On the Superlinear Convergence of Newton’s Method on Riemannian Manifolds

Teles A. Fernandes¹ · Orizon P. Ferreira² · Jinyun Yuan³

Received: 10 November 2016 / Accepted: 21 March 2017 / Published online: 28 March 2017
© Springer Science+Business Media New York 2017

Abstract In this paper, we study Newton’s method for finding a singularity of a differentiable vector field defined on a Riemannian manifold. Under the assumption of invertibility of the covariant derivative of the vector field at its singularity, we show that Newton’s method is well defined in a suitable neighborhood of this singularity. Moreover, we show that the sequence generated by Newton’s method converges to the solution with superlinear rate.

Keywords Riemannian manifold · Newton’s method · Local convergence · Superlinear rate

Mathematics Subject Classification 90C30 · 49M15 · 65K05

Communicated by Alexandru Kristály.

Teles A. Fernandes
 telesfernandes@uesb.edu.br

Orizon P. Ferreira
 orison@ufg.br

Jinyun Yuan
 jin@ufpr.br

¹ Universidade Estadual do Sudoeste da Bahia, Vitória da Conquista, BA 45083-900, Brazil
² IME, Universidade Federal de Goiás, Goiânia, GO 74001-970, Brazil
³ Universidade Federal do Paraná, Curitiba, PR 81531-980, Brazil
1 Introduction

Applications of the concepts of Riemannian geometry in optimization arise when optimization problems are formulated as problems of finding minimizers of real-valued functions defined on smooth nonlinear manifolds. Indeed, many optimization problems are naturally posed on Riemannian manifolds, which have specific underlying geometric and algebraic structures that can be exploited to significantly reduce the computational cost of obtaining the minimizer. For instance, to exploit the Riemannian geometric structure, it is preferable to treat certain constrained optimization problems as problems of finding singularities of gradient vector fields on Riemannian manifolds rather than using Lagrange multipliers or projection methods; see [1, 2]. Accordingly, constrained optimization problems are regarded as unconstrained ones from the viewpoint of Riemannian geometry. Early studies on this issue include [3–6]. Recent years have witnessed a growing interest in the development of geometric optimization algorithms that exploit the differential structure of nonlinear manifolds. Papers published on this topic include [7–14]. In the present paper, instead of focusing on finding singularities of gradient vector fields on Riemannian manifolds, which includes finding local minimizers, we consider the more general problem of finding singularities of vector fields.

It is well known that Newton’s method is a powerful tool for finding zeros of nonlinear functions in Banach spaces. Moreover, Newton’s method serves as a powerful theoretical tool with a wide range of applications in pure and applied mathematics; see, for example, [15–17]. These factors have motivated several studies to investigate the issue of generalizing Newton’s method from the linear to the Riemannian setting; see [2, 18–24], among others. It is worth noting that, in all these previous papers, the analysis of the Riemannian version of Newton’s method involves Lipschitz or Lipschitz-like conditions on the covariant derivative of the vector field. In fact, all these papers are concerned with the establishment of the quadratic rate of convergence of the method. It seems that some type of control on the covariant derivative of the vector field, in a suitable neighborhood of its singularity, is required for obtaining the quadratic convergence rate of the sequence generated by Newton’s method. Hence, the convergence analysis of Newton’s method for finding a singularity of a vector field in Riemannian manifolds under Lipschitz or Lipschitz-like conditions is well known. However, we also know that in the linear context, whenever the derivative of the function that defines the equation is nonsingular at the solution, Newton’s method shows local convergence with superlinear rate, [25, chapter 8, Theorem 8.1.10, p. 148]. To the best of our knowledge, the local superlinear convergence analysis of Newton’s method under a mild assumption, namely only invertibility of the covariant derivative of the vector field at its singularity, is a novel contribution of this paper.

The remainder of this paper is organized as follows. Section 2 presents the notations and basic results used in the paper. Section 3 describes the local superlinear convergence analysis of Newton’s method. Section 4 presents two concrete examples to support the main result. Finally, Sect. 5 concludes the paper.
2 Basic Definition and Auxiliary Results

In this section, we recall some notations, definitions, and basic properties of Riemannian manifolds used throughout the paper, which can be found in many introductory books on Riemannian geometry, for example [27] and [28].

For a smooth manifold \( M \), denote the tangent space of \( M \) at \( p \) by \( T_pM \) and the tangent bundle of \( M \) by \( TM = \bigcup_{p \in M} T_pM \). The corresponding norm associated with the Riemannian metric \( \langle \cdot, \cdot \rangle \) is denoted by \( \| \cdot \| \). The Riemannian distance between \( p \) and \( q \) in a finite-dimensional Riemannian manifold \( M \) is denoted by \( d(p,q) \), and it induces the original topology on \( M \). An open ball of radius \( r > 0 \) centered at \( p \) is defined as \( B_r(p) := \{ q \in M : d(p,q) < r \} \). Let \( \Omega \subset M \) be an open set, and let \( \mathcal{X}(\Omega) \) denote the space of \( C^1 \) vector fields on \( \Omega \). Let \( \nabla \) be the Levi-Civita connection associated with \( (M, \langle \cdot, \cdot \rangle) \). The covariant derivative of \( X \in \mathcal{X}(\Omega) \) determined by \( \nabla \) defines at each \( p \in \Omega \) a linear map \( \nabla X(p) : T_pM \to T_pM \) given by \( \nabla X(p) v := \nabla_Y X(p) \), where \( Y \) is a vector field such that \( Y(p) = v \). For \( f : M \to \mathbb{R} \), a twice-differentiable function the Riemannian metric induces the mappings \( f \mapsto \nabla f \) and \( f \mapsto \text{Hessian} f \), which associate its gradient and Hessian via the rules

\[
\langle \nabla f, X \rangle := df(X), \quad \langle \text{Hessian} f, X, X \rangle := d^2f(X,X), \quad \forall X \in \mathcal{X}(\Omega),
\]

respectively. Therefore, the last equalities imply that

\[
\text{Hessian} f X = \nabla_X \nabla f, \quad \forall X \in \mathcal{X}(\Omega).
\]

