A flexible framework for cubic regularization algorithms for non-convex optimization in function space

Anton Schiela

November 3, 2017

Abstract

We propose a cubic regularization algorithm that is constructed to deal with non-convex minimization problems in function space. It allows for a flexible choice of the regularization term and thus accounts for the fact that in such problems one often has to deal with more than one norm. Global and local convergence results are established in a general framework.

AMS MSC 2000: 49M37, 90C30

Keywords: non-convex optimization, optimization in function space, cubic regularization

1 Introduction

In broad terms, non-linear optimization algorithms rely on two types of models. A *local model* q_x at a given point x of the functional f to be minimized that can be treated with techniques of linear algebra, and a *rough parametrized model* for the remaining difference between q_x and the functional, i.e., the local error $f - q_x$. Usually such an error model is based on a norm $\|\cdot\|$. In classical trust region methods (cf. e.g. [4]) the error model is 0 inside a ball of varying radius around the current iterate (the trust-region) and $+\infty$ outside. In cubic regularization methods the error model is chosen according to the assumption that $f - q_x$ is of third order. Thus, a scaling of $\|\cdot\|^3$ by an algorithmic parameter (called ω in the following) is taken as a model for the error.

The reason for introducing such an error model (in contrast to a line-search approach) is the wish to transfer information about the $f - q_x$ attained at sampling points (i.e., at trial steps) to a whole neighborhood of the current iterate. The implicit assumption behind this reasoning is that the error indeed behaves more or less isotropically with respect to the chosen norm. Of course, this cannot be guaranteed in general, but in those cases where the error model predicts the actual error well, we expect a stable and efficient behavior of our algorithm. It is thus important, in particular for large scale problems, to choose the error model, and thus the underlying norm, carefully. Many state-of-the-art optimization codes incorporate this idea by allowing the use of a preconditioner for the problem, which in turn defines a problem related norm $\|\cdot\|$.

In this paper we consider a cubic error model. This idea is not new, and, to the best knowledge of the author, has first been proposed by Griewank [9] in an unpublished technical report. Independently, Weiser, Deuflhard, and Erdmann [20] proposed a cubic regularization in an attempt to generalize the works [6, 7] on convex optimization to the non-convex case. Focus was laid on the construction of estimates for the third order remainder term. Even more recently Cartis, Gould, and Toint proposed an algorithmic framework, similar to trust-region methods, but with a cubic regularization term, and provided detailed first and second order convergence analysis [2] and a complexity analysis [3]. Common idea of all these methods is the cubic regularization, but apart from this basic idea the proposed methods differ significantly.

Compared to trust-region methods cubic regularization methods have some appealing features which caused recent interest in them. First, as already indicated, the cubic term can straightforwardly be interpreted as a model for the local error $f - q_x$ as long as the model is second order consistent. At least in the smooth case Taylor expansion shows that this difference is of third order. This allows for an elegant update of the scaling parameter ω (cf. (26), below). Second, [3] have shown that cubic regularization methods exhibit better worst case complexity bounds than trust-region methods. Nevertheless, both classes of methods have many things in common, and even admit a unified convergence theory, as shown in [16].

Many large scale optimization problems are discretizations of problems with partial differential equations (PDEs). These may comprise problems of energy minimization, such as nonlinear elliptic problems, or problems from optimal control of PDEs. If one wants to apply cubic regularization methods in this setting, one is naturally led to versions that work in function space, and the need for a convergence theory in function space arises. In trust-region methods this topic is well understood and several works have been published that pursue this line of thought (cf. e.g. [15, 11, 19, 18]). Concerning cubic regularization a global convergence theory as performed in [2] may be lifted to function space in a relatively straightforward way. If f is continuously differentiable and the second order term is bounded above and below with respect to a norm $\|\cdot\|$, one ends up with a globally convergent cubic regularization method that is equipped with the error model $f - q_x \sim \|\cdot\|^3$. Thus, the error model is based on the same norm that is used to define directions of steepest descent and thus Cauchy points, which are needed to define acceptable search directions.

However, analysis reveals, as we will sketch in Section 2.2, that there are large classes of infinite dimensional problems where the behaviour of $f - q_x$ is not described adequately by the norm $\|\cdot\|$. The error term would appear highly anisotropic, when compared to the model $\|\cdot\|^3$. Rather, one can find different third order models, which we call $R_x(\cdot)$, that are much better suited as a model: $f - q_x \sim R_x(\cdot)$. This analytic structure calls for decoupling the norm $\|\cdot\|$ used to define directions of steepest descent, and the model $R_x(\cdot)$ for third order terms. In cubic regularization methods this can be realized easily. Of course, $R_x(\cdot)$ cannot be chosen arbitrarily, but has to be compatible with the problem. The aim of this paper is to find and analyse such compatibility conditions on the choice of $R_x(\cdot)$. Our framework employs two norms which are used to formulate the required regularity assumptions on f. Then conditions depending on these norms are imposed on $R_x(\cdot)$ which allow for a global and a local convergence analysis.

In particular, we will introduce our flexible analytic framework in Section 2. In particular, we discuss some examples to illustrate and motivate the abstract concepts. In Section 3 we develop our algorithmic framework. It resembles in a couple of points the classical trust region-like algorithms with the usual fraction of model decrease acceptance criterion and a fraction of Cauchy decrease condition. However, the latter condition has to be modified to take into account non-equivalence of norms. Also here it was our aim to leave as much flexibility for concrete implementations of algorithms, concerning updates of regularization parameters and computation of steps. Within this framework we show in Section 4 global and local convergence results.

We emphasize that the focus of this paper is to establish a framework for algorithms, rather than propose concrete implementations. In particular, we will only briefly address the issue of step computation and rather provide minimal requirements that acceptable steps have to fulfill. There are excellent candidates available in the literature (cf. e.g. [4, Chapt. 5] in a general context and [9, 20, 2] for cubic regularization) that certainly can be used within our framework. We would like to postpone the treatment of the arising algorithmic issues and also numerical testings to a future publication.

2 Functional analytic framework

Consider for a given function $f : X \to \mathbb{R}$ on a linear space X the minimization problem

$$\min_{x \in X} f(x).$$

Suppose that we can compute for each $x \in X$ a quadratic model, consisting of a linear functional $f'_x : X \to \mathbb{R}$ and a bilinear form $H_x : X \times X \to \mathbb{R}$:

$$q_x(\delta x) := f(x) + f'_x \delta x + \frac{1}{2} H_x(\delta x, \delta x). \tag{1}$$

Further, let us denote the error of the quadratic model as follows

$$w_x(\delta x) := f(x + \delta x) - q_x(\delta x) \tag{2}$$

with the help of a function $w_x : X \to \mathbb{R}$. Later, we will impose various smoothness conditions on f, i.e., make assumptions on the limiting behavior of w_x for small δx. Depending on the smoothness of the problem, $w_x(\lambda \delta x)$ may be of higher order locally, such as $o(\lambda)$, $o(\lambda^2)$ or even $O(\lambda^3)$ as $\lambda \to 0$. The last case motivates the construction of the following cubic model for f with parameter $\omega > 0$:

$$f(x + \delta x) - f(x) \approx m_x^\omega(\delta x) := f'_x \delta x + \frac{1}{2} H_x(\delta x, \delta x) + \frac{\omega}{6} R_x(\delta x). \tag{3}$$

Here R_x is a functional, which is homogenous of order 3:

$$R_x(\lambda \delta x) = |\lambda|^3 R_x(\delta x) \qquad \forall \lambda \in \mathbb{R} \tag{4}$$

and positive:

$$R_x(\delta x) > 0 \quad \forall \delta x \neq 0. \tag{5}$$

In (3) the parameter $\omega > 0$ is updated adaptively during the course of the algorithm in order to globalize the method. Comparison of (2) and (3) yields that $(\omega/6)R_x$ can be seen as a model for w_x. The subscript in R_x indicates that R_x may vary from point to point, as long as the conditions imposed below hold independently of x.

2.1 Assumptions for global and local convergence

If X is equipped with the norm $\|\cdot\|$, the classic cubic regularization method uses $R_x(\delta x) := \|\delta x\|^3$. However, in most function space problems an adequate analysis requires the use of several non-equivalent norms. There are a couple of different issues, which each on its own may require a separate choice. This is why we aim for a theoretical framework that is flexible with respect to choosing more than one norm.

Assumptions for global convergence. Let us collect the following set of assumptions for later reference, which are needed to show global convergence, i.e., $\liminf_{k\to\infty} \|f'_{x_k}\| = 0$ for our algorithm.

