Journal of Convex Analysis
Volume 25 (2018), No. 2, 341–370

Newton-Type Method for Solving Generalized Equations on Riemannian Manifolds

Samir Adly
Laboratoire Xlim, Université de Limoges,
123 Avenue Albert Thomas, 87060 Limoges, France
samir.adly@unilim.fr

Huynh Van Ngai
Department of Mathematics, University of Quy Nhon,
170 An Duong Vuong, Quy Nhon, Vietnam
ngaivn@yahoo.com

Nguyen Van Vu
Laboratoire Xlim, Université de Limoges,
123 Avenue Albert Thomas, 87060 Limoges, France
van-vu.nguyen@etu.unilim.fr

Dedicated to Antonino Maugeri on the occasion of his 70th birthday.

Received: November 29, 2015
Revised manuscript received: August 1, 2016
Accepted: August 2, 2016

This paper is devoted to the study of Newton-type algorithm for solving inclusions involving set-valued maps defined on Riemannian manifolds. We provide some sufficient conditions ensuring the existence as well as the quadratic convergence of Newton sequence. The material studied in this paper is based on Riemannian geometry as well as variational analysis, where metric regularity property is a key point.

Keywords: Riemannian manifold, generalized equation, Newton's method, metric regularity, variational inclusion.

1. Introduction

We consider the problem finding a solution of inclusion

$$0 \in f(x) + F(x). \tag{1}$$

Here, the variable x varies in a finite dimensional Riemannian manifold $\mathcal{M}$, $f : \mathcal{M} \longrightarrow \mathbb{R}^n$ is at least continuous, while $F : \mathcal{M} \rightrightarrows \mathbb{R}^n$ is a set-valued mapping. We suppose that the graph $\mathrm{Gr}(F)$ of F is closed with respect to the product topology on $\mathcal{M} \times \mathbb{R}^n$.

ISSN 0944-6532 / \$ 2.50

The general inclusion (1) covers many situations which have been studied widely in the literature. When $\mathcal{M}$ is a Euclidean space, (1) is nothing but the so-called generalized equation. In [3], the authors studied the (super-)linear convergence of a Newton-type iterative process for solving generalized equations. The Kantorovich approach and Smale's classical (α,γ)-theory was extended to generalized equations in [4]. If $F(x) \equiv 0$, problem (1) reduces to solve nonlinear equation $f(x) = 0$ on $\mathcal{M}$. In the case $F(x) \equiv K$ for a fixed cone $K \subset \mathbb{R}^n$, (1) coincides with the problem studied in [21].

The current work considers a scheme of Newton-type method to approximate a solution of (1). This is based on the well-known Josephy-Newton method applied to the generalized equation which was introduced in [8]. The strategy is that to start at a guess point x_0 nearby a solution, and generates a sequence (x_k, v_k) in the tangent bundle $T\mathcal{M}$ by the scheme

$$0 \in f(x_k) + \mathscr{D}f(x_k)(v_k) + (F \circ R_{x_k})(v_k), x_{k+1} = R_x(v_k). \tag{2}$$

In (2), $\mathscr{D}f(x) : T_x\mathcal{M} \longrightarrow \mathbb{R}^n$ and $R_x : T_x\mathcal{M} \longrightarrow \mathcal{M}$ are respectively the covariant derivative of f and the retraction at x (see Section 2). Observe that when $F \equiv 0$, (2) subsumes as a particular case of Newton-type methods for solving nonlinear equation on manifold $\mathcal{M}$ studied e.g. in [2, 6]. Furthermore, it might be viewed as an extension of algorithms for finding singularities of a vector field considered by the works [5, 9, 10] if f is replaced by a smooth vector field X, and F is the zero field $F(x) = 0_x$. In the same spirit, more other discussions about Newton-type method applying on smooth manifold can be found in [1, 19, 17, 20].

The paper is organized as follows. In section 2, we recall basic elements, notations and backgrounds on Riemannian geometry that will be useful in the sequel. Section 3 is devoted to the stability of the metric regularity property of the sum of two operators on Riemannian manifolds. In section 4, we prove the local and global convergence of the retraction Newton-type algorithm for solving (1).

2. Notions and backgrounds

Throughout this paper, we prefer to adopt the standard notations and different concepts used in [7, 15] dealing with the basic background from Riemannian geometry. For more details about these events, the readers are referred to [12, 1, 14, 18] and references therein.

Riemannian manifold, metric structure. All objects under consideration in this paper concerns finite dimensional smooth manifolds. The terminology smooth is meant to be differentiable at least beyond the order that appears. Elementary notions of differential geometry are assumed to be familiar.

A Riemannian manifold $\mathcal{M}$ of dimension m is a differentiable manifold of the same dimension which is endowed with a Riemannian metric g. If x is a point in $\mathcal{M}$, the norm induced by g in the tangent space $T_x\mathcal{M}$ of $\mathcal{M}$ at x is denoted by $\|\cdot\|_x$. By $\mathbb{B}_x(v,r)$ (resp. $\overline{\mathbb{B}}_x(v,r)$) we mean the open (resp. closed) ball in $T_x\mathcal{M}$ with center at $v \in T_x\mathcal{M}$ and radius $r > 0$. To mention about the open (closed) unit ball in $T_x\mathcal{M}$, we write $\mathbb{B}_x$ (resp. $\overline{\mathbb{B}}_x$).

Let $\chi\colon [a,b] \longrightarrow \mathcal{M}$ be a piecewise smooth curve, then its length is given by the quantity

$$\ell(\chi) = \int_a^b \|\chi'(t)\|_{\chi(t)}\, dt.$$

Recall that here $\chi'(t)$ stands for the tangent vector (or velocity) of χ at the instant t. (This is sometimes denoted in other ways, such as, $\dot{\chi}(t)$ or $\frac{d\chi}{dt}(t)$.) For two points x, y in the manifold $\mathcal{M}$, the Riemannian distance between x and y can be defined as follow

$$d_{\mathcal{M}}(x,y) = \inf\left\{ \ell(\chi) \;\middle|\; \begin{array}{l} \chi\colon [a,b] \subset \mathbb{R} \longrightarrow \mathcal{M} \ \text{ piecewise smooth;} \\ \chi(a) = x, \chi(b) = y \end{array} \right\}.$$

We will omit the subscript $\mathcal{M}$ when the manifold is fixed.
With the Riemannian distance $d_{\mathcal{M}}$, $\mathcal{M}$ becomes a metric space. Unless it has some other specification, the space $(\mathcal{M}, d_{\mathcal{M}})$ is always supposed to be complete. Completeness can be described via the famous Hofp–Rinow theorem (see [7, 15]). As usual, we denote the open and closed ball on $\mathcal{M}$ with center x and radius $r > 0$ by $\mathbb{B}_{\mathcal{M}}(x,r)$ and $\overline{\mathbb{B}}_{\mathcal{M}}(x,r)$, respectively.

Connection, covariant derivative. Let ∇ be the Levi-Civita connection on the manifold $\mathcal{M}$. Consider a smooth function $f\colon \mathcal{M} \longrightarrow \mathbb{R}$ and a vector field Y on $\mathcal{M}$. The covariant derivative of f with respect to Y is the function $\nabla_Y f := Yf$, where Yf indicates the action of Y on f (see [12, 15]). For $x \in \mathcal{M}$, the covariant derivative $\mathscr{D}f(x)\colon T_x\mathcal{M} \longrightarrow \mathbb{R}$ of f at x is a function which assigns to each $v \in T_x\mathcal{M}$ the value $\mathscr{D}f(x)(v) := (\nabla_Y f)(x) \in \mathbb{R}$, where Y is a vector field such that $Y_x = v$. Explicitly, we have

$$\mathscr{D}f(x)(v) = \chi'(0)(f) = (f \circ \chi)'(0) \tag{3}$$

for any smooth curve χ satisfying $\chi(0) = x$ and $\chi'(0) = v$.
Generally, if $f = (f_1, \dots, f_n) : \mathcal{M} \longrightarrow \mathbb{R}^n$ is a smooth map, then one defines

$$\mathscr{D}f(x) = \begin{pmatrix} \mathscr{D}f_1(x) \\ \vdots \\ \mathscr{D}f_n(x) \end{pmatrix}$$

as the covariant derivative of f at x. The covariant derivative $\mathscr{D}f(x)$ is a linear map from $T_x\mathcal{M}$ into $\mathbb{R}^n$ (cf.[6]). In the sequel, we use the following norm for the covariant derivative

$$\|\mathscr{D}f(x)\| := \sup_{\substack{u\in T_x\mathcal{M}\\ \|u\|_x\leqslant 1}} \|\mathscr{D}f(x)(u)\|_{\mathbb{R}^n}. \tag{4}$$

Vector transportation. We will need later the concept of vector transports. They provide a way to link between different tangent spaces of a manifold. Among them, parallel transportations are the most typical and important. Assume $\chi : [a,b] \longrightarrow \mathcal{M}$ is a smooth curve. By a vector field V along χ, we mean a smooth map such that $V(t) \in T_{\chi(t)}\mathcal{M}$. V is said to be parallel along χ if its covariant derivative $\nabla_{\chi'}V$ vanishes. Then, the parallel transport $P_{\chi,\chi(a),\chi(b)} : T_{\chi(a)}\mathcal{M} \longrightarrow T_{\chi(b)}\mathcal{M}$ is given by

$$P_{\chi,\chi(a),\chi(b)}(v) = V(b),$$

where V is the unique vector field along χ satisfying $\nabla_{\chi'}V = 0$ and $V(a) = v$. It is well-known from Riemannian geometry that $P_{\chi,\chi(a),\chi(b)}$ is a linear isometry. In particular, one always has

$$\left\|P_{\chi,\chi(a),\chi(b)}\right\|_{T_{\chi(a)}\mathcal{M},T_{\chi(b)}\mathcal{M}} = 1 \quad \text{and} \quad P^{-1}_{\chi,\chi(a),\chi(b)} = P_{\chi,\chi(b),\chi(a)}.$$

Retraction. Retraction (cf. [5, 17]) is a crucial object for our approach. In this paper, a retraction is meant to be a smooth map $R\colon T\mathcal{M} \longrightarrow \mathcal{M}$ from the tangent bundle into $\mathcal{M}$ so that its restriction R_x (retraction at x) of R onto each tangent space $T_x\mathcal{M}$ satisfies

- $R_x(0_x) = x$,
- $(dR_x)_{0_x} = id_{T_x\mathcal{M}}$, with the identification $T_{0_x}(T_x\mathcal{M}) \simeq T_x\mathcal{M}$.

On the preceding descriptions, $(dR_x)_{0_x}$ stands for the differential of R_x at the origin 0_x of $T_x\mathcal{M}$. If R is a retraction, and (x,v) is in $T\mathcal{M}$, then the map $\chi : t \longmapsto R_x(tv)$ defines a smooth curve on $\mathcal{M}$ satisfying $\chi(0) = x$ and $\chi'(0) = v$. There is a natural retraction reduced by geodesics on $\mathcal{M}$, known as exponential map of the tangent bundle [7]. Recall that a curve $\chi : I \longrightarrow \mathcal{M}$ is called a geodesic, if its acceleration $\nabla_{\chi'}\chi'$ is vanishing. The exponential map $\exp : T\mathcal{M} \longrightarrow \mathcal{M}$ can be defined by $\exp(x,v) = \chi(1,x,v)$, where $t \longmapsto \chi(t,x,v)$ is the unique geodesic such that $\chi(0,x,v) = x$ and $\chi'(0,x,v) = v$.

The exponential map has many important properties. At each point of a Riemannian manifold, there is a normal neighborhood for which $\exp_x$ is injective on some ball around the origin of tangent space $T_x\mathcal{M}$ (see [12, 7, 18]). In accordance with this property, we will establish here a similar result applied for any retraction. The proof is based on the same strategy as the case of exponential maps presented in the literature.

Proposition 2.1. *Let $\mathcal{M}$ be a complete m–dimensional Riemannian manifold, and let $R : T\mathcal{M} \longrightarrow \mathcal{M}$ be a retraction. Then for every $x \in \mathcal{M}$ there exists a neighborhood $W_x = \mathbb{B}_{\mathcal{M}}(x, \mathbf{r}_x)$ ($\mathbf{r}_x > 0$ depends on x) and a real number $C_R(x) > 0$ such that if $y \in W_x$ then $R_y : T_y\mathcal{M} \longrightarrow \mathcal{M}$ is injective in the ball $C_R(x)\mathbb{B}_y$ with $R_y\left(C_R(x)\mathbb{B}_y\right) \supset W_x$. We will call such a pair $(\mathbf{r}_x, C_R(x))$ as a R–normal pair at x.*

Proof. Let $(U, \mathbf{x})$ be a local coordinate of $\mathcal{M}$ at x. Then $\big(U \times U, (\mathbf{x}, \mathbf{x})\big)$ is a local coordinate on $\mathcal{M} \times \mathcal{M}$. Consider the map $H : TU \longrightarrow \mathcal{M} \times \mathcal{M}$ given by

$$H(z, v) = \big(z, R_z(v)\big).$$

Since $R_x(0_x) = x$ and $(dR_x)_{0_x} = id$, the matrix of $dH_{(x,0_x)}$ in the local coordinate above can be written as follow

$$\begin{pmatrix} I & 0 \\ * & I \end{pmatrix}.$$

Hence, we can apply an analogous argument as the proof of existence for normal neighborhood in Riemannian manifold and obtain the required conclusion. For shortness, one keeps in mind [15, Lemma 5.12] and/or [7, Theorem 3.7]. $\square$

Besides Proposition 2.1, we shall also need some other facts. The next statement will be useful throughout the rest of this paper.

Assumption 2.2. Given a retraction $R : T\mathcal{M} \longrightarrow \mathcal{M}$, an open subset $\Omega \subset \mathcal{M}$ and positive numbers $\varepsilon, \rho_1, \rho_2$. We say that R satisfies the uniform rate condition on Ω with respect to ratios ρ_1, ρ_2 and radius ε, written as $R \in$ $\mathbf{URC}(\rho_1, \rho_2, \varepsilon, \Omega)$, if

$$\rho_1 \left\|u - v\right\|_x \leqslant d\big(R_x(u), R_x(v)\big) \leqslant \rho_2 \left\|u - v\right\|_x \tag{5}$$

for all $x \in \Omega$ and $u, v \in \varepsilon\mathbb{B}_x$.