The norm of a linear map \( A : T_pM \to T_pM \) is defined by \( \| A \| := \sup \{ \| Av \| : v \in T_pM, \| v \| = 1 \} \). A vector field \( V \) along a differentiable curve \( \gamma \) in \( M \) is said to be parallel iff \( \nabla_{\gamma'} V = 0 \). If \( \gamma' \) itself is parallel, we say that \( \gamma \) is a geodesic. The restriction of a geodesic to a closed, bounded interval is called a geodesic segment. A geodesic segment joining \( p \) to \( q \) in \( M \) is said to be minimal iff its length is equal to \( d(p,q) \). If there exists a unique geodesic segment joining \( p \) to \( q \), then we denote it by \( \gamma_{pq} \). For each \( t \in [a, b] \), \( \nabla \) induces an isometry relative to \( \langle \cdot, \cdot \rangle \), \( P_{\gamma,a,t} : T_{\gamma(a)}M \to T_{\gamma(t)}M \), defined by \( P_{\gamma,a,t} v = V(t) \), where \( V \) is the unique vector field on \( \gamma \) such that

\[
\nabla_{\gamma'(t)} V(t) = 0, \quad V(a) = v,
\]

the so-called parallel transport along the geodesic segment \( \gamma \) joining the points \( \gamma(a) \) and \( \gamma(t) \). Further, note that \( P_{\gamma,b_1,b_2} \circ P_{\gamma,a,b_1} = P_{\gamma,a,b_2} \) and \( P_{\gamma,b,a} = P_{\gamma,a,b}^{-1} \). As long as there is no confusion, we will consider the notation \( P_{pq} \) instead of \( P_{\gamma,a,b} \) when \( \gamma \) is the unique geodesic segment joining \( p \) and \( q \). A Riemannian manifold is complete iff the geodesics are defined for all values of \( t \in \mathbb{R} \). The Hopf-Rinow’s theorem asserts that any pair of points in a complete Riemannian manifold \( M \) can be joined by a (not necessarily unique) minimal geodesic segment. Owing to the completeness of the Riemannian manifold \( M \), the exponential map \( \exp_p : T_pM \to M \) can be given by \( \exp_p v = \gamma(1) \) for each \( p \in M \), where \( \gamma \) is the geodesic defined by its starting point \( p \) and velocity \( v \) at \( p \). For \( p \in M \), the injectivity radius of \( M \) at \( p \) is defined by...
where $0_p$ denotes the origin of $T_p\mathbb{M}$ and $B_r(0_p) := \{ v \in T_p\mathbb{M} : \| v - 0_p \| < r \}$, is called neighborhood of injectivity of $p$.

**Remark 2.1** Let $\bar{p} \in \mathbb{M}$. The above definition implies that if $0 < \delta < i_{\bar{p}}$, then $\exp_{\bar{p}} B_\delta(0_{\bar{p}}) = B_\delta(\bar{p})$. Moreover, for all $p \in B_\delta(\bar{p})$, there exists a unique geodesic segment $\gamma$ joining $p$ to $\bar{p}$, which is given by $\gamma_{p\bar{p}}(t) = \exp_p(t \exp_{\bar{p}}^{-1} \bar{p})$, for all $t \in [0, 1]$.

Next, we present the number $K_p$, introduced in [26], which measures how fast the geodesics spread apart in $\mathbb{M}$. Let $i_p$, $p \in \mathbb{M}$, be the radius of injectivity of $\mathbb{M}$ at $p$ and define the quantity

$$\delta_p := \min\{1, i_p\}.$$  

Consider the quantity given by

$$K_p := \sup \left\{ \frac{d(\exp_q u, \exp_q v)}{\| u - v \|} : q \in B_{\delta_p}(p), u, v \in T_q\mathbb{M}, u \neq v, \| v \| \leq \delta_p, \| u - v \| \leq \delta_p \right\}. \quad (1)$$

**Remark 2.2** In particular, when $u = 0$, or, more generally, when $u$ and $v$ are on the same line passing through $0$, $d(\exp_q u, \exp_q v) = \| u - v \|$. Hence, $K_p \geq 1$ for all $p \in \mathbb{M}$. Moreover, when $\mathbb{M}$ has nonnegative sectional curvature, the geodesics spread apart less than the rays [27, chapter 5], i.e., $d(\exp_p u, \exp_p v) \leq \| u - v \|$, and, in this case, $K_p = 1$ for all $p \in \mathbb{M}$.

Let $X \in \mathcal{X}(\Omega)$ and $\bar{p} \in \Omega$. Assume that $0 < \delta < \delta_{\bar{p}}$. Since $\exp_{\bar{p}} B_\delta(0_{\bar{p}}) = B_\delta(\bar{p})$, there exists a unique geodesic joining each $p \in B_\delta(\bar{p})$ to $\bar{p}$. Moreover, using [23, equality 2.3], we obtain

$$X(p) = P_{\bar{p}p} X(\bar{p}) + P_{\bar{p}p} \nabla X(\bar{p}) \exp_{\bar{p}}^{-1} p + d(p, \bar{p}) r(p), \quad \lim_{p \to \bar{p}} r(p) = 0, \quad (2)$$

for each $p \in B_\delta(\bar{p})$. We end this section with the well-known Banach’s Lemma.

**Lemma 2.1** Let $B$ be a linear operator, and let $I_p$ be the identity operator in $T_p\mathbb{M}$. If $\| B - I_p \| < 1$, then $B$ is invertible and $\| B^{-1} \| \leq \frac{1}{1 - \| B - I_p \|}$.

Throughout the paper, $\mathbb{M}$ is a complete Riemannian manifold of finite dimension.
3 Superlinear Convergence of Newton’s Method

In this section, we study Newton’s method to find a point \( p \in \Omega \) satisfying the equation

\[
X(p) = 0, \quad (3)
\]

where \( X : \Omega \to TM \) is a differentiable vector field and \( \Omega \subset M \) is an open set. To solve (3), Newton’s method formally generates a sequence with an initial point \( p_0 \in \Omega \) as follows:

\[
p_{k+1} = \exp_{p_k} \left( -\nabla X (p_k)^{-1} X (p_k) \right), \quad k = 0, 1, \ldots \quad (4)
\]

Hereinafter, we assume that \( p_* \in \Omega \) is a solution of (3). Our aim is to prove that, under the assumption of nonsingularity of the covariant derivative at the solution \( p_* \), the iteration (4) starting in a suitable neighborhood of \( p_* \) is well defined and converges superlinearly to \( p_* \). To obtain this result, we begin by stating an important property of the parallel transport in our context. It is worth noting that, to ensure this property, we use the same ideas as those presented in the proof of [9, Lemma 2.4, item (iv)] for Hadamard manifolds, with some minor necessary technical adjustments to fit any Riemannian manifold.

