of rough paths on manifolds based on Gubinelli's [16] notions of "controlled" rough paths. As Gubinelli's perspective has proved extremely useful in the flat case (most notably see Hairer [17]), it is expected such a theory of controlled rough paths on manifolds can give new insights as well as applications to the existing literature. We now will present a brief summary of the results contained in this paper. The geometrical notation will follow closely that found in [11]. The reader may also refer to well-known works such as [10], [23], [22], and [27] for basic but essential geometric background.

1.1. Summary of results

Let M^d be a d-dimensional manifold, $\mathbf{X}_{s,t} := 1 + x_{s,t} + \mathbb{X}_{s,t}$ be a weak-geometric rough path in $W := \mathbb{R}^k$ with $1 \le p < 3$, see Definition 2.2. In order to define controlled rough paths on M it is necessary to make sense of the "increments" of paths in M and TM. This leads us to add two extra structures to M, namely a "logarithm" and a "parallelism". A **logarithm** is a smooth assignment, $\psi(m,n) \in T_mM$, for (m,n) near the diagonal of M which, in local coordinates, has the form $\psi(m,n) = n - m + O(|n-m|^2)$. Similarly a **parallelism** is an assignment of a linear transformation, $U(n,m) : T_mM \to T_nM$ for each $(n,m) \in M \times M$ near the diagonal which in local coordinates has the form $U(n,m) = I + O(|n-m|)$. See Definition 2.15 and Definition 2.16 for the precise definitions of a logarithm and a parallelism respectively, and Theorem 2.24 which asserts the description given above matches these definitions (when $M = \mathbb{R}^d$, one identifies all tangent spaces in which case one typically takes $U(m,n) = I$ and $\psi(m,n) = n - m$). The pair $\mathcal{G} := (\psi, U)$ is called a **gauge.**

A *rough path controlled by* $\mathbf{X}$ on M (see Definition 2.35) is a pair of continuous functions $y\colon [0,T] \to M$, and $y^\dagger\colon [0,T] \to L(W, TM)$ such that (somewhat imprecisely speaking), for all $0 \le s \le t \le T$,

1) $y_s^\dagger : W \to T_{y_s}M$,
2) $\psi(y_s, y_t) = y_s^\dagger\, x_{s,t} + O(|x_{s,t}|^2)$, and
3) $U(y_s, y_t)\, y_t^\dagger - y_s^\dagger = O(|x_{s,t}|)$,

where ψ is a logarithm on M and U is a parallelism on M. Alternatively one can define controlled rough paths locally via a chart ϕ by requiring (see Definition 2.40)

$$\phi(y_t) - \phi(y_s) - d\phi \circ y_s^\dagger x_{s,t} = O(|x_{s,t}|^2) \quad \text{and} \quad d\phi \circ y_t^\dagger - d\phi \circ y_s^\dagger = O(|x_{s,t}|).$$

It is shown in Theorem 2.45 that these two notions of controlled rough paths agree. Moreover, these manifold-valued rough paths may also be characterized as pairs $\mathbf{y} = (y, y^\dagger)$ whose "push-forwards" under smooth real-valued functions are controlled rough paths on $\mathbb{R}$ (See Theorem 2.57).

Two natural examples of manifold valued controlled rough paths are as follows. 1) If M^d is an embedded submanifold (see subsection 2.6) and the path $x_s \in W$ happens to lie in M (i.e., $x_s \in M$ for all s in $[0,T]$), then $(x_s, P(x_s))$ is an M-valued rough path controlled by $\mathbf{X}$ where $P(m)$ is orthogonal projection onto $T_m M$ (see Example 2.55). 2) If $f\colon W \to M^d \subseteq \mathbb{R}^{\bar{k}}$ is smooth, then $(f(x_s), f'(x_s))$ is a rough path controlled by $\mathbf{X}$ (see Example 2.56).

Now let $\mathcal{G} = (\psi, U)$ be a gauge, V be a Banach space, and $\mathbf{y} = (y, y^\dagger)$ be an M-valued controlled rough path as above. A pair of continuous functions $\alpha\colon [0,T] \to L(TM, V)$ and $\alpha^\dagger\colon [0,T] \to L(W \otimes TM, V)$ is a *U-controlled (rough) one-form along y* with values in a Banach space V provided (see Definition 3.1 for details);

1. $\alpha_s : T_{y_s} M \to V$ for all s,
2. $\alpha_s^\dagger : W \otimes T_{y_s} M \to V$ for all s,
3. $\alpha_t \circ U(y_t, y_s) - \alpha_s - \alpha_s^\dagger(x_{s,t} \otimes (\cdot)) = O(|x_{s,t}|^2)$, and
4. $\alpha_t^\dagger \circ (I \otimes U(y_t, y_s)) - \alpha_s^\dagger = O(|x_{s,t}|)$.

To abbreviate notation we write $\boldsymbol{\alpha}_s = \left(\alpha_s, \alpha_s^\dagger\right)$. As an example, if α is a smooth one-form on M with values in V (denoted $\Omega^1(M,V)$) and U is a parallelism, it is shown in Proposition 4.2 how to construct $\alpha_s^{\dagger U}$ so that $\boldsymbol{\alpha}_s^U = (\alpha_s := \alpha|_{T_{y_s}M}, \alpha_s^{\dagger U})$ is the **associated U-controlled (rough) one-form** along y.

Theorem 3.21 below constructs the integral, $\int \langle \boldsymbol{\alpha}, d\mathbf{y}^{\mathcal{G}} \rangle$, of $\boldsymbol{\alpha}$ along $\mathbf{y} = (y, y^\dagger)$ such that $\int_s^t \langle \boldsymbol{\alpha}, d\mathbf{y}^{\mathcal{G}} \rangle$ is well approximated by $\langle \boldsymbol{\alpha}_s, \mathbf{y}_{s,t}^{\mathcal{G}} \rangle$. Here $\mathbf{y}_{s,t}^{\mathcal{G}}$ is a "second" order increment of y whose first order contribution is $\psi(y_s, y_t)$ and higher order contributions are determined by the gauge, $\mathcal{G}$, in combination with $y^\dagger$ and $\mathbb{X}$ (see Definition 3.20 for details). This integral is a standard flat V-valued controlled rough path along $\mathbf{X}$ which, as the notation suggests, a priori depends on both the choice of logarithm ψ and parallelism U in the gauge, $\mathcal{G} = (\psi, U)$. However, it is shown in Corollary 3.30 that the resulting integral is in fact independent of the choice of logarithm and thus, in the end, only depends on the choice of parallelism, U. (The logarithm appears in both the first and second order contributions of $\mathbf{y}_{s,t}^{\mathcal{G}}$; ultimately there is a cancellation among these two contributions in such a way that the dependence of the integral on ψ cancels out in the limit; see Subsection 6.1 in the Appendix for intuition as to why this might be expected.)

In Theorem 3.32, it is shown that there are "natural" transformations relating all of the above structures under change of parallelism, $U \to \tilde{U}$, in such a way that the integral, $\int \langle \boldsymbol{\alpha}, d\mathbf{y}^{\mathcal{G}} \rangle$, is preserved. Although the integration of a general controlled one-form along a rough path does depend on the choice of parallelism, this dependence drops out in the special case that the controlled one-form is "induced" from a globally defined smooth one-form on M. In more detail if U is a parallelism and $\alpha \in \Omega^1(M,V)$, we let $\boldsymbol{\alpha}_s^U = (\alpha_s := \alpha|_{T_{y_s}M}, \alpha_s^{\dagger U} := \nabla^U_{y_s^\dagger(\cdot)} \alpha)$ be the associated U-controlled (rough) one-form along y. It is then shown in Theorem 4.3 that the

resulting integral, $\int \langle \boldsymbol{\alpha}^U, d\mathbf{y}^{\mathcal{G}} \rangle$, is in fact independent of both the parallelism, U, and the logarithm, ψ, used in the construction. A gauge independent formula for this integral is then given in Corollary 4.7. Because of these results, we are justified in writing $\int \alpha(d\mathbf{y})$ for the path integral of α against $\mathbf{y}$ whenever $\alpha \in \Omega^1(M, V)$ and $\mathbf{y}$ is a controlled rough path on M. In Theorem 4.9 (also see Proposition 3.6), we show that, whether we are using a smooth one-form $\alpha \in \Omega^1(M, V)$ or rough one-form $\boldsymbol{\alpha}$ along $\mathbf{y}$, the integrals $\int \alpha(d\mathbf{y})$ and $\int \langle \boldsymbol{\alpha}, d\mathbf{y}^{\mathcal{G}} \rangle$ satisfy a basic but useful associativity property.

It is shown in Theorem 4.15 that if $\alpha \in \Omega^1(\tilde{M}, V)$ is a smooth one-form on M and $f\colon M \to \tilde{M}$ is a smooth map between two manifolds, then

$$\int (f^*\alpha)(d\mathbf{y}) = \int \alpha(d(f_*\mathbf{y})),$$

where $f_*\mathbf{y} := \left(f \circ y, f_* \circ y_s^\dagger\right)$ is the "push-forward" of $\mathbf{y}$ by f (see Definition 4.11) and $f^*\alpha \in \Omega^1(M, V)$ is the pull-back of α.

In Section 5.1, we discuss the notion of a controlled rough path $\mathbf{y} = (y, y^\dagger)$ solving the rough differential equation (RDE)

$$d\mathbf{y}_t = F_{d\mathbf{X}}(y_t) \quad \text{with} \quad y_0 = \bar{y}_0$$

when $F\colon W \to \Gamma(TM)$ where $\Gamma(TM)$ denotes the smooth sections on TM. Essentially $\mathbf{y}$ will solve such an equation if the path y, when pushed forward by any smooth function f, has the correct "Taylor expansion" and $y^\dagger$ is the correct derivative, i.e., $y_s^\dagger = F_{(\cdot)}(y_s)$ (see Definition 5.2). Theorem 5.3 then compiles a list of alternative characterizations for solving an RDE both by approximating solutions and by relating them to familiar flat space rough integrals. Next, the existence and uniqueness of solutions are proved in Theorem 5.4 and Theorem 5.5. Lastly, in Theorem 5.9, we record what it means to solve an RDE when one takes the gauge perspective.

In a sequel to this paper, we will develop notions of parallel translation along a controlled rough path along with rough version of Cartan's rolling and unrolling maps in order to characterize all controlled rough paths on M.

Acknowledgments. The authors are very thankful to the two anonymous referees for their careful readings and many corrections and suggestions to this paper. Their input has substantially reduced the number of typos and (more importantly) has greatly improved the clarity of the paper.

2. Definitions of controlled rough paths with examples

2.1. Review of Euclidean space rough paths

The presentation here will be brief. For a more thorough development, the reader can refer to many sources, for example [14] or [15].

Throughout this paper, we denote $W = \mathbb{R}^k$. Let $1 \le p < 3$ and let

$$\Delta_{[S,T]} = \{(s,t) : S \le s \le t \le T\}. \tag{2.1}$$

Definition 2.1. A **control** ω is a continuous function $\omega\colon \Delta_{[0,T]} \to \mathbb{R}_+$ which is superadditive[1] and such that $\omega(s,s) = 0$ for all $s \in [0,T]$.

Definition 2.2. Let $\mathbf{X} = (x, \mathbb{X})$ where

$$x : [0,T] \to W \quad \text{and} \quad \mathbb{X} : \Delta_{[0,T]} \to W \otimes W$$

and are continuous. Then $\mathbf{X}$ is a p**-rough path with control** ω if

1. The Chen identity holds:

$$\mathbb{X}_{s,u} = \mathbb{X}_{s,t} + \mathbb{X}_{t,u} + x_{s,t} \otimes x_{t,u} \tag{2.2}$$

   for all $0 \le s \le t \le u \le T$, where $x_{s,t} := x_t - x_s$.

2. For all $0 \le s \le t \le T$,

$$|x_{s,t}| \le \omega(s,t)^{1/p} \quad \text{and} \quad |\mathbb{X}_{s,t}| \le \omega(s,t)^{2/p}. \tag{2.3}$$

Further, we say that $\mathbf{X}$ is **weak-geometric** if the symmetric part of $\mathbb{X}_{s,t}$ ($\operatorname{sym}(\mathbb{X}_{s,t})$) satisfies the relation

$$\operatorname{sym}(\mathbb{X}_{s,t}) = \frac{1}{2}\, x_{s,t} \otimes x_{s,t}.$$

Notation 2.3. Let $F_{s,t}$ and $G_{s,t}$ be a pair of functions into a normed space. When it is not important to keep careful track of constants we will often write $F_{s,t} \underset{i}{\approx} G_{s,t}$ (for any $i \in \mathbb{N}$) to indicate that there exists $C < \infty$ and $\delta > 0$ such that

$$|F_{s,t} - G_{s,t}| \le C\,\omega(s,t)^{i/p} \quad \text{for all } 0 \le s \le t \le T \text{ with } |t-s| \le \delta.$$

In this paper, V,$\tilde{V}$, and $\hat{V}$ will denote Banach spaces, and $L(V,\tilde{V})$ will denote the bounded linear transformations from V to $\tilde{V}$.

Example 2.4. If $x(t) \in C^\infty([0,T],V)$ is a smooth curve in V and

$$\mathbb{X}_{s,t} = \int_{s \le u \le v \le t} dx_u \otimes dx_v = \int_s^t x_{s,v} \otimes dx_v, \tag{2.4}$$

then $\mathbf{X} = (x, \mathbb{X})$ is a weak-geometric rough path controlled by $\omega(s,t) = |t-s|$. In this example we could take even take $p = 1$.

Definition 2.5. Let $\mathbf{X}$ be a p-rough path on $W \oplus W^{\otimes 2}$ with control ω. The continuous pair $\mathbf{y} := (y, y^\dagger) \in C([a,b],V) \times C([a,b], L(W,V))$ is a V-valued **rough path controlled by** $\mathbf{X}$ (denoted $\mathbf{y} \in CRP_{\mathbf{X}}([a,b],V)$) if there exists a C such that

1. $|y_t - y_s - y_s^\dagger\, x_{s,t}| \le C\,\omega(s,t)^{2/p}$, and
2. $|y_t^\dagger - y_s^\dagger| \le C\,\omega(s,t)^{1/p}$ for all $s \le t$ in $[0,T]$.

We denote $CRP_{\mathbf{X}}(V) := CRP_{\mathbf{X}}([0,T],V)$ for some fixed $T < \infty$.

[1] To say ω is superadditive means $\omega(s,t) + \omega(t,u) \le \omega(s,u)$ for all $0 \le s \le t \le u \le T$.

The approximations in Definition 2.5 are statements which only need to hold locally because of the following (easy) sewing lemma.

Lemma 2.6 (Sewing lemma). *Let*

$$\mathbf{y} := (y, y^\dagger) \in C([0,T], V) \times C([0,T], L(W,V))$$

and let $0 = t_0 < t_1 < \ldots < t_l = T$ *be a partition of* $[0,T]$ *such that* $\mathbf{y}|_{[t_i,t_{i+i}]}$ *is a rough path controlled by* $\mathbf{X}|_{[t_i,t_{i+1}]} := (x|_{[t_i,t_{i+1}]}, \mathbb{X}|_{\Delta_{[t_i,t_{i+1}]}})$ *for all* $0 \le i \le l-1$. *Then* $\mathbf{y}$ *is a rough path controlled by* $\mathbf{X}$.

Proof. Let C_i with $0 \le i \le l-1$ be such that

$$\left|y_t - y_s - y_s^\dagger\, x_{s,t}\right| \le C_i\, \omega(s,t)^{2/p} \quad \text{and} \quad |y_t^\dagger - y_s^\dagger| \le C_i\, \omega(s,t)^{1/p}$$

whenever $(s,t) \in \Delta_{[t_i,t_{i+1}]}$. Let $\tilde{C} := \sum_{i=0}^{l-1} C_i$. Then by a telescoping series argument and the fact that ω is increasing (because it is superadditive), it is clear that

$$|y_t^\dagger - y_s^\dagger| \le \tilde{C}\, \omega(s,t)^{1/p} \quad \forall\, (s,t) \in \Delta_{[0,T]}.$$

Now let $C = (2l-1)\,\tilde{C}$. If $(s,t) \in \Delta_{[0,T]}$ then there exists j and j^* such that $s \in [t_j, t_{j+1}]$ and $t \in [t_{j^*}, t_{j^*+1}]$ with $j \le j^*$. If $j = j^*$, then

$$\left|y_t - y_s - y_s^\dagger\, x_{s,t}\right| \le C\, \omega(s,t)^{2/p}$$

trivially. Otherwise, we have

$$\begin{aligned}
y_t - y_s - y_s^\dagger\, x_{s,t} &= (y_t - y_{t_{j^*}}) + (y_{t_{j+1}} - y_s) + \sum_{i=j+1}^{j^*-1} (y_{t_{i+1}} - y_{t_i}) \\
&\quad - y_s^\dagger\, x_{s,t_{j+1}} - y_s^\dagger\, x_{t_{j^*},t} - \sum_{i=j+1}^{j^*-1} y_s^\dagger\, x_{t_i,t_{i+1}} \\
&= (y_t - y_{t_{j^*}} - y_{t_{j^*}}^\dagger\, x_{t_{j^*},t}) + (y_{t_{j+1}} - y_s - y_s^\dagger\, x_{s,t_{j+1}}) \\
&\quad + [y_{t_{j^*}}^\dagger - y_s^\dagger]\, x_{t_{j^*},t} + \sum_{i=j+1}^{j^*-1} (y_{t_{i+1}} - y_{t_i} - y_{t_i}^\dagger x_{t_i,t_{i+1}}) \\
&\quad + \sum_{i=j+1}^{j^*-1} [y_s^\dagger - y_{t_i}^\dagger]\, x_{t_i,t_{i+1}}.
\end{aligned}$$

Taking absolute values and using the fact that ω is superadditive, we have that the absolute value of each term on the right (including those within the summations) is bounded by $\tilde{C}\, \omega(s,t)^{2/p}$. Thus

$$\left|y_t - y_s - y_s^\dagger\, x_{s,t}\right| \le (2l-1)\, \tilde{C}\, \omega(s,t)^{2/p} = C\, \omega(s,t)^{2/p}. \qquad \square$$

Utilizing the results of Theorem 1 in [16], Theorem 3.3.1 of [26] can be generalized to apply to the class of controlled rough paths. This generalization (after restricting to the case $p < 3$ and matching our notation) results in the following theorem.

Theorem 2.7. *Let $\mathbf{X}$ be a p-rough path on $W \oplus W^{\otimes 2}$ with control ω and let $(y, y^\dagger)$ be an $L(W, V)$-valued rough path controlled by $\mathbf{X}$. Then there exists a $z \in C([0,T], V)$ with $z_0 = 0$ and a $C \geq 0$ such that*

$$\left|z_t - z_s - y_s\, x_{s,t} - y_s^\dagger\, \mathbb{X}_{s,t}\right| \leq C\, \omega(s,t)^{3/p} \tag{2.5}$$

for all $s \leq t$ in $[0,T]$.

We will more commonly refer to the path z_t as $\int_0^t \langle \mathbf{y}_\tau, d\mathbf{X}_\tau \rangle$ and its increment, $z_{s,t} := z_t - z_s$, as $\int_s^t \langle \mathbf{y}_\tau, d\mathbf{X}_\tau \rangle$. Theorem 2.9 below is a generalization of Theorem 2.7, but before we state it, we will make a remark about certain identifications of spaces.

Remark 2.8. If $V, \tilde{V}$, and $\hat{V}$ are vector spaces, we can make the identification

$$L(V, L(\tilde{V}, \hat{V})) \cong L(V \otimes \tilde{V}, \hat{V})$$

via the map $\Xi : L(V, L(\tilde{V}, \hat{V})) \to L(V \otimes \tilde{V}, \hat{V})$ given by

$$\Xi(\alpha)[v \otimes \tilde{v}] = \alpha\langle v \rangle \langle \tilde{v} \rangle.$$

if $\alpha \in L(V, L(\tilde{V}, \hat{V}))$.

The proof of the following theorem (modulo a reparameterization) may be found in [16] or in [14], Remark 4.11.

Theorem 2.9. *Let $\mathbf{X}$ be a p-rough path on $W \oplus W^{\otimes 2}$ with control ω, let $(y, y^\dagger)$ be an V-valued rough path controlled by $\mathbf{X}$ and let $\boldsymbol{\alpha} = (\alpha, \alpha^\dagger)$ be an $L(V, \tilde{V})$-valued rough path controlled by $\mathbf{X}$, where $\alpha_s^\dagger \in L(W, L(V, \tilde{V})) \cong L(W \otimes V, \tilde{V})$. Then there exists a $z \in C([0,T], V)$ with $z_0 = 0$ and a $C > 0$ such that*

$$\left|z_t - z_s - \alpha_s\,(y_t - y_s) - \alpha_s^\dagger\,\left(I \otimes y_s^\dagger\right) \mathbb{X}_{s,t}\right| \leq C\, \omega(s,t)^{3/p} \tag{2.6}$$

for all $s \leq t$ in $[0,T]$. Moreover, if we let $z_s^\dagger := \alpha_s \circ y_s^\dagger$, then $\mathbf{z}_s := (z_s, z_s^\dagger)$ is a $\tilde{V}$-valued controlled rough path.

The path z_t in this case will be denoted $\int_0^t \langle \boldsymbol{\alpha}_\tau, d\mathbf{y}_\tau \rangle$ and we will typically summarize inequality (2.6) by writing

$$\int_s^t \langle \boldsymbol{\alpha}_\tau, d\mathbf{y}_\tau \rangle \underset{3}{\approx} \left\langle \boldsymbol{\alpha}_s, \mathbf{y}_{s,t}^{\mathbb{X}} \right\rangle := \alpha_s\, y_{s,t} + \alpha_s^\dagger\,\left(I \otimes y_s^\dagger\right) \mathbb{X}_{s,t} \tag{2.7}$$

wherein we let $\mathbf{y}_{s,t}^{\mathbb{X}}$ be the increment process defined by,

$$\mathbf{y}_{s,t}^{\mathbb{X}} := \left(y_{s,t}, \left(I \otimes y_s^\dagger\right) \mathbb{X}_{s,t}\right). \tag{2.8}$$

Notice that Theorem 2.7 does indeed follow from Theorem 2.9 upon replacing $\left(\alpha, \alpha^\dagger\right)$ by $(y, y^\dagger)$ and $(y, y^\dagger)$ by (x, I_W) in inequality (2.6).

Remark 2.10 (Motivations). In order to develop some intuition for the expression appearing on the right side of equation (2.7), suppose for the moment that all functions $\mathbf{X}$, $(y, y^\dagger)$, and $(\alpha, \alpha^\dagger)$ are smooth and let $\mathbb{X}$ be given by equation (2.4). In this case we want $z_{s,t}$ to be the usual integral $\int_s^t \alpha_\tau \dot{y}_\tau \, d\tau$ and to arrive at the expression in inequality (2.6) we look for an appropriate second order approximation to the integral. Since $p = 1$ now we may conclude

$$\alpha_{s,\tau} = \alpha_s^\dagger \, x_{s,\tau} + O((\tau - s)^2)$$

and

$$y_t - y_\tau = y_\tau^\dagger (x_t - x_\tau) + O((t-\tau)^2) \implies \dot{y}_\tau = y_\tau^\dagger \, \dot{x}_\tau.$$

We have the identity

$$\int_s^t \alpha_\tau \, dy_\tau = \int_s^t [\alpha_s + \alpha_{s,\tau}] \, \dot{y}_\tau \, d\tau = \alpha_s \, y_{s,t} + \int_s^t \alpha_{s,\tau} \, \dot{y}_\tau \, d\tau. \tag{2.9}$$

The last term on the right-hand side is approximated up to an error of size $O((t-s)^3)$ as follows:

$$\begin{aligned} \int_s^t \alpha_{s,\tau} \, \dot{y}_\tau \, d\tau &= \int_s^t \alpha_{s,\tau} \, y_\tau^\dagger \, \dot{x}_\tau \, d\tau = \int_s^t \alpha_s^\dagger \, x_{s,\tau} \, y_\tau^\dagger \, \dot{x}_\tau \, d\tau + O((t-s)^3) \\ &= \int_s^t \alpha_s^\dagger \, x_{s,\tau} \, y_s^\dagger \, \dot{x}_\tau \, d\tau + O((t-s)^3) \\ &= \alpha_s^\dagger (I \otimes y_s^\dagger) \int_s^t x_{s,\tau} \otimes \dot{x}_\tau d\tau + O((t-s)^3) \\ &= \alpha_s^\dagger (I \otimes y_s^\dagger) \, \mathbb{X}_{s,t} + O((t-s)^3). \end{aligned} \tag{2.10}$$

Combining (2.9) and (2.10) gives the approximate equality

$$\int_s^t \alpha_\tau \, dy_\tau = \alpha_s \, y_{s,t} + \alpha_s^\dagger \left(I \otimes y_s^\dagger\right) \mathbb{X}_{s,t} + O((t-s)^3).$$

Controlled rough paths are also useful in interpreting solutions to rough differential equations. Let $F \colon V \to L(W, V)$ be smooth where we will write $F(a)w$ as $F_w(a)$. We can then make sense of the rough differential equation

$$d\mathbf{y}_t = F_{d\mathbf{X}_t}(y_t) \tag{2.11}$$

with initial condition $y_0 = \bar{y}_0$. We will need a bit of notation regarding tensor products before we say what it means to solve such an equation.

Notation 2.11. If $\Xi : W \times W \to V$ is a bilinear form into a vector space V, by the universal property of tensor products, Ξ factors through a unique linear function $\Xi^\otimes$ on $W \otimes W$ such that $\Xi^\otimes(w \otimes \tilde{w}) = \Xi(w, \tilde{w})$ for a simple tensor $w \otimes \tilde{w}$. If $\mathbb{W} \in W \otimes W$ we will abuse notation and write

$$\Xi\left(w, \tilde{w}\right)|_{w \otimes \tilde{w} = \mathbb{W}} = \Xi\left(w \otimes \tilde{w}\right)|_{w \otimes \tilde{w} = \mathbb{W}} = \Xi^\otimes\left(\mathbb{W}\right),$$

where, to be precise, if $\mathbb{W} = \sum w_i \otimes \tilde{w}_i$ then

$$\Xi^\otimes\left(\mathbb{W}\right) = \sum \Xi\left(w_i, \tilde{w}_i\right).$$

We say the controlled rough path $\mathbf{y} = (y, y^\dagger)$ defined on[2] $I_0 = [0, T)$ or $I_0 = [0, T]$ solves equation (2.11) if for every $[0, b] \subseteq I_0$, we have

$$y_{s,t} \underset{3}{\approx} F_{x_{s,t}}(y_s) + \left(\partial_{F_w(y_s)} F_{\tilde{w}}\right)(y_s)|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}}$$
$$y_s^\dagger = F_\cdot(y_s)$$

for all $s, t \in [0, b]$. If in addition $y_0 = \bar{y}_0$, we say $\mathbf{y}$ solves equation (2.11) with initial condition $y_0 = \bar{y}_0$.

The existence and uniqueness of solutions (at least of the path y_s) to these differential equations (provided F is sufficiently regular) is due to Lyons [26]. Clearly if y_s is given, then $y_s^\dagger$ exists and is uniquely determined by $y_s^\dagger = F_\cdot(y_s)$. One may refer to Subsection 6.5 in the Appendix for more results regarding rough differential equations on Euclidean space.

2.2. Manifold-valued controlled rough paths

Let $M = M^d$ be a d-dimensional manifold, TM be its tangent space, and $\pi := \pi_{TM} : TM \to M$ be the natural projection map. Throughout, let $\mathbf{X} = (x, \mathbb{X})$ be a weak-geometric p-rough path on $[0, T]$ with values in $W \oplus W^{\otimes 2}$ and control ω.

The letters x and y will appear in this paper generally as paths, but occasionally they will refer to arbitrary points in Euclidean space. The context will allow the reader to identify their proper usage.

Notation 2.12. When $M = \mathbb{R}^d$ we will identify $T\mathbb{R}^d$ with $\mathbb{R}^d \times \mathbb{R}^d$ via

$$\mathbb{R}^d \times \mathbb{R}^d \ni (m, v) \to v_m := \frac{d}{dt}\Big|_0 (m + tv) \in T_m\mathbb{R}^d$$

and, by abuse of notation, we let $|v_m| = |v|$ when $|\cdot|$ is the standard Euclidean norm.

Notation 2.13. Whenever ϕ is a map, let $D(\phi)$ and $R(\phi)$ denote the domain and range of ϕ respectively. If $\phi \in C^\infty(M, \mathbb{R}^{d'})$ has open domain, let $d\phi : TD(\phi) \to \mathbb{R}^{d'}$ be defined by

$$d\phi(v_m) := \frac{d}{dt}\Big|_0 \phi(\sigma(t)) \in \mathbb{R}^{d'}, \tag{2.12}$$

where σ is such that $\sigma(0) = m \in D(\phi)$ and $\dot{\sigma}(0) = v_m \in T_mM$. Denote $d\phi_m := d\phi|_{T_mM}$. If $f \in C^\infty(M, \tilde{M})$ where $\tilde{M}$ is another manifold, we let f_* be the push-forward of f so that $f_* : TD(f) \to T\tilde{M}$ is defined by

$$f_*(v_m) := \frac{d}{dt}\Big|_0 f(\sigma(t)) \in T_{f(m)}\tilde{M},$$

where again $\dot{\sigma}(0) = v_m$. Analogously we let $f_{*m} = f_*|_{T_mM}$. Note that $\phi_*(v_m) = (\phi(m), d\phi(v_m)) = [d\phi(v_m)]_{\phi(m)}$.

[2]Here we allow that $\mathbf{y} \in CRP_{\mathbf{X}}(I_0, V)$ if it is an element of $CRP_{\mathbf{X}}(K, V)$ for every compact interval $K \in I_0$.

2.3. Gauges

Definition 2.14. Let $\mathcal{U}$ be an open set on M. An open set $\mathcal{D}^{\mathcal{U}} \subseteq M \times M$ is a **$\mathcal{U}$-diagonal domain** if it contains the diagonal of $\mathcal{U}$, that is, $\Delta^{\mathcal{U}} := \bigcup_{m\in\mathcal{U}}(m,m) \subseteq \mathcal{D}^{\mathcal{U}}$. A **local diagonal domain** is a $\mathcal{V}$-diagonal domain for some nonempty open $\mathcal{V} \subseteq M$.

If $\mathcal{U} = M$ we write $\mathcal{D} := \mathcal{D}^M$ and refer to $\mathcal{D}$ simply as a diagonal domain.

Throughout the paper, $\mathcal{D}$ will always denote a diagonal domain.

Definition 2.15. A smooth function $\psi : \mathcal{D} \to TM$ is called a **logarithm** if:

1. $\psi(m,n) \in T_mM$,
2. $\psi(m,m) = 0_m$,
3. $\psi(m,\cdot)_* |_{T_mM} = I_m$.

We also write ψ_m for $\psi(m,\cdot)$.

If the above holds for ψ defined on a local diagonal domain, we may refer to ψ as a **local logarithm**.