(i) Let $(X, \|\cdot\|)$ be a Hilbert space and X^* its normed dual. The *primary norm* $\|\cdot\|$ on X has to be strong enough that f is continuously differentiable on X. This means in particular that we have for each $x \in X$ the property:

$$\|f'_x\| := \sup_{\|\delta x\|=1} |f'_x \delta x| < \infty, \tag{6}$$

and, moreover that $x_k \to x_*$ in $(X, \|\cdot\|)$ implies $f'_{x_k} \to f'_{x_*}$ in X^*, i.e., w.r.t the norm (6).

(ii) The *secondary norm* $|\cdot|$ of X is used to describe possible non-convexity of the quadratic model q_x. We assume that $|\cdot|$ is weaker than $\|\cdot\|$:

$$\exists C < \infty : |v| \le C\|v\| \quad \forall v \in X.$$

With the help of our two norms we impose a condition of Gårding-type:

$$\exists \gamma > -\infty,\ \Gamma < \infty : \quad \gamma |v|^2 \le H_x(v,v) \le \Gamma \|v\|^2 \quad \forall v \in X. \tag{7}$$

He do *not* assume completeness of $(X, |\cdot|)$, which allows to choose $|\cdot|$ strictly weaker than $\|\cdot\|$. Hence, H_x is assumed to be bounded below in a weaker norm than it is bounded above. Similar conditions appear, for example, in the theory of linear monotone operators (cf. e.g. [22, Chap. 22]). In the next section we will discuss some examples, where this condition is fulfilled.

(iii) The main purpose of R_x is to compensate the possible non-convexity of the quadratic models and to model the remainder term. Thus, we impose the following flexible boundedness and coercivity condition (without a constant in the left inequality for simplicity):

$$|v|^3 \le R_x(v) \le C\|v\|^3 \qquad \forall v \in X. \tag{8}$$

Among these assumptions the only *standing assumptions* we will use is existence of f'_x in X^*. All other assumptions will be referenced later, when needed.

Lemma 2.1. *Consider a sequence x_k in X that converges to some limit x_* and sequence $\delta v_k \to 0$ in X. Assume that f is continuously differentiable and* (7) *holds. Then for the remainder term, defined in* (2) *we conclude*

$$\lim_{k\to\infty} \frac{w_{x_k}(\delta v_k)}{\|\delta v_k\|} = 0. \tag{9}$$

Proof. Concerning (9), we conclude from a standard result of analysis (cf. e.g. [14, Thm. 25.23], an application of the fundamental theorem of calculus) that

$$\lim_{k\to\infty} \|\delta v_k\| = 0 \quad \Rightarrow \quad \lim_{k\to\infty} \frac{f(x_k+\delta v_k) - f(x_k) - f'_{x_k}\delta v_k}{\|\delta v_k\|} = 0.$$

Moreover, if $\delta v_k \to 0$, then (7) implies $H_{x_k}(\delta v_k, \delta v_k) = o(\|\delta v_k\|)$. Hence

$$w_{x_k}(\delta v_k) = f(x_k+\delta v_k) - q_{x_k}(\delta v_k) = f(x_k+\delta v_k) - f(x_k) - f'_{x_k}\delta v_k + o(\delta v_k).$$

Combining these two results yields (9).

□

Our global convergence results will be proved by contradiction. The following lemma, which uses reflexivity of the Hilbert space X, will serve as a key facility to obtain this contradiction (see Theorem 4.3).

Lemma 2.2. *Let $x_k \in X$ be a sequence such that $|x_k| \to 0$ and $\|x_k\|$ is bounded. Then by reflexivity of X:*

$$x_k \rightharpoonup 0 \text{ weakly in } (X, \|\cdot\|).$$

Proof. Since $(X, \|\cdot\|)$ is reflexive, x_k has a weakly convergent subsequence, say $x_{k_j} \rightharpoonup x_*$. Since $|x_{k_j}| \to 0$ we conclude that for each $\varepsilon > 0$, x_{k_j} is eventually contained in a ball of $|\cdot|$-radius ε, and thus also x_* has to be contained in that ball. It follows that $x_* = 0$. This also shows that every possible weak accumulation point of our sequence is 0, so by a standard argument the whole sequence converges weakly to 0. □

Assumptions for fast local convergence. If we want to show fast local convergence we need the following additional assumptions, which strengthen assumption (i) and (ii) from the above list:

$(\mathrm{i})_{\mathrm{loc}}$ Setting $\delta x_k = x_{k+1} - x_k$ we need a second order approximation error estimate in (2):

$$\lim_{\|x_k - x_*\| \to 0} \frac{w_{x_k}(\delta x_k)}{\|\delta x_k\|^2} = 0, \tag{10}$$

close to a local minimizer, which is fulfilled in particular, if f is twice continuously differentiable and $H_x = f''_x$.

$(\mathrm{ii})_{\mathrm{loc}}$ Locally, we have to impose stronger assumptions on H_x. Close to a minimizer we assume in addition to (7) ellipticity of H_x with respect to the strong norm $\|\cdot\|$:

$$\exists \gamma > 0: \quad \gamma \|\delta x\|^2 \leq H_x(\delta x, \delta x). \tag{11}$$

2.2 Examples

To get a feeling for the peculiarities of the class of problems that fit into our framework, we will discuss a couple of typical examples. The purpose of Section 2.2.1 is to show that several norms appear naturally in non-convex optimization problems in function space, while the Section 2.2.2 illustrates, how R_x can be chosen, and why the additional flexibility of our new framework is beneficial. In Section 2.2.3 a further important class of problems is introduced that fall into our framework.

2.2.1 Two illustrative examples

The following well known simple examples serve as an illustration, why the choice of two norms in (7) is quite natural in infinite dimensional optimization.

In contrast to finite dimensional problems, where existence of minimizers of lower semi-continuous functions is obtained by the classical theorem of Weierstrass, infinite dimensional problems are notoriously hard to analyse. The main reason is the *lack of compactness* of closed and bounded sets in infinite dimensions. By turning to weak lower semi-continuity existence of minimizers can often still be attained, but this property only holds under certain convexity or compactness assumptions on the problem, or subtle combinations of both. As a rule, non-convex problems are tractable, as long as there are other, additional compactness results available, such as compact Sobolev embeddings, e.g., $E : H^1_0 \hookrightarrow L_p$ (cf. e.g. [1]). The interested reader is referred to the textbooks [8, 5] for a thorough introduction.

Let us consider the following two functionals:

$$\phi(v) := \int_0^1 \frac{1}{2} v(t)^2 \, dt, \qquad \psi(v) := \int_0^1 (v(t)^2 - 1)^2 \, dt.$$

We observe that $\phi : L_2(0,1) \to \mathbb{R}$ is convex, while $\psi : L_4(0,1) \to \mathbb{R}$ is non-convex (the integrand has minima at ± 1) and both functionals are non-negative.

The following minimization problem, which involves the first derivative $\dot{u} = du/dt$ is well defined in $H_0^1(0,1)$:

$$\min_{u \in H_0^1(0,1)} f(u) := \phi(\dot{u}) + \psi(Eu).$$

The non-convex part of this problem appears together with the compact Sobolev embedding E, which suffices to show weak lower semi-continutiy of f and conclude existence of minimizers.

In contrast, the following minimization problem, commonly attributed to Bolza, which is well defined on $W_0^{1,4}(0,1)$:

$$\min_{u \in W_0^{1,4}(0,1)} \tilde{f}(u) := \phi(Eu) + \psi(\dot{u})$$

does not admit a global minimizer as is well known.

In view of (7), let us consider the (formal) second derivatives of f and $\tilde{f}$, for example at $u_* = 0$:

$$f''_{u_*}(\delta u, \delta u) = \int_0^1 (\delta u')^2 - 4\delta u^2 \, dt \quad \Rightarrow \quad -4\|\delta u\|_{L_2}^2 \le f''_{u_*}(\delta u, \delta u) \le \|\delta u\|_{H^1}^2$$

$$\tilde{f}''_{u_*}(\delta u, \delta u) = \int_0^1 \delta u^2 - 4(\delta u')^2 \, dt \quad \Rightarrow \quad -4\|\delta u\|_{H^1}^2 \le \tilde{f}''_{u_*}(\delta u, \delta u) \le \|\delta u\|_{H^1}^2.$$

We observe that (7) is fulfilled with two different norms for f with the weaker norm measuring the non-convexity. We may set $|\cdot| = \|\cdot\|_{L_2}$ and $\|\cdot\| = \|\cdot\|_{H^1}$. For $\tilde{f}$ the choice of a weaker norm for the lower bound is not possible.