Lemma 3.1 Let \( \bar{p} \in \bar{M}, 0 < \delta < \delta_{\bar{p}} \) and \( u \in T_{\bar{p}}\bar{M} \). Then, the vector field \( F : B_{\delta}(\bar{p}) \to T\bar{M} \) defined by \( F(p) := P_{\bar{p}p}u \) is continuous.

Proof Assume that \( \bar{M} \) is \( n \)-dimensional. Let \( p \in B_{\delta}(\bar{p}) \) and \( \gamma_p \) be the unique geodesic segment joining \( \bar{p} \) to \( p \) (according to Remark 2.1). Let \( u \in T_{\bar{p}}\bar{M} \). From the definition of the parallel transport, there is a unique continuously differentiable vector field \( Y_p \) along \( \gamma_p \) such that \( Y_p (\gamma_p(0)) = u, \ Y_p (\gamma_p(1)) = P_{\bar{p}p}u \) and

\[
\nabla_{\gamma_p'(t)} Y_p (\gamma_p(t)) = 0, \quad \forall \ t \in [0, 1]; \quad (5)
\]

see [28, pag. 29]. The definition of \( \delta_{\bar{p}} \) implies that \( \varphi := \exp_{\bar{p}}^{-1} : B_{\delta}(\bar{p}) \to B_{\bar{p}}(0) \) is a diffeomorphism; hence, \( (B_{\delta}(\bar{p}), \varphi) \) is a local chart at \( \bar{p} \). For each \( j = 1, 2, \ldots, n \), define \( y^j : B_{\delta}(\bar{p}) \to \mathbb{R} \) by \( y^j = \pi^j \circ \varphi \), where \( \pi^j : T_{\bar{p}}\bar{M} \to \mathbb{R} \) is the projection defined by \( \pi^j(a_1, \ldots, a_j, \ldots, a_n) = a_j \) for all \( (a_1, \ldots, a_j, \ldots, a_n) \in T_{\bar{p}}\bar{M} \). Then, \( (B_{\delta}(\bar{p}), \varphi, y^j) \) is a local coordinate system at \( \bar{p} \). Let \( \{\partial/\partial y^j\} \) be the associated correspondent natural basis to \( (B_{\delta}(\bar{p}), \varphi, y^j) \). Since \( \gamma_p(t) \in B_{\delta}(\bar{p}) \) for all \( t \in [0, 1] \) and \( Y_p (\gamma_p(t)) \in T_{\gamma_p(t)}\bar{M} \), we can write

\[
Y_p (\gamma_p(t)) = \sum_j Y_p^j(t) \frac{\partial}{\partial y^j} |_{\gamma_p(t)} \quad \forall \ t \in [0, 1],
\]

where each coordinate function \( Y_p^j : [0, 1] \to \mathbb{R} \) is continuously differentiable for all \( j = 1, 2, \ldots, n \). For simplicity, we set \( y_p^j := y^j \circ \gamma_p(\cdot) \) for each \( j = 1, 2, \ldots, n \). Thus, (5) is equivalent to the ordinary differential equation:
\[
\frac{dY^k_p}{dt} + \sum_{i,j} \Gamma^k_{i,j}(\gamma_p) \frac{dy^i_p}{dt} Y^j_p = 0, \quad k = 1, 2, \ldots, n,
\]

where \( \Gamma^k_{i,j} \) are the Christoffel symbols of the connection \( \nabla \); see [28, pag. 29]. Hence, the last equality implies that \( \{Y^k_p : k = 1, 2, \ldots, n\} \) is the unique solution of the following system of \( p \)-parameter linear differential equations

\[
\frac{dY^k_p}{dt} = -\sum_j a^j_{k,j} Y^j_p, \quad k = 1, 2, \ldots, n,
\]

\[
\sum_j Y^j_p(0) \frac{\partial}{\partial y^j} |_{y^j(0)} = u,
\]

where, for \((k, j), k, j = 1, \ldots, n\), the continuous function \( a^j_{k,j} : [0, 1] \times B_\delta(\bar{p}) \to \mathbb{R} \) is given by

\[
a^j_{k,j}(t, p) = \sum_{i=1}^n \Gamma^k_{i,j}(\gamma_p(t)) \frac{dy^i_p(t)}{dt}.
\]

Thus, from the continuity on parameters for differential equations (see, e.g., [29, Theorem 10.7.1, pag. 353]), the solution \( \{Y^k(\cdot)\} \) is continuous on \([0, 1] \times B_\delta(\bar{p})\), and, equivalently, \( Y(\gamma(\cdot)) \) is continuous on \([0, 1] \times B_\delta(\bar{p})\). Furthermore, we have \( F(p) = P_{\bar{p}p} u = Y_p(\gamma_p(1)) \), for any \( p \in B_\delta(\bar{p}) \). Therefore, \( F \) is continuous on \( B_\delta(\bar{p}) \), and the proof is completed. \( \Box \)

Next, we present an immediate consequence of Lemma 3.1.

**Corollary 3.1** Let \( \bar{p} \in M, 0 < \delta < \delta_{\bar{p}} \) and \( u \in T_{\bar{p}}M \). If the vector field \( Z : B_\delta(\bar{p}) \to T_{\bar{p}}M \) is continuous at \( \bar{p} \), then the mapping \( G : B_\delta(\bar{p}) \to T_{\bar{p}}M \) defined by \( G(p) := P_{\bar{p}p} Z(p) \) is also continuous at \( \bar{p} \).

**Proof** Since the parallel transport is an isometry, it follows from the definition of the vector field \( G \) that

\[
\|G(p) - G(\bar{p})\| = \|Z(p) - P_{\bar{p}p}Z(\bar{p})\|.
\]

Considering that \( Z \) is continuous at \( \bar{p} \) and \( P_{\bar{p}p} = I_{\bar{p}} = P_{\bar{p}\bar{p}} P_{\bar{p}p} \), we conclude from Lemma 3.1 that

\[
\lim_{p \to \bar{p}} \|Z(p) - P_{\bar{p}p}Z(\bar{p})\| = 0.
\]

Therefore, the desired result follows by simple combination of the two last equalities. \( \Box \)
The next result ensures that if $\nabla X(p_*)$ is nonsingular, then there exists a neighborhood of $p_*$ where $\nabla X$ is also nonsingular. Moreover, in this neighborhood, $\nabla X^{-1}$ is bounded.

