Newton Method for Finding a Singularity of a Special Class of Locally Lipschitz Continuous Vector Fields on Riemannian Manifolds

Fabiana R. de Oliveira · Orizon P. Ferreira

Received: 8 October 2018 / Accepted: 12 March 2020 / Published online: 30 March 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract
We extend some results of nonsmooth analysis from the Euclidean context to the Riemannian setting. Particularly, we discuss the concepts and some properties, such as the Clarke generalized covariant derivative, upper semicontinuity, and Rademacher theorem, of locally Lipschitz continuous vector fields on Riemannian settings. In addition, we present a version of the Newton method for finding a singularity of a special class of locally Lipschitz continuous vector fields. For mild conditions, we establish the well-definedness and local convergence of the sequence generated using the method in a neighborhood of a singularity.

Keywords
Riemannian manifold · Locally Lipschitz continuous vector fields · Clarke generalized covariant derivative · Semismooth vector field · Regularity · Newton method

Mathematics Subject Classification 90C30 · 49J52 · 90C56

1 Introduction
During the last decade, there has been an increase in the number of studies that have proposed the extensions of the concepts and techniques of nonsmooth analysis, as well as nonsmooth methods, from Euclidean context to the Riemannian setting. The aforementioned studies include [1–7]. In addition, different methods, such as gradient...
sampling in [6] and trust region method in [4], have been proposed for solving nonsmooth optimization problems on the Riemannian manifolds setting. One of the main reasons for the increasing interest to develop theoretical and computational tools is the fact that nonsmooth optimization problems are observed in various applications, such as computer vision, robotics, and signal processing; see [2,8]. Although the interest in nonsmooth analysis in the Riemannian setting has increased, only a few studies exist on nonsmooth vector fields in this context; see [3,9]. Therefore, the development of a nonsmooth theory for vector fields might be of significant interest.

In this study, we discuss the concepts and some properties, such as the Clarke generalized covariant derivative, upper semicontinuity, and Rademacher theorem, of locally Lipschitz continuous vector fields on Riemannian settings. Moreover, a version of the Newton method for finding a singularity of locally Lipschitz continuous vector fields is presented. To present our method, we first define the Clarke generalized covariant derivative, which can be considered a natural generalization to the Riemannian setting of the Clarke generalized Jacobian, studied in [10]. The concept of the Clarke generalized covariant derivative has previously appeared in [3,9]. However, in this study, we show its existence using a version of the Rademacher theorem in the Riemannian setting, which is one of our contributions.

The Newton method is considerably popular for finding a singularity of a differentiable vector field, the origins of which go back to the work of [11]; see also [12–16]. This method became popular because of its attractive convergence properties under suitable assumptions. For instance, in all the previously cited works, the superlinear and/or quadratic local convergence of the sequence generated using the Newton method has been established under invertibility of the covariant derivative of the vector field at its singularity and/or Lipschitz-like conditions on the covariant derivative of the vector field. Recently, [17] established a local convergence analysis for the Newton method under the invertibility assumption of the covariant derivative of the vector field at its singularity; however, any Lipschitz-like conditions were necessary to perform the analysis. In the Newton method, the vector field is replaced by an approximation that depends on the current iterate, and then, the original problem is converted to an approximated problem, which can be solved more easily than the original problem. The solution of the approximated problem is taken as a new iterate, and the process is repeated. The success of the Newton method for finding a singularity of a differentiable vector field motivates us to study the Newton method for finding a singularity of a locally Lipschitz continuous vector field. The foundation of our method is similar to that of the classical case; however, in the approximated problem, we combine the exponential mapping on the manifold with an element of the Clarke generalized covariant derivative of the vector field, as the derivative covariant of a locally Lipschitz continuous vector field may not exist. Notably, when the vector field is continuously differentiable, our method reduces to the classical Newton method. From the theoretical viewpoint, we present the local convergence analysis of the proposed method under mild assumptions.

This paper is organized as follows. In Sect. 2, some notations and basic results are presented. In Sect. 3, we generalize some results of nonsmooth analysis to the Riemannian context; particularly, we establish the Rademacher theorem, introduce the Clarke generalized covariant derivative associated with a locally Lipschitz continuous
vector field and study its main properties. In Sect. 4, we describe the Newton method and establish its convergence theorems. In Sect. 5, we present a class of examples of locally Lipschitz continuous vector field satisfying the assumptions of the convergence theorems. We conclude this paper with some remarks in Sect. 6.

2 Notations and Auxiliary Results

Here, we recall some notations, definitions, and basic properties of Riemannian manifolds used throughout this paper. They can be found in many books on Riemannian geometry; see, for example [18–20].

A chart on an n-dimensional smooth manifold M is a pair (U, ϕ), where U denotes an open subset of the manifold M and the coordinate mapping ϕ : U → ˜U a smooth homeomorphism from U to an open subset ˜U = ϕ(U) ⊆ R^n. Let N and M be manifolds of finite dimension and F : N → M a continuous mapping. We say that F is smooth at p ∈ N, if there exist smooth charts (U, ϕ) containing p and (W, ψ) containing F(p) such that F(U) ⊆ W and the composite mapping ψ ◦ F o ϕ−1 : ϕ(U) → ψ(W) is smooth at ϕ(p). The definition of the smoothness of a map F : N → M at a point is independent of the choice of charts; see [20, Proposition 6.7, p. 61]. A diffeomorphism of manifolds is a bijective smooth mapping F : N → M whose inverse F−1 is also smooth. According to [20, Proposition 6.10, p. 63], coordinate maps are diffeomorphisms and, particularly, continuously differentiable. Let M be a Riemannian manifold with Riemannian metric denoted by ⟨·, ·⟩ and the corresponding norm by ∥·∥. The length of a piecewise smooth curve γ : [a, b] → M joining p to q in M, i.e., γ(a) = p and γ(b) = q, is denoted by ℓ(γ). The Riemannian distance between p and q is defined as follows:

d(p, q) = infγ∈Γp,q ℓ(γ),

where Γp,q denotes the set of all the piecewise smooth curves in M joining points p and q. This distance induces the original topology on M; that is, (M, d) is a complete metric space and the bounded and closed subsets are compact. The open and closed balls of radius r > 0, centered at p, are, respectively, defined by Br(p) := {q ∈ M : d(p, q) < r} and B̄r(p) := {q ∈ M : d(p, q) ≤ r}.

Consider M an n-dimensional smooth Riemannian manifold. Denote the tangent space at point p by TpM, the tangent bundle by TM := ∪p∈MTpM, and a vector field by a mapping X : M → TM such that X(p) ∈ TpM. Let γ be a curve joining points p and q in M, and let ∇ be the Levi–Civita connection associated to (M, ⟨·, ·⟩). For each t ∈ [a, b], ∇ induces a linear isometry between the tangent spaces Tγ(t)M and Tγ(t)M, relative to ⟨·, ·⟩, defined by Pγ,a,t v = Y(t), where Y denotes the unique vector field on γ such that Yγ(t) Y(t) = 0 and Y(a) = v. The aforementioned isometry is called parallel transport along the segment γ joining γ(a) to γ(t). It can be shown that Pγ,b,t o Pγ,a,b = Pγ,a,t and Pγ,t,a = Pγ,a,t−1, for all a ≤ b ≤ t. For simplicity and convenience, whenever there is no confusion, we consider the notation Pγ,p,q instead of Pγ,a,b, where γ denotes a segment joining p to q, with γ(a) = p and γ(b) = q. We use the short notation Ppq instead of Pγ,p,q whenever there exists a unique geodesic segment joining p to q. For any n-dimensional smooth manifold M, the tangent bundle TM has a natural topology and smooth structure, thereby making it a 2n-dimensional smooth manifold. With respect to this structure, the projection
\( \pi : TM \to M \) is smooth; see [21, Proposition 3.18, p. 66]. The standard Riemannian distance \( d_{TM} \) on the tangent bundle \( TM \) can be defined as follows: given \( u, v \in TM \), then \( d_{TM} \) is defined by

\[
d_{TM}(u, v) := \inf \left\{ \sqrt{\ell^2(\gamma) + \| P_{\gamma, \pi u, \pi v} u - v \|^2} : \gamma \in \Gamma_{\pi u, \pi v} \right\},
\]

