Convergence of Lie group integrators

Charles Curry*, Alexander Schmeding†

We relate two notions of local error for integration schemes on Riemannian homogeneous spaces, and show how to derive global error estimates from such local bounds. In doing so, we prove for the first time that the Lie-Butcher theory of Lie group integrators leads to global error estimates.

Keywords: Lie group methods, homogeneous Riemannian manifold, Gronwall inequality, Runge-Kutta Munthe-Kaas methods, commutator free methods

MSC2010: 65L20 (primary); 22F30, 53C30 (secondary)

Contents

Introduction

The 1990s saw rapid development in numerical methods for differential equations on manifolds that are intrinsic in the sense that they preserve the manifold structure by evolving using geometric operations such as group actions and exponentials, see [CG93, MK95, MKZ97, MK98, MK99]. The case where the manifold in question is a homogeneous space has received particular focus, as it allows the equation to be

*NTNU Trondheim, charles.curry@ntnu.no
†TU Berlin, schmeding@tu-berlin.de

phrased in terms of Lie group actions, with important consequences for both practical implementation and theoretical analysis of methods.

Two classes of method in particular, the Runge-Kutta Munthe-Kaas (RKMK) methods and commutator-free methods (a generalization of Crouch-Grossmann methods) [CMO03] are now widely used in geometric integration theory and have proven themselves in a wide range of problems (see e.g. [IMNZ00] for a survey).

The order theory of such methods is developed in Lie-Butcher theory, [Owr06,MW08, ELM15], which generalizes the order theory of numerical methods in $\mathbb{R}^n$ rooted in B-series to homogeneous spaces. Whilst the algebraic side of this field has reached considerable maturity, there remain notable analytic gaps which we aim to fill. Indeed, in this context, the Taylor expansion of the pullback action of the approximate flow on smooth functions is compared to that of the exact flow; a Lie-group integrator is then said to be of order p if the two expansion coincide up to order p. A method of order p hence obeys the following local estimate:

Suppose V is a C^{p+1} vector field on a manifold M, and let $\hat{y}$ be an approximation of the integral curve y of V, arising from a Lie-group integrator of order p. Then for all $f \in C^\infty(M)$,

$$|f(\hat{y}) - f(y)| \leq Ch^{p+1} \text{ for some constant } C. \tag{1}$$

There does not, however, seem to exist a fully satisfactory derivation of either local or global estimates at present. Indeed, to the authors knowledge, only partial results are available, such as [Fal00] cf. also the survey in [IMNZ00, Section 9] and most recently [CEOR18, Section 3]. The main issue here is twofold:

- local error estimates of the form (1) are not immediately relatable to the Riemannian distance.
- global error estimates are only available with additional assumptions on the vector fields or the geometry of the manifold.

Our first main result (Theorem 3.4) clarifies how the local estimate, given with respect to test functions, implies a comparable estimate involving the Riemannian distance as is often required [Fal00, IMNZ00]. Recall that the natural setting for Lie group methods is a homogeneous Riemannian manifold (M, g)[1].For the rest of this introduction we will always assume that (M, g) is such a manifold. Then our results subsume the following:

Theorem A *Let V be a C^{p+1} vector field such that $\hat{y}$ approximates the integral curve y of V up to order p. Then the local estimate*

$$d(y(h), \hat{y}(h)) \leq Ch^{p+1}$$

holds, where d is the geodesic distance on M.

This result settles the problem of local error estimates and clarifies the dependency on the different kinds of error estimates found in the literature. The second point is

[1]Every Lie group is a homogeneous Riemannian manifold. We recall the necessary results and facts on these manifolds in Section 1.

covered by Theorem 4.4, which shows that global error estimates follow from local estimates involving the Riemannian metric. As such it subsumes the following:

Theorem B *For a C^{p+1} vector field V we fix a sequence $\{\hat{y}_n\}_{i=1,\dots,n}$ approximating the integral curve of V through y_0 at a discrete set of times t_i with $h_i = t_{i+1} - t_i$, and $\max_i h_i = h$. If $\hat{y}$ obeys either of the local estimates above with exponent $p+1$, then we obtain the global estimate*

$$d(y_n, \hat{y}_n) \leq Ch^p,$$

We mention here that our results also clarify the dependency of the constants C on the parameters such as the vector field V (which we have deliberately suppressed in the above statements of our theorems). These results taken together give a fully rigorous analytic counterpart to the algebraic Lie-Butcher order theory.

The paper is organized as follows: we begin with a brief overview of Riemannian homogeneous spaces, fixing notation and stating some standard results. We follow this in §2 by a brief treatment of the local estimates obtained from Lie-Butcher theory, where for the sequel it is important that we establish estimates with explicit remainder terms. The passage from local estimates obtained via Lie-Butcher theory to local estimates using the Riemannian metric is then accomplished in §3. We conclude in §4 with a derivation of the global error estimate.

1. Preliminaries on Riemannian manifolds

In this section we fix the notation and general setting. All of the material here is standard and can be found in books on differential geometry and Riemannian geometry, e.g. [Lan99, Kli95, KN96]. We assume that the reader is familiar with basic concepts such as Riemannian metrics and associated concepts such as (Levi-Civita) connections and covariant derivatives.

1.1 We let $\mathbb{N} := \{1, 2, \dots\}$ denote the natural numbers and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. All manifolds in this paper are assumed to be paracompact and finite dimensional. By (M, g) we denote a Riemannian manifold, where we write the following for the data associated to g:

- g_m will be the inner product on $T_m M, m \in M$ with associated norm $\|\cdot\|_{g_m}$.
- ∇X will be the covariant derivative of a vector field X.
- $d\colon M \times M \to \mathbb{R}$ will be the geodesic length metric induced by g.

In general we will be working with a special class of Riemannian manifolds, arising as quotients of isometric Lie group actions: the so called homogeneous Riemannian manifolds (see e.g. [Bes08, B.7]).

1.2 By $\Lambda\colon G \times M \to M$ we denote a (left) Lie group action on (M, g) such that

1. the action Λ is transitive,
2. $\forall g \in G$, the map $\Lambda_g := \Lambda(g, \cdot)$ is a Riemannian isometry (g is left G-invariant).

Note that by the above (M, g) becomes a Riemannian homogeneous space, i.e. the isometry group $\mathrm{Iso}(M, g)$ acts transitively. This entails that as manifolds $M \cong G/H$ where H is a compact subgroup of G (where H can be identified as the stabiliser subgroup of a point). Denote by $\pi\colon G \to G/H \cong M$ the canonical quotient map and set $o := \pi(e)$ (where e is the identity element of G).

Finally note that Riemannian homogeneous spaces are geodesically complete [KN96, IV. Theorem 4.5], i.e. geodesics exist for all time.