**Lemma 3.2** Assume that $\nabla X$ is continuous at $p_*$. Then,

$$\lim_{p \to p_*} \left\| P_{pp_*} \nabla X(p) P_{p_*p} - \nabla X(p_*) \right\| = 0. \quad (6)$$

Moreover, if $\nabla X(p_*)$ is nonsingular, then there exists $0 < \delta < \delta_{p_*}$ such that $B_\delta(p_*) \subset \Omega$, and for each $p \in B_\delta(p_*)$, the following hold:

(i) $\nabla X(p)$ is nonsingular;
(ii) $\left\| \nabla X(p)^{-1} \right\| \leq 2 \left\| \nabla X(p_*)^{-1} \right\|$.

**Proof** Let $0 < \delta < \delta_{p_*}$ such that $B_\delta(p_*) \subset \Omega$. For each $u \in T_{p_*}M$, define $Z : B_\delta(p_*) \to T^* M$ by

$$Z(p) = \nabla X(p) P_{p_*p} u.$$ 

By applying Lemma 3.1, we conclude that $P_{p_*p}u$ is continuous on $B_\delta(p_*)$. Thus, because $\nabla X$ is continuous, $Z$ is also continuous on $B_\delta(p_*)$. Hence, using Corollary 3.1, we conclude that the mapping $F : B_\delta(p_*) \to T_{p_*}M$ defined by

$$F(p) = P_{pp_*} Z(p)$$

is also continuous at $p_*$. Considering that $P_{p_*p} = I_{p_*}$ as well as the definitions of the mappings $F$ and $Z$, we conclude that $\lim_{p \to p_*} F(p) = \nabla X(p_*) u$. Now, define the mapping

$$B_\delta(p_*) \ni p \mapsto \left[ P_{pp_*} \nabla X(p) P_{p_*p} - \nabla X(p_*) \right] \in \mathcal{L}(T_{p_*}M, T_{p_*}M),$$

where $\mathcal{L}(T_{p_*}M, T_{p_*}M)$ denotes the space consisting of all linear operator from $T_{p_*}M$ to $T_{p_*}M$. Since $\lim_{p \to p_*} F(p) = \nabla X(p_*)u$, for each $u \in T_{p_*}M$, the definition of $F$ implies that

$$\lim_{p \to p_*} \left[ P_{pp_*} \nabla X(p) P_{p_*p} - \nabla X(p_*) \right] u = 0, \quad u \in T_{p_*}M.$$ 

Owing to the fact that $T_{p_*}M$ is finite-dimensional and $\left[ P_{pp_*} \nabla X(p) P_{p_*p} - \nabla X(p_*) \right] \in \mathcal{L}(T_{p_*}M, T_{p_*}M)$, for each $p \in B_\delta(p_*)$, the latter equality implies that the equality $(6)$ holds. Now, we proceed with the proof of item (i). The equality $(6)$ implies that there exists $0 < \tilde{\delta} < \delta$ such that

$$\left\| P_{pp_*} \nabla X(p) P_{p_*p} - \nabla X(p_*) \right\| \leq \frac{1}{2 \left\| \nabla X(p_*)^{-1} \right\|}, \quad \forall p \in B_{\tilde{\delta}}(p_*).$$
Thus, from the last inequality and the property of the operator norm defined in \( \mathcal{L}(T_p\mathbb{M}, T_p\mathbb{M}) \), for all \( p \in B_\delta(p_*) \), we obtain

\[
\begin{align*}
\left\| \nabla X(p_*)^{-1} P_{p*} \nabla X(p) P_{p,p} - I_{p*} \right\| & \leq \left\| \nabla X(p_*)^{-1} \right\| \left\| P_{p*} \nabla X(p) P_{p,p} - \nabla X(p) \right\| \leq \frac{1}{2}. \\
(7)
\end{align*}
\]

Hence, from Lemma 2.1, we conclude that \( \nabla X(p_*)^{-1} P_{p*} \nabla X(p) P_{p,p} \) is a nonsingular operator for each \( p \in B_\delta(p_*) \). Owing to the fact that \( \nabla X(p_*) \) and the parallel transport are nonsingular, \( \nabla X(p) \) is also nonsingular for each \( p \in B_\delta(p_*) \), and the proof of the first item is completed. To prove item (ii), we first note that, from (7) and Lemma 2.1, it follows that for all \( p \in B_\delta(p_*) \),

\[
\left\| \left[ \nabla X(p_*)^{-1} P_{p*} \nabla X(p) P_{p,p} \right]^{-1} \right\| \leq \frac{1}{1 - \left\| \nabla X(p_*)^{-1} P_{p*} \nabla X(p) P_{p,p} - I_{p*} \right\|}.
\]

Since the parallel transport is an isometry, by combining (7) with the latter inequality, we obtain

\[
\left\| \nabla X(p) \right\| \leq 2.
\]

Thus, using the properties of the norm and since the parallel transport is an isometry, the last inequality implies that, for all \( p \in B_\delta(p_*) \), we have

\[
\left\| \nabla X(p)^{-1} \right\| \leq \left\| \nabla X(p)^{-1} P_{p,p} \nabla X(p_*) \right\| \left\| \nabla X(p_*)^{-1} \right\| \leq 2 \left\| \nabla X(p_*)^{-1} \right\|,
\]

which is the desired inequality in the second item. Thus, the proof of the lemma is completed. \( \square \)

Lemma 3.2 establishes the nonsingularity of \( \nabla X \) in a neighborhood of \( p_* \). It ensures that there exists a neighborhood of \( p_* \) where Newton’s iterate (4) is well defined, but it does not guarantee that it belongs to this neighborhood. In the next lemma, we will establish this fact. For stating the next result, we first define Newton’s iterate mapping \( N_X : B_\delta(p_*) \to \mathbb{M} \) by

\[
N_X(p) := \exp_p(-\nabla X(p)^{-1} X(p)),
\]

where \( \delta \) is given by Lemma 3.2.