Remark 2.3. It is obvious to see that (5) is trivial in the case $\mathcal{M} = \mathbb{R}^n$ and R is the usual translation $R_x(u) = x + u$. Another less trivial case will be showed in Example 2.4 below. When $R = \exp$ is the exponential, the left-hand side of (5) is satisfied on any Hadamard manifold (for instance, cf. [18], Chapter V, Proposition 4.5). If $R = \exp$ and $\Omega = \mathbb{B}_{\mathcal{M}}(\bar{x}, r)$, where r is small enough, the right-hand side of (5) is also valid in terms of local property [12, 6]. For general retractions, a condition ensures (5) appeared in the work [17], where retractions R_x are required to satisfy the equicontinuous derivative on Ω.

Example 2.4 (retraction on the unit sphere). Consider the m-unit sphere

$$\mathbb{S}^m := \big\{x \in \mathbb{R}^{m+1} : \|x\|^2 = x^T x = 1\big\},$$

where the vector is written in column form. $\mathbb{S}^m$ is endowed with the Riemannian metric

$$g_x(u, v) = \langle u, v\rangle_{T_x\mathbb{S}^m} := u^T v, \quad \forall u, v \in T_x\mathbb{S}^m = \big\{z \in \mathbb{R}^{m+1} : z^T x = 0\big\}. \tag{6}$$

The Riemannian distance associated to the metric above is given by

$$d(x,y) = \arccos\left(x^T y\right), \quad x, y \in \mathbb{S}^m. \tag{7}$$

According to [1], let us consider the retraction

$$R_x(u) = \frac{x+u}{\|x+u\|}, \quad x \in \mathbb{S}^m, u \in T_x\mathbb{S}^m. \tag{8}$$

Observe that we have

$$d\big(x, R_x(u)\big) = \arccos\left(\frac{x^T(x+u)}{\|x+u\|}\right) = \arccos\frac{1}{\sqrt{1+\|u\|^2}} = \arctan\|u\|.$$

So, R_x is injective on the whole space $T_x\mathbb{S}^m$ and $R_x(\varepsilon\mathbb{B}_x) = \mathbb{B}_{\mathbb{S}^m}(x, \arctan\varepsilon)$ for every $x \in \mathbb{S}^m$ and $\varepsilon > 0$.

Let $r > 0$ be small enough and $\Omega \subset \mathbb{S}^m$ is a closed set. Suppose that $x \in \Omega$ and $u, v \in r\mathbb{B}_x \subset T_x\mathbb{S}^m$ are fixed. For brevity, we put $\hat{d} := d\big(R_x(u), R_x(v)\big)$. From (7) and (8) we get

$$\begin{aligned}\cos\hat{d} = R_x(u)^T R_x(v) &= \frac{1}{\|x+u\|\,\|x+v\|}(x+u)^T(x+v) \\ &= \frac{1+u^T v}{\left(1+\|u\|^2\right)^{1/2}\left(1+\|v\|^2\right)^{1/2}}.\end{aligned}$$

Hence

$$\sin^2\hat{d} = 1-\cos^2\hat{d} = \frac{\|u-v\|^2 + \|u\|^2\|v\|^2 - \left(u^T v\right)^2}{\left(1+\|u\|^2\right)\left(1+\|v\|^2\right)}.$$

It is clear that $\|u\|^2\|v\|^2 - \left(u^T v\right)^2 \geqslant 0$, so

$$\sin^2\hat{d} \geqslant \frac{\|u-v\|^2}{\left(1+\|u\|^2\right)\left(1+\|v\|^2\right)} \geqslant \left(\frac{1}{1+r^2}\right)^2\|u-v\|^2. \tag{9}$$

On the other hand, we have $2u^T v = \|u\|^2 + \|v\|^2 - \|u-v\|^2$, which implies

$$\begin{aligned}\|u\|^2\|v\|^2 - \left(u^T v\right)^2 &= \|u\|^2\|v\|^2 - \frac{1}{4}\left(\|u\|^2+\|v\|^2-\|u-v\|^2\right)^2 \\ &= \frac{1}{2}\left(\|u\|^2+\|v\|^2\right)\|u-v\|^2 - \frac{1}{4}\left[\|u-v\|^4 + \left(\|u\|^2-\|v\|^2\right)^2\right] \\ &\leqslant \frac{1}{2}\left(\|u\|^2+\|v\|^2\right)\|u-v\|^2.\end{aligned}$$

Thus,

$$\sin^2\hat{d} \leqslant \frac{\left[1+\frac{1}{2}\left(\|u\|^2+\|v\|^2\right)\right]\|u-v\|^2}{\left(1+\|u\|^2\right)\left(1+\|v\|^2\right)} \leqslant \|u-v\|^2. \tag{10}$$

Note that

$$\hat{d} \leqslant d\big(x, R_x(u)\big) + d\big(x, R_x(v)\big) = \arctan \|u\| + \arctan \|v\| \leqslant \frac{\pi}{2}$$

whenever u, v are in the unit ball of $T_x\mathbb{S}^m$. Taking into account the following fact

$$0 < t \leqslant \frac{\pi}{2} \Longrightarrow \frac{2}{\pi} \leqslant \frac{\sin t}{t} \leqslant 1,$$

(9) and (10) give us

$$\frac{1}{1+r^2} \leqslant \frac{d\big(R_x(u), R_x(v)\big)}{\|u-v\|} \leqslant \frac{\pi}{2}, 0 \leqslant r \leqslant 1, u \neq v \in r\mathbb{B}_x.$$

Consequently,

$$R \in \mathbf{URC}\left(\rho_1(r), \rho_2(r), r, \Omega\right), \tag{11}$$

where $r \in (0,1]$, $\rho_1(r) = \frac{1}{1+r^2}$ and $\rho_2(r) = \frac{\pi}{2}$.

3. Stability of the metric regularity property on Riemannian manifolds

To study the convergence of schemes like (2), the concept of metric regularity for set-valued mapping [13] plays an important role in our analysis. Recall that a set-valued (or also multivalued) mapping $\Phi : X \rightrightarrows Y$ between two metric spaces X and Y is a correspondence assigning to each $x \in X$ a subset $\Phi(x) \subset Y$. Such a mapping is totally determined by its graph $\mathrm{Gr}(\Phi) := \big\{(x, y) \in X \times Y : y \in \Phi(x)\big\}$. The mapping Φ is said to be metrically regular on a subset $V \subset X \times Y$ if there exists a constant $\kappa > 0$ (a modulus of regularity) such that

$$\mathrm{dist}_X\big(x, \Phi^{-1}(y)\big) \leqslant \kappa \mathrm{dist}_Y\big(y, \Phi(x)\big), \quad \text{for all} \quad (x, y) \in V. \tag{12}$$

In (12), $\mathrm{dist}_X(\cdot,\cdot)$ and $\mathrm{dist}_Y(\cdot,\cdot)$ are respectively the distances in X and Y, while Φ^{-1} stands for the inverse of Φ, given by

$$x \in \Phi^{-1}(y) \Longleftrightarrow y \in \Phi(x).$$

When dealing with a property like (12), we will sometimes write $\kappa \in \mathrm{REG}\big(\Phi, V\big)$ to simplify notations. Moreover, we also write $G \in \mathscr{L}ip_\lambda(S)$ to indicate that the mapping (maybe set-valued) G is Lipschitz continuous on the set S with a modulus λ. In words, $G \in \mathscr{L}ip_\lambda(S)$ means

$$d_Y^{\mathcal{H}}\big(G(x), G(y)\big) \leqslant \lambda d_X(x, y), \quad x, y \in S, \tag{13}$$

where $d_Y^{\mathcal{H}}\big(C, D\big)$ is the Hausdorff distance [16] between two subsets C and D in Y

$$d_Y^{\mathcal{H}}\big(C, D\big) = \sup_{y \in Y} \left| d_Y\big(y, C\big) - d_Y\big(y, D\big) \right|. \tag{14}$$

We now present the main results of this section, which ensures the stability of metric regularity property. These are essential preliminaries for the convergence analysis at the end of this paper.

Proposition 3.1 (stability of local metric regularity). *Given a complete Riemannian manifold of dimension m, and a retraction $R : T\mathcal{M} \longrightarrow \mathcal{M}$. Fix a point $x \in \mathcal{M}$ and let $(\mathbf{r}_x, C_R(x))$ be a R–normal pair at x. Suppose that there exist some positive numbers $0 < \rho_1 \leqslant \rho_2$ satisfying $R \in \mathbf{URC}(\rho_1, \rho_2, \delta, W_x)$ with $W_x = \mathbb{B}_{\mathcal{M}}(x, \mathbf{r}_x)$ and $\delta = C_R(x)$. Pick some tangent vector $\bar{u} \in T_x\mathcal{M}$ and some positive constants r, s, r', s', λ, λ', σ and κ obeying the following relations*

$$\begin{cases} \theta = \rho_1^{-1}\rho_2\kappa\left(\lambda' + \lambda\mu_x\right) & < 1, \\ \rho_1^{-1}\rho_2\left(1 + \frac{1+\theta}{1-\theta}\right)r' + \frac{\kappa}{1-\theta}s' & \leqslant r, \\ \left(\rho_1^{-1}\rho_2\right)^2\left(\lambda' + \lambda\mu_x\right)\left(1 + \frac{1+\theta}{1-\theta}\right)r' + \left(1 + \frac{\theta}{1-\theta}\right)s' & \leqslant s, \\ \rho_1^{-1}\rho_2\left\|\bar{u}\right\|_x + \rho_1^{-1}\sigma + \rho_1^{-1}\rho_2\left(1 + \frac{1+\theta}{1-\theta}\right)r' + \frac{\kappa}{1-\theta}s' & \leqslant \delta, \\ \rho_2\left\|\bar{u}\right\|_x + \rho_2\left[\rho_1^{-1}\rho_2\left(1 + \frac{1+\theta}{1-\theta}\right)r' + \frac{\kappa}{1-\theta}s'\right] & \leqslant \mathbf{r}_x, \end{cases} \tag{15}$$

for $\mu_x = \|\mathscr{D}f(x)\|$. Assume that the mapping $\Phi_x = \mathscr{D}f(x)(\cdot) + (F \circ R_x)(\cdot)$ is metrically regular on a neighborhood $V_x := \mathbb{B}_x(\bar{u}, r) \times \mathbb{B}_{\mathbb{R}^n}(\bar{p}, s)$ of $(\bar{u}, \bar{p}) \in \mathrm{Gr}(\Phi_x)$ together with a modulus κ. Pick some point $y \in \mathbb{B}_{\mathcal{M}}(x, \sigma) \cap W_x$ such that, for each geodesic $\chi : [0, 1] \longrightarrow \mathcal{M}$ having $\chi(0) = x$, $\chi(1) = y$ and $\chi([0, 1]) \subset W_x$:

(i) $\Sigma_{\chi,y,x} = R_x^{-1} \circ R_y - P_{\chi,y,x} \in \mathscr{L}ip_\lambda(\delta\mathbb{B}_y)$;

(ii) the linear map $G_{\chi,x,y} := \mathscr{D}f(y) \circ P_{\chi,x,y} - \mathscr{D}f(x)$ *satisfies* $\|G_{\chi,x,y}\| \leqslant \lambda'$.

Set $\bar{v} = \left(R_y^{-1} \circ R_x\right)(\bar{u}) \in T_y\mathcal{M}$, $\bar{q} = \bar{p} - \mathscr{D}f(x)(\bar{u}) + \mathscr{D}f(y)(\bar{v}) \in \mathbb{R}^n$, and $\tau = \rho_1^{-1}\rho_2\frac{\kappa}{1-\theta}$. Then one has $\tau \in \mathrm{REG}\left(\Phi_y, V_y\right)$, where $\Phi_y = \mathscr{D}f(y)(\cdot) + (F \circ R_y)(\cdot)$ and $V_y := \mathbb{B}_y(\bar{v}, r') \times \mathbb{B}_{\mathbb{R}^n}(\bar{q}, s')$.

To prove Proposition 3.1, we need the following lemma.

Lemma 3.2. *Under the assumptions of Proposition* 3.1*, one has*

$$\Psi_{y,x} := \mathscr{D}f(y) - \mathscr{D}f(x) \circ \left(R_x^{-1} \circ R_y\right) \in \mathscr{L}ip_{\lambda'+\lambda\mu_x}(\delta\mathbb{B}_y).$$

Proof. Without loss of generality, we can focus on the case where W_x is a normal neighborhood at x. Let $\chi\colon [0, 1] \longrightarrow W_x$ be a geodesic having $\chi(0) = x$ and $\chi(1) = y$. It is not difficult to check that

$$\Psi_{y,x}(v) = \left(G_{\chi,x,y} \circ P_{\chi,x,y}^{-1}\right)(v) + \left(\mathscr{D}f(x) \circ \Sigma_{\chi,y,x}\right)(v),$$

for $v \in \delta\mathbb{B}_y$. As a result, we find

$$
\begin{aligned}
&\|\Psi_{y,x}(v') - \Psi_{y,x}(v)\|_{\mathbb{R}^n} \\
&= \left\|\left(G_{\chi,y,x} \circ P_{\chi,x,y}^{-1}\right)(v'-v) + \mathscr{D}f(x)\left(\Sigma_{\chi,y,x}(v') - \Sigma_{\chi,y,x}(v)\right)\right\|_{\mathbb{R}^n} \\
&\leqslant \|G_{\chi,y,x}\| \left\|P_{\chi,x,y}^{-1}\right\| \|v'-v\|_y + \|\mathscr{D}f(x)\| \|\Sigma_{\chi,y,x}(v') - \Sigma_{\chi,y,x}(v)\|_x \\
&\leqslant \lambda' \|v'-v\|_y + \mu_x \lambda \|v'-v\|_y .
\end{aligned}
$$

Hence, the conclusion of Lemma 3.2 follows. □

Proof of Proposition 3.1. Pick $(v,w) \in V_y$. We have to establish the estimation

$$
\operatorname{dist}\big(v, \Phi_y^{-1}(w)\big) \leqslant \tau \operatorname{dist}\big(w, \Phi_y(v)\big). \tag{16}
$$

We omitted the subscripts of spaces in (16), since these ones are determined by the objects upon which the corresponding distances act. Here and in what follows, we will use the common notation $\operatorname{dist}(\cdot,\cdot)$ for distance on any arbitrary metric space.