Many manifolds appearing in applications are homogeneous Riemannian manifolds as is recalled in the following example.

1.3 Example 1. Every Lie group G is a Riemannian homogeneous space, where $H = \{e\}$ is the identity subgroup and g is a left invariant Riemannian metric

2. Spheres [Bes08, B. Example 7.13] and projective spaces [Bes08, B. Example 7.1] are homogeneous Riemannian spaces (e.g. $\mathbb{S}^n \cong \mathrm{SO}(n)/\mathrm{SO}(n-1)$ where the Riemannian metric is induced by the biinvariant metric on $\mathrm{SO}(n)$)

We refer to the survey [IMNZ00] for a wealth of examples on numerical integrators on these spaces which can be treated in the framework of Lie-Butcher theory.

Let us remind readers who are not familiar with Riemannian geometry that many properties of Riemannian homogeneous spaces might be conveniently formulated as properties of the geodesic length metric d. We collect two important facts:

1.4 (Metric view of Riemannian manifolds) 1. The Riemannian manifold (M, g) is complete if and only if (M, d) is a complete metric space (this is part of the famous Hopf-Rinow theorem [Bes08, Theorem 1.65]).

2. A surjective smooth map $f\colon M \to M$ is an isometry, if and only if it is distance preserving: $d(f(x), f(y)) = d(x, y), \forall x, y \in M$ (see [Bes08, Theorem 1.75]).

Finally, we fix notation concerning vector fields.

1.5 (Vector fields and flows) We will denote by $\mathcal{X}^p(M)$ the space of all vector fields on M of class $p \in \mathbb{N} \cup \{\infty\}$ (writing $\mathcal{X}(M) := \mathcal{X}^\infty(M)$). For $V \in \mathcal{X}^p(M)$ we write

$$\mathrm{Fl}_0^V : \mathbb{R} \times M \supseteq \mathcal{D}(V) \to M$$

for the *flow associated to* V. The flow is defined by sending a pair $(t, x_0) \in \mathcal{D}(V)$ ($\mathcal{D}(V)$ is open subset of $\mathbb{R} \times M$) to the solution $y(t) := y_{x_0}(t) = \mathrm{Fl}_0^V(x_0, t)$ of the initial value problem

$$y' = V(y) \quad y(0) = x_0.$$

We let $\varphi_t := \mathrm{Fl}_0^V(t,\cdot)$ be the associated (local) diffeomorphism (assuming that $\mathcal{D}(V) \cap \{t\} \times M$ is non void).

Further, the usual conventions (cf. e.g. [Lan99, Section V]) for the derivative operation of a vector field on a C^k function f, i.e. $V(f)$ and the shorthand $V^k(f) = V(V(\cdots V(f)))$ are used throughout the text. As is common $C^p(M)$ denotes the set of real valued C^p-functions on the manifold M.

For $f \in C^p(M)$, we let $\varphi_h^*(f) = f \circ \varphi_h$ denote the pullback.

In the next section we turn now to Lie-Butcher theory. After a very brief primer on the most important concepts, we are interested in the description of Taylor expansions of exact and numerical solutions to differential equations on Riemannian manifolds.

2. Lie-Butcher theory and Lie Series estimates

This section is devoted to giving a precise version of the local estimates deriving from the Lie-Butcher order theory for Lie group integration methods developed in [MK98, MK99, MO99]. Only an extremely brief discussion of Lie-Butcher theory is provided to lay the foundation for the following computations. For a friendly introduction to the theory of Lie group integrators we refer the reader to [CMO14].

Let us first consider the Taylor expansion of an exact solution of a differential equation given by the vector field V tested against a C^∞ function.

2.1 Lemma (Lie Series) *Let $V \in \mathcal{X}^{p+1}(M)$. For any $f \in C^\infty(M)$, the pullback action of the flow of V has the Taylor series expansion*

$$\varphi_h^*(f)(x) = f(x) + \sum_{k=1}^{p} \frac{1}{k!} h^k V^k(f)(x) + \frac{1}{(p+1)!} h^{p+1} \varphi_t^* V^{p+1}(f)(x), \tag{2}$$

where $t < h$.

Proof. Assume that $(h,x) \in \mathcal{D}(V)$ (whence $(t,x) \in \mathcal{D}(V)$ for all $0 \le t \le h$). It is a standard result [Lan99, V.§5 Propositions 5.2 and 5.3] that

$$\frac{d}{dt}(\varphi_t^* f) = \varphi_t^* V(f).$$

Now $V(f) \in C^{p+1}(M)$, and each further application of V reduces the differentiability by one degree, so we can iterate this $p+1$ times to obtain $p+1$ derivatives of $\varphi_t^* f$. Fixing $x \in M$, we can view $\varphi_t^*(f)(x)$ as a function of t alone, and the given result follows immediately from Taylor's theorem. □

A key idea of Lie-Butcher theory is that a numerical scheme yields a Taylor series like expansion, the *Lie-Butcher series*, which can be compared to the Taylor expansion (2) of the exact solution. To give the Lie-Butcher series of an (numerical) approximation

of the exact flow, we require the notion of an elementary (covariant) differential. A suitably general setting to define these differentials is the following:

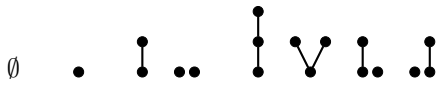
2.2 Suppose that $\triangleright\colon \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}(M), \quad X \triangleright Y := \tilde{\nabla}_X Y$ is a binary product defined through a flat, constant torsion Koszul connection.[2] This may be extended to give a product on the enveloping algebra by

$$\begin{aligned} (XY) \triangleright Z &= X \triangleright (Y \triangleright Z) - (X \triangleright Y) \triangleright Z \\ X \triangleright (YZ) &= (X \triangleright Y)Z + Y(X \triangleright Z), \end{aligned} \tag{3}$$

see [ELM15, CEMM18]. Note that we can not define a similar concept for vector fields of class C^p for $p \in \mathbb{N}$ as they do not form a Lie algebra (whose universal enveloping algebra the construction uses). However, given a connection as above covariant derivatives of C^p vector fields make sense if one takes care to account for the loss of differentiability.

Now we note that in every term of (2), the vector field V acts (up to $p+1$ times) as a derivation on the test function f. Following an idea by Cayley, this situation can conveniently be described using rooted trees:

2.3 (Trees) For $n \in \mathbb{N}_0$, a *rooted tree* of degree n is a finite oriented tree with n vertices. We distinguish one vertex without outgoing edges, the *root* of the tree. Any vertex can have arbitrarily many incoming edges, and any vertex other than the root has exactly one outgoing edge. Vertices with no incoming edges are called leaves. A planar rooted tree is a rooted tree together with an embedding in the plane [3]. A planar rooted forest is a finite ordered collection of planar rooted trees. Here the planar rooted forests are depicted up to order three (with $\emptyset$ being the empty tree):



We will see now, one one can construct differential operators from forests with edges using the binary product $\triangleright$.