**Lemma 3.3** Assume that \( \nabla X \) is continuous at \( p_* \) and \( \nabla X(p_*) \) is nonsingular. Then,

\[
\lim_{p \to p_*} \frac{d(N_X(p), p_*)}{d(p, p_*)} = 0.
\]

\( \square \) Springer
Proof Let $\bar{\delta}$ be given by Lemma 3.2 and $p \in B_{\bar{\delta}}(p_{*})$. Some algebraic manipulations show that

$$\nabla X(p)^{-1} X(p) + \exp^{-1}_{p} p^* = \nabla X(p)^{-1} \left[ X(p) - P_{p_{*}p} X(p_{*}) - P_{p_{*}p} \nabla X(p_{*}) \exp^{-1}_{p_{*}p} p + \left[ P_{p_{*}p} \nabla X(p_{*}) - \nabla X(p) P_{p_{*}p} \right] \exp^{-1}_{p} p \right].$$

Define $r(p) := \left[ X(p) - P_{p_{*}p} X(p_{*}) - P_{p_{*}p} \nabla X(p_{*}) \exp^{-1}_{p_{*}p} p \right] / d(p, p_{*})$ for $p \in B_{\bar{\delta}}(p_{*})$. From (2), we have $\lim_{p \to p_{*}} r(p) = 0$. Thus, using the above equality, the definition of $r$, $d(p, p_{*}) = \| \exp^{-1}_{p} p \|$, and some properties of the norm, we conclude that

$$\left\| \nabla X(p)^{-1} X(p) + \exp^{-1}_{p} p^* \right\| \leq \left\| \nabla X(p)^{-1} \left[ \|r(p)\| + \| P_{p_{*}p} \nabla X(p_{*}) - \nabla X(p) P_{p_{*}p} \| \right] d(p, p_{*}) \right\|. \quad (9)$$

Owing to (6) and $\lim_{p \to p_{*}} r(p) = 0$, the right-hand side of the last inequality tends to zero, as $p$ goes to $p_{*}$. Recalling that $\delta_{p_{*}} = \min\{1, i_{p_{*}}\}$, we can shrink $\bar{\delta}$, if necessary, to obtain

$$\left\| \nabla X(p)^{-1} X(p) + \exp^{-1}_{p} p^* \right\| \leq \delta_{p_{*}}, \quad \forall p \in B_{\bar{\delta}}(p_{*}).$$

Hence, from the definition of Newton’s iterate mapping $N_{X}$ in (8) and the definition of $K_{p_{*}}$ in (1), we have

$$d(N_{X}(p), p_{*}) \leq K_{p_{*}} \left\| -\nabla X(p)^{-1} X(p) - \exp^{-1}_{p} p^* \right\|, \quad \forall p \in B_{\bar{\delta}}(p_{*}).$$

Therefore, by combining (9) with the last inequality, we conclude that for all $p \in B_{\bar{\delta}}(p_{*})$,

$$\frac{d(N_{X}(p), p_{*})}{d(p, p_{*})} \leq 2K_{p_{*}} \left\| \nabla X(p_{*})^{-1} \left[ \|r(p)\| + \| P_{p_{*}p_{*}} \nabla X(p_{*}) P_{p_{*}p_{*}} - \nabla X(p_{*}) \| \right] \right\|. \quad \Box$$

By letting $p$ tend to $p_{*}$ in the last inequality, and by considering (6) and that $\lim_{p \to p_{*}} r(p) = 0$, the desired result follows.

Now, we are ready to establish our main result, whose proof is a combination of the two previous lemmas.
Theorem 3.1 Let $\Omega \subset \mathbb{M}$ be an open set, $X : \Omega \to T\mathbb{M}$ a differentiable vector field and $p_* \in \Omega$. Suppose that $p_*$ is a singularity of $X$, $\nabla X$ is continuous at $p_*$, and $\nabla X(p_*)$ is nonsingular. Then, there exists $\delta > 0$ such that, for all $p_0 \in B_\delta(p_*)$, the Newton sequence

$$p_{k+1} = \exp_{p_k} \left(-\nabla X(p_k)^{-1} X(p_k)\right), \quad k = 0, 1, \ldots$$

is well defined, contained in $B_\delta(p_*)$, and it converges superlinearly to $p_*$. 

Proof Let $\tilde{\delta}$ be given by Lemma 3.2. From Lemma 3.3, we can shrink $\tilde{\delta}$, if necessary, to conclude that

$$d(N_X(p), p_*) < \frac{1}{2} d(p, p_*), \quad \forall \ p \in B_{\tilde{\delta}}(p_*).$$

Thus, $N_X(p) \in B_{\tilde{\delta}}(p_*)$, for all $p \in B_{\tilde{\delta}}(p_*)$. Note that (8) and (10) imply that $\{p_k\}$ satisfies

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

which is indeed an equivalent definition of this sequence. Since $N_X(p) \in B_{\tilde{\delta}}(p_*)$ for all $p \in B_{\tilde{\delta}}(p_*)$, it follows from (12) and Lemma 3.2 item (i) that, for all $p_0 \in B_{\tilde{\delta}}(p_*)$, the Newton sequence $\{p_k\}$ is well defined and contained in $B_{\tilde{\delta}}(p_*)$. Moreover, using (11) and (12), we obtain

$$d(p_{k+1}, p_*) < \frac{1}{2} d(p_k, p_*), \quad k = 0, 1, \ldots$$

The latter inequality implies that $\{p_k\}$ converges to $p_*$. Thus, by combining (12) with Lemma 3.3, we conclude that

$$\lim_{k \to +\infty} \frac{d(p_{k+1}, p_*)}{d(p_k, p_*)} = 0.$$ 

Therefore, $\{p_k\}$ converges superlinearly to $p_*$, and the proof is completed.

Next, we present an application of Theorem 3.1 for finding the critical points of a twice-differentiable function defined on a Riemannian manifold.

Corollary 3.2 Let $\Omega \subset \mathbb{M}$ be an open set, $f : \Omega \to \mathbb{R}$ a twice-differentiable function and $p_* \in \Omega$. Suppose that $p_*$ is a critical point of $f$, Hess $f$ is continuous at $p_*$ and Hess $f(p_*)$ is nonsingular. Then, there exists $\tilde{\delta} > 0$ such that, for all $p_0 \in B_{\tilde{\delta}}(p_*)$, the Newton sequence

$$p_{k+1} = \exp_{p_k} \left(-\text{Hess } f(p_k)^{-1} \text{ grad } f(p_k)\right), \quad k = 0, 1, \ldots,$$

is well defined, contained in $B_{\tilde{\delta}}(p_*)$, and it converges superlinearly to $p_*$. 

