OPTIMIZATION METHODS ON RIEMANNIAN MANIFOLDS AND THEIR APPLICATION TO SHAPE SPACE

WOLFGANG RING* AND BENEDIKT WIRTH†

Abstract. We extend the scope of analysis for linesearch optimization algorithms on (possibly infinite-dimensional) Riemannian manifolds to the convergence analysis of the BFGS quasi-Newton scheme and the Fletcher–Reeves conjugate gradient iteration. Numerical implementations for exemplary problems in shape spaces show the practical applicability of these methods.

Key words. Riemannian optimization, BFGS quasi-Newton, Fletcher–Reeves conjugate gradient, shape space

AMS subject classifications. 90C48, 65K05, 49M37, 58E50

1. Introduction.

There are a number of problems that can be expressed as a minimization of a function $f : \mathcal{M} \to \mathbb{R}$ over a smooth Riemannian manifold $\mathcal{M}$. Applications range from linear algebra (e. g. principal component analysis or singular value decomposition, [1, 11]) to the analysis of shape spaces (e. g. computation of shape geodesics, [24]), see also the references in [3].

If the manifold $\mathcal{M}$ can be embedded in a higher-dimensional space or if it is defined via equality constraints, then there is the option to employ the very advanced tools of constrained optimization (see [16] for an introduction). Often, however, such an embedding is not at hand, and one has to resort to optimization methods designed for Riemannian manifolds. Even if an embedding is known, one might hope that a Riemannian optimization method performs more efficiently since it exploits the underlying geometric structure of the manifold. For this purpose, various methods have been devised, from simple gradient descent on manifolds [25] to sophisticated trust region methods [5]. The aim of this article is to extend the scope of analysis for these methods, concentrating on linesearch methods. In particular, we will consider the convergence of BFGS quasi-Newton methods and Fletcher–Reeves nonlinear conjugate gradient iterations on (possibly infinite-dimensional) manifolds, thereby filling a gap in the existing analysis. Furthermore, we apply the proposed methods to exemplary problems, showing their applicability to manifolds of practical interest such as shape spaces.

Early attempts to adapt standard optimization methods to problems on manifolds were presented by Gabay [9] who introduced a steepest descent, a Newton, and a quasi-Newton algorithm, also stating their global and local convergence properties (however, without giving details of the analysis for the quasi-Newton case). Udrişte [23] also stated a steepest descent and a Newton algorithm on Riemannian manifolds and proved (linear) convergence of the former under the assumption of exact linesearch. Fairly recently, Yang took up these methods and analysed convergence and convergence rate of steepest descent and Newton's method for Armijo step-size control [25].

*Institute for Mathematics and Scientific Computing, University of Graz, Heinrichstraße 36, A-8010 Graz, Austria, Wolfgang.Ring@uni-graz.at

†Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012, USA, Benedikt.Wirth@cims.nyu.edu

Benedikt Wirth was funded by the Austrian Science Fund (FWF) via the SFB "Mathematical Optimization and Applications in Biomedical Sciences".

In comparison to standard linesearch methods in vector spaces, the above approaches all substitute the linear step in the search direction by a step along a geodesic. However, geodesics may be difficult to obtain. In alternative approaches, the geodesics are thus often replaced by more general paths, based on so-called retractions (a retraction R_x is a mapping from the tangent space $T_x\mathcal{M}$ to the manifold $\mathcal{M}$ at x onto $\mathcal{M}$). For example, Hüper and Trumpf find quadratic convergence for Newton's method without step-size control [11], even if the Hessian is computed for a different retraction than the one defining the path along which the Newton step is taken. In the sequel, more advanced trust region Newton methods have been developed and analysed in a series of papers [6, 1, 5]. The analysis requires some type of uniform Lipschitz continuity for the gradient and the Hessian of the concatenations $f \circ R_x$, which we will also make use of. The global and local convergence analysis of gradient descent, Newton's method, and trust region methods on manifolds with general retractions is summarized in [2].

In [2], the authors also present a way to implement a quasi-Newton as well as a nonlinear conjugate gradient iteration, however, without analysis. Riemannian BFGS quasi-Newton methods, which generalize Gabay's original approach, have been devised in [18]. Like the above schemes, they do not rely on geodesics but allow more general retractions. Furthermore, the vector transport between the tangent spaces $T_{x_k}\mathcal{M}$ and $T_{x_{k+1}}\mathcal{M}$ at two subsequent iterates (which is needed for the BFGS update of the Hessian approximation) is no longer restricted to parallel transport. A similar approach is taken in [7], where a specific, non-parallel vector transport is considered for a linear algebra application. A corresponding global convergence analysis of a quasi-Newton scheme has been performed by Ji [12]. However, to obtain a superlinear convergence rate, specific conditions on the compatibility between the vector transports and the retractions are required (cf. section 3.1) which are not imposed (at least explicitly) in the above work.

In this paper, we pursue several objectives. First, we extend the convergence analysis of standard Riemannian optimization methods (such as steepest descent and Newton's method) to the case of optimization on infinite-dimensional manifolds. Second, we analyze the convergence rate of the Riemannian BFGS quasi-Newton method as well as the convergence of a Riemannian Fletcher–Reeves conjugate gradient iteration (as two representative higher order gradient-based optimization methods), which to the authors' knowledge has not been attempted before (neither in the finite- nor the infinite-dimensional case). The analysis is performed in the unifying framework of step-size controlled linesearch methods, which allows for rather streamlined proofs. Finally, we demonstrate the feasibility of these Riemannian optimization methods by applying them to problems in state-of-the-art shape spaces.

The outline of this article is as follows. Section 2 summarizes the required basic notions. Section 3 then introduces linesearch optimization methods on Riemannian manifolds, following the standard procedure for finite-dimensional vector spaces [16, 17] and giving the analysis of basic steepest descent and Newton's method as a prerequisite for the ensuing analysis of the BFGS scheme in section 3.1 and the nonlinear CG iteration in section 3.2. Finally, numerical examples on shape spaces are provided in section 4.

2. Notations. Let $\mathcal{M}$ denote a geodesically complete (finite- or infinite-dimensional) Riemannian manifold. In particular, we assume $\mathcal{M}$ to be locally homeomorphic to some separable Hilbert space $\mathcal{H}$ [14, Sec. 1.1], that is, for each $x \in \mathcal{M}$ there is some neighborhood $x \in U_x \subset \mathcal{M}$ and a homeomorphism ϕ_x from U_x into some open subset

of $\mathcal{H}$. Let $C^\infty(\mathcal{M})$ denote the vector space of smooth real functions on $\mathcal{M}$ (in the sense that for any $f \in C^\infty(\mathcal{M})$ and any chart (U_x, ϕ_x), $f \circ \phi_x^{-1}$ is smooth). For any smooth curve $\gamma : [0,1] \to \mathcal{M}$ we define the tangent vector to γ at $t_0 \in (0,1)$ as the linear operator $\dot\gamma(t_0) : C^\infty(\mathcal{M}) \to \mathbb{R}$ such that

$$\dot\gamma(t_0) f = \left(\frac{\mathrm{d}}{\mathrm{d}t} f \circ \gamma\right)(t_0) \quad \forall f \in C^\infty(\mathcal{M})\,.$$

The tangent space $T_x\mathcal{M}$ to $\mathcal{M}$ in $x \in \mathcal{M}$ is the set of tangent vectors $\dot\gamma(t_0)$ to all smooth curves $\gamma : [0,1] \to \mathcal{M}$ with $\gamma(t_0) = x$. It is a vector space, and it is equipped with an inner product $g_x(\cdot,\cdot) : T_x\mathcal{M} \times T_x\mathcal{M} \to \mathbb{R}$, the so-called Riemannian metric, which smoothly depends on x. The corresponding induced norm will be denoted $\|\cdot\|_x$, and we will employ the same notation for the associated dual norm.

Let $\mathcal{V}(\mathcal{M})$ denote the set of smooth tangent vector fields on $\mathcal{M}$, then we define a connection $\nabla : \mathcal{V}(\mathcal{M}) \times \mathcal{V}(\mathcal{M}) \to \mathcal{V}(\mathcal{M})$ as a map such that

$$\begin{aligned}
\nabla_{fX+gY} Z &= f\nabla_X Z + g\nabla_Y Z && \forall X, Y, Z \in \mathcal{V}(\mathcal{M}) \text{ and } f, g \in C^\infty(\mathcal{M})\,,\\
\nabla_Z(aX + bY) &= a\nabla_Z X + b\nabla_Z Y && \forall X, Y, Z \in \mathcal{V}(\mathcal{M}) \text{ and } a, b \in \mathbb{R}\,,\\
\nabla_Z(fX) &= (Zf)X + f\nabla_Z X && \forall X, Z \in \mathcal{V}(\mathcal{M}) \text{ and } f \in C^\infty(\mathcal{M})\,.
\end{aligned}$$

In particular, we will consider the Levi-Civita connection, which additionally satisfies the properties of absence of torsion and preservation of the metric,

$$\begin{aligned}
[X, Y] &= \nabla_X Y - \nabla_Y X && \forall X, Y \in \mathcal{V}(\mathcal{M})\,,\\
Xg(Y, Z) &= g(\nabla_X Y, Z) + g(Y, \nabla_X Z) && \forall X, Y, Z \in \mathcal{V}(\mathcal{M})\,,
\end{aligned}$$

where $[X, Y] = XY - YX$ denotes the Lie bracket and $g(X, Y) : \mathcal{M} \ni x \mapsto g_x(X, Y)$.

Any connection is paired with a notion of parallel transport. Given a smooth curve $\gamma : [0,1] \to \mathcal{M}$, the initial value problem

$$\nabla_{\dot\gamma(t)} v = 0,\ v(0) = v_0$$

defines a way of transporting a vector $v_0 \in T_{\gamma(0)}\mathcal{M}$ to a vector $v(t) \in T_{\gamma(t)}\mathcal{M}$. For the Levi-Civita connection considered here this implies constancy of the inner product $g_{\gamma(t)}(v(t), w(t))$ for any two vectors $v(t)$, $w(t)$, transported parallel along γ. For $x = \gamma(0)$, $y = \gamma(t)$, we will denote the parallel transport of $v(0) \in T_x\mathcal{M}$ to $v(t) \in T_y\mathcal{M}$ by $v(t) = T^{P_\gamma}_{x,y} v(0)$ (if there is more than one t with $y = \gamma(t)$, the correct interpretation will become clear from the context). As detailed further below, $T^{P_\gamma}_{x,y}$ can be interpreted as the derivative of a specific mapping $P_\gamma : T_x\mathcal{M} \to \mathcal{M}$.

We assume that any two points $x, y \in \mathcal{M}$ can be connected by a shortest curve $\gamma : [0,1] \to \mathcal{M}$ with $x = \gamma(0)$, $y = \gamma(1)$, where the curve length is measured as

$$L[\gamma] = \int_0^1 \sqrt{g_{\gamma(t)}(\dot\gamma(t), \dot\gamma(t))}\,\mathrm{d}t\,.$$

Such a curve is denoted a geodesic, and the length of geodesics induces a metric distance on $\mathcal{M}$,

$$\operatorname{dist}(x, y) = \min_{\substack{\gamma:[0,1]\to\mathcal{M}\\ \gamma(0)=x, \gamma(1)=y}} L[\gamma]\,.$$

If the geodesic γ connecting x and y is unique, then $\dot\gamma(0) \in T_x\mathcal{M}$ is denoted the logarithmic map of y with respect to x, $\log_x y = \dot\gamma(0)$. It satisfies $\operatorname{dist}(x, y) = \|\log_x y\|_x$.

Geodesics can be computed via the zero acceleration condition $\nabla_{\dot\gamma(t)}\dot\gamma(t) = 0$. The exponential map $\exp_x : T_x\mathcal{M} \to \mathcal{M}$, $v \mapsto \exp_x v$, is defined as $\exp_x v = \gamma(1)$, where γ solves the above ordinary differential equation with initial conditions $\gamma(0) = x$, $\dot\gamma(0) = v$. It is a diffeomorphism of a neighborhood of $0 \in T_x\mathcal{M}$ into a neighborhood of $x \in \mathcal{M}$ with inverse $\log_x$. We shall denote the geodesic between x and y by $\gamma[x; y]$, assuming that it is unique.

The exponential map $\exp_x$ obviously provides a (local) parametrization of $\mathcal{M}$ via $T_x\mathcal{M}$. We will also consider more general parameterizations, so-called retractions. Given $x \in \mathcal{M}$, a retraction is a smooth mapping $R_x : T_x\mathcal{M} \to \mathcal{M}$ with $R_x(0) = x$ and $\mathrm{D}R_x(0) = \mathrm{id}_{T_x\mathcal{M}}$, where $\mathrm{D}R_x$ denotes the derivative of R_x. The inverse function theorem then implies that R_x is a local homeomorphism. Besides $\exp_x$, there are various possibilities to define retractions. For example, consider a geodesic $\gamma : \mathbb{R} \to \mathcal{M}$ with $\gamma(0) = x$, parameterized by arc length, and define the retraction

$$P_\gamma : T_x\mathcal{M} \to \mathcal{M}\,, \quad v \mapsto \exp_{p_\gamma(v)}\left(T^{P_\gamma}_{x,p_\gamma(v)}\left[v - \pi_\gamma(v)\right]\right)\,,$$

where $p_\gamma(v) = \exp_x(\pi_\gamma(v))$ and π_γ denotes the orthogonal projection onto $\mathrm{span}\{\dot\gamma(0)\}$. This retraction corresponds to running along the geodesic γ according to the component of v parallel to $\dot\gamma(0)$ and then following a new geodesic into the direction of the parallel transported component of v orthogonal to $\dot\gamma(0)$.

We will later have to consider the transport of a vector from one tangent space $T_x\mathcal{M}$ into another one $T_y\mathcal{M}$, that is, we will consider isomorphisms $T_{x,y} : T_x\mathcal{M} \to T_y\mathcal{M}$. We are particularly interested in operators $T^{R_x}_{x,y}$ which represent the derivative $\mathrm{D}R_x(v)$ of a retraction R_x at $v \in T_x\mathcal{M}$ with $R_x(v) = y$ (where in case of multiple v with $R_x(v) = y$ it will be clear from the context which v is meant). In that sense, the parallel transport $T^{P_\gamma}_{x,y}$ along a geodesic γ connecting x and y belongs to the retraction P_γ. Another possible vector transport is defined by the variation of the exponential map, evaluated at the representative of y in $T_x\mathcal{M}$,

$$T^{\exp_x}_{x,y} = \mathrm{D}\exp_x(\log_x y)\,,$$

which maps each $v_0 \in T_x\mathcal{M}$ onto the tangent vector $\dot\gamma(0)$ to the curve $\gamma : t \mapsto \exp_x(\log_x y + tv_0)$. Also the adjoints $(T^{P_\gamma}_{y,x})^*, (T^{\exp_y}_{y,x})^*$ (defined by $g_y(v, T^*_{y,x}w) = g_x(T_{y,x}v, w)\ \forall v \in T_y\mathcal{M}, w \in T_x\mathcal{M}$) or inverses $(T^{P_\gamma}_{y,x})^{-1}, (T^{\exp_y}_{y,x})^{-1}$ can be considered for vector transport $T_{x,y}$. Note that $T^{P_\gamma}_{x,y}$ is an isometry with $T^{P_\gamma}_{x,y} = (T^{P_{\tilde\gamma}}_{y,x})^* = (T^{P_{\tilde\gamma}}_{y,x})^{-1}$, where γ is the geodesic connecting x with y and $\tilde\gamma(\cdot) = \gamma(-\cdot)$. Furthermore, $\log_x y$ is transported onto the same vector $T_{x,y}\log_x y = \dot\gamma(1)$ by $T^{P_\gamma}_{x,y}$, $T^{\exp_x}_{x,y}$ and their adjoints and inverses.

Given a smooth function $f : \mathcal{M} \to \mathbb{R}$, we define its (Fréchet-)derivative $\mathrm{D}f(x)$ at a point $x \in \mathcal{M}$ as an element of the dual space to $T_x\mathcal{M}$ via $\mathrm{D}f(x)v = vf$. The Riesz representation theorem then implies the existence of a $\nabla f(x) \in T_x\mathcal{M}$ such that $\mathrm{D}f(x)v = g_x(\nabla f(x), v)$ for all $v \in T_x\mathcal{M}$, which we denote the gradient of f at x. On $T_x\mathcal{M}$, define

$$f_{R_x} = f \circ R_x$$

for a retraction R_x. Due to $\mathrm{D}R_x(0) = \mathrm{id}$ we have $\mathrm{D}f_{R_x}(0) = \mathrm{D}f(x)$. Furthermore, $f_{R_x} = f_{R_y} \circ R_y^{-1} \circ R_x$ where R_y^{-1} exists and thus for $y = R_x(v)$,

$$\mathrm{D}f_{R_x}(v) = \mathrm{D}f_{R_y}(0)\mathrm{D}R_x(v) = \mathrm{D}f(y)T^{R_x}_{x,y}, \quad \nabla f_{R_x}(v) = (T^{R_x}_{x,y})^*\nabla f(y)\,.$$

We define the Hessian $\mathrm{D}^2 f(x)$ of a smooth f at x as the symmetric bilinear form $\mathrm{D}^2 f(x) : T_x\mathcal{M} \times T_x\mathcal{M} \to \mathbb{R}$, $(v, w) \mapsto g_x(\nabla_v \nabla f(x), w)$, which is equivalent to $\mathrm{D}^2 f(x) = \mathrm{D}^2 f_{\exp_x}(0)$. By $\nabla^2 f(x) : T_x\mathcal{M} \to T_x\mathcal{M}$ we denote the linear operator mapping $v \in T_x\mathcal{M}$ onto the Riesz representation of $\mathrm{D}^2 f(x)(v, \cdot)$. Note that if $\mathcal{M}$ is embedded into a vector space $\mathcal{X}$ and $f = \tilde{f}|_{\mathcal{M}}$ for a smooth $\tilde{f} : \mathcal{X} \to \mathbb{R}$, we usually have $\mathrm{D}^2 f(x) \neq \mathrm{D}^2 \tilde{f}(x)|_{T_x\mathcal{M} \times T_x\mathcal{M}}$ (which is easily seen from the example $\mathcal{X} = \mathbb{R}^n$, $\tilde{f}(x) = \|x\|^2_{\mathcal{X}}$, $\mathcal{M} = \{x \in \mathcal{X} : \|x\|_{\mathcal{X}} = 1\}$). For a smooth retraction R_x we do not necessarily have $\mathrm{D}^2 f_{R_x}(0) = \mathrm{D}^2 f(x)$, however, this holds at stationary points of f [1, 2].