where \( \Gamma_{\pi u, \pi v} \) denotes the set of all the piecewise smooth curves in \( M \) joining points \( \pi u \) to \( \pi v \), whose derivative is never zero; see [22, Appendix, p. 240]. A vector field \( Y \) along the smooth curve \( \gamma \) in \( M \) is parallel when \( \nabla_{\gamma'} Y = 0 \). If \( \gamma' \) is parallel, we say that \( \gamma \) is a geodesic. Because the geodesic equation \( \nabla_{\gamma'} \gamma' = 0 \) is a second-order nonlinear ordinary differential equation, the geodesic \( \gamma \) is determined using its position \( p \) and velocity \( v \) at \( p \). It is easy to check that \( \| \gamma' \| \) is constant. The restriction of a geodesic to a closed bounded interval is called a \textit{geodesic segment}. A geodesic segment joining \( p \) to \( q \) in \( M \) is minimal if its length is equal to \( d(p, q) \), and, in this case, it will be denoted by \( \gamma_{pq} \). A Riemannian manifold is complete if its geodesics \( \gamma(t) \) are defined for any value of \( t \in \mathbb{R} \). The Hopf–Rinow theorem asserts that any pair of points in a complete Riemannian manifold \( M \) can be joined by a (not necessarily unique) minimal geodesic segment. Hereinafter, \( M \) denotes an \textit{n-dimensional smooth and complete Riemannian manifold}. Because of the completeness of the Riemannian manifold \( M \), the exponential map at \( p \), \( \exp_p : T_p M \to M \) can be given by \( \exp_p v = \gamma(1) \), where \( \gamma \) denotes the geodesic defined by its position \( p \) and velocity \( v \) at \( p \), and \( \gamma(t) = \exp_p(tv) \) for any value of \( t \). The inverse of the exponential map (if exists) is denoted by \( \exp_p^{-1} \). Let \( p \in M \), the injectivity radius of \( M \) at \( p \) is defined by \( r_p := \sup\{ r > 0 : \exp_p|_{B_r(0)} \) is a diffeomorphism\}, where \( 0_p \) denotes the origin of the \( T_p M \), and \( B_r(0_p) := \{ v \in T_p M : \| v - 0_p \| < r \} \). A neighborhood \( \mathcal{W} \) of \( p \in M \) is a normal neighborhood of \( p \) if there exists a neighborhood \( \mathcal{U} \) of the origin in \( T_p M \) such that \( \exp_p : \mathcal{U} \to \mathcal{W} \) is a diffeomorphism. Furthermore, if \( \mathcal{W} \) is a normal neighborhood of each of its points, \( \mathcal{W} \) is a totally normal neighborhood.

\textbf{Remark 2.1} For \( \tilde{p} \in M \), the above definition implies that if \( 0 < \delta < r_{\tilde{p}} \), then \( \exp_{\tilde{p}} B_{\delta}(0) = B_{\delta}(\tilde{p}) \) is a totally normal neighborhood. Therefore, for all \( p, q \in B_\delta(\tilde{p}) \), there exists a unique geodesic segment \( \gamma \) joining \( p \) to \( q \), given by \( \gamma_{pq}(t) = \exp_p(t \exp_{p}^{-1} q) \) for all \( t \in [0, 1] \) and \( d(p, q) = \| \exp_p^{-1} q \| \).

Next, we present a quantity that plays an important role in the sequel; it was defined in [23].

\textbf{Definition 2.1} Let \( p \in M \) and \( r_p \) be the radius of injectivity of \( M \) at \( p \). We define the quantity as follows:

\[
K_p := \sup \left\{ \frac{d(\exp_q u, \exp_q v)}{\| u - v \|} : q \in B_{r_p}(p), \ u, v \in T_q M, \ u \neq v, \ \| v \| \leq r_p, \ \| u - v \| \leq r_p \right\}.
\]

In the following remark, we show that an estimative for the value of \( K_p \) can be found for Riemannian manifolds with nonnegative sectional curvature.
Remark 2.2 The number $K_p$ measures how fast the geodesics spread apart in $M$. Particularly, when $u = 0$ or, more generally, when $u$ and $v$ are on the same line through 0, then $d(\exp_q u, \exp_v v) = \|u - v\|$. Therefore, $K_p \geq 1$ for all $p \in M$. When $M$ has nonnegative sectional curvature, the geodesics spread apart less than the rays [24, Chapter 5], i.e., $d(\exp_p u, \exp_p v) \leq \|u - v\|$; in this case, $K_p = 1$ for all $p \in M$.

The directional derivative of $X$ at $p$ along the direction $v \in T_pM$ is defined by

$$\nabla X(p, v) := \lim_{t \downarrow 0} \frac{1}{t} \left[ P_{\exp_p((tv)_p)} X(\exp_p(tv)) - X(p) \right] \in T_pM,$$

whenever the limit exists, where $P_{\exp_p((tv)_p)}$ denotes the parallel transport along $\gamma(t) = \exp_p(tv)$. If this directional derivative exists for every $v$, the vector field $X$ is directionally differentiable at $p$. Let $\mathcal{X}(M)$ denote the space of the differentiable vector fields on $M$. For each $X \in \mathcal{X}(M)$, the covariant derivative of $X$ determined using the connection $\nabla$ defines at each $p \in M$ a linear map $\nabla X \colon T_pM \rightarrow T_pM$ given by $\nabla X(p)v := \nabla Y X(p)$, where $Y$ denotes a vector field such that $Y(p) = v$. Furthermore, $\nabla X(p, v) = \nabla X(p)v$; see [25, Proposition 3, p. 234]. To state the next result, we must define the norm of a linear mapping.

Definition 2.2 Let $p \in M$; the norm of a linear mapping $A : T_pM \rightarrow T_pM$ is defined by

$$|A| := \sup \left\{ \|Av\| : v \in T_pM, \|v\| = 1 \right\}.$$

We end this section with the well-known Banach lemma. For the proof of the lemma, see [26, Lemma 2.3.2, p. 45].

Lemma 2.1 Let $A$ and $B$ be linear operators in $T_pM$. If $A$ is nonsingular and $\|A^{-1}\| \|B - A\| < 1$, then $B$ is also nonsingular, and $\|B^{-1}\| \leq \|A^{-1}\|/(1 - \|A^{-1}\|\,(B - A))$.

3 Nonsmooth Analysis

In this section, we aim to extend some basic results of the nonsmooth analysis in linear context to the Riemannian setting. Particularly, we study the basic properties of the locally Lipschitz continuous vector fields in the Riemannian setting, namely a generalization of Rademacher theorem, and introduce the concept of the Clarke generalized covariant derivative to this new context. A comprehensive study of the nonsmooth analysis in linear context can be found in [10]. First, we define locally Lipschitz continuous vector fields; this concept was introduced in [27] for gradient vector fields, and its extension to general vector fields can be found in [22, p. 241].

Definition 3.1 A vector field $X$ on $M$ is Lipschitz continuous on $\Omega \subset M$, if there exists a constant $L > 0$ such that for all $p, q \in \Omega$ and all $\gamma$ geodesic segment joining $p$ to $q$, the following holds:
Given \( p \in M \), if there exists \( \delta > 0 \) such that \( X \) is Lipschitz continuous on \( B_\delta(p) \), then \( X \) is Lipschitz continuous at \( p \). Moreover, if for all \( p \in M \), \( X \) is Lipschitz continuous at \( p \), then \( X \) is locally Lipschitz continuous on \( M \).

Let \( d_{TM} \) be the Riemannian distance on \( TM \). Let us define the concept of Lipschitz continuity of vector field as a map between metric spaces \((M, d)\) and \((TM, d_{TM})\). The following is the formal definition:

**Definition 3.2** A vector field \( X \) on \( M \) is metrically Lipschitz continuous on \( \Omega \subseteq M \), if there exists a constant \( L > 0 \) such that \( d_{TM}(X(p), X(q)) \leq L d(p, q) \) for all \( p, q \in \Omega \). Given \( p \in M \), if there exists \( \delta > 0 \) such that \( X \) is metrically Lipschitz continuous on \( B_\delta(p) \), then \( X \) is metrically Lipschitz continuous at \( p \). Moreover, if for all \( p \in M \), \( X \) is metrically Lipschitz continuous at \( p \), then \( X \) is locally metric Lipschitz continuous on \( M \).

It is evident from the last definition that all metrically Lipschitz continuous vector fields are continuous. In the next result, we present a relationship between the above-mentioned two definitions.

**Theorem 3.1** If \( X \) is Lipschitz continuous with constant \( L > 0 \), then \( X \) is also metrically Lipschitz continuous with constant \( \sqrt{1 + L^2} \). Consequently, if \( X \) is locally Lipschitz continuous on \( M \), then \( X \) is also locally metric Lipschitz continuous on \( M \).