2.4 Definition (Elementary differentials) Let $\tau = B_+(\tau_1, \tau_2, \ldots, \tau_n)$ be a planar tree defined recursively by connecting the branches $\tau_1, \ldots, \tau_n$ from left to right onto a root.

[2] See [KN96, Chapters II and III] for basic informations on the connections used here. We write "$\tilde{\nabla}$" to emphasise that the connection will in general not be the Levi-Civita connection [KN96, IV.2] of the Riemannian manifold.

[3] A standard notion from graph theory, informally comprising a drawing of the tree in the plane such that branches (paths from a node to a leaf) do not cross. In practice, this introduces an order (left-to-right) on the set of branches starting from a given node.

We then define the *elementary (covariant) differentials* of V recursively by

$$V_\bullet = V,$$
$$V_\tau = (V_{\tau_1} \cdots V_{\tau_n}) \triangleright V,$$

where $\bullet$ is the tree with a single node and products are expanded via (3). This is extended to forests of trees $\tau_1 \cdots \tau_n$ by

$$V_{\tau_1 \tau_2}(f) = V_{\tau_1}\big(V_{\tau_2}(f)\big) - \big(V_{\tau_1} \triangleright V_{\tau_2}\big)(f),$$

see [ELM15, Section 3] or [CEMM18, 5.2].

2.5 Remark

1. For a forest τ of order p, the construction of the differential operator V_τ in Definition 2.4 requires one to take (recursively) at most p (covariant) derivatives of the vector field V. Thus if the order of the forest τ is smaller then p, we may consider the differential operator V_τ for any vector field $V \in \mathcal{X}^p(M)$.
2. Again by construction V_τ takes at most p derivatives of the function f if τ is a forest of order at most p.

The typical assumption of Lie-Butcher order theory is that a method of order p admits a Taylor expansion of the following type:

2.6 Assumption (Lie-Butcher series) *For any test function $f \in C^\infty(M)$, the pullback action of an approximate flow $\hat{y}$ of order p has a Taylor series expansion*

$$\hat{\varphi}_h^*(f)(x) = f(x) + \sum_{k=1}^{p} \frac{1}{k!} h^k V^k(f)(x) + \sum_{|\tau|=p+1} \alpha(\tau) h^{p+1} \hat{\varphi}_t^* V_\tau(f),$$

where $|\tau|$ is the number of nodes in the forest τ, α is a linear functional on forests and $\hat{\varphi}_h$ is the (local) diffeomorphism associated to the flow $\hat{y}$.

We emphasise here that due to Remark 2.5, it makes sense to consider the Assumption 2.6 for $V \in \mathcal{X}^{p+1}(M)$ for $p \in \mathbb{N}$ as the differential operators V_τ make sense in this regime.

2.7 Remark In a universal setting, the Lie-Butcher series of a method is often identified with a character $\alpha(\tau)$, i.e. a multiplicative linear functional on the space of planar forests with concatenation product, see [ELM15], in particular [ELM15, Section 3.3].

The Assumption 2.6 is then satisfied whenever $V \in \mathcal{X}^{p+1}(M)$, and the functional α agrees with the exact solution character $\frac{1}{\sigma(\tau)\tau!}$ on forests with p or fewer nodes (here σ is the internal symmetry factor and $\tau!$ is the planar forest factorial character, see [CEMM18] for definitions).

By comparing the Lie-Butcher series of an approximate flow $\hat{\varphi}$ to the Lie series of the exact flow φ, we obtain the first local estimate:

2.8 (Local estimate for smooth test functions) Suppose $V \in \mathcal{X}^{p+1}(M)$, and let $\hat{y}$ be an approximation of the integral curve y of V, of order p in the sense of Assumption 2.6. Then for all $f \in C^\infty(M)$ and $y_0 \in M$,

$$|f(\hat{y}) - f(y)| \leq C(V, f)h^{p+1} \tag{4}$$

2.9 Remark For later use we wish to be very explicit on how the constant $C(V, f)$ from 2.8 (4) depends on f. Comparing the Lie-Butcher series and the Lie series of the exact method we see that the remainder term is given (for an order p method) by

$$R(V, f)(x) = \frac{1}{(p+1)!}\varphi_t^* V^{p+1}(f)(x) - \sum_{|\tau|=p+1} \alpha(\tau)\hat{\varphi}_t^* V_\tau(f).$$

where $(t, x) \in \mathcal{D}(V)$, φ_t is the flow of the exact solution and $\hat{\varphi}_t$ the flow of the approximate solution. As $C(V, f)$ can be chosen to be any constant which dominates the norm of the remainder term, we need an estimate on this norm. Thus

$$|R(V, f)(x)| \lesssim \sup_{t\in[0,h]} \left(|V^{p+1}(f)(\varphi_t(x))| + \sum_{|\tau|=p+1} |V_\tau(f)(\hat{\varphi}_t(x))| \right),$$

where "$\lesssim$" denotes an inequality up to constants which neither depend on f nor on V. Now by definition (cf. Remark 2.5) V^{p+1} and V_τ act on smooth functions as differential operators whose order is at most $p + 1$. Thus the terms $V^{p+1}(f)(\varphi_t(x))$ and $V_\tau(f)(\hat{\varphi}_t(x))$ can be computed as linear combinations (depending on V, p, τ and the connection $\tilde{\nabla}$) of the partial derivatives of f up to order $p+1$. Further, we note that to obtain an estimate we only need to give an upper bound on all partial derivatives of f up to order $p + 1$ on the compact set $\{\varphi_t(x) \mid t \in [0, h]\} \cup \{\hat{\varphi}_t(x) \mid t \in [0, h]\}$.

3. Local estimates to local metric estimates

Our first step is to show that the condition above implies the weaker (but in some senses more natural) condition.

3.1 Definition (Local metric estimate) Let $\hat{y}$ approximate the integral curve y through y_0 at $y(h)$ of order $p \in \mathbb{N}$. Then there is a constant $C = C(\hat{y}, y)$ such that

$$d(y(h), \hat{y}(h)) \leq Ch^{p+1}. \tag{5}$$

Note that the condition (5) appeared in the stability analysis for Lie group methods in [IMNZ00, Section 9] and the earlier work by Faltinsen [Fal00]. However, there

the local metric estimate (5) is **assumed to hold** for a given method to enable the analysis, whereas we will deduce the validity of (5) in case the method satisfies local estimate (4).