If E is a vector bundle, we will denote the smooth sections of E by $\Gamma(E)$. We define $L(TM,TM)$ as the vector bundle $\tilde{E}$ over the manifold $M \times M$ such that $\tilde{E}_{(n,m)} = L(T_mM, T_nM)$ and

$$\tilde{E} = \bigcup \{\tilde{E}_{(n,m)} : n, m \in M\}.$$

Definition 2.16. A smooth section $U \in \Gamma(L(TM,TM))$ with domain $\mathcal{D}$ (i.e., $U(n,m) \in L(T_mM, T_nM)$ for all $(n,m) \in \mathcal{D}$) is called a **parallelism** if

$$U(m,m) = I_m.$$

If U is only defined on a local diagonal domain, we refer to U as a **local parallelism**.

Definition 2.17. We call the pair $\mathcal{G} := (\psi, U)$ (where ψ and U have common domain $\mathcal{D}$) a **gauge** on the manifold M. If $\mathcal{D}$ is replaced by a local diagonal domain, we call $\mathcal{G}$ a **local gauge**.

Example 2.18. If $M = \mathbb{R}^d$, the maps $\psi(x,y) = [y-x]_x$ and $U(x,y)v_y = v_x$ form the **standard gauge** on $\mathbb{R}^d$.

Example 2.19. One natural example of a gauge comes from any covariant derivative ∇ on TM. The construction is as follows. Choose an arbitrary Riemannian metric g on M. If $m, n \in M$ are "close enough", there is a unique vector v_m with minimum length such that $n = \exp_m^{\nabla}(v_m)$. We denote this vector by $\psi^{\nabla}(m,n) := (\exp_m^{\nabla})^{-1}(n)$ or by $\exp_m^{-1}(n)$ if ∇ is clear from the context. We further let

$$U^{\nabla}(n,m) := //_1 \left(t \to \exp_m \left(t \exp_m^{-1}(n)\right)\right),$$

where, for any smooth curve $\sigma : [0,1] \to M$, we let $//_s(\sigma) = //_s^{\nabla}(\sigma) : T_{\sigma(0)}M \to T_{\sigma(s)}M$ denote parallel translation along σ up to time $s \in [0,1]$. It is shown in Corollary 2.33 that there is a diagonal domain $\mathcal{D} \subseteq M \times M$ such that $\left(\psi^{\nabla}, U^{\nabla}\right)$ so defined is a gauge on $\mathcal{D}$.

Remark 2.20. We can also get a covariant derivative from a parallelism. If U is a parallelism, then we can define covariant derivative ∇^U on TM by

$$\nabla^U_{v_m}(Y) := \frac{d}{dt}\Big|_0 U(m,\sigma_t)\, Y(\sigma_t),$$

where $\dot{\sigma}(0) = v_m$ and Y is a vector field on M.

Remark 2.21. Although the definition of a gauge includes stipulating a U, if we have just ψ, we can define $U^{\psi}(n,m) := \psi(n,\cdot)_{*m}$ and set $\mathcal{G}^{\psi} := \left(\psi, U^{\psi}\right)$.

Remark 2.22. We may make a local gauge out of a chart ϕ. Indeed, we pull back the flat gauge in Example 2.18 to M to define

$$\begin{aligned}\psi^{\phi}(m,n) &:= (d\phi_m)^{-1}\left[\phi(n) - \phi(m)\right],\\ U^{\phi}(n,m) &:= (d\phi_n)^{-1}\, d\phi_m.\end{aligned}$$

This is a gauge which is also consistent with Remark 2.21 and $D(\psi^{\phi}) = D\left(U^{\phi}\right) = D(\phi) \times D(\phi)$.

Before moving on to controlled rough paths on manifolds, let us record the structure of the general gauge on $\mathbb{R}^d$.

Notation 2.23. If (ψ, U) is a local gauge on $\mathbb{R}^d$, then we write $\left(\bar{\psi}, \bar{U}\right)$ to mean the functions determined by the relations

$$\psi(x,y) = \left[\bar{\psi}(x,y)\right]_x \quad \text{and} \quad U(x,y)(v_y) = \left[\bar{U}(x,y)v\right]_x$$

so that $\bar{\psi}(x,y) \in \mathbb{R}^d$ and $\bar{U}(x,y) \in \operatorname{End}\left(\mathbb{R}^d\right)$.

Theorem 2.24. *If $\mathcal{G} = (\psi, U)$ is a local gauge on $\mathbb{R}^d$, for every open convex subset $\mathcal{V} \subseteq \mathbb{R}^d$ such that $\mathcal{V} \times \mathcal{V} \subseteq D(\mathcal{G})$, there exists smoothly varying functions $A(x,y) \in L\left((\mathbb{R}^d)^{\otimes 2}, \mathbb{R}^d\right)$ and $B(x,y) \in L(\mathbb{R}^d, \operatorname{End}(\mathbb{R}^d))$ defined for $(x,y) \in \mathcal{V} \times \mathcal{V}$ such that*

$$\bar{U}(x,y) = I + B(x,y)(y-x), \tag{2.13}$$

$$\bar{\psi}(x,y) = y - x + A(x,y)(y-x)^{\otimes 2}, \tag{2.14}$$

$$B(x,x) = D_2\bar{U}(x,x), \text{ and } A(x,x) = \frac{1}{2}\left(D_2^2\bar{\psi}\right)(x,x). \tag{2.15}$$

The converse holds as well.

Furthermore, we can find a smoothly varying function

$$C(x,y) \in L\left((\mathbb{R}^d)^{\otimes 3}, \mathbb{R}^d\right)$$

defined on $\mathcal{V} \times \mathcal{V}$ *such that*

$$C(x,x) = \frac{1}{6}\left(D_2^3\bar{\psi}\right)(x,x), \text{ and} \tag{2.16}$$

$$\bar{\psi}(x,y) = y - x + \frac{1}{2}\left(D_2^2\bar{\psi}\right)(x,x)(y-x)^{\otimes 2} + C(x,y)(y-x)^{\otimes 3}. \tag{2.17}$$

Proof. Let x, y be points in $\mathcal{V}$. Taylor's theorem with integral remainder applied to the second variable with x fixed gives,

$$\bar{U}(x,y) = I + \int_0^1 \left(D_2\bar{U}\right)(x, x + t(y-x))(y-x)\,dt$$

and

$$\bar{\psi}(x,y) = 0 + (D_2\bar{\psi})(x,x)(y-x) + \int_0^1 (D_2^2\bar{\psi})(x, x+t(y-x))(y-x)^{\otimes 2}(1-t)\,dt,$$

from which equations (2.13)–(2.15) follow with

$$B(x,y) = \int_0^1 \left(D_2\bar{U}\right)(x, x+t(y-x))\,dt \quad \text{and}$$
$$A(x,y) = \int_0^1 \left(D_2^2\bar{\psi}\right)(x, x+t(y-x))(1-t)\,dt.$$

The converse statement is easy to verify. The proof of equations (2.16) and (2.17) also follow by Taylor's theorem (now to third order), in which case

$$C(x,y) = \frac{1}{2}\int_0^1 \left(D_2^3\bar{\psi}\right)(x, x+t(y-x))(1-t)^2 dt. \qquad \square$$

Let $B_r(x) \subseteq \mathbb{R}^d$ be the open ball of radius r centered at x.

Remark 2.25. If ψ and $\tilde{\psi}$ are local logarithms on $\mathbb{R}^d$, it is easy to check using Theorem 2.24 that for all $\tilde{x} \in \mathbb{R}^d$, there exists an $r > 0$ and $C > 0$ such that $|\psi(x,y)| \le C|\tilde{\psi}(x,y)|$ for all $x, y \in B_r(\tilde{x})$.

We now wish to transfer these local results to the manifold setting. In order to do this we need to develop some notation for stating that two objects on a manifold are "close" up to some order. Let g be any smooth Riemannian metric on M.

Notation 2.26. We write d_g for the metric associated to g and define $|v_m|_g := \sqrt{g_m(v_m, v_m)}\ \forall\ v_m \in TM$. Further, we let $|\cdot|_{g,op}$ be the operator "norm" induced by $|\cdot|_g$ on $L(TM, V)$, i.e., if $f_m \in L(T_mM, V)$, then

$$|f_m|_{g,op} := \sup\{|f_m\langle v_m\rangle| : |v_m|_g = 1\}.$$

Definition 2.27. Let F, G be smooth TM (respectively, $L(TM, TM)$) valued functions with $\mathcal{W}$-diagonal domains. The expression

$$F(m,n) =_k G(m,n) \quad \text{on } \mathcal{W} \tag{2.18}$$

indicates that for every point in $w \in \mathcal{W}$, there exists an open $\mathcal{O}_w \subseteq M$ containing w such that $\mathcal{O}_w \times \mathcal{O}_w \subseteq D(F) \cap D(G)$ and a $C > 0$ such that

$$|F(m,n) - G(m,n)|_{g,[g,op]} \leq C\,(d_g(m,n))^k \tag{2.19}$$

for all $m, n \in \mathcal{O}_w$.

Occasionally we will omit the reference to $\mathcal{W}$ in which case we mean the condition (2.19) holds where it makes sense to hold.

Note that in (2.18), the reference to g is not explicit. In fact, the definition does not depend on the choice of g as all Riemannian metrics are locally equivalent. (See Corollary 6.6 in the Appendix for precise statement and proof of this standard fact.)

We may also use the $=_k$ notation to make statements in regards to other measures of distance.

Corollary 2.28. *Let $\mathcal{W}$ be an open subset of M and g and $\tilde{g}$ be any two Riemannian metrics on M. If $F(m,n) =_k G(m,n)$ on $\mathcal{W}$ (so that F and G have $\mathcal{W}$-diagonal domains), then for every local logarithm ψ and $w \in \mathcal{W}$ such that $(w,w) \in D(\psi)$, there exists an open $\mathcal{O}_w \subseteq \mathcal{W}$ containing w and $C > 0$ such that*

$$|F(m,n) - G(m,n)|_{g,[g,op]} \leq C\,|\psi(m,n)|_{\tilde{g}}^k \quad \forall\, m,n \in \mathcal{O}_w.$$

In particular, using the local logarithm $\psi(m,n) = (d\phi_m)^{-1}[\phi(n) - \phi(m)]$, we have that if $w \in D(\phi) \cap \mathcal{W}$, then there exists an $\mathcal{O}_w \subseteq D(\phi) \cap \mathcal{W}$ and a $C > 0$ such that

$$|F(m,n) - G(m,n)|_{g,[g,op]} \leq C\,|\phi(n) - \phi(m)|^k \quad \forall\, m,n \in \mathcal{O}_w.$$

Proof. The proof of the corollary will use Remark 2.25 and the local equivalence of any two Riemannian metrics, Corollary 6.6 in the Appendix. First we simplify matters by assuming that we are working in Euclidean space which may be accomplished by pushing the metric and functions forward using charts. Assuming this, we now derive a local inequality that holds for any two logarithms ψ and $\tilde{\psi}$ when $(w,w) \in D(\psi) \cap D(\tilde{\psi})$. Namely, there exist an open neighborhood, $\mathcal{O}_w$, of w such that

$$|\tilde{\psi}(m,n)|_g \leq C_1\,|\tilde{\psi}(m,n)| \leq C_2\,C_1\,|\psi(m,n)| \leq C_3\,C_2\,C_1\,|\psi(m,n)|_{\tilde{g}}$$

for all $(m,n) \in \mathcal{O}_w \times \mathcal{O}_w$ where the first and third inequality follow from Corollary 6.6 with one metric being the standard Euclidean metric and the other metric being g or $\tilde{g}$ respectively, and the second inequality is true by Remark 2.25. Thus, there exists a $\tilde{C}$ such that

$$\left|\tilde{\psi}(m,n)\right|_g \leq \tilde{C}\,|\psi(m,n)|_{\tilde{g}}\,.$$

Now let ∇^g be the Levi-Civita covariant derivative associated to g. By setting $\tilde{\psi}(m,n) = (\exp_m^{\nabla^g})^{-1}(n)$ and shrinking $\mathcal{O}_w$ if necessary to ensure that the function $\left(\exp_{(\cdot)}^{\nabla^g}\right)^{-1}(\cdot)$ is defined and injective on $\mathcal{O}_w \times \mathcal{O}_w$, we have that

$$\left|\left(\exp_m^{\nabla^g}\right)^{-1}(n)\right|_g \leq \tilde{C}\ |\psi(m,n)|_{\tilde{g}}\,.$$

In this setting, $d_g(m,n) = \left|(\exp_m^{\nabla^g})^{-1}(n)\right|_g$, and since $F(m,n) =_k G(m,n)$ on $\mathcal{W}$ (by shrinking $\mathcal{O}_w$ if necessary), we have

$$|F(m,n) - G(m,n)|_{g,[g,op]} \leq \hat{C}\,(d_g(m,n))^k \quad \forall\, m,n \in \mathcal{O}_w$$

for some $\hat{C}$. Thus, we have

$$|F(m,n) - G(m,n)|_{g,[g,op]} \leq \hat{C}(\tilde{C})^k\ |\psi(m,n)|_{\tilde{g}}^k\,.$$

which is the statement of the Corollary with $C := \hat{C}(\tilde{C})^k$. □

In the sequel, Corollary 2.28 will typically be used without further reference in order reduce the proof of showing $F(m,n) =_k G(m,n)$ in the manifold setting to a local statement about functions on convex neighborhoods in $\mathbb{R}^d$ equipped with the standard Euclidean flat metric structures. The first example of this strategy will already occur in the proof of Corollary 2.29 below. For a general parallelism it is not true that $U(n,m)^{-1} = U(m,n)$, yet $U(m,n)$ is always a very good approximation to $U(n,m)^{-1}$.

Corollary 2.29. *If U is a parallelism on a manifold M, then*

$$U(n,m)^{-1} =_2 U(m,n).$$

Proof. This is a local statement so we may use Corollary 2.28 to reduce to the case that M is a convex open subset of $\mathbb{R}^d$. We then may use Theorem 2.24 to learn

$$\bar{U}(n,m)^{-1} = (I + [B(n,m)(m-n)])^{-1} = I + [B(n,m)(n-m)] + O(|n-m|^2)$$

while

$$\bar{U}(m,n) = (I + [B(m,n)\,(n-m)])\,.$$

Subtracting these two equations shows

$$\bar{U}(n,m)^{-1} - \bar{U}(m,n) = [B(n,m) - B(m,n)](n-m) + O(|n-m|^2) = O(|n-m|^2)$$

wherein we have used $B(n,m) - B(m,n)$ vanishes for $m = n$ and therefore is of order $|m-n|$. □

2.3.1. A Covariant derivative gives rise to a gauge. Let ∇ be a covariant derivative on TM, and g be any fixed Riemannian metric on M. Let $G\colon TM \to M \times M$ be the function on TM defined by

$$G(v_m) := \left(m, \exp_m^{\nabla}(v_m)\right) \quad \text{for all } v_m \in D\left(G\right), \tag{2.20}$$

where $D\left(G\right)$ is the domain of G defined by

$$D\left(G\right) := \left\{v_m \in TM : t \to \exp_m^{\nabla}\left(tv_m\right) \text{ exists for } 0 \le t \le 1\right\}.$$

We will now develop a subset of $D\left(G\right)$ for which G is injective. For each $m \in M$, let Λ_m denote the set of $r > 0$ so that $B_r\left(0_m\right) \subseteq D\left(G\right)$, $\exp_m^{\nabla}\left(B_r\left(0_m\right)\right)$ is an open neighborhood of m in M, and $\exp_m^{\nabla} : B_r\left(0_m\right) \to \exp_m^{\nabla}\left(B_r\left(0_m\right)\right)$ is a diffeomorphism (here $B_r\left(0_m\right)$ is the open ball in T_mM centered at 0_m with radius r). The fact that Λ_m is not empty is a consequence of the inverse function theorem and the fact that $(\exp_m^{\nabla})_{*0_m} = I_{T_mM}$ is invertible. We now define $r_m := \sup \Lambda_m$ where $r_m = \infty$ is possible and allowed. A little thought shows that $\exp_m^{\nabla}\left(B_{r_m}\left(0_m\right)\right)$ is open and $\exp_m^{\nabla} : B_{r_m}\left(0_m\right) \to \exp_m^{\nabla}\left(B_{r_m}\left(0_m\right)\right)$ is a diffeomorphism, i.e., either $r_m = \infty$ or $r_m \in \Lambda_m$.

Let us now set $\mathcal{C}^* := \cup_{m\in M} B_{r_m}\left(0_m\right) \subseteq TM$ and let $G^*\colon \mathcal{C}^* \to M \times M$ be the map defined by

$$G^*(v_m) := \left(m, \exp_m^{\nabla}\left(v_m\right)\right) \quad \text{for all } v_m \in \mathcal{C}^*.$$

It is easy to verify that G^* is injective.

We will now build our domain $\mathcal{C}$ for which $G|_{\mathcal{C}}$ is diffeomorphic onto its range. First we need a simple local invertibility proposition.

Proposition 2.30. *Let G be the function defined in equation (2.20). Then for each $m \in M$, there exists open subsets $\mathcal{V}_m \subseteq TM$ and $\mathcal{W}_m \subseteq M$ such that $0_m \in \mathcal{V}_m$, $m \in \mathcal{W}_m$, and $G|_{\mathcal{V}_m} : \mathcal{V}_m \to \mathcal{W}_m \times \mathcal{W}_m$ is a diffeomorphism.*

Proof. As this a local result we may assume that $M = \mathbb{R}^d$ and identify TM with $M \times M$. The function $G\colon TM \to M \times M$ then takes on the form $G\left(x, v\right) = \left(x, \bar{G}\left(x, v\right)\right)$ where $\bar{G}\left(x, 0\right) = x$ and $\left(D_2\bar{G}\right)\left(x, 0\right) = I_M$ for all $x \in M$. A simple computation then shows

$$G'\left(x, 0\right) = \begin{bmatrix} I & 0 \\ I & I \end{bmatrix} \quad \text{for all } x \in M.$$

The result now follows by an application of the inverse function theorem. □

Notation 2.31. If $\mathcal{W}$ is an open subset of M and $\epsilon > 0$, let $\mathcal{U}(\mathcal{W}, \epsilon)$ be the open subset of TM defined by

$$\mathcal{U}(\mathcal{W}, \epsilon) := \{v \in \pi^{-1}(\mathcal{W}) \subseteq TM : |v|_g < \epsilon\}.$$

Theorem 2.32. *Let $\mathcal{C} := \bigcup \mathcal{U}(\mathcal{W}, \epsilon)$ where the union is taken over all open subsets $\mathcal{W} \subseteq M$ and $\epsilon > 0$ such that $\mathcal{U}(\mathcal{W}, \epsilon) \subseteq D(G)$ and $G|_{\mathcal{U}(\mathcal{W},\epsilon)} : \mathcal{U}(\mathcal{W}, \epsilon) \to G(\mathcal{U}(\mathcal{W}, \epsilon))$ is a diffeomorphism. Then $\mathcal{C}$ is an open subset of TM such that $\mathcal{D} := G(\mathcal{C})$ is open in $M \times M$, $G\colon \mathcal{C} \to \mathcal{D}$ is a diffeomorphism,*

$$\{0_m : m \in M\} \subseteq \mathcal{C} \subseteq \mathcal{C}^* \quad \text{and} \quad \Delta^M = \{(m, m) : m \in M\} \subseteq \mathcal{D}.$$

Proof. According to Proposition 2.30, for each $m \in M$ there exists an open neighborhood $\mathcal{W}$ of $m \in M$ and $\epsilon > 0$ so that $\mathcal{U}(\mathcal{W}, \epsilon) \subseteq D(G)$ and $G\colon \mathcal{U}(\mathcal{W}, \epsilon) \to G(\mathcal{U}(\mathcal{W}, \epsilon))$ is a diffeomorphism. From this it follows that $\{0_m : m \in \mathcal{W}\} \subseteq \mathcal{C}$ and $\mathcal{U}(\mathcal{W}, \epsilon) \subseteq \mathcal{C}^*$. As $m \in M$ was arbitrary we may conclude $\{0_m : m \in M\} \subseteq \mathcal{C} \subseteq \mathcal{C}^*$. It is now easily verified that $G(\mathcal{C}) = \cup G(\mathcal{U}(\mathcal{W}, \epsilon))$ is open, $G\colon \mathcal{C} \to G(\mathcal{C})$ is a surjective local diffeomorphism and hence is a diffeomorphism as $G|_{\mathcal{C}}$ is injective (since $G|_{\mathcal{C}^*}$ is injective). □

Corollary 2.33. *Continuing the notation used in Theorem 2.32, we have $\mathcal{D}$ is a diagonal domain and $\psi := G|_{\mathcal{C}}^{-1}\colon \mathcal{D} \to \mathcal{C} \subseteq TM$ is a logarithm. Moreover, if we define*

$$U(m, n) := //_1 \left(t \to \exp^{\nabla}(t\psi(m, n))\right)^{-1} : T_nM \to T_mM$$

for all $(m, n) \in \mathcal{D}$, then U is a parallelism on M.

Proof. The only thing that remains to be proven is that $U(m, n)$ is smoothly varying. This is a consequence of the fact that solutions to ordinary differential equations depend smoothly on their starting points and parameter in the vector fields. To be more explicit in this case, for $a \in \mathbb{R}^d$ let $B_a^{\nabla}(\mu) = \dot{u}(0)$ where $u(t) = //_t \left(\exp^{\nabla}((\cdot)\mu a)\right)\mu$ for μ in the frame bundle $GL(M)$ over M, so that B_a^{∇} are the ∇-horizontal vector fields. Now suppose that $w \in M$ is given and $O(m) : \mathbb{R}^d \to T_mM$ is a local frame defined for m in an open neighborhood $\mathcal{W}$ of w. For $v \in \pi^{-1}(\mathcal{W}) \cap \mathcal{C}$ let $\gamma(t) = \exp^{\nabla}(tv)$ and $u(t) := //_t(\gamma)\, O(\pi(v))$. We then have

$$\dot{\gamma}(t) = //_t(\gamma)\, v = u(t)\, O(\pi(v))^{-1} v \quad \text{and}$$
$$\frac{\nabla u}{dt} = 0 \quad \text{with } u(0) = O(\pi(v)).$$

These equations are equivalent to solving

$$\dot{u}(t) = B^{\nabla}_{O(\pi(v))^{-1}v}(u(t)) \quad \text{with } u(0) = O(\pi(v)), \tag{2.21}$$

in which case $\gamma(t) = \pi_{O(M)}(u(t))$ where $\pi_{O(M)}$ is the projection map from $O(M)$ to M. We now define $F(v) := u(1)$ provided $v \in \pi^{-1}(\mathcal{W}) \cap \mathcal{C}$. It then follows that $F\colon \pi^{-1}(\mathcal{W}) \cap \mathcal{C} \to GL(M)$ is smooth as the solutions to equation (2.21) depend smoothly on its starting point and parameter. From this we learn for $(m, n) \in G\left(\pi^{-1}(\mathcal{W}) \cap \mathcal{C}\right)$ that

$$U(n, m) = F(\psi(m, n))\, O(m)^{-1}$$

is a smooth function of (m, n). □

2.4. Controlled rough paths

Notation 2.34. Throughout the remainder of this paper, $\mathbf{y} := (y, y^\dagger)$ denotes a pair of continuous functions, $y \in C([0,T], M)$ and $y^\dagger \in C([0,T], L(W, TM))$, such that $y^\dagger_s \in L(W, T_{y_s}M)$ for all s.

Definition 2.35. Let (ψ, U) be a gauge. The pair $(y_s, y^\dagger_s)$ is (ψ, U)-**rough path controlled by X** if there exists a $C > 0$ and $\delta > 0$ such that

$$|\psi(y_s, y_t) - y^\dagger_s\, x_{s,t}|_g \le C\,\omega(s,t)^{2/p} \tag{2.22}$$

and

$$|U(y_s, y_t)\, y^\dagger_t - y^\dagger_s|_g \le C\,\omega(s,t)^{1/p} \tag{2.23}$$

hold whenever $0 \le s \le t \le T$ and $|t - s| \le \delta$. Occasionally we will refer to y_s as the path and $y^\dagger_s$ as the derivative process (or Gubinelli derivative).

Remark 2.36. In Definition 2.35 and in the definitions that follow, we use the convention that the δ is small enough to ensure that all of the expressions are well defined (in particular here it is small enough to ensure $(y_s, y_t) \in \mathcal{D}$).

Remark 2.37. Any path z_s in Euclidean space naturally gives rise to a two-parameter "increment process", namely $z_{s,t} = z_t - z_s$. If φ is any function such that $\varphi(z, \tilde{z}) \approx \tilde{z} - z$, then it makes sense to define $z^\varphi_{s,t} := \varphi(z_s, z_t)$. This serves as motivation for the following notation.

Notation 2.38. Given a gauge, $\mathcal{G} = (\psi, U)$, let $y^\psi_{s,t} := \psi(y_s, y_t)$ and $(y^\dagger)^U_{s,t} := U(y_s, y_t)\, y^\dagger_t - y^\dagger_s$. These will be referred to as the $\mathcal{G}$-**local increment processes of** $(y, y^\dagger)$.

Remark 2.39. With Notation 2.38, (2.22) becomes $|y^\psi_{s,t} - y^\dagger_s\, x_{s,t}| \le C\,\omega(s,t)^{2/p}$ and (2.23) becomes $|(y^\dagger)^U_{s,t}| \le C\,\omega(s,t)^{1/p}$.

Definition 2.35 gives one possible notion of a controlled rough path on a manifold. We can also define such an object without having to provide a metric or gauge by using charts on the manifold.

Definition 2.40. The pair $\mathbf{y}_s = (y_s, y^\dagger_s)$ is a **chart-rough path controlled by X** if for every chart ϕ on M and every $[a,b]$ such that $y([a,b]) \subseteq D(\phi)$ we have the existence of a $C_{\phi,a,b} \ge 0$ such that, for all $a \le s \le t \le b$,

$$|\phi(y_t) - \phi(y_s) - d\phi \circ y^\dagger_s\, x_{s,t}| \le C_{\phi,a,b}\,\omega(s,t)^{2/p} \tag{2.24}$$

and

$$|d\phi \circ y^\dagger_t - d\phi \circ y^\dagger_s| \le C_{\phi,a,b}\,\omega(s,t)^{1/p}. \tag{2.25}$$

We will denote $C_{\phi,a,b}$ by C_ϕ when no confusion is likely to arise.

Notation 2.41. If $(y_s, y_s^\dagger)$ is a chart rough path and ϕ is a chart as in Definition 2.40, we will write $\phi_* \mathbf{y}_s$ to mean

$$\phi_* \mathbf{y}_s := \phi_*(y_s, y_s^\dagger) := (\phi \circ y_s, d\phi \circ y_s^\dagger).$$

Note that as long as y remains away from the boundary of $D(\phi)$, then $\phi_* \mathbf{y}_s$ is a controlled rough path on $\mathbb{R}^d$. Another way to think of this is that a chart controlled rough path is one which pushes forward to a controlled rough path in $\mathbb{R}^d$.

Before moving on, we will make a few remarks.

Remark 2.42. If $y^\dagger$ is any function satisfying the conditions in either of Definitions 2.35 or 2.40, then $s \to y_s^\dagger$ is automatically continuous. For example, if $\left(y_s, y_s^\dagger\right)$ satisfies the conditions of a (ψ, U)-rough path in Definition 2.35, then the function $t \to U(y_s, y_t)\, y_t^\dagger$ is continuous at s and therefore the function $t \to y_t^\dagger = U(y_s, y_t)^{-1} U(y_s, y_t)\, y_t^\dagger$ is continuous at s.

Remark 2.43. If $M = \mathbb{R}^d$ and $\phi = I$ then the chart Definition 2.40 reduces to the usual Definition 2.5 of controlled rough paths. In this case, we identify all the tangent spaces with $\mathbb{R}^d$ and forget the base point in the derivative process.

Remark 2.44. A natural question is how the notions of controlled rough paths generalize in the case $p \geq 3$. While we do not attempt develop the theory in this paper, it is not difficult to reconcile the theory of higher order controlled rough paths (for example, see [14] for a definition) with Definition 2.40 for chart rough paths. However, the generalization of Definition 2.35 of gauge rough paths and what geometric structures are necessary for such a generalization is not immediately obvious; more work will be required to understand these concepts in addition to how results below will extend in this environment.

2.5. Chart and gauge CRP definitions are equivalent

Theorem 2.45. *Let $\mathbf{y} := (y, y^\dagger)$ be a pair of continuous functions as in Notation 2.34, let M be a manifold, and let $\mathcal{G} = (\psi, U)$ be any gauge on M . Then $\mathbf{y}$ is a chart controlled rough path (Definition 2.40) if and only if it is a (ψ, U)-controlled rough path (Definition 2.35).*

Corollary 2.46. *We have the equality of sets*

$$\{(\psi, U)\text{-rough paths}\} = \left\{(\tilde{\psi}, \tilde{U})\text{-rough paths}\right\}$$

for any gauges (ψ, U) and $(\tilde{\psi}, \tilde{U})$ on M.

Notation 2.47. Let $CRP_{\mathbf{X}}(M)$ be the collection of **controlled rough paths in** M, i.e., pairs of functions $\mathbf{y} = (y, y^\dagger)$ as in Notation 2.34 which satisfy either (and hence both) of Definitions 2.35 or 2.40.

We will prove Theorem 2.45 after assembling a number of preliminary results that will be needed in the proof and in the rest of the paper.

2.5.1. Results used in proof of Theorem 2.45. Our first result is a local version of Theorem 2.45.

Theorem 2.48. *Let $\mathcal{G} = (\psi, U)$ be a gauge on $\mathbb{R}^d$, $\mathbf{z} = (z, z^\dagger) \in C([a,b], \mathbb{R}^d) \times C([a,b], L(W, \mathbb{R}^d))$, and $\mathcal{W}$ be an open convex set such that $z([a,b]) \subseteq \mathcal{W}$ and $\mathcal{W} \times \mathcal{W} \subseteq D(\mathcal{G})$. Then $\mathbf{z} \in CRP_{\mathbf{X}}(\mathbb{R}^d)$ if and only if $\mathbf{z}$ is a (ψ, U)-rough path controlled by $\mathbf{X}$ with the choice $\delta := b - a$.*

Proof. Suppose $\mathbf{z} \in CRP_{\mathbf{X}}(\mathbb{R}^d)$. By Theorem 2.24,

$$\bar{\psi}(x,y) = y - x + A(x,y)(y-x)^{\otimes 2} \quad \forall\, x, y \in \mathcal{W}.$$

Clearly A is bounded if it is restricted to x, y in the convex hull of $z([a,b])$ (which is compact and contained in $\mathcal{W}$). Thus, for all such points, we have there exists a C_1 such that

$$|\bar{\psi}(x,y) - (y-x)| \le C_1\, |y-x|^2. \tag{2.26}$$

Taking $y = z_t$ and $x = z_s$ in this inequality shows

$$|\bar{\psi}(z_s, z_t) - z_{s,t}| \le C_1\, |z_t - z_s|^2. \tag{2.27}$$

Since $\mathbf{z} \in CRP_{\mathbf{X}}(\mathbb{R}^d)$, there exists a C_2 such that

$$|z_{s,t} - z_s^\dagger\, x_{s,t}| \le C_2\, \omega(s,t)^{2/p} \tag{2.28}$$

$$|z_{s,t}^\dagger| \le C_2\, \omega(s,t)^{1/p}. \tag{2.29}$$

By enlarging C_2 if necessary we may further conclude

$$|z_{s,t}| \le C_2\, \omega(s,t)^{1/p}. \tag{2.30}$$

Using equations (2.28) and (2.30) in equation (2.27) gives the existence of a $C_3 < \infty$ such that

$$|\bar{\psi}(z_s, z_t) - z_s^\dagger\, x_{s,t}| \le C_3\, \omega(s,t)^{2/p}.$$

By Theorem 2.24 once more, we have

$$\bar{U}(x,y) = I + B(x,y)(y-x). \tag{2.31}$$

As was the case for A, B is bounded on the convex hull of $z([a,b])$ so that there exists a C_4 such that

$$\left|\bar{U}(z_s, z_t)\, z_t^\dagger - z_s^\dagger\right| \le \left|z_{s,t}^\dagger\right| + C_4\, |z_{s,t}| \le (C_2 + C_4 C_2)\, \omega(s,t)^{1/p}.$$

Thus $\mathbf{z}$ is a (ψ, U)-rough path controlled by $\mathbf{X}$ with the choice $\delta := b - a$ where our $C := \max\{C_1, C_2(1 + C_4)\}$.