The bottom line is that compactness, an important principle on which existence of minimizers for non-convex problems rests, is closely related to the presence of two norms in (7), where the lower bound is measured in a weaker norm than the upper bound.

2.2.2 Semi-linear elliptic PDEs

In the following let $\Omega \subset \mathbb{R}^d$ $(1 \le d \le 3)$ be a smoothly bounded open domain. Further, let $H_0^1(\Omega)$ be the usual Sobolev space of weakly differentiable functions on Ω with zero boundary conditions. By the Sobolev embedding theorem there is a continuous embedding $H_0^1(\Omega) \hookrightarrow L_6(\Omega)$ for $d \le 3$. Further, denote by $v \cdot w$ the euclidean scalar product of $v, w \in \mathbb{R}^d$. We denote the spatial variable by $s \in \mathbb{R}^d$.

As a prototypical example we consider the following energy functional of a semi-linear elliptic PDE $f : H_0^1(\Omega) \to \mathbb{R}$:

$$f(u) := \int_\Omega \frac{1}{2} \nabla u(s) \cdot \nabla u(s) + a(u(s), s) \, ds.$$

Here $a : \mathbb{R} \times \Omega \to \mathbb{R}$ is a Carathéodory function that is twice continuously differentiable with respect to u. For more information on the theory of semi-linear PDEs and wider classed of problems we refer the reader to the textbooks [23, 12], variational approaches can be found e.g. in [8, 21, 5].

In the following discussion we will consider a second order model q_u, setting $H_u = f''_u$. We are looking for the solution of the minimization problem

$$\min_{u\in H_0^1(\Omega)} f(u).$$

Its (formal) first and second derivatives are given by:

$$\begin{aligned} f'_u\delta u &= \int_\Omega \nabla u \cdot \nabla \delta u + \frac{\partial}{\partial u} a(u,s)\delta u \, ds \\ f''_u(\delta u_1, \delta u_2) &= \int_\Omega \nabla \delta u_1 \cdot \nabla \delta u_2 + \frac{\partial^2}{\partial u^2} a(u,s)\delta u_1 \delta u_2 \, ds. \end{aligned}$$

Let us analyse these functionals. We may assume that $u \in H_0^1(\Omega)$, which implies that

$$\left| \int_\Omega \nabla u \cdot \nabla \delta u \, ds \right| \le c(u)\|\delta u\|_{H^1}$$

and similarly

$$0 \le \int_\Omega \nabla \delta u \cdot \nabla \delta u \, dx \le \|\delta u\|_{H^1}^2.$$

Let us now assume for simplicity that $\frac{\partial}{\partial u} a(u,\cdot) \in L_2$ and $\frac{\partial^2}{\partial u^2} a(u,\cdot) \in L_\infty$. We obtain the following estimates for the second parts of the derivatives via the Hölder inequality:

$$\begin{aligned} \left| \int_\Omega \frac{\partial}{\partial u} a(u,s)\delta u \, ds \right| &\le \|\frac{\partial}{\partial u} a(u,\cdot)\|_{L_2} \|\delta u\|_{L_2} \le c(u)\|\delta u\|_{H^1} \\ \left| \int_\Omega \frac{\partial^2}{\partial u^2} a(u,s)\delta u_1 \delta u_2 \, ds \right| &\le \|\frac{\partial^2}{\partial u^2} a(u,\cdot)\|_{L_\infty} \|\delta u_1\|_{L_2} \|\delta u_2\|_{L_2} \end{aligned}$$

Taking these estimates together, we obtain the following results:

$$\begin{aligned} |f'_u\delta u| &\le c(u)\|\delta u\|_{H^1} \\ c_0(u)\|\delta u\|_{L_2}^2 \le f''_u(\delta u, \delta u) &\le c_1(u)\|\delta u\|_{H^1}^2, \end{aligned}$$

where $c_0(u) > -\infty$ may be negative, and $c(u), c_1(u) < +\infty$ are positive. Our first observation is that f' and f'' can be bounded (from above) via a strong norm $\|\cdot\| := \|\cdot\|_{H^1}$, while it only takes a weaker norm $|\cdot| := \|\cdot\|_{L_2}$ to formulate a lower bound on f''.

Thus, the condition (8) on R_u reads in this example

$$\|\delta u\|_{L_2}^3 \le R_u(\delta u) \le C\|\delta u\|_{H^1}^3.$$

Let us take into account that $R_u(\delta u)$ should model the qualitative behaviour of the difference w_u of f and its quadratic model q_u (cf. (2)). Under the assumption that $\frac{\partial^2}{\partial u^2} a$ is Lipschitz w.r.t u with Lipschitz-constant ω, repeated application of the fundamental theorem of calculus yields

$$\begin{aligned} |w_u(\delta u)| = |f(u+\delta u) - q_u(\delta u)| &= \left| f(u+\delta u) - f(u) - f'_u\delta u - \frac{1}{2} f''_u(\delta u, \delta u) \right| \\ &= \left| \int_\Omega a(u+\delta u, s) - a(u,s) - \frac{\partial}{\partial u} a(u,s)\delta u - \frac{1}{2}\frac{\partial^2}{\partial u^2} a(u,s)\delta u^2 \, ds \right| \\ &\le \frac{\omega}{6} \int_\Omega |\delta u|^3 \, ds = \frac{\omega}{6}\|\delta u\|_{L_3}^3. \end{aligned}$$

In particular, we observe that $|w_u(\delta u)|$ only depends on the values of δu, but not on its derivatives. Thus, an appropriate choice for R_u in this context is

$$R_u(\delta u) := \|\delta u\|_{L_3}^3.$$

This is allowed in our flexible framework, while without the added flexibility we would have to choose

$$\tilde{R}_u(\delta u) = \|\delta u\|_{H^1}^3.$$

These two error models differ significantly. In particular for corrections δu that are highly oscillating (like for example $\delta u(s) := c\sin(\nu|s|)$, where the frequency ν is large), we get that $\tilde{R}_u(\delta u) \gg R_u(\delta u)$, while for smooth corrections δu we expect $\tilde{R}_u(\delta u) \approx R_u(\delta u)$. Thus, in view of our calculation, $\tilde{R}_u$ would tend to overestimate $|w_u|$ for rough corrections, so that $|w_u|$ may appear *highly anisotropic* w.r.t. $\tilde{R}_u(\delta u)$. This example also demonstrates that the norm, induced by a good preconditioner (which would be of H^1-type here) is not automatically a suitable norm for measuring remainder terms.

2.2.3 Nonlinear optimal control: black-box approach

The aim in (PDE constrained) optimal control is to minimize a cost functional subject to a (partial) differential equation as equality constraint. For an introduction into this topic, we refer to the textbooks [13, 17, 10, 18]. Usually, the optimization variable is divided into a *control* u which enters the differential equation as data, and the *state* y, which is the corresponding solution. This relation can be described by a nonlinear operator via $y = S(u)$. Elimination of y then yields an optimization problem of the following form:

$$\min_{u\in U} f(S(u), u).$$

This general problem, however, is hardly tractable theoretically, and thus, one restricts considerations often to the following special case:

$$f(S(u), u) = g_1(S(u)) + g_2(u) = g_1(S(u)) + \frac{\alpha}{2}\|u\|_U^2.$$

If, for example, S is the solution operator for a non-linear elliptic PDE, and $U = L_2(\Omega)$, then an appropriate choice of norms would be

$$\|v\| := \|v\|_{L_2(\Omega)} \qquad |v| := \|S'(u)v\|_{H_0^1}.$$

Here $|\cdot|$ depends on u, an issue that is encountered frequently. We will ignore this, however, for the sake of simplicity. Usually, $S'(u)$ is a compact linear operator, so that $|\cdot|$ is strictly weaker than $\|\cdot\|$.

Due to the special form of f, which only allows non-convexity in $g_1 \circ S$ we obtain, similarly as above the following estimates:

$$|f'_u \delta u| \le c(u)\|\delta u\|$$
$$c_0(u)|\delta u|^2 \le f''_u(\delta u, \delta u) \le c_1(u)\|\delta u\|^2.$$

3 Algorithmic framework

In this work we will follow to some extent the ideas of [20] and [2] and consider algorithms that are based on successive computation of trial steps, acceptance or rejection of these steps, and update of the model parameters. We consider in the following Algorithm 3.1, which consists of a simple outer iteration in which accepted steps δx_k are added to iterates x_k, and an inner loop, summarized here in the subroutine "CompAccStep". At a given point x "CompAccStep" computes an increasing sequence of parameters ω_i and corresponding trial steps δx_i until one of them is found acceptable. Moreover, it returns a parameter ω_{i+1}, which is used in the next call of "CompAccStep" as a starting value ω_0. The boolean variable ι_{mod} which appears in the algorithm is used in Condition 3.5 (below) to switch a modification of a "fraction-of-Cauchy-decrease" condition on or off.