Let $\mathcal{T}^n(T_x\mathcal{M})$ denote the vector space of n-linear forms $T : (T_x\mathcal{M})^n \to \mathbb{R}$ together with the norm $\|T\|_x = \sup_{v_1,\ldots,v_n \in T_x\mathcal{M}} \frac{T(v_1,\ldots,v_n)}{\|v_1\|_x \ldots \|v_n\|_x}$. We call a function $g : \mathcal{M} \to \bigcup_{x\in\mathcal{M}} \mathcal{T}^n(T_x\mathcal{M})$, $x \mapsto g(x) \in \mathcal{T}^n(T_x\mathcal{M})$, (Lipschitz) continuous at $\hat{x} \in \mathcal{M}$ if $x \mapsto g(x) \circ (T^{P_{\gamma[\hat{x};x]}}_{\hat{x},x})^n \in \mathcal{T}^n(T_{\hat{x}}\mathcal{M})$ is (Lipschitz) continuous at $\hat{x}$ (with Lipschitz constant $\limsup_{x\to\hat{x}} \left\| g(x) \circ (T^{P_{\gamma[\hat{x};x]}}_{\hat{x},x})^n - g(\hat{x}) \right\|_{\hat{x}} / \mathrm{dist}(x, \hat{x})$). Here, $\gamma[\hat{x}; x]$ denotes the shortest geodesic, which is unique in the neighborhood of $\hat{x}$. (Lipschitz) continuity on $U \subset \mathcal{M}$ means (Lipschitz) continuity at every $x \in U$ (with uniform Lipschitz constant). A function $f : \mathcal{M} \to \mathbb{R}$ is called n times (Lipschitz) continuously differentiable, if $\mathrm{D}^l f : \mathcal{M} \to \bigcup_{x\in\mathcal{M}} \mathcal{T}^l(T_x\mathcal{M})$ is (Lipschitz) continuous for $0 \leq l \leq n$.

3. Iterative minimization via geodesic linesearch methods. In classical optimization on vector spaces, linesearch methods are widely used. They are based on updating the iterate by choosing a search direction and then adding a multiple of this direction to the old iterate. Adding a multiple of the search direction obviously requires the structure of a vector space and is not possible on general manifolds. The natural extension to manifolds is to follow the search direction along a path. We will consider iterative algorithms of the following generic form.

ALGORITHM 3.1 (Linesearch minimization on manifolds).

Input: $f : \mathcal{M} \to \mathbb{R}$, $x_0 \in \mathcal{M}$, $k = 0$
  repeat
    choose a descent direction $p_k \in T_{x_k}\mathcal{M}$
    choose a retraction $R_{x_k} : T_{x_k}\mathcal{M} \to \mathcal{M}$
    choose a step length $\alpha_k \in \mathbb{R}$
    set $x_{k+1} = R_{x_k}(\alpha_k p_k)$
    $k \leftarrow k + 1$
  until x_{k+1} *sufficiently minimizes* f

Here, a descent direction denotes a direction p_k with $\mathrm{D}f(x_k)p_k < 0$. This property ensures that the objective function f indeed decreases along the search direction.

For the choice of the step length, various approaches are possible. In general, the chosen α_k has to fulfill a certain quality requirement. We will here concentrate on the so-called Wolfe conditions, that is, for a given descent direction $p \in T_x\mathcal{M}$, the chosen step length α has to satisfy

$$f(R_x(\alpha p)) \leq f(x) + c_1 \alpha \mathrm{D}f(x)p\,, \tag{3.1a}$$

$$\mathrm{D}f(R_x(\alpha p))T^{R_x}_{x,R_x(\alpha p)}p \geq c_2 \mathrm{D}f(x)p\,, \tag{3.1b}$$

where $0 < c_1 < c_2 < 1$. Note that both conditions can be rewritten as $f_{R_x}(\alpha p) \leq f_{R_x}(0) + c_1\alpha \mathrm{D}f_{R_x}(0)p$ and $\mathrm{D}f_{R_x}(\alpha p)p \geq c_2 \mathrm{D}f_{R_x}(0)p$, the classical Wolfe conditions

for minimization of f_{R_x}. If the second condition is replaced by

$$|\mathrm{D}f(R_x(\alpha p))T^{R_x}_{x,R_x(\alpha p)}p| \leq -c_2\mathrm{D}f(x)p\,, \tag{3.2}$$

we obtain the so-called strong Wolfe conditions. Most optimization algorithms also work properly if just the so-called Armijo condition (3.1a) is satisfied. In fact, the stronger Wolfe conditions are only needed for the later analysis of a quasi-Newton scheme. Given a descent direction p, a feasible step length can always be found.

PROPOSITION 3.2 (Feasible step length, e. g. [16, Lem. 3.1]). *Let $x \in \mathcal{M}$, $p \in T_x\mathcal{M}$ be a descent direction and $f_{R_x} : \mathrm{span}\{p\} \to \mathbb{R}$ continuously differentiable. Then there exists $\alpha > 0$ satisfying* (3.1) *and* (3.2).

Proof. Rewriting the Wolfe conditions as $f_{R_x}(\alpha p) \leq f_{R_x}(0) + c_1\alpha\mathrm{D}f_{R_x}(0)p$ and $\mathrm{D}f_{R_x}(\alpha p)p \geq c_2\mathrm{D}f_{R_x}(0)p$ ($|\mathrm{D}f_{R_x}(\alpha p)p| \leq -c_2\mathrm{D}f_{R_x}(0)p$, respectively), the standard argument for Wolfe conditions in vector spaces can be applied. □

Besides the quality of the step length, the convergence of linesearch algorithms naturally depends on the quality of the search direction. Let us introduce the angle θ_k between the search direction p_k and the negative gradient $-\nabla f(x_k)$,

$$\cos\theta_k = \frac{-\mathrm{D}f(x_k)p_k}{\|\mathrm{D}f(x_k)\|_{x_k}\,\|p_k\|_{x_k}}\,.$$

There is a classical link between the convergence of an algorithm and the quality of its search direction.

THEOREM 3.3 (Zoutendijk's theorem). *Given $f : \mathcal{M} \to \mathbb{R}$ bounded below and differentiable, assume the α_k in Algorithm 3.1 to satisfy* (3.1). *If the $f_{R_{x_k}}$ are Lipschitz continuously differentiable on* $\mathrm{span}\{p_k\}$ *with uniform Lipschitz constant L, then*

$$\sum_{k\in\mathbb{N}} \cos^2\theta_k\,\|\mathrm{D}f(x_k)\|^2_{x_k} < \infty\,.$$

Proof. The proof for optimization on vector spaces also applies here: Lipschitz continuity and (3.1b) imply

$$\alpha_k L\,\|p_k\|^2_{x_k} \geq (\mathrm{D}f_{R_{x_k}}(\alpha_k p_k) - \mathrm{D}f_{R_{x_k}}(0))p_k \geq (c_2-1)\mathrm{D}f(x_k)p_k\,,$$

from which we obtain $\alpha_k \geq (c_2-1)\mathrm{D}f(x_k)p_k/(L\,\|p_k\|^2_{x_k})$. Then (3.1a) implies

$$f(x_{k+1}) \leq f(x_k) - c_1\frac{1-c_2}{L}\cos^2\theta_k\,\|\mathrm{D}f(x_k)\|^2_{x_k}$$

so that the result follows from the boundedness of f by summing over all k. □

COROLLARY 3.4 (Convergence of generalized steepest descent). *Let the search direction in Algorithm 3.1 be the solution to $B_k(p_k, v) = -\mathrm{D}f(x_k)v\ \forall v \in T_{x_k}\mathcal{M}$, where the B_k are uniformly coercive and bounded bilinear forms on $T_{x_k}\mathcal{M}$ (the case $B_k(\cdot,\cdot) = g_{x_k}(\cdot,\cdot)$ yields the steepest descent direction). Under the conditions of Theorem 3.3,* $\|\mathrm{D}f(x_k)\|_{x_k} \to 0$.

Proof. Obviously, $\cos\theta_k = \frac{B_k(p_k,p_k)}{\|B_k(p_k,\cdot)\|_{x_k}\|p_k\|_{x_k}}$ is uniformly bounded above zero so that the convergence follows from Zoutendijk's theorem. □

For continuous $\mathrm{D}f$, the previous corollary implies that any limit point x^* of the sequence x_k is a stationary point of f. Hence, on finite-dimensional manifolds, if $\{x \in \mathcal{M} : f(x) \leq f(x_0)\}$ is bounded, x_k can be decomposed into subsequences each of which

converges against a stationary point. On infinite-dimensional manifolds, where only a weak convergence of subsequences can be expected, this is not true in general (which is not surprising given that not even existence of stationary points is granted without stronger conditions on f such as sequential weak lower semi-continuity). Moreover, the limit points may be non-unique. However, in the case of (locally) strictly convex functions we have (local) convergence against the unique minimizer by the following classical estimate.

PROPOSITION 3.5. *Let $U \subset \mathcal{M}$ be a geodesically star-convex neighborhood around $x^* \in \mathcal{M}$ (i. e. for any $x \in U$ the shortest geodesics from x^* to x lie inside U), and define*

$$V(U) = \{v \in \exp_{x^*}^{-1}(U) \,:\, \|v\|_{x^*} \leq \|w\|_{x^*} \,\forall w \in \exp_{x^*}^{-1}(\exp_{x^*}(v))\}\,.$$

Let f be twice differentiable on U with x^* being a stationary point. If $\mathrm{D}^2 f_{\exp_{x^*}}$ is uniformly coercive on $V(U) \subset T_{x^*}\mathcal{M}$, then there exists $m > 0$ such that for all $x \in U$,

$$\operatorname{dist}(x, x^*) \leq \frac{1}{m} \|\mathrm{D}f(x)\|_x \,.$$

Proof. By hypothesis, there is $m > 0$ with $\mathrm{D}^2 f_{\exp_{x^*}}(x)(v,v) \geq m \|v\|_{x^*}^2$ for all $x \in \exp_{x^*}^{-1}(U)$, $v \in T_{x^*}\mathcal{M}$. Thus, for $v = \log_{x^*} x$ we have $\|\mathrm{D}f(x)\|_x \geq \frac{|\mathrm{D}f(x)\log_x x^*|}{\|\log_x x^*\|_x} = \frac{|\mathrm{D}f_{\exp_{x^*}}(v)v|}{\|v\|_{x^*}} = \frac{\left|\int_0^1 \mathrm{D}^2 f_{\exp_{x^*}}(tv)(v,v)\,\mathrm{d}t\right|}{\|v\|_{x^*}} \geq m\|v\|_x = m\operatorname{dist}(x, x^*)$. (Note that coercivity along $\gamma[x; x^*]$ would actually suffice.) □

Hence, if f is smooth with $\mathrm{D}f(x^*) = 0$ and $\mathrm{D}^2 f(x^*) = \mathrm{D}^2 f_{\exp_{x^*}}(0)$ coercive (so that there exists a neighborhood $U \subset \mathcal{M}$ of x^* such that $\mathrm{D}^2 f_{\exp_{x^*}}$ is uniformly coercive on $V(U)$), then by the above proposition and Corollary 3.4 we thus have $x_k \to x^*$ if $\{x \in \mathcal{M} \,:\, f(x) \leq f(x_0)\} \subset U$.

REMARK 3.6. *Note the difference between steepest descent and an approximation of the so-called gradient flow, the solution to the ODE*

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -\nabla f(x)\,.$$

While the gradient flow aims at a smooth curve along which the direction at each point is the steepest descent direction, the steepest descent linesearch tries to move into a fixed direction as long as possible and only changes its direction to be the one of steepest descent if the old direction does no longer yield sufficient decrease. Therefore, the steepest descent typically takes longer steps in one direction.

REMARK 3.7. *The condition on $f_{R_{x_k}}$ in Theorems 3.2 and 3.3 may be untangled into conditions on f and R_{x_k}. For example, one might require f to be Lipschitz continuously differentiable and $T^{R_{x_k}}_{x_k, x_{k+1}}|_{\operatorname{span}\{p_k\}}$ to be uniformly bounded, which is the case for $R_{x_k} = \exp_{x_k}$ or $R_{x_k} = P_{\gamma[x_k; x_{k+1}]}$, for example.*

As a method with only linear convergence, steepest descent requires many iterations. Improvement can be obtained by choosing in each iteration the Newton direction p_k^N as search direction, that is, the solution to

$$\mathrm{D}^2 f_{R_{x_k}}(0)(p_k^N, v) = -\mathrm{D}f(x_k)v \quad \forall v \in T_{x_k}\mathcal{M}\,. \tag{3.3}$$

Note that the Newton direction is obtained with regard to the retraction R_{x_k}. The fact that $\mathrm{D}^2 f_{R_{x^*}}(0) = \mathrm{D}^2 f(x^*)$ at a stationary point x^* and the later result that the

Hessian only needs to be approximated (Proposition 3.11) suggests that $\mathrm{D}^2 f(x_k)$ or a different approximation could be used as well to obtain fast convergence. In contrast to steepest descent, Newton's method is invariant with respect to a rescaling of f and in the limit allows a constant step size and quadratic convergence as shown in the following sequence of propositions which can be transferred from standard results (e. g. [16, Sec. 3.3]).

PROPOSITION 3.8 (Newton step length). *Let f be twice differentiable and $\mathrm{D}^2 f_{R_x}(0)$ continuous in x. Consider Algorithm 3.1 and assume the α_k to satisfy (3.1) with $c_1 \leq \frac{1}{2}$ and the p_k to satisfy*

$$\lim_{k\to\infty} \frac{\left\| \mathrm{D}f(x_k) + \mathrm{D}^2 f_{R_{x_k}}(0)(p_k,\cdot) \right\|_{x_k}}{\|p_k\|_{x_k}} = 0\,. \tag{3.4}$$

Furthermore, let the $f_{R_{x_k}}$ be twice Lipschitz continuously differentiable on $\operatorname{span}\{p_k\}$ with uniform Lipschitz constant L. If x^* with $\mathrm{D}f(x^*)=0$ and $\mathrm{D}^2 f(x^*)$ bounded and coercive is a limit point of x_k so that $x_k \to_{k\in\mathcal{I}} x^*$ for some $\mathcal{I}\subset\mathbb{N}$, then $\alpha_k = 1$ would also satisfy (3.1) (independent of whether $\alpha_k = 1$ is actually chosen) for sufficiently large $k\in\mathcal{I}$.

Proof. Let $m>0$ be the lowest eigenvalue belonging to $\mathrm{D}^2 f(x^*) = \mathrm{D}^2 f_{R_{x^*}}(0)$. The continuity of the second derivative implies the uniform coercivity $\mathrm{D}^2 f_{R_{x_k}}(0)(v,v) \geq \frac{m}{2}\|v\|_{x_k}^2$ for all $k\in\mathcal{I}$ sufficiently large. From $\mathrm{D}^2 f_{R_{x_k}}(0)(p_k - p_k^N,\cdot) = \mathrm{D}f(x_k) + \mathrm{D}^2 f_{R_{x_k}}(0)(p_k,\cdot)$ we then obtain $p_k - p_k^N = o(\|p_k\|_{x_k})$. Furthermore, condition (3.4) implies $\lim_{k\to\infty} |\mathrm{D}f(x_k)p_k + \mathrm{D}^2 f_{R_{x_k}}(0)(p_k,p_k)| / \|p_k\|_{x_k}^2 = 0$ and thus

$$0 = \limsup_{k\to\infty} \frac{\mathrm{D}f(x_k)p_k}{\|p_k\|_{x_k}^2} + \frac{\mathrm{D}^2 f_{R_{x_k}}(0)(p_k,p_k)}{\|p_k\|_{x_k}^2} \geq \limsup_{k\to\infty,k\in\mathcal{I}} \frac{\mathrm{D}f(x_k)p_k}{\|p_k\|_{x_k}^2} + \frac{m}{2} \quad\Rightarrow\quad -\mathrm{D}f(x_k)p_k / \|p_k\|_{x_k}^2 \geq \frac{m}{4} \tag{3.5}$$

for $k\in\mathcal{I}$ sufficiently large. Due to $\|\mathrm{D}f(x_k)\|_{x_k} \to_{k\in\mathcal{I}} 0$ we deduce $\|p_k\|_{x_k} \to_{k\in\mathcal{I}} 0$.