The set $\Phi_y(v)$ is obviously closed. So, (16) is easy when $\operatorname{dist}\big(w, \Phi_y(v)\big) = 0$. Consider the case $\eta = \operatorname{dist}\big(w, \Phi_y(v)\big) > 0$. First, we set $v_0 = v$ and $y_0 = R_y(v_0)$. Note that $\bar{v} = R_y^{-1}(R_x(\bar{u})) \in C_R(x)\mathbb{B}_y$, the triangle inequality and Assumption 2.2 give us

$$
\begin{aligned}
\|\bar{v}\|_y &\leqslant \rho_1^{-1} d\big(y, R_y(\bar{v})\big) = \rho_1^{-1} d\big(y, R_x(\bar{u})\big) \leqslant \rho_1^{-1}\left(d(y,x) + d\big(x, R_x(\bar{u})\big)\right) \\
&\leqslant \rho_1^{-1}\left(d(x,y) + \rho_2 \|\bar{u}\|_x\right) \leqslant \rho_1^{-1}\sigma + \rho_1^{-1}\rho_2 \|\bar{u}\|_x .
\end{aligned}
$$

This implies

$$
\|v_0\|_y \leqslant \|v_0 - \bar{v}\|_y + \|\bar{v}\|_y < r' + \rho_1^{-1}\sigma + \rho_1^{-1}\rho_2 \|\bar{u}\|_x < \delta.
$$

Since $R \in \mathbf{URC}(\rho_1, \rho_2, \delta, W_x)$, we have

$$
\begin{aligned}
d\big(y_0, x\big) &\leqslant d\big(y_0, R_y(\bar{v})\big) + d\big(R_y(\bar{v}), x\big) = d\big(R_y(v_0), R_y(\bar{v})\big) + d\big(R_x(\bar{u}), x\big) \\
&\leqslant \rho_2 \|v_0 - \bar{v}\|_y + \rho_2 \|\bar{u}\|_x < \rho_2 r' + \rho_2 \|\bar{u}\|_x \\
&< \rho_2 \|\bar{u}\|_x + \rho_2 \left[\rho_1^{-1}\rho_2\left(1 + \frac{1+\theta}{1-\theta}\right) r' + \frac{\kappa}{1-\theta} s'\right] \leqslant \mathbf{r}_x.
\end{aligned}
$$

Thus, $y_0 = R_y(v_0) \in W_x \subset R_x(\delta\mathbb{B}_x)$. Let us choose $u_0 \in \delta\mathbb{B}_x$ with $R_x(u_0) = y_0$ and put

$$
w_0 = w + \mathscr{D}f(x)(u_0) - \mathscr{D}f(y)(v_0) \in \mathbb{R}^n.
$$

We claim that $(u_0, w_0) \in V_x$. Indeed, due to Assumption 2.2 and the choice of u_0 we get

$$
\begin{aligned}
\|u_0 - \bar{u}\|_x &\leqslant \rho_1^{-1} d\big(R_x(u_0), R_x(\bar{u})\big) = \rho_1^{-1} d\big(R_y(v_0), R_y(\bar{v})\big) \\
&\leqslant \rho_1^{-1}\rho_2 \|v_0 - \bar{v}\|_y ,
\end{aligned} \tag{17}
$$

which yields $\|u_0 - \bar{u}\|_x < \rho_1^{-1}\rho_2 r' < r$. Taking into account $R_x(u_0) = R_y(v_0)$ and $R_x(\bar{u}) = R_y(\bar{v})$, we deduce

$$\begin{aligned} \|w_0 - \bar{p}\|_{\mathbb{R}^n} &= \|w - \Psi_{y,x}(v_0) - [\bar{q} - \Psi_{y,x}(\bar{v})]\|_{\mathbb{R}^n} \\ &\leqslant \|\Psi_{y,x}(v_0) - \Psi_{y,x}(\bar{v})\|_{\mathbb{R}^n} + \|w - \bar{q}\|_{\mathbb{R}^n} \\ &\leqslant (\lambda' + \lambda\mu_x)\,\|v_0 - \bar{v}\| + \|w - \bar{q}\|_{\mathbb{R}^n} . \end{aligned} \tag{18}$$

In view of (18), we obtain

$$\begin{aligned} \|w_0 - \bar{p}\|_{\mathbb{R}^n} &< (\lambda' + \lambda\mu_x) r' + s' \\ &< \left(\rho_1^{-1}\rho_2\right)^2 (\lambda' + \lambda\mu_x) \left(1 + \frac{1+\theta}{1-\theta}\right) r' + \left(1 + \frac{\theta}{1-\theta}\right) s' \leqslant s. \end{aligned}$$

Hence, the inclusion $(u_0, v_0) \in V_x$ is clear. By invoking the fact $\kappa \in \mathrm{REG}(\Phi_x, V_x)$, we find

$$\begin{aligned} \operatorname{dist}\bigl(u_0, \Phi_x^{-1}(w_0)\bigr) &\leqslant \kappa \operatorname{dist}\bigl(w_0, \Phi_x(u_0)\bigr) \\ &= \kappa \operatorname{dist}\bigl(w_0, \mathscr{D}f(x)(u_0) + (F \circ R_x)(u_0)\bigr) \\ &= \kappa \operatorname{dist}\bigl(w, \mathscr{D}f(y)(v_0) + (F \circ R_y)(v_0)\bigr) \\ &= \kappa \operatorname{dist}\bigl(w, \mathscr{D}f(y)(v) + (F \circ R_y)(v)\bigr) = \kappa\eta. \end{aligned}$$

So we can select a vector u_1 in $\Phi_x^{-1}(w_0)$ which satisfies

$$\|u_0 - u_1\|_x = \operatorname{dist}\bigl(u_0, \Phi_x^{-1}(w_0)\bigr) \leqslant \kappa\eta.$$

To continue, let us set $y_1 = R_x(u_1)$. Inasmuch as $\kappa \in \mathrm{REG}(\Phi_x, V_x)$, one has

$$\operatorname{dist}\bigl(\bar{u}, \Phi_x^{-1}(w_0)\bigr) \leqslant \kappa \operatorname{dist}\bigl(w_0, \Phi_x(\bar{u})\bigr) \leqslant \kappa\,\|w_0 - \bar{p}\|_{\mathbb{R}^n} .$$

Consequently,

$$\begin{aligned} \|u_1 - u_0\|_x &= \operatorname{dist}\bigl(u_0, \Phi_x^{-1}(w_0)\bigr) \\ &\leqslant \|u_0 - \bar{u}\|_x + \operatorname{dist}\bigl(\bar{u}, \Phi_x^{-1}(w_0)\bigr) \\ &\leqslant \|u_0 - \bar{u}\|_x + \kappa\,\|w_0 - \bar{p}\|_{\mathbb{R}^n} . \end{aligned} \tag{19}$$

By induction hypothesis, suppose that the tangent vectors $u_0, \ldots, u_k \in \mathbb{B}_x(\bar{u}, r)$, $v_0, \ldots, v_{k-1} \in \delta\mathbb{B}_y \subset T_y\mathcal{M}$ are given, and they obey the following conditions

- $R_y(v_j) = R_x(u_j)$, $j = 0, 1, \ldots, k-1$;
- $u_{j+1} \in \Phi_x^{-1}(w_j)$, with $w_j = w + \mathscr{D}f(x)(u_j) - \mathscr{D}f(y)(v_j)$, $j = 0, 1, \ldots, k-1$;
- $\|u_j - u_{j+1}\|_x \leqslant \theta^j\,\|u_1 - u_0\|_x$, for $\theta = \rho_1^{-1}\rho_2\kappa(\lambda' + \lambda\mu_x) < 1$ and $j = 0, 1, \ldots, k-1$.

Thanks to the triangle inequality, we can write

$$\begin{aligned}\|u_j - \bar{u}\|_x &\leqslant \sum_{i=0}^{j-1} \|u_i - u_{i+1}\|_x + \|u_0 - \bar{u}\|_x \leqslant \sum_{i=0}^{j-1} \theta^i \|u_1 - u_0\|_x + \|u_0 - \bar{u}\|_x \\ &= \frac{1-\theta^j}{1-\theta} \|u_1 - u_0\|_x + \|u_0 - \bar{u}\|_x \leqslant \frac{1}{1-\theta} \|u_1 - u_0\|_x + \rho_1^{-1}\rho_2 \|v_0 - \bar{v}\|_y .\end{aligned}$$

Since $\rho_1^{-1}\rho_2 \geqslant 1$, from (17), (18) and (19) we obtain

$$\begin{aligned}&\|u_j - \bar{u}\|_x \leqslant \\ &\leqslant \left\{\rho_1^{-1}\rho_2 + \frac{1}{1-\theta}\left[\rho_1^{-1}\rho_2 + \kappa(\lambda' + \lambda\mu_x)\right]\right\} \|v_0 - \bar{v}\|_x + \frac{\kappa}{1-\theta} \|w - \bar{q}\|_{\mathbb{R}^n} \\ &\leqslant \rho_1^{-1}\rho_2 \left(1 + \frac{1+\theta}{1-\theta}\right) \|v_0 - \bar{v}\|_x + \frac{\kappa}{1-\theta} \|w - \bar{q}\|_{\mathbb{R}^n} .\end{aligned}$$

The latter implies

$$\begin{aligned}\|u_j\|_x &\leqslant \|u_j - \bar{u}\|_x + \|\bar{u}\|_x \\ &\leqslant \rho_1^{-1}\rho_2 \left(1 + \frac{1+\theta}{1-\theta}\right) \|v_0 - \bar{v}\|_x + \frac{\kappa}{1-\theta} \|w - \bar{q}\|_{\mathbb{R}^n} + \|\bar{u}\|_x \\ &< \rho_1^{-1}\rho_2 \left(1 + \frac{1+\theta}{1-\theta}\right) r' + \frac{\kappa}{1-\theta} s' + \|\bar{u}\|_x \leqslant \delta.\end{aligned}$$

Next, involving again Assumption 2.2, we derive

$$\begin{aligned}d\big(R_x(u_k), x\big) &\leqslant d\big(R_x(u_k), R_x(\bar{u})\big) + d\big(R_x(\bar{u}), x\big) \leqslant \rho_2 \|u_k - \bar{u}\|_x + \rho_2 \|\bar{u}\|_x \\ &\leqslant \rho_2 \left[\rho_1^{-1}\rho_2 \left(1 + \frac{1+\theta}{1-\theta}\right) \|v_0 - \bar{v}\|_x + \frac{\kappa}{1-\theta} \|w - \bar{q}\|_{\mathbb{R}^n}\right] + \rho_2 \|\bar{u}\|_x \\ &< \rho_2 \left[\rho_1^{-1}\rho_2 \left(1 + \frac{1+\theta}{1-\theta}\right) r' + \frac{\kappa}{1-\theta} s'\right] + \rho_2 \|\bar{u}\|_x \leqslant \mathbf{r}_x.\end{aligned}$$

This means $y_k = R_x(u_k) \in W_x$. Selecting $v_k \in \delta\mathbb{B}_y$ with $R_y(v_k) = R_x(u_k)$ and

$$w_k = w + \mathscr{D}f(x)(u_k) - \mathscr{D}f(y)(v_k) = w - \Psi_{y,x}(v_k),$$

we get

$$\begin{aligned}\|w_k - \bar{p}\|_{\mathbb{R}^n} &= \|w - \Psi_{y,x}(v_k) - [\bar{q} - \Psi_{y,x}(\bar{v})]\|_{\mathbb{R}^n} \\ &\leqslant \|\Psi_{y,x}(v_k) - \Psi_{y,x}(\bar{v})\|_{\mathbb{R}^n} + \|w - \bar{q}\|_{\mathbb{R}^n} \\ &\leqslant (\lambda' + \lambda\mu_x) \|v_k - \bar{v}\|_y + \|w - \bar{q}\|_{\mathbb{R}^n} \\ &< (\lambda' + \lambda\mu_x) \|v_k - \bar{v}\|_y + s'.\end{aligned}$$

Moreover, it holds that

$$\begin{aligned}\|v_k-\bar v\|_y &\leqslant \rho_1^{-1}d\big(R_y(v_k),R_y(\bar v)\big)=\rho_1^{-1}d\big(R_x(u_k),R_x(\bar u)\big)\leqslant \rho_1^{-1}\rho_2\,\|u_k-\bar u\|_x\\ &\leqslant \rho_1^{-1}\rho_2\left\{\rho_1^{-1}\rho_2\left(1+\frac{1+\theta}{1-\theta}\right)\|v_0-\bar v\|_x+\frac{\kappa}{1-\theta}\|w-\bar q\|_{\mathbb{R}^n}\right\}\\ &< \rho_1^{-1}\rho_2\left\{\rho_1^{-1}\rho_2\left(1+\frac{1+\theta}{1-\theta}\right)r'+\frac{\kappa}{1-\theta}s'\right\}.\end{aligned}$$

Hence,

$$\begin{aligned}\|w_k-\bar p\|_{\mathbb{R}^n} &< \rho_1^{-1}\rho_2(\lambda'+\lambda\mu_x)\left\{\rho_1^{-1}\rho_2\left(1+\frac{1+\theta}{1-\theta}\right)r'+\frac{\kappa}{1-\theta}s'\right\}+s'\\ &=\big(\rho_1^{-1}\rho_2\big)^2(\lambda'+\lambda\mu_x)\left(1+\frac{1+\theta}{1-\theta}\right)r'+\left(1+\frac{\theta}{1-\theta}\right)s'\leqslant s.\end{aligned}$$

In other words, (u_k,w_k) belongs to V_x. Including the hypothesis of metric regularity property once more, we find

$$\begin{aligned}&\operatorname{dist}\big(u_k,\Phi_x^{-1}(w_k)\big)\leqslant \kappa\operatorname{dist}\big(w_k,\Phi_x(u_k)\big)\leqslant \kappa\,\|w_k-w_{k-1}\|_{\mathbb{R}^n}\\ &=\kappa\,\|w-\Psi_{y,x}(v_k)-[w-\Psi_{y,x}(v_{k-1})]\|_{\mathbb{R}^n}=\kappa\,\|\Psi_{y,x}(v_k)-\Psi_{y,x}(v_{k-1})\|_{\mathbb{R}^n}\\ &\leqslant \kappa(\lambda'+\lambda\mu_x)\,\|v_k-v_{k-1}\|_y\,.\end{aligned}$$