Proof By letting $X = \text{ grad } f$, the result follows by applying Theorem 3.1.
Before concluding this section, we present an important property of the parallel transport. When \( M \) is a complete and finite-dimensional Riemannian manifold, Lemma 3.1 enable us to obtain the continuity of the vector field \( B_\delta(\bar{p}) \ni p \mapsto P_\delta Z(p) \), where \( \delta \leq \delta_\bar{p} \); see Corollary 3.1. On the other hand, when \( M \) is a Hadamard manifold, i.e., a complete simply connected Riemannian manifold of nonpositive sectional curvature, we know that \( i_\bar{p} = +\infty \), \( \phi := \exp_{\bar{p}}^{-1} : M \to T_{\bar{p}}M \) is a diffeomorphism, and \( (M, \phi) \) is a global chart for \( M \). Therefore, by following the same idea as that of the proof of Corollary 3.1, we can prove the following generalization of [9, Lemma 2.4, item (iv)]:

**Corollary 3.3** Let \( M \) be a Hadamard manifold. If \( Z : M \to TM \) is a continuous vector field, then the mapping \( M \ni (p, q) \mapsto P_{p,q} Z(p) \) is continuous.

**Proof** First, from [9, Lemma 2.4, item (iv)], we have \( \lim_{q \to \bar{q}} P_{\bar{p}q} Z(\bar{p}) = P_{\bar{p}\bar{q}} Z(\bar{p}) \). Then, to prove the result, it is sufficient to show that the following equality holds:

\[
\lim_{(p, q) \to (\bar{p}, \bar{q})} \left\| P_{pq} Z(p) - P_{\bar{p}q} Z(\bar{p}) \right\| = 0. \quad (13)
\]

Indeed, after some algebraic manipulations, and considering that \( P_{\bar{p}q} = P_{\bar{q}q} P_{\bar{p}\bar{q}} \), \( P_{\bar{q}q} P_{\bar{q}q} = I_{\bar{q}} \), \( P_{\bar{p}q} P_{\bar{p}p} = I_{\bar{q}} \), and the parallel transport is a isometry, we have

\[
\left\| P_{pq} Z(p) - P_{\bar{p}q} Z(\bar{p}) \right\| = \left\| Z(p) - P_{\bar{p}p} Z(\bar{p}) \right\|.
\]

Since [9, Lemma 2.4, item (iv)] implies that \( \lim_{p \to \bar{p}} \left\| Z(p) - P_{\bar{p}p} Z(\bar{p}) \right\| = 0 \), from the last equality, (13) holds and the proof is completed. \( \square \)

### 4 Examples

In this section, we present two concrete examples. We show that the main result can be applied to find the singularity of differentiable vector fields defined on a sphere and on a cone of symmetric positive definite matrices.

**Example 4.1** Let \( (\cdot, \cdot) \) be the Euclidian inner product, with the corresponding norm denoted by \( \| \cdot \| \). The \( n \)-dimensional Euclidean sphere and its tangent hyperplane at a point \( p \) are, respectively, denoted by

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

\[
T_p S^n := \left\{ v \in \mathbb{R}^{n+1} : (p, v) = 0 \right\}.
\]

Let \( M := (S^n, (\cdot, \cdot)) \). For \( p \in \mathbb{M} \), the unique segment of the geodesic connecting \( p \) to \(-p\), starting at \( p \) with velocity \( v \) at \( p \), is given by \( \gamma(t) = \cos(t)p + \sin(t)v \) for all \( t \in [0, \pi] \). Thus, for \( p \in \mathbb{M} \), \( \exp_p : T_p \mathbb{M} \to \mathbb{M} \) is given by

\[
\exp_p v = \begin{cases} 
\cos(\|v\|) p + \sin(\|v\|) \frac{v}{\|v\|}, & v \neq 0 \in T_p \mathbb{M} / \{0\}; \\
0, & v = 0.
\end{cases}
\]
Let $X$ and $Y$ be vector fields in $\mathbb{M}$. Then, using [30], we can prove that the Levi-Civita connection of $\mathbb{M}$ is given by

$$\nabla_Y X(p) = \left[ I - pp^T \right] X'(p) Y,$$

(15)

where $X'(p)$ is the usual derivative of $X$ at $p$ and $I$ denotes the $(n + 1) \times (n + 1)$ identity matrix. Hence, for $f : \mathbb{M} \to \mathbb{R}$ a twice-differentiable function, by using the Euclidian inner product and (15), the gradient and Hessian of $f$ at $p$ are, respectively, given by

$$\text{grad } f(p) = f'(p) - \langle f'(p), p \rangle p,$$
$$\text{Hess } f(p) = \left[ I - pp^T \right] \left[ f''(p) - \langle f'(p), p \rangle I \right],$$

(16)

where $f'(p)$ and $f''(p)$ are the Euclidean gradient and Hessian of $f$ at $p$, respectively. Hence, it follows from (14) and (15) that Newton’s method for finding $p \in \mathbb{M}$ such that $X(p) = 0$ where $X : \mathbb{M} \to T\mathbb{M}$ is a differentiable vector field is given by

$$p_{k+1} = \cos (\|v_k\|) p_k + \sin (\|v_k\|) \frac{v_k}{\|v_k\|}, \quad k = 0, 1, \ldots,$$

where $v_k \neq 0$ is the unique solution of the linear equation

$$\left[ I - pp_k^T \right] X'(p_k) v_k = -X(p_k).$$

Moreover, for $f : \mathbb{M} \to \mathbb{R}$ a twice-differentiable function, by using (16), Newton’s method for finding $p \in \mathbb{M}$ such that $\text{grad } f(p) = 0$ is given by

$$p_{k+1} = \cos (\|v_k\|) p_k + \sin (\|v_k\|) \frac{v_k}{\|v_k\|}, \quad k = 0, 1, \ldots,$$

(17)

where $v_k \neq 0$ is the unique solution of the linear equation

$$\left[ I - pp_k^T \right] \left[ f''(p_k) - \langle f'(p_k), p_k \rangle I \right] v_k = -f'(p_k) + \langle f'(p_k), p_k \rangle p_k.$$  

(18)

In particular, letting $A$ be a $n \times n$ positive definite symmetric matrix and $f : \mathbb{S}^n \to \mathbb{R}$ be the Rayleigh quotient function, i.e., $f(p) = p^T Ap$. Then, after some calculations, (18) becomes

$$\left[ \left( I - pp_k^T \right) B_k \right] v_k = -B_k p_k,$$

(19)

where $B_k = 2 \left[ A - p_k^T A p_k I \right]$. Therefore, from Corollary 3.2, we conclude that there exists a neighborhood of a global minimizer $p^*$ of the Rayleigh quotient function, such that the sequence defined by (17) and (19) converges superlinearly to $p^*$. We remark that the minimum value of the Rayleigh quotient function is the minimum eigenvalue of $A$. 