By Taylor's theorem, $f_{R_{x_k}}(p_k) = f_{R_{x_k}}(0) + \mathrm{D}f_{R_{x_k}}(0)p_k + \frac{1}{2}\mathrm{D}^2 f_{R_{x_k}}(q_k)(p_k,p_k)$ for some $q_k \in [0,p_k]$ so that

$$\begin{aligned} f_{R_{x_k}}(p_k) - f(x_k) - \frac{1}{2}\mathrm{D}f(x_k)p_k &= \frac{1}{2}(\mathrm{D}f(x_k)p_k + \mathrm{D}^2 f_{R_{x_k}}(q_k)(p_k,p_k)) \\ &= \frac{1}{2}\Big[\Big(\mathrm{D}f(x_k)p_k + \mathrm{D}^2 f_{R_{x_k}}(0)(p_k^N,p_k)\Big) + \mathrm{D}^2 f_{R_{x_k}}(0)(p_k - p_k^N,p_k) \\ &\qquad + \Big(\mathrm{D}^2 f_{R_{x_k}}(q_k) - \mathrm{D}^2 f_{R_{x_k}}(0)\Big)(p_k,p_k)\Big] \\ &\leq o(\|p_k\|_{x_k}^2) \end{aligned}$$

which implies feasibility of $\alpha_k = 1$ with respect to (3.1a) for $k\in\mathcal{I}$ sufficiently large. Also,

$$|\mathrm{D}f_{R_{x_k}}(p_k)p_k| = \left| \mathrm{D}f(x_k)p_k + \mathrm{D}^2 f_{R_{x_k}}(0)(p_k,p_k) + \int_0^1 \Big(\mathrm{D}^2 f_{R_{x_k}}(tp_k) - \mathrm{D}^2 f_{R_{x_k}}(0)\Big)(p_k,p_k)\,\mathrm{d}t \right| = o(\|p_k\|_{x_k}^2)$$

which together with (3.5) implies $\mathrm{D}f_{R_{x_k}}(p_k)p_k \geq c_2 \mathrm{D}f(x_k)p_k$ for sufficiently large $k\in\mathcal{I}$, that is, (3.1b) for $\alpha_k = 1$. ∎

LEMMA 3.9. *Let $U\subset\mathcal{M}$ be open and retractions $R_x : T_x\mathcal{M}\to\mathcal{M}$, $x\in U$, have equicontinuous derivatives at x in the sense*

$$\forall\varepsilon>0 \exists\delta>0 \forall x\in U : \|v\|_x < \delta \Rightarrow \left\| T_{x,R_x(v)}^{P_{\gamma[x;R_x(v)]}} \mathrm{D}R_x(0) - \mathrm{D}R_x(v) \right\| < \varepsilon\,.$$

Then for any $\varepsilon > 0$ there is an $\varepsilon' > 0$ such that for all $x \in U$ and $v, w \in T_x\mathcal{M}$ with $\|v\|_x, \|w\|_x < \varepsilon'$,

$$(1-\varepsilon)\|w-v\|_x \leq \operatorname{dist}(R_x(v), R_x(w)) \leq (1+\varepsilon)\|w-v\|_x .$$

Proof. For $\varepsilon > 0$ there is $\delta > 0$ such that for any $\tilde{x} \in U$, $\|v\|_{\tilde{x}} < \delta$ implies $\left\| T^{P_{\gamma[\tilde{x};R_{\tilde{x}}(v)]}}_{\tilde{x},R_{\tilde{x}}(v)} - \mathrm{D}R_{\tilde{x}}(v) \right\| < \varepsilon$. From this we obtain for $v, w \in T_x\mathcal{M}$ with $\|v\|_x, \|w\|_x < \delta$ that

$$\begin{aligned}
\operatorname{dist}(R_x(v), R_x(w)) &\leq \int_0^1 \|\mathrm{D}R_x(v+t(w-v))(w-v)\|_{R_x(v+t(w-v))}\,\mathrm{d}t \\
&\leq \|w-v\|_x + \int_0^1 \left\| \left[\mathrm{D}R_x(v+t(w-v)) - T^{P_{\gamma[x;R_x(v+t(w-v))]}}_{x,R_x(v+t(w-v))}\right](w-v) \right\|_{R_x(v+t(w-v))}\,\mathrm{d}t \\
&\leq (1+\varepsilon)\|w-v\|_x .
\end{aligned}$$

Furthermore, for δ small enough, the shortest geodesic path between $R_x(v)$ and $R_x(w)$ can be expressed as $t \mapsto R_x(p(t))$, where $p : [0,1] \to T_x\mathcal{M}$ with $p(0) = v$ and $p(1) = w$. Then, for $\|v\|_x, \|w\|_x < (1-\varepsilon)\frac{\delta}{2} =: \varepsilon'$,

$$\begin{aligned}
\operatorname{dist}(R_x(v), R_x(w)) &= \int_0^1 \left\| \mathrm{D}R_x(p(t))\frac{\mathrm{d}p(t)}{\mathrm{d}t} \right\|_{R_x(p(t))} \mathrm{d}t \\
&\geq \int_0^1 \left\| \frac{\mathrm{d}p(t)}{\mathrm{d}t} \right\|_x \mathrm{d}t - \int_0^1 \left\| \left[\mathrm{D}R_x(p(t)) - T^{P_{\gamma[x;R_x(p(t))]}}_{x,R_x(p(t))}\right]\frac{\mathrm{d}p(t)}{\mathrm{d}t} \right\|_{R_x(p(t))} \mathrm{d}t \\
&\geq (1-\varepsilon)\int_0^1 \left\| \frac{\mathrm{d}p(t)}{\mathrm{d}t} \right\|_x \mathrm{d}t \geq (1-\varepsilon)\|w-v\|_x ,
\end{aligned}$$

where we have used $\|p(t)\|_x < \delta$ for all $t \in [0,1]$, since otherwise one could apply the above estimate to the segments $[0, t_1)$ and $(t_2, 1]$ with $t_1 = \inf\{t : \|p(t)\|_x \geq \delta\}$, $t_2 = \sup\{t : \|p(t)\|_x \geq \delta\}$, which yields $\operatorname{dist}(R_x(v), R_x(w)) \geq (1-\varepsilon)\int_{[0,t_1)\cup(t_2,1]} \left\| \frac{\mathrm{d}p(t)}{\mathrm{d}t} \right\|_x \mathrm{d}t \geq (1-\varepsilon)2(\delta - \varepsilon') = (1-\varepsilon^2)\delta > (1+\varepsilon)\|w-v\|_x$, contradicting the first estimate. □

Proposition 3.10 (Convergence of Newton's method). *Let f be twice differentiable and $\mathrm{D}^2 f_{R_x}(0)$ continuous in x. Consider Algorithm 3.1 where $p_k = p_k^N$ as defined in (3.3), α_k satisfies (3.1) with $c_1 \leq \frac{1}{2}$, and $\alpha_k = 1$ whenever possible. Assume x_k has a limit point x^* with $\mathrm{D}f(x^*) = 0$ and $\mathrm{D}^2 f(x^*)$ bounded and coercive. Furthermore, assume that in a neighborhood U of x^*, the $\mathrm{D}R_{x_k}$ are equicontinuous in the above sense and that the $f_{R_{x_k}}$ with $x_k \in U$ are twice Lipschitz continuously differentiable on $R_{x_k}^{-1}(U)$ with uniform Lipschitz constant L. Then $x_k \to x^*$ with*

$$\lim_{k\to\infty} \frac{\operatorname{dist}(x_{k+1}, x^*)}{\operatorname{dist}^2(x_k, x^*)} \leq C$$

for some $C > 0$.

Proof. There is a subsequence $(x_k)_{k\in\mathcal{I}}$, $\mathcal{I} \subset \mathbb{N}$, with $x_k \to_{k\in\mathcal{I}} x^*$. By Proposition 3.8, $\alpha_k = 1$ for $k \in \mathcal{I}$ sufficiently large. Furthermore, by the previous lemma there is $\epsilon > 0$ such that $\frac{1}{2}\|w-v\|_{x_k} \leq \operatorname{dist}(R_{x_k}(v), R_{x_k}(w)) \leq \frac{3}{2}\|w-v\|_{x_k}$ for all $w, v \in T_{x_k}\mathcal{M}$ with $\|v\|_{x_k}, \|w\|_{x_k} < \epsilon$. Hence, for $k \in \mathcal{I}$ large enough such that

$\operatorname{dist}(x_k, x^*) < \frac{\epsilon}{4}$, there exists $R_{x_k}^{-1}(x^*)$, and

$$\begin{aligned}\left\|\mathrm{D}^2 f_{R_{x_k}}(0)(p_k - R_{x_k}^{-1}(x^*), \cdot)\right\|_{x_k} &= \left\|(\mathrm{D} f_{R_{x_k}}(R_{x_k}^{-1}(x^*)) - \mathrm{D} f(x_k)) - \mathrm{D}^2 f_{R_{x_k}}(0)(R_{x_k}^{-1}(x^*), \cdot)\right\|_{x_k} \\ &= \left\|\int_0^1 \left(\mathrm{D}^2 f_{R_{x_k}}(tR_{x_k}^{-1}(x^*)) - \mathrm{D}^2 f_{R_{x_k}}(0)\right)(R_{x_k}^{-1}(x^*), \cdot)\,\mathrm{d}t\right\|_{x_k} \leq L \left\|R_{x_k}^{-1}(x^*)\right\|_{x_k}^2\end{aligned}$$

which implies $\left\|p_k - R_{x_k}^{-1}(x^*)\right\|_{x_k} \leq 2\frac{L}{m}\left\|R_{x_k}^{-1}(x^*)\right\|_{x_k}^2$ for the smallest eigenvalue m of $\mathrm{D}^2 f(x^*)$ (using the same argument as in the proof of Proposition 3.8). The previous lemma then yields the desired convergence rate (note from the proof of Proposition 3.8 that p_k tends to zero) and thus also convergence of the whole sequence. □

The Riemannian Newton method was already proposed by Gabay in 1982 [9]. Smith proved quadratic convergence for the more general case of applying Newton's method to find a zero of a one-form on a manifold with the Levi-Civita connection [21] (the method is stated for an unspecified affine connection in [4]), using geodesic steps. Yang rephrased the proof within a broader framework for optimization algorithms, however, restricting to the case of minimizing a function (which corresponds to the method introduced above, only with geodesic retractions) [25]. Hüper and Trumpf show quadratic convergence even if the retraction used for taking the step is different from the retraction used for computing the Newton direction [11]. A more detailed overview is provided in [2, Sec. 6.6]. While all these approaches were restricted to finite-dimensional manifolds, we here explicitly include the case of infinite-dimensional manifolds.

A superlinear convergence rate can also be achieved if the Hessian in each step is only approximated, which is particularly interesting with regard to our aim of also analyzing a quasi-Newton minimization approach.

PROPOSITION 3.11 (Convergence of approximated Newton's method). *Let the assumptions of the previous proposition hold, but instead of the Newton direction p_k^N consider directions p_k which only satisfy* (3.4)*. Then we have superlinear convergence,*

$$\lim_{k\to\infty} \frac{\operatorname{dist}(x_{k+1}, x^*)}{\operatorname{dist}(x_k, x^*)} = 0\,.$$

Proof. From the proof of Proposition 3.8 we know $\left\|p_k - p_k^N\right\|_{x_k} = o(\|p_k\|_{x_k})$. Also, the proof of Proposition 3.10 shows $\left\|p_k^N - R_{x_k}^{-1}(x^*)\right\|_{x_k} \leq 2\frac{L}{m}\left\|R_{x_k}^{-1}(x^*)\right\|_{x_k}^2$. Thus we obtain

$$\left\|p_k - R_{x_k}^{-1}(x^*)\right\|_{x_k} \leq \left\|p_k - p_k^N\right\|_{x_k} + \left\|p_k^N - R_{x_k}^{-1}(x^*)\right\|_{x_k} \leq o(\|p_k\|_{x_k}) + 2\frac{L}{m}\left\|R_{x_k}^{-1}(x^*)\right\|_{x_k}^2\,.$$

This inequality first implies $\|p_k\|_{x_k} = O(\left\|R_{x_k}^{-1}(x^*)\right\|_{x_k})$ and then $\left\|p_k - R_{x_k}^{-1}(x^*)\right\|_{x_k} = o(\left\|R_{x_k}^{-1}(x^*)\right\|_{x_k})$ so that the result follows as in the proof of Proposition 3.10. □

REMARK 3.12. *Of course, again the conditions on $f_{R_{x_k}}$ in the previous analysis can be untangled into separate conditions on f and the retractions. For example, one might require f to be twice Lipschitz continuously differentiable and the retractions R_x to have uniformly bounded second derivatives for x in a neighborhood of x^* and arguments in a neighborhood of $0 \in T_x\mathcal{M}$. For example, $R_{x_k} = \exp_{x_k}$ or $R_{x_k} = P_{\gamma[x_k;x_{k+1}]}$ satisfy these requirements if the manifold $\mathcal{M}$ is well-behaved near x^*. (On a two-dimensional manifold, the second derivative of the exponential map deviates the stronger from zero, the larger the Gaussian curvature is in absolute value. On*

higher-dimensional manifolds, to compute the directional second derivative of the exponential map in two given directions, it suffices to consider only the two-dimensional submanifold which is spanned by the two directions, so the value of the directional second derivative depends on the sectional curvature belonging to these directions. If all sectional curvatures of $\mathcal{M}$ are uniformly bounded near x^*—which is not necessarily the case for infinite-dimensional manifolds, this implies also the boundedness of the second derivatives of the exponential map so that $R_{x_k} = \exp_{x_k}$ or $R_{x_k} = P_{\gamma[x_k;x_{k+1}]}$ indeed satisfy the requirements.)

3.1. BFGS quasi-Newton scheme. A classical way to retain superlinear convergence without computing the Hessian at every iterate consists in the use of quasi-Newton methods, where the objective function Hessian is approximated via the gradient information at the past iterates. The most popular method, which we would like to transfer to the manifold setting here, is the BFGS rank-2-update formula. Here, the search direction p_k is chosen as the solution to

$$B_k(p_k, \cdot) = -\mathrm{D}f(x_k)\,, \tag{3.6}$$

where the bilinear forms $B_k : (T_{x_k}\mathcal{M})^2 \to \mathbb{R}$ are updated according to

$$\begin{aligned} s_k &= \alpha_k p_k = R_{x_k}^{-1}(x_{k+1}) \\ y_k &= \mathrm{D}f_{R_{x_k}}(s_k) - \mathrm{D}f_{R_{x_k}}(0) \\ B_{k+1}(T_k v, T_k w) &= B_k(v,w) - \frac{B_k(s_k,v)B_k(s_k,w)}{B_k(s_k,s_k)} + \frac{(y_k v)(y_k w)}{y_k s_k} \quad \forall v, w \in T_{x_k}\mathcal{M}\,. \end{aligned}$$

Here, $T_k \equiv T_{x_k,x_{k+1}}$ denotes some linear map from $T_{x_k}\mathcal{M}$ to $T_{x_{k+1}}\mathcal{M}$, which we obviously require to be invertible to make B_{k+1} well-defined.

REMARK 3.13. *There are more possibilities to define the BFGS update, for example, using $s_k = -R_{x_{k+1}}^{-1}(x_k)$, $y_k = \mathrm{D}f_{R_{x_{k+1}}}(0) - \mathrm{D}f_{R_{x_{k+1}}}(-s_k)$, and a corresponding update formula for B_{k+1} (which looks as above, only with the vector transport at different places). The analysis works analogously, and the above choice only allows the most elegant notation.*

Actually, the formulation from the above remark was already introduced by Gabay [9] (for geodesic retractions and parallel transport) and resumed by Absil et al. [2, 18] (for general retractions and vector transport). A slightly different variant, which ignores any kind of vector transport, was provided in [12], together with a proof of convergence. In contrast to these approaches, we also consider infinite-dimensional manifolds and prove convergence as well as superlinear convergence rate.

LEMMA 3.14. *Consider Algorithm 3.1 with the above BFGS search direction and Wolfe step size control, where $\|T_k\|$ and $\|T_k^{-1}\|$ are uniformly bounded and the $f_{R_{x_k}}$ are assumed smooth. If B_0 is bounded and coercive, then*

$$y_k s_k > 0$$

and B_k is bounded and coercive for all $k \in \mathbb{N}$.

Proof. Assume B_k to be bounded and coercive. Then p_k is a descent direction, and (3.1b) implies the curvature condition $y_k s_k \geq (c_2 - 1)\alpha_k \mathrm{D}f_{R_{x_k}}(0)p_k > 0$ so that B_{k+1} is well-defined.

Let us denote by $\hat{y}_k = \nabla f_{R_{x_k}}(s_k) - \nabla f_{R_{x_k}}(0)$ the Riesz representation of y_k. Furthermore, by the Lax–Milgram lemma, there is $\hat{B}_k : T_{x_k}\mathcal{M} \to T_{x_k}\mathcal{M}$ with $B_k(v,w) =$

$g_{x_k}(v, \hat{B}_k w)$ for all $v, w \in T_{x_k}\mathcal{M}$. Obviously,

$$\hat{B}_{k+1} = T_k^{-*}\left[\hat{B}_k - \frac{B_k(s_k,\cdot)\hat{B}_k s_k}{B_k(s_k,s_k)} + \frac{y_k(\cdot)\hat{y}_k}{y_k s_k}\right] T_k^{-1}$$

If $H_k = \hat{B}_k^{-1}$, then by the Sherman–Morrison formula,

$$H_{k+1} = T_k\left[J^* H_k J + \frac{g_{x_k}(s_k,\cdot)s_k}{y_k s_k}\right] T_k^*, \quad J = \left(\mathrm{id} - \frac{g_{x_k}(s_k,\cdot)\hat{y}_k}{y_k s_k}\right), \tag{3.7}$$

is the inverse of $\hat{B}_{k+1}$. Its boundedness is obvious. Furthermore, H_{k+1} is coercive. Indeed, let $M = \left\|\hat{B}_k\right\|$, then for any $v \in T_{x_k}\mathcal{M}$ with $\|v\|_{x_k} = 1$ and $w = v - \frac{g_{x_k}(s_k,v)}{y_k s_k}\hat{y}_k$ we have

$$g_{x_{k+1}}(T_k^{-*}v, H_{k+1}T_k^{-*}v) = g_{x_k}(w, H_k w) + \frac{g_{x_k}(s_k,v)^2}{y_k s_k} \geq \frac{1}{M}\|w\|_{x_k}^2 + \frac{g_{x_k}(s_k,v)^2}{y_k s_k}$$

which is strictly greater than zero since the second summand being zero implies the first one being greater than or equal to $\frac{1}{M}$. In finite dimensions this already yields the desired coercivity, in infinite dimensions it still remains to show that the right-hand side is uniformly bounded away from zero for all unit vectors $v \in T_{x_k}\mathcal{M}$. Indeed,

$$\frac{1}{M}\|w\|_{x_k}^2 + \frac{g_{x_k}(s_k,v)^2}{y_k s_k} = \frac{1}{M}\left[1 - 2\frac{g_{x_k}(s_k,v)}{y_k s_k}g_{x_k}(\hat{y}_k, v) + \left(\frac{g_{x_k}(s_k,v)}{y_k s_k}\right)^2 \|\hat{y}_k\|_{x_k}^2\right] + \frac{g_{x_k}(s_k,v)^2}{y_k s_k}$$

is a continuous function in $(g_{x_k}(s_k,v), g_{x_k}(\hat{y}_k,v))$ which by Weierstrass' theorem takes its minimum on $[-\|s_k\|_{x_k}, \|s_k\|_{x_k}] \times [-\|y_k\|_{x_k}, \|y_k\|_{x_k}]$. This minimum is the smallest eigenvalue of H_{k+1} and must be greater than zero.