On the other hand,

$$\begin{aligned}\|v_k-v_{k-1}\|_y &\leqslant \rho_1^{-1}d\big(R_y(v_k),R_y(v_{k-1})\big)=\rho_1^{-1}d\big(R_x(u_k),R_x(u_{k-1})\big)\\ &\leqslant \rho_1^{-1}\rho_2\,\|u_k-u_{k-1}\|_x\,,\end{aligned}$$

which implies that

$$\operatorname{dist}\big(u_k,\Phi_x^{-1}(w_k)\big)\leqslant \theta\,\|u_k-u_{k-1}\|_x\leqslant \theta^k\,\|u_1-u_0\|_x\,.$$

However, $\Phi_x^{-1}(w_k)$ is a closed subset of a finite dimension space, so it must contain at least a vector u_{k+1} such that

$$\|u_k-u_{k+1}\|_x=\operatorname{dist}\big(u_k,\Phi_x^{-1}(w_k)\big)\leqslant \theta^k\,\|u_1-u_0\|_x\,.$$

As a result,

$$\begin{aligned}\|u_{k+1}-\bar u\|_x &\leqslant \sum_{j=0}^{k}\|u_{j+1}-u_j\|_x+\|u_0-\bar u\|_x\\ &\leqslant \sum_{j=0}^{k}\theta^j\,\|u_1-u_0\|_x+\|u_0-\bar u\|_x\\ &\leqslant \frac{1}{1-\theta}\|u_1-u_0\|_x+\|u_0-\bar u\|_x\,.\end{aligned}$$

According to (18) and (19) it is possible to write

$$\begin{aligned}\|u_1 - u_0\|_x &\leqslant \|u_0 - \bar{u}\|_x + \kappa \|w_0 - \bar{p}\|_{\mathbb{R}^n} \\ &\leqslant [\rho_1^{-1}\rho_2 + \kappa(\lambda' + \lambda\mu_x)] \|v_0 - \bar{v}\|_y + \kappa \|w - \bar{q}\|_{\mathbb{R}^n} \\ &< \rho_1^{-1}\rho_2(1+\theta)r' + \kappa s'.\end{aligned}$$

Notice that $\|u_0 - \bar{u}\|_x \leqslant \rho_1^{-1}\rho_2 \|v_0 - \bar{v}\|_y < \rho_1^{-1}\rho_2 r'$, we deduce

$$\|u_{k+1} - \bar{u}\|_x < \rho_1^{-1}\rho_2 \left(1 + \frac{1+\theta}{1-\theta}\right) r' + \frac{\kappa}{1-\theta} s' \leqslant r.$$

In particular, $u_{k+1} \in \mathbb{B}_x(\bar{u}, r)$, the sequences (u_k) and (v_k) are now well-defined by induction.

From the construction above, both (u_k) and (v_k) are Cauchy sequences. Thus, there are $u^* \in T_x\mathcal{M}$ and $v^* \in T_y\mathcal{M}$ such that $u_k \to u^*$ and $v_k \to v^*$. Passing to the limit in the inclusion

$$w + \mathscr{D}f(x)(u_k) - \mathscr{D}f(y)(v_k) = w_k \in \Phi_x(u_{k+1})$$

we conclude

$$w + \mathscr{D}f(x)(u^*) - \mathscr{D}f(y)(v^*) \in \Phi_x(u^*).$$

Equivalently, $w \in \mathscr{D}f(y)(v^*) + (F \circ R_x)(u^*)$. But from $R_x(u_k) = R_y(v_k)$ we also get $R_x(u^*) = R_y(v^*)$, which gives us

$$w \in \mathscr{D}f(y)(v^*) + (F \circ R_y)(v^*) = \Phi_y(v^*).$$

Consequently,

$$\begin{aligned}\operatorname{dist}\big(v, \Phi_y^{-1}(w)\big) &\leqslant \|v - v^*\|_y = \|v_0 - v^*\|_y = \left\|\sum_{k\geqslant 0}(v_k - v_{k+1})\right\|_y \\ &\leqslant \sum_{k\geqslant 0} \|v_k - v_{k+1}\|_y \leqslant \sum_{k\geqslant 0} \rho_1^{-1}\rho_2 \|u_k - u_{k+1}\|_x \\ &\leqslant \rho_1^{-1}\rho_2 \sum_{k\geqslant 0} \theta^k \|u_1 - u_0\|_x = \rho_1^{-1}\rho_2 \frac{1}{1-\theta} \|u_1 - u_0\|_x \\ &\leqslant \rho_1^{-1}\rho_2 \frac{1}{1-\theta} \kappa\eta.\end{aligned}$$

This is exactly (16). The proof is thereby complete. □

A semi-local version of Proposition 3.1 will be useful for the study of algorithm (2). The next statement is in this sense.

Proposition 3.3 (stability of semi-local metric regularity). *Let $\mathcal{M}$, R, $\mathbf{r}_x$, W_x, $\delta = C_R(x)$ and ρ_1, ρ_2 be as in the statement of Proposition 3.1. Let $f\colon \mathcal{M} \longrightarrow \mathbb{R}^n$ be a given smooth map and $F\colon \mathcal{M} \rightrightarrows \mathbb{R}^n$ be a multivalued mapping having closed graph. Fix a point $x \in \mathcal{M}$. Suppose that the mapping $\Phi_x(\cdot) := \mathscr{D}f(x)(\cdot) + (F \circ R_x)(\cdot)$ is metrically regular on the set*

$$V_{r,s}(\Phi_x) := \big\{(u,w) \in T_x\mathcal{M} \times \mathbb{R}^n : \|u\|_x \leqslant r, \operatorname{dist}\big(w, \Phi_x(u)\big) \leqslant s\big\}$$

with a modulus $\kappa > 0$. Consider some positive numbers r', s', σ, λ and λ' such that

$$\begin{cases} \theta = \rho_1^{-1}\rho_2\kappa(\lambda' + \lambda\mu_x) < 1, \\ \rho_1^{-1}\sigma + \rho_1^{-1}\rho_2 r' + \frac{\kappa}{1-\theta}s' < \min\{\delta, r\}, \\ s' \leqslant s, \\ \rho_1^{-1}\rho_2\sigma + \rho_1^{-1}\rho_2^2 r' + \rho_2\frac{\kappa}{1-\theta}s' < \mathbf{r}_x. \end{cases} \tag{20}$$

Pick some $y \in W_x$ so that $d(y,x) \leqslant \sigma$ and the conditions (i) and (ii) of Proposition 3.1 hold. Then the multivalued mapping $\Phi_y(\cdot) := \mathscr{D}f(y)(\cdot) + (F \circ R_y)(\cdot)$ satisfies $\tau \in \operatorname{REG}\big(\Phi_y, V_{r',s'}(\Phi_y)\big)$ for $\tau = \rho_1^{-1}\rho_2\frac{\kappa}{1-\theta}$ and

$$V_{r',s'}(\Phi_y) := \big\{(v,w) \in T_y\mathcal{M} \times \mathbb{R}^n : \|v\|_x \leqslant r', \operatorname{dist}\big(w, \Phi_y(v)\big) \leqslant s'\big\}.$$

Proof. We fix some $(v,w) \in V_{r',s'}(\Phi_y)$ with $\eta = \operatorname{dist}\big(w, \Phi_y(v)\big) > 0$ and look for a point $v^* \in \Phi_y^{-1}(w)$ such that

$$\|v - v^*\|_x \leqslant \tau\eta. \tag{21}$$

For this goal, let us set $v_0 = v$, and $y_0 = R_y(v_0)$. Since $\|v_0\|_y \leqslant r' \leqslant \rho_1^{-1}\rho_2 r' < \delta$, the fact $R \in \mathbf{URC}(\rho_1, \rho_2, \delta, W_x)$ can be used. This yields

$$\begin{aligned} d(R_y(v_0), x) &\leqslant d(R_y(v_0), y) + d(y,x) \leqslant \rho_2\|v_0\|_y + \sigma \leqslant \rho_2 r' + \sigma \\ &< \rho_1^{-1}\rho_2\sigma + \rho_1^{-1}\rho_2^2 r' + \rho_2\frac{\kappa}{1-\theta}s' < \mathbf{r}_x, \end{aligned}$$

which implies $y_0 = R_y(v_0) \in W_x$. Thus, there exists (unique) $u_0 \in \delta\mathbb{B}_x$ with $R_x(u_0) = R_y(v_0)$. By setting

$$w_0 = w + \mathscr{D}f(x)(u_0) - \mathscr{D}f(y)(v_0) \in \mathbb{R}^n,$$

we are going to verify $(u_0, w_0) \in V_{r,s}(\Phi_x)$. Indeed, according to Assumption 2.2

$$\begin{aligned} \|u_0\|_x &\leqslant \rho_1^{-1} d\big(R_x(u_0), x\big) = \rho_1^{-1} d\big(R_y(v_0), x\big) \leqslant \rho_1^{-1}\left(d\big(R_y(v_0), y\big) + d\big(y,x\big)\right) \\ &\leqslant \rho_1^{-1}\Big(\rho_2\|v_0\|_y + \sigma\Big) = \rho_1^{-1}\rho_2\|v_0\|_y + \rho_1^{-1}\sigma, \end{aligned} \tag{22}$$

and therefore

$$\|u_0\|_x \leqslant \rho_1^{-1}\rho_2 r' + \rho_1^{-1}\sigma < \rho_1^{-1}\sigma + \rho_1^{-1}\rho_2 r' + \frac{\kappa}{1-\theta}s' \leqslant r.$$

In addition, one has

$$\begin{aligned}\operatorname{dist}\big(w_0, \Phi_x(u_0)\big) &= \operatorname{dist}\big(w_0 - \mathscr{D}f(x)(u_0), (F \circ R_x)(u_0)\big)\\ &= \operatorname{dist}\big(w - \mathscr{D}f(y)(v_0), (F \circ R_y)(v_0)\big)\\ &= \operatorname{dist}\big(w, \mathscr{D}f(y)(v_0) + (F \circ R_y)(v_0)\big)\\ &= \operatorname{dist}\big(w, \Phi_y(v)\big) \leqslant s' \leqslant s.\end{aligned}$$

So, the inclusion $(u_0, w_0) \in V_{r,s}(\Phi_x)$ is now clear. Since $\kappa \in \operatorname{REG}\big(\Phi_x, V_{r,s}(\Phi_x)\big)$, we get

$$\operatorname{dist}\big(u_0, \Phi_x^{-1}(w_0)\big) \leqslant \kappa \operatorname{dist}\big(w_0, \Phi_x(u_0)\big) = \kappa \operatorname{dist}\big(w, \Phi_y(v)\big) = \kappa\eta.$$

Selecting $u_1 \in \Phi_x^{-1}(w_0)$ that $\|u_0 - u_1\|_x = \operatorname{dist}\big(u_0, \Phi_x^{-1}(w_0)\big)$, we conclude

$$\|u_0 - u_1\|_x \leqslant \kappa\eta.$$

With respect to the inductive step, assume that for some $k \geqslant 1$ we have found $u_1, \dots, u_k$ in $T_x\mathcal{M}$ and $v_0, \dots, v_{k-1}$ in $\delta\mathbb{B}_y \subset T_y\mathcal{M}$ such that

- $R_x(u_j) = R_y(v_j)$, $j = 0, 1, \dots, k-1$;
- $u_{j+1} \in \Phi_x^{-1}(w_j)$ for $w_j = w + \mathscr{D}f(x)(u_j) - \mathscr{D}f(y)(v_j)$ and $j \leqslant k-1$;
- $\|u_j - u_{j+1}\|_x \leqslant \theta^j \|u_j - u_{j+1}\|_x$, $j \leqslant k-1$.

In the tangent space $T_x\mathcal{M}$ the triangle inequality tells us

$$\begin{aligned}\|u_j - u_0\|_x &\leqslant \sum_{i=0}^{j-1} \|u_{i+1} - u_i\|_x \leqslant \sum_{i=0}^{j-1} \theta^i \|u_1 - u_0\|_x = \frac{1-\theta^j}{1-\theta} \|u_1 - u_0\|_x\\ &\leqslant \frac{1}{1-\theta} \|u_1 - u_0\|_x, \quad j = 1, \dots, k.\end{aligned}$$

Taking into account (22), $\|u_0\|_x \leqslant \rho_1^{-1}\rho_2 \|v_0\|_y + \rho_1^{-1}\sigma$, which yields

$$\begin{aligned}\|u_j\|_x &\leqslant \|u_j - u_0\|_x + \|u_0\|_x \leqslant \frac{1}{1-\theta}\kappa\eta + \rho_1^{-1}\rho_2 \|v_0\|_y + \rho_1^{-1}\sigma\\ &\leqslant \rho_1^{-1}\sigma + \rho_1^{-1}\rho_2 r' + \frac{\kappa}{1-\theta}s' < \min\{\delta, r\}, \quad j = 1, \dots, k.\end{aligned}$$

From Assumption 2.2, we find

$$d\big(R_x(u_k), x\big) \leqslant \rho_2 \|u_k\|_x \leqslant \rho_1^{-1}\rho_2\sigma + \rho_1^{-1}\rho_2^2 r' + \rho_2\frac{\kappa}{1-\theta}s' < \mathbf{r}_x.$$

Hence, there exists $v_k \in \delta\mathbb{B}_y$ with $R_y(v_k) = R_x(u_k)$. By setting

$$w_k = w + \mathscr{D}f(x)(u_k) - \mathscr{D}f(y)(v_k) \in \mathbb{R}^n,$$

we claim $(u_k, w_k) \in V_{r,s}(\Phi_x)$. In fact, we have known $\|u_k\|_x \leqslant r$. Besides that, let us now estimate the quantity $\operatorname{dist}\big(w_1, \Phi_x(u_k)\big)$ as follows

$$\begin{aligned}\operatorname{dist}\big(w_k, \Phi_x(u_k)\big) &\leqslant \|w_k - w_{k-1}\|_{\mathbb{R}^n} = \|w - \Psi_{y,x}(v_k) - [w - \Psi_{y,x}(v_{k-1})]\|_{\mathbb{R}^n} \\ &= \|\Psi_{y,x}(v_k) - \Psi_{y,x}(v_{k-1})\|_{\mathbb{R}^n} \leqslant (\lambda' + \lambda\mu_x)\|v_k - v_{k-1}\|_y .\end{aligned}$$

In the preceding estimations, we used the relations

$$\Psi_{y,x}(v_j) = \mathscr{D}f(y)(v_j) - \mathscr{D}f(x)(u_j), \quad j = 0, 1, \ldots, k.$$