Springer
Example 4.2 Let $\mathbb{P}^n$ be the set of symmetric matrices, and let $\mathbb{P}^n_{++}$ be the cone of symmetric positive definite matrices of order $n \times n$. Following Rothaus [31], let $\mathbb{M} := (\mathbb{P}^n_{++}, (,))$ be the Riemannian manifold endowed with the Riemannian metric induced by the Euclidean Hessian of $\psi(A) = -\ln \det A$, namely

$$
\langle U, V \rangle = \text{tr}(V \psi''(A) U) = \text{tr}(V A^{-1} U A^{-1}), \quad A \in \mathbb{M}, \quad U, V \in T_A \mathbb{M}, \tag{20}
$$

where $\text{tr}(A)$ denotes the trace of matrix $A \in \mathbb{P}^n$ and $T_A \mathbb{M} \approx \mathbb{P}^n$. In this case, the unique geodesic segment connecting any $A, B \in \mathbb{M}$ is given by $\gamma(t) = A^{1/2} (A^{-1/2} B A^{-1/2})^t A^{1/2}$, for all $t \in [0, 1]$; see [32]. More precisely, $\mathbb{M}$ is a Hadamard manifold; see, for example, [33, Theorem 1.2, p. 325]. From this equality, we have $\gamma'(0) = A^{1/2} \ln (A^{-1/2} B A^{-1/2}) A^{1/2}$. Thus, for $A \in \mathbb{M}$, $\exp_A : T_A \mathbb{M} \to \mathbb{M}$ is given by

$$
\exp_A V = A^{1/2} e^{(A^{-1/2} V A^{-1/2})} A^{1/2}, \quad \tag{21}
$$

Let $X$ and $Y$ be vector fields in $\mathbb{M}$. Then, by using [28, Theorem 1.2, page 28], we can prove that the Levi-Civita connection of $\mathbb{M}$ is given by

$$
\nabla_Y X (A) = X' Y - \frac{1}{2} \left[ Y A^{-1} X + X A^{-1} Y \right], \tag{22}
$$

where $X'$ denotes the Euclidean derivative of $X$ and $A \in \mathbb{M}$. Therefore, from (20) and (22), the gradient and hessian are, respectively, given by:

$$
\text{grad} f(A) = Af’(A) A, \quad \text{Hess} f(A) V = Af''(A) VA + \frac{1}{2} \left[ V f'(A) A + AF'(A) V \right], \quad V \in T_A \mathbb{M}, \tag{23}
$$

where $f'(A)$ and $f''(A)$ are the Euclidean gradient and Hessian of $f$ at $A$, respectively. In this case, by using (21) and (22), Newton’s method for finding $A \in \mathbb{M}$ such that $X(A) = 0$, where $X : \mathbb{M} \to T \mathbb{M}$ is a differentiable vector field, is given by

$$
A_{k+1} = A_k^{1/2} e^{\left( A_k^{-1/2} V_k A_k^{-1/2} \right)} A_k^{1/2}, \quad k = 0, 1, \ldots ,
$$

where $V_k$ is the unique solution of the linear equation

$$
X'(A_k) V_k - \frac{1}{2} \left[ V_k A_k^{-1} X(A_k) + X(A_k) A_k^{-1} V_k \right] = -X(A_k).
$$

Moreover, for $f : \mathbb{M} \to \mathbb{R}$, a twice-differentiable function, by using (23), Newton’s method for finding $A \in \mathbb{M}$ such that $\text{grad} f(A) = 0$ is given by

$$
A_{k+1} = A_k^{1/2} e^{\left( A_k^{-1/2} V_k A_k^{-1/2} \right)} A_k^{1/2}, \quad k = 0, 1, \ldots , \tag{24}
$$
where $V_k$ is the unique solution of the linear equation

$$A_k f''(A_k) V_k A_k + \frac{1}{2} \left[ V_k f'(A_k) A_k + A_k f'(A_k) V_k \right] = -A_k f'(A_k) A_k. \tag{25}$$

In particular, letting $f : \mathbb{P}_{++}^n \rightarrow \mathbb{R}$ be defined by $f(A) = \ln \det A + \text{tr} A^{-1}$, where $\det A$ and $\text{tr} A$ denote the determinant and the trace of the matrix $A$, respectively. Then, the Euclidean gradient and Hessian are given, respectively, by

$$f'(A_k) = A_k^{-1} - A_k^{-2} \quad \text{and} \quad f''(A_k) V_k = -A_k^{-1} V_k A_k^{-1} + A_k^{-1} V_k A_k^{-2} + A_k^{-2} V_k A_k^{-1}.$$  

Therefore, substituting the last equalities into (25), we obtain

$$A_k V_k + V_k A_k = 2(A_k^2 - A_k^3), \tag{26}$$

We remark that $\text{grad } f(A) = A - \hat{I}$, where $\hat{I}$ is the $n \times n$ identity matrix, and the global minimizer of $f$ is $\hat{I}$. Hence, from Corollary 3.2 we conclude that there exists a neighborhood of the $\hat{I}$ such that the sequence defined by (24) and (26) converges superlinearly to $\hat{I}$.

5 Conclusions

In this paper, under nonsingularity of the covariant derivative of the vector field at its zero and without any additional conditions on this derivative, we established the superlinear local convergence of Newton’s method on a finite-dimensional Riemannian manifold. It is worth noting that we assumed that the Riemannian manifold is finite-dimensional in order to establish this result, at least two times, namely in Lemma 3.1 and Lemma 3.2. On the other hand, we know that Newton’s method converges superlinearly on infinite-dimensional Banach spaces under nonsingularity of the derivative of the operator at its zero. Since many important problems arise in infinite-dimensional Riemannian manifold as problems of finding singularities of vector fields, e.g., see [34,35], it would be interesting to extend the above results to such manifolds.