Concerning the naming of the indices we use the convention that an index k always refers to the outer loop, while the index i is connected to the inner loop at a given point x.

Algorithm 3.1.
Input: initial guess x_0, initial parameter ω_0
 $k \leftarrow -1$, $\iota_{\mathrm{mod}} \leftarrow 0$
 repeat *(outer loop)*
  $k \leftarrow k+1$
  $(\delta x_k,\, \omega_{k+1,0},\, \iota_{\mathrm{mod}},\, \omega_k)$=CompAccStep$(x_k,\, \omega_{k,0},\, \iota_{\mathrm{mod}})$
  $x_{k+1} = x_k + \delta x_k$ (δx_k *is an accepted step for the model* $m_{x_k}^{\omega_k}$)
 until convergence test satisfied
Output: x_{k+1}

subroutine CompAccStep$(x,\, \omega_0,\, \iota_{\mathrm{mod}})$
Input: current iterate x, current parameter ω_0, current value of ι_{mod}
 $i \leftarrow -1$
 Specify at x the quantities f'_x, H_x, R_x
 Compute a direction Δx^C that satisfies Condition 3.4
 repeat *(inner loop)*
  $i \leftarrow i+1$
  Create model function $m_x^{\omega_i}$
  Compute δx_i^C along the direction Δx^C (cf. Condition 3.4)
  if $\frac{\omega_i R_x(\delta x_i^C)}{\|\delta x_i^C\|^2} \geq C_{\mathrm{mod}}$ **then** $\iota_{\mathrm{mod}} \leftarrow 1$
  Compute δx_i that satisfies Condition 3.5 (which depends on ι_{mod})
  Compute ω_{i+1} as described in Section 3.4
 until δx_i satisfies Condition 3.7 with $m_x^{\omega_i}$.
 if $\frac{\omega_i R_x(\delta x_i^C)}{\|\delta x_i^C\|^2} < C_{\mathrm{mod}}$ **then** $\iota_{\mathrm{mod}} \leftarrow 0$
Output: $(\delta x_i,\, \omega_{i+1},\, \iota_{\mathrm{mod}},\, \omega_i)$

This general algorithm offers room for a large variety of implementations. They may differ in the way δx is computed, ω is updated, and iterates are accepted. In this section we will discuss the main features of this algorithm and show that the subroutine "CompAcc-Step" terminates after finitely many iteration. Global and local convergence properties of the overall algorithm are discussed in Section 4.

In practical implementations, a convergence test (last line of the outer loop) may check, whether $\|f'_{x_{k+1}}\|$ is sufficiently small. For our theoretical purpose, where we consider a possibly infinite sequence of iterates, it is sufficient to check whether $f'_{x_{k+1}} = 0$.

3.1 Directional model minimizers

As a minimal requirement, we suppose that any trial step δx computed in Algorithm 3.1 minimizes m_x^ω (defined in (3)) on span$\{\delta x\}$. We call any such a correction δx a *directional minimizer* of m_x^ω. Directional minimizers are easy to compute and have nice properties.

Existence of a minimizer of m_x^ω in X may not hold due to a possible lack of $\|\cdot\|$-coercivity (consider e.g. the case $m_x^\omega(\delta x) = f_x'\delta x + \omega/6|\delta x|^3$). However, directional minimizers of m_x^ω always exist.

Lemma 3.2. *For a directional minimizer δx of m_x^ω it holds $f_x'\delta x \le 0$ and*

$$0 = f_x'\delta x + H_x(\delta x, \delta x) + \frac{\omega}{2}R_x(\delta x), \tag{12}$$

$$m_x^\omega(\delta x) = \frac{1}{2}f_x'\delta x - \frac{\omega}{12}R_x(\delta x) \tag{13}$$

$$= -\frac{1}{2}H_x(\delta x, \delta x) - \frac{\omega}{3}R_x(\delta x). \tag{14}$$

Proof. Since the term $\frac{1}{2}H_x(\delta x, \delta x) + \omega/6R_x(\delta x)$ is the same for $+\delta x$ and $-\delta x$ it follows that $m_x^\omega(-\delta x) < m_x^\omega(\delta x)$ if $f_x'\delta x > 0$. Hence, a directional minimizer of m_x^ω satisfies $f_x'\delta x \le 0$.

As first order optimality conditions for a minimizer δx of m_x^ω we compute:

$$0 = (m_x^\omega)'(\delta x)v = f_x'v + H_x(\delta x, v) + \frac{\omega}{6}R_x'(\delta x)v \quad \forall v \in \mathrm{span}\{\delta x\}. \tag{15}$$

and thus, by homogeneity (4) of R_x we conclude $R_x'(\delta x)v = 3R_x(\delta x)v$ and thus (12). Inserting this into the definition of m_x^ω, we obtain (13) – (14). □

The following basic property is a simple consequence:

Lemma 3.3. *Denote by $\delta x(\omega) = \lambda(\omega)\Delta x$ directional model minimizers along a fixed direction Δx for varying $\omega > 0$. We have*

$$\lim_{\omega\to\infty} \|\delta x(\omega)\| = \lim_{\omega\to\infty} \lambda(\omega) = 0. \tag{16}$$

Proof. Fix $\omega_0 > 0$ and denote the corresponding directional minimizer in our direction by Δx. For any other $\omega > 0$ we have $\delta x(\omega) = \lambda\Delta x$ with some $\lambda > 0$. Inserting this into (12) and dividing by λ we obtain the following quadratic equation for λ:

$$0 = f_x'\Delta x + H_x(\Delta x, \Delta x)\lambda + \frac{\omega}{2}R_x(\Delta x)\lambda^2 \tag{17}$$

that depends on a parameter $\omega > 0$, and which is of the form $0 = c + 2b\lambda + a\omega\lambda^2$ with $c = f_x'\Delta x < 0$ and $a\omega > 0$. This yields exacly one non-negative solution

$$\lambda(\omega) = \frac{1}{a\omega}\left(-b + \sqrt{b^2 + |c|a\omega}\right)$$

Considering the limit $\omega \to \infty$ we find that (16) holds. □

3.2 Acceptable steps

In addition to being a directional minimizer, a trial step δx has to satisfy a "fraction of Cauchy decrease" type condition. Classically, this involves the explicit computation of a direction of steepest descent Δx^{SD} in each step of the outer loop. Its purpose is to establish a link between primal quantities δx and dual quantities f'_x. We emphasize that steepest descent directions depend on the choice of the norm $\|\cdot\|$.

In many cases the analytically straightforward choice of $\|\cdot\|$ will lead to a rather expensive computation of Δx^{SD}. For example, if $\|\cdot\|_{H^1}$ is used, then Δx^{SD} has to be computed from f'_x via the solution of an elliptic partial differential equation. It is sufficient, however, to compute directions of significant descent:

Condition 3.4. *Let $1 \geq \mu > 0$ be fixed. We compute at the beginning of each inner loop a fixed direction Δx^C that satisfies*

$$f'_x \Delta x^C \leq -\mu \|f'_x\| \|\Delta x^C\|, \tag{18}$$

and call Δx^C a direction of significant descent. *For given $\omega > 0$ the directional minimizer of m_x^ω in direction of Δx^C is called* quasi Cauchy step δx^C.

Often such steps are much cheaper to compute via a *preconditioner* than exact steepest descent directions. The corresponding parameter μ does not need to be specified explicitely. Note that δx^C results from a *scaling* of Δx^C (cf. Lemma 3.3):

$$\delta x^C = \lambda(\omega)\Delta x^C, \text{ for some } \lambda(\omega) > 0.$$

In our flexible framework R_x can be chosen quite independently of $\|\cdot\|$. This results in a modification of the classical Cauchy decrease condition. This modification penalizes irregular search directions, i.e., directions, where $\|\delta x\|^3 \gg R_x(\delta x)$ and thus avoids that iterates leave $(X, \|\cdot\|)$.

However, since such a modification might exclude useful search directions, we will only employ it to enforce global convergence in difficult cases. To this end we introduce the logic variable $\iota_{\mathrm{mod}} \in \{0,1\}$ which is set to 1, if during an inner loop the relation

$$\frac{\omega_i R_x(\delta x_i^C)}{\|\delta x_i^C\|^2} \geq C_{\mathrm{mod}} \qquad \text{for some constant } C_{\mathrm{mod}} > 0 \tag{19}$$

holds, and it is set to 0, if after acceptance of a step the converse relation holds (cf. Algorithm 3.1). We will show in the course of our analysis that the left hand side of this inequality tends to ∞ (thus $\iota_{\mathrm{mod}} = 1$ eventually), if global convergence is delayed, while it tends to 0 (thus $\iota_{\mathrm{mod}} = 0$ eventually) in case of fast local convergence.