For the converse direction, suppose $\mathbf{z}$ is a (ψ, U)-rough path controlled by $\mathbf{X}$ with the choice $\delta := b - a$ as in Definition 2.35. From equation (2.26) and the triangle inequality we have

$$|y - x| \le C_1\, |y-x|^2 + |\bar{\psi}(x,y)|.$$

Taking $x = z_s$ and $y = z_t$ in this inequality and using Definition 2.35 we may find $C_2 < \infty$ such that

$$|z_{s,t}| \le C_1\, |z_{s,t}|^2 + |\psi(z_s, z_t)| \le C_1\, |z_{s,t}|^2 + C_2\, \omega(s,t)^{1/p}$$

for all $s \le t$ in $[a,b]$. By the uniform continuity of z on $[a,b]$, there exists $\epsilon > 0$ such that $C_1|z_{s,t}| \le \frac{1}{2}$ when $|t-s| \le \epsilon$ which combined with the previous inequality implies

$$|z_{s,t}| \le 2\, C_2\, \omega(s,t)^{1/p} \quad \text{when } |t-s| \le \epsilon.$$

For general $a \le s \le t \le b$ we may write $z_{s,t}$ as a sum of at most $n \le (b-a)/\epsilon$ increments whose norms are bounded by $2C_2\, \omega(s,t)^{1/p}$ wherein we have repeatedly used the estimate above along with the monotonicity of ω resulting from superadditivity. Thus we conclude, with $C_3 := 2C_2(b-a)/\epsilon < \infty$, that

$$|z_{s,t}| \le C_3\, \omega(s,t)^{1/p} \quad \forall\, s,t \in [a,b].$$

This estimate along with the inequality in equation (2.26) gives

$$|\bar{\psi}(z_s, z_t) - z_{s,t}| \le C_1\, |z_{s,t}|^2 \le C_1\, C_3^2\, \omega(s,t)^{2/p} \quad \forall\, s,t \in [a,b].$$

The previous inequality along with the assumption that $\mathbf{z}$ is a (ψ, U)-rough path shows there exists $C_4 < \infty$ such that

$$\left|z_{s,t} - z_s^\dagger\, x_{s,t}\right| \le \left|z_{s,t} - \bar{\psi}(z_s, z_t)\right| + |\bar{\psi}(z_s, z_t) - z_s^\dagger\, x_{s,t}| \le C_4\, \omega(s,t)^{2/p}.$$

From equation (2.31), there exists a C_5 such that

$$|z_{s,t}^\dagger| \le \left|U(z_s, z_t) z_t^\dagger - z_s^\dagger\right| + C_5\, |z_{s,t}|.$$

This inequality along with the assumption that $\mathbf{z}$ is a (ψ, U)-rough path shows there exists $C_6 < \infty$ such that $|z_{s,t}^\dagger| \le C_6\, \omega(s,t)^{1/p}$ for all $a \le s \le t \le b$. Thus we have shown $\mathbf{z} \in CRP_{\mathbf{X}}(\mathbb{R}^d)$. □

The rest of this section is now devoted to a number of "stitching" arguments which will be used to piece together a number of local versions of Theorem 2.45 over subintervals as described in Theorem 2.48 into the full global version as stated in Theorem 2.45. For the rest of this section let $\mathcal{X}$ be a topological space and $0 \le S < T < \infty$.

Lemma 2.49. *If $y\colon [S,T] \to \mathcal{X}$ is continuous and $y([S,T]) \subseteq \bigcup_{\alpha \in A} \mathcal{O}_\alpha$, where $\{\mathcal{O}_\alpha\}_{\alpha \in A}$ is a collection of open subsets of $\mathcal{X}$, then there exists a partition of $[S,T]$, $S = t_0 < t_1 < \ldots < t_l = T$, and $\alpha_i \in A$ such that for all i less than l, we have*

$$y([t_i, t_{i+1}]) \subseteq \mathcal{O}_{\alpha_i}$$

Proof. Define

$$T^* := \sup\{t : S \le t \le T,\ \text{the conclusion of the lemma holds for } [S,t]\}.$$

Note that trivially $T^* > S$. For sake of contradiction, suppose $T^* < T$. Then there exists an $\epsilon > 0$ such that $T^* + \epsilon < T$, $T^* - \epsilon > S$ and $y(T^* - \epsilon, T^* + \epsilon) \subset \mathcal{O}_{\alpha^*}$ for some α^*. But the condition of the theorem holds for $T^* - \epsilon$ for some partition P. By appending P with $T^* + \lambda\epsilon$ with $\lambda \in (-1, 1]$ we have that $T^* \geq T^* + \epsilon$ which is absurd. Thus, we must have that $T^* = T$. □

Definition 2.50. The set $\{a_i, b_i\}_{i=0}^l \subset [S, T]$ is an **interlaced cover of** $[S, T]$ if $S = a_0 < a_1 < b_0 < a_2 < b_1 < a_3 < b_2 < \cdots < a_l < b_{l-1} < b_l = T$. Let $y\colon [S, T] \to \mathcal{X}$. The set $\{a_i, b_i\}_{i=0}^l$ is an **interlaced cover for** y if $\{a_i, b_i\}_{i=0}^l$ is an interlaced cover of $[S, T]$ and $y(a_{i+1}) \neq y(b_i)$ for all i less than l.

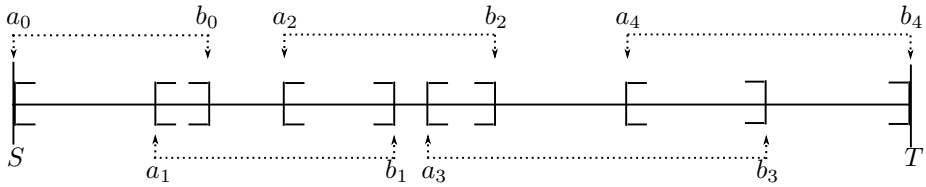

FIGURE 1. An interlaced cover of $[S, T]$.

Corollary 2.51. *Suppose $y\colon [S, T] \to \mathcal{X}$ is continuous and $y([S, T]) \subseteq \bigcup_{\alpha \in A} \mathcal{O}_\alpha$, where $\{\mathcal{O}_\alpha\}_{\alpha \in A}$ is a collection of open sets $\mathcal{O}_\alpha$. There exists an interlaced cover for y, $\{a_i, b_i\}_{i=0}^l$, such that $y([a_i, b_i]) \subseteq \mathcal{O}_{\alpha_i}$. Note that for such a setup, this implies $y([a_{i+1}, b_i]) \subseteq \mathcal{O}_{\alpha_i} \cap \mathcal{O}_{\alpha_{i+1}}$*

Proof. The first step will be a technical one to get rid of unnecessary endpoints. Let t'_i and α'_i be as given in Lemma 2.49. Then clearly $y(t'_i) \in \mathcal{O}_{\alpha'_{i-1}} \cap \mathcal{O}_{\alpha'_i}$ for all $1 \leq i < l'$. Starting with t'_1, we check if $y([t'_0, t'_1]) \subseteq \mathcal{O}_{\alpha_1}$. In the case it is, we may renumber our partition after removing t'_1 and $\mathcal{O}_{\alpha'_0}$ to get a new set of t'_j and α'_j which still satisfy the result of the lemma. Continuing this process inductively, we may assume that we have such a set $\{t_i, \alpha_i\}_{i=0}^l$ such that $y([t_i, t_{i+1}])$ is not contained in $\mathcal{O}_{\alpha_{i+1}}$.

To construct the desired interlaced cover, we define $b_i := t_{i+1}$ for all $i \leq l := l' - 1$ and $a_0 := S$. Note for now that this means $y([b_{i-1}, b_i]) \subseteq \mathcal{O}_{\alpha_i}$. Then we define the "lower end" stopping time T_i for all $i > 0$ by the formula

$$T_i := \inf \left\{ t < b_i : y([t, b_i]) \subseteq \mathcal{O}_{\alpha_{i+1}} \right\}.$$

By construction and because we refined our partition, $b_{i-1} \leq T_i < b_i$. It is clear that $y(T_i) \neq y(b_i)$ by the continuity of y. Thus, there exists a time T_i^* such that $T_i < T_i^*$ and $y(T_i^*) \neq y(b_i)$. Define

$$a_{i+1} := T_i^* \quad \text{for all } 0 < i < n.$$

Since $y([b_{i-1}, b_i]) \subseteq \mathcal{O}_{\alpha_i}$ and $a_i > b_{i-1}$, we have that $y([a_i, b_i]) \subseteq \mathcal{O}_{\alpha_i}$. □

Since the following patching trick will be used multiple times in later proofs, we will prove it here in more generality to avoid too much indexing notation later.

Lemma 2.52. *Let ω be a control and $\{a_i, b_i\}_{i=0}^{l}$ be an interlaced cover of $[S,T]$ such that $\omega(a_{i+1}, b_i) > 0$ for all $i < n$. Let $\theta > 0$ and $F\colon D \to [0,\infty)$ be a bounded function such that $D \subseteq \Delta_{[S,T]}$ and for each $1 \le i \le l$ there exists $C_i < \infty$ such that*

$$F(s,t) \le C_i\, \omega(s,t)^{\theta} \quad \text{for all } (s,t) \in \Delta_{[a_i,b_i]} \cap D.$$

Then there exists a $\tilde{C} < \infty$ such that

$$F(s,t) \le \tilde{C}\, \omega(s,t)^{\theta} \quad \forall\, (s,t) \in D. \tag{2.32}$$

Proof. Let

$$\begin{aligned} m &:= \min\{\omega(a_{i+1}, b_i)^{\theta} : 0 \le i < n\}, \\ C &:= \max\{C_i : 0 \le i \le n\}, \text{ and} \\ M &:= \sup\{F(s,t) : (s,t) \in D\} < \infty, \end{aligned}$$

and then define $\tilde{C} := \max\{M/m, C\}$. We claim that inequality (2.32) holds.

If there exists an i such that $s,t \in [a_i, b_i] \cap D$, then (2.32) holds trivially. Otherwise, let i^* be the largest i such that $s \in [a_i, b_i]$. Then $s < a_{i^*+1}$ and $t > b_{i^*}$. However this says that $[s,t] \supset [a_{i^*+1}, b_{i^*}]$ so that

$$F(s,t) \le M = \frac{M}{m}\, m \le \tilde{C}\, \omega\left(a_{i^*+1}, b_{i^*}\right)^{\theta} \le \tilde{C}\, \omega(s,t)^{\theta}. \qquad \square$$

2.5.2. Proof of Theorem 2.45. The recurring strategy here will be to localize appropriately to work in the $\mathbb{R}^d$ case so that we may apply Theorem 2.48. We must choose these localizations carefully so that we may patch the estimates together (with two different strategies) using the lemmas above. One method of patching is a bit more involved than the other; therefore we will present it more formally.

Remark 2.53 (Proof strategy). Let $y\colon [a,b] \to M$ be the first component of $(y, y^{\dagger})$ where $\left(y, y^{\dagger}\right)$ is either a (ψ, U)-controlled rough path or chart controlled rough path. Also suppose for each $m \in y([a,b])$, we are given an open neighborhood, $\mathcal{W}_m \subseteq M$, of m. By Corollary 2.51, there exists an interlaced cover for y, $\{a_i, b_i\}_{i=1}^{l}$ and $\{m_i\}_{i=1}^{l}$ such that $y\left([a_i, b_i]\right) \subseteq \mathcal{W}_{m_i}$ and $\omega\left(a_{i+1}, b_i\right) > 0$. Thus, if $F\colon D \to [0,\infty)$ is a bounded function such that $D \subseteq \Delta_{[a,b]}$, then in order to prove that

$$F(s,t) \le C\, \omega(s,t)^{\theta} \quad \forall\, (s,t) \in D, \tag{2.33}$$

it suffices to prove that for each $1 \le i \le l$, there exists $C_i < \infty$ such that

$$F(s,t) \le C_i\, \omega(s,t)^{\theta} \quad \text{for all } (s,t) \in \Delta_{[a_i,b_i]} \cap D.$$

Therefore in attempting to prove an assertion in the form of inequality (2.33), we may assume, without loss of generality, that $y\left([a,b]\right) \subseteq \mathcal{W}$ where the $\mathcal{W}$ will have nice properties dependent on our setting.

The proof of Theorem 2.45 will consist of two steps:

1. If gauge conditions of (2.22) and (2.23) hold for some $C > 0$ and $\delta > 0$, then the chart conditions of (2.24) and (2.25) hold. We will reduce this to the $\mathbb{R}^d$ case immediately, then use Lemma 2.6 to patch the estimates together.
2. If the chart condition of (2.24) and (2.25) hold, then gauge condition of (2.22) and (2.23) hold for an appropriately chosen δ. Here we will first show which local estimates we need to satisfy to use Remark 2.53 and then reduce to the $\mathbb{R}^d$ case.

In simple terms, step 1 is "localize then patch" and step 2 is "cut nicely, localize, then patch".

Proof of Theorem 2.45.

Step 1. Definition 2.35 $\implies$ Definition 2.40.

We will first assume that the gauge definition holds, i.e., that there exists a $\delta > 0$ and a $C_1 > 0$ such that

$$\left|\psi(y_s, y_t) - y_s^\dagger\, x_{s,t}\right|_g \leq C_1\, \omega(s,t)^{2/p} \tag{2.34}$$

and

$$|U(y_s, y_t)\, y_t^\dagger - y_s^\dagger|_g \leq C_1\, \omega(s,t)^{1/p}$$

hold for all $0 \leq s \leq t \leq T$ such that $|t - s| \leq \delta$. Let ϕ be a chart on M and let $[a,b]$ be such that $y([a,b]) \subseteq D(\phi)$. If we define

$$\begin{aligned}
\psi^\phi(x,y) &:= \phi_* \psi\left(\phi^{-1}(x), \phi^{-1}(y)\right), \\
U^\phi(x,y) &:= \phi_* U\left(\phi^{-1}(x), \phi^{-1}(y)\right) \circ \left(\phi_*^{-1}\right)_{\phi(y)}, \\
\left(z_s, z_s^\dagger\right) &:= \phi_*\left(\mathbf{y}_s\right) = \left(\phi(y_s), d\phi \circ y_s^\dagger\right),
\end{aligned}$$

then it is clear that there exists a $C_2 = C_2\left(\phi_*\right)$ such that

$$|\bar{\psi}^\phi(z_s, z_t) - z_s^\dagger\, x_{s,t}| \leq C_2\, \omega(s,t)^{2/p} \tag{2.35}$$

$$|\bar{U}^\phi(z_s, z_t) z_t^\dagger - z_s^\dagger| \leq C_2\, \omega(s,t)^{1/p} \tag{2.36}$$

for all $a \leq s \leq t \leq b$ such that $t - s \leq \delta$ where $\left(\psi^\phi, U^\phi\right)$ is a local gauge on $\mathbb{R}^d$ and $\left(\bar{\psi}^\phi, \bar{U}^\phi\right)$ is consistent with Notation 2.23. Thus $(z, z^\dagger)$ is a $\left(\psi^\phi, U^\phi\right)$-rough path controlled by $\mathbf{X}$. Finally we need to use this information to show there exists a $C_{\phi,a,b}$ such that

$$|z_t - z_s - z_s^\dagger\, x_{s,t}| \leq C_{\phi,a,b}\, \omega(s,t)^{2/p}. \tag{2.37}$$

and

$$|z_t^\dagger - z_s^\dagger| \leq C_{\phi,a,b}\, \omega(s,t)^{1/p} \tag{2.38}$$

for all s, t such that $a \leq s \leq t \leq b$.

In light of the sewing Lemma 2.6 and Lemma 2.49, we only need to show that for each $u \in [a,b]$, the inequalities (2.37) and (2.38) hold with $C_{\phi,a,b}$ replaced with C_u for all $s, t \in (u - \delta_u, u + \delta_u) \cap [a,b]$ such that $s \leq t$ for some $\delta_u > 0$.

For any $u \in [a,b]$, let $\mathcal{W}_u$ be an open convex set of z_u such that $\mathcal{W}_u \times \mathcal{W}_u \subseteq D(\psi^\phi)$. We then choose $\delta_u > 0$ to be such that $z\left([u-\delta_u, u+\delta_u] \cap [a,b]\right) \subseteq \mathcal{W}_u$ and $2\delta_u \leq \delta$. However, now we are in the setting of Theorem 2.48 and are therefore finished with this step.

Step 2. Definition 2.40 $\implies$ Definition 2.35.

Suppose that the chart item (2.24) holds. We must prove that there exists a $\delta, C > 0$ such that

$$\left|\psi(y_s, y_t) - y_s^\dagger\, x_{s,t}\right|_g \leq C\,\omega(s,t)^{2/p} \quad \text{and} \quad \left|U(y_s, y_t)\, y_t^\dagger - y_s^\dagger\right|_g \leq C\,\omega(s,t)^{1/p}$$

for all $s \leq t$ such that $|t-s| \leq \delta$.

We choose δ such that $|t-s| \leq \delta$ for $0 \leq s \leq t \leq T$ implies that both $|\psi(y_s, y_t)|_g$ and $|U(y_s, y_t)|_g$ make sense and are bounded. Around every point m of $y([0,T])$, there exists an open $\mathcal{O}_m$ containing m and such that $\mathcal{O}_m \times \mathcal{O}_m \subseteq \mathcal{D}$. Additionally there exists a chart ϕ^m such that $m \in D\left(\phi^m\right)$. By considering an open ball around $\phi^m(m)$ in $R\left(\phi^m\right)$ and shrinking the radius, we may assume that $\mathcal{V}_m := D\left(\phi^m\right) \subseteq \mathcal{O}_m$ and the range, $\mathcal{W}_m := \phi\left(\mathcal{V}_m\right)$, of ϕ^m is convex. Since $\{\mathcal{V}_m\}_{m \in y([0,T])}$ is an open cover of $y\left([0,T]\right)$, we may use this cover along with $D = \{(s,t) : 0 \leq s \leq t \leq T \text{ and } |t-s| \leq \delta\}$ to employ the proof strategy in Remark 2.53. We will do this twice, with $F(s,t) = |\psi(y_s, y_t) - y_s^\dagger\, x_{s,t}|_g$ in the first iteration and $F(s,t) = |U(y_s, y_t)\, y_t^\dagger - y_s^\dagger|_g$ in the second; this will reduce us to considering the case where there exists a single chart ϕ such that $y([0,T]) \subseteq D(\phi)$, $D(\phi) \times D(\phi) \subseteq \mathcal{D}$ and $\mathcal{W} = R(\phi)$ is convex.

Now that we have reduced to a single chart ϕ, we may define (ψ^ϕ, U^ϕ) and the path $(z, z^\dagger)$ as in Step 1. Then $z([0,T]) \subseteq \mathcal{W}$ and $\mathcal{W} \times \mathcal{W} \subseteq D(\psi^\phi) = D(U^\phi)$. However, by Theorem 2.48 we have that the proper estimates hold because $\mathbf{z}$ is a (ψ^ϕ, U^ϕ)-rough path controlled by $\mathbf{X}$. Therefore, we are finished by patching using Remark 2.53. □

In situations in which we are given a covariant derivative ∇ on a manifold, by Example 2.19, we have an equivalent definition:

Example 2.54. The pair $\left(y_s, y_s^\dagger\right)$ is an element of $CRP_{\mathbf{X}}(M)$ if and only if there exists a C such that

$$|(\exp_{y_s}^\nabla)^{-1}(y_t) - y_s^\dagger\, x_{s,t}|_g \leq C\,\omega(s,t)^{2/p} \tag{2.39}$$

and

$$|U^\nabla(y_s, y_t) y_t^\dagger - y_s^\dagger|_g \leq C\,\omega(s,t)^{1/p}, \tag{2.40}$$

where $(\exp_m^\nabla)^{-1}$ and $U^\nabla(n,m)$ are defined as in Example 2.19 and the inequalities hold when (y_s, y_t) are in the domain $\mathcal{D}$ as given in Theorem 2.32. In particular, on a Riemannian manifold we can use this definition with the Levi-Civita covariant derivative.

Before providing yet another equivalent definition of controlled rough paths on manifolds, we will present some examples.

2.6. Examples of controlled rough paths

Recall $\mathbf{X} = (x, \mathbb{X})$ is a weak-geometric rough path with values in $W \oplus W^{\otimes 2}$ where $W = \mathbb{R}^k$. The results here will rely on basic approximations found in the Appendix, Section 6.

Example 2.55. Let $M^d \subseteq W$ be an embedded submanifold and for every $m \in M^d$, let $P(m)$ be the orthogonal projection onto the tangent space $T_m M$. Suppose $x_s \in M^d$ for all s in $[0, T]$. Then $(x_s, P(x_s)) \in CRP_{\mathbf{X}}(M)$.

Proof. We will use the gauge as given in Example 2.54 where the ∇ is the Levi-Civita covariant derivative from the induced metric from Euclidean space. Verifying that $P(x_s)$ lives in the correct space is trivial.

Next, to show inequality 2.39 is satisfied, we use item (1) of Lemma 6.4, which says

$$\exp_m^{-1}(\tilde{m}) = P(m)(\tilde{m} - m) + O(|\tilde{m} - m|^3) \quad \text{for all } m \in M^d.$$

Letting $m = x_s$ and $\tilde{m} = x_t$, we are done.

Inequality (2.40) is also satisfied easily; the first equality in item (2) of Lemma 6.4 implies $U^{\nabla}(\tilde{m}, m) = P(m) + O(|\tilde{m} - m|)$. Thus

$$P(x_t) - U^{\nabla}(x_t, x_s) P(x_s) \underset{1}{\approx} P(x_t) - P(x_s)P(x_s) = P(x_t) - P(x_s) \underset{1}{\approx} 0. \qquad \square$$

The next example will be proved in more generality in Section 4.2. However, we find it instructive to prove it without charts and in the embedded context where the reader may be more comfortable.

Example 2.56. Let f be a smooth function from W to an embedded manifold $\tilde{M}^d \subseteq \mathbb{R}^{\tilde{k}}$. Then $(f(x_s), f'(x_s)) \in CRP_{\mathbf{X}}(\tilde{M})$.

Proof. Again we will use the Levi-Civita covariant derivative $\tilde{\nabla}$ from the embedded metric. First we note that $f'(x_s)$ lives in the correct space as $R(f) \subseteq \tilde{M}^d$.

To show inequality (2.39) holds one can use the fact that $(f((x_s), f'(x_s))$ is a controlled rough path in the embedded space or Taylor's theorem to see that

$$f(x_t) - f(x_s) - f'(x_s)(x_t - x_s) \underset{2}{\approx} 0$$

which easily implies

$$P(f(x_s))\left[f(x_t) - f(x_s) - f'(x_s)(x_t - x_s)\right] \underset{2}{\approx} 0.$$

But again by Lemma 6.4,

$$\begin{aligned} P(f(x_s))[f(x_t) &- f(x_s) - f'(x_s)(x_t - x_s)] \\ &= P(f(x_s))[f(x_t) - f(x_s)] - f'(x_s)(x_t - x_s) \\ &\underset{2}{\approx} (\exp^{\tilde{\nabla}}_{f(x_s)})^{-1}(f(x_t)) - f'(x_s)(x_t - x_s). \end{aligned}$$

Thus

$$\left(\exp^{\tilde{\nabla}}_{f(x_s)}\right)^{-1}(f(x_t)) - f'(x_s)(x_t - x_s) \underset{2}{\approx} 0.$$

Lastly, to show inequality (2.40), we have

$$f'(x_t) - f'(x_s) \underset{1}{\approx} 0$$

and therefore

$$\begin{aligned} 0 &\underset{1}{\approx} P(f(x_t))[f'(x_t) - f'(x_s)] = f'(x_t) - P(f(x_t))f'(x_s) \\ &\underset{1}{\approx} f'(x_t) - U^{\tilde{\nabla}}(f(x_t), f(x_s))f'(x_s), \end{aligned}$$

wherein we have used $P(f(x_t))f'(x_t) = f'(x_t)$ in the second line and Lemma 6.4 in the last. Thus $(f(x_s), f'(x_s)) \in CRP_{\mathbf{X}}(\tilde{M})$ □

2.7. Smooth function definition of CRP

In the spirit of semi-martingales on manifolds (see for example Chapter III of [13] or [18], [12], [19]), we can define controlled rough paths on manifolds as elements which, when composed with any smooth function, give rise to a one-dimensional controlled rough path on flat space. More precisely we have the following theorem.

Theorem 2.57. $\mathbf{y} = (y, y^\dagger) \in CRP_{\mathbf{X}}(M)$ *if and only if for every* $f \in C^\infty(M)$,

$$f_*\mathbf{y} = \left(f(y), df \circ y^\dagger\right) \in CRP_{\mathbf{X}}(\mathbb{R}).$$

Proof. The proof that $\mathbf{y} \in CRP_{\mathbf{X}}(M)$ implies that $f_*\mathbf{y} \in CRP_{\mathbf{X}}(\mathbb{R})$ for every $f \in C^\infty(M)$ will be deferred to the more general case proved in Proposition 4.10 (in which case we consider the codomain of f to be a manifold $\tilde{M}$).

To prove the converse, let ϕ be a chart and $0 \le a < b \le T$ be such that $y([a,b]) \subseteq D(\phi)$ and let $\mathcal{O} \subset M$ be an open set such that $\bar{\mathcal{O}}$ is compact and

$$y([a,b]) \subseteq \mathcal{O} \subseteq \bar{\mathcal{O}} \subseteq D(\phi).$$

Then by using a cutoff function we can manufacture global functions $f^i \in C^\infty(M)$ which agree with the coordinates ϕ^i on $\mathcal{O}$. The assumption that

$$f^i_*\mathbf{y} \in CRP_{\mathbf{X}}([a,b], \mathbb{R})$$

is a controlled rough path for $1 \le i \le d$ then shows the inequalities in (2.24) and (2.25) of Definition 2.40 hold. □

3. Integration of controlled one-forms

In the flat case, a controlled rough path with values in an appropriate Euclidean spaces can be integrated against another controlled rough path (see Theorem 2.9) provided their controlling rough path $\mathbf{X}$ is the same. The integral in this case is another rough path controlled by $\mathbf{X}$. We can do something similar on manifolds, though it will be necessary to add some extra structure. As usual let $\mathbf{y}_s = (y_s, y_s^\dagger)$ be a controlled rough path on M controlled by $\mathbf{X} = (x, \mathbb{X}) \in W \oplus W^{\otimes 2}$. Let V be a Banach space.

3.1. Controlled one-forms along a rough path

Let U be a parallelism on M.

Definition 3.1. The pair $(\alpha_s, \alpha_s^\dagger)$ is a V-valued U-controlled (rough) one-form along y_s if

(1) $\alpha_s \in L\left(T_{y_s}M, V\right)$,

(2) $\alpha_s^\dagger \in L\left(W \otimes T_{y_s}M, V\right)$,

(3) $\alpha_t \circ U(y_t, y_s) - \alpha_s - \alpha_s^\dagger\left(x_{s,t} \otimes (\cdot)\right) \underset{2}{\approx} 0$,

(4) $\alpha_t^\dagger \circ \left(I \otimes U\left(y_t, y_s\right)\right) - \alpha_s^\dagger \underset{1}{\approx} 0$.

By items (3) and (4), we mean these hold if $|t-s| < \delta$ for some $\delta > 0$ to ensure the expressions make sense.

Remark 3.2. For the sake of clarity, by item (3) of Definition 3.1, we mean that if s, t are close, then there exists a C such that

$$\left|\alpha_t \circ U(y_t, y_s) - \alpha_s - \alpha_s^\dagger\left(x_{s,t} \otimes (\cdot)\right)\right|_{g,op} \leq C\,\omega(s,t)^{2/p}.$$

For item (4), we mean for s, t close, there exists a C such that

$$\left|\alpha_t^\dagger \circ \left(w \otimes U(y_t, y_s)\right) - \alpha_s^\dagger\left(w \otimes (\cdot)\right)\right|_{g,op} \leq C\,|w|\,\omega(s,t)^{1/p}$$

for all $w \in W$. By Corollary 6.6, it does not matter which Riemannian metric g we choose here.

Notation 3.3. Let $CRP_y^U(M,V)$ denote those $\boldsymbol{\alpha}_s := (\alpha_s, \alpha_s^\dagger)$ satisfying Definition 3.1. We refer to $CRP_y^U(M,V)$ as a **space of U-controlled one-forms along $\mathbf{y}$**.

Remark 3.4. If $M = \mathbb{R}^d$ and $U = I$ and we identify $T_{y_s}M$ with $\mathbb{R}^d$ then Definition 3.1 reduces to the flat case definition of a $L\left(\mathbb{R}^d, V\right)$-valued rough path controlled by $\mathbf{X}$.

Remark 3.5. Note that (3) and (4) of Definition 3.1 force continuity of both α_s and $\alpha_s^\dagger$.

We can take linear combinations of elements of $CRP_y^U(M,V)$ to form other elements in $CRP_y^U(M,V)$. The following proposition, whose simple proof is left to the reader, shows how to construct more non-trivial examples of elements in $CRP_y^U(M,V)$.

Proposition 3.6. *If V and $\tilde{V}$ are Banach spaces, $\boldsymbol{\alpha} \in CRP_y^U(M,V)$ and*

$$\mathbf{f} = (f, f^\dagger) \in CRP_{\mathbf{X}}\big(\mathrm{Hom}(V, \tilde{V})\big),$$

then

$$(\mathbf{f}\boldsymbol{\alpha})_s := \left(f_s\alpha_s, f_s^\dagger\alpha_s + f_s\,\alpha_s^\dagger\right) \in CRP_y^U(M, \tilde{V}).$$

where by $f_s^\dagger\alpha_s$ we mean $f_s^\dagger\left((\cdot) \otimes \alpha_s(\cdot)\right)$.