Acknowledgements The work was supported by FAPEG, UESB, and CNPq Grants 305158/2014-7 and 408151/2016-1.

References

1. Smith, S.T.: Optimization techniques on Riemannian manifolds, Hamiltonian and gradient flows, algorithms and control. In: Bloch, A. (ed.) Fields Institute Communication, vol. 3, pp. 113–136. American Mathematical Society, Providence, RI (1994)
2. Absil, P.A., Mahony, R., Sepulchre, R.: Optimization Algorithms on Matrix Manifolds. Princeton University Press, Princeton, NJ (2008)
3. Luenberger, D.G.: The gradient projection method along geodesics. Manag. Sci. 18, 620–631 (1972)
4. Gabay, D.: Minimizing a differentiable function over a differential manifold. J. Optim. Theory Appl. 37(2), 177–219 (1982)
5. Rapcsák, T.: Minimum problems on differentiable manifolds. Optimization 20(1), 3–13 (1989)
6. Udrişte, C.: Convex Functions and Optimization Methods on Riemannian Manifolds, Mathematics and its Applications, vol. 297. Kluwer Academic Publishers, Dordrecht (1994)
7. Absil, P.A., Amodio, L., Meyer, G.: Two Newton methods on the manifold of fixed-rank matrices endowed with Riemannian quotient geometries. Comput. Stat. 29(3–4), 569–590 (2014)
8. Huang, W., Gallivan, K.A., Absil, P.A.: A Broyden class of quasi-Newton methods for Riemannian optimization. SIAM J. Optim. 25(3), 1660–1685 (2015)
9. Li, C., López, G., Martín-Márquez, V.: Monotone vector fields and the proximal point algorithm on Hadamard manifolds. J. Lond. Math. Soc. 79(3), 663–683 (2009)
10. Li, C., Wang, J.: Newton’s method for sections on Riemannian manifolds: generalized covariant $\alpha$-theory. J. Complex. 24(3), 423–451 (2008)
11. Hu, Y., Li, C., Yang, X.: On convergence rates of linearized proximal algorithms for convex composite optimization with applications. SIAM J. Optim. 26(2), 1207–1235 (2016)
12. Ring, W., Wirth, B.: Optimization methods on Riemannian manifolds and their application to shape space. SIAM J. Optim. 22(2), 596–627 (2012)
13. Manton, J.H.: A framework for generalising the Newton method and other iterative methods from Euclidean space to manifolds. Numer. Math. 129(1), 91–125 (2015)
14. Bittencourt, T., Ferreira, O.: Kantorovich’s theorem on Newton’s method under majorant condition in Riemannian manifolds. J. Glob. Optim. (2016). doi:10.1007/s10898-016-0472
15. Adler, R.L., Dedieu, J.P., Margulies, J.Y., Martens, M., Shub, M.: Newton’s method on Riemannian manifolds and a geometric model for the human spine. IMA J. Numer. Anal. 22(3), 359–390 (2002)
16. Nash, J.: The imbedding problem for Riemannian manifolds. Ann. Math. 2(63), 20–63 (1956)
17. Moser, J.: A new technique for the construction of solutions of nonlinear differential equations. Proc. Natl. Acad. Sci. USA 47, 1824–1831 (1961)
18. Li, C., Wang, J.H., Dedieu, J.P.: Smale’s point estimate theory for Newton’s method on Lie groups. J. Complex. 25(2), 128–151 (2009)
19. Argyros, I.K., Hilout, S.: Newton’s method for approximating zeros of vector fields on Riemannian manifolds. J. Appl. Math. Comput. 29(1–2), 417–427 (2009)
20. Schulz, V.H.: A Riemannian view on shape optimization. Found. Comput. Math. 14(3), 483–501 (2014)
21. Wang, J.H., Li, C.: Kantorovich’s theorems for Newton’s method for mappings and optimization problems on Lie groups. IMA J. Numer. Anal. 31(1), 322–347 (2011)
22. Ferreira, O.P., Silva, R.C.M.: Local convergence of Newton’s method under a majorant condition in Riemannian manifolds. IMA J. Numer. Anal. 32(4), 1696–1713 (2012)
23. Ferreira, O.P., Svaiter, B.F.: Kantorovich’s theorem on Newton’s method in Riemannian manifolds. J. Complex. 18(1), 304–329 (2002)
24. 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)
25. Ortega, J.M.: Numerical Analysis. A Second Course. Computer Science and Applied Mathematics. Academic Press, New York (1972)
26. 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)
27. do Carmo, M.P.: Riemannian Geometry. Mathematics: Theory & Applications. Birkhäuser, Boston (1992)
28. Sakai, T.: Riemannian Geometry, Translations of Mathematical Monographs, vol. 149. American Mathematical Society, Providence, RI (1996)
29. Dieudonné, J.: Foundations of Modern Analysis, vol. 10-I. Academic Press, New York (1969). Enlarged and Corrected Printing, Pure and Applied Mathematics
30. Ferreira, O.P., Iusem, A.N., Németh, S.Z.: Concepts and techniques of optimization on the sphere. TOP 22(3), 1148–1170 (2014)
31. Rothaus, O.S.: Domains of positivity. Abh. Math. Sem. Univ. Hambg. 24, 189–235 (1960)
32. Nesterov, Y.E., Todd, M.J.: On the Riemannian geometry defined by self-concordant barriers and interior-point methods. Found. Comput. Math. 2(4), 333–361 (2002)
33. Lang, S.: Fundamentals of Differential Geometry, Graduate Texts in Mathematics, vol. 191. Springer, New York (1999)
34. Milnor, J.: Morse theory. Based on lecture notes by M. Spivak and R. Wells. In: Annals of Mathematics Studies, No. 51. Princeton University Press, Princeton (1963)
35. Kristály, A., Rădulescu, V.D., Varga, C.G.: Variational Principles in Mathematical Physics, Geometry, and Economics, Encyclopedia of Mathematics and its Applications, vol. 136. Cambridge University Press, Cambridge (2010)