Condition 3.5. *Let $1 \geq \underline{\beta} > 0$ be fixed and δx^C be the quasi Cauchy step of m_x^ω. For a given directional minimizer δx define*

$$\sigma := \frac{R_x(\delta x^C)}{\|\delta x^C\|^3} \cdot \frac{\|\delta x\|^3}{R_x(\delta x)}.$$

$$\beta := \begin{cases} \underline{\beta} & \text{if} \quad \iota_{\mathrm{mod}} = 0 \\ \underline{\beta}\max\{1, \sqrt{\sigma}\} & \text{if} \quad \iota_{\mathrm{mod}} = 1. \end{cases} \tag{20}$$

Then choose δx as a directional minimizer of m_x^ω, such that

$$m_x^\omega(\delta x) \leq \beta m_x^\omega(\delta x^C). \tag{21}$$

The criterion (20) reduces to $\beta = \underline{\beta}$, if either $\delta x = \delta x^C$, so that δx^C is always acceptable, or $R_x(\cdot) = \|\cdot\|^3$, which is the classical case.

Lemma 3.6. *The following inequality holds for δx^C, as defined in Condition 3.4:*

$$\mu\|f_x'\|\|\delta x^C\| \le H_x(\delta x^C,\delta x^C) + \frac{\omega}{2}R_x(\delta x^C). \tag{22}$$

Let δx be a directional minimizer that satisfies Condition 3.5 with $\iota_{\rm mod}=1$. Then there is $c(\mu,\underline{\beta})>0$, such that

$$\frac{R_x(\delta x)}{\|\delta x\|} \ge c\frac{R_x(\delta x^C)}{\|\delta x^C\|}. \tag{23}$$

Proof. From (12) and (18) we conclude (22):

$$\mu\|f_x'\|\|\delta x^C\| \overset{(18)}{\le} |f_x'\delta x^C| \overset{(12)}{=} H_x(\delta x^C,\delta x^C) + \frac{\omega}{2}R_x(\delta x^C).$$

To show (23) assume first that $f_x'\delta x + \beta|f_x'\delta x^C| \le 0$. Then

$$\|f_x'\|\|\delta x\| \ge |f_x'\delta x| = -f_x'\delta x \ge \beta\mu\|f_x'\|\|\delta x^C\|,$$

and thus $\|\delta x\| \ge \beta\mu\|\delta x^C\|$. Inserting (20) we obtain

$$\left(\frac{R_x(\delta x)}{\|\delta x\|}\right)^{1/2} \ge \underline{\beta}\mu\left(\frac{R_x(\delta x^C)}{\|\delta x^C\|}\right)^{1/2}$$

which implies (23) in this case.

Otherwise, we use (13) for δx, (21), and (13) for δx^C to compute

$$\begin{aligned}\frac{\omega}{6}R_x(\delta x) &= f_x'\delta x - 2m_x^\omega(\delta x) \ge f_x'\delta x - 2\beta m_x^\omega(\delta x^C)\\ &= \underbrace{f_x'\delta x + \beta|f_x'\delta x^C|}_{\ge 0} + \beta\frac{\omega}{6}R_x(\delta x^C) \ge \beta\frac{\omega}{6}R_x(\delta x^C),\end{aligned}$$

and thus, $R_x(\delta x) \ge \beta R_x(\delta x^C)$. Inserting once again (20) we get

$$\left(\frac{R_x(\delta x)}{\|\delta x\|}\right)^{3/2} \ge \underline{\beta}\left(\frac{R_x(\delta x^C)}{\|\delta x^C\|}\right)^{3/2}.$$

and thus also (23). □

3.3 Acceptance of trial steps

After a directional minimizer of our model has been computed and serves as a trial step, we have to decide, whether this trial step is acceptable as an optimization step. For this purpose we impose the following relative acceptance criterion, which is well known and popular in trust-region methods. To this end, let us define the ratio of decrease in f and in the model m_x^ω:

$$\eta := \frac{f(x+\delta x)-f(x)}{m_x^\omega(\delta x)} = \frac{f(x+\delta x)-f(x)}{m_x^\omega(\delta x)-m_x^\omega(0)}. \tag{24}$$

Recall that we have chosen m_x^ω in a way that $m_x^\omega(0)=0$. Since $m_x^\omega(\delta x)<0$, we see that $\eta>0$ yields a decrease of f, and $\eta=0$ means that f has remained constant. This yields the following classical condition:

Condition 3.7. *Choose $\underline{\eta}\in]0,1[$. A trial step δx is accepted, if it satisfies the condition*

$$\eta \ge \underline{\eta}. \tag{25}$$

Otherwise it is rejected.

3.4 Adaptive choice of ω

In this section we discuss the adaptive choice of the sequence of regularization parameters ω_i in $m_x^{\omega_i}$, needed in subroutine "CompAccSteps" at a given point x. We assume that ω_i and δx_i have already been computed, and the update ω_{i+1} has to be made.

In [2] the choice of regularization parameters is made according to a classification of the steps into "unsuccessful", "successful" and "very successful". Here we follow [20] and base our considerations on the idea that $\omega_i/6\, R_x$ should be a third order model for the difference $f - q_x$, which leads to the assignment:

$$\omega_{i,\text{raw}} := \frac{6(f(x+\delta x_i) - q_x(\delta x_i))}{R_x(\delta x_i)} \overset{(2)}{=} \frac{6w_x(\delta x_i)}{R_x(\delta x_i)}. \tag{26}$$

Of course, $\omega_{i,\text{raw}}$ cannot be used directly as ω_{i+1} in our algorithm. We have to introduce some safeguard restrictions.

In order to guarantee positivity of ω_{i+1} and to avoid oscillatory behavior, we assume that the algorithm provides restrictions on updates $\omega_i \to \omega_{i+1}$ to guarantee:

$$\begin{aligned} &\omega_{i+1} > 0 \\ &\omega_{i+1} \le \frac{1}{\rho}\left(\omega_i + C_\omega + 2\frac{C_{f'}|f'_x\delta x_i| + |H_x(\delta x_i, \delta x_i)|}{R_x(\delta x_i)}\right) \\ &\qquad\qquad\qquad \text{for constants } 0 < \rho < 1, C_{f'} \ge 0, C_\omega \ge 0. \end{aligned} \tag{27}$$

Positivity of ω is, of course, a basic requirement which guarantees that the term R_x is present throughout the computation. The second condition (27) inhibits that ω is increased too quickly, with the result that the next trial step has to be chosen much shorter than the previous one. However, in a certain range (corresponding to C_ω) the increase can be performed freely. If R_x is much smaller than the remaining terms of $m_x^{\omega_i}$ a fast increase of ω is also possible. Technically, this restriction enters into the global convergence proof in (54), below.

The following theory will cover algorithms that respect these restrictions, and increase ω after a rejected trial step, according to

$$\text{After rejected trial step:} \quad \omega_{i+1} \ge \min\{\omega_{i,\text{raw}}, \rho^{-1}\omega_i\} \qquad (\rho \text{ as defined in (27)}) \tag{28}$$

Any algorithm that does not allow an increase like this, is likely to get stuck in an inner loop. Technically this condition is used at the beginning of the proof of Theorem 3.8. Next, we impose the restriction on our algorithm that after an increase of ω, ω_{i+1} should not be chosen larger than $\omega_{i,\text{raw}}$:

$$\omega_{i+1} \le \max\{\omega_i, \omega_{i,\text{raw}}\}.$$

By (26) this implies the following estimate:

$$\omega_{i+1} \ge \omega_i \quad \Rightarrow \quad \omega_{i+1}R_x(\delta x_i) \le 6w_x(\delta x_i). \tag{29}$$

To obtain fast local convergence under weak assumptions on R_x we do not increase ω if η_i (defined via (24) for δx_i) is very close to 1. Let us choose $\overline{\eta} \in [\underline{\eta}, 1[$ and state

$$\text{If } \ \eta_i > \overline{\eta} \ \text{ then} \qquad \omega_{i+1} \le \omega_i. \tag{30}$$

The following simple example update that satisfies all these requirements is the following rule (where we could replace the simple upper bound $\rho^{-1}\omega_i$ by the right hand side of (27)):

- If $\eta_i \leq \overline{\eta}$: $\omega_{i+1} = \max\{\rho\omega_i, \min\{\rho^{-1}\omega_i, \omega_{i,\mathrm{raw}}\}\}$ for some $0 < \rho < 1$
- If $\eta_i > \overline{\eta}$: $\omega_{i+1} = \omega_i$.