Both the boundedness and coercivity of $T_k^{-1}H_{k+1}T_k^{-*}$ imply boundedness and coercivity of $\hat{B}_{k+1}$ and thus of B_{k+1}. ∎

By virtue of the above theorem, the BFGS search direction is well-defined for all iterates. It satisfies the all-important secant condition

$$B_{k+1}(T_k s_k, \cdot) = y_k T_k^{-1}$$

which basically has the interpretation that $B_{k+1} \circ (T_k)^2$ is supposed to be an approximation to $\mathrm{D}^2 f_{R_{x_k}}(s_k)$. Since we also aim at $B_{k+1} \approx \mathrm{D}^2 f_{R_{x_{k+1}}}(0)$, the choice $T_k = T_{x_k,x_{k+1}}^{R_{x_k}}$ seems natural in view of the fact $\mathrm{D}^2 f_{R_x}(R_x^{-1}(x^*)) = \mathrm{D}^2 f_{R_{x^*}}(0) \circ (T_{x,x^*}^{R_x})^2$ for a stationary point x^*. However, we were only able to establish the method's convergence for isometric T_k (see Proposition 3.15).

For the actual implementation, instead of solving (3.6) one applies H_k to $-\nabla f(x_k)$. By (3.7), this entails computing the transported vectors $T_{k-1}^*\nabla f(x_k), T_{k-2}^* T_{k-1}^* \nabla f(x_k), \ldots, T_0^* T_1^* \ldots T_{k-1}^* \nabla f(x_k)$ and their scalar products with the s_i and y_i, $i = 0, \ldots, k-1$, as well as the transports of the s_i and y_i. The convergence analysis of the scheme can be transferred from standard analyses (e. g. [17, Prop. 1.6.17-Thm. 1.6.19] or [16, Thm. 6.5] and [12] for slightly weaker results). Due to their technicality, the corresponding proofs are deferred to the appendix.

Proposition 3.15 (Convergence of BFGS descent). *Consider Algorithm 3.1 with BFGS search direction and Wolfe step size control, where the T_k are isometries.*

Assume the $f_{R_{x_k}}$ to be uniformly convex on the $f(x_0)$-sublevel set of f, that is, there are $0 < m < M < \infty$ such that

$$m \|v\|_{x_k}^2 \leq \mathrm{D}^2 f_{R_{x_k}}(p)(v,v) \leq M \|v\|_{x_k}^2 \quad \forall v \in T_{x_k}\mathcal{M}$$

for $p \in R_{x_k}^{-1}(\{x \in \mathcal{M} : f(x) \leq f(x_0)\})$. If B_0 is symmetric, bounded and coercive, then there exists a constant $0 < \mu < 1$ such that

$$f(x_k) - f(x^*) \leq \mu^{k+1}(f(x_0) - f(x^*))$$

for the minimizer $x^* \in \mathcal{M}$ of f.

COROLLARY 3.16. *Under the conditions of the previous proposition and if $f_{\exp_{x^*}}$ is uniformly convex in the sense*

$$m \|v\|_{x^*}^2 \leq \mathrm{D}^2 f_{\exp_{x^*}}(p)(v,v) \leq M \|v\|_{x^*}^2 \quad \forall v \in T_{x^*}\mathcal{M}$$

if $f_{\exp_{x^*}}(p) \leq f(x_0)$, then

$$\operatorname{dist}(x_k, x^*) \leq \sqrt{\frac{M}{m}}\sqrt{\mu}^{k+1}\operatorname{dist}(x_0, x^*) \quad \forall k \in \mathbb{N}\,.$$

Proof. From $f(x) - f(x^*) = \frac{1}{2}\mathrm{D}^2 f_{\exp_{x^*}}(t \log_{x^*} x)(\log_{x^*} x, \log_{x^*} x)$ for some $t \in [0,1]$ we obtain $\frac{m}{2}\operatorname{dist}^2(x,x^*) \leq f(x) - f(x^*) \leq \frac{M}{2}\operatorname{dist}^2(x,x^*)$. ∎

REMARK 3.17. *The isometry condition on T_k is for example satisfied for the parallel transport from x_k to x_{k+1}. The Riesz representation of y_k can be computed as $(T_{x_k,x_{k+1}}^{R_{x_k}})^* \nabla f(x_{k+1}) - \nabla f(x_k)$. For $R_{x_k} = P_{\gamma[x_k;x_{k+1}]}$, this means simple parallel transport of $\nabla f(x_{k+1})$.*

For a superlinear convergence rate, it is well-known that in infinite dimensions one needs particular conditions on the initial Hessian approximation B_0 (it has to be an approximation to the true Hessian in the nuclear norm, see e. g. [19]). On manifolds we furthermore require a certain type of consistency between the vector transports as shown in the following proposition, which modifies the analysis from [16, Thm. 6.6]. For ease of notation, define the averaged Hessian $G_k = \int_0^1 \mathrm{D}^2 f_{R_{x_k}}(t s_k)\,\mathrm{d}t$ and let the hat in $\hat{G}_k$ denote the Lax–Milgram representation of G_k. The proof of the following is given in the appendix.

PROPOSITION 3.18 (BFGS approximation of Newton direction). *Consider Algorithm 3.1 with BFGS search direction and Wolfe step size control. Let x^* be a stationary point of f with bounded and coercive $\mathrm{D}^2 f(x^*)$. For k large enough, assume $R_{x_k}^{-1}(x^*)$ to be defined and assume the existence of isomorphisms $T_{*,k} : T_{x^*}\mathcal{M} \to T_{x_k}\mathcal{M}$ with $\|T_{*,k}\|, \left\|T_{*,k}^{-1}\right\| < C$ uniformly for some $C > 0$ and*

$$\left\|T_{x_k,x^*}^{R_{x_k}} T_{*,k} - \mathrm{id}\right\| \underset{k\to\infty}{\longrightarrow} 0\,, \quad \left\|T_{*,k+1}^{-1} T_k T_{*,k} - \mathrm{id}\right\|_1 < b\beta^k$$

for some $b > 0$, $0 < \beta < 1$, where $\|\cdot\|_1$ denotes the nuclear norm (the sum of all singular values, $\|T\|_1 = \operatorname{tr}\sqrt{T^*T}$). Finally, assume $\left\|T_{*,0}^* \hat{B}_0 T_{*,0} - \nabla^2 f(x^*)\right\|_1 < \infty$ for a symmetric, bounded and coercive B_0 and let

$$\sum_{k=0}^{\infty} \left\|T_{*,k}^* \hat{G}_k T_{*,k} - \nabla^2 f(x^*)\right\| < \infty \quad \text{and} \quad \left\|\nabla^2 f_{R_{x_k}}(R_{x_k}^{-1}(x^*)) - \nabla^2 f_{R_{x_k}}(0)\right\| \underset{k\to\infty}{\longrightarrow} 0\,.$$

Then

$$\lim_{k\to\infty} \frac{\left\| \mathrm{D}f(x_k) + \mathrm{D}^2 f_{R_{x_k}}(0)(p_k,\cdot)\right\|_{x_k}}{\|p_k\|_{x_k}} = 0$$

so that Proposition 3.11 can be applied.

COROLLARY 3.19 (Convergence rate of BFGS descent). *Consider Algorithm 3.1 with BFGS search direction and Wolfe step size control, where $c_1 \leq \frac{1}{2}$, $\alpha_k = 1$ whenever possible, and the T_k are isometries. Let x^* be a stationary point of f with bounded and coercive $\mathrm{D}^2 f(x^*)$, where we assume $f_{\exp_{x^*}}$ to be uniformly convex on its $f(x_0)$-sublevel set. Assume that in a neighborhood U of x^*, the $\mathrm{D}R_{x_k}$ are equicontinuous (in the sense of Lemma 3.9) and that the $f_{R_{x_k}}$ with $x_k \in U$ are twice Lipschitz continuously differentiable on $R_{x_k}^{-1}(U)$ with uniform Lipschitz constant L and uniformly convex on the $f(x_0)$-sublevel set of f. Furthermore, assume the existence of isomorphisms $T_{*,k} : T_{x^*}\mathcal{M} \to T_{x_k}\mathcal{M}$ with $\|T_{*,k}\|$ and $\left\|T_{*,k}^{-1}\right\|$ uniformly bounded and*

$$\left\|T_{x_k,x^*}^{R_{x_k}} T_{*,k} - \mathrm{id}\right\|, \left\|T_{*,k+1}^{-1} T_k T_{*,k} - \mathrm{id}\right\|_1 \leq c \max\{\mathrm{dist}(x_k,x^*), \mathrm{dist}(x_{k+1},x^*)\}$$

for some $c > 0$. If $\hat{B}_0$ is bounded and coercive with $\left\|T_{*,0}^* \hat{B}_0 T_{*,0} - \nabla^2 f(x^*)\right\|_1 < \infty$, then $x_k \to x^*$ with

$$\lim_{k\to\infty} \frac{\mathrm{dist}(x_{k+1},x^*)}{\mathrm{dist}(x_k,x^*)} = 0\,.$$

Proof. The conditions of Corollary 3.16 are satisfied so that we have $\mathrm{dist}(x_k,x^*) < b\beta^k$ for some $b > 0$, $0 < \beta < 1$. As in the proof of Proposition 3.10, the equicontinuity of the retraction variations then implies the well-definedness of $R_{x_k}^{-1}(x^*)$ for k large enough. Also,

$$\left\|T_{x_k,x^*}^{R_{x_k}} T_{*,k} - \mathrm{id}\right\|\,,\ \left\|T_{*,k+1}^{-1} T_k T_{*,k} - \mathrm{id}\right\|_1 < cb\beta^k \underset{k\to\infty}{\longrightarrow} 0\,.$$

Furthermore, $\left\|\nabla^2 f_{R_{x_k}}(R_{x_k}^{-1}(x^*)) - \nabla^2 f_{R_{x_k}}(0)\right\| \leq L\|R_{x_k}^{-1}(x^*)\|_{x_k} = O(\mathrm{dist}(x_k,x^*)) \underset{k\to\infty}{\longrightarrow} 0$ follows from the uniform Lipschitz continuity of $\nabla^2 f_{R_{x_k}}$ and Lemma 3.9. Likewise,

$$\left\|G_k \circ (T_{*,k})^2 - \mathrm{D}^2 f(x^*)\right\| \leq \left\|T_{*,k}^* \nabla^2 f_{R_{x_k}}(0) T_{*,k} - \nabla^2 f(x^*)\right\| + \|T_{*,k}\|^2 \left\|G_k - \mathrm{D}^2 f_{R_{x_k}}(0)\right\|\,.$$

The second term is bounded by $\|T_{*,k}\|^2 L \left\|R_{x_k}^{-1}(x_{k+1})\right\|_{x_k}$ and thus by some constant times β^k due to the equivalence between the retractions and the exponential map near x^* (Lemma 3.9). The first term on the right-hand side can be bounded as in the proof of Proposition 3.18 which also yields a constant times β^k so that Proposition 3.18 can be applied. Proposition 3.11 then implies the result. ∎

On finite-dimensional manifolds, the operator norm $\|\cdot\|$ is equivalent to the nuclear norm $\|\cdot\|_1$, and the condition on B_0 reduces to B_0 being bounded and coercive. In that case, if the manifold $\mathcal{M}$ has bounded sectional curvatures near x^*, $f_{\exp_{x^*}}$ is uniformly convex, and f twice Lipschitz continuously differentiable, then the above conditions are for instance satisfied with $R_{x_k} = \exp_{x_k}$ or $R_{x_k} = P_{\gamma[x_k;x_{k+1}]}$ and T_k as well as $T_{*,k}$ being standard parallel transport.

Note that the nuclear norm bound on $T_{*,k+1}^{-1} T_k T_{*,k} - \mathrm{id}$ is quite a strong condition, and it is not clear whether it could perhaps be relaxed. If for illustration we imagine T_k and $T_{*,k}$ to be simple parallel transport, then the bound implies that the manifold behaves almost like a hyperplane (except for a finite number of dimensions).

3.2. Fletcher–Reeves nonlinear conjugate gradient scheme. Other gradient-based minimization methods with good convergence properties in practice include nonlinear conjugate gradient schemes. Here, in each step the search direction is chosen conjugate to the old search direction in a certain sense. Such methods often enjoy superlinear convergence rates. For example, Luenberger has shown

$$\|x_{k+n} - x^*\| = O(\|x_k - x^*\|^2)$$

for the Fletcher–Reeves nonlinear conjugate gradient iteration on flat n-dimensional manifolds. However, such analyses seem quite special and intricate, and the understanding of these methods seems not yet as advanced as of the quasi-Newton approaches, for example. Hence, we will only briefly transfer the convergence analysis for the standard Fletcher–Reeves approach to the manifold case. Here, the search direction is chosen according to

$$\begin{aligned} p_k &= -\nabla f(x_k) + \beta_k T_{x_{k-1},x_k}^{R_{x_{k-1}}} p_{k-1}\,, \\ \beta_k &= \frac{\|\mathrm{D}f(x_k)\|_{x_k}^2}{\|\mathrm{D}f(x_{k-1})\|_{x_{k-1}}^2}\,, \end{aligned}$$

(with $\beta_0 = 0$) and the step length α_k satisfies the strong Wolfe condition.

REMARK 3.20. *If (as is typically done) the nonlinear conjugate gradient iteration is restarted every K steps with $p_{iK} = -\nabla f(x_{iK})$, $i \in \mathbb{N}$, then Theorem 3.3 directly implies* $\liminf_{k\to\infty} \|\mathrm{D}f(x_k)\|_{x_k} = 0$. *Hence, if f is strictly convex in the sense of Proposition 3.5, we obtain convergence of x_k against the minimizer.*

We immediately have the following classical bound (e. g. [16, Lem. 5.6]).

LEMMA 3.21. *Consider Algorithm 3.1 with Fletcher–Reeves search direction and strong Wolfe step size control with $c_2 < \frac{1}{2}$. Then*

$$-\frac{1}{1-c_2} \le \frac{\mathrm{D}f(x_k)p_k}{\|\mathrm{D}f(x_k)\|_{x_k}^2} \le \frac{2c_2 - 1}{1 - c_2} \quad \forall k \in \mathbb{N}\,.$$

Proof. The result is obvious for $k = 0$. In order to perform an induction, note

$$\frac{\mathrm{D}f(x_{k+1})p_{k+1}}{\|\mathrm{D}f(x_{k+1})\|_{x_{k+1}}^2} = -1 + \beta_{k+1}\frac{\mathrm{D}f(x_{k+1})T_{x_k,x_{k+1}}^{R_{x_k}} p_k}{\|\mathrm{D}f(x_{k+1})\|_{x_{k+1}}^2} = -1 + \frac{\mathrm{D}f_{R_{x_k}}(\alpha_k p_k)p_k}{\|\mathrm{D}f(x_k)\|_{x_k}^2}\,.$$

The strong Wolfe condition (3.2) thus implies

$$-1 + c_2\frac{\mathrm{D}f(x_k)p_k}{\|\mathrm{D}f(x_k)\|_{x_k}^2} \le \frac{\mathrm{D}f(x_{k+1})p_{k+1}}{\|\mathrm{D}f(x_{k+1})\|_{x_{k+1}}^2} \le -1 - c_2\frac{\mathrm{D}f(x_k)p_k}{\|\mathrm{D}f(x_k)\|_{x_k}^2}$$

from which the result follows by the induction hypothesis. ∎

This bound allows the application of a standard argument (e. g. [16, Thm. 5.7]) to obtain $\liminf_{k\to\infty} \|\mathrm{D}f(x_k)\|_{x_k} = 0$. Thus, as before, if f is strictly convex in the sense of Proposition 3.5, the sequence x_k converges against the minimizer.

PROPOSITION 3.22 (Convergence of Fletcher–Reeves CG iteration). *Under the conditions of the previous lemma and if* $\left\|T^{R_{x_k}}_{x_k,x_{k+1}} p_k\right\|_{x_{k+1}} \leq \|p_k\|_{x_k}$ *for all* $k \in \mathbb{N}$, *then* $\liminf_{k\to\infty} \|\mathrm{D}f(x_k)\|_{x_k} = 0$.