Our next goal is to define an integral of $\boldsymbol{\alpha}_s$ along $\mathbf{y}_s$. However, this integral will depend on a choice of parallelism and for this reason we need to introduce the "compatibility tensor" which measures the difference between two parallelisms.

3.2. The compatibility tensors

Definition 3.7. The **compatibility tensor** $S^{\tilde{U},U} \in \Gamma\left(L\left(TM \otimes TM\right), TM\right)$ of two parallelisms $\tilde{U}$ and U on M is the defined by

$$S_m^{\tilde{U},U} := d[U(\cdot,m)^{-1}\tilde{U}(\cdot,m)]_m.$$

In more detail if $v_m, w_m \in T_mM$, then

$$S_m^{\tilde{U},U}[v_m \otimes w_m] = v_m[x \to U(x,m)^{-1}\tilde{U}(x,m)w_m].$$

Remark 3.8. There are actually multiple ways to define $S_m^{\tilde{U},U}$. For example, we have on simple tensors

$$\begin{aligned} S_m^{\tilde{U},U}(v_m \otimes w_m) &= d[U(m,\cdot)\tilde{U}(m,\cdot)^{-1}w_m]_m v_m \\ &= (\nabla_{v_m}[\tilde{U}(\cdot,m) - U(\cdot,m)])w_m = (\nabla_{v_m}[U(m,\cdot) - \tilde{U}(m,\cdot)])w_m, \end{aligned} \tag{3.1}$$

where ∇ is any covariant derivative on M (the last line can be interpreted as ∇_{v_m} acting on $[U(m,\cdot)-\tilde{U}(m,\cdot)]\mathbf{W}$, where $\mathbf{W}$ is any smooth section such that $\mathbf{W}(m) = w_m$). Similar to the proofs of Corollary 2.29 above and Theorem 3.15 below, the identities in (3.1) are straightforward to prove by employing charts to reduce them to Euclidean space identities.

Example 3.9. If ∇ and $\tilde{\nabla}$ are two covariant derivatives on TM, $U = U^{\nabla}$, $\tilde{U} = U^{\tilde{\nabla}}$, and $A \in \Omega^1\left(\mathrm{End}(TM)\right)$ such that $\nabla = \tilde{\nabla} + A$, then

$$S_m^{\tilde{U},U}\left(v_m \otimes w_m\right) = A(v_m)w_m \in T_mM.$$

Indeed,

$$\begin{aligned} v_m[U(\cdot,m)^{-1}\tilde{U}(\cdot,m)w_m] &= \nabla_{v_m}[\tilde{U}(\cdot,m)w_m] \\ &= \tilde{\nabla}_{v_m}[\tilde{U}(\cdot,m)w_m] + A(v_m)\tilde{U}(m,m)w_m = 0 + A(v_m)w_m = A(v_m)w_m. \end{aligned}$$

Example 3.10 (Converse of Example 3.9). If U and $\tilde{U}$ are two parallelisms on M and $\nabla = \nabla^U$ and $\tilde{\nabla} = \nabla^{\tilde{U}}$ are the corresponding covariant derivatives on TM (as in Remark 2.20), then

$$\nabla_{v_m} = \tilde{\nabla}_{v_m} + S_m^{\tilde{U},U}\left(v_m \otimes (\cdot)\right) \quad \forall\, v_m \in T_mM.$$

The verification is as follows. If Y is a vector-field on M and σ_t is such that $\dot{\sigma}_0 = v_m$, we have

$$\begin{aligned} \nabla_{v_m}Y - \tilde{\nabla}_{v_m}Y &:= \frac{d}{dt}\Big|_0 [U(m,\sigma_t) - \tilde{U}(m,\sigma_t)]Y(\sigma_t) \\ &= (\nabla_{v_m}[U(m,\cdot) - \tilde{U}(m,\cdot)])Y(m) + 0 \cdot \nabla_{v_m}Y = S_m^{\tilde{U},U}(v_m \otimes Y(m)), \end{aligned}$$

wherein we have used equation (3.1) for the last equality.

Lemma 3.11. *If U, $\tilde{U}$ and $\hat{U}$ are three parallelisms, then*

$$S^{\hat{U},U} = S^{\hat{U},\tilde{U}} + S^{\tilde{U},U} \quad \text{and} \quad S^{\tilde{U},U} = -S^{U,\tilde{U}}.$$

Proof. For $v_m, w_m \in T_mM$, an application of the product rules shows

$$\begin{aligned} S_m^{\hat{U},U}(v_m \otimes w_m) &= v_m \left[U(\cdot,m)^{-1}\hat{U}(\cdot,m)w_m\right] \\ &= v_m \left[\left[U(\cdot,m)^{-1}\tilde{U}(\cdot,m)\right]\left[\tilde{U}(\cdot,m)^{-1}\hat{U}(\cdot,m)\right]w_m\right] \\ &= S_m^{\hat{U},\tilde{U}}(v_m \otimes w_m) + S_m^{\tilde{U},U}(v_m \otimes w_m). \end{aligned}$$

The second statement follows from the first by letting $\hat{U} = U$. □

Notation 3.12. If $\mathcal{G} := (\psi, U)$ is a gauge, we let $S^{\mathcal{G}} := S^{\psi_*,U}$ be the compatibility tensor between U^{ψ} and U, where $U^{\psi}(m,n) := \psi(m,\cdot)_{*n}$ as in Remark 2.21.

If we have a covariant derivative ∇ on M, then as in Example 2.19 we have the choice of gauge $\mathcal{G} = (\psi, U) = \left((\exp^{\nabla})^{-1}, U^{\nabla}\right)$. In this case, the tensor $S_m^{\mathcal{G}}$ is a more familiar object.

Lemma 3.13. *If $\psi = (\exp^{\nabla})^{-1}$ and $U = U^{\nabla}$, then*

$$S_m^{\mathcal{G}} = \frac{1}{2}\, T_m^{\nabla},$$

where T^{∇} is the torsion tensor of ∇.

Proof. By transferring the covariant derivative and functions using charts, we may assume we are working on Euclidean space. In this case, by equation (6.15) and Corollary 6.8, we have

$$\begin{aligned} S_m^{\mathcal{G}}((m,v)\otimes(m,w)) &= \left(\nabla_{(m,v)}[U^{\nabla}(m,\cdot) - (\exp_m^{\nabla})_{*\cdot}^{-1}]\right) w \\ &= [\partial_{(m,v)} + A_m\langle v\rangle][U^{\nabla}(m,\cdot) - (\exp_m^{\nabla})_{*\cdot}^{-1}]\, w \\ &= (U^{\nabla}(m,\cdot))'(m)[v\otimes w] - \left((\exp_m^{\nabla})^{-1}\right)''(m)[v\otimes w] \\ &\quad + A_m\langle v\rangle\langle w\rangle - A_m\langle v\rangle\langle w\rangle \\ &= A_m\langle v\rangle\langle w\rangle - \frac{1}{2}A_m\langle v\rangle\langle w\rangle - \frac{1}{2}A_m\langle w\rangle\langle v\rangle \\ &= \frac{1}{2}\left[A_m\langle v\rangle\langle w\rangle - A_m\langle w\rangle\langle v\rangle\right] = \frac{1}{2}T_m^{\nabla}((m,v)\otimes(m,w)). \square \end{aligned}$$

Here is one last example of a gauge and its compatibility tensor.

Proposition 3.14. *Let G be a Lie group and ∇ be the left covariant derivative on TG uniquely determined by requiring the left invariant vector fields to be covariantly constant, i.e., $\nabla\tilde{A} = 0$ for all $A \in \mathfrak{g}$. Then for g near k,*

$$U^{\nabla}(g,k) = //(k \to g) = L_{gk^{-1}*}, \tag{3.2}$$

and

$$\psi^{\nabla}(k,g) = \left(\exp_k^{\nabla}\right)^{-1}(g) = k \cdot \log\left(k^{-1}g\right), \tag{3.3}$$

where $L_g : G \to G$ is left multiplication by $g \in G$ and log *is the local inverse of the map $A \to e^A$. Moreover the compatibility tensor for this gauge is given by*

$$S(\xi_g, \eta_g) = -\frac{1}{2} L_{g*}\left[\theta(\xi_g), \theta(\eta_g)\right] \quad \textit{for all } \xi_g, \eta_g \in T_gG, \tag{3.4}$$

where θ is the Maurer–Cartan form on G defined by $\theta(\xi) := L_{g^{-1}*}\xi \in \mathfrak{g} := T_eG$ for all $\xi \in T_gG$.

Proof. The torsion of ∇ is given by

$$T(\tilde{A}, \tilde{B}) = \nabla_{\tilde{A}}\tilde{B} - \nabla_{\tilde{B}}\tilde{A} - [\tilde{A}, \tilde{B}] = -\widetilde{[A, B]}$$

or equivalently as

$$T(\xi_g, \eta_g) = -L_{g*}[\theta(\xi_g), \theta(\eta_g)] \quad \text{for all } \xi_g, \eta_g \in T_gG.$$

Equation (3.4) follows from the above formula along with the result in Lemma 3.13.

If $\xi(t)$ is a path TG above $\sigma(t) \in G$ it may be written as $\xi(t) = L_{\sigma(t)*}\theta(\xi(t))$. Since $L_{\sigma(t)*}$ is parallel translation, it follows that

$$\frac{\nabla\xi(t)}{dt} = L_{\sigma(t)*}\frac{d}{dt}\theta(\xi(t)).$$

Thus $\xi(t) \in TG$ is parallel if and only if $\theta(\xi(t))$ is constant for all t. If σ is a general curve in G, we may conclude

$$//\left(\sigma|_{[s,t]}\right) = L_{\sigma(t)*}L_{\sigma(s)^{-1}*} = L_{\sigma(t)\sigma(s)^{-1}*}$$

and therefore U^{∇} is given as in equation (3.2).

By definition, a curve $\sigma(t) \in G$ is a geodesic if and only if $\dot{\sigma}(t)$ is parallel, i.e., if and only if $\theta(\dot{\sigma}(t)) = A$ for some $A \in \mathfrak{g}$. That is $\dot{\sigma}(t) = \tilde{A}(\sigma(t))$ with $\sigma(0) = k \in G$. The solution to this equation is $\sigma(t) = ke^{tA}$ and hence we have shown that $\exp_k^{\nabla}(k \cdot A) = ke^A$. So setting $g = ke^A$ and solving for A gives $A = \log(k^{-1}g)$, and the formula for ψ^{∇} in equation (3.3) now follows. □

The last three results of this subsection show how the compatibility tensor allows us to compare two different parallelisms and two different logarithms on M.

Theorem 3.15. *Suppose that U and $\tilde{U}$ are two parallelisms on M and ψ is a logarithm on M, then*

$$U(m,n)\,\tilde{U}(m,n)^{-1} =_2 I + S_m^{\tilde{U},U}\left(\psi(m,n) \otimes (\cdot)\right). \tag{3.5}$$

Proof. By using charts it suffices to prove the theorem when $M = \mathbb{R}^d$. By Taylor's theorem (see Theorem 2.24),

$$\begin{aligned} U(m,n) &=_2 I + [(D_2U)(m,m)(n-m)] \quad \text{and} \\ \tilde{U}(m,n) &=_2 I + [(D_2\tilde{U})(m,m)(n-m)], \end{aligned}$$

and therefore

$$\begin{aligned} &U(m,n)\,\tilde{U}(m,n)^{-1} \\ &\quad =_2 (I + [(D_2U)(m,m)(n-m)])\,(I - [(D_2\tilde{U})(m,m)(n-m)]) \\ &\quad =_2 I + [((D_2U)(m,m) - (D_2\tilde{U})(m,m))(n-m)]. \end{aligned} \tag{3.6}$$

However, by equation (3.1) we have

$$S_m^{\tilde{U},U} = (D_2U)(m,m) - (D_2\tilde{U})(m,m). \tag{3.7}$$

Using this identity back in (3.6) shows

$$U(m,n)\,\tilde{U}(m,n)^{-1} =_2 I + S_m^{\tilde{U},U}\left([n-m]_m \otimes (\cdot)\right)$$

from which (3.5) follows because $\psi(m,n) =_2 [n-m]_m$. □

Corollary 3.16. *If $\mathcal{G} = (\psi, U)$ is a gauge on M, then*

$$\psi\left(n,\cdot\right)_{*m} =_2 U(n,m)\left[I + S_m^{\mathcal{G}}\left(\psi(m,n)\otimes(\cdot)\right)\right]. \tag{3.8}$$

In particular,

$$\psi\left(y_t,\cdot\right)_{*y_s} \underset{2}{\approx} U\left(y_t,y_s\right)\left[I + S_{y_s}^{\mathcal{G}}\left(\psi(y_s,y_t)\otimes(\cdot)\right)\right]. \tag{3.9}$$

Proof. Theorem 3.15 implies

$$U(m,n)\,\psi\left(m,\cdot\right)_{*n}^{-1} =_2 I + S_m^{\mathcal{G}}\left(\psi(m,n)\otimes(\cdot)\right),$$

while Corollary 2.29 shows

$$U(m,n)^{-1} =_2 U(n,m) \quad \text{and} \quad \psi\left(m,\cdot\right)_{*n}^{-1} =_2 \psi\left(n,\cdot\right)_{*m}.$$

Equation (3.8) now easily follows from the last two displayed equations. The second statement follows by patching. □

Lastly we may use the compatibility tensor to compare two logarithms.

Proposition 3.17. *Suppose that ψ and $\tilde{\psi}$ are two logarithms on a manifold M. Then the compatibility tensor, $S^{\psi_*,\tilde{\psi}_*}$ is symmetric and*

$$\psi(m,n) - \tilde{\psi}(m,n) =_3 \frac{1}{2}\,S_m^{\tilde{\psi}_*,\psi_*,}\left(\psi(m,n)\otimes\psi(m,n)\right). \tag{3.10}$$

Proof. As usual it suffices to prove this result when $M = \mathbb{R}^d$ in which case we omit the base points of tangent vectors. From (3.7) with $U(x,y) = \psi'_x(y)$ and $\tilde{U}(x,y) = \tilde{\psi}'_x(y)$, we see that

$$S_x^{\tilde{\psi}_*,\psi_*} = \psi''_x(x) - \tilde{\psi}''_x(x) \tag{3.11}$$

which is symmetric since mixed partial derivatives commute. Then by Taylor's theorem and (3.11),

$$\begin{aligned}\psi(x,y) - \tilde{\psi}(x,y) &= \frac{1}{2}\,[\psi''_x(x) - \tilde{\psi}''_x(x)](y-x)^{\otimes 2} + O(|y-x|^3)\\ &= \frac{1}{2}\,S_x^{\tilde{\psi}_*,\psi_*}(\psi(x,y)^{\otimes 2}) + O(|y-x|^3),\end{aligned}$$

wherein we have also used $(y-x)^{\otimes 2} =_3 \psi(x,y)^{\otimes 2}$. □

Remark 3.18. If ∇ is any covariant derivative on TM, then

$$S_m^{\tilde{\psi}_*,\psi_*} = \left[\nabla d(\psi(m,\cdot) - \tilde{\psi}(m,\cdot))\right]_m = \mathrm{Hess}_m^{\nabla}(\psi_m - \tilde{\psi}_m),$$

where $\mathrm{Hess}_m^{\nabla} f := [\nabla df]_m$. By choosing ∇ to be torsion free we again see that $S_m^{\tilde{\psi}_*,\psi_*}$ is a symmetric tensor.

3.3. *U*-controlled rough integration

Our next goal is to construct "the" integral, $\int \langle \boldsymbol{\alpha}, d\mathbf{y}\rangle$, where $\mathbf{y} \in CRP_{\mathbf{X}}(M)$ and $\boldsymbol{\alpha} \in CRP_y^U(M,V)$. We begin with the following proposition in the smooth category which is meant to motivate the definitions to come.

Proposition 3.19. *Assume (in this proposition only) that all functions,* $\mathbf{y}_s$, $\boldsymbol{\alpha}_s$, *and* x_s *are smooth,* $p = 1$, *and* $\omega(s,t) = |t-s|$. *Further assume* $\mathbf{y}$ *(respectively* $\boldsymbol{\alpha}$*) still satisfy the estimates of being controlled rough path (along* $\mathbf{y}$*). Then*

$$\int_s^t \alpha_\tau\, \dot{y}_\tau\, d\tau = \alpha_s[\psi(y_s,y_t) + S_{y_s}^{\mathcal{G}}(y_s^\dagger \otimes y_s^\dagger\, \mathbb{X}_{s,t})] + \alpha_s^\dagger (I \otimes y_s^\dagger)\, \mathbb{X}_{s,t} + O((t-s)^3). \tag{3.12}$$

Proof. Our assumptions give

$$\begin{aligned}&\psi(y_s,y_t) = y_s^\dagger\, x_{s,t} + O((t-s)^2) \implies \dot{y}_s = y_s^\dagger\, \dot{x}_s,\\ &\alpha_t\, U(y_t,y_s) = \alpha_s + \alpha_s^\dagger x_{s,t} + O((t-s)^2),\\ &U(y_s,y_t)\, y_t^\dagger = y_s^\dagger + O(t-s), \text{ and}\\ &\alpha_t^\dagger (I \otimes U(y_t,y_s)) = \alpha_s^\dagger + O(t-s).\end{aligned}$$

We start with the identity

$$\begin{aligned}
\int_s^t \alpha_\tau \dot{y}_\tau \, d\tau &= \int_s^t \alpha_\tau \, U(y_\tau, y_s) \, U(y_\tau, y_s)^{-1} \, \dot{y}_\tau \, d\tau \\
&= \int_s^t [\alpha_s + \alpha_s^\dagger \, x_{s,\tau} + O((\tau - s)^2)] \, U(y_\tau, y_s)^{-1} \, \dot{y}_\tau \, d\tau \\
&= \int_s^t \alpha_s \, U(y_\tau, y_s)^{-1} \, \dot{y}_\tau d\tau + \int_s^t \alpha_s^\dagger \, x_{s,\tau} \, U(y_\tau, y_s)^{-1} \, \dot{y}_\tau \, d\tau + O((t-s)^3) \\
&= \int_s^t \alpha_s \, U(y_s, y_\tau) \, \dot{y}_\tau \, d\tau + \int_s^t \alpha_s^\dagger \, x_{s,\tau} \, U(y_s, y_\tau) \, \dot{y}_\tau \, d\tau + O((t-s)^3) \\
&=: A + B + O((t-s)^3),
\end{aligned}$$

wherein we have used Corollary 2.29 in order to show it is permissible to replace $U(y_\tau, y_s)^{-1}$ by $U(y_s, y_\tau)$ above. The B term is then easily estimated as

$$\begin{aligned}
B &= \int_s^t \alpha_s^\dagger \, x_{s,\tau} \, U(y_s, y_\tau) \, \dot{y}_\tau \, d\tau = \int_s^t \alpha_s^\dagger \, x_{s,\tau} \, U(y_s, y_\tau) y_\tau^\dagger \, \dot{x}_\tau \, d\tau \\
&= \int_s^t \alpha_s^\dagger \, x_{s,\tau} \, y_s^\dagger \, \dot{x}_\tau \, d\tau + O((t-s)^3) = \alpha_s^\dagger (I \otimes y_s^\dagger) \, \mathbb{X}_{s,t} + O((t-s)^3).
\end{aligned}$$

The estimate of the A term to order $O((t-s)^3)$ requires more care. For this term we use

$$\frac{d}{dt}\psi(y_s, y_t) = \psi(y_s, \cdot)_{*y_t} \dot{y}_t \implies \dot{y}_t = \psi(y_s, \cdot)_{*y_t}^{-1} \frac{d}{dt}\psi(y_s, y_t)$$

and (from Theorem 3.15) that

$$U\left(y_s, y_\tau\right) \psi(y_s, \cdot)_{*y_\tau}^{-1} =_2 I + S_{y_s}^{\mathcal{G}}\left(\psi\left(y_s, y_\tau\right) \otimes (\cdot)\right)$$

in order to conclude

$$\begin{aligned}
A &:= \int_s^t \alpha_s \, U(y_s, y_\tau) \, \dot{y}_\tau \, d\tau = \int_s^t \alpha_s \, U(y_s, y_\tau) \, \psi(y_s, \cdot)_{*y_\tau}^{-1} \, \frac{d}{d\tau}\psi(y_s, y_\tau) \, d\tau \\
&= \int_s^t \alpha_s \, [I + S_{y_s}^{\mathcal{G}}(\psi(y_s, y_\tau) \otimes (\cdot))] \, \frac{d}{d\tau}\psi(y_s, y_\tau) \, d\tau + O(|t-s|^3) \\
&= \alpha_s(\psi(y_s, y_t)) + \alpha_s \int_s^t S_{y_s}^{\mathcal{G}}\left(\psi(y_s, y_\tau) \otimes \frac{d}{d\tau}\psi(y_s, y_\tau)\right) d\tau + O(|t-s|^3) \\
&= \alpha_s(\psi(y_s, y_t)) + \alpha_s \int_s^t S_{y_s}^{\mathcal{G}}(y_s^\dagger \, x_{s,\tau} \otimes y_s^\dagger \, \dot{x}_\tau) \, d\tau + O(|t-s|^3) \\
&= \alpha_s(\psi(y_s, y_t)) + \alpha_s \, S_{y_s}^{\mathcal{G}}(y_s^\dagger \otimes y_s^\dagger \, \mathbb{X}_{s,t}) + O(|t-s|^3).
\end{aligned}$$

Putting this all together proves (3.12). □

The following definition is motivated by the right-hand side of equation (3.12).

Definition 3.20 ($(\mathcal{G},\mathbf{y})$-integrator). Given a gauge $\mathcal{G} := (\psi, U)$ and a path $\mathbf{y} \in CRP_{\mathbf{X}}(M)$, the $(\mathcal{G},\mathbf{y})$**-integrator** is the increment process

$$\mathbf{y}^{\mathcal{G}}_{s,t} := \left(\psi(y_s,y_t) + S^{\mathcal{G}}_{y_s}\left(y_s^{\dagger\otimes 2}\,\mathbb{X}_{s,t}\right), \left(I\otimes y_s^{\dagger}\right)\mathbb{X}_{s,t}\right) \in T_{y_s}M \times [W\otimes T_{y_s}M]\,.$$

Moreover, for $\boldsymbol{\alpha} \in CRP^U_y(M,V)$ (see Notation 3.3), let

$$\tilde{z}_{s,t} := \left\langle \boldsymbol{\alpha}_s, \mathbf{y}^{\mathcal{G}}_{s,t}\right\rangle = \alpha_s\left(\psi\left(y_s,y_t\right) + S^{\mathcal{G}}_{y_s}\left(y_s^{\dagger\otimes 2}\mathbb{X}_{s,t}\right)\right) + \alpha_s^{\dagger}\left(I\otimes y_s^{\dagger}\right)\mathbb{X}_{s,t}, \tag{3.13}$$

which is defined for $(s,t)\in\Delta_{[0,T]}$ with $|t-s|<\delta$ for some sufficiently small $\delta>0$.

Recall that a two-parameter function $F\colon \Delta_{[0,T]} \to V$ is an almost additive functional if there exists a $\theta>1$, a control $\tilde{\omega}(s,t)$ and a $C>0$ such that

$$|F_{s,u} - F_{s,t} - F_{t,u}| \le C\,\tilde{\omega}(s,t)^{\theta} \tag{3.14}$$

for all $0\le s\le t\le u\le T$.

Theorem 3.21. *Let $\mathcal{G} := (\psi,U)$ be a gauge, $\boldsymbol{\alpha}\in CRP^U_y(M,V)$, and $\tilde{z}_{s,t}$ be as in Definition 3.20. Then there exists a unique $\mathbf{z} = (z,z^{\dagger}) \in CRP_{\mathbf{X}}(V)$ such that $z_0=0$, $z_{s,t}\underset{3}{\approx}\tilde{z}_{s,t}$, and $z_s^{\dagger} = \alpha_s\circ y_s^{\dagger}$. We denote this unique controlled rough path by $\int\langle\boldsymbol{\alpha},d\mathbf{y}^{\mathcal{G}}\rangle$, i.e.,*

$$\int_s^t \left\langle \boldsymbol{\alpha},d\mathbf{y}^{\mathcal{G}}\right\rangle := \left[\int\left\langle\boldsymbol{\alpha},d\mathbf{y}^{\mathcal{G}}\right\rangle\right]^1_{s,t} \underset{3}{\approx} \left\langle \boldsymbol{\alpha}_s, \mathbf{y}^{\mathcal{G}}_{s,t}\right\rangle \quad \textit{and} \quad \left[\int\left\langle\boldsymbol{\alpha},d\mathbf{y}^{\mathcal{G}}\right\rangle\right]^{\dagger}_s = \alpha_s\circ y_s^{\dagger}.$$

Proof. By Theorem 3.25 below, $\tilde{z}_{s,t} := \left\langle \boldsymbol{\alpha}_s, \mathbf{y}^{\mathcal{G}}_{s,t}\right\rangle$ is an almost additive functional and therefore by Theorem 3.3.1 of [26] there exists a unique additive functional $z_{s,t}$ such that $z_{s,t}\underset{3}{\approx}\tilde{z}_{s,t}$. Moreover,

$$z_{s,t}\underset{3}{\approx}\tilde{z}_{s,t}\underset{2}{\approx}\alpha_s\left(\psi(y_s,y_t)\right)\underset{2}{\approx}\alpha_s\left(y_s^{\dagger}\,x_{s,t}\right),$$

which shows that $\mathbf{z}_s := \left(z_s, \alpha_s\circ y_s^{\dagger}\right)$ is indeed a controlled rough path with values in V. □

Example 3.22. In the case that $U = U^{\psi}$, so that

$$\alpha_t\circ\left(\psi_{y_t}\right)_{*y_s} - \alpha_s - \alpha_s^{\dagger}\left(x_{s,t}\otimes(\cdot)\right)\underset{2}{\approx}0,$$

we have that $\mathbf{y}^{\mathcal{G}}_{s,t} := \left(\psi\left(y_s,y_t\right), \left(I\otimes y_s^{\dagger}\right)\mathbb{X}_{s,t}\right)$ and so

$$\int_s^t\left\langle\boldsymbol{\alpha},d\mathbf{y}^{\mathcal{G}^{\psi}}\right\rangle \underset{3}{\approx}\alpha_s\left(\psi(y_s,y_t)\right) + \alpha_s^{\dagger}\left(I\otimes y_s^{\dagger}\right)\mathbb{X}_{s,t}.$$

Example 3.23. If $\mathcal{G}^{\nabla} = \left((\exp^{\nabla})^{-1}, U^{\nabla}\right)$, then by Lemma 3.13, we have that

$$\int_s^t\left\langle\boldsymbol{\alpha},d\mathbf{y}^{\mathcal{G}^{\nabla}}\right\rangle \underset{3}{\approx}\alpha_s\left(\exp^{-1}_{y_s}(y_t)\right) + \alpha_s^{\dagger}\left(I\otimes y_s^{\dagger}\right)\mathbb{X}_{s,t} + \alpha_s\Big(\frac{1}{2}\,T^{\nabla}_{y_s}\circ y_s^{\dagger\otimes 2}\,\mathbb{X}_{s,t}\Big).$$

Remark 3.24. The $(\mathcal{G}, \mathbf{y})$-integrator $\mathbf{y}^{\mathcal{G}}_{s,t}$ is helpful in easing notation so that the integral is simply written $\int_s^t \langle \boldsymbol{\alpha}, d\mathbf{y}^{\mathcal{G}} \rangle$. A more honest notation for this integral would be

$$\int_s^t \left\langle \left(\alpha, \alpha^\dagger\right), d\left(y^\psi, \mathbb{X}\right) \right\rangle_{\mathcal{S}^{\mathcal{G}}_{y^\dagger}},$$

where $\mathcal{S}^{\mathcal{G}}_{y^\dagger}(s)$ is the block matrix defined by

$$\mathcal{S}^{\mathcal{G}}_{y^\dagger}(s) := \begin{pmatrix} I & S^{\mathcal{G}}_{y_s} \circ \left(y^\dagger_s\right)^{\otimes 2} \\ 0 & I \otimes y^\dagger_s \end{pmatrix}$$

and $\langle \cdot, \cdot \rangle_{\mathcal{S}^{\mathcal{G}}_{y^\dagger}}$ is the "inner product" given by the matrix $\mathcal{S}^{\mathcal{G}}_{y^\dagger}$. When s is close to t, we have

$$\begin{aligned} \int_s^t \left\langle \left(\alpha, \alpha^\dagger\right), d\left(y^\psi, \mathbb{X}\right) \right\rangle_{\mathcal{S}^{\mathcal{G}}_{y^\dagger}} &\approx \left(\alpha_s, \alpha^\dagger_s\right) \begin{pmatrix} I & S^{\mathcal{G}}_{y_s} \circ \left(y^\dagger_s\right)^{\otimes 2} \\ 0 & I \otimes y^\dagger_s \end{pmatrix} \begin{pmatrix} y^\psi_{s,t} \\ \mathbb{X}_{s,t} \end{pmatrix} \\ &= \alpha_s \left(\psi(y_s, y_t) + S^{\mathcal{G}}_{y_s}\left(y^{\dagger \otimes 2}_s \mathbb{X}_{s,t}\right)\right) + \alpha^\dagger_s \left(I \otimes y^\dagger_s\right) \mathbb{X}_{s,t}. \end{aligned}$$

3.4. Almost additivity result

The following theorem was the key ingredient in the proof of Theorem 3.21 on the existence of rough path integration in the manifold setting.

Theorem 3.25 (Almost additivity). *Let $\mathcal{G} := (\psi, U)$ be any gauge. If $\boldsymbol{\alpha} \in CRP^U_y(M, V)$, then $\tilde{z}_{s,t} \in V$ defined as in Definition 3.20 is an almost additive functional.*

The proof of Theorem 3.25 will be given after Corollary 3.28 which states that logarithms are "almost additive". We first need a couple of lemmas. Recall from Definition 2.15 that $\psi_x = \psi(x, \cdot)$.