In our framework we deliberately dispense with a-priori restrictions like $\underline{\omega} \leq \omega_i$ for some very small lower bound $0 < \underline{\omega} \ll 1$. Such restrictions can, and should of course, be added in finite precision arithmetic.

3.5 Finite termination of inner loops

Next, we show that each inner loop, i.e., the subroutine "CompAccStep" of our algorithm accepts a finite ω after finitely many updates and thus terminates finitely. Hence, in the following we consider fixed x and a sequence ω_i of parameters and δx_i of trial steps, computed by the subroutine "CompAccStep".

Theorem 3.8. *Assume that f is Fréchet differentiable at x and $f'_x \neq 0$. Moreover, assume that the left inequalities of (7) and (8) hold at x. Then:*

(i) If a trial step δx_i is rejected, then $\omega_{i+1} \geq \min\{\rho^{-1}, (1-\underline{\eta})/2 + 1\}\omega_i > \omega_i$.

(ii) The inner loop terminates successfully after finitely many iterations.

Proof. In view of (28), assume that $\omega_{i+1} < \rho^{-1}\omega_i$. Then violation of (25) implies

$$\begin{aligned}
\omega_{i+1} \geq \omega_{i,\mathrm{raw}} &= \frac{6}{R_x(\delta x_i)}(f(x+\delta x_i) - q_x(\delta x_i)) \\
&= \frac{6}{R_x(\delta x_i)}\Big(f(x+\delta x_i) - f(x) - m_x^{\omega_i}(\delta x_i) + \frac{\omega_i}{6}R_x(\delta x_i)\Big) \\
&> \frac{6}{R_x(\delta x_i)}(\underline{\eta}-1)m_x^{\omega_i}(\delta x_i) + \omega_i \\
&= \frac{6}{R_x(\delta x_i)}(1-\underline{\eta})\Big(-1/2 f'_x\delta x_i + \frac{\omega_i}{12}R_x(\delta x_i)\Big) + \omega_i \\
&\geq ((1-\underline{\eta})/2 + 1)\omega_i.
\end{aligned}$$

Hence, (i) is shown: each rejection is followed by an increase of ω by at least a fixed factor.

Next, assume for contradiction that (25) fails infinitely often during successive updates from ω_i to ω_{i+1}. Then, we have $\lim_{i\to\infty}\omega_i = \infty$ so that Lemma 3.3 yields for the quasi Cauchy steps (which are all scalings of a fixed direction: $\delta x_i^C = \lambda(\omega_i)\Delta x^C$)

$$\lim_{i\to\infty} \|\delta x_i^C\| = \lim_{i\to\infty} \lambda(\omega_i) = 0$$

and thus

$$\begin{aligned}
\liminf_{i\to\infty} \frac{\omega_i R_x(\delta x_i^C)}{2\|\delta x_i^C\|} &\overset{(22)}{\geq} \mu\|f'_x\| - \lim_{i\to\infty}\frac{H_x(\delta x_i^C, \delta x_i^C)}{\|\delta x_i^C\|} \\
&= \mu\|f'_x\| - \lim_{i\to\infty}\lambda(\omega_i)\frac{H_x(\Delta x^C, \Delta x^C)}{\|\Delta x^C\|} = \mu\|f'_x\| > 0
\end{aligned}$$

since $\|f'_x\| \neq 0$. This implies that

$$\lim_{i\to\infty}\frac{\omega_i R_x(\delta x_i^C)}{\|\delta x_i^C\|^2} = \infty$$

and thus $\iota_{\text{mod}} = 1$ by (19) for sufficiently large i. It thus follows from (23) that there exists a constant $W_0 > 0$ such that for the following positive sequence

$$\frac{\omega_i R_x(\delta x_i)}{\|\delta x_i\|} \geq W_0 > 0 \quad \forall i \in \mathbb{N}. \tag{31}$$

Since $\omega_{i+1} > \omega_i$, (29) holds for ω_{i+1}. Thus, we conclude that

$$0 < W_0 \overset{(31)}{\leq} \frac{\omega_i R_x(\delta x_i)}{\|\delta x_i\|} < \frac{\omega_{i+1} R_x(\delta x_i)}{\|\delta x_i\|} \overset{(29)}{\leq} \frac{6 w_x(\delta x_i)}{\|\delta x_i\|}, \tag{32}$$

and hence by Fréchet differentiability and Lemma 2.1 (with $x_k = x$ the constant sequence) there is $D_0 > 0$ such that $\|\delta x_i\| \geq D_0$, and hence, by (32) also

$$\omega_i R_x(\delta x_i) \geq W_0 D_0 > 0. \tag{33}$$

With this, we compute from (12) (using $f'_x \delta x_i = -|f'_x \delta x_i|$)

$$\frac{|f'_x \delta x_i|}{\omega_i R_x(\delta x_i)} = \frac{H_x(\delta x_i, \delta x_i)}{\omega_i R_x(\delta x_i)} + \frac{1}{2}.$$

By (7), if the middle term including the Hessian has a negative contribution it vanishes asymptotically:

$$\lim_{i\to\infty} \left| \frac{\min\{H_x(\delta x_i, \delta x_i), 0\}}{\omega_i R_x(\delta x_i)} \right| \overset{(7)}{\leq} \lim_{i\to\infty} \frac{|\gamma||\delta x_i|^2}{\omega_i R_x(\delta x_i)} \overset{(8)}{\leq} \lim_{i\to\infty} \frac{|\gamma|}{\omega_i^{2/3} (\omega_i R_x(\delta x_i))^{1/3}} \overset{(33)}{=} 0.$$

Thus we conclude that

$$\liminf_{i\to\infty} \frac{|f'_x \delta x_i|}{\omega_i R_x(\delta x_i)} \geq \frac{1}{2} > 0 \tag{34}$$

and thus via (31) that also

$$\liminf_{i\to\infty} \frac{|f'_x \delta x_i|}{\|\delta x_i\|} \geq \frac{W_0}{2} > 0. \tag{35}$$

However, as a consequence of (33) and (8) we have:

$$\omega_i R_{x_i}(\delta x_i) \overset{(8)}{\geq} \omega_i R_x(\delta x_i)^{2/3} |\delta x_i| \overset{(33)}{\geq} (W_0 D_0)^{2/3} \omega_i^{1/3} |\delta x_i|. \tag{36}$$

Thus, by (34) we conclude

$$\lim_{i\to\infty} \frac{\|\delta x_i\|}{|\delta x_i|} \geq \lim_{i\to\infty} \frac{|f'_x \delta x_i|}{|\delta x_i| \|f'_x\|} \overset{(34)}{\geq} \lim_{i\to\infty} \frac{1}{2} \frac{\omega_i R_{x_i}(\delta x_i)}{|\delta x_i| \|f'_x\|} \overset{(36)}{\geq} \frac{1}{2} \frac{(W_0 D_0)^{2/3}}{\|f'_x\|} \lim_{i\to\infty} \omega_i^{1/3} = \infty.$$

Thus, $\lim_{i\to\infty} |\delta x_i|/\|\delta x_i\| = 0$ and by Lemma 2.2 we conclude weak convergence $\delta x_i/\|\delta x_i\| \rightharpoonup 0$ in $(X, \|\cdot\|)$. This, however, implies

$$\lim_{k\to\infty} \frac{|f'_x \delta x_i|}{\|\delta x_i\|} = 0.$$

in contradiction to (35). □

4 Convergence Theory

In this section we will establish first order global convergence, and second order local convergence results. In the following we will consider the sequence x_k, generated by Algorithm 3.1, corresponding derivatives $f'_{x_k} \in X^*$, and accepted corrections δx_k. We denote by ω_k that parameter for which δx_k has been computed as an accepted directional minimizer of $m^{\omega_k}_{x_k}$. At the end of subroutine "CompAccStep", after the inner loop, this parameter appears as ω_i, to be distinguished from the update ω_{i+1} that corresponds to $\omega_{k+1,0}$ in the outer loop and is computed after acceptance of δx_k. Similarly, δx^C_k denotes a quasi-Cauchy step for $m^{\omega_k}_{x_k}$ and η_k is defined by (24) for δx_k.