Proof. The inequality of the previous lemma can be multiplied by $\frac{\|\mathrm{D}f(x_k)\|_{x_k}}{\|p_k\|_{x_k}}$ to yield

$$\frac{1-2c_2}{1-c_2}\frac{\|\mathrm{D}f(x_k)\|_{x_k}}{\|p_k\|_{x_k}} \leq \cos\theta_k \leq \frac{1}{1-c_2}\frac{\|\mathrm{D}f(x_k)\|_{x_k}}{\|p_k\|_{x_k}}\,.$$

Theorem 3.3 then implies $\sum_{k=0}^{\infty} \frac{\|\mathrm{D}f(x_k)\|^4_{x_k}}{\|p_k\|^2_{x_k}} < \infty$. Furthermore, by (3.2) and the previous lemma, $|\mathrm{D}f(x_k)T^{R_{x_{k-1}}}_{x_{k-1},x_k} p_{k-1}| \leq -c_2\mathrm{D}f(x_{k-1})p_{k-1} \leq \frac{c_2}{1-c_2}\|\mathrm{D}f(x_{k-1})\|^2_{x_{k-1}}$ so that

$$\begin{aligned}
\|p_k\|^2_{x_k} &= \|\nabla f(x_k)\|^2_{x_k} + 2\beta_k \mathrm{D}f(x_k)T^{R_{x_{k-1}}}_{x_{k-1},x_k} p_{k-1} + \beta_k^2 \left\|T^{R_{x_{k-1}}}_{x_{k-1},x_k} p_{k-1}\right\|^2_{x_k} \\
&\leq \|\nabla f(x_k)\|^2_{x_k} + \frac{2c_2}{1-c_2}\beta_k \|\mathrm{D}f(x_{k-1})\|^2_{x_{k-1}} + \beta_k^2 \|p_{k-1}\|^2_{x_{k-1}} \\
&= \frac{1+c_2}{1-c_2}\|\nabla f(x_k)\|^2_{x_k} + \frac{\|\mathrm{D}f(x_k)\|^4_{x_k}}{\|\mathrm{D}f(x_{k-1})\|^4_{x_{k-1}}}\|p_{k-1}\|^2_{x_{k-1}} \leq \frac{1+c_2}{1-c_2}\|\mathrm{D}f(x_k)\|^4_{x_k}\sum_{j=0}^{k}\|\mathrm{D}f(x_j)\|^{-2}_{x_j}\,,
\end{aligned}$$

where the last step follows from induction and $p_0 = -\nabla f(x_0)$. However, if we now assume $\|\mathrm{D}f(x_k)\|_{x_k} \geq \gamma$ for some $\gamma > 0$ and all $k \in \mathbb{N}$, then this implies $\|p_k\|^2_{x_k} \leq \frac{1+c_2}{1-c_2}\|\mathrm{D}f(x_k)\|^4_{x_k}\frac{k}{\gamma^2}$ so that $\sum_{k=0}^{\infty}\frac{\|\mathrm{D}f(x_k)\|^4_{x_k}}{\|p_k\|^2_{x_k}} \geq \frac{1-c_2}{1+c_2}\gamma^2\sum_{k=0}^{\infty}\frac{1}{k} = \infty$, contradicting Zoutendijk's theorem. ∎

REMARK 3.23. $R_{x_k} = \exp_{x_k}$ *and* $R_{x_k} = P_{\gamma[x_k;x_{k+1}]}$ *both satisfy the conditions required for convergence (the iterations for both retractions coincide). If the vector transport associated with the retraction increases the norm of the transported vector, convergence can no longer be guaranteed.*

REMARK 3.24. *If in the iteration we instead use*

$$\beta_k = \frac{\mathrm{D}f(x_k)(\nabla f(x_k) - \nabla f_{R_{x_k}}(R^{-1}_{x_k}(x_{k-1})))}{\|\mathrm{D}f(x_{k-1})\|^2_{x_{k-1}}} \quad \text{or} \quad \beta_{k+1} = \frac{\mathrm{D}f_{R_{x_k}}(\alpha_k p_k)(\nabla f_{R_{x_k}}(\alpha_k p_k) - \nabla f(x_k))}{\|\mathrm{D}f(x_k)\|^2_{x_k}}$$

we obtain a Polak–Ribière variant of a nonlinear CG iteration. If there are $0 < m < M < \infty$ *such that*

$$m\|v\|^2_{x_k} \leq \mathrm{D}^2 f_{R_{x_k}}(p)(v,v) \leq M\|v\|^2_{x_k} \quad \forall v \in T_{x_k}\mathcal{M}$$

whenever $f(R_{x_k}(p)) \leq f(x_0)$, *if* $\left\|(T^{R_{x_k}}_{x_k,x_{k+1}})^{-1}\right\|$ *(respectively* $\left\|T^{R_{x_k}}_{x_k,x_{k+1}}\right\|$*) is uniformly bounded, and if* α_k *is obtained by exact linesearch, then for all* $k \in \mathbb{N}$ *one can show* $\cos\theta_k \geq \frac{1}{1+\frac{M}{m}\left\|(T^{R_{x_k}}_{x_k,x_{k+1}})^{-1}\right\|} =: \rho$ *(respectively* $\cos\theta_k \geq \frac{1}{1+\frac{M}{m}\left\|T^{R_{x_k}}_{x_k,x_{k+1}}\right\|}$*) which implies at least linear convergence,*

$$\mathrm{dist}(x_k, x^*) \leq \sqrt{\frac{2}{m}[f(x_0) - f(x^*)]}\sqrt{1-\frac{m}{M}\rho^2}^{\,k}\,.$$

Here, the proofs of 1.5.8 and 1.5.9 from [17] can be directly transferred with $\lambda_i := \alpha_i$, $h_i := p_i$, $g_i := \nabla f(x_i)$, $H(x_i + s\lambda_i h_i) := \nabla^2 f_{R_{x_i}}(s\alpha_i p_i)$ *and replacing* g_{i+1} *by* $\mathrm{D}f_{x_i}(\alpha_i p_i)$ *in (19a),* $\langle g_i, h_{i-1}\rangle$ *by* $\mathrm{D}f_{R_{x_{i-1}}}(\alpha_{i-1}p_{i-1})p_{i-1}$, *and* $\langle H_i h_i, g_{i+1}\rangle$ *by either* $\langle (T^{R_{x_i}}_{x_i,x_{i+1}})^{-*} H_i h_i, g_{i+1}\rangle$ *or* $\langle H_i h_i, \mathrm{D}f_{R_{x_i}}(\alpha_i p_i)\rangle$ *in the denominator of (19d). 1.5.8(b) follows from Theorem 3.3 similarly to the proof of Proposition 3.15.*

4. Numerical examples. In this section we will first consider simple optimization problems on the two-dimensional torus to illustrate the influence of the Riemannian metric on the optimization progress. Afterwards we turn to an active contour model and a simulation of truss shape deformations as exemplary optimization problems to prove the efficiency of Riemannian optimization methods also in the more complex setting of Riemannian shape spaces.

4.1. The rôle of metric and vector transport. The minimum of a functional on a manifold is independent of the manifold metric and the chosen retractions. Consequently, exploiting the metric structure does not necessarily aid the optimization process, and one could certainly impose different metrics on the same manifold of which some are more beneficial for the optimization problem at hand than others. The optimal pair of a metric and a retraction would be such that one single gradient descent step already hits the minimum. However, the design of such pairs requires far more effort than solving the optimization problem in a suboptimal way.

Often, a certain metric and retraction fit naturally to the optimization problem at hand. For illustration, consider the two-dimensional torus, parameterized by

$$(\varphi, \psi) \mapsto y(\varphi, \psi) = ((r_1 + r_2\cos\varphi)\cos\psi, (r_1 + r_2\cos\varphi)\sin\psi, r_2\sin\varphi)\ , \quad \varphi, \psi \in [-\pi, \pi]\,,$$

$r_1 = 2$, $r_2 = \frac{3}{5}$. The corresponding metric shall be induced by the Euclidean embedding. As objective functions let us consider the following, both expressed as functions on the parametrization domain,

$$\begin{aligned} f_1(\varphi, \psi) &= a(1 - \cos\varphi) + b[(\psi + \pi/2) \bmod (2\pi) - \pi]^2\,, \\ f_2(\varphi, \psi) &= a(1 - \cos\varphi) + b\mathrm{dist}(\varphi, \Phi_\psi)^2\,, \end{aligned}$$

where we use $(a, b) = (1, 40)$ and for all $\psi \in [-\pi, \pi]$, Φ_ψ is a discrete set such that $\{(\varphi, \psi) \in [-\pi, \pi]^2 : \varphi \in \Phi_\psi\}$ describes a shortest curve winding five times around the torus and passing through $(\varphi, \psi) = (0, 0)$ (compare Figure 4.1, right). Both functions may be slightly altered so that they are smooth all over the torus.

f_1 exhibits a narrow valley that is aligned with geodesics of the parametrization domain, while the valleys of f_2 follow a geodesic path on the torus. Obviously, an optimization based on the (Euclidean) metric and (straight line) geodesic retractions of the parametrization domain is much better in following the valley of f_1 than an optimization based on the actual torus metric (Figure 4.1). For f_2 the situation is reverse. This phenomenon is also reflected in the iteration numbers until convergence (Table 4.1). It is not very pronounced for methods which converge after only few iterations, but it is very noticeable especially for gradient descent and nonlinear conjugate gradient iterations.

Of course, the qualitative convergence behavior stays the same (such as linear convergence for gradient descent or superlinear convergence for the BFGS method), independent of the employed retractions. Therefore it sometimes pays off to choose rapidly computable retractions over retractions that minimize the number of optimization steps. We might for example minimize f_1 or f_2 using the torus metric but

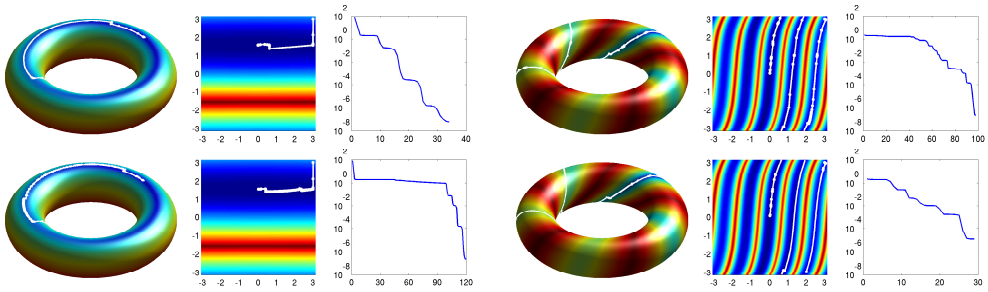

FIG. 4.1. *Minimization of f_1 (left) and f_2 (right) on the torus via the Fletcher–Reeves algorithm. The top row performs the optimization in the parametrization domain, the bottom row shows the result for the Riemannian optimization. For each case the optimization path is shown on the torus and in the parametrization domain (the color-coding from blue to red indicates the function value) as well as the evolution of the function values. Obviously, optimization in the parametrization domain is more suitable for f_1, whereas Riemannian optimization with the torus metric is more suitable for f_2.*

	objective f_1		objective f_2	
	parametr. metric	torus metric	parametr. metric	torus metric
gradient descent	213	890	9537	2212
Newton descent	22	28	38	29
BFGS quasi-Newton	16	28	55	69
Fletcher–Reeves NCG	34	120	98	29

TABLE 4.1

Iteration numbers for minimization of f_1 and f_2 with different methods. The iteration is stopped as soon as $\|(\frac{\partial f_i}{\partial \varphi}, \frac{\partial f_i}{\partial \psi})\|_{\ell^2} < 10^{-3}$. As retractions we use the exponential maps with respect to the metric of the parametrization domain and of the torus, respectively. The BFGS and NCG method employ the corresponding parallel transport.

non-geodesic retractions $R_x : T_x\mathcal{M} \to \mathcal{M}$, $v \mapsto y(y^{-1}(x) + \mathrm{D}y^{-1}v)$ (a straight step in the parametrization domain, where y was the parameterization). The resulting optimization will behave similarly to the optimization based on the (Euclidean) metric of the parametrization domain, and indeed, the Fletcher–Reeves algorithm requires 47 minimization steps for f_1 and 83 for f_2.

As a final discussion based on the illustrative torus example, consider the conditions on the vector transport T_k for the BFGS method. It is an open question whether the isometry of T_k is really needed for global convergence in Proposition 3.15. On compact manifolds such as the torus this condition is not needed since the sequence of iterates will always contain a converging subsequence so that global convergence follows from Proposition 3.18 instead of Proposition 3.15. A counterexample, showing the necessity of isometric transport, will likely be difficult to obtain. On the other hand, we can numerically validate the conditions on T_k to obtain superlinear convergence. Figure 4.2 shows the evolution of the distance between the current iterate x_k and $x^* = \arg\min f_2$ for the BFGS method with varying T_k. In particular, we take T_k to be parallel transport concatenated with a rotation by some angle ω that depends on $\mathrm{dist}(x_k, x^*)$. The term $\|T_{*,k+1}^{-1} T_k T_{*,k} - \mathrm{id}\|$ in Proposition 3.18 and Corollary 3.19 then scales (at least) like ω. Apparently, the superlinear convergence seems to be retained for $\omega \sim \mathrm{dist}(x_k, x^*)$ as well as $\omega \sim \sqrt{\mathrm{dist}(x_k, x^*)}$, which is better than predicted by Corollary 3.19. (This is not surprising, though,

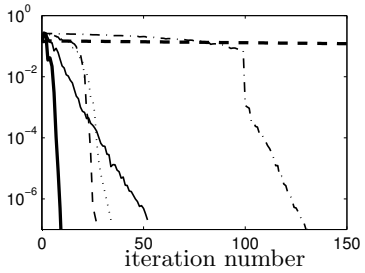

FIG. 4.2. *Evolution of distance d to the global minimizer $(\varphi, \psi) = (0, 0)$ of f_2 for different methods, starting from $(\varphi, \psi) = (\frac{1}{10}, \frac{1}{10})$. The thick dashed and solid line belongs to Riemannian gradient descent and Riemannian BFGS descent (with T_k being parallel transport), respectively. The thin lines show the evolution for Riemannian BFGS descent, where T_k is parallel transport concatenated with a rotation by ω for $\omega = \frac{1}{100}$ (solid), $\omega = \frac{d}{10}$ (dashed), $\omega = \frac{2}{5}\sqrt{d}$ (dotted), and $\omega = \frac{1}{5\log(-\log d)}$ (dot-dashed).*

since $\|T_{*,k+1}^{-1} T_k T_{*,k} - \mathrm{id}\| \sim \sqrt{\mathrm{dist}(x_k, x^*)}$ combined with superlinear convergence of x_k can still satisfy the conditions of Proposition 3.18.) However, for constant ω and $\omega \sim \frac{1}{\log(-\log(\mathrm{dist}(x_k, x^*)))}$, convergence is indeed only linear. Nevertheless, convergence is still much faster than for gradient descent.

4.2. Riemannian optimization in the space of smooth closed curves. Riemannian optimization on the Stiefel manifold has been applied successfully and efficiently to several linear algebra problems, from low rank approximation [7] to eigenvalue computation [18]. The same concepts can be transferred to efficient optimization methods in the space of closed smooth curves, using the shape space and description of curves introduced by Younes et al. [26]. They represent a curve $c : [0, 1] \to \mathbb{C} \equiv \mathbb{R}^2$ by two functions $e, g : [0, 1] \to \mathbb{R}$ via

$$c(\theta) = c(0) + \frac{1}{2} \int_0^\theta (e + ig)^2 \, \mathrm{d}\vartheta \, .$$

The conditions that the curve be closed, $c(1) = c(0)$, and of unit length, $1 = \int_0^1 |c'(\theta)| \, \mathrm{d}\theta$, result in the fact that e and g are orthonormal in $L^2([0, 1])$, thus (e, g) forms an element of the Stiefel manifold

$$\mathbf{St}(2, L^2([0,1])) = \left\{ (e, g) \in L^2([0,1]) \, : \, \|e\|_{L^2([0,1])} = \|g\|_{L^2([0,1])} = 1, (e, g)_{L^2([0,1])} = 0 \right\} .$$

The Riemannian metric of the Stiefel manifold can now be imposed on the the space of smooth closed curves with unit length and fixed base point, which was shown to be equivalent to endowing this shape space with a Sobolev-type metric [26].

For general closed curves we follow Sundaramoorthi et al. [22] and represent a curve c by an element $(c_0, \rho, (e, g))$ of $\mathbb{R}^2 \times \mathbb{R} \times \mathbf{St}(2, L^2([0, 1]))$ via

$$c(\theta) = c_0 + \frac{\exp \rho}{2} \int_0^\theta (e + ig)^2 \, \mathrm{d}\vartheta \, .$$

c_0 describes the curve base point and $\exp \rho$ its length. (Note that Sundaramoorthi et al. choose c_0 as the curve centroid. Choosing the base point instead simplifies the notation a little and yields the same qualitative behavior.) Sundaramoorthi et al.

have shown the corresponding Riemannian metric to be equivalent to the very natural metric

$$g_{[c]}(h,k) = h^t \cdot k^t + \lambda_l h^l k^l + \lambda_d \int_{[c]} \frac{\mathrm{d}h^d}{\mathrm{d}s} \cdot \frac{\mathrm{d}k^d}{\mathrm{d}s}\,\mathrm{d}s$$

on the tangent space of curve variations $h,k : [c] \to \mathbb{R}^2$, where $[c]$ is the image of $c : [0,1] \to \mathbb{R}^2$, s denotes arclength, and $\lambda_l, \lambda_d > 0$. Here, h^t and h^l are the Gâteaux derivatives of the curve centroid (in our case the base point) and the logarithm of the curve length for curve variation in direction h, and $h^d = h^t + h^l(c - c_0)$ (analogous for k). By [10] this yields a geodesically complete shape space in which there is a closed formula for the exponential map [22], lending itself for Riemannian optimization. (Note that simple L^2-type metrics in the space of curves can in general not be used since the resulting spaces usually are degenerate [15]: They exhibit paths of arbitrarily small length between any two curves.)

To illustrate the efficiency of Riemannian optimization in this context we consider the task of image segmentation via active contours without edges as proposed by Chan and Vese [8]. For a given gray scale image $u : [0,1]^2 \to \mathbb{R}$ we would like to minimize the objective functional

$$f([c]) = a_1 \left(\int_{\mathrm{int}[c]} (u_i - u)^2 \,\mathrm{d}x + \int_{\mathrm{ext}[c]} (u_e - u)^2 \,\mathrm{d}x \right) + a_2 \int_{[c]} \mathrm{d}s\,,$$

where $a_1, a_2 > 0$, u_i and u_e are given gray values, and $\mathrm{int}[c]$ and $\mathrm{ext}[c]$ denote the interior and exterior of $[c]$. The first two terms indicate that $[c]$ should enclose the image region where u is close to u_i and far from u_e, while the third term acts as a regularizer and measures the curve length.

We interpret the curve c as an element of the above Riemannian manifold $\mathbb{R}^2 \times \mathbb{R} \times \mathbf{St}(2, L^2([0,1]))$ and add an additional term to the objective functional that prefers a uniform curve parametrization. The objective functional then reads

$$f(c_0,\rho,(e,g)) = a_1 \left(\int_{\mathrm{int}[(c_0,\rho,(e,g))]} (u_i - u)^2 \,\mathrm{d}x + \int_{\mathrm{ext}[(c_0,\rho,(e,g))]} (u_e - u)^2 \,\mathrm{d}x \right) + a_2 \exp(\rho) + a_3 \int_0^1 (e^2+g^2)^2 \,\mathrm{d}\vartheta\,,$$

where we choose $(a_1, a_2, a_3) = (50, 1, 1)$. For numerical implementation, e and g are discretized as piecewise constant functions on an equispaced grid over $[0,1]$, and the image u is given as pixel values on a uniform quadrilateral grid, where we interpolate bilinearly between the nodes.