Lemma 3.26. *If U and $\tilde{U}$ are two parallelisms on M, then*

$$S^{\tilde{U},U}_{y_t} \circ U(y_t, y_s)^{\otimes 2} \underset{1}{\approx} U(y_t, y_s) \circ S^{\tilde{U},U}_{y_s}.$$

Proof. By the usual patching arguments it suffices to prove this lemma for $M = \mathbb{R}^d$. In the Euclidean space setting the identity is trivial to prove since $U(n, m) =_1 I$ and $S^{\tilde{U},U}_n =_1 S^{\tilde{U},U}_m$. □

Lemma 3.27. *Let K be a compact, convex set in $\mathbb{R}^d$. If ψ is a logarithm with domain $\mathcal{D}$ and $K \times K \subseteq \mathcal{D}$, then there exists a C_K such that, for all $x, y, z \in K$,*

$$|\psi'_y(x)\psi(x, y) + \psi(y, z) - \psi'_y(x)\psi(x, z)| \leq C_K \max\{|\psi(x, y)|, |\psi(y, z)|, |\psi(x, z)|\}^3.$$

Proof. We will use the notation $|x, y, z| := \max\{|y - x|, |z - y|, |z - x|\}$ and write $f(x, y, z) =_k g(x, y, z)$ if and only if $f(x, y, z) = g(x, y, z) + O(|x, y, z|^k)$. Since ψ

is zero on the diagonal and $\psi'_y(y) = \mathrm{id}$ for all y, it follow from Taylor's theorem (or see Theorem 2.24) that

$$\psi'_y(x) =_2 \mathrm{id} + \psi''_y(y)\,(x-y) \quad \text{and}$$

$$\psi(x,y) =_3 (y-x) + \frac{1}{2}\,\psi''_x(x)(y-x)^{\otimes 2} =_3 (y-x) + \frac{1}{2}\,\psi''_y(y)(y-x)^{\otimes 2}. \tag{3.15}$$

From these approximations we learn

$$\psi(x,y) - \psi(x,z) =_3 y - z + \frac{1}{2}\,\psi''_y(y)\big[(y-x)^{\otimes 2} - (z-x)^{\otimes 2}\big]$$

and

$$\begin{aligned}
&\psi'_y(x)\psi(x,y) - \psi'_y(x)\psi(x,z)\\
&\quad =_3 \big[\mathrm{id} + \psi''_y(y)\,(x-y)\otimes(\cdot)\big]\,(\psi(x,y)-\psi(x,z))\\
&\quad =_3 y - z + \frac{1}{2}\,\psi''_y(y)\big[(y-x)^{\otimes 2} - (z-x)^{\otimes 2}\big] + \psi''_y(y)\,[(x-y)\otimes(y-z)]\,.
\end{aligned}$$

As simple calculation now shows, with $a = y - x$ and $b = y - z$, that

$$\frac{1}{2}\,\big[(y-x)^{\otimes 2} - (z-x)^{\otimes 2}\big] + (x-y)\otimes(y-z) = -\frac{1}{2}\,\big[b^{\otimes 2} + b\otimes a - a\otimes b\big]\,.$$

Since $\psi''_y(y)a\otimes b = \psi''_y(y)b\otimes a$ (mixed partial derivatives commute), the last two displayed equations give

$$\begin{aligned}
\psi'_y(x)\psi(x,y) - \psi'_y(x)\psi(x,z) &=_3 y - z - \frac{1}{2}\,\psi''_y(y)b^{\otimes 2}\\
&= -\big[(z-y) + \frac{1}{2}\,\psi''_y(y)(z-y)^{\otimes 2}\big] =_3 -\psi\,(y,z)\,.
\end{aligned}$$

The bounds derived above are uniform over a compact set K. Due to (3.15), we may replace $O\big(|x,y,z|^3\big)$ with $O\big(\max\{|\psi(x,y)|, |\psi(y,z)|, |\psi(x,z)|\}^3\big)$. □

Corollary 3.28. *If $\big(y_s, y_s^\dagger\big)$ is a controlled rough path and ψ is a logarithm, there exists C_ψ, $\delta_\psi > 0$ such that if $0 \le s \le t \le u \le T$ and $u - s \le \delta_\psi$, then*

$$\big|\psi(y_t, y_u) - \psi\,(y_t, \cdot)_{*y_s}\,[\psi\,(y_s, y_u) - \psi\,(y_s, y_t)]\,\big|_g \le C_\psi\,\omega(s,u)^{3/p}.$$

Proof. Around every point in $y([0,T])$, using our usual techniques, we can find a neighborhood $\mathcal{W}$ such that $\mathcal{W}\times\mathcal{W} \subseteq \mathcal{D}$ and maps to a convex open set by a chart. We can then use Remark 2.53 with a slightly modified version (which includes three variables instead of two) of Lemma 2.52 to create a global estimate. We can then choose a δ such that $u - s \le \delta$ forces the path to lie within one of these sets $\mathcal{W}$. Therefore, it suffices to prove the estimate locally. However, we can push forward the metric and ψ to a convex set on Euclidean space. The rest follows from the Lemma 3.27 and the fact that $|\psi(y_s, y_t)| \le C\,\omega(s,t)^{1/p}$ for all $|t-s| \le \delta$ for some $C < \infty$ and $\delta > 0$. □

3.5. Proof of Theorem 3.25

Proof of Theorem 3.25. Let $0 \le s \le t \le u \le T$. Throughout this proof, we will use the notation $\underset{i}{\approx}$ with respect to the times s and u. To prove the statement, we need to show $\tilde{z}_{s,t} + \tilde{z}_{t,u} \underset{3}{\approx} \tilde{z}_{s,u}$. We begin by working on the three terms for $\tilde{z}_{t,u}$ in the following equation:

$$\tilde{z}_{t,u} = \alpha_t\left(\psi(y_t, y_u)\right) + \alpha_t^\dagger\big(I \otimes y_t^\dagger\big)\mathbb{X}_{t,u} + \alpha_t\big(S^{\mathcal{G}}_{y_t} \circ y_t^{\dagger\otimes 2}\,\mathbb{X}_{t,u}\big). \tag{3.16}$$

Using Corollary 3.28 followed by Corollary 3.16 we find

$$\begin{aligned}
&\alpha_t\left(\psi(y_t, y_u)\right)\\
&\quad\underset{3}{\approx} \alpha_t\psi\left(y_t, \cdot\right)_{*y_s}\left[\psi\left(y_s, y_u\right) - \psi(y_s, y_t)\right]\\
&\quad\underset{3}{\approx} \alpha_t\, U(y_t, y_s)\left[I + S^{\mathcal{G}}_{y_s}\left(\psi(y_s, y_t) \otimes (\cdot)\right)\right]\left[\psi\left(y_s, y_u\right) - \psi(y_s, y_t)\right]\\
&\quad\underset{3}{\approx} \left[\alpha_s + \alpha_s^\dagger\, x_{s,t} \otimes (\cdot)\right]\left[I + S^{\mathcal{G}}_{y_s}\left(\psi(y_s, y_t) \otimes (\cdot)\right)\right]\left[\psi\left(y_s, y_u\right) - \psi(y_s, y_t)\right]\\
&\quad\underset{3}{\approx} \alpha_s\left[I + S^{\mathcal{G}}_{y_s}\left(\psi(y_s, y_t) \otimes (\cdot)\right)\right]\left[\psi\left(y_s, y_u\right) - \psi(y_s, y_t)\right]\\
&\qquad + \alpha_s^\dagger\, x_{s,t} \otimes \left[\psi\left(y_s, y_u\right) - \psi(y_s, y_t)\right].
\end{aligned}$$

Combining this equation with the estimates

$$\psi(y_s, y_t) \underset{2}{\approx} y_s^\dagger x_{s,t} \quad \text{and} \quad \psi\left(y_s, y_u\right) - \psi\left(y_s, y_t\right) \underset{2}{\approx} y_s^\dagger\left[x_{s,u} - x_{s,t}\right] = y_s^\dagger\, x_{t,u}$$

then shows

$$\begin{aligned}
\alpha_t\left(\psi(y_t, y_u)\right) &\underset{3}{\approx} \alpha_s\left[\psi\left(y_s, y_u\right) - \psi(y_s, y_t)\right]\\
&\quad + \alpha_s S^{\mathcal{G}}_{y_s}\left(y_s^\dagger\right)^{\otimes 2} x_{s,t} \otimes x_{t,u} + \alpha_s^\dagger\left(I \otimes y_s^\dagger\right) x_{s,t} \otimes x_{t,u}.
\end{aligned} \tag{3.17}$$

By the definitions of $CRP_{\mathbf{X}}(M)$ and $CRP_y^U(M, V)$ we have

$$\begin{aligned}
\alpha_t^\dagger(I \otimes y_t^\dagger)\mathbb{X}_{t,u} &\underset{3}{\approx} \alpha_t^\dagger\left(I \otimes U\left(y_t, y_s\right) y_s^\dagger\right)\mathbb{X}_{t,u}\\
&= \alpha_t^\dagger\left(I \otimes U(y_t, y_s)\right)\left(I \otimes y_s^\dagger\right)\mathbb{X}_{t,u} \underset{3}{\approx} \alpha_s^\dagger\left(I \otimes y_s^\dagger\right)\mathbb{X}_{t,u}.
\end{aligned} \tag{3.18}$$

Lastly, by the definitions of $CRP_{\mathbf{X}}(M)$ and $CRP_y^U(M, V)$ along with Lemma 3.26 with $\tilde{U}(m, n) = (\psi_m)_{*n}$, we have

$$\begin{aligned}
\alpha_t\big(S^{\mathcal{G}}_{y_t} \circ y_t^{\dagger\otimes 2}\,\mathbb{X}_{t,u}\big) &\underset{3}{\approx} \alpha_t\left(S^{\mathcal{G}}_{y_t} \circ U(y_t, y_s)^{\otimes 2} \circ y_s^{\dagger\otimes 2}\,\mathbb{X}_{t,u}\right)\\
&\underset{3}{\approx} \alpha_t\left(U(y_t, y_s) \circ S^{\mathcal{G}}_{y_s} \circ y_s^{\dagger\otimes 2}\,\mathbb{X}_{t,u}\right) \underset{3}{\approx} \alpha_s\left(S^{\mathcal{G}}_{y_s} \circ y_s^{\dagger\otimes 2}\,\mathbb{X}_{t,u}\right).
\end{aligned} \tag{3.19}$$

Adding $\tilde{z}_{t,u}$ in equation (3.16) to

$$\tilde{z}_{s,t} = \alpha_s\left(\psi\left(y_s, y_t\right)\right) + \alpha_s^\dagger\left(I \otimes y_s^\dagger\right)\mathbb{X}_{s,t} + \alpha_s\left(S^{\mathcal{G}}_{y_s} \circ y_s^{\dagger\otimes 2}\mathbb{X}_{s,t}\right)$$

while making use of equations (3.17)–(3.19) and Chen's identity in (2.2) shows

$$\tilde{z}_{s,t} + \tilde{z}_{t,u} \underset{3}{\approx} \alpha_s\left(\psi\left(y_s, y_u\right)\right) + \alpha_s^\dagger\left(I \otimes y_s^\dagger\right)\mathbb{X}_{s,u} + \alpha_s\left(S^{\mathcal{G}}_{y_s} \circ y_s^{\dagger\otimes 2}\mathbb{X}_{s,u}\right) = \tilde{z}_{s,u}. \quad \square$$

3.6. A map from $CRP_y^U(M,V)$ to $CRP_y^{\tilde{U}}(M,V)$

Suppose that $\mathcal{G} = (\psi, U)$ and $\tilde{\mathcal{G}} = (\tilde{\psi}, \tilde{U})$ are two gauges on M. Generally, if $\boldsymbol{\alpha} := (\alpha, \alpha^\dagger) \in CRP_y^U(M,V)$, there is no reason to expect it also to be an element of $CRP_y^{\tilde{U}}(M,V)$. However, the main theorem (Theorem 3.32) of this section shows there is a "natural" bijection between $CRP_y^U(M,V)$ and $CRP_y^{\tilde{U}}(M,V)$ which preserves the notions of integration. The following proposition is needed in the proof of Theorem 3.32 and moreover motivates the statement of the theorem.

Proposition 3.29. *If $\mathcal{G} = (\psi, U)$ and $\tilde{\mathcal{G}} = (\tilde{\psi}, \tilde{U})$ are two gauges on M and $\mathbf{y} = (y, y^\dagger) \in CRP_{\mathbf{X}}(M)$, then*

$$\mathbf{y}_{s,t}^{\mathcal{G}} \underset{3}{\approx} \mathbf{y}_{s,t}^{\tilde{\mathcal{G}}} + \big(S_{y_s}^{\tilde{U},U}\big((y_s^\dagger)^{\otimes 2}\,\mathbb{X}_{s,t}\big), 0\big), \tag{3.20}$$

where $\mathbf{y}_{s,t}^{\mathcal{G}}$ and $\mathbf{y}_{s,t}^{\tilde{\mathcal{G}}}$ are as in Definition 3.20.

Proof. From Proposition 3.17,

$$\begin{aligned}\psi(y_s,y_t) - \tilde{\psi}(y_s,y_t) &\underset{3}{\approx} \frac{1}{2}\, S_{y_s}^{\tilde{\psi}_*,\psi_*}\left(\psi(y_s,y_t)\otimes\psi\left(y_s,y_t\right)\right)\\ &\underset{3}{\approx} \frac{1}{2}\, S_{y_s}^{\tilde{\psi}_*,\psi_*}\left(\left(y_s^\dagger\otimes y_s^\dagger\right)[x_{s,t}\otimes x_{s,t}]\right) = S_{y_s}^{\tilde{\psi}_*,\psi_*}\big((y_s^\dagger)^{\otimes 2}\,\mathbb{X}_{s,t}\big),\end{aligned}$$

wherein we have used $S_{y_s}^{\tilde{\psi}_*,\psi_*}$ is symmetric and $\mathbf{X} = (x, \mathbb{X})$ is a weak-geometric rough path for the last equality. Making use of this estimate it now follows that

$$\begin{aligned}\mathbf{y}_{s,t}^{\mathcal{G}} - \mathbf{y}_{s,t}^{\tilde{\mathcal{G}}} &= \big(\psi(y_s,y_t) - \tilde{\psi}(y_s,y_t) + \big(S_{y_s}^{\mathcal{G}} - S_{y_s}^{\tilde{\mathcal{G}}}\big)\big((y_s^\dagger)^{\otimes 2}\,\mathbb{X}_{s,t}\big), 0\big)\\ &\underset{3}{\approx} \big(\big(S_{y_s}^{\tilde{\psi}_*,\psi_*} + S_{y_s}^{\mathcal{G}} - S_{y_s}^{\tilde{\mathcal{G}}}\big)\big((y_s^\dagger)^{\otimes 2}\,\mathbb{X}_{s,t}\big), 0\big).\end{aligned} \tag{3.21}$$

On the other hand, by Lemma 3.11,

$$S^{\tilde{\psi}_*,\psi_*} = S^{\tilde{\psi}_*,\tilde{U}} + S^{\tilde{U},\psi_*} = S^{\tilde{\psi}_*,\tilde{U}} + S^{\tilde{U},U} + S^{U,\psi_*} = S^{\tilde{\mathcal{G}}} - S^{\mathcal{G}} + S^{\tilde{U},U},$$

which combined with (3.21) gives (3.20). □

Corollary 3.30. *The integral, $\int \langle \boldsymbol{\alpha}, d\mathbf{y}^{\mathcal{G}}\rangle$ only depends on the choice of parallelism U, and not on the logarithm used to make the gauge $\mathcal{G} = (\psi, U)$.*

Proof. From Proposition 3.29 with $U = \tilde{U}$, it follows that

$$\int_s^t \langle \boldsymbol{\alpha}, d\mathbf{y}^{\mathcal{G}}\rangle \underset{3}{\approx} \langle \boldsymbol{\alpha}_s, \mathbf{y}_{s,t}^{\mathcal{G}}\rangle \underset{3}{\approx} \langle \boldsymbol{\alpha}_s, \mathbf{y}_{s,t}^{\tilde{\mathcal{G}}}\rangle \underset{3}{\approx} \int_s^t \langle \boldsymbol{\alpha}, d\mathbf{y}^{\tilde{\mathcal{G}}}\rangle,$$

from which it follows that the two additive functionals $\int\langle \boldsymbol{\alpha}, d\mathbf{y}^{\mathcal{G}}\rangle$ and $\int\langle \boldsymbol{\alpha}, d\mathbf{y}^{\tilde{\mathcal{G}}}\rangle$ must be equal. □

Remark 3.31. Corollary 3.30 should not come as a surprise; the definition of a controlled rough path $\mathbf{y}$ does not depend on the choice of gauge $\mathcal{G}$ while the definition of $\boldsymbol{\alpha}$ only depends only on a parallelism. Thus, if such an integral is to exist uniquely for each $\boldsymbol{\alpha}$ and $\mathbf{y}$, it cannot depend on a choice of logarithm.

Although the integral is independent of logarithm, such a geometrical device is necessary to write the integral approximation in (3.13); the reader may refer to Subsection 6.1 in the Appendix to see why this is the case even in the smooth category.

If $\boldsymbol{\alpha} = \big(\alpha, \alpha^\dagger\big) \in CRP_y^U(M,V)$ and $U \neq \tilde{U}$, then

$$\big\langle \boldsymbol{\alpha}_s, \mathbf{y}^{\mathcal{G}}_{s,t}\big\rangle \underset{3}{\approx} \big\langle \boldsymbol{\alpha}_s, \mathbf{y}^{\tilde{\mathcal{G}}}_{s,t} + \big(S^{\tilde{U},U}_{y_s}\big((y^\dagger_s)^{\otimes 2}\,\mathbb{X}_{s,t}\big), 0\big)\big\rangle = \langle \tilde{\boldsymbol{\alpha}}_s, \mathbf{y}^{\tilde{\mathcal{G}}}_{s,t}\rangle, \tag{3.22}$$

where $\tilde{\boldsymbol{\alpha}}_s$ is defined in (3.23) below. The identity in (3.22) suggests the following theorem.

Theorem 3.32. *The map*

$$\boldsymbol{\alpha}_s = (\alpha_s, \alpha^\dagger_s) \to \tilde{\boldsymbol{\alpha}}_s := \big(\tilde{\alpha}_s, \tilde{\alpha}^\dagger_s\big) := \big(\alpha_s, \alpha^\dagger_s + \alpha_s S^{\tilde{U},U}_{y_s} y^\dagger_s \otimes I\big) \tag{3.23}$$

is a bijection from $CRP_y^U(M,V)$ to $CRP_y^{\tilde{U}}(M,V)$ such that

$$\int \big\langle \boldsymbol{\alpha}, d\mathbf{y}^{\mathcal{G}}\big\rangle = \int \big\langle \tilde{\boldsymbol{\alpha}}, d\mathbf{y}^{\tilde{\mathcal{G}}}\big\rangle. \tag{3.24}$$

Proof. The only thing that is really left to prove here is the assertion that $\tilde{\boldsymbol{\alpha}} \in CRP_y^{\tilde{U}}(M,V)$. First we prove that item (3) of Definition 3.1 holds for $\tilde{\boldsymbol{\alpha}}$.

From Theorem 3.15 with $m = y_s$ and $n = y_t$, we find

$$U(y_s, y_t)\,\tilde{U}(y_s, y_t)^{-1} \underset{2}{\approx} I + S^{\tilde{U},U}_{y_s}\left(\psi(y_s, y_t) \otimes (\cdot)\right),$$

and then combining this result with Corollary 2.29 shows

$$\tilde{U}(y_t, y_s) \underset{2}{\approx} U\left(y_t, y_s\right)\big[I + S^{\tilde{U},U}_{y_s}\left(\psi(y_s, y_t) \otimes (\cdot)\right)\big]. \tag{3.25}$$

From this equation and the fact that $\boldsymbol{\alpha} \in CRP_y^U(M,V)$, we learn

$$\begin{aligned}
\alpha_t \tilde{U}(y_t, y_s) - \alpha_s &\underset{2}{\approx} \alpha_t\, U(y_t, y_s)\big[I + S^{\tilde{U},U}_{y_s}\left(\psi(y_s, y_t) \otimes (\cdot)\right)\big] - \alpha_s \\
&\underset{2}{\approx} \big(\alpha_s + \alpha^\dagger_s x_{s,t}\big)\big[I + S^{\tilde{U},U}_{y_s}\left(\psi(y_s, y_t) \otimes (\cdot)\right)\big] - \alpha_s \\
&\underset{2}{\approx} \alpha^\dagger_s\, x_{s,t} + \alpha_s S^{\tilde{U},U}_{y_s}\big(y^\dagger_s\, x_{s,t} \otimes (\cdot)\big) = \tilde{\alpha}^\dagger_s\left(x_{s,t} \otimes (\cdot)\right),
\end{aligned}$$

as desired.

Next we check item (4) of Definition 3.1. We are given

$$\begin{aligned}
0 &\underset{1}{\approx} \alpha^\dagger_t \circ (I \otimes U(y_t, y_s)) - \alpha^\dagger_s \\
&= \tilde{\alpha}^\dagger_t \circ (I \otimes \tilde{U}(y_t, y_s)) - \tilde{\alpha}^\dagger_s - \alpha_t \circ S^{\tilde{U},U}_{y_t} \circ (y^\dagger_t \otimes U(y_t, y_s)) + \alpha_s \circ S^{\tilde{U},U}_{y_s} \circ (y^\dagger_s \otimes I)
\end{aligned}$$

wherein we have used that $U(y_s, y_t) \underset{1}{\approx} \tilde{U}(y_s, y_t)$ (for example, see (3.25)). We therefore must show the last line is approximately 0. However, by Lemma 3.26, we have $S_{y_t}^{\tilde{U},U} \circ U(y_t, y_s)^{\otimes 2} \underset{1}{\approx} U(y_t, y_s) \circ S_{y_t}^{\tilde{U},U}$. Thus

$$\begin{aligned}
&\alpha_t \circ S_{y_t}^{\tilde{U},U} \circ (y_t^\dagger \otimes U(y_t, y_s)) - \alpha_s \circ S_{y_s}^{\tilde{U},U} \circ (y_s^\dagger \otimes I) \\
&\quad \underset{1}{\approx} \alpha_t \circ S_{y_t}^{\tilde{U},U} \circ (U(y_t, y_s) y_s^\dagger \otimes U(y_t, y_s)) - \alpha_s \circ S_{y_s}^{\tilde{U},U} \circ (y_s^\dagger \otimes I) \\
&\quad \underset{1}{\approx} [\alpha_t \circ U(y_t, y_s) - \alpha_s][S_{y_s}^{\tilde{U},U} \circ (y_s^\dagger \otimes I)] \underset{1}{\approx} 0.
\end{aligned}$$

□

4. Integrating one-forms along a CRP

For the next result we extend (in the usual way) the covariant derivative ∇^U, defined in Remark 2.20 acting on vector fields, to a covariant derivative acting on one-forms which we continue to denote by ∇^U. In more detail, if $\alpha \in \Omega^1(M, V)$, $Y \in \Gamma(TM)$, and $v_m \in T_m M$, then $\nabla_{v_m}^U \alpha$ is determined by the product rule, i.e.,

$$v_m [\alpha (Y)] = \left(\nabla_{v_m}^U \alpha\right) (Y(m)) + \alpha_m \left(\nabla_{v_m}^U Y\right). \tag{4.1}$$

Since

$$\begin{aligned}
v_m [\alpha (Y)] &= \frac{d}{dt}\Big|_0 \alpha \left(U(m, \sigma_t)^{-1} U(m, \sigma_t) Y(\sigma_t)\right) \\
&= \frac{d}{dt}\Big|_0 \alpha \left(U(m, \sigma_t)^{-1} Y(m)\right) + \alpha_m \left(\nabla_{v_m}^U Y\right),
\end{aligned}$$

where σ_t is a path in M such that $\dot{\sigma}_0 = v_m$, it follows that $\nabla_{v_m}^U \alpha$ may be computed by

$$\nabla_{v_m}^U \alpha = \frac{d}{dt}\Big|_0 [\alpha \circ U(m, \sigma_t)^{-1}]. \tag{4.2}$$

Moreover, by Corollary 2.29, we may alternatively write (4.2) as

$$\nabla_{v_m}^U \alpha = \frac{d}{dt}\Big|_0 [\alpha \circ U(\sigma_t, m)].$$

Lemma 4.1. *Let V be a Banach space and U be a parallelism on M. If $\alpha \in \Omega^1(M, V)$ is a V-valued smooth one-form on M, then*

$$\alpha_n \circ U(n, m) - \alpha_m =_2 \nabla_{\psi(m,n)}^U \alpha,$$

where ∇^U is the covariant derivative defined in Remark 2.20 and ψ is any logarithm.

Proof. To prove the lemma, we note this is a local result and we therefore may assume $M = \mathbb{R}^d$. Then by Taylor's theorem,

$$\begin{aligned}
\alpha_n \circ U(n, m) &= \alpha_m + D[\alpha_{(\cdot)} \circ U(\cdot, m)](m)(n - m) + O(|n - m|^2) \\
&= \alpha_m + \nabla_{(n-m)_m}^U \alpha + O(|n - m|^2) = \alpha_m + \nabla_{\psi(m,n)}^U \alpha + O(|\psi(m, n)|^2).
\end{aligned}$$

□

Suppose that $\alpha \in \Omega^1(M,V)$ is a V-valued one-form and U is a parallelism on M. We wish to take $\alpha_s^U = \alpha_{y_s} := \alpha|_{T_{y_s}M}$. Making use of Lemma 4.1, we find

$$\alpha_t^U \circ U(y_t, y_s) - \alpha_s \underset{2}{\approx} \nabla^U_{\psi(y_s,y_t)}\alpha \underset{2}{\approx} \nabla^U_{y_s^\dagger x_{s,t}}\alpha, \tag{4.3}$$

and this computation suggests the following proposition.

Proposition 4.2. *Suppose that $\alpha \in \Omega^1(M,V)$ is a V-valued one-form and U is a parallelism on M, then*

$$\boldsymbol{\alpha}_s^{(y,U)} := \left(\alpha_{y_s}, \alpha_s^{\dagger(y,U)}\right) := \left(\alpha|_{T_{y_s}M}, \nabla^U_{y_s^\dagger(\cdot)}\alpha\right) \in CRP_y^U(M,V).$$

Proof. In light of how $\boldsymbol{\alpha}_s^{y,U}$ has been defined and of equation (4.3), we need only verify item (4) in Definition 3.1 is satisfied. To this end, suppose that $w \in W$, then

$$\begin{aligned} \alpha_t^{\dagger(y,U)} \circ (I \otimes U(y_t,y_s))(w \otimes (\cdot)) &= \left(\nabla^U_{y_t^\dagger w}\alpha\right) U(y_t,y_s) \\ &\underset{1}{\approx} \left(\nabla^U_{U(y_t,y_s)y_s^\dagger w}\alpha\right) U(y_t,y_s), \end{aligned} \tag{4.4}$$

wherein we have used inequality (2.23) along with Corollary 2.29 in the last line. Since for $v_m \in T_mM$ the function $F(n) := (\nabla^U_{U(n,m)v_m}\alpha)U(n,m) \in L(T_mM,V)$ is smooth, it follows by Taylor's theorem that $F(n) =_1 F(m)$ which translates to

$$\left(\nabla^U_{U(n,m)v_m}\alpha\right) U(n,m) =_1 \nabla^U_{v_m}\alpha.$$

Taking $m = y_s$, $n = y_t$, and $v_m = y_s^\dagger w$ in this estimates shows

$$\left(\nabla^U_{U(y_t,y_s)y_s^\dagger w}\alpha\right) U(y_t,y_s) \underset{1}{\approx} \nabla^U_{y_s^\dagger w}\alpha$$

which combined with (4.4) completes the proof. □

Theorem 4.3. *If $\alpha \in \Omega^1(M,V)$ is a V-valued one-form, then $\int \left\langle \boldsymbol{\alpha}^{(y,U)}, d\mathbf{y}^{\mathcal{G}} \right\rangle$ is independent of any choice of gauge $\mathcal{G} = (\psi, U)$ on M. In the future we denote this integral more simply as $\int \langle \boldsymbol{\alpha}, d\mathbf{y} \rangle$.*

Proof. Suppose that U and $\tilde{U}$ are two parallelisms. According to Theorem 3.32, it suffices to show

$$\alpha_s^{\dagger\left(y,\tilde{U}\right)} = \alpha_s^{\dagger(y,U)} + \alpha_{y_s} S^{\tilde{U},U}_{y_s}\left[y_s^\dagger \otimes I\right]. \tag{4.5}$$

We will see that (4.5) is a fairly direct consequence of Example 3.10 which, when translated to the language of forms (see (4.1)), states

$$\nabla_{v_m}\alpha = \tilde{\nabla}_{v_m}\alpha - \alpha \circ S^{\tilde{U},U}_m\left(v_m \otimes (\cdot)\right). \tag{4.6}$$

So for $w \in W$, we have

$$\alpha_s^{\dagger(y,\tilde{U})} w = \tilde{\nabla}_{y_s^\dagger w}\alpha = \nabla_{y_s^\dagger w}\alpha + \alpha_{y_s} S^{\tilde{U},U}_m(y_s^\dagger w \otimes (\cdot)) = \alpha_s^{\dagger(y,U)} w + \alpha_{y_s} S^{\tilde{U},U}_m(y_s^\dagger w \otimes (\cdot))$$

which proves (4.5). □

Let us now record a number of possible different expressions for computing $\int_s^t \alpha(d\mathbf{y})$ depending on the choice of gauge we make.

Proposition 4.4. *Let $\mathcal{G} = (\psi, U)$ be a gauge. There exists a $\delta > 0$ such that for $s < t$ and $t - s < \delta$, the approximation*

$$\Big[\int \alpha(d\mathbf{y})\Big]_{s,t}^1 \underset{3}{\approx} \alpha_{y_s}\left(\psi(y_s, y_t)\right) + \left[\left(\nabla^U_{(\cdot)}\alpha\right)_{y_s} + \alpha_{y_s} \circ S^{\mathcal{G}}_{y_s}\right] \circ y_s^{\dagger\otimes 2}\, \mathbb{X}_{s,t}$$

holds.

In the case that we take $U = U^\psi$, we get a slightly simpler formula.

Corollary 4.5. *Let ψ be a logarithm. There exists a $\delta > 0$ such that for $s < t$ and $t - s < \delta$, the approximation*

$$\Big[\int \alpha(d\mathbf{y})\Big]_{s,t}^1 \underset{3}{\approx} \alpha_{y_s}\left(\psi(y_s, y_t)\right) + d\big(\alpha_{(\cdot)} \circ \left(\psi_{(\cdot)}\right)_{*y_s}\big)_{y_s} \circ y_s^{\dagger\otimes 2}\, \mathbb{X}_{s,t}$$

holds.

Example 4.6. Let ∇ be a covariant derivative on M. There exists a $\delta > 0$ such that for $s < t$ and $t - s < \delta$, the approximation

$$\Big[\int \alpha(d\mathbf{y})\Big]_{s,t}^1 \underset{3}{\approx} \alpha_{y_s}\big(\left(\exp^\nabla_{y_s}\right)^{-1}(y_t)\big) + \Big[\left(\nabla\alpha\right)_{y_s} + \frac{1}{2}\,\alpha_{y_s} \circ T^\nabla_{y_s}\Big] \circ y_s^{\dagger\otimes 2}\, \mathbb{X}_{s,t}$$

holds. Indeed this follows immediately from Proposition 4.4, Lemma 3.13, and the fact that

$$(\nabla\alpha)_{y_s}(v_m, w_m) := v_m[\alpha(\boldsymbol{W})] - \alpha(\nabla_{v_m}\boldsymbol{W}) = \; d(\alpha_{(\cdot)} \circ \boldsymbol{W}(\cdot))_{y_s}(v_m) - \alpha(\nabla_{v_m}\boldsymbol{W})$$

where $\boldsymbol{W}$ is any vector field such that $\boldsymbol{W}(m) = w_m$. Choosing $\boldsymbol{W} = U^\nabla(\cdot, m)\, w_m$, we have

$$\nabla_{v_m}\boldsymbol{W} = \nabla_{v_m} U^\nabla(\cdot, m)\, w_m = 0$$

by the definition of parallel translation.