In the whole section we use that the computed quantities satisfy Conditions 3.4, 3.5, and 3.7 and the update conditions from Section 3.4. Moreover, we assume throughout the basic properties, introduced in the beginning of Section 2. From the assumptions Section 2.1 we only use existence of $f'_x \in X^*$ throughout, as well as all assumptions, needed to show finite termination of the inner loops, i.e., existence of the sequence x_k. All other assumptions will be referenced explicitly, when needed.

In the whole section we exclude the trivial case that $f'_{x_k} = 0$, for some k, which leads to finite termination of our algorithm. Moreover, we may assume that the sequence of computed function values $f(x_k)$ is bounded from below. Otherwise our algorithm, which enforces descent, fulfills its purpose of minimization by generating a sequence x_k with $\lim_{k\to\infty} f(x_k) = -\infty$.

4.1 Global Convergence

Under mild assumptions we will show that our algorithm cannot converge to non-stationary points, while slightly stronger assumptions yield convergence of derivatives to 0. Our technique will be to derive a contradiction to the case that x_k converges to a non-stationary point, so that in particular $\|f'_{x_k}\|$ remains bounded away from zero.

Lemma 4.1. *Assume that Algorithm 3.1 generates an infinite sequence x_k such that $f(x_k)$ is bounded from below. Then*

$$\sum_{k=0}^{\infty} \|f'_{x_k}\| \|\delta x^C_k\| < \infty \tag{37}$$

$$\sum_{k=0}^{\infty} \omega_k R_{x_k}(\delta x^C_k) < \infty, \qquad \sum_{k=0}^{\infty} \omega_k R_{x_k}(\delta x_k) < \infty. \tag{38}$$

Proof. We use (25) and (21) (using only $\beta \geq \underline{\beta}$) to compute

$$\begin{aligned} f(x_{k+1}) - f(x_k) &\overset{(25)}{\leq} \underline{\eta} m^{\omega_k}_{x_k}(\delta x_k) \overset{(12)}{=} \underline{\eta}\left(\frac{1}{2} f'_{x_k}\delta x_k - \frac{\omega_k}{12} R_{x_k}(\delta x_k)\right) \\ &\overset{(21)}{\leq} \underline{\eta}\beta m^{\omega_k}_{x_k}(\delta x^C_k) \leq \underline{\eta}\,\underline{\beta} m^{\omega_k}_{x_k}(\delta x^C_k) \overset{(12)}{=} \underline{\eta}\,\underline{\beta}\left(\frac{1}{2} f'_{x_k}\delta x^C_k - \frac{\omega_k}{12} R_{x_k}(\delta x^C_k)\right) \\ &\leq \frac{\underline{\eta}\,\underline{\beta}}{2} f'_{x_k}\delta x^C_k \overset{(18)}{\leq} -\frac{\mu\underline{\eta}\,\underline{\beta}}{2}\|f'_{x_k}\|\|\delta x^C_k\| \leq 0. \end{aligned}$$

By monotonicity and boundedness

$$\sum_{k=0}^{\infty} f(x_{k+1}) - f(x_k) = \inf_k f(x_k) - f(x_0) > -\infty,$$

and by our chain of inequalities we conclude (37) and (38). □

The main observation here is that $\|f'_{x_k}\|\|\delta x_k^C\| \to 0$, and to obtain $\|f'_{x_k}\| \to 0$ it remains to prevent $\|\delta x_k^C\|$ from becoming too small, compared to $\|f'_{x_k}\|$. The extraordinary role of δx_k^C has its origin in the acceptance criterion (21), which compares all steps to the quasi Cauchy steps.

To obtain a quick understanding of the situation, take a look at (22) and observe the following relation:

$$\mu\|f'_{x_k}\| \overset{(22)}{\leq} \frac{H_{x_k}(\delta x_k^C, \delta x_k^C)}{\|\delta x_k^C\|} + \frac{\omega_k}{2}\frac{R_{x_k}(\delta x_k^C)}{\|\delta x_k^C\|}.$$

The undesired case is that $\|f'_{x_k}\|$ is bounded away from zero, which in turn implies that $\lim_{k\to\infty}\|\delta x_k^C\| = 0$ by (37). Taking into account the upper bounds (7) for H_x and (8) for R_x, we see that

$$\limsup_{k\to\infty} \frac{H_{x_k}(\delta x_k^C, \delta x_k^C)}{\|\delta x_k^C\|} \leq 0, \tag{39}$$

$$\lim_{k\to\infty} \frac{R_{x_k}(\delta x_k^C)}{\|\delta x_k^C\|} = 0. \tag{40}$$

In fact for all global convergence results we may replace the upper bounds of (7) and (8) by these weaker assumptions.

In view of (39) and (40), which exclude that our iteration is stalled by H_x and R_x being overly large, the "bad case" can only happen, if ω_k is increased too rapidly. Under smoothness assumptions on f that imply boundedness of ω_k (a global Lipschitz condition on $H_x = f''(x)$) we would be finished at this point. To cover the more general case, we have to invest some more theoretical work. Let us start with collecting some simple consequences of $\|f'_{x_k}\|$ being bounded away from 0:

Lemma 4.2. *Assume that the sequence $f(x_k)$ is bounded from below. For fixed $\nu > 0$ assume that the following set of indices is infinite:*

$$\mathcal{L}^\nu := \{k : \|f'_{x_k}\| \geq \nu\}.$$

Assume that (39) *and* (40) *hold for δx_k^C along the sequence of iterates x_k for $k \in \mathcal{L}^\nu$. Then*

$$\lim_{k\in\mathcal{L}^\nu\to\infty} \frac{\omega_k R_{x_k}(\delta x_k^C)}{\|\delta x_k^C\|^2} = \infty \tag{41}$$

$$\lim_{k\in\mathcal{L}^\nu\to\infty} \omega_k = \infty. \tag{42}$$

Let δv_k be any directional minimizer of $m_{x_k}^{\omega_k}$ that satisfies Condition 3.5 with $\iota_{\mathrm{mod}} = 1$. Then

$$\inf_{k\in\mathcal{L}^\nu} \frac{\omega_k R_x(\delta v_k)}{\|f'_{x_k}\|\|\delta v_k\|} > 0 \tag{43}$$

For iteration $k \in \mathcal{L}^\nu$ let δx_k be the accepted step. Eventually, (for sufficiently large k) δx_k satisfies Condition 3.5 with $\iota_{\mathrm{mod}} = 1$ and it holds:

$$\sum_{k\in\mathcal{L}^\nu} \|\delta x_k\| < \infty. \tag{44}$$

Proof. From $\|f'_{x_k}\| \geq \nu$ we get $\|\delta x_k^C\| \to 0$ due to (37) and thus, via (22) and (39):

$$\liminf_{k\in\mathcal{L}^\nu\to\infty} \frac{\omega_k R_{x_k}(\delta x_k^C)}{2\|f'_{x_k}\|\|\delta x_k^C\|} \overset{(22)}{\geq} \mu - \limsup_{k\in\mathcal{L}^\nu\to\infty} \frac{H_{x_k}(\delta x_k^C, \delta x_k^C)}{\|f'_{x_k}\|\|\delta x_k^C\|} \overset{(39)}{\geq} \mu. \tag{45}$$

Since $\lim_{k\in\mathcal{L}^\nu\to\infty}\|\delta x_k^C\| = 0$ and $\|f'_{x_k}\| \geq \nu$ for $k \in \mathcal{L}^\nu$ multiplication of (45) by $\|f'_{x_k}\|/\|\delta x_k^C\|$ implies (41). By (45) we compute, using again $\lim_{k\in\mathcal{L}^\nu\to\infty}\|\delta x_k^C\| = 0$ and (40):

$$2\|f'_{x_k}\|\omega_k^{-1} \overset{(45)}{\leq} c\frac{R_{x_k}(\delta x_k^C)}{\|\delta x_k^C\|} \overset{(40)}{\to} 0.$$

From that (42) follows from $\|f'_{x_k}\| \geq \nu$ for $k \in \mathcal{L}^\nu$.

By (23) we also have

$$\liminf_{k\to\infty} \frac{\omega_k R_{x_k}(\delta v_k)}{\|f'_{x_k}\|\|\delta v_k\|} \geq 2c\mu > 0$$

and thus (43), because all members of the sequence are strictly positive.

Due to (41) $\frac{\omega_k R_{x_k}(\delta x_k^C)}{\|\delta x_k^C\|^2} \geq C_{\mathrm{mod}}$ eventually, i.e. for all $k \in \mathcal{L}^\nu$ sufficiently large. In subroutine "CompAccStep" this relation is evaluated before δx_k is computed, and thus $\iota_{\mathrm{mod}} = 1$ when δx_k is computed.