Figure 4.3 shows the curve evolution for a particular example. Obviously, the natural metric ensures that the correct curve positioning, scaling, and deformation take place quite independently. Corresponding iteration numbers are shown in Table 4.2. In one case we employed geodesic retractions with parallel transport (simple formulae for which are based on the matrix exponential [22]), in the other case we used

$$\begin{aligned} R_{(c_0,\rho,(e,g))}(\delta c_0, \delta\rho, (\delta e, \delta g)) &= (c_0 + \delta c_0, \rho + \delta\rho, \Pi_{\mathbf{St}(2,L^2([0,1]))}(e + \delta e, g + \delta g)\,, \\ T_{(c_0^1,\rho^1,(e^1,g^1)),(c_0^2,\rho^2,(e^2,g^2))}(\delta c_0, \delta\rho, (\delta e, \delta g)) &= (\delta c_0, \delta\rho, \Pi_{T_{(e^2,g^2)}\mathbf{St}(2,L^2([0,1]))}(\delta e, \delta g))\,, \end{aligned}$$

where Π_S denotes the orthogonal projection onto $S \subset (L^2([0,1]))^2$. Due to the closed-form solution of the exponential map there is hardly any difference in computational costs. Riemannian BFGS and nonlinear conjugate gradient iteration yield much faster

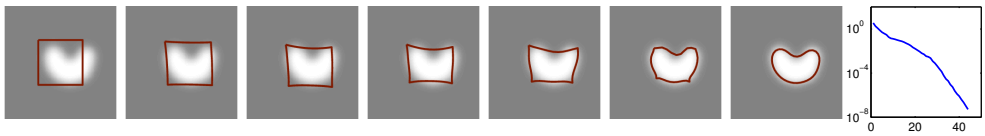

FIG. 4.3. *Curve evolution during BFGS minimization of f. The curve is depicted at steps 0, 1, 2, 3, 4, 7, and after convergence. Additionally we show the evolution of the function value $f(c_k) - \min_c f(c)$.*

	non-geodesic retr.	geodesic retr.
gradient flow	4207	4207
gradient descent	1076	1064
BFGS quasi-Newton	44	45
Fletcher–Reeves NCG	134	220

TABLE 4.2

Iteration numbers for minimization of f with different methods. The iteration is stopped as soon as the derivative of the discretized functional f has ℓ^2-norm less than 10^{-3}. For the gradient flow discretization we employ a stepsize of 0.001, which is roughly the largest stepsize for which the curve stays within the image domain during the whole iteration.

convergence than gradient descent. Gradient descent with step size control in turn is faster than gradient flow (which in numerical implementations corresponds to gradient descent with a fixed small step size), which is the method employed in [22], so that the use of higher order methods such as BFGS or NCG will yield a substantial gain in computation time.

Figure 4.4 shows experiments for different weights inside the metric. We vary λ_d and the ratio $\frac{\lambda_d}{\lambda_l}$ by the factor 16. Obviously, a larger λ_d ensures a good curve positioning and scaling before starting major deformations. Small λ_d has the reverse effect. The ratio between λ_d and $\frac{\lambda_d}{\lambda_l}$ decides whether first the scaling or the positioning is adjusted.

To close this example, Figure 4.5 shows the active contour segmentation on the widely used cameraman image. Since the image contains rather sharp discontinuities, the derivatives of the objective functional exhibit regions of steep variations. Nevertheless, the NCG and BFGS method stay superior to gradient descent: While for the top example in Figure 4.5 the BFGS method needed 46 steps, the NCG iteration needed 2539 and the gradient descent 8325 steps (the iteration was stopped as soon as the derivative of the discretized objective functional reached an ℓ^2-norm less than 10^{-2}). The cameraman example also shows the limitations of the above shape space in the context of segmentation problems. Since we only consider closed curves with a well-defined interior and exterior, we can only segment simply connected regions. As soon as the curve self-intersects, we thus have to stop the optimization (compare Figure 4.5 bottom).

4.3. Riemannian optimization in the space of truss shapes. Kilian et al. introduced a Riemannian shape space in [13], in which shapes are represented by meshes with fixed connectivity. Each mesh consist of a number of vertices connected by thin rods (e. g. the edges in a triangulated surface). The number of nodes as well as their connectivity stay the same throughout the shape space so that each shape can be identified with an n-tuple $\mathcal{S} \in (\mathbb{R}^3)^n$ of node positions in $\mathbb{R}^3$.

Let us denote the set of mesh edges by $\mathcal{E}$, i. e., for two node indices $p, q \in \mathbb{N}$, let

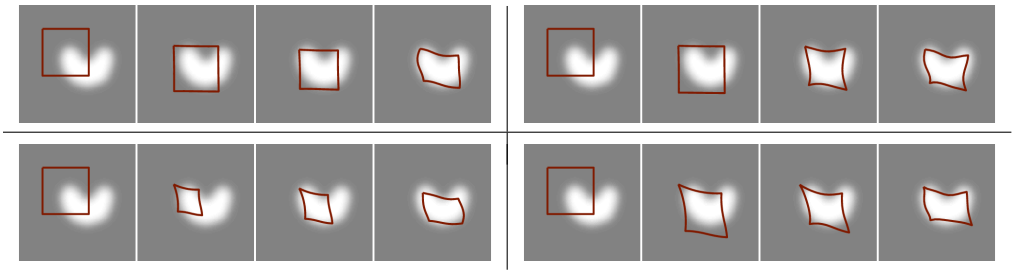

FIG. 4.4. *Steps 0, 1, 3, 5 of the BFGS optimization, using different weights in the shape space metric (top row: $\lambda_d = 16$; bottom row: $\lambda_d = 1$; left column: $\frac{\lambda_d}{\lambda_l} = 16$; right column: $\frac{\lambda_d}{\lambda_l} = 1$).*

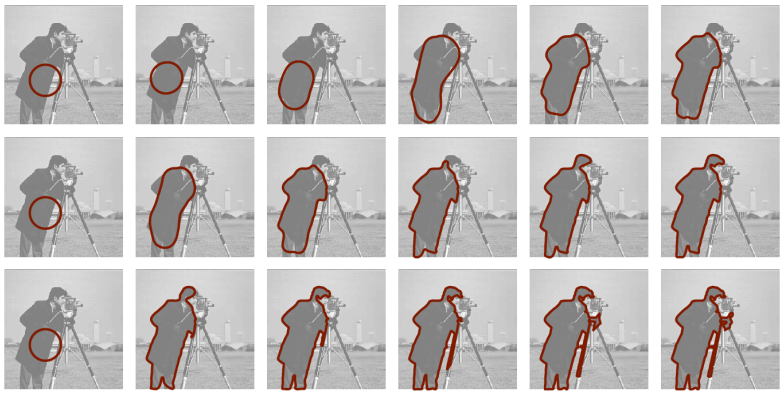

FIG. 4.5. *Segmentation of the cameraman image with different parameters (using the BFGS iteration and $\lambda_l = \lambda_d = 1$). Top: $(a_1, a_2, a_3) = (50, 3 \cdot 10^{-1}, 10^{-3})$, steps 0, 1, 5, 10, 20, 46 are shown. Middle: $(a_1, a_2, a_3) = (50, 8 \cdot 10^{-2}, 10^{-3})$, steps 0, 10, 20, 40, 60, 116 are shown. Bottom: $(a_1, a_2, a_3) = (50, 10^{-2}, 10^{-3})$, steps 0, 50, 100, 150, 200, 250 are shown. The curves were reparameterized every 70 steps. The bottom iteration was stopped as soon as the curve self-intersected.*

$(p, q) \in \mathcal{E}$ express the fact that the corresponding nodes are connected by a rod. Their positions in a shape $\mathcal{S}$ will be denoted $\mathcal{S}_p$ and $\mathcal{S}_q$. As already mentioned, $(\mathbb{R}^3)^n$ is interpreted as the manifold $\mathcal{M}$ of shapes so that the tangent space to $\mathcal{M}$ at a shape $\mathcal{S}$ is given by $T_{\mathcal{S}}\mathcal{M} = (\mathbb{R}^3)^n$, the vector space of all node position variations. For $\mathcal{S} \in \mathcal{M}$, Kilian et al. introduce the Riemannian metric

$$g_{\mathcal{S}}(\cdot,\cdot) : T_{\mathcal{S}}\mathcal{M} \times T_{\mathcal{S}}\mathcal{M} \to \mathbb{R}, \quad g_{\mathcal{S}}(v, w) = v \cdot w + \beta \sum_{(p,q)\in\mathcal{E}} \left((v_p - v_q) \cdot (\mathcal{S}_p - \mathcal{S}_q)\right) \left((w_p - w_q) \cdot (\mathcal{S}_p - \mathcal{S}_q)\right) ,$$

where the dot denotes the Euclidean inner product and v_p, v_q, w_p, w_q are the displacements of nodes p and q, respectively. The weight $\beta > 0$ specifies the penalization of isometry violations: For a curve $t \mapsto \mathcal{S}(t) \in \mathcal{M}$, $(\dot{\mathcal{S}}_p - \dot{\mathcal{S}}_q) \cdot (\mathcal{S}_p - \mathcal{S}_q)$ is the rate at which the squared distance between node p and q changes. Hence, the above sum over $\mathcal{E}$ vanishes for v or w being the velocity of a curve in $\mathcal{M}$ along which the shapes stay isometric to each other (i. e. all edges keep their initial length).

Given an objective function f, one can choose whether to neglect the Riemannian structure and perform a standard optimization in Euclidean space $(\mathbb{R}^3)^n$ or whether to perform a truly Riemannian optimization on $\mathcal{M}$ as proposed in this article. Of course, as discussed in Section 4.1, using the Riemannian structure of $\mathcal{M}$ does only

make sense if it is sufficiently compatible with the objective function. Here, let us consider the following exemplary energy: Assume a meshed cuboid to be given such as the top left shape in Figure 4.6, and assume its rods to be elastic. If both ends are rotated relative to each other, all rods deform, which costs elastic energy. As objective functional, we choose this elastic energy plus a potential that causes the twisting,

$$f(\mathcal{S}) = \sum_{(p,q)\in\mathcal{E}} (|\mathcal{S}_p - \mathcal{S}_q| - |\mathcal{S}_p^0 - \mathcal{S}_q^0|)^2 + \alpha \sum_{p\in\mathcal{N}} (\mathcal{S}_{p,23} - (x_2^p, x_3^p))^2 \,,$$

where $\mathcal{S}^0$ is the initial cuboid, $\mathcal{N}$ is the set of nodes at both cuboid ends, $\mathcal{S}_{p,23}$ denotes the second and third coordinate of the point $\mathcal{S}_p$, and x_i^p are prescribed coordinate values. In our experiments we employ $\alpha = 10^4$. This energy models the situation that for a given shape we seek a deformed version which satisfies certain constraints but is as isometric to the original as possible.

Note that the above energy is not necessarily suited for optimization in a Riemannian metric. It is an elastic energy which compares each shape $\mathcal{S}$ with a reference configuration $\mathcal{S}^0$, independent of any path in $\mathcal{M}$ connecting $\mathcal{S}$ with $\mathcal{S}^0$. Nevertheless, the optimization benefits considerably from exploiting the above Riemannian metric, as can be seen in Figure 4.6. The Euclidean gradient descent in $(\mathbb{R}^3)^n$ completely ignores that we are looking for a near-isometric deformation of the initial shape, while the Riemannian gradient descent with geodesic retractions produces only near-isometric shapes right from the start. The energy decrease of the purely Euclidean method is so slow that it definitely pays off to employ the Riemannian method, despite the additional costs for computing a geodesic in each step (without code optimization, each step of the Riemannian method takes about 60 times as long as a Euclidean step). The runtime is considerably improved if we use a Riemannian gradient descent with non-geodesic retractions $R_{\mathcal{S}}(\delta\mathcal{S}) = \mathcal{S} + \delta\mathcal{S}$, however, the intermediate shapes are visually less appealing (since infinitesimal rotations are extended linearly in each step, producing an upscaling of shape parts which can hardly be remedied based on gradients with respect to the isometry-enforcing metric).

Without additional cost we can perform a geodesic nonlinear conjugate gradient iteration, which strongly outperforms the gradient descent. (From Figure 4.6 one can also see that it outperforms the non-Riemannian NCG- or BFGS-method in terms of iterations, though one has to admit that with non-optimized code, each step of the Riemannian method takes roughly 60 times longer due to the higher computational costs per step.) However, the use of a BFGS quasi-Newton method is quite restricted due to the additional costs of computing vector transports. Thus, for experiments with the BFGS method let us restrict to a two-dimensional example. This time, the potential part in the objective energy shall induce a bending deformation (Figure 4.7). Table 4.3 shows iteration numbers for different shape discretizations. Newton's method seems rather independent of the shape resolution. Likewise, if the Hessian approximation in the BFGS method is initialized with the true Hessian, the iteration numbers of the geodesic BFGS method stay also roughly constant. However, if the Hessian approximation is initialized with the Riemannian metric, then the iteration numbers increase with refined discretizations. This is related to the condition on the initial Hessian approximation $\hat{B}_0$ for superlinear convergence (Proposition 3.18). While the norm of $\nabla^2 f(x^*) - T_{*,0}^* \hat{B}_0 T_{*,0}$ stays roughly the same for different discretizations, its nuclear norm increases with the number of degrees of freedom.

5. Conclusion. We analysed the convergence of the BFGS quasi-Newton method and the Fletcher–Reeves conjugate gradient iteration on Riemannian manifolds. These

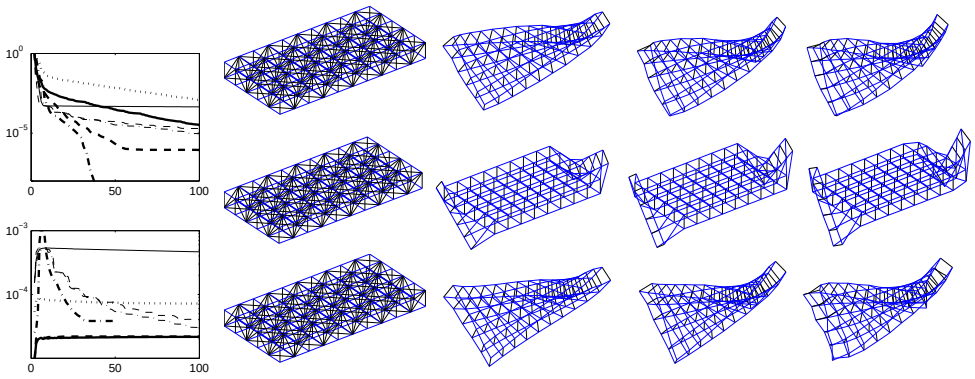

FIG. 4.6. *Left: Objective function f (top) and elastic energy (bottom) during minimization of f via geodesic gradient descent on $\mathcal{M}$ (thick solid), Euclidean gradient descent in $(\mathbb{R}^3)^n$ (solid), gradient descent on $\mathcal{M}$ with non-geodesic retractions $R_{\mathcal{S}}(\delta\mathcal{S}) = \mathcal{S} + \delta\mathcal{S}$ (dotted), geodesic and non-Riemannian NCG (thick and thin dashed), and geodesic and non-Riemannian BFGS iteration (thick and thin dash-dotted). In the top graph, the energy values have been shifted by $-\min f$. In the bottom graph, the curve for the NCG method lies on top of the curve for geodesic gradient descent. Note that the BFGS method ends up in a local minimum and thus does not achieve the least elastic energy. Right: First to third and 200th shape of the geodesic (top), the Euclidean (middle), and the non-geodesic gradient descent (bottom). Only for the first shape all involved edges are displayed.*

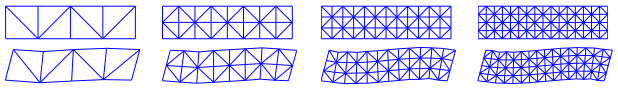

FIG. 4.7. *Initial and final shape of the optimization for different discretization resolutions.*

methods compute a search direction based on the search directions at past iterates. Since these previous search directions lie in different tangent spaces, the methods require a vector transport between the different tangent spaces. The BFGS method converges superlinearly, if this vector transport is isometric and the transport along a triangle is sufficiently close to the identity (more precisely, near the optimum, the vector transport between two consecutive iterates must be sufficiently approximable by a vector transport from one iterate to the optimum and back from the optimum to the other iterate). On compact manifolds, the isometry condition can be dropped, but it is unclear whether it might also be unnecessary on non-compact manifolds.

The Riemannian optimization methods were applied to example problems in the context of Riemannian shape spaces, where they proved to be appropriate optimization tools. To the authors' knowledge, such applications have not been considered before, and it seems that there is some potential to be exploited in this area by using Riemannian linesearch algorithms instead of for instance the commonly used gradient flow. Of course, there is still much room to improve the computational efficiency of the algorithms even in the rather simple shape spaces considered here. In particular, arising questions are how to most efficiently approximate geodesics and parallel transport or how to devise rapidly computable retractions which fit to the given optimization problems. Furthermore, it will be interesting to test the methods on a greater variety of (perhaps more elaborate and diverse) shape spaces.

Appendix A. Proofs of BFGS convergence and convergence rate.