To prove this result we introduce a family of smooth functions which will allow us to deduce the local metric estimate (5) from the local estimate for smooth test functions 4. To this end, we need smooth functions controlling the geodesic distance. The constructions of these functions in Lemma 3.2 and Lemma 3.3 below is somewhat technical, hence we postpone it to Appendix A. We need a smooth function which allows us to control the geodesic distance for points "far away" from o.

3.2 Lemma *For $\varepsilon > 0$ and (M, g) a connected[4] complete Riemannian manifold, there exists $F_\varepsilon \in C^\infty(M)$ with the following properties.*

1. *$F_\varepsilon(o) = 0$ and $F_\varepsilon(x) \geq 0$, $\forall x \in M$,*
2. *if $d(x, o) > \varepsilon$, then $F_\varepsilon(x) \geq d(x, o)$.*

Note that the Riemannian distance from a fixed point is in general only continuous. Due to the cut locus of the Riemannian manifold, the distance $d(o, \cdot)$ does not become smooth if we consider it only on an open set away from o. This phenomenon prevents us from simply "smoothing out" the Riemannian distance at o to obtain the desired smooth function. Furthermore, the geodesic distance $d(\cdot, o)$ is also non smooth at o whence we need smooth functions controlling the distance near o.

3.3 Lemma *Choose $\varepsilon > 0$ such that the closure of the metric ball $B^d_\varepsilon(o)$ is contained in a manifold chart (U, φ). Then there is $N \in \mathbb{N}$ and a family $\{f_n\}_{1 \leq n \leq N} \subseteq C^\infty(M)$ with the following properties*

1. *$f_n(o) = 0$,*
2. *if $d(x, o) < \varepsilon$ then there is $1 \leq n_x \leq N$ such that $f_n(x) \geq d(x, o)$.*

Note that the functions $\{f_n\}_n$ are also allowed to take negative values (which they will take on a neighborhood of o!). This is unavoidable if one wants to obtain smooth functions which dominate the distance (at least in some directions) and are 0 at o. We now have all technical tools at our disposal to prove the main result of this section:

3.4 Theorem *Let (M, g) be a homogeneous Riemannian manifold with G-invariant Riemannian metric and $V \in \mathcal{X}^{p+1}(M)$. Assume that $\hat{y}$ approximates the integral curve $y = \mathrm{Fl}^V_0(\cdot, y_0)$ up to order p as in Assumption 2.6. Then $\hat{y}$ satisfies the local metric estimate (5).*

[4] Here connectedness is only needed to make sense of condition 2 in the statement, as for a non-connected manifold it is customary to set $d(x, y) = \infty$ if x, y are from different connected components. Assuming that M is connected is no essential restriction as we will only compare curves lying in the same connected component.

Proof. We proceed in several steps to obtain the estimate from the family of smooth functions constructed in Lemma 3.3 and Lemma 3.2. To this end, let $y\colon [0,h] \to M$ be the integral curve $y(t) := \mathrm{Fl}_0^V(t, y_0)$ of a C^{p+1}-vector field V defined (at least) on the compact interval $[0,h]$ with $h > 0$. By standard arguments [Lan99, Section IV] y is a C^{p+2}-mapping. The construction of the functions (see Step 2) does not depend on h and work for any choice of h as long as the integral curve exists up to time h. However, the estimates in Step 3 depend on h as they are taken over the interval $[0,h]$.

Step 1: Smooth shift from $y(t)$ to o: Since M is a homogeneous space, there exists a C^{p+2}-curve $\tilde{y}\colon [0,h] \to G$ which lifts y, i.e. $\pi \circ \tilde{y} = y$.[5] Note that the lift is non-unique but the estimates will not depend on the choices made. By construction, $\Lambda_{\tilde{y}(t))^{-1}}\colon M \to M$ is the left action of the Lie group element $(\tilde{y}(t))^{-1}$ on the homogeneous space M, see 1.2. Since $\tilde{y}$ lifts the curve y, we have $\Lambda_{\tilde{y}(t))^{-1}}(y(t)) = o, \quad \forall t \in [0,h]$. Using left invariance of the metric, for every $t \in [0,h]$ the map $\Lambda_{\tilde{y}(t)^{-1}}$ is an isometry, whence for $x \in M$ and $t \in [0,h]$,

$$d(x, y(t)) = d(\Lambda_{\tilde{y}(t)^{-1}}(x)), \Lambda_{\tilde{y}(t)^{-1}}(y(t))) = d(\Lambda_{\tilde{y}(t)^{-1}}(x), o). \tag{6}$$

Step 2: A family of smooth comparison functions. Choose $\varepsilon > 0$ such that the closure of the ball $B := B_\varepsilon^d(o)$ is contained in a chart (U, φ). By Lemma 3.3 we obtain a family of smooth functions $\{f_n\}_{1 \le n \le N}$ (where we can choose $N = 2^{\dim M}$) which controls the geodesic distance on B such that every f_n vanishes in o. Then Lemma 3.2 applied for the same ε yields $F_\varepsilon \in C^\infty(M)$ which controls the Riemannian distance $d(\cdot, o)$ outside of B and satisfies $F_\varepsilon(o) = 0$.

We construct now the functions which will yield the necessary estimates:

$$\begin{aligned} \omega_n &\colon [0,h] \times M \to \mathbb{R}, \quad (t,x) \mapsto f_n \circ \Lambda_{y(t)^{-1}}(x), \quad 1 \le n \le N \\ \omega_{N+1} &\colon [0,h] \times M \to \mathbb{R}, \quad (t,x) \mapsto F_\varepsilon \circ \Lambda_{y(t)^{-1}}(x). \end{aligned}$$

By construction the ω_n are $p+2$-times continuously differentiable with respect to t and every of these differentials is smooth with respect to x. Thus we obtain continuous maps into the space $C^\infty(M)$ endowed with the compact open C^∞-topology via[6]

$$\omega_n^\vee \colon [0,h] \to C^\infty(M), \omega^\vee(t) := \omega_n(t, \cdot) \quad 1 \le n \le N+1.$$

Recall that the compact open C^∞ topology is generated by the family of seminorms which control the growth of a function and up to finitely many of its derivatives on some compact subset in M (cf. e.g. [HS17] for more on topologies for $C^\infty(M)$). Since $\omega_n^\vee$ is continuous, $\omega_n^\vee([0,h])$ is compact whence every continuous seminorm of $C^\infty(M)$

[5] Here we use that a homogeneous space is a principal H-bundle, whence a C^{p+2}-curve admits a C^{p+2} horizontal lift, cf. e.g. [OR04, Chapter 5.1].