**Proof** Because \( M \) is a complete manifold, \( \pi X(p) = p \) and \( \pi X(q) = q \), it follows from (1) that

\[
d_{TM}(X(p), X(q)) \leq \sqrt{d^2(p, q) + \| P_{\gamma, p, q} X(p) - X(q) \|^2}, \quad \forall p, q \in M,
\]  

where \( \gamma \) denotes the minimal geodesic segment joining \( p \) to \( q \). Because \( X \) is Lipschitz continuous with constant \( L > 0 \) from Definition 3.1, we have \( \| P_{\gamma, p, q} X(p) - X(q) \| \leq L d(p, q) \) for all \( p, q \in M \). Therefore, inequality (3) becomes \( d_{TM}(X(p), X(q)) \leq \sqrt{1 + L^2} d(p, q) \) for all \( p, q \in M \). Consequently, using Definition 3.2, we conclude that \( X \) is metrically Lipschitz continuous with constant \( \sqrt{1 + L^2} \). Therefore, the proof of the first part is complete. The proof of the second part is similar to that of the first part. \( \Box \)

In the next definition, we present the notion of sets of measure zero to manifolds \([19,21]\).

**Definition 3.3** A subset \( E \subseteq M \) has measure zero in \( M \) if for each smooth chart \((U, \varphi)\) for \( M \), the subset \( \varphi(E \cap U) \subseteq \mathbb{R}^n \) has an \( n \)-dimensional measure zero.

Let \( X \) be a locally Lipschitz continuous vector field on \( M \). Throughout this paper, \( D_X \) is the set defined as follows: \( D_X := \{ p \in M : X \text{ is differentiable at } p \} \). Although locally Lipschitz continuous vector fields are, generally, nondifferentiable, they are
almost everywhere differentiable with respect to the Riemannian measure (see the concept of the Riemannian measure in [19, p. 61]); that is, the set \( M \setminus D_X \) has measure zero. The aforementioned result follows from the Rademacher theorem. A version of this theorem for locally Lipschitz continuous vector fields is given in the following.

**Theorem 3.2** If \( X \) is a locally Lipschitz continuous vector field on \( M \), then \( X \) is almost everywhere differentiable on \( M \).

**Proof** Because \( M \) is an \( n \)-dimensional smooth manifold, \( TM \) is a \( 2n \)-dimensional smooth manifold. First, Theorem 3.1 implies that \( X \) is continuous. Let \((U, \varphi)\) and \((W, \psi)\) be smooth charts for \( M \) and \( TM \), respectively, such that \( X(U) \subseteq W \) and consider the composite mapping \( \psi \circ X \circ \varphi^{-1} : \varphi(U) \to \psi(W) \). We prove that the mapping \( \psi \circ X \circ \varphi^{-1} \) is locally Lipschitz continuous on \( \varphi(U) \). According to [20, Proposition 6.10, p. 63], we obtain that the coordinate mappings \( \varphi^{-1} : \varphi(U) \to U \) and \( \psi : W \to \psi(W) \) are diffeomorphisms and, particularly, continuously differentiable. Consider \( z \in \varphi(U) \) and \( \rho > 0 \) such that \( B_\rho[z] \subset \varphi(U) \). Because \( B_\rho[z] \) is compact and because the derivative of \( \varphi^{-1} \) is continuous in \( B_\rho[z] \), from the mean value inequality (see [1, Theorem 2.14]), there exists \( L_1 > 0 \) such that \( d(\varphi^{-1}(x), \varphi^{-1}(y)) \leq L_1 d(x, y) \), for all \( x \) and \( y \in B_\rho[z] \), where \( d \) denotes the Euclidean distance in \( \mathbb{R}^n \). However, Theorem 3.1 implies that \( X \) is locally metric Lipschitz continuous on \( \varphi(U) \); subsequently, upon shrinking \( \rho > 0 \), if necessary, we conclude that there exists \( L_2 > 0 \) such that \( d_{TM}(X \circ \varphi^{-1}(x), X \circ \varphi^{-1}(y)) \leq L_2 d(\varphi^{-1}(x), \varphi^{-1}(y)) \), for all \( x \) and \( y \in B_\rho[z] \). Because \( X(\varphi^{-1}(B_\rho[z])) \) is compact and because the derivative of \( \psi \) is continuous in \( X(\varphi^{-1}(B_\rho[z])) \), again using the mean value inequality (see [1, Theorem 2.14]), there exists \( L_3 > 0 \) such that

\[
\tilde{d}(\psi \circ X \circ \varphi^{-1}(x), \psi \circ X \circ \varphi^{-1}(y)) \leq L_3 d_{TM}(X \circ \varphi^{-1}(x), X \circ \varphi^{-1}(y)),
\]

for all \( x \) and \( y \in B_\rho[z] \), where \( \tilde{d} \) denotes the Euclidean distance in \( \mathbb{R}^{2n} \). Combining the last three inequalities, we obtain that

\[
\tilde{d}(\psi \circ X \circ \varphi^{-1}(x), \psi \circ X \circ \varphi^{-1}(y)) \leq \tilde{L} \tilde{d}(x, y), \quad \forall x, y \in B_\rho[z],
\]

where \( \tilde{L} = L_1 L_2 L_3 > 0 \). Accordingly, \( \psi \circ X \circ \varphi^{-1} \) is locally Lipschitz continuous on \( \varphi(U) \subseteq \mathbb{R}^n \). Therefore, using the Rademacher theorem (see [28, Theorem 2, p. 81]), we obtain that \( \psi \circ X \circ \varphi^{-1} \) is almost everywhere differentiable on \( \varphi(U) \). Because charts \((U, \varphi)\) and \((W, \psi)\) are arbitrary, \( X \) is almost everywhere differentiable on \( M \).

Next, we introduce the **Clarke generalized covariant derivative** of a locally Lipschitz continuous vector field and explore some of its properties. For comprehensively studying the Clarke generalized Jacobian in linear space, see [10].

**Definition 3.4** The Clarke generalized covariant derivative of a locally Lipschitz continuous vector field \( X \) is a set-valued mapping \( \partial X : M \rightrightarrows TM \) defined as

\[
\partial X(p) := \text{conv} \left\{ H \in \mathcal{L}(T_p M) : \exists \{p_k\} \subset D_X, \lim_{k \to +\infty} p_k = p, H = \lim_{k \to +\infty} P_{p_k} p \nabla X(p_k) \right\}.
\]
where “conv” denotes the convex hull and $\mathcal{L}(T_p M)$ the vector space that comprises all the linear operators from $T_p M$ to $T_p M$.

From Definition 3.4 and [17, Corollary 3.1], it is evident that if $X$ is differentiable near $p$, and if its covariant derivative is continuous at $p$, then $\partial X(p) = \{ \nabla X(p) \}$. Otherwise, $\partial X(p)$ could contain other elements that are different from $\nabla X(p)$, even if $X$ is differentiable at $p$ (see [10, Example 2.2.3]). In the following, we shall show important results of the Clarke generalized covariant derivative. Particularly, we shall show that the set $\partial X(p)$ is nonempty for all $p \in M$ and that $\partial X$ is locally bounded and closed, which is a generalization of [10, Proposition 2.6.2, items (a), (b) and (c), p. 70]. The aforementioned results will be considerably useful throughout this paper. Similar results have previously been extended to the functions defined in $M$ (see [29, Theorem 2.9]).

**Proposition 3.1** Let $X$ be a locally Lipschitz continuous vector field on $M$. The following statements are valid for any $p \in M$:

(i) $\partial X(p)$ is a nonempty, convex, and compact subset of $\mathcal{L}(T_p M)$;
(ii) the set-valued mapping $\partial X : M \rightrightarrows T M$ is locally bounded; that is, for all $\delta > 0$, there exists a $L > 0$ such that for all $q \in B_{\delta}(p)$ and $V \in \partial X(q)$, $\|V\| \leq L$ holds;
(iii) the mapping $\partial X$ is upper semicontinuous at $p$; that is, for every scalar $\epsilon > 0$, there exists a $0 < \delta < r_p$, and such that for all $q \in B_{\delta}(p)$,

$$P_{q\delta} \partial X(q) \subset \partial X(p) + B_{\epsilon}(0),$$

where $B_{\epsilon}(0) := \{ V \in \mathcal{L}(T_p M) : \|V\| < \epsilon \}$. Consequently, the set-valued mapping $\partial X$ is closed at $p$; that is, if $\lim_{k \to +\infty} p_k = p$, $V_k \in \partial X(p_k)$ for all $k = 0, 1, \ldots$, and $\lim_{k \to +\infty} P_{p_k p} V_k = V$, then $V \in \partial X(p)$.