On the other hand, we conclude from the fact $R \in \mathbf{URC}(\rho_1, \rho_2, \delta, W_x)$ that

$$\|v_k - v_{k-1}\|_y \leqslant \rho_1^{-1} d\big(R_y(v_k), R_y(v_{k-1})\big) \leqslant \rho_1^{-1}\rho_2 \|u_k - u_{k-1}\|_x .$$

Hence,

$$\begin{aligned}\operatorname{dist}\big(w_k, \Phi_x(u_k)\big) &\leqslant (\lambda' + \lambda\mu_x)\rho_1^{-1}\rho_2 \|u_k - u_{k-1}\|_x \\ &\leqslant (\lambda' + \lambda\mu_x)\rho_1^{-1}\rho_2\theta^{k-1} \|u_1 - u_0\|_x \\ &\leqslant (\lambda' + \lambda\mu_x)\rho_1^{-1}\rho_2\theta^{k-1}\kappa\eta = \theta^k\eta \leqslant \theta^k s' < s.\end{aligned}$$

Invoking the hypothesis $\kappa \in \operatorname{REG}\big(\Phi_x, V_{r,s}(\Phi_x)\big)$, we obtain

$$\begin{aligned}\operatorname{dist}\big(u_k, \Phi_x^{-1}(w_k)\big) &\leqslant \kappa \operatorname{dist}\big(w_k, \Phi_x(u_k)\big) \leqslant \kappa(\lambda' + \lambda\mu_x)\rho_1^{-1}\rho_2\theta^{k-1} \|u_1 - u_0\|_x \\ &= \theta^k \|u_1 - u_0\|_x .\end{aligned}$$

Since the tangent space $T_x\mathcal{M}$ is finite dimensional, the projection $u_{k+1} \in \Phi_x^{-1}(w_k)$ of u_k onto $\Phi_x^{-1}(w_k)$ satisfies

$$\|u_k - u_{k+1}\|_x = \operatorname{dist}\big(u_k, \Phi_x^{-1}(w_k)\big) \leqslant \theta^k \|u_1 - u_0\|_x ,$$

and the construction goes on. So, (u_k) and (v_k) are totally defined. The rest of the proof is similar to the one of Proposition 3.1. □

Remark 3.4. In the case when the set-valued part $F(x)$ does not depend on x (i.e., $F(x) \equiv K$), one has

$$\Phi_y \circ P_{\chi,x,y} = G_{\chi,x,y} + \Phi_x, \tag{23}$$

for any arbitrary geodesic segment $\chi : [0,1] \longrightarrow \mathcal{M}$ such that $\chi(0) = x$ and $\chi(1) = y$. Inasmuch as $P_{\chi,x,y}$ is linearly isometric, for each subset $V \subset T_y\mathcal{M} \times \mathbb{R}^n$ the fact $\tau \in \operatorname{REG}\big(\Phi_y, V\big)$ is equivalent to $\tau \in \operatorname{REG}\big(\Phi_y \circ P_{\chi,x,y}, V'\big)$, in which

$$V' = \big\{(u, w) \in T_x\mathcal{M} \times \mathbb{R}^n : (P_{\chi,x,y}u, w) \in V\big\}.$$

It is well-known from the literature that under condition imposed on Φ_x and Lipschitz modulus of $G_{\chi,x,y}$, the right-hand side of (23) is also metrically regular. Therefore, the suppositions related to the maps $\Sigma_{\chi,y,x}$ in both Propositions 3.1 and 3.3 might be omitted. And we simply need $\|G_{\chi,x,y}\| \leqslant \lambda' < \kappa^{-1}$. Of course, the corresponding parameters as well as the region of metric regularity may be different from the old ones in Propositions 3.1 and 3.3 above.

4. Convergence of Newton-type algorithms

We begin with some crucial hypotheses.

Standing assumptions. Let $L_1, L_2 : [0,+\infty) \longrightarrow [0,+\infty)$ be some nondecreasing continuous function with $L_1(0) = L_2(0) = 0$ and Ω be a nonempty open subset of $\mathcal{M}$. We consider the conditions below:

(A1) Given $x \in \Omega$ and a geodesic $\chi\colon [0,1] \longrightarrow \Omega$ with $\chi(0) = x$. If $(\mathbf{r}_x, C_R(x))$ is a normal pair, and $\chi([0,1]) \subset \mathbb{B}_{\mathcal{M}}(x, \mathbf{r}_x)$ one has $\varSigma_{\chi,y,x} \in \mathscr{L}ip_{L_1(\ell(\chi))}\big(C_R(x)\mathbb{B}_y\big)$, where $y = \chi(1)$ and $\varSigma_{\chi,y,x} = R_x^{-1} \circ R_y - P_{\chi,y,x}$.

(A2) If $\Theta\colon [0,1] \longrightarrow \mathcal{M}$ is a geodesic in Ω joining $z = \Theta(0)$ to $z' = \Theta(1)$ and $G_{\Theta,x,y} := \mathscr{D}f(x) - \mathscr{D}f(y) \circ P_{\Theta,x,y}$, then $\|G_{\Theta,x,y}\| \leqslant L_2(\ell(\Theta))$.

Under the conditions (A1) and (A2) we have the following statement.

Proposition 4.1. *Given $\bar z \in \Omega$ and one R–normal pair $(\bar{\mathbf{r}}, \bar C)$ at $\bar z$. Suppose that*

- *the retraction segment $c(t) = R_z(tu)$ belongs to $\mathbb{B}_{\mathcal{M}}(\bar z, \bar{\mathbf{r}})$,*
- *$\mathbb{B}_{\mathcal{M}}(\bar z, \bar{\mathbf{r}})$ is a convex neighborhood at $\bar z$ (cf. [7, 15]).*

Define

$$\varLambda_R(f,x,u) = \big\|f\big(R_x(u)\big) - [f(x) + \mathscr{D}f(x)(u)]\big\|_{\mathbb{R}^n}, \tag{24}$$

then one has

$$\begin{aligned} &\varLambda_R(f,z,u)) \leqslant \\ &\leqslant \|u\|_z \int_0^1 \Big\{\bar\mu L_1\big(d(z,c(t))\big) + \big[L_1\big(d(z,c(t))\big) + 1\big]\, L_2\big(d(z,c(t))\big)\Big\}\, dt, \end{aligned} \tag{25}$$

where $\bar\mu = \|\mathscr{D}f(\bar z)\|$.

Proof. Denoting $\Theta_t, \chi_t : [0,1] \longrightarrow \mathcal{M}$ the minimizing geodesics with $\Theta_t(0) = z$, $\Theta_t(1) = R_z(tu)$, $\chi_t(0) = \bar z$ and $\chi_t(1) = R_z(tu)$. For simplicity we will use the notation $c_t = c(t)$, $Q^t = (dR_z)_{tu}$, $P^t = P_{\Theta_t,z,c(t)}$, $\bar P^t = P_{\chi_t,\bar z,c(t)}$. Then $\varSigma_{\Theta_t,z,c_t} = R_{c(t)}^{-1} \circ R_z - P^t$ and $G_{\Theta_t,z,c_t} = \mathscr{D}f(x) - \mathscr{D}f(c_t)P^t$. Note that $tu = R_z^{-1}(c(t)) \in \bar C\mathbb{B}_z$, from (A1) and (A2) one gets

$$\big\|d\left(\varSigma_{\Theta_t,z,c_t}\right)_{tu}\big\| \leqslant L_1\big(\ell(\Theta_t)\big), \|G_{\Theta_t,z,c_t}\| \leqslant L_2\big(\ell(\Theta_t)\big).$$

As a consequence of the choice of Θ_t it holds that $\ell(\Theta_t) = d\big(z, c(t)\big)$. Combining this with the following relation

$$d\left(\varSigma_{\Theta_t,z,c_t}\right)_{tu} = \left(dR_{c(t)}^{-1}\right)_{c(t)} \circ (dR_z)_{tu} - P^t = Q^t - P^t,$$

we obtain

$$\big\|Q^t - P^t\big\| \leqslant L_1\big(\ell(\Theta_t)\big) = L_1\left(d\big(z, c(t)\big)\right).$$

Thanks to the expressions $G_{\chi_t,\bar z,c_t} = \mathscr{D}f(\bar z) - \mathscr{D}f(c_t)P_{\chi_t,\bar z,c_t}$, we obtain

$$\begin{aligned}\|\mathscr{D}f(c_t)\| &\leqslant \|-G_{\chi_t,\bar z,c_t} + \mathscr{D}f(\bar z)\| \left\|(P_{\chi_t,\bar z,c_t})^{-1}\right\| \leqslant \|G_{\chi_t,\bar z,c_t}\| + \|\mathscr{D}f(\bar z)\| \\ &\leqslant L_2\big(\ell(\chi_t)\big) + \|\mathscr{D}f(\bar z)\| = L_2\big(d\big(\bar z, c(t)\big)\big) + \bar\mu.\end{aligned}$$

The map $h = f \circ c$ is smooth and

$$\begin{aligned}h'(t) &= \mathscr{D}f(c_t)\left(Q^t(u)\right) = \mathscr{D}f(c_t)\left(Q^t(u) - P^t(u)\right) + \mathscr{D}f(c_t)\left(P^t u\right) \\ &= \mathscr{D}f(c_t)\left(Q^t(u) - P^t(u)\right) - G_{\Theta_t,x,c_t}(u) + \mathscr{D}f(x)(u).\end{aligned}$$

Consequently,

$$\begin{aligned}\Lambda_R(f,z,u) &= \|h(1) - h(0) - \mathscr{D}f(x)(u)\| = \left\|\int_0^1 h'(t)\, dt - \mathscr{D}f(x)(u)\right\| \\ &\leqslant \int_0^1 \left(\, \|\mathscr{D}f(c_t)\| \left\|Q^t - P^t\right\| + \|G_{\Theta_t,x,c_t}\|\,\right) \|u\|_z \; dt.\end{aligned}$$

Hence, the conclusion of Proposition 4.1 follows. □

We are now ready to present the main theorems of this section.

Theorem 4.2 (local convergence). *Given a complete Riemannian manifold $\mathcal{M}$ of dimension m and a retraction $R\colon T\mathcal{M} \longrightarrow \mathcal{M}$. Let $f\colon \mathcal{M} \longrightarrow \mathbb{R}^n$ be a smooth map, and $F\colon \mathcal{M} \rightrightarrows \mathbb{R}^n$ be a closed multivalued mapping. Let $x^* \in \mathcal{M}$ be a solution of problem* (1)*, and $(\mathbf{r}^*, C^*)$ be a R–normal pair at x^*. Assume that*

(i) $R \in \mathbf{URC}\big(\rho_1, \rho_2, C^*, W^*\big)$ *for some* $0 < \rho_1 \leqslant \rho_2$ *and* $W^* = \mathbb{B}_{\mathcal{M}}(x^*, \mathbf{r}^*)$*;*

(ii) *the statements* (A1) *and* (A2) *hold for some real functions L_1, L_2 together with $\Omega = W^*$;*

(iii) *the mapping* $\Phi^* := \mathscr{D}f(x^*)(\cdot) + (F \circ R_{x^*})(\cdot)$ *satisfies* $\tau \in \mathrm{REG}\big(\Phi^*, V^*\big)$ *where* $V^* = r\mathbb{B}_{x_*} \times \mathbb{B}_{\mathbb{R}^n}\big(-f(x^*), s\big)$*;*

(iv) L_1 *and* L_2 *are of class* C^1*;*

(v) $2K^* r \leqslant 1$*, with*

$$\begin{cases} K^* = (\rho_1^{-1}\rho_2)^2\big\{\mu^* K_1(r) + [L_1(r) + 1]K_2(r)\big\}, \\ K_j(r) = \sup_{0\leqslant t\leqslant r}\left|L_j'(t)\right|, j = 1, 2, \\ \mu^* = \|\mathscr{D}f(x^*)\|\,. \end{cases}$$

We also assume in addition that W^* is a convex neighborhood at x^*. Set

$$\sigma = \min\left\{\frac{\rho_1}{\rho_2}r, \rho_1 r, \frac{\rho_1\tau^*}{2K^*r}s, \frac{\rho_1}{1+2K^*r}C^*, \frac{\rho_2}{2\rho_1 K^* r}\mathbf{r}^*\right\} > 0, \tag{26}$$

then for any $x \in \mathbb{B}_{\mathcal{M}}(x^*, \sigma)$, there is a sequence (x_k) generated by scheme (2) which starts at $x_0 = x$ and converges quadratically to x^*. More precisely, one has

$$d(x_{k+1}, x^*) \leqslant \frac{\rho_2}{2\rho_1 r}\big[d(x_k, x^*)\big]^2, k = 0, 1, \ldots \tag{27}$$

Under the suppositions of Theorem 4.2

$$L_j(t) \leqslant L_j(0) + K_j(r)t = K_j(r)t, j = 1, 2 \tag{28}$$

for every t in the interval $[0, r]$.