[6] Functions with the differentiability exhibited by ω_n are called $C^{p+2,\infty}$-functions in [AS15]. Indeed that ω_n is $C^{p+2,\infty}$ follows from the chain rules in ibid. The continuity of $\omega_n^\vee$ into the locally convex space $C^\infty(M)$ is a consequence of the exponential law [AS15, Theorem B] which even shows that $\omega_n^\vee$ is a C^{p+2} map. Since continuity is sufficient for our purposes we do not need to explain what differentiable functions into the (non normable!) space $C^\infty(M)$ are.

is bounded on the image of $\omega_n^\vee$. Summarising: The growth of (up to finitely many) derivatives of the functions $\omega_n^\vee(t)$ on a given compact set can be uniformly bounded in t. Finally, by construction $\omega_n^\vee(t)(y(t)) = 0$ for all $1 \leq n \leq N+1$ and $t \in [0,h]$.

Step 3: The metric estimate. By construction, the functions obtained via Lemma 3.3 and 3.2 control the geodesic distance of a point to o. Hence (6) implies that for $d(\hat{y}(t), y(t)) \geq \varepsilon$ the function $\omega_{N+1}^\vee(t)$ controls $d(\hat{y}(t), y(t))$, while for $d(\hat{y}(t), y(t)) < \varepsilon$ one of the functions $\omega_n^\vee(t), 1 \leq n \leq N$ controls $d(\hat{y}(t), y(t))$. We deduce that for every $t \in [0,h]$ there is some index n_t such that

$$d(\hat{y}(t), y(t)) \leq \omega_{n_t}^\vee(t)(\hat{y}(t)) = |\omega_{n_t}^\vee(t)(\hat{y}(t)) - \underbrace{\omega_{n_t}^\vee(t)(y(t))}_{=0}| \leq C_{\omega_{n_t}^\vee(t)} t^{p+1} \tag{7}$$

where the last inequality follows from the local estimate (4) and $C_{\omega_{n_t}^\vee}$ depends on $\omega_{n_t}^\vee(t)$ (where the other dependencies do not matter here). Since we are after a global estimate independent of $\omega_n^\vee$ and t we have to recall how these constants depend on $\omega_n^\vee$. From Remark 2.9 we know that up to some constant A (depending on the Lie-Butcher series, V and the initial conditions but not on the smooth function), the constants $C_{\omega_n^\vee(t)}$ can be bounded by the partial derivatives up to order $p+1$ of $\omega_n^\vee(t)$ on the compact set

$$K := \{y(t) \mid t \in [0,h]\} \cup \{\hat{y}(t) \mid t \in [0,h]\} \subseteq M.$$

In other words, we have to control $\sup_{t\in[0,h]} \|\omega_n^\vee(t)\|_{p+1,K}$, where $\|\cdot\|_{p+1,K}$ measures the (sum of) absolute values of partial derivatives on K up to order $p+1$.

Following Step 2, we know that there is a uniform bound in t, i.e.

$$R := \sup_{1\leq n\leq N+1} \sup_{t\in[0,h]} \|\omega_n^\vee(t)\|_{p+1,K} < \infty$$

and thus $\sup_{t\in[0,h]} C_{\omega_{n_t}^\vee(t)} \leq AR =: C$. Hence, from (7) we conclude that

$$d(\hat{y}(t), y(t)) \leq Ct^{p+1} \leq Ch^{p+1} \quad \forall t \in [0,h]. \quad \square$$

Note that the main point in the proof of Theorem 3.4 was to establish a uniform bound independent of t. We remark that the constant C obtained for the metric estimate still depends on the choices we made in the proof (e.g. the choice of $\varepsilon > 0$). Thus the proof is a pure existence proof without any claim of optimality of C. Indeed, if one chooses ε very small one should expect C to become bigger as it is derived from an estimate of the derivatives of smooth functions which involve cut-off functions confined to the ε-ball.

4. Local to global estimates

In this section we prove our second main result, a global error estimate for the Lie group methods. In the last chapter we have seen that Lie group methods satisfy a (local) metric estimate with respect to the geodesic metric. We apply now a suitable version

of a Gronwall type estimate for Riemannian manifolds which was first established in [KSSV06]. Let us recall its statement for easy reference:

4.1 For $X \in \mathcal{X}^p(M)$ (and $p \in \mathbb{N} \cup \{\infty\}$) the covariant derivative induces continuous linear maps

$$\nabla X(p)\colon (T_pM, \|\cdot\|_p) \to (T_pM, \|\cdot\|_p), \quad Y_p \mapsto \nabla_{Y_p} X, \quad p \in M,$$

(cf. [Kli95, Section 1.5] and [Lan99, Section VIII, in particular VIII, §2 Lemma 2.3]).[7] The operator norm of these mappings will be denoted by $\|\nabla X(p)\|_g$.

4.2 (From [KSSV06, Corollary 1.6]) Let (M, g) be a connected and complete Riemannian manifold, $V \in \mathcal{X}(M)$ and $p_0, q_0 \in M$. Let S be a minimizing geodesic segment connecting p_0 and q_0 (cf. [Kli95, Section 2.1 and Theorem 2.1.3]). Choose $T > 0$ such that the flow Fl^X of the vector field X is defined on $[0, T] \times S$.
Then the integral curves $\varphi(t) := \mathrm{Fl}^X_t(p_0), \psi(t) := \mathrm{Fl}^X_t(q_0)$ *with initial value* p_0 *(resp.* q_0*) satisfy the Gronwall type estimate*

$$d(\varphi(t), \psi(t)) \leq d(p_0, q_0) e^{C_T t}, \quad t \in [0, T] \tag{8}$$

where $C_T = \sup\{\|\nabla X(p)\|_g \mid p \in \mathrm{Fl}^X([0, T] \times S)\}$.

The Gronwall type estimate exhibited in 4.2 has been established in [KSSV06] only for smooth vector fields. We wish to obtain a similar estimate for $\mathcal{X}^p(M)$ vector fields ($p \in \mathbb{N} \cup \{\infty\}$).

4.3 Remark (Estimate 4.2 holds for $X \in \mathcal{X}^p(M), p \in \mathbb{N}$.) To prove that the Gronwall estimate (8) holds also with lower differentiability of the vector field, note that [KSSV06, Corollary 1.6] is a direct consequence of [KSSV06, Theorem 1.4]. The proof of said theorem uses the differentiability class of the vector field V only through an application of [KSSV06, Proposition 1.1] in the proof. Hence if [KSSV06, Proposition 1.1] holds for vector fields in $\mathcal{X}^p(M)$ for $p \in \mathbb{N}$ we are done.