**Proof** To prove item (i), we define the auxiliary set as follows:

$$\partial_B X(p) := \left\{ H \in \mathcal{L}(T_p M) : \exists \{ p_k \} \subset D_X, \lim_{k \to +\infty} p_k = p, H = \lim_{k \to +\infty} P_{p_k p} \nabla X(p_k) \right\}.$$ 

Because $T_p M$ is finite dimensional and because $\partial X(p)$ is the convex hull in $\mathcal{L}(T_p M)$ of set $\partial_B X(p)$, $\partial X(p)$ must be convex. Our next goal is to prove that $\partial X(p)$ is compact. Because the convex hull of a compact set be compact, it is sufficient to prove that $\partial_B X(p)$ is bounded and closed. Our first task is to prove that $\partial_B X(p)$ is bounded. Therefore, take $p \in D_X$ and $v \in T_p M$. Because $\nabla X(p) v = \nabla X(p, v)$ using (2), i.e., the fact that $X$ is locally Lipschitz continuous on $M$ and the definition of the exponential mapping, we obtain that

$$\|\nabla X(p) v\| = \lim_{t \downarrow 0} \frac{1}{t} \left\| P_{\exp_p(t v)} p X(p) \exp_p(t v) - X(p) \right\| \leq L \|v\|,$$

where $L > 0$ denotes the Lipschitz constant of $X$ around $p$. Thus, using Definition 2.2, $\|\nabla X(p)\| \leq L$, implying that $\partial_B X(p)$ is bounded. To prove that $\partial_B X(p)$
is closed, let \( \{ H_\ell \} \) be a sequence in \( \partial_B X(p) \) such that \( \lim_{\ell \to +\infty} H_\ell = H \). Because \( \{ H_\ell \} \subset \partial_B X(p) \), there exists a sequence \( \{ p_{k,\ell} \} \) such that \( \lim_{k \to +\infty} p_{k,\ell} = p \), and \( \lim_{k \to +\infty} P_{p_{k,\ell}} \nabla X(p_{k,\ell}) = H_\ell \), for each fixed \( \ell \). Thus, \( \lim_{k \to +\infty} p_{k,k} = p \) and \( \lim_{k \to +\infty} P_{p_{k,k}} \nabla X(p_{k,k}) = H \), and then, \( H \in \partial_B X(p) \). Consequently \( \partial_B X(p) \) is a compact set. To prove that \( \partial X(p) \) is a nonempty set, first note that Theorem 3.2 implies that \( X \) is almost everywhere differentiable on \( M \); that is, the set \( M \setminus \mathcal{D}_X \) has measure zero. According to [21, Proposition 6.8, p. 128], \( \mathcal{D}_X \) is dense in \( M \). Accordingly, for any fixed point \( p \in M \), there exists \( \{ p_k \} \subset \mathcal{D}_X \) that converges to \( p \). Because \( \nabla X \) is bounded in norm by the Lipschitz constant and because the parallel transport is an isometry, the sequence \( \{ P_{p_k} \nabla X(p_k) \} \) must have at least one accumulation point, and, therefore, \( \partial X(p) \) is a nonempty set. To prove item (iii), take \( \delta > 0 \), \( p \in M \), and \( L > 0 \) the Lipschitz constant of \( X \) around \( p \). The same argument used to prove item (i) shows that \( \| \nabla X(\tilde{q}) \| \leq L \) for all \( \tilde{q} \in B_\delta(p) \cap \mathcal{D}_X \). Let \( q \in B_\delta(p) \) and consider \( V \in \partial X(q) \). Therefore, there exist \( H_1, \ldots, H_m \in \partial_B X(q) \) and \( \alpha_1, \ldots, \alpha_m \in [0, 1] \) such that \( V = \sum_{\ell=1}^m \alpha_\ell H_\ell \) and \( \sum_{\ell=1}^m \alpha_\ell = 1 \). Because \( H_1, \ldots, H_m \in \partial_B X(q) \), there exists \( \{ q_{k,\ell} \} \subset B_\delta(p) \cap \mathcal{D}_X \) with \( \lim_{k \to +\infty} q_{k,\ell} = q \) such that \( V = \sum_{\ell=1}^m \alpha_\ell \lim_{k \to +\infty} P_{q_{k,\ell}} \nabla X(q_{k,\ell}) \). Because \( \{ q_{k,\ell} \} \subset B_\delta(p) \cap \mathcal{D}_X \), we have \( \| \nabla X(q_{k,\ell}) \| \leq L \). Therefore, because the parallel transport is an isometry, we conclude that

\[
\| V \| = \left\| \sum_{\ell=1}^m \alpha_\ell \lim_{k \to +\infty} P_{q_{k,\ell}} \nabla X(q_{k,\ell}) \right\| \leq \sum_{\ell=1}^m \alpha_\ell \lim_{k \to +\infty} \left\| P_{q_{k,\ell}} \nabla X(q_{k,\ell}) \right\| \leq L,
\]

which is the desired inequality. To prove item (iii), suppose, by contradiction, that for a given \( \epsilon > 0 \) and all \( 0 < \delta < r_p \), there exists \( q \in B_\delta(p) \) such that \( P_{q_p} \partial X(q) \not\subset \partial X(p) + B_\epsilon(0) \). Accordingly, there exists a sequence \( \{ q_k \} \subset \mathcal{D}_X \) such that \( \lim_{k \to +\infty} q_k = p \) and \( P_{q_k} \nabla X(q_k) \not\in \partial X(p) + B_\epsilon(0) \). However, item (ii) implies that \( \partial X \) is locally bounded. Because the parallel transport is an isometry, \( \{ P_{q_k} \nabla X(q_k) \} \) is bounded. Therefore, we can extract \( \{ P_{q_{k,\ell}} \nabla X(q_{k,\ell}) \} \), which is a convergent subsequence of \( \{ P_{q_k} \nabla X(q_k) \} \); let us say that \( \{ P_{q_{k,\ell}} \nabla X(q_{k,\ell}) \} \) converges to some \( H \). Using Definition 3.4, we obtain that \( H \in \partial X(p) \), which is a contradiction. Therefore, \( \partial X \) is upper semicontinuous at \( p \). The last part of item (iii) is an immediate consequence of the first part, and the proof is complete. \( \square \)

## 4 Newton Method

Here, we present the Newton method for finding a singularity of a vector field \( X \) on \( M \), i.e., to solve the following problem:

\[
\text{find } p \in M \text{ such that } X(p) = 0,
\]

where \( X \) denotes a locally Lipschitz continuous vector field on \( M \). We study the local properties of the sequence generated using the method. In the following, we formally state the Newton method to solve problem (4).

\( \copyright \) Springer
Algorithm 1: Newton method

Step 0. Let $p_0 \in M$ be given, and set $k = 0$.

Step 1. If $X(p_k) = 0$, stop.

Step 2. Choose a $V_k \in \partial X(p_k)$ and compute

$$p_{k+1} = \exp_{p_k}(-V_{k}^{-1}X(p_k)).$$

Step 3. Set $k \leftarrow k + 1$, and go to Step 1.

This method is a natural extension to the Riemannian setting of the Newton method introduced in [30]. Notably, to guarantee the well-definedness of the method, there are two issues that must be considered in each iteration $k$. The Clarke generalized covariant derivative $\partial X(p_k)$ must be nonempty, which has previously been proven in item $(i)$ of Proposition 3.1, and all $V_k \in \partial X(p_k)$ must be nonsingular. In the following section, we study the well-definedness and convergence properties of the Newton method.

4.1 Local Convergence Analysis

Here, we present the local convergence analysis of Algorithm 1. To this end, we assume that $p_* \in M$ is a solution of problem (4). First, we show that under some assumptions, the sequence generated using this algorithm, starting from a suitable neighborhood of $p_*$, is well defined and converges to $p_*$ with the rate of order $1 + \mu$. We first introduce the concept of regularity.

Definition 4.1 A vector field $X$ on $M$ is regular at $p \in M$ if all $V_p \in \partial X(p)$ are nonsingular. If $X$ is regular at every point of $\Omega \subseteq M$, $X$ is regular on $\Omega$.