Proof. Pick $x \in \mathbb{B}_{\mathcal{M}}(x^*, \sigma)$ and we set $x_0 = x$, $\sigma_0 = d(x^*, x_0) < \sigma$. If $x = x^*$, then the proof is trivial. We focus on the case where $x \neq x^*$. To find the next iteration, we shall make use of Proposition 3.1. Let $\chi_0 \colon [0, 1] \longrightarrow W^*$ be a minimizing geodesic which links $x^* = \chi_0(0)$ and $x_0 = \chi_0(1)$. The hypothesis *(ii)* of Theorem 4.2 tells us

$$\varSigma_{\chi_0, x_0, x^*} \in \mathscr{L}ip_{L_1(\sigma_0)}(W^*), \|G_{\chi_0, x^*, x_0}\| \leqslant L_2(\sigma_0).$$

Choose $\bar{u} = 0_{x^*}$, $\bar{p} = -f(x^*)$, $\bar{v} = R_{x_0}^{-1}(x^*) \in C^*\mathbb{B}_{x_0}$ and $\bar{q} = -f(x^*) + \mathscr{D}f(x_0)(\bar{v})$. Using condition *(i)* of Theorem 4.2

$$\|\bar{v}\|_{x_0} \leqslant \rho_1^{-1} d\big(R_{x_0}(\bar{v}), x_0\big) = \rho_1^{-1} d(x^*, x_0) = \rho_1^{-1}\sigma_0 < \rho_1^{-1}\sigma \leqslant r.$$

Since

$$d\big(x^*, R_{x_0}(t\bar{v})\big) = d\big(R_{x_0}(\bar{v}), R_{x_0}(t\bar{v})\big) \leqslant \rho_2 \left\|(1 - t)\bar{v}\right\|_{x_0} \leqslant \rho_2\rho_1^{-1}\sigma \leqslant r$$

whenever $0 \leqslant t \leqslant 1$, the retraction segment $\Theta_0(t) = R_{x_0}(t\bar{v})$ lies into W^*. Moreover, $d\big(x_0, \Theta_0(t)\big) \leqslant \rho_2 t \left\|\bar{v}\right\|_{x_0}$ by assumption *(i)*. Thus, applying Proposition 4.1 for $\bar{z} = x^*$ and $z = x_0$

$$\begin{aligned}
&\varLambda_R(f, x_0, \bar{v}) \leqslant \\
&\leqslant \left\|\bar{v}\right\|_{x_0} \int_0^1 \Big\{\mu^* L_1\left(\rho_2 t \left\|\bar{v}\right\|_{x_0}\right) + \left[L_1\left(\rho_2 t \left\|\bar{v}\right\|_{x_0}\right) + 1\right] L_2\left(\rho_2 t \left\|\bar{v}\right\|_{x_0}\right)\Big\}\, dt \\
&= \frac{1}{\rho_2} \int_0^{\rho_2 \|\bar{v}\|_{x_0}} \Big\{\mu^* L_1(t) + \left[L_1(t) + 1\right] L_2(t)\Big\}\, dt \\
&\leqslant \frac{1}{\rho_2} \int_0^{\rho_1^{-1}\rho_2\sigma_0} \Big\{\mu^* L_1(t) + \left[L_1(t) + 1\right] L_2(t)\Big\}\, dt.
\end{aligned}$$

Taking into account $\rho_1^{-1}\rho_2\sigma_0 < \rho_1^{-1}\rho_2\sigma \leqslant r$, we get $L_j(t) \leqslant K_j(r)t$ for $j = 1, 2$

and $t \in [0, \rho_1^{-1}\rho_2\sigma_0]$. As a result,

$$\begin{aligned} \Lambda_R(f, x_0, \bar{v}) &\leqslant \frac{1}{\rho_2} \int_0^{\rho_1^{-1}\rho_2\sigma_0} \Big\{ \mu^* L_1(t) + [L_1(t)+1]\, L_2(t) \Big\}\, dt \\ &\leqslant \frac{1}{\rho_2} \big\{ \mu^* K_1(r) + [L_1(r)+1] K_2(r) \big\} \int_0^{\rho_1^{-1}\rho_2\sigma_0} t\, dt \\ &= \frac{\rho_2}{\rho_1^2} \big\{ \mu^* K_1(r) + [L_1(r)+1] K_2(r) \big\} \sigma_0^2 = \frac{1}{2\rho_2\tau^*} K^* \sigma_0^2. \end{aligned}$$

Hence

$$\|f(x^*) - f(x_0) - \mathscr{D}f(x_0)(\bar{v})\|_{\mathbb{R}^n} = \Lambda_R(f, x_0, \bar{v}) \leqslant \frac{1}{2\rho_2\tau^*} K^* \sigma_0^2. \tag{29}$$

Let us now test the following evaluations

$$\begin{cases} \theta_0 = \rho_1^{-1}\rho_2\tau^*[L_2(\sigma_0) + L_1(\sigma_0)\mu^*] & < 1, \\ \frac{\tau^*}{1-\theta_0}\left[\frac{1}{2\rho_1\tau^*}K^*\sigma_0^2\right] & < r, \\ \left(1 + \frac{\theta_0}{1-\theta_0}\right)\left[\frac{1}{2\rho_1\tau^*}K^*\sigma_0^2\right] & < s, \\ \rho_1^{-1}\sigma_0 + \frac{\tau^*}{1-\theta_0}\left[\frac{1}{2\rho_1\tau^*}K^*\sigma_0^2\right] & < C^*, \\ \rho_2\frac{\tau^*}{1-\theta_0}\left[\frac{1}{2\rho_1\tau^*}K^*\sigma_0^2\right] & < \mathbf{r}_x. \end{cases} \tag{30}$$

Indeed, the first relation in (30) is valid from the hypothesis, because of

$$\theta_0 \leqslant \rho_1^{-1}\rho_2\tau^*\sigma_0[K_2(r) + K_1(r)\mu^*] \leqslant K^*\sigma_0 \leqslant K^* r \leqslant \frac{1}{2}.$$

Besides that, one has $K^*\sigma_0 \leqslant K^* r$, $\frac{1}{1-\theta_0} \leqslant \frac{1}{1-K^*r} \leqslant 2$ and $\frac{\theta_0}{1-\theta_0} \leqslant \frac{K^*r}{1-K^*r} \leqslant 1$. So, (30) follows from the suppositions (26) and *(v)*.

Returning to the main proof: From (30), we can take some constants $r_0 > 0$ and $s_0 > \frac{1}{2\rho_1\tau^*}K^*\sigma_0^2$ with

$$\begin{cases} \rho_1^{-1}\rho_2\left(1 + \frac{1+\theta_0}{1-\theta_0}\right) r_0 + \frac{\tau^*}{1-\theta_0} s_0 & < r, \\ \left(\rho_1^{-1}\rho_2\right)^2 [L_2(\sigma_0) + L_2(\sigma_0)\mu^*]\left(1 + \frac{1+\theta_0}{1-\theta_0}\right) r_0 + \left(1 + \frac{\theta_0}{1-\theta_0}\right) s_0 & < s, \\ \rho_1^{-1}\sigma_0 + \rho_1^{-1}\rho_2\left(1 + \frac{1+\theta_0}{1-\theta_0}\right) r_0 + \frac{\tau^*}{1-\theta_0} s_0 & < C^*, \\ \rho_2\left[\rho_1^{-1}\rho_2\left(1 + \frac{1+\theta_0}{1-\theta_0}\right) r_0 + \frac{\tau^*}{1-\theta_0} s_0\right] & < \mathbf{r}_x. \end{cases}$$

Now, applying Proposition 3.1 for data $\bar{u}$, $\bar{p}$, $\bar{v}$ and $\bar{q}$ above, the mapping $\Phi_0(\cdot) = \mathscr{D}f(x_0)(\cdot) + (F \circ R_{x_0})(\cdot)$ satisfies the relation $\tau_0 \in \mathrm{REG}\big(\Phi_0, V_0\big)$, where $\tau_0 = \rho_1^{-1}\rho_2\frac{\tau^*}{1-\theta_0}$ and $V_0 = \mathbb{B}_{x_0}(\bar{v}, r_0) \times \mathbb{B}_{\mathbb{R}^n}(\bar{q}, s_0)$. If we pick $w_0 = -f(x_0)$, then

$$w_0 - \bar{q} = f(x^*) - [f(x_0) + \mathscr{D}f(x_0)(\bar{v})].$$

It follows from (29) that $(\bar{v}, w_0) \in V_0$. Thus, we deduce

$$\begin{aligned}\operatorname{dist}\big(\bar{v}, \varPhi_0^{-1}(w_0)\big) &\leqslant \tau_0 \operatorname{dist}\big(w_0, \varPhi_0(\bar{v})\big) \leqslant \tau_0 \left\|w_0 - \bar{q}\right\|_{\mathbb{R}^n} = \varLambda_R(f, x_0, \bar{v})\\ &\leqslant \rho_1^{-1}\rho_2 \frac{\tau^*}{1-\theta_0}\left[\frac{1}{2\rho_2\tau^*}K^*\sigma_0^2\right] \leqslant \rho_1^{-1}K^*\sigma_0^2.\end{aligned}$$

From this, the closed set $\varPhi_0^{-1}(w_0)$ contains a vector v_0 such that

$$\left\|\bar{v} - v_0\right\|_{x_0} = \operatorname{dist}\big(\bar{v}, \varPhi_0^{-1}(w_0)\big) \leqslant \rho_1^{-1}K^*\sigma_0^2.$$

By the triangle inequality in $T_{x_0}\mathcal{M}$, we find

$$\begin{aligned}\left\|v_0\right\|_{x_0} &\leqslant \left\|\bar{v} - v_0\right\|_{x_0} + \left\|\bar{v}\right\|_{x_0} \leqslant \rho_1^{-1}K^*\sigma_0^2 + \rho_1^{-1}\sigma_0\\ &\leqslant \rho_1^{-1}\big(K^*\sigma + 1\big)\sigma_0 < \rho_1^{-1}\big(K^*r + 1\big)\sigma < C^*.\end{aligned}$$

Consequently,

$$\begin{aligned}d(x_1, x^*) = d\big(R_{x_0}(v_0), R_{x_0}(\bar{v})\big) &\leqslant \rho_2 \left\|v_0 - \bar{v}\right\|_{x_0} \leqslant \rho_2\rho_1^{-1}K^*\sigma_0^2\\ &\leqslant \frac{\rho_2}{2\rho_1 r}[d(x_0, x^*)]^2.\end{aligned}$$

Taking into account $\sigma \leqslant \frac{\rho_1}{\rho_2}r$ and $d(x_0, x^*) < \sigma$, we get $\frac{\rho_2}{2\rho_1 r}d(x_0, x^*) \leqslant \frac{1}{2}$. This tells us $x_1 \in \mathbb{B}_{\mathcal{M}}(x^*, \sigma)$. Therefore, we can use x_1 as a new starting point, and go on.

Repeating this procedure, the sequence (x_k) generated by (2) is well-defined for which (27) is fulfilled. Exploiting (27) many times, we arrive at

$$d(x_k, x^*) \leqslant \left[\frac{\rho_2}{2\rho_1 r}d(x_0, x^*)\right]^{2^k-1} d(x_0, x^*) \leqslant \left(\frac{1}{2}\right)^{2^k-1} d(x_0, x^*),$$

which means $x_k \to x^*$ quadratically. Hence, the proof is done. □

In the statement of Theorem 4.2, the behavior of f and F around the solution x^* is a key point. There are many situations in practice where the informations around x^* are not available. So, a global version might be quite useful and significant. The next theorem will be in this sense.

Theorem 4.3 (global analysis). *Let $\mathcal{M}$, R, f and F as in Theorem 4.2. Given a point $x \in \mathcal{M}$ and let $(\mathbf{r}_x, C_x)$ be a R-normal pair at x so that $W_x = \mathbb{B}_{\mathcal{M}}(x, \mathbf{r}_x)$ is a convex neighborhood. We require in addition that $R \in$ $\mathbf{URC}\big(\rho_1, \rho_2, C_x, W_x\big)$ with $0 < \rho_1 \leqslant \rho_2$. Assume the following statements are fulfilled:*

(i) $\Phi(\cdot) := \mathscr{D}f(x)(\cdot) + (F \circ R_x)(\cdot)$ *satisfying* $\tau \in \operatorname{REG}\big(\Phi, V_{r,s}(\Phi)\big)$ *for some* $\tau > 0$, $r > 0$, $s > 0$;

(ii) $\operatorname{dist}\big(0, f(x)+F(x)\big) < \min\left\{s, \frac{1}{2\rho_1\tau}C_x\right\}$;

(iii) *the assumptions* (A1) *and* (A2) *are valid with some* C^1 *functions* L_1, L_2;

(iv) $\alpha = 2K\beta \leqslant 1$, *where*

$$\begin{cases} K = \left(\rho_1^{-1}\rho_2\right)^2 \tau\left\{\mu_x K_1(r) + [L_1(r)+1]K_2(r)\right\} > 0, \\ \beta = \rho_1\tau\operatorname{dist}\big(0, f(x)+F(x)\big), \\ K_1(r) = \sup_{0\leqslant t\leqslant r}\left|L_1'(t)\right|, K_2(r) = \sup_{0\leqslant t\leqslant r}\left|L_2'(t)\right|, \end{cases}$$

and $\mu_x = \|\mathscr{D}f(x)\|$;

(v) $\eta\beta \leqslant \min\left\{\rho_1^2\rho^{-1}r, \rho_1\rho_2^{-1}r, \rho_1^2\rho_2^{-1}C_x, \rho_1^2\rho_2^{-2}\mathbf{r}_x\right\}$, *for* $\eta = \frac{2}{1+\sqrt{1-\alpha}}$.

Then, there exists a solution x^* *of problem* $0 \in f(x)+F(x)$ *such that*

$$d(x, x^*) \leqslant \rho_2\rho_1^{-1}\beta\eta \leqslant r. \tag{31}$$

For such a solution, it has a selection of sequence (x_k) *generated by* (2) *and converges to* x^*. *More precisely, one has*

$$\begin{cases} d(x_k, x^*) \leqslant \rho_2\rho_1^{-1}\left(\frac{4\sqrt{1-\alpha}}{\alpha}\frac{\lambda^{2^k}}{1-\lambda^{2^k}}\right)\beta, & \text{if } \alpha < 1, \\ d(x_k, x^*) \leqslant \rho_2\rho_1^{-1}2^{-k+1}\beta & \text{if } \alpha = 1, \end{cases} \tag{32}$$

where $\lambda = \frac{1-\sqrt{1-\alpha}}{1+\sqrt{1-\alpha}} \leqslant 1$.

It is clear to see that the case $\beta(\tau, x) = 0$ is trivial. From now, we assume $\beta > 0$. In order to prove Theorem 4.3, the next technical lemma will be essential.