Reexamining the proof of [KSSV06, Proposition 1.1]: One needs differentiability of the flow map $\mathrm{Fl}^V_0(t, x_0)$ (cf. 1.5 for details on the notation). Namely, the existence of iterated derivatives of the type $\partial_t \partial_{x_0} \mathrm{Fl}^V$ and $\partial_t \partial_{x_0} \mathrm{Fl}^V$ are required. If these second mixed partial derivatives (with respect to t and x_0) exist, then the proof can be carried out exactly as presented in [KSSV06, p. 134-135].

The case $p \geq 2$. For $p \geq 2$ by standard ODE arguments (cf. e.g. [Lan99, Section IV] or [AS15, Proposition 5.13]) also Fl^V_0 is a C^p map whence the iterated differentials exist.

[7] Covariant derivatives are often only defined for smooth vector fields. However, the $\nabla_X V$ makes sense for vector fields from $\mathcal{X}^p(M)$ (for $p \in \mathbb{N}$ using that (M, g) is smooth) if one accounts for the loss of differentiability.

The case $p = 1$. Looking closer, the derivatives still exist for $p = 1$: We call a continuous $f\colon M \times N \to L$ between (finite dimensional) manifolds a $C^{1,1}$ map if (in charts) the partial derivative with respect to M exists, is continuous and f and the derivative are again continuously differentiable with respect to the N variable (cf. [AS15, Definition 5.11]). The concept of $C^{1,1}$ (or the general $C^{r,s}$ maps developed in [AS15]) captures exactly existence of the mixed partial derivatives needed. Now for $V \in \mathcal{X}^1(M)$, [AS15, Proposition 5.13] asserts that the flow Fl_0^V is indeed of class $C^{1,1}$.

Summing up, the Gronwall estimate 4.2 holds $V \in \mathcal{X}^p(M)$ where $p \in \mathbb{N} \cup \{\infty\}$.

4.4 Theorem (Global error estimate) *Consider a vector field $V \in \mathcal{X}^{p+1}(M)$ with $p \in \mathbb{N}_0 \cup \{\infty\}$ on a Riemannian homogeneous space (M, g) together with a sequence $\{\hat{y}_n\}_{i=1,\dots,n}$ approximating the integral curve of V through y_0 at a discrete set of times t_i with $h_i = t_{i+1} - t_i$, and $\max_i h_i = h$. If $\hat{y}$ obeys either of the local estimates (4) or (5) with exponent $p+1$, then we obtain the global estimate*

$$d(y_n, \hat{y}_n) \leq Ch^p,$$

where C is a constant depending only on V, y_0 and T.

Proof. The proof follows the standard "Lady Windemere's fan" argument, as per [HNW08, Section 2.3]. Indeed, for $i = 1, \dots, n$, define the local error

$$e_i = d\big(\hat{y}_i, \varphi_h(\hat{y}_{i-1})\big),$$

and the transported local error

$$E_i = d\big(\varphi_{T_i}(\hat{y}_i), \varphi_{T_{i-1}}(\hat{y}_{i-1})\big),$$

where $T_i = T - t_i$. From 4.2 and Remark 4.3 (adapting [KSSV06, Corollary 1.6]) the errors are related by

$$E_i \leq e^{C_T T_i} e_i$$

We then use the local metric estimate (5), if necessary invoking Theorem 3.4 to justify this, obtaining

$$e_i \leq C_i h_i^{p+1}.$$

The Lady Windemere's fan estimate then concludes the argument; indeed taking $C = \max_i C_i$ we have

$$\begin{aligned} d(y_n, \hat{y}_n) &\leq \sum_{i=1}^{n} E_i \\ &\leq h^p C\big(h_0 e^{C_T T_1} + h_1 e^{C_T T_2} + \dots\big) \\ &\leq h^p \frac{C}{C_T}\big(e^{C_T T} - 1\big) \end{aligned}$$

□

4.5 Remark Global error estimates for discrete gradient descent methods were recently obtained in [CEOR18] and the methods in ibid. are very similar to the ones used to derive Theorem 4.4. Though ibid. concerns itself with discrete gradient methods, it is not hard to see that the arguments given there are universal, i.e. could be adapted to analyse the convergence of general numerical methods.

The key difference between our approach and the analysis in [CEOR18] is in the basic setting: Studying Lie group integrators we are working by default in a complete Riemannian manifold. On non complete Riemannian manifolds, the local to global argument using a Gronwall inequality 4.2 breaks down (see [KSSV06, Example 1]). Without completeness of the manifold the argument holds in general only for complete vector fields. Thus the local to global result [CEOR18, Theorem 2] is derived only for complete and smooth vector fields (though on a possibly non complete manifold, see also [CEOR18, Remark after Theorem 2] on how to relax the completeness condition). Note that in light of Remark 4.3 the proofs developed in ibid. hold up if one replaces smooth vector fields with C^1-vector fields.

A. Auxiliary constructions

In this appendix we collect several auxiliary results which enable us to construct smooth functions needed in the estimates. We begin with a technical Lemma concerning partitions of unity with some desirable properties:

A.1 Lemma *Let M be a paracompact finite dimensional manifold, $o \in M$ and B be an open o-neighborhood. There exists a locally finite open cover $\{U_i\}_{i\in I}$ of M, such that $I = J \cup \{i_o\}$ and the following holds:*

1. *i_o is the unique index such that $o \in U_{i_o}$,*
2. *$U_{i_o} \subseteq B$,*
3. *every U_i is connected and relatively compact,*

Proof. Since M is locally compact, we can choose a connected manifold chart (U_{i_o}, φ_{i_o}) and compact o-neighborhoods C_1, C_2 of o such that the following inclusions hold:

$$o \in C_1 \subseteq U_{i_o} \subseteq \overline{U}_{i_o} \subseteq C_2^\circ \subseteq C_2 \subseteq B$$

(where $\overline{U}_{i_o}$ is the closure and C_2° the interior). Then $U := M \setminus C_1$ is open and metrisable, whence paracompact. Following [Lan99, II, §3 Theorem 3.3] there is a locally finite cover of U by charts $(U_j, \varphi_j)_{j\in J'}$ such that U_j is connected and relatively compact. Let us now throw out all elements of the cover which are contained in U_{i_o}, i.e. define $J := \{j \in J' \mid U_j \cap M \setminus U_{i_o} \neq \emptyset\}$ and set $I := J \cup \{i_o\}$. By construction $o \in U_i$ for $i \in I$ if and only if $i = i_o$, $U_{i_o} \subseteq B$ and every U_i is connected and relatively compact. To prove that $\{U_i\}_{i\in I}$ is a locally finite cover of M, we observe that $\{U_i\}_{i\in I}$ covers M by construction. Now $K := C_2 \setminus U_{i_o} \subseteq U$ is compact, whence only finitely

many elements of the locally finite cover $\{U_j\}_{j\in J}$ intersect it. This means that only finitely many of the sets $U_j, j \in J$ intersect U_{i_o}, whence $\{U_i\}_{i\in I}$ is locally finite. □

We now prove Lemma 3.2 whose statement we repeat here for convenience.