In the following, we study the behavior of the Newton method for a special class of vector field in a neighborhood of a regular point. To that end, we assume that $X$ is a locally Lipschitz continuous vector field on $M$. Consider the following condition: A1. Let $\bar{p} \in M$, $0 < \delta < r_{\bar{p}}$, $X$ be regular on $B_\delta(\bar{p})$, $\lambda_{\bar{p}} \geq \max\{\|V_{\bar{p}}^{-1}\| : V_{\bar{p}} \in \partial X(\bar{p})\}$, and $\epsilon > 0$ with $\epsilon \lambda_{\bar{p}} < 1$. For all $p \in B_\delta(\bar{p})$ and $V_p \in \partial X(p)$ the following hold:

$$\|V_p^{-1}\| \leq \frac{\lambda_{\bar{p}}}{1 - \epsilon \lambda_{\bar{p}}},$$

$$\|X(\bar{p}) - P_{\bar{p}}\left[ X(p) + V_p \exp_p^{-1} \bar{p} \right]\| \leq \epsilon d(p, \bar{p})^{1+\mu}, \quad 0 \leq \mu \leq 1.$$  

The above-mentioned assumption guarantees, particularly, that $X$ is regular in a neighborhood of $\bar{p}$ and, consequently, that the Newton iteration mapping is well defined. Let $0 < \delta < r_{\bar{p}}$ be given by the above-mentioned assumption and $N_X : B_\delta(\bar{p}) \Rightarrow M$ be the Newton iteration mapping for $X$ defined by

$$N_X(p) := \left\{ \exp_p(-V_p^{-1}X(p)) : V_p \in \partial X(p) \right\}.$$
Therefore, one can apply a single Newton iteration on any \( p \in B_\delta(\hat{p}) \) to obtain \( N_X(p) \), which may be not included in \( B_\delta(\hat{p}) \). Therefore, it is sufficient to guarantee the well-definedness of only one iteration. In the following result, we establish that Newtonian iterations may be repeated indefinitely in a suitable neighborhood of \( \hat{p} \).

Lemma 4.1 Suppose that \( p_* \in M \) is a solution of problem (4), \( X \) satisfies A1 with \( \hat{p} = p_* \), and the constants \( \epsilon > 0, 0 < \delta < r_{p_*} \) and \( 0 \leq \mu \leq 1 \) satisfy \( \epsilon \lambda_{p_*} (1 + \delta \mu K_{p_*}) < 1 \). Then, there exists \( \hat{\delta} > 0 \) such that \( X \) is regular on \( B_{\hat{\delta}}(p_*) \) and

\[
d \left( \exp_p(-V_p^{-1}X(p)), p_* \right) \leq \frac{\epsilon \lambda_{p_*} K_{p_*}}{1 - \epsilon \lambda_{p_*}} d(p, p_*)^{1+\mu}, \tag{8}
\]

for all \( p \in B_{\hat{\delta}}(p_*) \) and for all \( V_p \in \partial X(p) \). Consequently, \( N_X \) is well defined on \( B_{\hat{\delta}}(p_*) \) and we have \( N_X(p) \subset B_{\hat{\delta}}(p_*) \) for all \( p \in B_{\hat{\delta}}(p_*) \).

**Proof** Because \( \epsilon > 0, 0 < \delta < r_{p_*} \) and \( 0 \leq \mu \leq 1 \) are constants such that \( X \) satisfies A1, \( X(p_*) = 0 \), and because the parallel transport is an isometry, we conclude that

\[
\left\| V_p^{-1}X(p) + \exp_p^{-1} p_* \right\| \leq \left\| V_p^{-1}\right\| \left\| X(p_*) - P_{pp_*} \left[ X(p) + V_p \exp_p^{-1} p_* \right] \right\|
\leq \frac{\epsilon \lambda_{p_*} K_{p_*}}{1 - \epsilon \lambda_{p_*}} d(p, p_*)^{1+\mu}, \tag{9}
\]

for all \( p \in B_{\hat{\delta}}(p_*) \) and \( V_p \in \partial X(p) \). Therefore, the inequality (9) implies that there exists \( 0 < \hat{\delta} < \delta \) such that \( \| V_p^{-1}X(p) + \exp_p^{-1} p_* \| \leq r_{p_*} \) for all \( p \in B_{\hat{\delta}}(p_*) \) and \( V_p \in \partial X(p) \). Since \( \| \exp_p^{-1} p_* \| = d(p, p_*) < r_{p_*} \), we use Definition 2.1 with \( p = p_* \), \( q = p \), \( u = -V_p^{-1}X(p) \) and \( v = \exp_p^{-1} p_* \) to obtain that

\[
d \left( \exp_p(-V_p^{-1}X(p)), p_* \right) \leq K_{p_*} \left\| -V_p^{-1}X(p) - \exp_p^{-1} p_* \right\|,
\]

for all \( p \in B_{\hat{\delta}}(p_*) \) and \( V_p \in \partial X(p) \). Therefore, the combination of the last inequality with (9) yields (8). Because \( 0 < \hat{\delta} < \delta \) and the vector field \( X \) is regular on \( B_{\hat{\delta}}(p_*) \), \( N_X \) is well defined on \( B_{\hat{\delta}}(p_*) \). Moreover, because \( \epsilon \lambda_{p_*} (1 + \delta \mu K_{p_*}) < 1 \) and \( 0 < \hat{\delta} < \delta \), from inequality (8), \( d \left( \exp_p(-V_p^{-1}X(p)), p_* \right) < d(p, p_*) \) for all \( p \in B_{\hat{\delta}}(p_*) \) and \( V_p \in \partial X(p) \). Therefore, we obtain \( N_X(p) \subset B_{\hat{\delta}}(p_*) \) for all \( p \in B_{\hat{\delta}}(p_*) \), and the proof of the lemma is complete. ∎

We now establish the main result of this section; its proof is a direct application of Lemma 4.1.

**Theorem 4.1** Suppose that \( p_* \in M \) is a solution of problem (4), \( X \) satisfies A1 with \( \hat{p} = p_* \), and the constants \( \epsilon > 0, 0 < \delta < r_{p_*} \) and \( 0 \leq \mu \leq 1 \) satisfy \( \epsilon \lambda_{p_*} (1 + \delta \mu K_{p_*}) < 1 \). Then, there exists \( 0 < \hat{\delta} < \delta \) such that for each \( p_0 \in B_{\hat{\delta}}(p_*) \), \( \{ p_k \} \) in Algorithm 1 is well defined, belongs to \( B_{\hat{\delta}}(p_*) \), and converges to \( p_* \) with order
$1 + \mu$ as follows:

$$d(p_{k+1}, p_*) \leq \frac{\epsilon \lambda_{p_*} K_{p_*}}{1 - \epsilon \lambda_{p_*}} d(p_k, p_*)^{1+\mu}, \quad k = 0, 1, \ldots$$  \hspace{1cm} (10)

**Proof** The definition of Newton mapping $N_X$ implies that the sequence generated using Algorithm 1 is equivalently stated as

$$p_{k+1} \in N_X(p_k), \quad k = 0, 1, \ldots$$  \hspace{1cm} (11)

Therefore, by using (11), we can apply Lemma 4.1 to conclude that there exists $0 < \tilde{\delta} < \delta$ such that if $p_0 \in B_{\tilde{\delta}}(p_*) \setminus \{p_*\}$, then $\{p_k\}$ in Algorithm 1 is well defined, belongs to $B_{\tilde{\delta}}(p_*)$, and satisfies (10). Because $\{p_k\}$ belongs to $B_{\tilde{\delta}}(p_*)$ and because $\epsilon \lambda_{p_*} (1 + \delta \mu K_{p_*}) < 1$, we obtain from (10) that

$$d(p_{k+1}, p_*) < \frac{\epsilon \lambda_{p_*} \tilde{\delta} \mu K_{p_*}}{1 - \epsilon \lambda_{p_*}} d(p_k, p_*) < d(p_k, p_*), \quad k = 0, 1, \ldots$$

Therefore, $\{p_k\}$ converges to $p_*$ with order $1 + \mu$ as (10). \hspace{1cm} $\square$

**Remark 4.1** If $\mu = 0$ in Theorem 4.1, then (10) holds for any $\epsilon > 0$ that satisfies $\epsilon \lambda_{p_*} (1 + K_{p_*}) < 1$, independent of $\tilde{\delta}$. Therefore, (10) implies that $\{p_k\}$ converges superlinearly to $p_*$. 