Lemma 4.4. *Let us consider the quadratic polynomial* $\omega(t) = \frac{1}{2}Kt^2 - t + \beta$. *The Newton iterative sequence applied for the equation* $\omega(t) = 0$ *reads*

$$t_0 = 0, \quad t_{k+1} = t_k - \omega(t_k)^{-1}\omega(t_k). \tag{33}$$

Then (t_k) *is well-defined, strictly increasing and converges to the smallest root* $t^* = \frac{1-\sqrt{1-\alpha}}{K}$ *of* ω. *Furthermore, the error bounds below are inherited:*

- *if* $\alpha < 1$, *then*

$$\begin{cases} t^* - t_k \leqslant \frac{4\sqrt{1-\alpha}}{\alpha}\frac{\lambda^{2^k}}{1-\lambda^{2^k}}(t_1 - t_0) = \frac{4\sqrt{1-\alpha}}{\alpha}\frac{\lambda^{2^k}}{1-\lambda^{2^k}}\beta, \\ \frac{2(t_{k+1}-t_k)}{1+\sqrt{1+4\lambda^{2^k}\left(1+\lambda^{2^k}\right)^{-2}}} \leqslant t^* - t_k \leqslant \lambda^{2^{k-1}}(t_k - t_{k-1}); \end{cases} \tag{34}$$

- *if* $\alpha = 1$, *then*

$$\begin{cases} t^* - t_k \leqslant 2^{-k+1}(t_1 - t_0) = 2^{-k+1}\beta, \\ 2\left(\sqrt{2}-1\right)(t_{k+1} - t_k) \leqslant t^* - t_k \leqslant t_k - t_{k-1}. \end{cases} \tag{35}$$

Proof of Lemma 4.4. The existence of (t_k) and the error bounds (34), (35) can be found in [11]. For the monotonicity, let us define the function $\varphi : t \in (-\infty, t^*) \longmapsto t - \omega'(t)^{-1}\omega(t)$. By a simple computation $\varphi'(t) = \frac{\omega''(t)\omega(t)}{[\omega'(t)]^2} = K\frac{\omega(t)}{[\omega'(t)]^2}$. The polynomial ω is positive on the interval $I = (-\infty, t^*)$, so $\varphi'(t) > 0$. Since $t_{k+1} = \varphi(t_k)$, we obtain the necessary conclusion by induction. $\square$

Remark 4.5. From Lemma 4.4 we have

$$t_k < t^* = \frac{1-\sqrt{1-\alpha}}{K} = \frac{\alpha}{K(1+\sqrt{1-\alpha})} = \frac{2\beta}{1+\sqrt{1-\alpha}} = \eta\beta \leqslant r \tag{36}$$

for all k. Moreover, if $\alpha = 1$, error bounds in (35) yield

$$\delta_k = t_{k+1} - t_k \leqslant \frac{1}{2\left(\sqrt{2}-1\right)}(t^* - t_k) \leqslant \frac{1}{2\left(\sqrt{2}-1\right)}2^{-k+1}\beta. \tag{37}$$

When $\alpha < 1$, taking into account (34), we deduce

$$\begin{aligned}
\delta_k = t_{k+1} - t_k &\leqslant \frac{1}{2}\left(1 + \sqrt{1 + \frac{4\lambda^{2^k}}{\left(1+\lambda^{2^k}\right)^2}}\right)(t^* - t_k) \\
&\leqslant \frac{1}{2}\left(1 + \sqrt{1 + \frac{4\lambda}{(1+\lambda)^2}}\right)(t^* - t_k) = \frac{1+\sqrt{1+\alpha}}{2}(t^* - t_k) \\
&\leqslant \frac{1+\sqrt{1+\alpha}}{2}\frac{4\sqrt{1-\alpha}}{\alpha}\frac{\lambda^{2^k}}{1-\lambda^{2^k}}\beta \leqslant \frac{1+\sqrt{1+\alpha}}{2}\frac{4\sqrt{1-\alpha}}{\alpha}\frac{\lambda^2}{1-\lambda^2}\beta \\
&= \frac{\alpha\left(1+\sqrt{1+\alpha}\right)}{2\left(1+\sqrt{1-\alpha}\right)^2}\beta.
\end{aligned}$$

In particular,

$$\delta_k \leqslant \frac{1+\sqrt{2}}{2}\beta = \frac{1+\sqrt{2}}{2}\rho_1\tau \operatorname{dist}\bigl(0, f(x) + F(x)\bigr). \tag{38}$$

Now, we begin to prove Theorem 4.3.

Proof of Theorem 4.3. We shall construct by induction the sequence (x_k) satisfying

$$d(x_k, x_{k+1}) \leqslant \rho_2\rho_1^{-1}\left(t_{k+1} - t_k\right), k = 0, 1, \ldots \tag{39}$$

At the initial step, we set $x_0 = x$, $\Phi_0 = \Phi$, $r_0 = r$, $s_0 = s$, $\tau_0 = \tau$. Then $\tau_0 \in \operatorname{REG}\bigl(\Phi_0, V_{r_0,s_0}(\Phi_0)\bigr)$. Put $z_0 = -f(x_0)$, one has

$$\operatorname{dist}\bigl(z_0, \Phi_0(0_{x_0})\bigr) = \operatorname{dist}\bigl(-f(x_0), F(x_0)\bigr) = \operatorname{dist}\bigl(0, f(x) + F(x)\bigr) < s_0,$$

which implies $(0_{x_0}, z_0) \in V_{r_0,s_0}(\Phi_0)$. Since $\tau_0 \in \mathrm{REG}\big(\Phi_0, V_{r_0,s_0}(\Phi_0)\big)$, we get

$$\mathrm{dist}\big(0_{x_0}, \Phi_0^{-1}(z0)\big) \leqslant \tau_0 \mathrm{dist}\big(z_0, \Phi_0(0_{x_0})\big) = \tau \mathrm{dist}\big(0, f(x) + F(x)\big) = \rho_1^{-1}\beta.$$

Hence, there is a vector v_0 in $\Phi_0^{-1}(z_0)$ with

$$\|v_0\|_{x_0} = \mathrm{dist}\big(0_{x_0}, \Phi_0^{-1}(z_0)\big) \leqslant \rho_1^{-1}\beta = \rho_1^{-1}(t_1 - t_0). \tag{40}$$

To continue, we pick $x_1 = R_{x_0}(v_0)$. Observe that $z_0 = -f(x_0) \in \Phi_0(v_0)$, or equivalently,

$$-f(x_0) \in \mathscr{D}f(x_0)(v_0) + (F \circ R_{x_0})(v_0).$$

Therefore, x_1 verifies (2). Furthermore, note that $\eta = \frac{2}{1+\sqrt{1-2K\beta}} > 1$, one has

$$\|v_0\|_{x_0} \leqslant \rho_1^{-1}\beta < \rho_1^{-1}\eta\beta \leqslant \rho_1\rho_2^{-1}C_x \leqslant C_x.$$

By virtue of $R \in \mathbf{URC}\big(\rho_1, \rho_2, C_x, W_x\big)$, we deduce

$$d(x_1, x_0) = d\big(R_{x_0}(v_0), x_0\big) \leqslant \rho_2 \|v_0\|_{x_0} \leqslant \rho_2\rho_1^{-1}\beta = \rho_2\rho_1^{-1}(t_1 - t_0), \tag{41}$$

which means that (39) is valid for $k = 0$.

Passing to the inductive step, given k ($k \geqslant 1$) points $x_1, \ldots, x_k$ in $\mathcal{M}$ and k corresponding tangent vectors $v_0 \in T_{x_0}\mathcal{M}$, ..., $v_{k-1} \in T_{x_{k-1}}\mathcal{M}$ for which the following relations hold

- $x_{j+1} = R_{x_j}(v_j)$, $j = 0, \ldots, k-1$;
- $v_j \in \Phi_j^{-1}(z_j)$, where $z_j = -f(x_j)$, $\Phi_j(\cdot) = \mathscr{D}f(x_j)(\cdot) + \big(F \circ R_{x_j}\big)(\cdot)$, $j \leqslant k-1$;
- $\|v_j\|_{x_j} \leqslant \rho_1^{-1}(t_{j+1} - t_j) := \rho_1^{-1}\delta_j$, $j \leqslant k-1$;
- $d(x_j, x_{j+1}) \leqslant \rho_2\rho_1^{-1}(t_{j+1} - t_j)$, $j \leqslant k-1$.

We also denote $\Phi_k(\cdot) = \mathscr{D}f(x_k)(\cdot) + (F \circ R_{x_k})(\cdot)$ and $z_k = -f(x_k)$. The triangle inequality in $\mathcal{M}$ tells us

$$d(x_j, x) = d(x_j, x_0) \leqslant \sum_{i=0}^{j-1} d(x_i, x_{i+1}) \leqslant \sum_{i=0}^{j-1} \rho_2\rho_1^{-1}(t_{i+1} - t_i) = \rho_2\rho_1^{-1}t_j.$$

So, by (36) $d(x_j, x) \leqslant \rho_2\rho_1^{-1}t_k < \rho_2\rho_1^{-1}t^* = \rho_2\rho_1^{-1}\eta\beta \leqslant \mathbf{r}_x$.

This means $x_j \in W_x$ for $j \leqslant k$. We are now going to estimate the value of quantity $\mathrm{dist}\big(z_k, \Phi_k(0_{x_k})\big)$. From the definition of z_k and Φ_k, it holds that

$$\mathrm{dist}\big(z_k, \Phi_k(0_{x_k})\big) = \mathrm{dist}\big(-f(x_k), F(x_k)\big).$$

But by the induction hypothesis

$$-f(x_{k-1}) = z_{k-1} \in \Phi_{k-1}(v_{k-1}) = \mathscr{D}f(x_{k-1})(v_{k-1}) + F(x_k)$$

we obtain

$$\begin{aligned}\operatorname{dist}\big(z_k, \Phi_k(0_{x_k})\big) &\leqslant \|-f(x_k) - [-f(x_{k-1}) - \mathscr{D}f(x_{k-1})(v_{k-1})]\|_{\mathbb{R}^n} \\ &= \varLambda_R(f, x_{k-1}, v_{k-1}).\end{aligned}$$

This suggests using Proposition 4.1. We can see that for $t \in [0,1]$

$$\begin{aligned}d\big(x, R_{x_{k-1}}(tv_{k-1})\big) &\leqslant d(x, x_{k-1}) + d\big(x_{k-1}, R_{x_{k-1}}(tv_{k-1})\big) \\ &\leqslant \rho_2\rho_1^{-1}t_{k-1} + t\rho_2 \|v_{k-1}\|_{x_{k-1}} \\ &\leqslant \rho_2\rho_1^{-1}t_{k-1} + \rho_2\rho_1^{-1}(t_k - t_{k-1}) \\ &= \rho_2\rho_1^{-1}t_k < \rho_2\rho_1^{-1}t^* = \rho_2\rho_1^{-1}\eta\beta < \mathbf{r}_x,\end{aligned}$$

so that the retraction segment $R_{x_{k-1}}(tv_{k-1})$ $(0 \leqslant t \leqslant 1)$ is in W_x. Invoking Proposition 4.1 and taking into account

$$d\big(x_{k-1}, R_{x_{k-1}}(tv_{k-1})\big) \leqslant t\rho_2 \|v_{k-1}\|_{x_{k-1}} \leqslant \rho_2\rho_1^{-1}\delta_{k-1}t,$$

we find

$$\begin{aligned}&\varLambda_R(f, x_{k-1}, v_{k-1}) \leqslant \\ &\leqslant \rho_1^{-1}\delta_{k-1}\int_0^1 \Big\{\mu_x L_1(\rho_2\rho_1^{-1}\delta_{k-1}t) + [L_1(\rho_2\rho_1^{-1}\delta_{k-1}t) + 1]L_2(\rho_2\rho_1^{-1}\delta_{k-1}t)\Big\}\, dt \\ &= \frac{1}{\rho_2}\int_0^{\rho_2\rho_1^{-1}\delta_{k-1}} \Big\{\mu_x L_1(t) + [L_1(t)+1]L_2(t)\Big\}\, dt \\ &\leqslant \frac{1}{\rho_2}\int_0^{\rho_2\rho_1^{-1}\delta_{k-1}} [\mu_x K_1(r) + [L_1(r)+1]K_2(r)]\, t\, dt \\ &= \frac{\rho_2}{2\rho_1^2}\Big\{\mu_x K_1(r) + [L_1(r)+1]K_2(r)\Big\}\delta_{k-1}^2.\end{aligned}$$

As a result,

$$\operatorname{dist}\big(z_k, \Phi_k(0_{x_k})\big) \leqslant \frac{\rho_2}{2\rho_1^2}\Big\{\mu_x K_1(r) + [L_1(r)+1]K_2(r)\Big\}\delta_{k-1}^2 \leqslant \frac{1}{2\rho_1\tau}K\delta_{k-1}^2. \tag{42}$$

For the goal of applying Proposition 3.3, let us check that

$$\begin{cases} \theta_k = \rho_1^{-1}\rho_2\tau[L_2(t_k) + L_1(t_k)\mu_x] < 1, \\ \rho_2\rho_1^{-2}t_k + \frac{\tau}{1-\theta_k}\left(\frac{1}{2\rho_1\tau}K\delta_{k-1}^2\right) < \min\{C_x, r\}, \\ \frac{1}{2\rho_1\tau}K\delta_{k-1}^2 < s, \\ \rho_2^2\rho_1^{-2}t_k + \rho_2\frac{\tau}{1-\theta_k}\left(\frac{1}{2\rho_1\tau}K\delta_{k-1}^2\right) < \mathbf{r}_x. \end{cases} \tag{43}$$

Indeed, since $t_k < t^* = \eta\beta \leqslant r$ (see (36)), it holds that $L_1(t_k) \leqslant K_1(r)t_k$ and $L_2(t_k) \leqslant K_2(r)t_k$. Thus,

$$\theta_k \leqslant \rho_1^{-1}\rho_2\tau[K_2(r) + K_1(r)\mu_x]t_k \leqslant Kt_k < Kt^* = 1 - \sqrt{1-\alpha} \leqslant 1,$$

and the first relation in (43) is done. For the second one, we expand the polynomial ω with center at t_{k-1}

$$\omega(t_k) = \omega(t_{k-1}) + \omega'(t_{k-1})\delta_{k-1} + \omega''(t_{k-1})\frac{\delta_{k-1}^2}{2} = \frac{1}{2}K\delta_{k-1}^2.$$

Consequently,

$$\begin{aligned}\rho_2\rho_1^{-2}t_k + \frac{\tau}{1-\theta_k}\left(\frac{1}{2\rho_1\tau}K\delta_{k-1}^2\right) &\leqslant \rho_2\rho_1^{-2}t_k + \frac{1}{1-Kt_k}\left(\frac{\rho_2}{2\rho_1^2}K\delta_{k-1}^2\right)\\ &= \rho_2\rho_1^{-2}\big[t_k - \omega'(t_k)^{-1}\omega(t_k)\big] = \rho_1^{-1}\big(t_k+\delta_k\big)\\ &= \rho_2\rho_1^{-2}t_{k+1} < \rho_2\rho_1^{-2}t^* = \rho_2\rho_1^{-2}\eta\beta \leqslant \min\big\{r, C_x\big\}.\end{aligned}$$

Next, we use (38) and combine this relation with the initial condition *(ii)*

$$\begin{aligned}\frac{1}{2\rho_1\tau}K\delta_{k-1}^2 &\leqslant \frac{1}{2\rho_1}\tau^{-1}K\left(\frac{1+\sqrt{2}}{2}\right)^2\beta^2 = \frac{1}{4\rho_1}\left(\frac{1+\sqrt{2}}{2}\right)^2\tau^{-1}\alpha\beta\\ &= \left(\frac{1+\sqrt{2}}{4}\right)^2\alpha\,\mathrm{dist}\big(0, f(x)+F(x)\big) < s.\end{aligned}$$

Finally, since

$$\begin{aligned}\rho_2^2\rho_1^{-2}t_k + \rho_2\frac{\tau}{1-\theta_k}\left(\frac{1}{2\rho_1\tau}K\delta_{k-1}^2\right) &\leqslant \rho_2^2\rho_1^{-2}\left(t_k + \frac{1}{1-Kt_k}\omega(t_k)\right)\\ &= \rho_2^2\rho_1^{-2}(t_k+\delta_k) = \rho_2^2\rho_1^{-2}t_k < \rho_2^2\rho_1^{-2}t^* = \rho_2^2\rho_1^{-2}\eta\beta \leqslant \mathbf{r}_x,\end{aligned}$$

the last inequality in (43) is valid due to the supposition *(v)* of Theorem 4.3. In summary, (43) is completely fulfilled.