A.2 Lemma *For $\varepsilon > 0$ and (M, g) a connected complete Riemannian manifold, there exists $F_\varepsilon \in C^\infty(M)$ with the following properties.*

1. *$F_\varepsilon(o) = 0$ and $F_\varepsilon(x) \geq 0$, $\forall x \in M$,*
2. *if $d(x, o) \geq \varepsilon$, then $F_\varepsilon(x) \geq d(x, o)$.*

Proof. Let $\varepsilon > 0$ and denote by $B := B^d_\varepsilon(o)$ the metric ball of radius ε around o. Apply now Lemma A.1 with the above choice of B to obtain a locally finite open cover $\{U_i\}_{i\in I}$ of M with a unique element U_{i_o} such that $o \in U_{i_o} \subseteq B$. Following [Lan99, II, §3 Corollary 3.8] we pick a smooth partition of unity $\{\chi_i\}_{i\in I}$ subordinate to the cover $\{U_i\}_{i\in I}$. Note that by construction of the cover, we must have $\chi_{i_0}(o) = 1$. Define the constants $M_j := \max\{\varepsilon, \sup_{y\in\overline{U}_j} d(o, y)\}$ for $j \in J$. By compactness of $\overline{U}_j$ and continuity of the Riemannian distance (follows from [Kli95, Theorem 1.9.5]), the M_j are finite. Hence we can build a family of smooth function:

$$f_i(x) := \begin{cases} \varepsilon(1 - \chi_{i_0}(x)) & \text{for } i = i_o \\ M_j\chi_j(x) & \text{for } i = j \in J \end{cases} \qquad i \in I.$$

Observe now that since the $\{\chi_i\}_{i\in I}$ is a partition of unity, their supports form a locally finite family $\{\mathrm{supp}\chi_i\}_{i\in I}$. We deduce that the family of supports for the functions f_i is also locally finite, whence we can define a smooth function

$$F_\varepsilon(x) := \sum_{i\in I} f_i(x) \quad x \in M$$

which satisfies $F_\varepsilon(o) = 0$ and $F_\varepsilon(x) \geq 0$ for all $x \in M$. If $x \in M\backslash B$, there is a finite non empty $L_x \subseteq J$ such that $x \in U_i$ if and only if $i \in L_x$. In particular $\sum_{i\in L_x} \chi_i(x) = 1$ and as $x \in U_i$ for every $i \in L_x$ by construction one has $d(x, o) \leq M_i$ for all $i \in L_x$. Thus we deduce that

$$d(x, o) \leq \min_{i\in L_x}\{M_i\} \sum_{i\in L_x} \chi_i(x) \leq \sum_{i\in L_x} M_i\chi_i(x) \leq \sum_{i\in L_x\cup\{i_o\}} f_i(x) = F_\varepsilon(x).$$

□

Finally, we construct a family of smooth functions which allows us to obtain estimates on the Riemannian distance for points close to o. This is Lemma 3.3 whose statement we repeat for the readers convenience.

A.3 Lemma *Let $\varepsilon > 0$ be sufficiently small that the closure of the metric ball $B^d_\varepsilon(o)$ is contained in a manifold chart (U, φ). Then there is $N \in \mathbb{N}$ and a family $\{f_n\}_{1\leq n\leq N} \subseteq C^\infty(M)$ with the following properties*

1. $f_n(o) = 0$,
2. *if* $d(x,o) < \varepsilon$ *then there is* $1 \leq n_x \leq N$ *such that* $f_n(x) \geq d(x,o)$.

Proof. As a homogeneous Riemannian manifold, (M,g) is complete, see 1.4. Thus the closed and bounded set $K := \overline{B^d_\varepsilon(o)}$ is compact by the Hopf-Rinow theorem [Bes08, Theorem 1.65]. We may assume without loss of generality that $\varphi(o) = 0$. Now by standard arguments [8] for every smooth Riemannian manifold the charts are locally bi-Lipschitz to Euclidean space. Since $K \subseteq U$ is compact, we may (after shrinking U if necessary) assume that φ is bi-Lipschitz with respect to the euclidean distance d_2 on $\mathbb{R}^n$ and the geodesic distance on U, i.e.

$$d_2(\varphi(x),\varphi(y)) \lesssim d(x,y) \lesssim d_2(\varphi(x),\varphi(y)), \quad \forall x,y \in U, \tag{9}$$

where "$\lesssim$" is used to denote an inequality up to a (multiplicative) constant. Using the equivalence of norms on $\mathbb{R}^n$, we now replace the euclidean distance d_2 in (9) by the distance d_1, induced by the ℓ^1-norm $\|x\|_1 := \sum_{i=1}^n |x_i|$. We claim now, that there is $N \in \mathbb{N}$ and a family of smooth functions $\{P_n\}_{1\leq n\leq N} \subseteq C^\infty(\varphi(U))$ which satisfy the following properties for all $1 \leq n \leq N$:

1. $P_n(\varphi(o)) = H_n(0) = 0$
2. if $x \in \varphi(K)$ then there exists $1 \leq n_x \leq N$ such that $\|x\| = d_1(x,\varphi(o)) \leq P_{n_x}(x)$.

If this were true, then the proof can be finished as follows: Let L be the (smallest) Lipschitz constant such that $d(x,y) \leq L d_1(\varphi(x),\varphi(y)), \quad \forall x,y \in U$. Since U is an open neighborhood of K, we can choose a smooth cut-off function $\xi\colon M \to [0,1]$ such that $\xi|_K \equiv 1$ and $\xi|_{M\setminus U} \equiv 0$. Then we set

$$f_n\colon M \to \mathbb{R},\ x \mapsto \begin{cases} L\xi(x)\cdot P_n \circ \varphi(x) & \text{if } x \in U \\ 0 & \text{otherwise.} \end{cases},$$

Clearly we have $f_n \in C^\infty(M)$ and $f_n(o) = 0$ for all $1 \leq n \leq N$. If $d(x,o) < \varepsilon$, then $x \in K$, whence there is $n_x := n_{\varphi(x)}$ as in property 2. of the family $\{P_n\}_n$ such that

$$f_{n_x}(x) = L\underbrace{\xi(x)}_{=1}\cdot P_{n_{\varphi(x)}} \circ \underbrace{\varphi(x)}_{\in\varphi(K)} \geq L d_1(\varphi(x),\varphi(o)) \geq d(x,o).$$