Proof. [Proof of Proposition 3.15] The eigenvalue estimates of Corollary 3.4 are

node number	Newton descent	BFGS with Hessian	BFGS without Hessian
10	3	4	57
27	3	5	70
52	7	6	102
85	7	6	118

TABLE 4.3

Steps until convergence for the different discretizations from Figure 4.7, using Newton's method, the BFGS iteration with B_0 being the true Hessian, and the BFGS iteration with B_0 being the metric. The last method was not started from the initial shape depicted in Figure 4.7, but rather from a shape close to the final one. Iterations were stopped as soon as the ℓ^2-norm of ∇f decreased below a fixed threshold. To make the iteration numbers comparable, the terms inside the objective function were rescaled so as to be independent of the discretization.

replaced by estimates on the trace and determinant of the $\hat{B}_k$ (the Lax–Milgram representations of the B_k). Define the angle θ_k, the Rayleigh quotient q_k, and an averaged Hessian G_k as

$$\cos\theta_k = \frac{B_k(s_k,s_k)}{\|s_k\|_{x_k}\,\|B_k(s_k,\cdot)\|_{x_k}}\,,\quad q_k = \frac{B_k(s_k,s_k)}{\|s_k\|_{x_k}^2}\,,\quad G_k = \int_0^1 \mathrm{D}^2 f_{R_{x_k}}(ts_k)\,\mathrm{d}t\,.$$

Obviously, $G_k(s_k,\cdot) = y_k$ and $m\,\|v\|_{x_k}^2 \le G_k(v,v) \le M\,\|v\|_{x_k}^2$ for all $v \in T_{x_k}\mathcal{M}$. Let $\hat{G}_k$ be the Lax–Milgram representation of G_k, then

$$\frac{\|y_k\|_{x_k}^2}{y_k s_k} = \frac{G_k(\sqrt{\hat{G}_k}s_k, \sqrt{\hat{G}_k}s_k)}{\left\|\sqrt{\hat{G}_k}s_k\right\|_{x_k}^2} \le M\,,\quad \frac{y_k s_k}{\|s_k\|_{x_k}^2} = \frac{G_k(s_k,s_k)}{\|s_k\|_{x_k}^2} \ge m\,.$$

Furthermore, let $\hat{y}_k$ be the Riesz representation of y_k and define $A_{k+1} = T_k\cdots T_0\sqrt{\hat{B}_0^{-1}}T_0^*\cdots T_k^*$, then $A_0\hat{B}_0A_0 = \mathrm{id}$ and thus, by induction, $A_k\hat{B}_kA_k - \mathrm{id}$ is of trace class for all k. In particular, $\mathrm{tr}(A_k\hat{B}_kA_k - \mathrm{id})$ and $\det(A_k\hat{B}_kA_k)$ are well-defined with

$$\mathrm{tr}(A_{k+1}\hat{B}_{k+1}A_{k+1} - \mathrm{id}) = \mathrm{tr}\left(A_k\left[\hat{B}_k - \frac{B_k(s_k,\cdot)\hat{B}_ks_k}{B_k(s_k,s_k)} + \frac{y_k(\cdot)\hat{y}_k}{y_ks_k}\right]A_k - \mathrm{id}\right)$$

$$= \mathrm{tr}(A_k\hat{B}_kA_k - \mathrm{id}) - \frac{\left\|A_k\hat{B}_ks_k\right\|_{x_k}^2}{B_k(s_k,s_k)} + \frac{\|A_k\hat{y}_k\|_{x_k}^2}{y_ks_k} \le \mathrm{tr}(A_k\hat{B}_kA_k - \mathrm{id}) - \frac{q_k}{\|\hat{B}_0\|\cos^2\theta_k} + \left\|\hat{B}_0^{-1}\right\| M\,,$$

$$\det(A_{k+1}\hat{B}_{k+1}A_{k+1}) = \det\left(A_k\left[\hat{B}_k - \frac{B_k(s_k,\cdot)\hat{B}_ks_k}{B_k(s_k,s_k)} + \frac{y_k(\cdot)\hat{y}_k}{y_ks_k}\right]A_k\right)$$

$$= \det\left(A_k\sqrt{\hat{B}_k}[I - \mathfrak{s}\otimes\mathfrak{s} + \mathfrak{y}\otimes\mathfrak{y}]\sqrt{\hat{B}_k}A_k\right) = \det(A_k\hat{B}_kA_k)\det(I - \mathfrak{s}\otimes\mathfrak{s} + \mathfrak{y}\otimes\mathfrak{y})$$

$$= \det(A_k\hat{B}_kA_k)(1+\lambda_1)(1+\lambda_2) = \det(A_k\hat{B}_kA_k)\frac{y_ks_k}{B_k(s_k,s_k)} = \det(A_k\hat{B}_kA_k)\frac{y_ks_k}{q_k\,\|s_k\|_{x_k}^2} \ge \frac{m}{q_k}\det(A_k\hat{B}_kA_k)\,.$$

where we have used the abbreviations $\mathfrak{s} = \sqrt{\hat{B}_k}s_k/\sqrt{B_k(s_k,s_k)}$, $\mathfrak{y} = \hat{B}_k^{-\frac{1}{2}}y_k\sqrt{y_ks_k}$, $\lambda_{1,2} = g_{x_k}(\mathfrak{s},\mathfrak{s}) + g_{x_k}(\mathfrak{y},\mathfrak{y}) \pm \sqrt{(g_{x_k}(\mathfrak{s},\mathfrak{s}) + g_{x_k}(\mathfrak{y},\mathfrak{y}))^2 - g_{x_k}(\mathfrak{s},\mathfrak{y})^2}$ (the two nonzero

eigenvalues of $-\mathfrak{s}\otimes\mathfrak{s}+\mathfrak{y}\otimes\mathfrak{y}$) as well as the fact that isometries leave the determinant and trace unchanged, the definition $\det(I+B)=\Pi_i(1+\lambda_i(B))$ for trace class operators B [20, Chp. 3] (where $\lambda_i(B)$ are the corresponding eigenvalues), and the product rule for determinants [20, Thm. 3.5]. Next, define $\Psi(\hat{B})=\operatorname{tr}(\hat{B}-\mathrm{id})-\log\det\hat{B}$, which is non-negative for positive definite $\hat{B}$ by Lidskii's theorem [20, Thm. 3.7]. By induction,

$$0\le\Psi(A_{k+1}\hat{B}_{k+1}A_{k+1})\le\Psi(A_0\hat{B}_0A_0)$$
$$+\sum_{i=0}^{k}\left((\left\|\hat{B}_0^{-1}\right\|M-\log m-1)+\log(\|\hat{B}_0\|\cos^2\theta_i)+\left[1-\frac{q_i}{\|\hat{B}_0\|\cos^2\theta_i}+\log\frac{q_i}{\|\hat{B}_0\|\cos^2\theta_i}\right]\right).$$

We now show that for any $0<r<1$ there are constants $\kappa,\rho,\sigma>0$ such that for any $k\in\mathbb{N}$ there are at least $\lfloor r(k+1)\rfloor$ indices i in $\{0,\dots,k\}$ with $\cos\theta_i\ge\kappa$ and $\rho\le\frac{q_i}{\cos\theta_i}\le\sigma$. Let us abbreviate

$$\eta_i=-\log(\|\hat{B}_0\|\cos^2\theta_i)-l_i=-\log(\|\hat{B}_0\|\cos^2\theta_i)-\left[1-\frac{q_i}{\|\hat{B}_0\|\cos^2\theta_i}+\log\frac{q_i}{\|\hat{B}_0\|\cos^2\theta_i}\right]$$

and denote the $\lfloor r(k+1)\rfloor$th smallest among $\eta_0,\dots,\eta_k$ by η^k. Now let $\mathcal{I}=\{i\in\{0,\dots,k\}:\eta_i>\eta^k\}$. For all $\lfloor r(k+1)\rfloor$ indices $i\in\{0,\dots,k\}\setminus\mathcal{I}$ we have

$$\eta_i\le\eta^k\le\frac{1}{|\mathcal{I}|}\sum_{j\in\mathcal{I}}\eta_j\le\frac{1}{1-r}\frac{1}{k+1}\sum_{j=0}^{k}\eta_j\le\frac{1}{1-r}\left(\frac{\Psi(\mathrm{id})}{k+1}+(\left\|\hat{B}_0^{-1}\right\|M-\log m-1)\right)\le c$$

for some constant c, where the second last inequality comes from the estimate for $\Psi(A_{k+1}\hat{B}_{k+1}A_{k+1})$ above. Hence, $\log(\|\hat{B}_0\|\cos^2\theta_i)$ and l_i must be bounded from below for these indices i, from which we first obtain $\sqrt{\|\hat{B}_0\|}\ge\sqrt{\|\hat{B}_0\|}\cos\theta_i\ge\exp\frac{-c}{2}=:\sqrt{\|\hat{B}_0\|}\kappa$ and then also boundedness of $\frac{q_i}{\cos\theta_i}$ above and below by some σ and ρ (due to the concavity and continuity of $g(x)=1-x+\log x$ with $g(x)\to-\infty$ for $x\to0$ and $x\to\infty$). Furthermore, for the same indices i, (3.1b) implies $-c_2\mathrm{D}f(x_i)p_i\ge-\mathrm{D}f_{R_{x_i}}(\alpha_ip_i)p_i=-\mathrm{D}f(x_i)p_i-\alpha_i\int_0^1\mathrm{D}^2f_{R_{x_i}}(t\alpha_ip_i)(p_i,p_i)\,\mathrm{d}t\ge-\mathrm{D}f(x_i)p_i-\alpha_iM\|p_i\|_{x_i}^2$ so that

$$\alpha_i\ge-\frac{\mathrm{D}f(x_i)p_i}{\|p_i\|_{x_i}^2}\frac{1-c_2}{M}=\frac{1-c_2}{M}q_i\ge\frac{1-c_2}{M}\kappa\rho\,.$$

Equation (3.1a) then implies

$$f(x_i)-f(x_{i+1})\ge-\alpha_ic_1\mathrm{D}f(x_i)p_i\ge\frac{1-c_2}{M}\kappa\rho c_1\frac{\cos^2\theta_i}{q_i}\|\mathrm{D}f(x_i)\|_{x_i}^2\ge\frac{1-c_2}{M}\kappa\rho c_1\frac{\kappa}{\sigma}2m(f(x_i)-f(x^*))\,,$$

where in the last step we used $f(x)-f(x^*)\le\frac{1}{2m}\|\mathrm{D}f(x)\|_x^2$, which follows from separately minimizing for $\hat{x}\in\mathcal{M}$ the right-hand and the left-hand side of the inequality

$$f(\hat{x})-f(x)=\mathrm{D}f(x)R_x^{-1}(\hat{x})+\frac{1}{2}\mathrm{D}^2f_{R_x}(tR_x^{-1}(\hat{x}))(R_x^{-1}(\hat{x}),R_x^{-1}(\hat{x}))\ge\mathrm{D}f(x)R_x^{-1}(\hat{x})+\frac{m}{2}\left\|R_x^{-1}(\hat{x})\right\|_x^2$$

with some $t\in[0,1]$. Due to the monotonicity of the $f(x_k)$ and the above estimates for the $\lfloor r(k+1)\rfloor$ indices i with $\eta_i\le\eta^k$, we arrive at

$$f(x_k)-f(x^*)\le\left(1-\frac{1-c_2}{M}c_1\frac{\kappa^2\rho}{\sigma}2m\right)^{\lfloor r(k+1)\rfloor}(f(x_0)-f(x^*))\quad\forall k\in\mathbb{N}\,.$$

(From the non-negativity of $f(x_k)-f(x^*)$ and $f(x_0)-f(x^*)$ we additionally see that the bounds σ and ρ satisfy $\frac{\sigma}{\rho} \geq \frac{1-c_2}{M} c_1 \kappa^2 2m$.) □

Proof. [Proof of Proposition 3.18] Let us define $G = \nabla^2 f(x^*)$ as well as

$$\tilde{s}_k = G^{\frac{1}{2}} T_{*,k}^{-1} s_k \,, \quad \tilde{y}_k = G^{-\frac{1}{2}} T_{*,k}^* \hat{y}_k \,, \quad \tilde{B}_k = G^{-\frac{1}{2}} T_{*,k}^* \hat{B}_k T_{*,k} G^{-\frac{1}{2}} \,, \quad T_{G,k} = G^{\frac{1}{2}} T_{*,k}^{-1} T_k^{-1} T_{*,k+1} G^{-\frac{1}{2}} \,,$$

where (as before) $\hat{y}_k$ and $\hat{B}_k$ are the Riesz and Lax–Milgram representation of y_k and B_k. Note

$$\tilde{B}_{k+1} = T_{G,k}^* \left[\tilde{B}_k - \frac{g_{x^*}(\tilde{B}_k \tilde{s}_k, \cdot) \tilde{B}_k \tilde{s}_k}{g_{x^*}(\tilde{B}_k \tilde{s}_k, \tilde{s}_k)} + \frac{g_{x^*}(\tilde{y}_k, \cdot) \tilde{y}_k}{g_{x^*}(\tilde{y}_k, \tilde{s}_k)} \right] T_{G,k}$$

as well as $\mathrm{tr}(T_{G,k}^* B T_{G,k}) = \mathrm{tr} B + \mathrm{tr}((T_{G,k}^* - \mathrm{id})B) + \mathrm{tr}(B(T_{G,k} - \mathrm{id})) + \mathrm{tr}((T_{G,k}^* - \mathrm{id})B(T_{G,k} - \mathrm{id}))$ for any trace class operator $B : T_{x^*}\mathcal{M} \to T_{x^*}\mathcal{M}$ so that

$$\mathrm{tr}(T_{G,k}^* B T_{G,k}) \leq \mathrm{tr} B + \|B\|_1 \left(2 \left\| G^{\frac{1}{2}} \right\| \left\| G^{-\frac{1}{2}} \right\| \delta_k + \|G\| \left\| G^{-1} \right\| \delta_k^2 \right) \leq \mathrm{tr} B + C \|B\|_1 \delta_k \tag{A.1}$$

with $\delta_k = \left\| T_{*,k}^{-1} T_k^{-1} T_{*,k+1} - \mathrm{id} \right\|$, where C is chosen large enough so that the above holds for all k. In the above we have used that trace and nuclear norm can be expressed as $\mathrm{tr} B = \sum_i g_{x^*}(e_i, Be_i)$ and $\|B\|_1 = \sum_i |g_{x^*}(e_i, Be_i)|$ for any orthonormal basis $(e_i)_i$ of $T_{x^*}\mathcal{M}$ so that $\mathrm{tr} B \leq \|B\|_1$ and $\|AB\|_1 = \sum_i |g_{x^*}(e_i, ABe_i)| \leq \|A\| \sum_i |g_{x^*}(e_i, Be_i)| = \|A\| \|B\|_1$ for any bounded operator $A : T_{x^*}\mathcal{M} \to T_{x^*}\mathcal{M}$. Furthermore, let $\hat{G}_k$ be the Lax–Milgram representation of G_k from the proof of Proposition 3.15, then

$$\|\tilde{y}_k - \tilde{s}_k\|_{x^*} = \left\| G^{-\frac{1}{2}} (T_{*,k}^* \hat{G}_k T_{*,k} - G) G^{-\frac{1}{2}} \tilde{s}_k \right\|_{x^*} \leq c \epsilon_k \|\tilde{s}_k\|_{x^*}$$

with $c = \left\| G^{-1} \right\|$ and $\epsilon_k = \left\| T_{*,k}^* \hat{G}_k T_{*,k} - G \right\|$. We deduce

$$\tilde{m}_k := \frac{g_{x^*}(\tilde{y}_k, \tilde{s}_k)}{\|\tilde{s}_k\|_{x^*}^2} \geq 1 - c\epsilon_k \,, \quad \tilde{M}_k := \frac{\|\tilde{y}_k\|_{x^*}^2}{g_{x^*}(\tilde{y}_k, \tilde{s}_k)} \leq 1 + c\epsilon_k$$

for k large enough, which follows from $\|\tilde{y}_k\|_{x^*}^2 - 2 g_{x^*}(\tilde{y}_k, \tilde{s}_k) + \|\tilde{s}_k\|_{x^*}^2 = \|\tilde{y}_k - \tilde{s}_k\|_{x^*}^2 \leq c^2 \epsilon_k^2 \|\tilde{s}_k\|_{x^*}^2$ as well as $|\|\tilde{y}_k\|_{x^*} - \|\tilde{s}_k\|_{x^*}| \leq \|\tilde{y}_k - \tilde{s}_k\|_{x^*} \leq c\epsilon_k \|\tilde{s}_k\|_{x^*}$ so that $(1 - c\epsilon_k) \|\tilde{s}_k\|_{x^*} \leq \|\tilde{y}_k\|_{x^*} \leq (1 + c\epsilon_k) \|\tilde{s}_k\|_{x^*}$.

Next, let us introduce

$$\tilde{q}_k = \frac{g_{x^*}(\tilde{s}_k, \tilde{B}_k \tilde{s}_k)}{\|\tilde{s}_k\|_{x^*}^2} \,, \quad \cos\tilde{\theta}_k = \frac{g_{x^*}(\tilde{s}_k, \tilde{B}_k \tilde{s}_k)}{\|\tilde{s}_k\|_{x^*} \left\| \tilde{B}_k \tilde{s}_k \right\|_{x^*}} \,.$$

Using $\left\| T_{G,k}^* B T_{G,k} \right\|_1 = \left\| B + (T_{G,k}^* - \mathrm{id})B + B(T_{G,k} - \mathrm{id}) + (T_{G,k}^* - \mathrm{id})B(T_{G,k} - \mathrm{id}) \right\|_1 \leq \|B\|_1 + \left\| (T_{G,k}^* - \mathrm{id})B \right\|_1 + \|B(T_{G,k} - \mathrm{id})\|_1 + \left\| (T_{G,k}^* - \mathrm{id})B(T_{G,k} - \mathrm{id}) \right\|_1 \leq \|B\|_1 + \|B\|_1 \left(2 \left\| G^{\frac{1}{2}} \right\| \left\| G^{-\frac{1}{2}} \right\| \delta_k + \|G\| \left\| G^{-1} \right\| \delta_k^2 \right) \leq \|B\|_1 + C \|B\|_1 \delta_k$ for any trace class operator B, the nuclear norm of $\tilde{B}_{k+1} - \mathrm{id}$ can

be bounded as follows,

$$\begin{aligned}\left\|\tilde{B}_{k+1}-\mathrm{id}\right\|_1 &= \left\|T_{G,k}^*\left[\tilde{B}_k-\mathrm{id}-\frac{g_{x^*}(\tilde{B}_k\tilde{s}_k,\cdot)\tilde{B}_k\tilde{s}_k}{g_{x^*}(\tilde{B}_k\tilde{s}_k,\tilde{s}_k)}+\frac{g_{x^*}(\tilde{y}_k,\cdot)\tilde{y}_k}{g_{x^*}(\tilde{y}_k,\tilde{s}_k)}\right]T_{G,k}+T_{G,k}^*T_{G,k}-\mathrm{id}\right\|_1\\ &\le \left\|T_{G,k}^*T_{G,k}-\mathrm{id}\right\|_1+(1+C\delta_k)\left(\left\|\tilde{B}_k-\mathrm{id}\right\|_1+\frac{\tilde{q}_k}{\cos^2\tilde{\theta}_k}+\tilde{M}_k\right)\\ &\le \left\|\tilde{B}_0-\mathrm{id}\right\|_1\prod_{i=0}^{k}(1+C\delta_i)+\sum_{j=0}^{k}\left[\left\|T_{G,j}^*T_{G,j}-\mathrm{id}\right\|_1+(1+C\delta_j)\left(\frac{\tilde{q}_j}{\cos^2\tilde{\theta}_j}+\tilde{M}_j\right)\right]\prod_{i=j+1}^{k}(1+C\delta_i)\,.\end{aligned}$$