### 4.1.1 Local Convergence for Semismooth Vector Fields

Here, we present a local convergence theorem for the Newton method, to find a singularity of semismooth vector fields. Semismoothness in the Euclidean setting was originally introduced by Mifflin [31] for scalar-valued functions and, subsequently, extended by Qi and Sun [30] for vector-valued functions. The extension of semismoothness to Riemannian settings was presented in [32], and it will play an important role in this section. As occur in the Euclidean context, semismooth vector fields are, in general, nonsmooth. However, we shall show that the Newton method is still applicable and converges locally with superlinear rate to a regular solution. Before formally stating the concept of semismoothness in the Riemannian setting, let us first show that locally Lipschitz continuous vector fields are regular near regular points. The following is the statement of the result:

**Lemma 4.2** Let $X$ be a locally Lipschitz continuous vector field on $M$. Assume that $X$ is regular at $p_* \in M$ and let $\lambda_{p_*} \geq \max \{\|V^{-1}_p\| : V_{p_*} \in \partial X(p_*)\}$. Then, for every $\epsilon > 0$ that satisfies $\epsilon \lambda_{p_*} < 1$, there exists $0 < \delta < r_{p_*}$ such that $X$ is regular on $B_{\delta}(p_*)$ and

$$\|V_p^{-1}\| \leq \frac{\lambda_{p_*}}{1 - \epsilon \lambda_{p_*}}, \quad \forall \ p \in B_{\delta}(p_*), \ \forall \ V_p \in \partial X(p).$$  \hspace{1cm} (12)
Proof Let \( \epsilon > 0 \) such that \( \epsilon \lambda_{p_s} < 1 \). Because \( X \) is locally Lipschitz, from Proposition 3.1(iii), there exists a \( 0 < \delta < r_{p_s} \) such that for all \( p \in B_\delta(p_s) \) holds \( P_{pp_s} \partial X(p) \subset \partial X(p_s) + \{ V \in T_{p_s}M : \| V \| < \epsilon \} \); that is,

\[
\partial X(p) \subset \{ V \in T_{p}M : \| P_{pp_s}V - V_{p_s} \| < \epsilon, \text{ for some } V_{p_s} \in \partial X(p_s) \},
\]

for all \( p \in B_\delta(p_s) \). The aforementioned inclusion implies that for each \( p \in B_\delta(p_s) \) and \( V_p \in \partial X(p) \), there exists nonsingular \( V_{p_s} \in \partial X(p_s) \) such that \( \| V_{p_s}^{-1} \| \| P_{pp_s}V_p - V_{p_s} \| < \epsilon \lambda_{p_s} < 1 \). Therefore, considering that the parallel transport is an isometry, it follows from Lemma 2.1 that \( V_p \) is nonsingular and

\[
\| V_{p_s}^{-1} \| \leq \frac{\| V_{p_s}^{-1} \|}{1 - \| P_{pp_s}V_p - V_{p_s} \|}.
\]

Therefore, considering that \( \| V_{p_s}^{-1} \| \leq \lambda_{p_s} \) and \( \| P_{pp_s}V_p - V_{p_s} \| < \epsilon \), the inequality (12) holds. \( \square \)

Let us present a class of vector fields satisfying the assumption A1, i.e., semismooth vector fields and \( \mu \)-order semismooth vector fields. There exist, in the Euclidean context, several equivalent definitions of the concept of semismoothness; see [30]; see also [33, Definition 7.4.2, p. 677]. In this paper, we extend to the Riemannian settings the concept of semismoothness adopted in [34, p. 411].

Definition 4.2 A vector field \( X \) on \( M \), which is Lipschitz continuous at \( p_s \) and directionally differentiable at \( p \in B_\delta(p_s) \) for all directions from \( T_pM \), is semismooth at \( p_s \in M \) if for every \( \epsilon > 0 \) there exists \( 0 < \delta < r_{p_s} \) such that

\[
\| X(p_s) - P_{pp_s}[X(p) + V_p \exp^{-1} p_s] \| \leq \epsilon d(p, p_s),
\]

for all \( p \in B_\delta(p_s) \) and for all \( V_p \in \partial X(p) \). The vector field \( X \) is \( \mu \)-order semismooth at \( p_s \in M \), for \( 0 < \mu \leq 1 \), if there exist \( \epsilon > 0 \) and \( 0 < \delta < r_{p_s} \) such that

\[
\| X(p_s) - P_{pp_s}[X(p) + V_p \exp^{-1} p_s] \| \leq \epsilon d(p, p_s)^{1+\mu},
\]

(13)

for all \( p \in B_\delta(p_s) \) and for all \( V_p \in \partial X(p) \).

Next, we state and prove the local convergence result for the Newton method, to find a singularity of semismooth and \( \mu \)-order semismooth vector fields.

Theorem 4.2 Let \( X \) be a locally Lipschitz continuous vector field on \( M \) and \( p_s \in M \) be a solution of problem (4). Assume that \( X \) is semismooth and regular at \( p_s \). Then, there exists a \( \delta > 0 \) such that for each \( p_0 \in B_\delta(p_s) \setminus \{ p_s \} \), \( \{ p_k \} \) generated using Algorithm 1 is well defined, belongs to \( B_\delta(p_s) \), and converges superlinearly to \( p_s \). Additionally, if \( X \) is \( \mu \)-order semismooth at \( p_s \), then the convergence of \( \{ p_k \} \) to \( p_s \) is of order \( 1 + \mu \).
Proof Because $X$ is semismooth and regular at $p^* \in M$, we can take $\lambda_{p^*} \geq \max\{\|V_{p^*}\| : V_{p^*} \in \partial X(p^*)\}$. Take $\epsilon > 0$ that satisfies $\epsilon \lambda_{p^*}(1 + K_{p^*}) < 1$. Therefore, using Lemma 4.2 and Definition 4.2, we can take $\delta > 0$ such that (6) and (7) hold for $\mu = 0$. Therefore, Assumption A1 holds with $\tilde{p} = p^*$, for all $p \in B_\delta(p^*)$ and $\mu = 0$. Therefore, using Theorem 4.1, we obtain that there exists $0 < \hat{\delta} < \delta$ such that every sequence $\{p_k\}$ generated using Algorithm 1 with $p_0 \in B_{\hat{\delta}}(p^*) \setminus \{p^*\}$ belongs to $B_{\hat{\delta}}(p^*)$ and satisfies (10). Therefore, we have

$$
\frac{d(p_{k+1}, p^*)}{d(p_k, p^*)} \leq \frac{\epsilon \lambda_{p^*} K_{p^*}}{1 - \epsilon \lambda_{p^*}}, \quad k = 0, 1, \ldots
$$

Because the last inequality holds for any $\epsilon$ such that $0 < \epsilon < 1/(\lambda_{p^*}(1 + K_{p^*}))$, $\{p_k\}$ converges superlinearly to $p^*$. The proof of the second part is similar to that of the first part. For a given $\epsilon > 0$ with $\epsilon \lambda_{p^*} < 1$, take $\delta > 0$ that satisfies $\epsilon \lambda_{p^*}(1 + \delta \mu K_{p^*}) < 1$ such that (12) and (13) hold. Accordingly, we can apply Theorem 4.1 and the proof follows.

In the following remark, particularly, we show that by incorporating some adjustments, Theorem 4.2 reduces to some well-known results.

Remark 4.2 The Newton method and its variants are fairly efficient for finding zero on nonlinear functions in Euclidean settings, as they have excellent convergence rate in the neighborhood of a zero. It was shown in [30] that for some class of nonsmooth functions, namely semismooth functions, the convergence of the Newton method was guaranteed. The above theorem enables us conclude that the generalization of the Newton method from the linear context to Riemannian settings for finding the singularities of semismooth vector fields still preserves its main convergence properties. Notably, if $X$ is continuously differentiable, then Theorem 4.2 reduces to [17, Theorem 3.1]. If $M = \mathbb{R}^n$, Theorem 4.2 reduces to the first part of [30, Theorem 3.2]; see also [33, Theorem 7.5.3, p. 693]. Finally, if $X$ is continuously differentiable and $M = \mathbb{R}^n$, the above-mentioned theorem reduces to the first part of [35, Proposition 1.4.1, p. 90].

5 Some Examples

Here, we present a class of examples of locally Lipschitz continuous vector fields on the sphere that satisfies Assumption A1. To that end, we first present some basic definitions regarding the geometry of the sphere. For further details, see [36,37] and the references therein.

Let $\langle \cdot, \cdot \rangle$ be the usual inner product on $\mathbb{R}^{n+1}$, with corresponding norm denoted by $\| \cdot \|$. The $n$-dimensional Euclidean sphere and its tangent hyperplane at a point $p$ are denoted, respectively, by

$$
S^n := \left\{ p = (p_1, \ldots, p_{n+1}) \in \mathbb{R}^{n+1} : \|p\| = 1 \right\},
$$

$$
T_p S^n := \left\{ v \in \mathbb{R}^{n+1} : \langle p, v \rangle = 0 \right\}.
$$
Let $I$ denote the $(n+1) \times (n+1)$ identity matrix. The projection onto the tangent hyperplane, $T_p\mathbb{S}^n$, is the linear mapping defined by $I - pp^T : \mathbb{R}^{n+1} \to T_p\mathbb{S}^n$, where $p^T$ denotes the transpose of vector $p$. Let $\Omega$ denote an open set in $\mathbb{R}^{n+1}$ such that $\mathbb{S}^n \subset \Omega$, and $Y : \Omega \to \mathbb{R}^{n+1}$ be any $\mu$-order semismooth mapping for $0 \leq \mu \leq 1$; several examples can be found in [33,34,38]. Accordingly, we define the vector field $X : \mathbb{S}^n \to \mathbb{R}^{n+1}$ as follows:

$$X(p) := (I - pp^T)Y(p).$$

Notably, $X(p) \in T_p\mathbb{S}^n$ for all $p \in \mathbb{S}^n$. The Clarke generalized covariant derivative of $X$ at $p$ is given by

$$\partial X(p) := \left( I - pp^T \right) \partial Y(p) - p^T Y(p) I,$$

where $\partial Y(p)$ denotes the Clarke generalized covariant derivative of $Y$ at $p$. Therefore, all $V_p \in \partial X(p)$ is a linear mapping $V_p : T_p\mathbb{S}^n \to T_p\mathbb{S}^n$ given by $V_p := (I - pp^T) \tilde{V}_p - p^T Y(p) I$, where $\tilde{V}_p \in \partial Y(p)$. Because $Y$ is a locally Lipschitz continuous mapping, from Rademacher theorem (see [28, Theorem 2, p. 81]), $Y$ is almost everywhere differentiable. Because $I - pp^T$ is a continuously differentiable mapping, $X$ is almost everywhere differentiable. Using the fundamental theorem of calculus in the Riemannian setting (see [39]), because of the fact that $\partial Y(p)$ is locally bounded and the continuity of $Y$, we can prove that $X$ is also a locally Lipschitz continuous vector field. Assume that $X$ is regular at $\tilde{p} \in \Omega$ and let $\lambda_{\tilde{p}} \geq \max\{ ||V_p^{-1}|| : V_p \in \partial X(\tilde{p}) \}$. Accordingly, from Lemma 4.2 for each $\epsilon > 0$ that satisfies $\epsilon \lambda_{\tilde{p}} < 1$, there exists $0 < \delta < \pi$ (where $\pi$ denotes the injectivity radius of $\mathbb{S}^n$) such that $X$ is regular on $B_{\delta}(\tilde{p})$, and for all $p \in B_{\delta}(\tilde{p})$ and $V_p \in \partial X(p)$, the following holds:

$$||V_p^{-1}|| \leq \frac{\lambda_{\tilde{p}}}{1 - \epsilon \lambda_{\tilde{p}}}.$$

Therefore, it is implied that inequality (6) holds. However, because $X$ is a composition of $\mu$-order semismooth mappings for $0 \leq \mu \leq 1$, $X$ is $\mu$-order semismooth for $0 \leq \mu \leq 1$; see [38, Proposition 1.74, p. 54]. Hence, using Definition 4.2, inequality (7) holds with $q = \tilde{p}$. Therefore, the projected vector field $X$ satisfies Assumption $A1$ with $q = \tilde{p}$. In the following, we present two concrete examples.

**Example 5.1** Let $Y : \mathbb{R}^2 \to \mathbb{R}^2$ be a semismooth mapping defined by $Y(p) := Ap - |p| - b$ with matrix $A = \text{diag}(4, 3)$ and vector $b = (b_1, b_2) \in \mathbb{R}^2$, where $\text{diag}(p_1, p_2)$ denotes the $2 \times 2$ diagonal matrix with the $(i,i)$-th entry equal to $p_i$, $i = 1, 2$. Take $\tilde{p} = (0, 1) \in \mathbb{S}$ and note that $Y(\tilde{p}) = (0, 0)$ for $b = (0, 2)$. The Clarke derivative of $Y$ at $\tilde{p}$ is given by $\partial Y(\tilde{p}) = \{ \text{diag}(d, 2) : d \in [3, 5] \}$. Define $X(p) := (I - pp^T)Y(p)$ as the vector field on $\mathbb{S}$. Therefore, using (14), we have $\partial X(\tilde{p}) = \{ V_{\tilde{p}} := \text{diag}(d - 2 + b_2, -2 + b_2) : d \in [3, 5] \}$. Notably, for all $b = (b_1, b_2) \in \mathbb{R}^2$ such that $b_2 \neq 2$ and $b_2 \neq 2 - d$, all $V_{\tilde{p}} \in \partial X(\tilde{p})$ are nonsingular as a linear mapping $V_{\tilde{p}} : T_{\tilde{p}}\mathbb{S} \to T_{\tilde{p}}\mathbb{S}$, where the tangent hyperplane at $\tilde{p}$ is given by $T_{\tilde{p}}\mathbb{S} := \{ v := (v_1, 0) \in \mathbb{R}^2 : v_1 \in \mathbb{R} \}$. Therefore, using Definition 4.1, $X$ is regular at $\tilde{p} =$
(0, 1). Let \( \lambda_{\bar{p}} \geq \max\{\|V_{\bar{p}}^{-1}\| : V_{\bar{p}} \in \text{partial} X(\bar{p})\} \). Because \( X \) is a locally Lipschitz continuous vector field, using Lemma 4.2 for every \( \epsilon > 0 \) that satisfies \( \epsilon \lambda_{\bar{p}} < 1 \), there exists \( 0 < \delta < \pi \) such that \( X \) is regular on \( B_\delta(\bar{p}) \), and for all \( p \in B_\delta(\bar{p}) \) and \( V_p \in \partial X(p) \), the following hold: \( \|V_p^{-1}\| \leq \lambda_{\bar{p}}/(1-\epsilon \lambda_{\bar{p}}) \). Because \( X \) is a semismooth vector field, it satisfies Assumption A1 with \( \bar{p} = (0, 1) \) for all \( p \in B_\delta(\bar{p}) \) and \( \mu = 0 \). Particularly, \( X \) satisfies the assumptions of Theorem 4.2 with \( p_* = (0, 1) \) and \( \mu = 0 \).

In the following, we present an example of a 1-order semismooth vector field.

**Example 5.2** Let \( \tau > 0 \) be a parameter and \( Y_\tau : \mathbb{R}^2 \to \mathbb{R}^2 \) be a 1-order semismooth mapping defined by \( Y_\tau(p) := \|p - \bar{p}\|/\tau + 2(p - \bar{p}) \) (see [38, Example 1.68, p. 52] and [38, Proposition 1.73, p. 53]), where \( \bar{p} \in \mathbb{R}^2 \) and \( \| \cdot \| \) is the Euclidean norm. Take \( \bar{p} \in \mathbb{S} \) and note that \( Y_\tau(\bar{p}) = (0, 0) \). Particularly, the Clarke derivative of \( Y \) at \( \bar{p} \) is given by

\[
\partial Y_\tau(\bar{p}) = \left\{ \tau \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} + \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} : \|w\| \leq 1, \|v\| \leq 1 \right\}.
\]

Define \( X_\tau(p) := (I - pp^T)Y_\tau(p) \) as the vector field on the one-dimensional sphere \( \mathbb{S} \). Using (14) and \( Y(\bar{p}) = (0, 0) \), we have \( \partial X_\tau(\bar{p}) = (I - \bar{p}\bar{p}^T) \partial Y_\tau(\bar{p}) \). Thus, letting \( \tau = 0 \), we obtain \( \partial X_0(\bar{p}) = 2(I - \bar{p}\bar{p}^T) \). Notably, all \( V_{\bar{p}} \in \partial X_0(\bar{p}) \) are nonsingular as a linear mapping \( V_{\bar{p}} : T_{\bar{p}}\mathbb{S} \to T_{\bar{p}}\mathbb{S} \), where the tangent hyperplane at \( \bar{p} \) is given by \( T_{\bar{p}}\mathbb{S} := \{ v \in \mathbb{R}^2 : \langle v, \bar{p} \rangle = 0 \} \). Notably, \( V_{\bar{p}}v = 2v \) for all \( v \in T_{\bar{p}}\mathbb{S} \). Therefore, there exists \( \hat{\tau} > 0 \) such that \( V_{\bar{p}} \in \partial X_{\hat{\tau}}(\bar{p}) \) are nonsingular as a linear mapping \( V_{\bar{p}} : T_{\bar{p}}\mathbb{S} \to T_{\bar{p}}\mathbb{S} \), and using Definition 4.1, we obtain that \( X_{\hat{\tau}} \) is regular at \( \bar{p} \). Let \( \lambda_{\bar{p}} \geq \max\{\|V_{\bar{p}}^{-1}\| : V_{\bar{p}} \in \partial X_{\hat{\tau}}(\bar{p})\} \). Because \( X_{\hat{\tau}} \) is a locally Lipschitz continuous vector field, using Lemma 4.2 for each \( \epsilon > 0 \) that satisfies \( \epsilon \lambda_{\bar{p}} < 1 \), there exists \( 0 < \delta < \pi \) such that \( X_{\hat{\tau}} \) is regular on \( B_\delta(\bar{p}) \), and for all \( p \in B_\delta(\bar{p}) \) and \( V_p \in \partial X_{\hat{\tau}}(p) \) the following hold: \( \|V_p^{-1}\| \leq \lambda_{\bar{p}}/(1-\epsilon \lambda_{\bar{p}}) \). However, because \( (I - pp^T) \) is continuously differentiable, we conclude on the basis of [33, Proposition 7.4.5, p. 682] that \( (I - pp^T) \) is 1-order semismooth at \( \bar{p} \). Therefore, using [38, Proposition 1.74, p. 54], we obtain that \( X_{\hat{\tau}} \) is a 1-order semismooth vector field. Therefore, Assumption A1 holds for \( \mu = 1 \). Because \( X_{\hat{\tau}} \) is a 1-order semismooth vector field, it satisfies Assumption A1 with \( q = \bar{p} \) for all \( p \in B_\delta(\bar{p}) \) and \( \mu = 1 \). Particularly, \( X \) satisfies the assumptions of Theorem 4.2 with \( p_* = \bar{p} \) and \( \mu = 1 \).