Proof of the claim: We have to construct smooth functions which satisfy properties 1. and 2. To this end, consider for $1 \leq k \leq n$ the smooth (linear) functions

$$p_{k,0}\colon \mathbb{R}^n \to \mathbb{R}, \quad (x_1,\dots,x_n) \mapsto x_k, \qquad p_{k,1}\colon \mathbb{R}^n \to \mathbb{R}, \quad p_{k,1}(x) := -p_{k,0}(x)$$

Construct for every multiindex $\alpha = (\alpha_1,\dots,\alpha_n) \in \{0,1\}^n$ a function

$$P_\alpha(x) := \sum_{i=1}^n p_{i,\alpha_i}(x), \quad x \in \mathbb{R}^n$$

[8] see eg. the answer by Benoît Kloeckner at https://mathoverflow.net/a/236851/

Set $N := 2^n$ and choose an arbitrary order of the multiindices α (naming the ith α^i, to define the desired family $P_n := P_{\alpha^n}|_U$ for $1 \leq n \leq N$. Obviously $P_n(0) = 0$ and from the construction it is clear that the P_n satisfy property 2. □

References

[AS15] Alzaareer, H. and Schmeding, A. *Differentiable mappings on products with different degrees of differentiability in the two factors.* Expo. Math. **33** (2015)(2):184–222. doi:10.1016/j.exmath.2014.07.002

[Bes08] Besse, A. L. *Einstein manifolds.* Classics in Mathematics (Springer-Verlag, Berlin, 2008). Reprint of the 1987 edition

[CEMM18] Curry, C., Ebrahimi-Fard, K., Manchon, D. and Munthe-Kaas, H. Z. *Planarly branched rough paths and rough differential equations on homogeneous spaces* 2018. arXiv:1804.08515v3

[CEOR18] Celledoni, E., Eidnes, S., Owren, B. and Ringholm, T. *Energy preserving methods on Riemannian manifolds* 2018. arXiv:1805.07578v1

[CG93] Crouch, P. E. and Grossman, R. *Numerical integration of ordinary differential equations on manifolds.* J. Nonlinear Sci. **3** (1993)(1):1–33. doi:10.1007/BF02429858

[CMO03] Celledoni, E., Marthinsen, A. and Owren, B. *Commutator-free lie group methods.* Future Gener. Comput. Syst. **19** (2003)(3):341–352. doi:10.1016/S0167-739X(02)00161-9. URL http://dx.doi.org/10.1016/S0167-739X(02)00161-9

[CMO14] Celledoni, E., Marthinsen, H. k. and Owren, B. *An introduction to Lie group integrators—basics, new developments and applications.* J. Comput. Phys. **257** (2014)(part B):1040–1061. doi:10.1016/j.jcp.2012.12.031

[ELM15] Ebrahimi-Fard, K., Lundervold, A. and Munthe-Kaas, H. Z. *On the Lie enveloping algebra of a post-Lie algebra.* J. Lie Theory **25** (2015)(4):1139–1165

[Fal00] Faltinsen, S. *Backward error analysis for Lie-group methods.* BIT **40** (2000)(4):652–670. doi:10.1023/A:1022336301001

[HNW08] Hairer, E., Nørsett, S. and Wanner, G. *Solving Ordinary Differential Equations I: Nonstiff Problems.* Springer Series in Computational Mathematics (Springer Berlin Heidelberg, 2008)

[HS17] Hjelle, E. O. and Schmeding, A. *Strong topologies for spaces of smooth maps with infinite-dimensional target.* Expo. Math. **35** (2017)(1):13–53. doi:10.1016/j.exmath.2016.07.004

[IMNZ00] Iserles, A., Munthe-Kaas, H. Z., Nørsett, S. P. and Zanna, A. *Lie-group methods*. In *Acta numerica, 2000*, *Acta Numer.*, vol. 9, pp. 215–365 (Cambridge Univ. Press, Cambridge, 2000). doi:10.1017/S0962492900002154

[Kli95] Klingenberg, W. P. A. *Riemannian geometry*, *De Gruyter Studies in Mathematics*, vol. 1 (Walter de Gruyter & Co., Berlin, 1995), second edn. doi:10.1515/9783110905120

[KN96] Kobayashi, S. and Nomizu, K. *Foundations of differential geometry. Vol 1 and Vol. II*. Wiley Classics Library (John Wiley & Sons, Inc., New York, 1996). Reprint of the 1969 original, A Wiley-Interscience Publication

[KSSV06] Kunzinger, M., Schichl, H., Steinbauer, R. and Vickers, J. A. *Global Gronwall estimates for integral curves on Riemannian manifolds*. Rev. Mat. Complut. **19** (2006)(1):133–137. doi:10.5209/rev_REMA.2006.v19.n1.16639

[Lan99] Lang, S. *Fundamentals of differential geometry*, *Graduate Texts in Mathematics*, vol. 191 (Springer-Verlag, New York, 1999). doi:10.1007/978-1-4612-0541-8

[MK95] Munthe-Kaas, H. *Lie-Butcher theory for Runge-Kutta methods*. BIT **35** (1995)(4):572–587. doi:10.1007/BF01739828

[MK98] Munthe-Kaas, H. *Runge-Kutta methods on Lie groups*. BIT **38** (1998)(1):92–111. doi:10.1007/BF02510919

[MK99] Munthe-Kaas, H. *High order Runge-Kutta methods on manifolds*. In *Proceedings of the NSF/CBMS Regional Conference on Numerical Analysis of Hamiltonian Differential Equations (Golden, CO, 1997)*, vol. 29, pp. 115–127 (1999). doi:10.1016/S0168-9274(98)00030-0

[MKZ97] Munthe-Kaas, H. and Zanna, A. *Numerical integration of differential equations on homogeneous manifolds*. In *Foundations of computational mathematics (Rio de Janeiro, 1997)*, pp. 305–315 (Springer, Berlin, 1997)

[MO99] Marthinsen, A. and Owren, B. *Runge-Kutta methods adapted to manifolds and based on rigid frames*. BIT **39** (1999)(1):116–142. doi:10.1023/A:1022325426017

[MW08] Munthe-Kaas, H. Z. and Wright, W. M. *On the Hopf algebraic structure of Lie group integrators*. Found. Comput. Math. **8** (2008)(2):227–257. doi:10.1007/s10208-006-0222-5

[OR04] Ortega, J.-P. and Ratiu, T. S. *Momentum maps and Hamiltonian reduction*, *Progress in Mathematics*, vol. 222 (Birkhäuser Boston, Inc., Boston, MA, 2004). doi:10.1007/978-1-4757-3811-7

[Owr06] Owren, B. *Order conditions for commutator-free lie group methods* **39** (2006):5585