4.1. Integration of a one-form using charts

It is easy to see that by independence of gauges, the integral of a one-form along $\left(y_s, y_s^\dagger\right)$ is an object which we only need to compute locally. As mentioned in Remark 2.22 we have an example of a local gauge by using a chart. Plugging this formula into the integral approximation from Corollary 4.5, we get the following.

Corollary 4.7. *Let ϕ be a chart on M. For all $a, b \in [0, T]$ such that $y[a, b] \subset D(\phi)$, we have that the approximation*

$$\begin{aligned}\Big[\int \alpha(d\mathbf{y})\Big]_{s,t}^1 \underset{3}{\approx} \; & \alpha_{y_s}\big(\left(d\phi_{y_s}\right)^{-1}\left[\phi(y_t) - \phi(y_s)\right]\big) \\ & + d\big(\alpha_{(\cdot)} \circ \left(d\phi_{(\cdot)}\right)^{-1} d\phi_{y_s}\big)_{y_s} \circ y_s^{\dagger\otimes 2}\, \mathbb{X}_{s,t}\end{aligned} \tag{4.7}$$

holds for all $s < t \in [a, b]$.

Although this formula looks a bit complicated, it may be reduced to something that makes more sense. First, note that

$$\alpha_m \circ (d\phi_m)^{-1} = \left[\left(\phi^{-1}\right)^* \alpha \right]_{\phi(m)}.$$

Thus we can reduce the right-hand side of (4.7) to

$$\begin{aligned}
&\left[(\phi^{-1})^*\alpha\right]_{\phi(y_s)} (\phi(y_t) - \phi(y_s)) + d\Big(\left[(\phi^{-1})^*\alpha\right]_{\phi(\cdot)} d\phi_{y_s} \Big)_{y_s} \circ y_s^{\dagger\otimes 2}\, \mathbb{X}_{s,t} \\
&\quad = \left[(\phi^{-1})^*\alpha\right]_{\phi(y_s)} (\phi(y_t) - \phi(y_s)) + \left[(\phi^{-1})^*\alpha\right]'_{\phi(y_s)} \left[d\phi_{y_s} \circ y_s^{\dagger}\right]^{\otimes 2} \mathbb{X}_{s,t}.
\end{aligned}$$

Now, if we recall Notation 2.41, we see that this is approximately equal to another rough integral. More precisely,

$$\left[\int \alpha(d\mathbf{y}) \right]^1_{s,t} \underset{3}{\approx} \left[\int (\phi^{-1})^* \, \alpha \, (d\phi_* \mathbf{y}) \right]^1_{s,t}.$$

However, additive functionals are unique up to this order, so in fact

$$\left[\int \alpha(d\mathbf{y}) \right]^1_{s,t} = \left[\int (\phi^{-1})^* \, \alpha \, (d\phi_* \mathbf{y}) \right]^1_{s,t},$$

which is a relation which should hold under any reasonable integral. This is summarized in the following theorem, which gives us an alternative way of defining this integral.

Theorem 4.8. *The integral, $\int \alpha(d\mathbf{y})$, is the unique V-valued rough path controlled by $\mathbf{X}$ on $[0,T]$ starting at 0 determined by*

(1) $\left[\int \alpha(d\mathbf{y})\right]^1_{s,t} = \left[\int \left((\phi^{-1})^*\alpha\right)(d\phi_*\mathbf{y})\right]^1_{s,t}$ *for any chart and $s < t \in [0,T]$ such that $y([s,t]) \subset D(\phi)$*

(2) $\left[\int \alpha(d\mathbf{y})\right]^\dagger_s = \alpha_{y_s} \circ y_s^\dagger$.

(*See Theorem 4.15 below for a more general version of this theorem.*)

A notion of associativity for the developed integration theory holds:

Theorem 4.9 (Associativity theorem). *Let $\mathbf{y} \in CRP_{\mathbf{X}}(M)$, $\alpha \in \Omega^1(M,V)$, and let $K\colon M \to L(V,\tilde{V})$ be a smooth function so that $K\alpha \in \Omega^1(M,\tilde{V})$. If $\mathbf{z} = \int \alpha(d\mathbf{y}) \in CRP(V)$, then*

$$\int (K\alpha)(d\mathbf{y}) = \int \langle K_*(\mathbf{y}), d\mathbf{z}\rangle \quad \Big(=: \int \Big\langle K_*(\mathbf{y}), d\int \alpha(d\mathbf{y}) \Big\rangle \Big),$$

where $K_*(\mathbf{y}) = \left(K(y), K_{*y} y^\dagger\right) \in CRP_{\mathbf{X}}(\mathrm{Hom}(V,V'))$.

Moreover, if $\mathbf{f}$ and $\mathbf{f}\boldsymbol{\alpha} := \left(f_s\,\alpha_s, f_s^\dagger\,(I \otimes \alpha_s) + f_s \alpha_s^\dagger\right)$ are as in Proposition 3.6 and $\mathbf{z} = (z, z^\dagger) = \int \left\langle \boldsymbol{\alpha}, d\mathbf{y}^{\mathcal{G}} \right\rangle$, then

$$\int \langle \mathbf{f}, d\mathbf{z}\rangle = \int \left\langle \mathbf{f}\boldsymbol{\alpha}, d\mathbf{y}^{\mathcal{G}} \right\rangle,$$

or in other words,

$$\int \left\langle \mathbf{f}, d\int \left\langle \boldsymbol{\alpha}, d\mathbf{y}^{\mathcal{G}} \right\rangle \right\rangle = \int \left\langle \mathbf{f}\boldsymbol{\alpha}, d\mathbf{y}^{\mathcal{G}} \right\rangle .$$

Proof. We prove the second statement first; we have the approximations

$$\begin{aligned}
&\left[\int \left\langle \mathbf{f}\boldsymbol{\alpha}, d\mathbf{y}^{\mathcal{G}} \right\rangle\right]^1_{s,t} \\
&\quad \underset{3}{\approx} f_s\, \alpha_s \left(\psi(y_s,y_t) + \mathcal{S}^{\mathcal{G}}_{y_s}\left(y_s^{\dagger\otimes 2}\, \mathbb{X}_{s,t}\right)\right) + \left[\left(f_s^\dagger \left(I\otimes \alpha_s\right) + f_s \alpha_s^\dagger\right)\right]\left(I\otimes y_s^\dagger\right)\mathbb{X}_{s,t} \\
&\quad = f_s \left(\alpha_s \left(\psi(y_s,y_t) + \mathcal{S}^{\mathcal{G}}_{y_s}\left(y_s^{\dagger\otimes 2}\, \mathbb{X}_{s,t}\right)\right) + \alpha_s^\dagger \left(I\otimes y_s^\dagger\right)\mathbb{X}_{s,t}\right) + f_s^\dagger \left(I\otimes \alpha_s y_s^\dagger\right)\mathbb{X}_{s,t} \\
&\quad \underset{3}{\approx} f_s\left(z_{s,t}\right) + f_s^\dagger \left(I\otimes z_s^\dagger\right)\mathbb{X}_{s,t} \underset{3}{\approx} \left[\int \langle \mathbf{f}, d\mathbf{z}\rangle\right]^1_{s,t}.
\end{aligned}$$

As the first and last terms of this equation are additive functionals, they must be equal.

Additionally,

$$\left[\int \left\langle \mathbf{f}\boldsymbol{\alpha}, d\mathbf{y}^{\mathcal{G}} \right\rangle\right]^\dagger_s = f_s\, \alpha_s\left(y_s^\dagger\right) = f_s z_s^\dagger = \left[\int \langle \mathbf{f}, d\mathbf{z}\rangle\right]^\dagger_s.$$

Thus, the two controlled rough paths are equal.

The first statement of the theorem, with our current toolset, can be proved in two different ways. We can reduce the result to a special case of the result proved above, or, by using the chart definitions of integration along a one-form, we can reduce it to its validity in the flat case. The first method is quick but may hide the concept of what is happening. We therefore provide both proofs.

Method 1. Letting $\mathcal{G} = (\psi, U)$ be any gauge, we define $\mathbf{f} := (f, f^\dagger) \in CRP_{\mathbf{X}}(\mathrm{Hom}(V,\tilde{V}))$ by the formula

$$f_s := K(y_s) \quad \text{and} \quad f_s^\dagger := K_{*y_s} y_s^\dagger$$

and $\boldsymbol{\alpha}^{(y,U)}$ as in Proposition 4.2 (see Proposition 4.10 below to see why $\mathbf{f} \in CRP_{\mathbf{X}}(\mathrm{Hom}(V,\tilde{V}))$). Then by the statement already proved, we have

$$\int \left\langle \mathbf{f}\boldsymbol{\alpha}^{(y,U)}, d\mathbf{y}^{\mathcal{G}} \right\rangle = \int \langle \mathbf{f}, d\mathbf{z}\rangle , \tag{4.8}$$

where $\mathbf{z} = \int \left\langle \boldsymbol{\alpha}^{(y,U)}, d\mathbf{y}^{\mathcal{G}} \right\rangle = \int \alpha\,(d\mathbf{y})$. The right-hand side in equation (4.8) is simply $\int \langle K_*(\mathbf{y}), d\mathbf{z}\rangle$, while the $\mathbf{f}\boldsymbol{\alpha}^{(y,U)}$ term on the left-hand side can be recognized as $(K\alpha)^{(y,U)}$. Indeed, by the product rule with ∇^U, we have

$$\begin{aligned}
(K\alpha)^{(y,U)}_s &= \left(K(y_s)\alpha|_{T_{y_s}M}, \nabla^U_{y_s^\dagger(\cdot)}\left[K(\cdot)\alpha\right]\right) = \left(K\alpha|_{T_{y_s}M}, K_{*y_s} y_s^\dagger \alpha + K(y_s)\nabla^U_{y_s^\dagger(\cdot)}\alpha\right) \\
&= \left(f_s\, \alpha_s, f_s^\dagger \alpha + f_s\, \alpha_s^{\dagger(y,U)}\right) = \mathbf{f}\boldsymbol{\alpha}^{(y,U)}.
\end{aligned}$$

Thus

$$\int (K\alpha)(d\mathbf{y}) := \int \left\langle (K\alpha)^{(y,U)}_s, d\mathbf{y}^{\mathcal{G}} \right\rangle = \int \left\langle \mathbf{f}\boldsymbol{\alpha}^{(y,U)}, d\mathbf{y}^{\mathcal{G}} \right\rangle = \int \langle K_*(\mathbf{y}), d\mathbf{z}\rangle .$$

Method 2. By a simple patching argument, this is really a local result and hence using the chart definitions of integration it suffices to check this result in the case M is an open subset of $\mathbb{R}^d$. First we check the derivative processes. From the definitions we have

$$z_s^\dagger = \alpha_{y_s} \circ y_s^\dagger \quad \text{and} \quad \Big[\int (K\alpha)(d\mathbf{y})\Big]_s^\dagger = (K\alpha)_{y_s} \circ y_s^\dagger = K(y_s)\, \alpha_{y_s} \circ y_s^\dagger = K(y_s)\, z_s^\dagger.$$

Thus

$$\Big[\int (K\alpha)(d\mathbf{y})\Big]_s^\dagger = K(y_s)\, z_s^\dagger.$$

On the other hand,

$$\Big[\int \langle K_*(\mathbf{y}), d\mathbf{z}\rangle\Big]_s^\dagger = [K(y)]_s z_s^\dagger = K(y_s)\, z_s^\dagger$$

Similarly for the paths

$$z_{s,t} \underset{3}{\approx} \alpha\left(y_{s,t}\right) + \alpha'_{y_s} y_s^{\dagger\otimes 2}\, \mathbb{X}_{s,t},$$

and so

$$\begin{aligned}\Big[\int (K\alpha)(d\mathbf{y})\Big]_{s,t}^1 &\underset{3}{\approx} (K\alpha)_{y_s}\, y_{s,t} + (K\alpha)'_{y_s}\, y_s^{\dagger\otimes 2}\, \mathbb{X}_{s,t} \\ &= K(y_s)\, \alpha_{y_s} y_{s,t} + K(y_s)\, \alpha'_{y_s} y_s^{\dagger\otimes 2}\, \mathbb{X}_{s,t} + \left[K'_{y_s}\left(y_s^\dagger(\cdot) \otimes \alpha y_s^\dagger(\cdot)\right)\right] \mathbb{X}_{s,t} \\ &\underset{3}{\approx} K(y_s)\, z_{s,t} + K'_{y_s}\left(y_s^\dagger \otimes z_s^\dagger\right) \mathbb{X}_{s,t}.\end{aligned}$$

On the other hand,

$$\Big[\int \langle K_*(\mathbf{y}), d\mathbf{z}\rangle\Big]_{s,t}^1 \underset{3}{\approx} K(y_s) z_{s,t} + [K_*(\mathbf{y})]_s^\dagger z_s^\dagger \mathbb{X}_{s,t} = K(y_s) z_{s,t} + K'_{y_s}(y_s^\dagger \otimes z_s^\dagger)\mathbb{X}_{s,t}.$$

Comparing these expressions completes the proof. □

4.2. Push-forwards of controlled rough paths

Let $M = M^d$ and $\tilde{M} = \tilde{M}^{\tilde{d}'}$ be manifolds. Let $f\colon M \to \tilde{M}$ be smooth and suppose $\mathbf{y}_s = \left(y_s, y_s^\dagger\right) \in CRP_{\mathbf{X}}(M)$. In Definition 4.11 below, we are going to give a definition of the push-forward of $\mathbf{y}$ by f which generalizes Example 2.56.

Proposition 4.10. *The pair* $\left(f(y_s), f_* \circ y_s^\dagger\right)$ *is an element of* $CRP_{\mathbf{X}}(\tilde{M})$.

Proof. Suppose $\tilde{\phi}$ is a chart on $\tilde{M}$ such that $f \circ y\left([a,b]\right) \subseteq D(\tilde{\phi})$. We must show that

$$\left|\tilde{\phi} \circ f(y_t) - \tilde{\phi} \circ f(y_s) - d\tilde{\phi} \circ f_* y_s^\dagger\, x_{s,t}\right| \le C_{\tilde{\phi},a,b}\, \omega(s,t)^{2/p} \tag{4.9}$$

and

$$\left|d\tilde{\phi} \circ f_* y_t^\dagger - d\tilde{\phi} \circ f_* y_s^\dagger\right| \le C_{\tilde{\phi},a,b}\, \omega(s,t)^{1/p} \tag{4.10}$$

hold for some $C_{\tilde{\phi},a,b}$ for all $s \le t$ in $[a,b]$. We can again use our proof strategy outlined in Remark 2.53 to treat this problem in nice neighborhoods. We leave it to the reader to follow the pattern of earlier proofs to see that we can assume without loss of generality that there is a chart ϕ on M such that $y([a,b]) \subseteq D(\phi)$ and $R(\phi)$ is convex. Which these simplifications, we note that $(z_s, z_s^\dagger) := (\phi(y_s), d\phi \circ y_s^\dagger)$ is a controlled rough path on $R(\phi)$ and the function $F\colon = \tilde{\phi} \circ f \circ \phi^{-1} : R(\phi) \to R(\tilde{\phi})$ is a map between Euclidean spaces. Therefore Inequalities (4.9) and (4.10) reduce to the fact that the pair $\left(F(z_s), F'(z_s) \circ z_s^\dagger\right)$ is a controlled rough path in $\mathbb{R}^{\tilde{d}}$ (which is trivial by applying Taylor's theorem after we check that we get the correct terms); indeed, by a simple computation, we have

$$\begin{aligned} F'(z_s) \circ z_s^\dagger &= d\tilde{\phi} \circ f_* \circ \left(d\phi^{-1}\right)_{z_s} \circ d\phi_{y_s} \circ y_s^\dagger \\ &= d\tilde{\phi} \circ f_* \circ \left(d\phi_{y_s}\right)^{-1} \circ d\phi_{y_s} \circ y_s^\dagger = d\tilde{\phi} \circ f_* y_s^\dagger, \end{aligned}$$

and clearly $F(z_s) = \tilde{\phi} \circ f(y_s)$. □

Definition 4.11. The **push-forward** of $\mathbf{y}$ denoted by $f_*\mathbf{y}$ or $f_*(y, y^\dagger)$ is the rough path controlled by $\mathbf{X}$ with path $f(y_s)$ and derivative process $f_* \circ y_s^\dagger$. If $\tilde{M} = \mathbb{R}^{\tilde{d}}$, we will abuse notation and write $f_*\mathbf{y}_s$ to mean $\left(f(y_s), df \circ y_s^\dagger\right)$ (i.e., we forget the base point on the derivative process).

Remark 4.12. The push-forward operation on elements in $CRP_{\mathbf{X}}(M)$ is clearly covariant, i.e., if $f\colon M \to N$ and $g\colon N \to P$ are two smooth maps between manifolds M, N, and P, then $(g \circ f)_*(\mathbf{y}) = g_*(f_*(\mathbf{y}))$.

This definition is consistent with how we defined the integral of a one-form along a controlled rough path in the sense that we have a fundamental theorem of calculus. Let V be a Banach space.

Theorem 4.13. *Let $\mathbf{y}_s = \left(y_s, y_s^\dagger\right) \in CRP_{\mathbf{X}}(M)$ and f be a smooth function from M to V. Then*

$$f(y_s) - f(y_0) = \left[\int df[d\mathbf{y}]\right]_{0,s}^1,$$

where df is interpreted as a one-form. Since we have $df \circ y_s^\dagger = \left[\int df\,[d\mathbf{y}]\right]_s^\dagger$, we have the equality

$$f_*(y, y^\dagger) - (f(y_0), 0) = \int df(d\mathbf{y}).$$

Proof. Although there are ways to do this proof without much machinery, we find it more instructive to work on a Riemannian manifold with the Levi-Civita covariant derivative. Since we have proved that the integral is independent of choice of metric, it does not matter which one we pick. With this in mind, we have the approximation

$$\left[\int df\,[d\mathbf{y}]\right]_{s,t}^1 \underset{3}{\approx} df_{y_s}\left(\exp_{y_s}^{-1}(y_t)\right) + (\nabla df)_{y_s}\left[y_s^{\dagger\otimes 2}\,\mathbb{X}_{s,t}\right],$$

and as ∇df is symmetric, it follows that

$$\begin{aligned}\left[\int df\left[d\mathbf{y}\right]\right]_{s,t}^{1} &\underset{3}{\approx} df_{y_s}\left(\exp_{y_s}^{-1}\left(y_t\right)\right)+\frac{1}{2}\left(\nabla df\right)_{y_s}\left[y_s^{\dagger\otimes 2}\left(x_{s,t}\otimes x_{s,t}\right)\right]\\ &\underset{3}{\approx} df_{y_s}\left(\exp_{y_s}^{-1}\left(y_t\right)\right)+\frac{1}{2}\left(\nabla df\right)_{y_s}\left[\exp_{y_s}^{-1}(y_t)^{\otimes 2}\right]\underset{3}{\approx} f(y_t)-f(y_s).\end{aligned}$$

The last approximation above follows from Taylor's theorem on manifolds (Theorem 6.3 in the Appendix). Note here that $f\left(y_t\right)-f(y_s)$ is additive so that

$$\left[\int df\left[d\mathbf{y}\right]\right]_{s,t}^{1}=f\left(y_t\right)-f(y_s). \qquad \square$$

Remark 4.14. If $M\subseteq$ is an embedded submanifold of $W=\mathbb{R}^k$, $\left(y_s,y_s^{\dagger}\right)\in CRP_{\mathbf{X}}(M)$, $I\colon M\to W$ denotes the identity (or embedding) map, and $\left(z_s,z_s^{\dagger}\right):=I_*\left(y_s,y_s^{\dagger}\right)$, then we have

$$z_s=y_s \quad \text{and} \quad z_s^{\dagger}=\pi_2\circ y_s^{\dagger},$$

where π_2 is the projection of the tangent vector component (i.e., it forgets the base point). We can associate to it a unique rough path $(y,\mathbb{Y})$ in W such that

$$\left(z_s^{\dagger}\otimes z_s^{\dagger}\right)\mathbb{X}_{s,t}\underset{3}{\approx}\mathbb{Y}_{s,t}.$$

In this case, this is a rough path in the embedded sense (see [3]) since

$$\left[I(y_s)\otimes Q(y_s)\right]\left[\mathbb{Y}\right]_{s,t}\underset{3}{\approx}\left[I(y_s)\otimes Q(y_s)\right]\left[z_s^{\dagger}\otimes z_s^{\dagger}\right]\mathbb{X}_{s,t}=0$$

as $Q(y_s)\circ z_s^{\dagger}=0$, where $Q=I-P$ and $P(x)$ is orthogonal projection onto the tangent space at x.

Lastly, we have a relation between push-forwards of paths and pull-backs of one-forms.

Theorem 4.15 (Push me-pull me). *Let $f\colon M\to\tilde{M}$, let $\mathbf{y}_s=\left(y_s,y_s^{\dagger}\right)\in CRP_{\mathbf{X}}(M)$ and let $\alpha\in\Omega^1(\tilde{M},V)$. Then*

$$\left[\int f^*\alpha(d\mathbf{y})\right]^1=\left[\int\alpha\left(d\left(f_*\mathbf{y}\right)\right)\right]^1. \tag{4.11}$$

Moreover,

$$\int f^*\alpha(d\mathbf{y})=\int\alpha\left(d\left(f_*\mathbf{y}\right)\right).$$

Proof. This is a statement we only have to prove locally. Indeed for each $s\in[0,T]$, there are charts ϕ^s and $\tilde{\phi}^s$ on M and $\tilde{M}$ respectively such that $y_s\in D(\phi^s)$ and $f(y_s)\in D(\tilde{\phi}^s)$ which are open. We take $\mathcal{U}_s:=f^{-1}(D(\tilde{\phi}^s))\cap D(\phi^s)$ and shrink it if necessary so that $\mathcal{V}_s=\phi(\mathcal{U}_s)$ is convex. Thus if we can prove that equation (4.11) holds whenever $y([a,b])\subseteq\mathcal{U}$ such that $\phi(\mathcal{U})$ is convex and such that $f(y([a,b]))\subseteq D(\tilde{\phi})$, we will be done. We do this now.

By Theorem 4.8, the fact that pull-backs are contravariant, and that push-forwards are covariant, we have

$$\begin{aligned}
&\Big[\int f^*\alpha(d\mathbf{y})\Big]^1_{s,t} = \Big[\int (\phi^{-1})^* f^*\alpha(d\phi_*\mathbf{y})\Big]^1_{s,t} = \Big[\int (f\circ\phi^{-1})^*\alpha(d\phi_*\mathbf{y})\Big]^1_{s,t}\\
&= \Big[\int (\tilde\phi^{-1}\circ\tilde\phi\circ f\circ\phi^{-1})^*\alpha(d\phi_*\mathbf{y})\Big]^1_{s,t} = \Big[\int (\tilde\phi\circ f\circ\phi^{-1})^*((\tilde\phi^{-1})^*\alpha)(d\phi_*\mathbf{y})\Big]^1_{s,t}\\
&= \Big[\int (\tilde\phi^{-1})^*\alpha(d((\tilde\phi\circ f\circ\phi^{-1})_*\phi_*\mathbf{y}))\Big]^1_{s,t},
\end{aligned}$$

where the last step is just (4.11) on Euclidean space. This is a simple computation (for example, see the appendix of [3]). Thus, we have

$$\begin{aligned}
\Big[\int f^*\alpha(d\mathbf{y})\Big]^1_{s,t} &= \Big[\int (\tilde\phi^{-1})^*\alpha(d((\tilde\phi\circ f\circ\phi^{-1})_*\phi_*\mathbf{y}))\Big]^1_{s,t}\\
&= \Big[\int (\tilde\phi^{-1})^*\alpha(d(\tilde\phi_*(f_*\mathbf{y})))\Big]^1_{s,t} = \Big[\int \alpha(d(f_*\mathbf{y}))\Big]^1_{s,t}.
\end{aligned}$$

The fact that

$$\Big[\int f^*\alpha(d\mathbf{y})\Big]^\dagger = \Big[\int \alpha(d(f_*\mathbf{y}))\Big]^\dagger$$

is trivial. □

5. Rough differential equations

Before discussing rough differential equations on a manifold, we will give an equivalent condition for a controlled rough path $\mathbf{z} \in CRP_{\mathbf{X}}\left(\mathbb{R}^d\right)$ to satisfy the RDE approximation on a compact interval in the flat case using logarithms.

For the next proposition, let ψ be a logarithm on $\mathbb{R}^d$ such that $\psi(x,y) = \left(x, \bar\psi(x,y)\right)$.

Proposition 5.1. *Let $z\colon [a,b] \to \mathbb{R}^d$ be a path and let $\mathcal{W} \subseteq \mathbb{R}^d$ be an open convex set such that $z([a,b]) \subseteq \mathcal{W}$ and $\mathcal{W}\times\mathcal{W} \subseteq D(\psi)$. Then*

$$z_{s,t} \underset{3}{\approx} F_{x_{s,t}}(z_s) + \left(\partial_{F_w(z_s)} F_{\tilde w}\right)(z_s)|_{w\otimes\tilde w = \mathbb{X}_{s,t}} \tag{5.1}$$

if and only if

$$\bar\psi(z_s,z_t) \underset{3}{\approx} F_{x_{s,t}}(z_s) + \left(\partial_{F_w(z_s)}\left[\bar\psi'_{z_s}(\cdot)F_{\tilde w}(\cdot)\right]\right)(z_s)|_{w\otimes\tilde w = \mathbb{X}_{s,t}}. \tag{5.2}$$

Proof. If $z_\cdot$ satisfies (5.1), then from (2.17) of Theorem 2.24 with $y = z_t$ and $x = z_s$ we find

$$\bar\psi(z_s,z_t) = z_{s,t} + \frac{1}{2}\bar\psi''_x(x)\left(z_{s,t}\right)^{\otimes 2} + C(z_s,z_t)\left(z_{s,t}\right)^{\otimes 3} \tag{5.3}$$

$$\underset{3}{\approx} F_{x_{s,t}}(z_s) + \left(\partial_{F_w(z_s)} F_{\tilde w}\right)(z_s)|_{w\otimes\tilde w = \mathbb{X}_{s,t}} + \frac{1}{2}\bar\psi''_{z_s}(z_s)\left[F_{x_{s,t}}(z_s)\right]^{\otimes 2}, \tag{5.4}$$

wherein C is a smooth function and we have made use of the fact that $z_{s,t} \underset{1}{\approx} 0$. By the product rule and the fact that ψ is a logarithm it follows that

$$\begin{aligned}
&\left(\partial_{F_w(z_s)}\left[\bar{\psi}'_{z_s}(\cdot)F_{\tilde{w}}(\cdot)\right]\right)(z_s)\\
&\qquad = \bar{\psi}''_{z_s}(z_s)F_w(z_s)\otimes F_{\tilde{w}}(z_s) + \bar{\psi}'_{z_s}(z_s)\left(\partial_{F_w(z_s)}F_{\tilde{w}}\right)(z_s)\\
&\qquad = \bar{\psi}''_{z_s}(z_s)F_w(z_s)\otimes F_{\tilde{w}}(z_s) + \left(\partial_{F_w(z_s)}F_{\tilde{w}}\right)(z_s). \qquad (5.5)
\end{aligned}$$

Since $\mathbf{X}$ is a weak-geometric rough path and $\bar{\psi}''_{z_s}(z_s)$ is symmetric, we also have

$$\bar{\psi}''_{z_s}(z_s)F_w(z_s)\otimes F_{\tilde{w}}(z_s)|_{|w\otimes\tilde{w}=\mathbb{X}_{s,t}} = \frac{1}{2}\,\bar{\psi}''_{z_s}(z_s)\left[F_{x_{s,t}}(z_s)\right]^{\otimes 2},$$

which combined with (5.5) shows

$$\begin{aligned}
&\left(\partial_{F_w(z_s)}\left[\bar{\psi}'_{z_s}(\cdot)F_{\tilde{w}}(\cdot)\right]\right)(z_s)|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}}\\
&\qquad = \left(\partial_{F_w(z_s)}F_{\tilde{w}}\right)(z_s)|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}} + \frac{1}{2}\bar{\psi}''_{z_s}(z_s)\left[F_{x_{s,t}}(z_s)\right]^{\otimes 2}. \qquad (5.6)
\end{aligned}$$

Equation (5.2) now follows directly from (5.4) and (5.6).

Conversely, now assume that (5.2) holds. From (5.2) and the fact that $\mathbf{X}$ is a rough path there exists $C_1 < \infty$ such that $\left|\bar{\psi}(z_s,z_t)\right| \le C_1\,\omega(s,t)^{1/p}$. Combining this observation with (5.3) easily implies $z_{s,t} \underset{1}{\approx} 0$. Indeed, by uniform continuity, there exists a $\delta > 0$ such that if $|t-s| \le \delta$, we have

$$\begin{aligned}
|z_{s,t}| &\le |\bar{\psi}\left(z_s,z_t\right)| + \left|\frac{1}{2}\,\psi''_{z_s}(z_s)\left(z_{s,t}\right)^{\otimes 2} + C(z_s,z_t)\left(z_{s,t}\right)^{\otimes 3}\right|\\
&\le C_1\,\omega(s,t)^{1/p} + \frac{1}{2}\,|z_{s,t}|\,.
\end{aligned}$$

By using an argument similar to the proof of Theorem 2.48 we can bootstrap these local inequalities to prove the existence of a $C_2 < \infty$ such that $|z_{s,t}| \le C_2\,\omega(s,t)^{1/p}$ for $a \le s \le t \le b$.