Notably, in the literature, there exist other examples of semismooth vector field; see, for example [6].

**6 Conclusions**

We studied the concept and some properties of locally Lipschitz continuous vector fields. Notably, the Rademacher theorem is an essential tool to ensure the existence of the Clarke generalized covariant derivative. Additionally, a version of the Newton method for finding a singularity of these vector fields was proposed. Assuming the
conditions of regularity and semismoothness, the well-definedness and local convergence of the method were established. We expect that the results this paper might aid in the establishment of new results and methods of nonsmooth analysis to the Riemannian context, for example, both the mean value theorem and inexact and globalized versions of the Newton method.

Acknowledgements This work was supported by CAPES, FAPEG, and CNPq Grants 408151/2016-1 and 302473/2017-3.

References

1. Azagra, D., Ferrera, J., López-Mesas, F.: Nonsmooth analysis and Hamilton–Jacobi equations on Riemannian manifolds. J. Funct. Anal. 220(2), 304–361 (2005)
2. Bačák, M., Bergmann, R., Steidl, G., Weinmann, A.: A second order nonsmooth variational model for restoring manifold-valued images. SIAM J. Sci. Comput. 38(1), A567–A597 (2016)
3. Ghahraei, E., Hosseini, S., Pouryayevali, M.R.: Pseudo–Jacobian and characterization of monotone vector fields on Riemannian manifolds. J. Convex Anal. 24(1), 149–168 (2017)
4. Grohs, P., Hosseini, S.: Nonsmooth trust region algorithms for locally Lipschitz functions on Riemannian manifolds.IMA J. Numer. Anal. 36(3), 1167–1192 (2016)
5. Hosseini, S., Pouryayevali, M.R.: Nonsmooth optimization techniques on Riemannian manifolds. J. Optim. Theory Appl. 158(2), 328–342 (2013)
6. Hosseini, S., Uschmajew, A.: A Riemannian gradient sampling algorithm for nonsmooth optimization on manifolds. SIAM J. Optim. 27(1), 173–189 (2017)
7. Ledyae, Y.S., Zhu, Q.J.: Nonsmooth analysis on smooth manifolds. Trans. Am. Math. Soc. 359(8), 3687–3732 (2007)
8. Bergmann, R., Weinmann, A.: A second-order TV-type approach for inpainting and denoising higher dimensional combined cyclic and vector space data. J. Math. Imaging Vision 55(3), 401–427 (2016)
9. Rampazzo, F., Sussmann, H.J.: Commutators of flow maps of nonsmooth vector fields. J. Differ. Equ. 232(1), 134–175 (2007)
10. Clarke, F.H.: Optimization and Nonsmooth Analysis, Classics in Applied Mathematics, vol. 5, 2nd edn. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (1990)
11. Shub, M.: Some Remarks on Dynamical Systems and Numerical Analysis. In: Dynamical Systems and Partial Differential Equations (Caracas, 1984), pp. 69–91. Universidad Simon Bolivar, Caracas (1986)
12. Gabay, D.: Minimizing a differentiable function over a differential manifold. J. Optim. Theory Appl. 37(2), 177–219 (1982)
13. Li, C., Wang, J.: Newton’s method on riemannian manifolds: Smale’s point estimate theory under the $\gamma$-condition. IMA J. Numer. Anal. 26(2), 228–251 (2006)
14. Smith, S.T.: Optimization techniques on Riemannian manifolds. In: Hamiltonian and gradient flows, algorithms and control. Fields Inst. Commun., vol. 3, pp. 113–136. American Mathematical Society, Providence, RI (1994)
15. Udište, C.: Convex Functions and Optimization Methods on Riemannian Manifolds, Mathematics and Its Applications, vol. 297. Kluwer Academic Publishers Group, Dordrecht (1994)
16. Wang, J.H.: Convergence of Newton’s method for sections on Riemannian manifolds. J. Optim. Theory Appl. 148(1), 125–145 (2011)
17. Fernandes, T.A., Ferreira, O.P., Yuan, J.: On the superlinear convergence of Newton’s method on Riemannian manifolds. J. Optim. Theory Appl. 173(3), 828–843 (2017)
18. Lang, S.: Differential and Riemannian manifolds, Graduate Texts in Mathematics, vol. 160, 3rd edn. Springer, New York (1995)
19. Sakai, T.: Riemannian Geometry, Translations of Mathematical Monographs, vol. 149. American Mathematical Society, Providence, RI (1996). Translated from the 1992 Japanese original by the author
20. Tu, L.W.: An Introduction to Manifolds, 2nd edn. Universitext. Springer, New York (2011)
21. Lee, J.M.: Introduction to Smooth Manifolds. Graduate Texts in Mathematics, vol. 218, 2nd edn. Springer, New York (2013)
22. Canary, R.D., Epstein, D., Marden, A.: Fundamentals of Hyperbolic Geometry: Selected Expositions, London Mathematical Society Lecture Note Series, vol. 328. Cambridge University Press, Cambridge (2006)
23. Dedieu, J.P., Priouret, P., Malajovich, G.: Newton’s method on Riemannian manifolds: covariant alpha theory. IMA J. Numer. Anal. 23(3), 395–419 (2003)
24. do Carmo, M.P.: Riemannian Geometry. Mathematics: Theory and Applications. Birkhäuser Boston, Inc., Boston, MA (1992). Translated from the second Portuguese edition by Francis Flaherty
25. Spivak, M.: A Comprehensive Introduction to Differential Geometry. Vol. II, 3rd edn. Publish or Perish, Inc., Wilmington, DE (1999)
26. Ortega, J.M., Rheinboldt, W.C.: Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York (1970)
27. da Cruz Neto, J.X., de Lima, L.L., Oliveira, P.R.: Geodesic algorithms in Riemannian geometry. Balkan J. Geom. Appl. 3(2), 89–100 (1998)
28. Evans, L.C., Gariepy, R.F.: Measure Theory and Fine Properties of Functions. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL (1992)
29. Hosseini, S., Pouryayevali, M.R.: Generalized gradients and characterization of epi-Lipschitz sets in Riemannian manifolds. Nonlinear Anal. 74(12), 3884–3895 (2011)
30. Qi, L., Sun, J.: A nonsmooth version of Newton’s method. Math. Program. 58(3), 353–367 (1993)
31. Mifflin, R.: Semismooth and semiconvex functions in constrained optimization. SIAM J. Control Optim. 15(6), 959–972 (1977)
32. Lageman, C., Helmke, U., Manton, J.H.: The nonsmooth Newton method on Riemannian manifolds. Unpublished Notes (https://scholar.lib.vt.edu/MTNS/Papers/108.pdf)
33. Facchinei, F., Pang, J.S.: Finite-Dimensional Variational Inequalities and Complementarity Problems. Springer Series in Operations Research, vol. II. Springer, New York (2003)
34. Dontchev, A.L., Rockafellar, R.T.: Implicit functions and solution mappings. 2nd edn. Springer Series in Operations Research and Financial Engineering. Springer, New York (2014). A view from variational analysis
35. Bertsekas, D.P.: Nonlinear Programming. Athena Scientific Optimization and Computation Series, 3rd edn. Athena Scientific, Belmont, MA (2016)
36. Ferreira, O.P., Iusem, A.N., Németh, S.Z.: Projections onto convex sets on the sphere. J. Optim. Theory Appl. 57(3), 663–676 (2013)
37. Ferreira, O.P., Iusem, A.N., Németh, S.Z.: Concepts and techniques of optimization on the sphere. TOP 22(3), 1148–1170 (2014)
38. Izmailov, A.F., Solodov, M.V.: Newton-Type Methods for Optimization and Variational Problems. Springer Series in Operations Research and Financial Engineering. Springer, Cham (2014)
39. Ferreira, O.P., Svaiter, B.F.: Kantorovich’s theorem on Newton’s method in Riemannian manifolds. J. Complex. 18(1), 304–329 (2002)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.