According to (43), we can select some positive numbers r_k and s_k such that

$$\begin{cases}\frac{1}{2\rho_1\tau}K\delta_{k-1}^2 < s_k < s,\\ \rho_2\rho_1^{-2}t_k + \rho_1^{-1}\rho_2 r_k + \frac{\tau}{1-\theta_k}s_k < \min\{r, C_x\},\\ \rho_1^{-2}\rho_2^2 t_k + \rho_1^{-1}\rho_2^2 r_k + \rho_2\frac{\tau}{1-\theta_k}s_k < \mathbf{r}_x.\end{cases}$$

Recalling $d(x_k, x) \leqslant \rho_2\rho_1^{-1}t_k$, Proposition 3.3 asserts that $\tau_k \in \mathrm{REG}\big(\Phi_k, V_k\big)$, where

$$\tau_k = \rho_1^{-1}\rho_2\frac{\tau}{1-\theta_k} \leqslant \rho_1^{-1}\rho_2\frac{\tau}{1-Kt_k}, V_k = V_{r_k,s_k}(\Phi_k).$$

Thanks to (42), the inclusion $(0_{x_k}, z_k) \in V_k$ is true. This guarantees for the estimations below

$$\begin{aligned}\operatorname{dist}\big(0_{x_k}, \Phi_k^{-1}(z_k)\big) &\leqslant \tau_k \operatorname{dist}\big(z_k, \Phi_k(0_{x_k})\big)\\ &\leqslant \rho_1^{-1}\rho_2 \frac{\tau}{1-Kt_k}\frac{\rho_2}{2\rho_1^2}\Big\{\mu_x K_1(r) + [L_1(r)+1]K_2(r)\Big\}\delta_{k-1}^2\\ &= \rho_1^{-1}\left(\frac{1}{1-Kt_k}\frac{1}{2}K\delta_{k-1}^2\right) = \rho_1^{-1}\big[-\omega'(t_k)^{-1}\omega(t_k)\big]\\ &= \rho_1^{-1}(t_{k+1}-t_k).\end{aligned}$$

Because $T_{x_k}\mathcal{M}$ is finite dimensional, there exists $v_k \in \Phi_k^{-1}(z_k)$ with

$$\|v_k\|_{x_k} = \operatorname{dist}\big(0_{x_k}, \Phi_k^{-1}(z_k)\big) \leqslant \rho_1^{-1}(t_{k+1}-t_k).$$

And we update now $x_{k+1} = R_{x_k}(v_k)$ as the next iteration.
The selection of v_k tells us $-f(x_k) = z_k \in \Phi_k(v_k)$, which is equivalent to

$$-f(x_k) \in \mathscr{D}f(x_k)(v_k) + \left(F \circ R_{x_k}\right)(v_k).$$

As a consequence, (x_k, v_k) is generated by (2). Using (38) once more, from hypothesis *(ii)* we arrive at

$$\|v_k\|_{x_k} \leqslant \frac{1+\sqrt{2}}{2}\tau\rho_1 \operatorname{dist}\big(0, f(x)+F(x)\big) < C_x,$$

and hence

$$d(x_{k+1}, x_k) = d\left(R_{x_k}(v_k), x_k\right) \leqslant \rho_2 \|v_k\|_{x_k} \leqslant \rho_2\rho_1^{-1}\left(t_{k+1}-t_k\right).$$

The construction of sequence (x_k) is now completed by induction.

Since (t_k) is convergent, and $d(x_k, x_{k+1}) \leqslant \rho_2\rho_1^{-1}\left(t_{k+1}-t_k\right)$, the sequence (x_k) is Cauchy. So, x_k must converge to some x^* in $\mathcal{M}$. The error bounds in (31) and (32) follows from (34), (35) and the following

$$d(x_k, x^*) \leqslant \sum_{j\geqslant k} d(x_j, x_{j+1}) \leqslant \rho_2\rho_1^{-1}\sum_{j\geqslant k}(t_{j+1}-t_j) = \rho_2\rho_1^{-1}\left(t^*-t_k\right).$$

To finish the proof, we claim that x^* solves (1). In fact, it follows from the preceding construction that

$$0 \in f(x_k) + \mathscr{D}f(x_k)(v_k) + F(x_{k+1}). \tag{44}$$

Fixing an index k which is large enough. If $\chi_k : [0,1] \longrightarrow W_x$ is a minimizing geodesic joining $x^* = \chi_k(0)$ to $x_k = \chi_k(1)$, then $\ell(\chi_k) = d(x^*, x_k)$. Using the hypothesis *(iii)* of Theorem 4.3,

$$\|G_{\chi_k, x^*, x_k}\| \leqslant L_2(\ell(\chi_k)) = L_2\big(d(x^*, x_k)\big) \leqslant K_2(r) d(x^*, x_k).$$

Note that $G_{\chi_k,x^*,x_k} = -\mathscr{D}f(x_k)P_{\chi_k,x^*,x_k} + \mathscr{D}f(x^*)$, so

$$\begin{aligned}\|\mathscr{D}f(x_k)\| &= \left\|-G_{\chi_k,x^*,x_k}P^{-1}_{\chi_k,x^*,x_k} + \mathscr{D}f(x^*)P^{-1}_{\chi_k,x^*,x_k}\right\| \\ &\leqslant \|G_{\chi_k,x^*,x_k}\| \left\|P^{-1}_{\chi_k,x^*,x_k}\right\| + \|\mathscr{D}f(x^*)\| \left\|P^{-1}_{\chi_k,x^*,x_k}\right\| \\ &= \|G_{\chi_k,x^*,x_k}\| + \|\mathscr{D}f(x^*)\| \leqslant K_2(r)d(x^*,x_k) + \|\mathscr{D}f(x^*)\| .\end{aligned}$$

Thus, we obtain

$$\begin{aligned}\|\mathscr{D}f(x_k)(v_k)\|_{\mathbb{R}^n} &\leqslant \|\mathscr{D}f(x_k)\| \, \|v_k\|_{x_k} \\ &\leqslant \left[K_2(r)d(x^*,x_k) + \|\mathscr{D}f(x^*)\|\right]\rho_1^{-1}\left(t_{k+1} - t_k\right).\end{aligned}$$

Hence, $\mathscr{D}f(x_k)(v_k)$ converges to 0 in $\mathbb{R}^n$ as $k \to \infty$. Finally, by passing to the limit in (44), we get $0 \in f(x^*) + F(x^*)$. □

Remark 4.6. In the proofs of Theorems 4.2 and 4.3, the existence and the convergence of a Newton sequence (x_k) require the assumptions (A1) and (A2) which are imposed on covariant derivative $\mathscr{D}f$ as well as the retraction R. Assumption (A2) usually concerns with the so-called "Lipschitz continuity" for $\mathscr{D}f$ (see, e.g., [10, 9]). If the higher order covariant derivatives $\mathscr{D}^k f$ (cf. [6, 21]) are involved in the hypotheses of those theorems, we obtain (A2) from the informations on $\mathscr{D}^k f$. Therefore, we can achieve some new versions of Theorems 4.2 and 4.3 under conditions of type Kantorovich and/or Smale.
On the other hand, if we restrict our consideration on the case $F(x) \equiv C$, and $R = \exp$ is the exponential, then Theorems 4.2 and 4.3 can be slightly improved. Indeed, as noticed in Remark 3.4, the stability of metric regularity property can be obtained by using only (A2). Besides that, when $R = \exp$, the essential Proposition 4.1 holds with $L_1(t) \equiv 0$ and $L_2(t)$ is the function in (A2) since $P_{\chi,\chi(0),\chi(t)}\left(\chi'(0)\right) = \chi'(t)$ for any arbitrary geodesic χ. Consequently, Kantorovich-type versions of Theorems 4.2 and 4.3 like in [21] can be recovered.
Now, we keep assuming $F(x) \equiv C$, while (A2) is replaced by the following

(A3) For each x and y in Ω so that $R_x^{-1}(y)$ is a singleton, one has

$$\left\|G^R_{x,y}\right\| = \left\|\mathscr{D}f\left(x\right) - \mathscr{D}f\left(y\right)T^R_{x,y}\right\| \leqslant L\left(d\big(x,y\big)\right). \tag{45}$$

where $T^R_{x,y}$ is the vector transport defined by

$$T^R_{x,y}(v) = \left(dR_x\right)_u(v) = \frac{d}{dt}\big\{R_x(u+tv)\big\}\bigg|_{t=0}, u = R_x^{-1}(y).$$

We refer to [1, Section 8.1] and [17] for more details about transportation of the form $T^R_{x,y}$. Then, by considering f along the retraction curve $\chi(t) = R_z(tu)$, the estimation (25) in Proposition 4.1 simply reads

$$\varLambda_R(f,z,u) \leqslant \|u\|_z \int_0^1 L\left(d\big(z,\chi(t)\big)\right)dt. \tag{46}$$

So, by removing all conditions related to L_1 and putting L instead of L_2 in both Theorems 4.2 and 4.3, we obtain Kantorovich-type results for algorithm (2), which might be viewed as extensions of the corresponding ones studied in [21].

5. Concluding Remarks

In this paper, we have studied a Newton-type algorithm for solving generalized equations in a finite dimensional Riemannian manifold. We provided sufficient conditions that guaranteed the existence and the quadratic convergence of the sequence generated by the extended Newton's method. This work can be considered one of the first that studied Newton's type algorithms for set-valued inclusions of the form (1), where the metric regularity concept plays an important role. Due to the large number of applications in the calculus of variations or PDE's e.g., it would be interesting to consider the infinite dimensional Riemannian manifolds modeled on a Banach or a Hilbert space.

References

[1] P.-A. Absil, R. Mahony, R. Sepulchre: Optimization Algorithms on Matrix Manifolds, Princeton University Press (2008).

[2] R. L. Adler, J. Dedieu, J. Y. Margulies, M. Martens, M. Shub: Newton's method on Riemannian manifolds and a geometric model for the human spine, IMA Journal of Numerical Analysis 22(3) (2002) 359–390.

[3] S. Adly, R. Cibulka, H. V. Ngai: Newton's method for solving inclusions using set-valued approximations, SIAM Journal on Optimization 25(1) (2015) 159–184.

[4] S. Adly, H. V. Ngai, V. V. Nguyen: Newton's method for solving generalized equation: Kantorovich's and Smale's approaches, J. Math. Analysis Appl. 439(1) (2016) 396–418.

[5] F. Alvarez, J. Bolte, J. Munier: A unifying local convergence result for Newton's method in Riemannian manifolds, Foundations of Computational Mathematics 8(2) (2008) 197–226.

[6] J.-P. Dedieu, P. Priouret, G. Malajovich: Newton's method on Riemannian manifolds: covariant alpha theory, IMA Journal of Numerical Analysis 23(3) (20039 395–419.

[7] M. P. do Carmo: Riemannian Geometry, Mathematics: Theory and Applications, Birkhäuser (1992).

[8] A. L. Dontchev, R. T. Rockafellar: Implicit Functions and Solution Mappings: A View from Variational Analysis, Springer Monographs in Mathematics, Springer, New York (2009).

[9] O. P. Ferreira, R. C. M. Silva: Local convergence of Newton's method under a majorant condition in Riemannian manifolds, IMA Journal of Numerical Analysis 32(4) (2012) 1696–1713.

[10] O. P. Ferreira, B. F. Svaiter: Kantorovich's theorem on Newton's method in Riemannian manifolds, Journal of Complexity 18(1) (2002) 304–329.

[11] W. B. Gragg, R. A. Tapia: Optimal error bounds for the Newton-Kantorovich theorem, SIAM Journal on Numerical Analysis 11(1) (1974) 10–13.

[12] S. Helgason: Differential Geometry, Lie Groups, and Symmetric Spaces, Pure and Applied Mathematics, Series of Monographs and Textbooks, Academic Press (1978).

[13] A. D. Ioffe: Metric regularity and subdifferential calculus, Russian Math. Surveys 55(3) (2000) 501–558.

[14] S. Lang: Fundamentals of Differential Geometry, Graduate Texts in Mathematics 191, Springer, New York (1999).

[15] J. M. Lee: Riemannian manifolds: An Introduction to Curvature, Graduate Texts in Mathematics 176, Springer, New York (1997).

[16] A. Papadopoulos: Metric Spaces, Convexity and Nonpositive Curvature, IRMA Lectures in Mathematics and Theoretical Physics 6, European Mathematical Society (2005).

[17] W. Ring, B. Wirth: Optimization methods on Riemannian manifolds and their application to shape space, SIAM Journal on Optimization 22(2) (2012) 596–627.

[18] T. Sakai: Riemannian Geometry, Translations of Mathematical Monographs 149, American Mathematical Society (1996).

[19] S. T. Smith: Optimization techniques on Riemannian manifolds, in: Hamiltonian and Gradient Flows, Algorithms and Control, Fields Institute Communications 3, American Mathematical Society (1994) 113–136.

[20] C. Udrişte: Convex Functions and Optimization Methods on Riemannian Manifolds, Mathematics and Its Applications 297, Springer Netherlands (1994).

[21] J.-H. Wang, S. Huang, C. Li: Extended Newton's method for mappings on Riemannian manifolds with values in a cone, Taiwanese Journal of Mathematics 13(2B) (2009) 633–656.