From equations (5.3) and (5.2),

$$\begin{aligned}
z_{s,t} &= \bar{\psi}(z_s,z_t) - \frac{1}{2}\,\bar{\psi}''_{z_s}(z_s)\left(\psi\left(z_s,z_t\right)\right)^{\otimes 2} + C(z_s,z_t)\left(z_{s,t}\right)^{\otimes 3}\\
&\underset{3}{\approx} F_{x_{s,t}}(z_s) + \left(\partial_{F_w(z_s)}\left[\bar{\psi}'_{z_s}(\cdot)F_{\tilde{w}}(\cdot)\right]\right)(z_s)|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}} - \frac{1}{2}\,\bar{\psi}''_{z_s}(z_s)\left(F_{x_{s,t}}(z_s)\right)^{\otimes 2}\\
&= F_{x_{s,t}}(z_s) + \left(\partial_{F_w(z_s)}F_{\tilde{w}}\right)(z_s),
\end{aligned}$$

wherein we have used (5.6) for the last equality. □

5.1. RDEs on a manifold

We now move to the manifold case. Let $F\colon M \to L\left(W, TM\right)$ be smooth such that $F(m) \in L\left(W, T_mM\right)$. Alternatively we can think of $F\colon W \to \Gamma(TM)$ where the map $w \to F_w(\cdot)$ is linear. We wish to give meaning to the differential equation

$$d\mathbf{y}_t = F_{d\mathbf{X}_t}(y_t) \qquad (5.7)$$

with initial condition $y_0 = \bar{y}_0$. To do this, first recall that any vector field can be transferred to Euclidean space by using charts. If $\mathcal{U} \subseteq D(\phi)$ where ϕ is a chart and $\mathcal{V} := \phi(\mathcal{U})$ then

$$F^{\phi} := d\phi \circ \left(F \circ \phi^{-1}\right)$$

is a vector field on $\mathcal{V}$ (which does not carry the base point). If $\mathbf{y}_t$ is to "solve" (5.7) then $\mathbf{z}_t := \phi_* \mathbf{y}_t$ should solve the differential equation

$$d\mathbf{z}_t = F^{\phi}_{d\mathbf{X}_t}(z_t). \tag{5.8}$$

In the Euclidean case, equation (5.8) is satisfied if

$$\begin{aligned} z_t &\underset{3}{\approx} z_s + F^{\phi}_{x_{s,t}}(z_s) + \left(\partial_{F^{\phi}_w(z_s)} F^{\phi}_{\tilde{w}}\right)(z_s)|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}}, \\ z^{\dagger}_s &= F^{\phi}_{(\cdot)}(z_s). \end{aligned} \tag{5.9}$$

By writing out equation (5.9) we have

$$\begin{aligned} \phi(y_t) &\underset{3}{\approx} \phi(y_s) + d\phi \circ F_{x_{s,t}}(y_s) + \left(\partial_{d\phi\circ F_w(y_s)} d\phi \circ \left(F_{\tilde{w}} \circ \phi^{-1}\right)\right)(\phi(y_s))\,|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}} \\ &= \phi(y_s) + d\phi \circ F_{x_{s,t}}(y_s) + F_w(y_s)\left[d\phi \circ F_{\tilde{w}}\right]|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}}. \end{aligned} \tag{5.10}$$

We note that F is linear with its range in the algebra of differential operators, we can extend it uniquely to $\mathcal{F}$ which acts on the tensor algebra $T(\mathbb{R}^n)$. In that case, we may write (5.10) more concisely as

$$\phi(y_t) \underset{3}{\approx} \phi(y_s) + \left(\mathcal{F}_{\mathbf{X}_{s,t}}\phi\right)(y_s). \tag{5.11}$$

This approximation will be satisfied for our solution to a rough differential equation on a manifold. However, we will opt to define our solution in a coordinate-free but equivalent way.

Definition 5.2. A controlled rough path $\mathbf{y} = (y, y^{\dagger})$ on $I_0 = [0,T]$ or $[0,T)$ solves (5.7) if $y^{\dagger}_s = F_{(\cdot)}(y_s)$ and for every $f \in C^{\infty}(M)$ and $[a,b] \subseteq I_0$, the approximation

$$f(y_t) - f(y_s) \underset{3}{\approx} \left(\mathcal{F}_{\mathbf{X}_{s,t}} f\right)(y_s)$$

holds for $a \le s \le t \le b$.

If in addition $y_0 = \bar{y}_0$, we say $\mathbf{y}$ solves (5.7) with initial condition $y_0 = \bar{y}_0$.

While this is an intuitive definition, there are many workable characterizations of solving a rough differential equation. Before presenting a few more, we note that if $\alpha \in \Omega^1(M,V)$ and $F\colon M \to L(W, TM)$ is smooth, then the composition $\alpha \circ F_{(\cdot)}$ is a smooth map from M to V. Given $\mathbf{y} \in CRP_{\mathbf{X}}(M)$, we can then define the push-forward $\left[\alpha \circ F_{(\cdot)}\right]_* \mathbf{y} \in CRP_{\mathbf{X}}(L(W,V))$. Recall from Theorem 2.7 that we can define the integral increment

$$\int_s^t \left\langle ([\alpha \circ F_{(\cdot)}]_* \mathbf{y})_\tau, d\mathbf{X}_\tau \right\rangle.$$

With this idea in mind, we now give other characterizations of solving (5.7).

Theorem 5.3. *Let y be a path in M on I_0 with $y_s^\dagger = F_\cdot(y_s)$. Let $\mathbf{y} = (y, y^\dagger) \in CRP_{\mathbf{X}}(M)$. The following are equivalent.*

(1) *For every chart ϕ with $a, b \in I_0$ such that $y([a,b]) \subseteq D(\phi)$ the approximation*

$$\phi(y_t) \underset{3}{\approx} \phi(y_s) + d\phi \circ F_{x_{s,t}}(y_s) + F_w(y_s)\left[d\phi \circ F_{\tilde{w}}\right]|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}} \tag{5.12}$$

holds $a \le s \le t \le b$; that is,

$$\phi(y_t) - \phi(y_s) = \int_s^t \left\langle \left(\left[d\phi \circ F_{(\cdot)}\right]_* \mathbf{y}\right)_\tau, d\mathbf{X}_\tau \right\rangle$$

for $a \le s \le t \le b$.

(2) *If V is a Banach space, $\alpha \in \Omega^1(M, V)$, and $[a,b]$ is such that $[a,b] \subseteq I_0$ then*

$$\int_s^t \alpha(d\mathbf{y}) \underset{3}{\approx} \alpha\left(F_{x_{s,t}}(y_s)\right) + F_w(y_s)\left[\alpha \circ F_{\tilde{w}}\right]|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}}$$

for $a \le s \le t \le b$; that is,

$$\int_s^t \alpha(d\mathbf{y}) = \int_s^t \left\langle \left(\left[\alpha \circ F_{(\cdot)}\right]_* \mathbf{y}\right)_\tau, d\mathbf{X}_\tau \right\rangle$$

for $a \le s \le t \le b$.

(3) *$\mathbf{y}$ solves (5.7); that is,*

$$f(y_t) - f(y_s) = \int_s^t \left\langle \left(\left[df \circ F_{(\cdot)}\right]_* \mathbf{y}\right)_\tau, d\mathbf{X}_\tau \right\rangle$$

for every $f \in C^\infty(M)$.

Proof. We will only prove the approximations in each case, that is the first statement of each item. The second statements are immediate from the definitions.

(1) $\implies$ (2). We assume that $\mathbf{y}$ satisfies the approximation in (5.12) for any chart. Let $[a,b] \subseteq I_0$ be given. For every $m \in y([a,b])$, we have there exists a chart ϕ^m with open domain $\mathcal{V}_m := D(\phi^m)$ containing m whose range $R(\phi^m)$ is convex. We may now use our patching strategy outlined in Remark 2.53 with the cover $\{V_m\}_{m\in y([a,b])}$ applied to the function

$$(s,t) \longrightarrow \int_s^t \alpha(d\mathbf{y}) - \alpha\left(F_{x_{s,t}}(y_s)\right) - F_w(y_s)\left[\alpha \circ F_{\tilde{w}}\right]|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}}$$

to reduce to the case where $y([a,b])$ is contained in the domain of a single chart.

With this reduction, we can further reduce to the flat case by defining $\mathbf{z}_t := (\phi(y_t), F_\cdot(y_s))$ and $F^\phi := d\phi\left(F \circ \phi^{-1}\right)$ and showing

$$\begin{aligned}
&\int_s^t \alpha(d\mathbf{y}) - \alpha\left(F_{x_{s,t}}(y_s)\right) - F_w(y_s)\left[\alpha \circ F_{\tilde{w}}\right]|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}} \\
&\quad = \int_s^t \left((\phi^{-1})^*\alpha\right)(d\mathbf{z}) - \left((\phi^{-1})^*\alpha\right)_{z_s}\left(F^\phi_{x_{s,t}}(z_s)\right) \\
&\qquad - \left(\partial_{F^\phi_w(z_s)}\left[\left(\phi^{-1}\right)^*\alpha \circ F^\phi_{\tilde{w}}\right]\right)(z_s)|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}}.
\end{aligned}$$

The above equality is true due to the following three identities:

$$\int_s^t \alpha(d\mathbf{y}) = \int_s^t \left(\left(\phi^{-1}\right)^* \alpha\right)(d\mathbf{z}), \tag{5.13}$$

$$\alpha\left(F_{x_{s,t}}(y_s)\right) = \left(\left(\phi^{-1}\right)^* \alpha\right)_{z_s} \left(F^{\phi}_{x_{s,t}}(z_s)\right), \text{ and} \tag{5.14}$$

$$F_w(y_s)\left[\alpha \circ F_{\tilde{w}}\right] = \left(\partial_{F^{\phi}_w(z_s)}\left[\left(\phi^{-1}\right)^* \alpha \circ F^{\phi}_{\tilde{w}}\right]\right)(z_s). \tag{5.15}$$

Equation (5.13) is true by Theorem 4.15. The differential geometric identities in (5.14) and (5.15) are simply a matter of unwinding the definitions.

(2) $\implies$ (3). By letting $\alpha = df$ and using Theorem 4.13, we have

$$\begin{aligned} f(y_t) - f(y_s) &= \int_s^t df\,(d\mathbf{y}) \\ &\underset{3}{\approx} df\left(F_{x_{s,t}}(y_s)\right) + F_w(y_s)\left[df \circ F_{\tilde{w}}\right]|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}} = \left(\mathcal{F}_{\mathbf{X}_{s,t}} f\right)(y_s). \end{aligned}$$

(3) $\implies$ (1). We leave it to the reader to work through the details of this step which follow exactly as in the proof of Theorem 2.57 by letting f^i be the coordinates of ϕ. □

By Theorem 6.10 in the Appendix, we see that a solution to a rough differential equation in flat space does actually satisfy equation (5.7). Moreover, we immediately get local existence of solutions:

Theorem 5.4. *Let $F\colon W \to \Gamma(TM)$ be linear and let $\bar{y}_0$ be a point in M. There exists a local in time solution to the differential equation (5.7) with initial condition $y_0 = \bar{y}_0$.*

Proof. Let ϕ be any chart such that $\bar{y}_0 \in D(\phi)$. Then there exists a solution on some time interval $[0,\tau]$ in $R(\phi)$ to the differential equation

$$d\mathbf{z}_t = F^{\phi}_{d\mathbf{X}_t}(z_t)$$

with initial condition $z_0 = \phi(\bar{y}_0)$. If $\tilde{\phi}$ is any other chart such that $[a,b] \subseteq [0,\tau]$ and $y([a,b]) \subseteq D(\tilde{\phi})$, then the transition map $\tilde{\phi} \circ \phi^{-1}$ has a domain containing $z([a,b])$. It is easy to check that

$$F^{\tilde{\phi}} = \left(F^{\phi}\right)^{\tilde{\phi}\circ\phi^{-1}},$$

and by Corollary 6.13, after unraveling the notation, we have

$$\tilde{\phi}(y_t) \underset{3}{\approx} \tilde{\phi}(y_s) + d\tilde{\phi} \circ F_{x_{s,t}}(y_s) + F_w(y_s)\left[d\tilde{\phi} \circ F_{\tilde{w}}\right]|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}}.$$

Thus satisfying the rough differential equation approximation in one chart is sufficient prove that it hold in all charts. □

Solutions to rough differential equations will be unique on the intersection of their time domain up to some possible explosion time. This is stated more precisely in the following theorem.

Theorem 5.5. *Let $T > 0$. There is unique solution $\mathbf{y}_t \in CRP_{\mathbf{X}}(M)$ to $d\mathbf{y}_t = F_{d\mathbf{X}_t}(y_t)$ with initial condition $y_0 = \bar{y}_0$ existing either on all of $[0,T]$ or on $[0,\tau)$ for some $\tau < T$ such that the closure of $\{y_t : 0 \le t < \tau\}$ is not compact.*

Proof. This proof follows the strategy of the proof of Theorem 4.2 in [3]. First we will show that we can always concatenate a solution $\mathbf{y}$ provided it has not exploded yet:

Suppose there exists a $\mathbf{y}$ solving $d\mathbf{y}_t = F_{d\mathbf{X}_t}(y_t)$ with initial condition $y_0 = \bar{y}_0$ on $[0,\tau)$. If there exists a compact $K \subseteq M$ such that $\{y_t : 0 \le t < \tau\} \subseteq K$, then there is a sequence of increasing times $t_n \in [0,\tau)$ such that $t_n \to \tau$ and $y_\infty := \lim_{n\to\infty} y(t_n)$ exists and is in K. We can now choose a chart ϕ such that the closure of $D(\phi)$ is compact and such that $y_\infty \in D(\phi)$. Let $\mathbf{z}_t$ and a be such that $\mathbf{z}_t := \phi_*\mathbf{y}$ on some time interval $[a,\tau)$ such that $y([a,\tau)) \subseteq D(\phi)$. By appealing to Lemma 6.9 in the Appendix, there exists an $\epsilon > 0$ and a $U \subseteq D(\phi)$ containing y_∞ such that for any $s \in [\tau-\epsilon,\tau]$ and $\bar{z} \in U$, there exists $\tilde{\mathbf{z}} \in CRP_{\mathbf{X}}\left(\mathbb{R}^d\right)$ defined on $[s,\tau+\epsilon]$ which solves

$$d\tilde{\mathbf{z}}_t = F^{\phi}_{x_{s,t}}(\tilde{z}_t) \quad \text{with} \quad \tilde{z}_s = \bar{z}.$$

Letting n be sufficiently large, we have that $t_n \in [\tau-\epsilon,\tau]$ and we let $\tilde{\mathbf{z}}$ be the solution to $d\tilde{\mathbf{z}}_t = F^{\phi}_{x_{s,t}}(\tilde{z}_t)$ with initial condition $\tilde{z}_s = z(t_n)$. Then we can concatenate $\mathbf{z}$ and $\tilde{\mathbf{z}}$ in the sense of Lemma 2.6. By pulling these back to the manifold by ϕ^{-1}, we now have a solution $\tilde{\mathbf{y}}$ on M which is defined on $[0,\tau+\epsilon]$.

With the preceding fact shown, we may now prove the theorem. We define

$$\tau := \sup\{T_0 \in (0,T) : \exists \mathbf{y} \text{ solving } d\mathbf{y}_t = F_{d\mathbf{X}_t}(y_t) \text{ with } y_0 = \bar{y}_0\}.$$

We can then for any $t < \tau$ define $\mathbf{y}_t := \hat{\mathbf{y}}_t$ where $\hat{\mathbf{y}}_t$ is any solution to $d\mathbf{y}_t = F_{d\mathbf{X}_t}(y_t)$ with initial condition $y_0 = \bar{y}_0$. By the uniqueness of solutions to rough differential equations on flat space and the fact that we can cover any portion of the path with the domain of a chart, we know that $\mathbf{y}_t$ is well defined, and in fact satisfies $d\mathbf{y}_t = F_{d\mathbf{X}_t}(y_t)$ on all of $[0,\tau)$. If the closure of $\{y_t : 0 \le t < \tau\}$ is compact, then from what we showed above, we can produce a solution $\tilde{\mathbf{y}}$ which is defined on $[0,\tau+\epsilon]$ for some $\epsilon > 0$. In this case, τ must be T and $\tilde{\mathbf{y}}|_{[0,T]}$ is a solution defined on all of $[0,T]$. □

Definition 5.6. Let $f\colon M \to N$ be a smooth map between manifolds. Let $F\colon W \to \Gamma(TM)$ and $\tilde{F}\colon W \to \Gamma(TN)$ be linear. We say F and $\tilde{F}$ are **f-related dynamical systems** if

$$f_*F_w = \tilde{F}_w \circ f \quad \text{for all } w \in W.$$

As in the flat case and shown in the Appendix in Theorem 6.12, we have a relation between dynamical systems. The proof is no different in the manifold case, and so we omit it.

Theorem 5.7. *Suppose $f\colon M \to N$ is a smooth map between manifolds and let $F\colon W \to \Gamma(TM)$ and $\tilde{F}\colon W \to \Gamma(TN)$ be f-related dynamical systems. If $\mathbf{y}$ solves the initial value problem equation (5.7), then $\tilde{\mathbf{y}}_t := \left(\tilde{y}_t, \tilde{y}_s^\dagger\right) := f_*\tilde{\mathbf{y}}_t$ solves*

$$d\tilde{\mathbf{y}}_t = \tilde{F}_{d\mathbf{X}_t}\left(\tilde{y}_t\right) \quad \textit{with} \quad \tilde{y}_0 = f\left(\bar{y}_0\right).$$

5.1.1. RDEs from the gauge perspective. Following the theme of Theorem 2.45, we also have a way to view a solution to a differential equation using the gauge perspective. Let ψ be a logarithm on M with diagonal domain $\mathcal{D}$.

Theorem 5.8. *Let y be a path in M on I_0 with $y_s^\dagger = F_\cdot(y_s)$. Let $\mathbf{y} = (y, y^\dagger)$. Then $\mathbf{y}$ solves (5.7) if and only if for every a, b such that $[a,b] \subseteq I_0$, there exists a $\delta > 0$ such that*

$$\psi(y_s, y_t) \underset{3}{\approx} F_{x_{s,t}}(y_s) + F_w(y_s)\left[\left(\psi_{y_s}\right)_* F_{\tilde{w}}\right]|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}}, \tag{5.16}$$

provided $a \le s \le t \le b$ and $t - s < \delta$.

Proof. This proof will be similar to the proof of Theorem 2.45.

First we show the condition of Theorem 5.8 implies that $\mathbf{y}$ solves (5.7). Let ϕ be a chart and let $[a,b]$ be such that $y([a,b]) \subseteq D(\phi)$. By defining

$$\begin{aligned} z_s &:= \phi(y_s), \\ \psi^\phi(x,y) &:= \phi_*\psi\left(\phi^{-1}(x), \phi^{-1}(y)\right), \\ F_w^\phi(x) &:= d\phi\left(F_w\left(\phi^{-1}(x)\right)\right), \end{aligned}$$

and denoting $\psi^\phi(x,y) = \left(x, \bar{\psi}^\phi(x,y)\right)$, (5.16), once pushed forward by ϕ, can be written as

$$\bar{\psi}^\phi(z_s, z_t) \underset{3}{\approx} F_{x_{s,t}}^\phi(z_s) + \left(\partial_{F_w^\phi(z_s)}\left[\bar{\psi}_{z_s}^{\phi\prime}(\cdot)F_{\tilde{w}}^\phi(\cdot)\right]\right)(z_s)|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}}$$

provided $a \le s \le t \le b$ and $t - s < \delta$. We then must prove that z solves equation (5.9) for all $a \le s \le t \le b$. However, by appealing to Lemma 2.49 and Lemma 6.14 of the Appendix, we only need to prove (5.9) holds for every u in $[a,b]$ for $s \le t$ in $(u - \delta_u, u + \delta_u) \cap [a,b]$ for some δ_u. We do this now.

For any $u \in [a,b]$, let $\mathcal{W}_u$ be an open convex set of z_u such that $\mathcal{W}_u \times \mathcal{W}_u \subseteq D(\psi^\phi)$. We then choose $\delta_u > 0$ such that $z\left([u - \delta_u, u + \delta_u] \cap [a,b]\right) \subseteq \mathcal{W}_u$ and $2\delta_u \le \delta$. We are now in the setting of Proposition 5.1 and have therefore shown $\mathbf{y}$ solves equation (5.7).

For the reverse implication, let $[a,b] \subseteq I_0$ be given. Choose $\delta > 0$ such that $|t - s| \le \delta$ for $a \le s \le t \le b$ implies that $|\psi(y_s, y_t)|_g$ is bounded. Around every point m of $y([a,b])$, there exists an open $\mathcal{O}_m$ containing m such that $\mathcal{O}_m \times \mathcal{O}_m \subseteq \mathcal{D}$. Additionally for each m there exists a chart ϕ^m such that $m \in D\left(\phi^m\right)$, $D\left(\phi^m\right) \subseteq \mathcal{O}_m$, and $\mathcal{W}_m := R\left(\phi^m\right)$ is convex. We may now use Remark 2.53 with the cover $\{\mathcal{V}_m\}_{m\in y([a,b])}$ and $D = \{(s,t) : a \le s \le t \le b \text{ and } |t - s| \le \delta\}$ with the function

$$(s,t) \longrightarrow \psi(y_s, y_t) - F_{x_{s,t}}(y_s) - F_w(y_s)\left[\left(\psi_{y_s}\right)_* \circ F_{\tilde{w}}\right]|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}}.$$

Doing this, we have reduced to considering the case of our path being contained in the domain of a single chart ϕ such that $D(\phi) \times D(\phi) \subseteq \mathcal{D}$ and $R(\phi)$ is convex. By using the same definitions above for z_s, F^ϕ, and ψ^ϕ, we reduce proving

$$\psi(y_s, y_t) \underset{3}{\approx} F_{x_{s,t}}(y_s) + F_w(y_s)\left[(\psi_{y_s})_* \circ F_{\tilde{w}}\right]|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}}$$

to the flat case

$$\bar{\psi}^\phi(z_s, z_t) \underset{3}{\approx} F^\phi_{x_{s,t}}(z_s) + \left(\partial_{F^\phi_w(z_s)}\left[\bar{\psi}^{\phi\prime}_{z_s}(\cdot)F^\phi_{\tilde{w}}(\cdot)\right]\right)(z_s)|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}}.$$

This is now in the setting of Proposition 5.1 and hence we are finished. □

Akin to the integral formulas, there is also a characterization of solving a differential equation which involves a gauge (ψ, U).

Theorem 5.9. $\mathbf{y} = (y, y^\dagger)$ *on* I_0 *solves* (5.7) *if and only if* $y^\dagger_s = F_{(\cdot)}(y_s)$ *and for all* $[a, b] \subseteq I_0$*, there exists a* $\delta > 0$ *such that* $|t - s| \leq \delta$*, and* $a \leq s \leq t \leq b$ *implies*

$$\begin{aligned}\psi(y_s, y_t) &\underset{3}{\approx} F_{x_{s,t}}(y_s)\\ &\quad + \left(-S^{\psi_*,U}_{y_s}\left[F_w(y_s)\otimes F_{\tilde{w}}(y_s)\right] + F_w(y_s)\left[U(y_s,\cdot)F_{\tilde{w}}\right]\right)|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}}.\end{aligned}$$

Proof. This follows immediately from the product rule:

$$\begin{aligned}F_w(y_s)\left[(\psi_{y_s})_* F_{\tilde{w}}\right] &= F_w(y_s)\left[(\psi_{y_s})_{*(\cdot)} U(y_s,\cdot)^{-1}U(y_s,\cdot)F_{\tilde{w}}\right]\\ &= -S^{\psi_*,U}_{y_s}\left[F_w(y_s)\otimes F_{\tilde{w}}(y_s)\right] + F_w(y_s)\left[U(y_s,\cdot)F_{\tilde{w}}\right]\square\end{aligned}$$

Example 5.10. If ∇ is a covariant derivative, then $\mathbf{y}$ on I_0 solves (5.7) if and only if $y^\dagger_s = F(y_s)$ and

$$\exp^{-1}_{y_s}(y_t) \underset{3}{\approx} F_{x_{s,t}}(y_s) + \left(\nabla_{F_w(y_s)}F_{\tilde{w}}\right) - \frac{1}{2}T^\nabla\left[F_w(y_s)\otimes F_{\tilde{w}}(y_s)\right]|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}}$$

for s and t close.

6. Appendix

6.1. Why logarithms

By definition, a manifold M is a topological space which locally "looks" like Euclidean space. The locally looks like statement means there are charts which allow us to identify sufficiently small open subsets of M with open subsets of $\mathbb{R}^d$. The manifold is smooth means these charts are consistent in the sense that the identification of open sets of M with those of $\mathbb{R}^d$ by two different charts are smoothly related to one another. Because of this very definition, at the end of the day, one has to use charts to make definitions and perform calculations.

Example 6.1. Suppose that α is a one-form on M and $\sigma(s)$ is a smooth path in M. One may think that no charts are involved in computing

$$\int_\sigma \alpha = \int_0^1 \alpha\left(\sigma'(s)\right) ds,$$

but this is not the case! In fact charts were used to define what it means for $\sigma(s)$ to be smooth, the notion of the tangent and cotangent bundles (hence the definition of α), and the very meaning of $\sigma'(s)$. When all is said and done,

$$\alpha\left(\sigma'(s)\right) := \alpha^{\phi}_{(\phi\circ\sigma(s))}\left(\phi\circ\sigma\right)'(s), \tag{6.1}$$

where

$$\alpha^{\phi}_x v := \alpha\left(\phi_*^{-1}\left(v_x\right)\right) \tag{6.2}$$

is the one-form α "read" in this chart. It is then verified that the definitions have been arranged so that the right side of (6.1) is independent of the choice of chart ϕ. Thus (assuming $\sigma\left([0,1]\right)$ is contained in the domain of ϕ for simplicity),

$$\int_\sigma \alpha := \int_0^1 \alpha^{\phi}_{(\phi\circ\sigma(s))}\left(\phi\circ\sigma\right)'(s) ds$$

is in fact independent of the chart, ϕ.

Example 6.2. Let us now suppose that $\sigma(s)$ is only α-Hölder continuous (as tested in a chart) for some $\alpha > 1/2$ (to keep it simple). In this case $\sigma'(s)$ need not exist. Nevertheless we would still like to define

$$\int_0^1 \alpha\left(d\sigma(s)\right).$$

It is natural (assuming $\sigma\left([0,1]\right)$ is contained in the domain of ϕ), to **define**

$$\int_0^1 \alpha\left(d\sigma(s)\right) := \int_0^1 \alpha^{\phi}_{(\phi\circ\sigma(s))}\, d\left(\phi\circ\sigma\right)(s), \tag{6.3}$$

where the latter integral is now the Young's integral. Of course we must now show the above definition is independent of the chart ϕ used in its description. Let us expand on this point a bit to see how logarithms enter.

Using the definition of the Young's integral, (6.3) may be written more explicitly as

$$\int_0^1 \alpha\left(d\sigma(s)\right) := \lim_{|\Pi|\to 0}\sum_i \alpha^{\phi}_{(\phi\circ\sigma(s_{i-1}))}\left[\phi\circ\sigma\left(s_i\right) - \phi\circ\sigma\left(s_{i-1}\right)\right], \tag{6.4}$$

where $\Pi = \{0 = s_0 < \cdots < s_n = 1\}$ is a partition of $[0,1]$ and $|\Pi|$ is the mesh size of Π. From (6.2), we have

$$\alpha^{\phi}_{(\phi\circ\sigma(s_{i-1}))}\left[\phi\circ\sigma\left(s_i\right) - \phi\circ\sigma\left(s_{i-1}\right)\right] = \alpha\big(\phi_*^{-1}\left[\phi\circ\sigma\left(s_i\right) - \phi\circ\sigma\left(s_{i-1}\right)\right]_{\phi\circ\sigma(s_{i-1})}\big).$$

If we then define

$$\psi^{\phi}(m,n) := \phi_*^{-1}\left[\phi(n)-\phi(m)\right]_{\phi(m)}$$

for all m,n in the domain of ϕ, we can rewrite (6.4) as

$$\int_0^1 \alpha\left(d\sigma(s)\right) := \lim_{|\Pi|\to 0}\sum_i \alpha\left(\psi^{\phi}\left(\sigma\left(s_{i-1}\right),\sigma\left(s_i\right)\right)\right).$$

The function $\psi^{\phi}(m,n)\in T_mM$ is an example of a logarithm associated to a chart ϕ. In this Young's setting, the fact that the definition in (6.3) is well defined (i.e., chart independent) is equivalent to the statement that

$$\lim_{|\Pi|\to 0}\sum_i \alpha\left(\psi\left(\sigma\left(s_{i-1}\right),\sigma\left(s_i\right)\right)\right) \tag{6.5}$$

is independent of the choice of logarithm, ψ.

Conclusions:

1. Independent of whether σ is smooth or not, one must use charts or equivalently logarithms to define $\int_\sigma \alpha$.
2. The fact that $\int_\sigma \alpha$ is well defined (i.e., chart independent) can be related to the assertion that an expression for $\int_\sigma \alpha$ as in (6.5) (or a more elaborate counterpart such as the formula found in Corollary 4.5) is independent of the choice of logarithm.

6.2. Taylor expansion on a Riemannian manifold

Let (M,g) be a Riemannian manifold, ∇ be the Levi-Civita covariant derivative, $\exp(tv)$ be the geodesic flow, and $//_t(\sigma)$ denote parallel translation relative to ∇. Recall that Taylor's formula with integral remainder states for any smooth function g on $[0,1]$, that

$$G(1)=\sum_{k=0}^{n}\frac{1}{k!}G^{(k)}(0)+\frac{1}{n!}\int_0^1 G^{(n+1)}(t)(1-t)^n dt. \tag{6.6}$$

We now apply this result to $G(t):=f\left(\exp_m(tv)\right)$ where $f\in C^\infty(M)$, $v\in T_mM$ and $m\in M$. To this end let $\sigma(t):=\exp(tv)$ so that $\nabla\dot\sigma(t)/dt=0$. It then follows that

$$\begin{aligned}
\dot G(t) &= df\left(\dot\sigma(t)\right)=df_{\sigma(t)}\left(\dot\sigma(t)\right),\\
\ddot G(t) &= \frac{d}{dt}df_{\sigma(t)}\left(\dot\sigma(t)\right)=\left(\nabla_{\dot\sigma(t)}df\right)\left(\dot\sigma(t)\right)+df_{\sigma(t)}\left(\frac{\nabla}{dt}\dot\sigma(t)\right)\\
&= \left(\nabla_{\dot\sigma(t)}df\right)\left(\dot\sigma(t)\right)=\left(\nabla df\right)\left(\dot\sigma(t)\otimes\dot\sigma(t)\right)\\
&\vdots
\end{aligned}$$

$$G^{(k)}(t)=\left(\nabla^{k-1}df\right)\left(\dot\sigma(t)^{\otimes k}\right)=\left(\nabla^{k-1}df\right)\Big(\overbrace{\dot\sigma(t)\otimes\cdots\otimes\dot\sigma(t)}^{k\text{ times}}\Big). \tag{6.7}$$

Therefore we may conclude that

$$f\left(\exp_m(v)\right) = G(1) = \sum_{k=0}^{n} \frac{1}{k!} G^{(k)}(0)$$

$$= f(x) + \sum_{k=1}^{n} \frac{1}{k!} \left(\nabla^{k-1} df\right)\left(v^{\otimes k}\right) + \frac{1}{n!} \int_0^1 \left(\nabla^n df\right)\left(\dot{\sigma}(t)^{\otimes(n+1)}\right)(1-t)^n dt. \tag{6.8}$$

Letting $n = \exp_m(v)$ in this formula then gives the following version of Taylor's theorem on a manifold.