Now use $\prod_{i=j}^{k}(1+C\delta_i)\le\prod_{i=0}^{\infty}(1+C\delta_i)=\exp(\sum_{i=0}^{\infty}\log(1+C\delta_i))\le\exp(C\sum_{i=0}^{\infty}\delta_i)<\infty$ as well as $\left\|T_{G,k}^*T_{G,k}-\mathrm{id}\right\|_1=\left\|G^{-\frac12}(T^*GT-G)G^{-\frac12}\right\|_1\le\left\|G^{-1}\right\|\,\|(T-\mathrm{id})^*G(T-\mathrm{id})+G(T-\mathrm{id})+(T-\mathrm{id})^*G\|_1$ $\le\|G\|\,\left\|G^{-1}\right\|\,\|T\|\,(2+b\beta^k)b\beta^k$ for $T\equiv T_{*,k}^{-1}T_k^{-1}T_{*,k+1}$ to obtain

$$\left\|\tilde{B}_{k+1}-\mathrm{id}\right\|_1\le\zeta\left(1+\sum_{j=0}^{k}\left(\frac{\tilde{q}_j}{\cos^2\tilde{\theta}_j}+\tilde{M}_j\right)\right)$$

for some $\zeta>0$. We now bound the trace and determinant of $\tilde{B}_{k+1}$, similarly to the proof of Proposition 3.15. By (A.1) and induction,

$$\begin{aligned}\mathrm{tr}(\tilde{B}_{k+1}-\mathrm{id}) &= \mathrm{tr}\left[T_{G,k}^*T_{G,k}-\mathrm{id}+T_{G,k}^*\left[\tilde{B}_k-\mathrm{id}-\frac{g_{x^*}(\tilde{B}_k\tilde{s}_k,\cdot)\tilde{B}_k\tilde{s}_k}{g_{x^*}(\tilde{B}_k\tilde{s}_k,\tilde{s}_k)}+\frac{g_{x^*}(\tilde{y}_k,\cdot)\tilde{y}_k}{g_{x^*}(\tilde{y}_k,\tilde{s}_k)}\right]T_{G,k}\right]\\ &\le \mathrm{tr}(T_{G,k}^*T_{G,k}-\mathrm{id})+\left(\mathrm{tr}(\tilde{B}_k-\mathrm{id})-\frac{\tilde{q}_k}{\cos^2\tilde{\theta}_k}+\tilde{M}_k\right)+C\delta_k\left(\left\|\tilde{B}_k-\mathrm{id}\right\|_1+\frac{\tilde{q}_k}{\cos^2\tilde{\theta}_k}+\tilde{M}_k\right)\\ &\le \mathrm{tr}(\tilde{B}_0-\mathrm{id})+\sum_{j=0}^{k}\left[\left\|T_{G,j}^*T_{G,j}-\mathrm{id}\right\|_1-\frac{\tilde{q}_j}{\cos^2\tilde{\theta}_j}(1-C\delta_j)+\tilde{M}_j(1+C\delta_j)+C\delta_j\zeta\left(1+\sum_{i=0}^{j-1}\left(\frac{\tilde{q}_i}{\cos^2\tilde{\theta}_i}+\tilde{M}_i\right)\right)\right]\\ &\le \eta+\sum_{j=0}^{k}\left[-\frac{\tilde{q}_j}{\cos^2\tilde{\theta}_j}\left(1-C\delta_j-C\zeta\sum_{i=j+1}^{k}\delta_i\right)+\tilde{M}_j\left(1+C\delta_j+C\zeta\sum_{i=j+1}^{k}\delta_i\right)\right]\,,\end{aligned}$$

$$\log\det\tilde{B}_{k+1}=\log\det(T_{G,k}^*T_{G,k})+\log\left[\det\tilde{B}_k\frac{\tilde{m}_k}{\tilde{q}_k}\right]=\log\det\tilde{B}_0+\sum_{j=0}^{k}\left(\log\det(T_{G,j}^*T_{G,j})+\log\frac{\tilde{m}_j}{\tilde{q}_j}\right)\,,$$

where η is a constant. Thus, for $\Psi(\tilde{B}_{k+1})=\mathrm{tr}(\tilde{B}_{k+1}-\mathrm{id})-\log\det\tilde{B}_{k+1}\ge0$ we

obtain

$$
\begin{aligned}
0 \leq \Psi(\tilde{B}_{k+1}) &\leq \eta - \log\det \tilde{B}_0 + \sum_{j=0}^{k} \left(\log\cos^2\tilde{\theta}_j - \log\det(T_{G,j}^* T_{G,j}) \right) \\
&\quad + \sum_{j=0}^{k} \left[\left(1 - \frac{\tilde{q}_j}{\cos^2\tilde{\theta}_j} \right) \left(1 - C\zeta \sum_{i=j}^{k} \delta_i \right) + \log \frac{\tilde{q}_j}{\cos^2\tilde{\theta}_j} \right] \\
&\quad + \sum_{j=0}^{k} \left((\tilde{M}_j - 1) \left(1 + C\zeta \sum_{i=j}^{k} \delta_i \right) + 2C\zeta \sum_{i=j}^{k} \delta_i - \log \tilde{m}_j \right) \\
&\leq \vartheta + \sum_{j=K}^{k} \left(c\epsilon_j \lambda + 2C\zeta \sup_{i\geq j} \left\| T_{*,i}^{-1} T_i^{-1} T_{*,i+1} \right\| b \frac{\beta^j}{1-\beta} + 2c\epsilon_j + \left\| T_{G,j}^{-1} T_{G,j}^{-*} \right\| \left\| T_{G,j}^* T_{G,j} - \mathrm{id} \right\|_1 \right) \\
&\quad + \sum_{j=K}^{k} \left[\left(1 - \frac{\tilde{q}_j}{\cos^2\tilde{\theta}_j} \right) \left(1 - C\zeta \sum_{i=j}^{k} \delta_i \right) + \log \frac{\tilde{q}_j}{\cos^2\tilde{\theta}_j} \right] + \sum_{j=K}^{k} \log\cos^2\tilde{\theta}_j \,,
\end{aligned}
$$

where $\vartheta > 0$ is some constant (depending on $K \in \mathbb{N}$), $\sum_{i=j}^{k} \delta_i \leq \sum_{i=j}^{k} \left\| T_{*,i}^{-1} T_i^{-1} T_{*,i+1} \right\| b\beta^i \leq b \sup_{i\geq j} \left\| T_{*,i}^{-1} T_i^{-1} T_{*,i+1} \right\| \frac{\beta^j}{1-\beta}$, $\lambda = 1 + C\zeta \sum_{i=0}^{\infty} \delta_i$, and we have used the above estimates on $\tilde{m}_j$ and $\tilde{M}_j$, $\det T = (\det T^{-1})^{-1} \geq \exp(-\left\| T^{-1} - \mathrm{id} \right\|_1) \geq \exp(-\left\| T^{-1} \right\| \left\| T - \mathrm{id} \right\|_1)$ for $T \equiv T_{G,j}^* T_{G,j}$ [20, Lem. 3.3], as well as $-2a \leq \log(1-a)$ for small $a > 0$. The first row of the right-hand side is uniformly bounded above for all k by some constant γ due to $\sum_{j=0}^{\infty} \epsilon_j < \infty$, $\sum_{j=0}^{\infty} \beta^j < \infty$, and $\left\| T_{G,j}^{-1} T_{G,j}^{-*} \right\| \leq \left\| T_{G,j}^{-1} \right\|^2 \leq \|G\| \left\| G^{-1} \right\| (1 + b\beta^j)$. Denote by $\mathcal{K}$ the set of indices $k \in \mathbb{N}$ for which the square-bracketed term is negative. The above inequality then can be transformed (using $\alpha(1-a) + \log a \leq \alpha - 1 - \log\alpha$ for all $\alpha, a > 0$) into

$$
\begin{aligned}
&- \sum_{j\in\mathcal{K}\cap\{K,\ldots,k\}} \left[\left(1 - \frac{\tilde{q}_j}{\cos^2\tilde{\theta}_j} \right) \left(1 - C\zeta \sum_{i=j}^{k} \delta_i \right) + \log \frac{\tilde{q}_j}{\cos^2\tilde{\theta}_j} \right] - \sum_{j=K}^{k} \log\cos^2\tilde{\theta}_j \\
&\leq \gamma + \sum_{j\in\{K,\ldots,k\}\setminus\mathcal{K}} \left[\left(1 - \frac{\tilde{q}_j}{\cos^2\tilde{\theta}_j} \right) \left(1 - C\zeta \sum_{i=j}^{k} \delta_i \right) + \log \frac{\tilde{q}_j}{\cos^2\tilde{\theta}_j} \right] \leq \gamma + \sum_{j=K}^{k} \left[-C\zeta \sum_{i=j}^{k} \delta_i - \log\left(1 - C\zeta \sum_{i=j}^{k} \delta_i \right) \right] \\
&\leq \gamma + \sum_{j=K}^{\infty} \left[C\zeta \sum_{i=j}^{\infty} \delta_i \right]^2 \leq \gamma + \left(\frac{C\zeta b}{1-\beta} \right)^2 \sum_{j=K}^{\infty} \left\| T_{*,j}^{-1} T_j^{-1} T_{*,j+1} \right\|^2 \beta^{2j} < \infty \qquad \text{(A.2)}
\end{aligned}
$$

for K large enough and γ the constant introduced above. Here, we have used

$$
-C\zeta \sum_{i=j}^{k} \delta_i - \log\left(1 - C\zeta \sum_{i=j}^{k} \delta_i \right) \leq -C\zeta \sum_{i=j}^{\infty} \delta_i - \log\left(1 - C\zeta \sum_{i=j}^{\infty} \delta_i \right) = \frac{1}{2} \left[C\zeta \sum_{i=j}^{\infty} \delta_i \right]^2 + O\left(\left[C\zeta \sum_{i=j}^{\infty} \delta_i \right]^3 \right)
$$

where the last equality derives from Taylor expansion of $x \mapsto -x - \log(1-x)$ around $x = 0$. From the limit $k \to \infty$ in (A.2) we deduce

$$
\cos^2\tilde{\theta}_k \underset{k\to\infty}{\longrightarrow} 1 \,, \quad \tilde{q}_k \underset{\substack{k\to\infty \\ k\in\mathcal{K}}}{\longrightarrow} 1 \,.
$$

The limit $\sum_{i=j}^{\infty} \delta_i \to_{j\to\infty} 0$ and the definition of $\mathcal{K}$ furthermore imply

$$\tilde{q}_k \underset{\substack{k\to\infty \\ k\notin\mathcal{K}}}{\longrightarrow} 1\,.$$

We now obtain

$$\frac{\left\|\mathrm{D}f(x_k) + \mathrm{D}^2 f_{R_{x_k}}(0)(p_k,\cdot)\right\|_{x_k}}{\|p_k\|_{x_k}} = \frac{\left\|(\mathrm{D}^2 f_{R_{x_k}}(0) - B_k)(s_k,\cdot)\right\|_{x_k}}{\|s_k\|_{x_k}} \leq \frac{\left\|(T_{*,k}^{-*} G T_{*,k}^{-1} - \nabla^2 f_{R_{x_k}}(0))s_k\right\|_{x_k}}{\|s_k\|_{x_k}}$$
$$+\frac{\left\|T_{*,k}^{-*} G^{\frac{1}{2}}(\mathrm{id} - \tilde{B}_k)\tilde{s}_k\right\|_{x_k}}{\left\|T_{*,k} G^{-\frac{1}{2}}\tilde{s}_k\right\|_{x_k}} \leq \left\|T_{*,k}^{-*} G T_{*,k}^{-1} - \nabla^2 f_{R_{x_k}}(0)\right\| + \left\|T_{*,k}^{-1}\right\|^2 \|G\| \frac{\left\|(\mathrm{id} - \tilde{B}_k)\tilde{s}_k\right\|_{x^*}}{\|\tilde{s}_k\|_{x^*}}\,.$$

The last fraction equals $\sqrt{\frac{\tilde{q}_k^2}{\cos^2\tilde{\theta}_k} - 2\tilde{q}_k + 1}$ and thus converges to zero for $k \to \infty$. The first term on the right-hand side can be bounded above by

$$\left\|\nabla^2 f_{R_{x_k}}(R_{x_k}^{-1}(x^*)) - \nabla^2 f_{R_{x_k}}(0)\right\| + \left\|\nabla^2 f_{R_{x_k}}(R_{x_k}^{-1}(x^*)) - (T_{*,k} T_{x_k,x^*}^{R_{x_k}})^{-*} \nabla^2 f_{R_{x_k}}(R_{x_k}^{-1}(x^*))(T_{*,k} T_{x_k,x^*}^{R_{x_k}})^{-1}\right\|$$

of which the first term converges to zero and the second can be rewritten as

$$\left\|(T_{x_k,x^*}^{R_{x_k}})^* G T_{x_k,x^*}^{R_{x_k}} - T_{*,k}^{-*} G T_{*,k}^{-1}\right\| \leq \left\|T_{*,k}^{-1}\right\|^2 \left\|(T_{x_k,x^*}^{R_{x_k}} T_{*,k})^* G (T_{x_k,x^*}^{R_{x_k}} T_{*,k}) - G\right\|$$
$$\leq \left\|T_{*,k}^{-1}\right\|^2 \left\|T_{x_k,x^*}^{R_{x_k}} T_{*,k} - \mathrm{id}\right\| \|G\| \left(2 + \left\|\mathrm{id} - T_{x_k,x^*}^{R_{x_k}} T_{*,k}\right\|\right) \underset{k\to\infty}{\longrightarrow} 0\,,$$

where we have used $A^*GA - G = (\mathrm{id} - A)^* G(\mathrm{id} - A) - G(\mathrm{id} - A) - (\mathrm{id} - A)^* G$. □

REFERENCES

[1] P.-A. Absil, C. G. Baker, and K. A. Gallivan. Trust-region methods on Riemannian manifolds. *Found. Comput. Math.*, 7(3):303–330, 2007.

[2] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, Princeton, NJ, 2008.

[3] P.-A. Absil, R. Mahony, and R. Sepulchre. Optimization on manifolds: Methods and with applications. In *Recent Advances in Optimization and its Applications in Engineering. The 14th Belgian-French-German Conference on Optimization*, pages 125–144. Springer, 2010.

[4] Roy L. Adler, Jean-Pierre Dedieu, Joseph Y. Margulies, Marco Martens, and Mike Shub. Newton's method on Riemannian manifolds and a geometric model for the human spine. *IMA J. Numer. Anal.*, 22(3):359–390, 2002.

[5] C. G. Baker, P.-A. Absil, and K. A. Gallivan. An implicit trust-region method on Riemannian manifolds. *IMA J. Numer. Anal.*, 28(4):665–689, 2008.

[6] Christopher G. Baker. *Riemannian manifold trust-region methods with applications ot eigenproblems*. PhD thesis, Florida State Universtity, 2008.

[7] Ian Brace and Jonathan H. Manton. An improved BFGS-on-manifold algorithm for computing weighted low rank approximations. In *Proceedings of the 17th International Symposium on Mathematical Theory of Networks and Systems*, pages 1735–1738, 2006.

[8] T. F. Chan and L. A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, 2001.

[9] D. Gabay. Minimizing a differentiable function over a differential manifold. *J. Optim. Theory Appl.*, 37(2):177–219, 1982.

[10] Philipp Harms and Andrea Mennucci. Geodesics in infinite dimensional Stiefel and Grassmann manifolds. 2010. submitted.

[11] Knut Hüper and Jochen Trumpf. Newton-like methods for numerical optimization on manifolds. In *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers*, volume 1, pages 136–139. IEEE, 2004.

[12] Huibo Ji. *Optimization approaches on smooth manifolds*. PhD thesis, Australian National University, 2007.

[13] Martin Kilian, Niloy J. Mitra, and Helmut Pottmann. Geometric modeling in shape space. In *ACM Transactions on Graphics*, volume 26, pages #64, 1–8, 2007.

[14] Wilhelm P. A. Klingenberg. *Riemannian geometry*, volume 1 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, second edition, 1995.

[15] Peter W. Michor and David Mumford. Riemannian geometries on spaces of plane curves. *J. Eur. Math. Soc. (JEMS)*, 8(1):1–48, 2006.

[16] Jorge Nocedal and Stephen J. Wright. *Numerical optimization*. Springer Series in Operations Research and Financial Engineering. Springer, New York, second edition, 2006.

[17] Elijah Polak. *Optimization*, volume 124 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1997. Algorithms and consistent approximations.

[18] Chunhong Qi, Kyle A. Gallivan, and P.-A. Absil. Riemannian BFGS algorithm with applications. In *Recent Advances in Optimization and its Applications in Engineering. The 14th Belgian-French-German Conference on Optimization*, pages 183–192. Springer, 2010.

[19] Ekkehard W. Sachs. Broyden's method in Hilbert space. *Math. Programming*, 35(1):71–82, 1986.

[20] Barry Simon. *Trace ideals and their applications*, volume 120 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, second edition, 2005.

[21] S. T. Smith. Optimization techniques on Riemannian manifolds. *Fields Institute Communications*, 3:113–135, 1994.

[22] G. Sundaramoorthi, A.C.G. Mennucci, S. Soatto, and A. Yezzi. A new geometric metric in the space of curves, and applications to tracking deforming objects by prediction and filtering. *SIAM Journal on Imaging Sciences*, 2010. accepted.

[23] Constantin Udrişte. *Convex functions and optimization methods on Riemannian manifolds*, volume 297 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1994.

[24] Benedikt Wirth, Leah Bar, Martin Rumpf, and Guillermo Sapiro. A continuum mechanical approach to geodesics in shape space. *International Journal of Computer Vision*, 93:293–318, 2011.

[25] Y. Yang. Globally convergent optimization algorithms on Riemannian manifolds: uniform framework for unconstrained and constrained optimization. *J. Optim. Theory Appl.*, 132(2):245–265, 2007.

[26] Laurent Younes, Peter W. Michor, Jayant Shah, and David Mumford. A metric on shape space with explicit geodesics. *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl.*, 19(1):25–57, 2008.