Theorem 6.3. *Let $f \in C^\infty(M)$ and $m, n \in M$ with $d_g(m,n)$ sufficiently small so that there exists a unique $v \in T_m M$ such that $|v|_{g_m} \le d(m,n)$ and $n = \exp_m(v)$. Then we have*

(6.9)

$$f(n) = f(m) + \sum_{k=1}^{n} \frac{1}{k!} \left(\nabla^{k-1} df\right)\left(v^{\otimes k}\right) + \frac{1}{n!} \int_0^1 \left(\nabla^n df\right)\left(\dot{\sigma}(t)^{\otimes(n+1)}\right)(1-t)^n dt$$

$$= f(m) + \sum_{k=1}^{n} \frac{1}{k!} (\nabla^{k-1} df)\left([\exp_m^{-1}(n)]^{\otimes k}\right) + \frac{1}{n!} \int_0^1 (\nabla^n df)\left(\dot{\sigma}(t)^{\otimes(n+1)}\right)(1-t)^n dt,$$

where $\sigma(t) = \exp_m(tv)$. In particular, since $|\dot{\sigma}(t)|_g = |v|_g = d_g(m,n)$, it follows that

$$f(n) = f(m) + \sum_{k=1}^{n} \frac{1}{k!} \left(\nabla^{k-1} df\right)\left(\left[\exp_m^{-1}(n)\right]^{\otimes k}\right) + O\left(d(m,n)^{n+1}\right). \tag{6.10}$$

Lemma 6.4. *Let M be an embedded submanifold of $W = \mathbb{R}^k$ and $P(m) : W \to T_m M$ be orthogonal projection onto the tangent space. If $m, n \in M$ are close, then*

(1) $P(m)[\exp_m^{-1}(n) - (n-m)] = O(|n-m|^3)$.
Moreover, $\exp_m^{-1}(n) - (n-m) = O(|n-m|^2)$

(2) $U^\nabla(n,m) = P(m) + dP(\exp_m^{-1}(n)) + O(|n-m|^2) = P(n) + O(|n-m|^2)$

(3) $P(n) - P(m) = dP(\exp_m^{-1}(n)) + O(|n-m|^2)$.
Here $U^\nabla(n,m)$ refers to the parallelism defined in Example 2.19.

Proof. We will denote $v := \exp_m^{-1}(n) \in T_m M$ and $\sigma(t) = \exp_m(tv)$.

For (1), we have by Taylor expansion on manifolds (Theorem 6.3) that

$$G(n) = G(m) + dG(v) + \frac{1}{2}(\nabla dG)(v \otimes v) + \frac{1}{2} \int_0^1 \left(\nabla^2 dG\right)\left(\dot{\sigma}(t)^{\otimes 3}\right)(1-t)^2 dt,$$

where $G \in C^\infty(M, W)$. Letting $G(m) = m$ as a function into W, we have

$$n = m + \exp_m^{-1}(n) + \frac{1}{2}(\nabla P)(v \otimes v) + O(|v|_g^3).$$

Rearranging, we have

$$\exp_m^{-1}(n) - (n-m) = -\frac{1}{2}\,(\nabla P)\,(v \otimes v) + O(|v|_g^3), \tag{6.11}$$

so that

$$P(m)\left[\exp_m^{-1}(n) - (n-m)\right] = -\frac{1}{2}\,P(m)\,(\nabla P)\,(v \otimes v) + O(|v|_g^3).$$

Note that $(\nabla P)\,(v \otimes v) = dP\,(v)\,v = dP\,(v)\,P(m)v$. Using the identities $dPQ - PdQ = 0$ and $dP = -dQ$, where $Q = I - P$, we get that $PdPP = 0$. Thus we have

$$P(m)[\exp_m^{-1}(n) - (n-m)] = O(|v|^3).$$

Lastly, in a small neighborhood around m, $|v|_g = |m-n| + o(|m-n|)$, so that

$$P(m)[\exp_m^{-1}(n) - (n-m)] = O(|n-m|^3).$$

The fact that $\exp_m^{-1}(n) - (n-m) = O(|n-m|^2)$ is immediate from (6.11).

For (3), we use Taylor's theorem again this time with $G\colon M \to L(W,W)$ defined by $G(n) := P(n)$ to see that

$$P(n) - P(m) = dP(\exp_m^{-1}(n)) + O(|v|^2).$$

As before, this is equivalent to $P(n) - P(m) = dP(\exp_m^{-1}(n)) + O(|m-n|^2)$.

Lastly, for (2), Taylor applied to $G_m : M \to L(T_mM, \mathbb{R}^N)$ defined by $G_m(n) = U^\nabla(n,m)$ gives

$$U^\nabla(n,m) - P(m) = dG_m(\exp_m^{-1}(n)) + O(|m-n|^2).$$

But

$$\begin{aligned} dG_m(\exp_m^{-1}(n)) &= \frac{d}{dt}|_0\, U(\sigma(t),m) = -dQ(\dot\sigma(t))|_0 \\ &= -dQ(\exp_m^{-1}(n)) = dP(\exp_m^{-1}(n)). \end{aligned}$$

Thus we have

$$U^\nabla(n,m) = P(m) + dP(\exp_m^{-1}(n)) + O(|m-n|^2)$$

which is the first equality of (2). The second equality follows trivially from this and (3). □

6.3. Equivalence of Riemannian metrics on compact sets

Proposition 6.5. *Let $\pi\colon E \to N$ be a real rank $d < \infty$ vector bundle over a finite dimensional manifold N. Further suppose that E is equipped with smoothly varying fiber inner product g and let $S_g := \{\xi \in E : g\,(\xi,\xi) = 1\}$ be a sub-bundle of E. Then for any compact $K \subseteq N$, $\pi^{-1}\,(K) \cap S_g$ is a compact sets.*

Proof. We wish to show that every sequence $\{\xi_l\}_{l=1}^\infty \subset \pi^{-1}\,(K) \cap S_g$ has a convergent subsequence. Since $\{\pi\,(\xi_l)\}_{l=1}^\infty$ is a sequence in K, by passing to a subsequence

if necessary we may assume that $m := \lim_{l\to\infty} \pi(\xi_l)$ exists in K. By passing to a further subsequence if necessary we may assume that $\{\xi_l\}_{l=1}^{\infty} \in \pi^{-1}(K_0) \cap S_g$, where K_0 is a compact neighborhood of m which is contained in an open neighborhood U over which E is trivializable and hence we may now assume that $\pi^{-1}(U) = U \times \mathbb{R}^d$ and that $\xi_l = (n_l, v_l)$ where $\lim_{l\to\infty} n_l = m \in K_0$.

Let S^{d-1} denote the standard Euclidean unit sphere inside of $\mathbb{R}^d$. The function, $F\colon U \times S^{d-1} \to (0,\infty)$ defined by $F(n,v) = g((n,v),(n,v))$ is smooth and hence has a minimum $c > 0$ and a maximum, $C < \infty$ on the compact set, $K \times S^{d-1}$. Therefore by a simple scaling argument we conclude that

$$c|v|^2 \le g((n,v),(n,v)) \le C|v|^2 \quad \forall\, n \in K \text{ and } v \in \mathbb{R}^d. \tag{6.12}$$

From the lower bound in inequality (6.12) and the assumption that $1 = g(\xi_l, \xi_l)$ it follows that $|v_l|_{\mathbb{R}^d} \le 1/\sqrt{c}$ for all l and therefore has a convergent sub-sequence $\{v_{l_k}\}_{k=1}^{\infty}$. This completes the proof as $\{\xi_{l_k} = (n_{l_k}, v_{l_k})\}_{k=1}^{\infty}$ is convergent as well. □

Corollary 6.6. *If $g, \tilde{g}$ are two Riemannian metrics on TM, $K \subseteq M$ is compact, then there exists $0 < c_K, C_K < \infty$ such that*

$$c_K |v|_{\tilde{g}_m} \le |v|_{g_m} \le C_K |v|_{\tilde{g}_m} \quad \forall\, v \in \pi^{-1}(K). \tag{6.13}$$

In other words, all Riemannian metrics are equivalent when restricted to compact subsets, $K \subset M$.

Proof. The function, $F\colon TM \to [0,\infty)$, defined by $F(v) := g(v,v)$ is smooth and positive when restricted to $S_{\tilde{g}} \cap \pi^{-1}(K)$ which is compact by Proposition 6.5. Therefore there exists $0 < c_K < C_K < \infty$ such that $c_K^2 \le g(v,v) \le C_K^2$ for all $v \in S_{\tilde{g}} \cap \pi^{-1}(K)$, from which inequality (6.13) follows by a simple scaling argument. □

6.4. Covariant derivatives on Euclidean space

On $\mathbb{R}^d$ every covariant derivative takes the form $\nabla_{(x,v)} = \partial_v + A_x\langle v\rangle$ where $A\colon \mathbb{R}^d \to L\left(\mathbb{R}^d, L\left(\mathbb{R}^d, \mathbb{R}^d\right)\right)$. If $\sigma_x^v(t) = \exp_x(tv)$ where $\exp = \exp^{\nabla}$, we have, by definition,

$$\begin{aligned} \partial_{\dot\sigma_x^v(t)}\dot\sigma_x^v &= -A_{\sigma_x^v(t)}\langle \dot\sigma_x^v(t)\rangle\, \dot\sigma_x^v(t), \\ \dot\sigma_x^v(0) &= v, \\ \sigma_x^v(0) &= x. \end{aligned}$$

In particular, if $f_x = \exp_x(\cdot)$ plugging in at $t = 0$ we get

$$f_x''(0)[v \otimes v] = -A_x\langle v\rangle\, v.$$

Now if we denote $G_x := \exp_x^{-1}(\cdot)$ and by differentiating $f_x \circ G_x$ twice, we get that

$$G_x''(x)[v \otimes v] = A_x\langle v\rangle\, v.$$

Indeed we have

$$\begin{aligned} 0 &= (f_x \circ G_x)''(x) = [f_x'(G_x(x))\, G_x'(x)]' \\ &= f_x''(G_x(x))\,[G_x'(x) \otimes G_x'(x)] + f_x'(G_x(x))\, G_x''(x). \end{aligned}$$

Since $G_x(x) = 0$, $G_x'(x) = I$, and $f_x'(0) = I$ we have

$$f_x''(0) = -G_x''(x).$$

Parallel translation $U^\nabla(\sigma_x^v(t), x)$ solves

$$\begin{aligned} \frac{d}{dt} U^\nabla(\sigma_x^v(t), x) &= -A_{\sigma_x^v(t)} \langle \dot\sigma_x^v(t) \rangle\, U^\nabla(\sigma_x^v(t), x), \\ U^\nabla(x, x) &= I. \end{aligned}$$

Again, using $t = 0$ we have that if $\tilde G_x = U^\nabla(\cdot, x)$, then

$$\tilde G_x'(x) v = -A_x \langle v \rangle.$$

To summarize, we have

$$\left(\exp_x^{-1}\right)''(x)\,[v \otimes v] = A_x \langle v \rangle\, v \tag{6.14}$$

and

$$\left(U^\nabla(\cdot, x)\right)'(x) v = -A_x \langle v \rangle.$$

Since $\left(\exp_x^{-1}\right)''(x)$ is symmetric, we have that

$$\begin{aligned} \left(\exp_x^{-1}\right)''(x)\,[v \otimes w] &= \frac{1}{2}\left(\exp_x^{-1}\right)''(x)\,(v \otimes w + w \otimes v) \\ &\quad + \frac{1}{2}\left(\exp_x^{-1}\right)''(x)\,(v \otimes w - w \otimes v) \\ &= \frac{1}{2}\left(\exp_x^{-1}\right)''(x)\,(v \otimes w + w \otimes v) \\ &= \frac{1}{2} A_x (v \otimes w + w \otimes v) = \frac{1}{2}\left(A_x \langle v \rangle\, w + A_x \langle w \rangle\, v\right). \end{aligned} \tag{6.15}$$

Another way of saying this is that $\left(\exp_x^{-1}\right)''(x)$ equals the symmetric part of A_x. By using this fact and Taylor's theorem, we get the following result.

Lemma 6.7. *If $\nabla_{(x,v)} = \partial_v + A_x \langle v \rangle$ is a covariant derivative on $\mathbb{R}^d$, then*

$$\left(\exp_x^\nabla\right)^{-1}(y) - (y - x) - \frac{1}{2} A_x \langle y - x \rangle \langle y - x \rangle = O\big(|y - x|^3\big),$$

$$U^\nabla(y, x) - I + A_x \langle y - x \rangle = O\big(|y - x|^2\big), \tag{6.16}$$

where $|x - y|$ is small enough for these terms to make sense.

Corollary 6.8. *If $\nabla_{(x,v)} = \partial_v + A_x \langle v \rangle$ is a covariant derivative on $\mathbb{R}^d$, then*

$$U^{\nabla}(y,x) - I - A_y \langle x - y \rangle = O(|y-x|^2),$$

where $|x-y|$ is small enough for these terms to make sense. In particular, we have

$$(U^{\nabla}(x,\cdot))'(x)\, v = A_x \langle v \rangle.$$

Proof. This is immediate after expanding $A_{(\cdot)}$ about x in the direction $y - x$ in (6.16) with Taylor's theorem. □

6.5. Rough differential equation results in Euclidean space

The following lemma (which is Corollary 2.17 in [3] and was proved using Theorem 10.14 of [15]) proves useful in the manifold case.

Lemma 6.9. *Let $U \subseteq \mathbb{R}^d$ be an open set and U_1 be a precompact open set whose closure is contained in U. There exists a $\delta > 0$ such that for all $(\bar{z}_0, t_0) \in U_1 \times [0,T]$, the rough differential equation*

$$d\mathbf{z}_t = F_{d\mathbf{X}_t}(z_t) \quad \text{with} \quad z_{t_0} = \bar{z}_0$$

has a unique solution $\mathbf{z} \in CRP_{\mathbf{X}}(\mathbb{R}^d)$ which is defined on $[t_0, t_0 + \delta \wedge T]$, with $z_t \in U$ for all $t \in [t_0, t_0 + \delta \wedge T]$.

We now state an equivalent condition for the path $\mathbf{z}$ to solve equation (2.11).

Theorem 6.10. *Let $U \subseteq \mathbb{R}^d$ be open such and $\mathbf{z} = (z, z^{\dagger}) \in CRP_{\mathbf{X}}(\mathbb{R}^d)$ defined on I_0 such that $z(I_0) \subseteq U$. Then $\mathbf{z}$ solves (2.11) if and only if $z_s^{\dagger} = F_{\cdot}(z_s)$ and for every $[a,b] \subseteq I_0$, Banach space V, and $\alpha \in \Omega^1(U,V)$, the approximation*

$$\int_s^t \alpha(d\mathbf{z}) \underset{3}{\approx} \alpha_{z_s}\left(F_{x_{s,t}}(z_s)\right) + \left(\partial_{F_w(z_s)}\left[\alpha \circ F_{\tilde{w}}\right]\right)(z_s)|_{w \otimes \tilde{w} = \mathbb{X}_{s,t}}$$

holds.

Proof. This is proved in Theorem 4.5 in [3] by letting $M = U$ but included here for completeness. To prove the "if" direction, it suffices to let $\alpha = d(I_U)$ and notice that

$$\int_s^t d(I_U)(d\mathbf{z}) = z_t - z_s$$

by Theorem 4.13, and that $d(I_U)_u(\tilde{u}) = \tilde{u}$ so that

$$d(I_U)_{z_s}\left(F_{x_{s,t}}(z_s)\right) = F_{x_{s,t}}(z_s)$$

and

$$\left(\partial_{F_w(z_s)}\left[d(I_U) \circ F_{\tilde{w}}\right]\right)(z_s) = \left(\partial_{F_w(z_s)} F_{\tilde{w}}\right)(z_s).$$

To prove the "only if" direction, by definition we have

$$z_{s,t} \underset{3}{\approx} F_{x_{s,t}}(z_s) + \left(\partial_{F_w(z_s)} F_{\tilde{w}}\right)(z_s)|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}}$$

and

$$\int_s^t \alpha(d\mathbf{z}) \underset{3}{\approx} \alpha_{z_s}(z_{s,t}) + \alpha'_{z_s}\left(F_\cdot(z_s) \otimes F_\cdot(z_s)\mathbb{X}_{s,t}\right).$$

Combining these approximations, we have

$$\begin{aligned}\int_s^t \alpha(d\mathbf{z}) &\underset{3}{\approx} \alpha_{z_s}(z_{s,t}) + \alpha'_{z_s}\left(F_\cdot(z_s) \otimes F_\cdot(z_s)\mathbb{X}_{s,t}\right)\\ &\underset{3}{\approx} \alpha_{z_s}\left(F_{x_{s,t}}(z_s) + \left(\partial_{F_w(z_s)} F_{\tilde{w}}\right)(z_s)\right) + \alpha'_{z_s}\left(F_w(z_s) \otimes F_{\tilde{w}}(z_s)\right)|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}}\\ &= \alpha_{z_s}\left(F_{x_{s,t}}(z_s)\right) + \left(\partial_{F_w(z_s)}\left[\alpha \circ F_{\tilde{w}}\right]\right)(z_s)|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}},\end{aligned}$$

where the last equality follows from the calculation

$$\begin{aligned}\left(\partial_{F_w(z_s)}\left[\alpha \circ F_{\tilde{w}}\right]\right)(z_s) &= \left(\partial_{F_w(z_s)}\left[\alpha_{z_s} \circ F_{\tilde{w}}(\cdot)\right]\right)(z_s) + \left(\partial_{F_w(z_s)}\alpha_{(\cdot)} \circ F_{\tilde{w}}(z_s)\right)(z_s)\\ &= \alpha_{z_s}\left(\left(\partial_{F_w(z_s)} F_{\tilde{w}}\right)(z_s)\right) + \alpha'_{z_s}\left(F_w(z_s) \otimes F_{\tilde{w}}(z_s)\right)\end{aligned}$$

□

Theorem 6.12 below is useful in showing that a solution to an RDE in the flat case satisfies our manifold Definition 5.2. Let U and $\tilde{U}$ be open sets for the remainder of this subsection.

Definition 6.11. Let $f\colon U \subseteq \mathbb{R}^d \to \tilde{U} \subseteq \mathbb{R}^{\tilde{d}}$ be a smooth map. Let $F\colon U \to L(W, \mathbb{R}^d)$ and $\tilde{F}\colon \tilde{U} \to L(W, \mathbb{R}^{\tilde{d}})$ be smooth. We say F and $\tilde{F}$ are **f-related dynamical systems** if

$$f'(x)F_w(x) = \tilde{F}_w \circ f(x) \text{ for all } w \in W.$$

Theorem 6.12. *Suppose $f\colon U \subseteq \mathbb{R}^d \to \tilde{U} \subseteq \mathbb{R}^{\tilde{d}}$ is a smooth map and let $F\colon U \to L(W, \mathbb{R}^d)$ and $\tilde{F} : \tilde{U} \to L(W, \mathbb{R}^{\tilde{d}})$ be f-related dynamical systems. If $\mathbf{z}$ solves*

$$d\mathbf{z}_t = F_{d\mathbf{X}_t}(z_t)$$

with initial condition $z_0 = \bar{z}_0$, then $\tilde{\mathbf{z}}_t := \left(\tilde{z}_t, \tilde{z}_s^\dagger\right) := f_*\mathbf{z}_t$ solves

$$d\tilde{\mathbf{z}}_t = \tilde{F}_{d\mathbf{X}_t}(\tilde{z}_t)$$

with initial condition $\tilde{z}_0 = f(\bar{z}_0)$.

Proof. We have, by letting $\alpha := df$ in Theorem 6.10,

$$\begin{aligned}\tilde{z}_{s,t} &= f(z_t) - f(z_s)\\ &\underset{3}{\approx} f'(z_s)F_{x_{s,t}}(z_s) + \partial_{F_w(z_s)}\left[f'(\cdot)F_{\tilde{w}}(\cdot)\right](z_s)|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}}\\ &\underset{3}{\approx} \tilde{F}_{x_{s,t}}(\tilde{z}_s) + \left(\partial_{F_w(z_s)}\tilde{F}_{\tilde{w}} \circ f\right)(z_s)|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}}\\ &\underset{3}{\approx} \tilde{F}_{x_{s,t}}(\tilde{z}_s) + \tilde{F}'_{\tilde{w}}\left(f(z_s)\right)f'(z_s)F_w(z_s)|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}}\\ &\underset{3}{\approx} \tilde{F}_{x_{s,t}}(\tilde{z}_s) + \tilde{F}'_{\tilde{w}}\left(f(z_s)\right)\tilde{F}_w \circ f(z_s)|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}}\\ &\underset{3}{\approx} \tilde{F}_{x_{s,t}}(\tilde{z}_s) + \left(\partial_{\tilde{F}_w(\tilde{z}_s)}\tilde{F}_{\tilde{w}}\right)(\tilde{z}_s)|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}}.\end{aligned}$$

Additionally,

$$\tilde{z}_t^\dagger = f'(z_t)\, z_t^\dagger = f'(z_t)\, F_{(\cdot)}(z_t) = \tilde{F}_{(\cdot)}(\tilde{z}_t)\,. \qquad \square$$

Corollary 6.13. *Let $\phi\colon U \subseteq \mathbb{R}^d \to \tilde{U} \subseteq \mathbb{R}^d$ be a diffeomorphism with $\phi(z(I_0)) \subseteq U$. Then $\mathbf{z}$ on I_0 solves*

$$d\mathbf{z}_t = F_{d\mathbf{X}_t}(z_t)$$

with initial condition $z_0 = \bar{z}_0$ if and only if $\tilde{\mathbf{z}} := \phi_* \mathbf{z}$ on I_0 solves

$$d\tilde{\mathbf{z}}_t = F^{\phi}_{d\mathbf{X}_t}(\tilde{z}_t)$$

with initial condition $\tilde{z}_0 = \phi(\bar{z}_0)$ where $F^\phi := d\phi \circ \left(F \circ \phi^{-1}\right)$.

Proof. This follows from Theorem 6.12 by seeing that F is ϕ-related to F^ϕ. $\square$

This last lemma helps patch solutions in the manifold case.

Lemma 6.14. *Let $z \in C([0,T], V)$ and let $0 = t_0 < t_1 < \cdots < t_l = T$ be a partition of $[0,T]$. If*

$$z_{s,t} \underset{3}{\approx} F_{x_{s,t}}(z_s) + \left(\partial_{F_w(z_s)} F_{\tilde{w}}\right)(z_s)|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}} \tag{6.17}$$

holds for all $t_i \le s \le t \le t_{i+1}$ and $0 \le i < l$, then (6.17) holds for $0 \le s \le t \le T$.

In particular, if $\mathbf{z}_t$ solves $d\mathbf{z}_t = F_{d\mathbf{X}_t}(z_t)$ with $z_0 = \bar{z}_0$ on $[0,\tau]$ and $\tilde{\mathbf{z}}_t$ solves $d\tilde{\mathbf{z}}_t = F_{d\mathbf{X}_t}(\tilde{z}_t)$ with $\tilde{z}_\tau = z_\tau$ on $[\tau, T]$, then the concatenation of $\mathbf{z}_t$ and $\tilde{\mathbf{z}}_t$ in the sense of Lemma 2.6 solves $d\mathbf{z}_t = F_{d\mathbf{X}_t}(z_t)$ with $z_0 = \bar{z}_0$ on $[0,T]$.

Proof. This proof is identical to the proof of Lemma A.2 in [3], adapted here with different notation. We will only prove it in the case of two subintervals. First note that

$$F_w(y) = F_w(x) + F_w'(x)(y-x) + O(|w|\,|y-x|^2)$$

and

$$\left(\partial_{F_w(y)} F_{\tilde{w}}\right)(y) = \left(\partial_{F_w(x)} F_{\tilde{w}}\right)(x) + O\left(|w|\,|\tilde{w}|\,|y-x|\right)$$

by Taylor's theorem and the fact that $w \to F_w$ is linear. Using these facts, we have

$$\begin{aligned}
&z_{s,t} = z_{s,\tau} + z_{\tau,t} \\
&\underset{3}{\approx} F_{x_{s,\tau}}(z_s) + \left(\partial_{F_w(z_s)} F_{\tilde{w}}\right)(z_s)|_{w\otimes\tilde{w}=\mathbb{X}_{s,\tau}} + F_{x_{\tau,t}}(z_\tau) + (\partial_{F_w(z_\tau)} F_{\tilde{w}})(z_\tau)|_{w\otimes\tilde{w}=\mathbb{X}_{\tau,t}} \\
&\underset{3}{\approx} F_{x_{s,t}}(z_s) + F'_{x_{\tau,t}}(z_s)(z_{s,\tau}) + (\partial_{F_w(z_s)} F_{\tilde{w}})(z_s)|_{w\otimes\tilde{w}=\mathbb{X}_{s,\tau}} + (\partial_{F_w(z_s)} F_{\tilde{w}})(z_s)|_{w\otimes\tilde{w}=\mathbb{X}_{\tau,t}} \\
&\underset{3}{\approx} F_{x_{s,t}}(z_s) + F'_{x_{\tau,t}}(z_s)\left(F_{x_{s,\tau}}(z_s)\right) + \left(\partial_{F_w(z_s)} F_{\tilde{w}}\right)(z_s)|_{w\otimes\tilde{w}=\mathbb{X}_{s,\tau}+\mathbb{X}_{\tau,t}} \\
&= F_{x_{s,t}}(z_s) + \left(\partial_{F_w(z_s)} F_{\tilde{w}}\right)(z_s)|_{w\otimes\tilde{w}=\mathbb{X}_{s,\tau}+\mathbb{X}_{\tau,t}+x_{s,\tau}\otimes x_{\tau,t}} \\
&= F_{x_{s,t}}(z_s) + \left(\partial_{F_w(z_s)} F_{\tilde{w}}\right)(z_s)|_{w\otimes\tilde{w}=\mathbb{X}_{s,t}}.
\end{aligned}$$

$\square$

References

[1] Blanchet, J., Chen, X. and Dong, J.: ϵ-Strong simulation for multidimensional stochastic differential equations via rough path analysis. *Ann. Appl. Probab.* **27** (2017), no. 1, 275–336.

[2] Boedihardjo, H., Geng, X. and Qian, Z.: Quasi-sure existence of Gaussian rough paths and large deviation principles for capacities. *Osaka J. Math.* **53** (2016), no. 4, 941–970.

[3] Cass, T., Driver, B. and Litterer, C.: Constrained rough paths. *Proc. Lond. Math. Soc. (3)* **111** (2015), no. 6, 1471–1518.

[4] Cass, T. and Friz, P.: Densities for rough differential equations under Hörmander's condition. *Ann. of Math. (2)* **171** (2010), no. 3, 2115–2141.

[5] Cass, T., Friz, P. and Victoir, N.: Non-degeneracy of Wiener functionals arising from rough differential equations. *Trans. Amer. Math. Soc.* **361** (2009), no. 6, 3359–3371.

[6] Cass, T., Hairer, M., Litterer, C. and Tindel, S.: Smoothness of the density for solutions to Gaussian rough differential equations. *Ann. Probab.* **43** (2015), no. 1, 188–239.

[7] Cass, T., Litterer, C. and Lyons, T.: Rough paths on manifolds. In *New trends in stochastic analysis and related topics,* 33–88. Interdiscip. Math. Sci. 12, World Sci. Publ., Hackensack, NJ, 2012.

[8] Cass, T., Litterer, C. and Lyons, T.: Integrability and tail estimates for Gaussian rough differential equations. *Ann. Probab.* **41** (2013), no. 4, 3026-3050.

[9] Delarue, F. and Diel, R.: Rough paths and 1d SDE with a time dependent distributional drift: application to polymers. *Probab. Theory Related Fields* **165** (2016), no. 1-2, 1–63.

[10] do Carmo, M.: *Riemannian geometry.* Mathematics: Theory and Applications, Birkhäuser, Boston, MA, 1992.

[11] Driver, B.: Curved Wiener space analysis. In *Real and stochastic analysis,* 43–198. Trends Math., Birkhäuser, Boston, 2004.

[12] Elworthy, K.: *Stochastic differential equations on manifolds.* London Mathematical Society Lecture Note Series 70, Cambridge University Press, Cambridge-New York, 1982.

[13] Émery, M.: *Stochastic calculus in manifolds.* Universitext, Springer-Verlag, Berlin, 1989.

[14] Friz, P. and Hairer, M.: *A course on rough paths.* Universitext, Springer, Cham, Berlin, 2014.

[15] Friz, P. and Victoir, N.: *Multidimensional stochastic processes as rough paths.* Cambridge Studies in Advanced Mathematics 120, Cambridge University Press, Cambridge 2010.

[16] Gubinelli, M.: Controlling rough paths. *J. Funct. Anal.* **216** (2004), no. 1, 86–140.

[17] Hairer, M.: A theory of regularity structures. *Invent. Math.* **198** (2014), no. 2, 269–504.

[18] Hsu, E.: *Stochastic analysis on manifolds.* Graduate Studies in Mathematics 38, American Mathematical Society, Providence, RI, 2002.

[19] Ikeda, N. and Watanabe, S.: *Stochastic differential equations and diffusion processes*. Second edition. North-Holland Mathematical Library 24, North-Holland, Amsterdam; Kodansha, Tokyo, 1989.

[20] Inahama, Y.: Short time kernel asymptotics for rough differential equation driven by fractional Brownian motion. *Electron. J. Probab.* **21** (2016), Paper no. 34, 29 pp.

[21] Kelly, D. and Melbourne, I.: Smooth approximation of stochastic differential equations. *Ann. Probab.* **44** (2014), no. 1, 479–520.

[22] Kobayashi, S. and Nomizu, K.: *Foundations of differential geometry. Vol. II.* Interscience Tracts in Pure and Applied Mathematics 15, Vol. II, Interscience Publishers John Wiley & Sons, New York-London-Sydney, 1969.

[23] Kobayashi, S. and Nomizu, K.: *Foundations of differential geometry. Vol. I.* Wiley Classics Library, John Wiley & Sons, New York, 1996.

[24] Lyons, T.: Differential equations driven by rough signals. I. An extension of an inequality of L. C. Young. *Math. Res. Lett.* **1** (1994), no. 4, 451–464.

[25] Lyons, T.: The interpretation and solution of ordinary differential equations driven by rough signals. In *Stochastic analysis (Ithaca, NY, 1993)*, 115–128. Amer. Math. Soc., Providence, RI, 1995.

[26] Lyons, T.: Differential equations driven by rough signals. *Rev. Mat. Iberoamericana* **14** (1998), no. 2, 215–310.

[27] Spivak, M.: *A comprehensive introduction to differential geometry. Vol. I.* Second edition. Publish or Perish, Wilmington, Del., 1979.

Received June 21, 2015; revised April 17, 2016.

Bruce K. Driver: Department of Mathematics, 0112 University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0112, USA.
E-mail: bdriver@ucsd.edu

Jeremy S. Semko: Department of Mathematics, 0112 University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0112, USA.
E-mail: jeremy.semko@gmail.com

The research of Bruce K. Driver was supported in part by NSF Grant DMS-0739164. The author also gratefully acknowledges the generosity and hospitality of the Imperial College Mathematics Department, where part of this research was carried out, and was supported by EPSRC grant EP/M00516X/1 and a Nelder Visiting Fellowship during the Fall of 2014. The research of Jeremy S. Semko was supported in part by NSF Grant DMS-0739164.