Now (44) follows from (38) and (43) via the computation

$$\begin{aligned}\sum_{k\in\mathcal{L}^\nu} \|f'_{x_k}\|\|\delta x_k\| &= \sum_{k\in\mathcal{L}^\nu} \omega_k R_{x_k}(\delta x_k)\left(\frac{\omega_k R_{x_k}(\delta x_k)}{\|f'_{x_k}\|\|\delta x_k\|}\right)^{-1} \\ &\leq \sum_{k\in\mathcal{L}^\nu} \omega_k R_{x_k}(\delta x_k)\left(\inf_{k\in\mathcal{L}^\nu}\frac{\omega_k R_{x_k}(\delta x_k)}{\|f'_{x_k}\|\|\delta x_k\|}\right)^{-1} < \infty,\end{aligned}$$

and the fact that $\|f'_{x_k}\|$ is bounded away from 0. □

An important conclusion of this lemma is that if $\mathcal{L}^\nu = \mathbb{N}$, then by (44) x_k is a Cauchy sequence in $(X, \|\cdot\|)$, and thus x_k converges to some limit x_*.

Up to now, the smoothness of f and the lower bound in the Gårding inequality (7) did not enter our considerations. In the following theorem, which is the main step of our study, we will take this and the safeguard restrictions (29) and (27) on the update of ω into account.

We are interested in the case that our algorithm converges to a non-stationary point. We show in this case that the following set is infinite:

$$\begin{aligned}\mathcal{I} := \{k \in \mathbb{N} : &\text{ subroutine "CompAccStep" at } x_k \\ &\text{ computes at least one rejected trial step }\}.\end{aligned}$$

In what follows, we will need for $k \in \mathcal{I}$ the *last rejected trial step* in the call of subroutine "CompAccStep" at x_k. We will denote this trial step by $\delta x_{k,\ell}$ with corresponding $\omega_{k,\ell}$ and quasi Cauchy step $\delta x_{k,\ell}^C$. In the context of the subroutine, if δx_i is returned as acceptable correction δx_k, then $(\delta x_{k,\ell}, \omega_{k,\ell}, \delta x_{k,\ell}^C) = (\delta x_{i-1}, \omega_{i-1}, \delta x_{i-1}^C)$.

Theorem 4.3. *Let $x_* \in X$. Assume that f is Fréchet differentiable in a neighborhood of x_* and f' is continuous at x_*. Suppose that x_k converges to x_*. Assume further that along x_k* (39) *and* (40) *hold for δx_k^C. Further, assume that the left inequalities of* (7) *and* (8) *hold along x_k, $k \in \mathcal{I}$ for $\delta x_{k,\ell}$. Then $f'_{x_*} = 0$.*

Proof. Assume for contradiction that $f'_{x_*} \neq 0$. Since $x_k \to x_*$ we also have $\|\delta x_k\| \to 0$ and the sequence $f(x_k)$ is bounded from below by $f(x_*)$ and also $f'_{x_k} \to f'_{x_*}$ in X^*, so that the positive sequence $\|f'_{x_k}\|$ converges to a non-zero value and thus

$$\exists \nu > 0, C_0 < \infty : \nu \leq \|f'_{x_k}\| \leq C_0. \tag{46}$$

Then by (39) Lemma 4.2 applies for $\mathcal{L}^\nu = \mathbb{N}$ for $\nu > 0$, and in particular ω_k is increased infinitely many times due to (42). Moreover, due to (41) we have eventually $\iota_{\mathrm{mod}} = 1$ for the computation of all trial steps. These facts will be used throughout the proof.

An increase of ω can occur in two cases: first, after an accepted trial step δx_k, second after a rejected trial step. In the first case, i.e., $\omega_{k+1,0} \geq \omega_k$ in Algorithm 3.1, we compute by (29):

$$\omega_k R_{x_k}(\delta x_k) \leq \omega_{k+1,0} R_{x_k}(\delta x_k) \overset{(29)}{\leq} 6 w_{x_k}(\delta x_k)$$

and thus, by (43) we conclude that there exists $c_1 > 0$, such that:

$$\frac{6 w_{x_k}(\delta x_k)}{\|f'_{x_k}\|\|\delta x_k\|} \geq \frac{\omega_k R_{x_k}(\delta x_k)}{\|f'_{x_k}\|\|\delta x_k\|} \overset{(43)}{\geq} c.$$

It follows that there is a constant $\tilde{W}_0 > 0$ independent of k, such that

$$\frac{w_{x_k}(\delta x_k)}{\|\delta x_k\|} \geq \tilde{W}_0 := \nu c > 0 \tag{47}$$

for every k after which ω_k was increased. However, since $\|\delta x_k\| \to 0$ by assumption, Lemma 2.1 implies that the left hand side of (47) tends to 0 along the sequence $x_k \to x_*$. Thus, the *first case* can only happen finitely many times.

Hence, the *second case* must occur infinitely many times, i.e., there must be infinitely many rejected trial steps, and thus infinitely many calls of "CompAccStep" in which a trial step is rejected. This means that $\mathcal{I}$ is a set of infinite size. The lower bound of (7) implies:

$$\exists \gamma_\ell > -\infty : \quad \frac{H_{x_k}(\delta x_{k,\ell}, \delta x_{k,\ell})}{|\delta x_{k,\ell}|^2} \geq \gamma_\ell \quad \forall k \in \mathcal{I}. \tag{48}$$

We divide the remaining argumentation into 3 steps. In the following we will consider $k \in \mathcal{I}$ only.

Step 1: For the inner loop k at x_k consider the last rejected trial step $\delta x_{k,\ell}$ with corresponding regularization parameter $\omega_{k,\ell}$. Recall that $\delta x_{k,\ell}$, like every trial step, is a directional minimizer of $m_{x_k}^{\omega_{k,\ell}}$. After rejection of $\delta x_{k,\ell}$ the next regularization parameter corresponds to the final accepted trial step in this loop δx_k and is thus denoted by ω_k. Let $\delta x^C_{k,\ell}$ be the quasi Cauchy step for $\omega_{k,\ell}$ and δx^C_k the quasi Cauchy step for ω_k. Since δx^C_k and $\delta x^C_{k,\ell}$ both point in the same direction, and since by Theorem 3.8 $\omega_k \geq \omega_{k,\ell}$, we have

$$\delta x^C_{k,\ell} = \lambda \delta x^C_k \quad \text{with} \quad \lambda \geq 1.$$

Then by (23) and (43) for $\delta v_k = \delta x^C_k$, taking into account (46), we get constants $c, \tilde{c} > 0$, such that

$$\begin{aligned} \frac{\omega_k R_{x_k}(\delta x_{k,\ell})}{\|\delta x_{k,\ell}\|} &\overset{(23)}{\geq} c\frac{\omega_k R_{x_k}(\delta x^C_{k,\ell})}{\|\delta x^C_{k,\ell}\|} = c\frac{\omega_k R_{x_k}(\lambda \delta x^C_k)}{\|\lambda \delta x^C_k\|} \\ &= c\frac{\omega_k \lambda^3 R_{x_k}(\delta x^C_k)}{\lambda \|\delta x^C_k\|} \geq c\frac{\omega_k R_{x_k}(\delta x^C_k)}{\|\delta x^C_k\|} \overset{(43)}{\geq} \tilde{c}\|f'_{x_k}\| \geq \tilde{c}\nu > 0. \end{aligned} \tag{49}$$

By (29) this implies that there is a constant $W_0 > 0$, such that

$$\frac{6w_{x_k}(\delta x_{k,\ell})}{\|\delta x_{k,\ell}\|} \ge \frac{\omega_k R_{x_k}(\delta x_{k,\ell})}{\|\delta x_{k,\ell}\|} \ge W_0 := \tilde{c}\nu > 0. \tag{50}$$

This in turn implies by Lemma 2.1 that there is a constant $D_0 > 0$, such that for these *rejected* trial steps $\|\delta x_{k,\ell}\| \ge D_0$ and thus

$$\omega_k R_{x_k}(\delta x_{k,\ell}) \ge W_0 D_0 > 0, \tag{51}$$

in contrast to (38), which holds for *accepted* trial steps.

Step 2: In this step we will show that there are k_0 and M_0 such that

$$\frac{|f'_{x_k}\delta x_{k,\ell}|}{\omega_k R_{x_k}(\delta x_{k,\ell})} \ge M_0 > 0 \qquad \forall k \in \mathcal{I}:\ k \ge k_0. \tag{52}$$

Here our safeguard restriction for the update $\omega_{k,\ell} \to \omega_k$ (27) comes into play, which reads now:

$$\frac{\rho}{2}\omega_k R_x(\delta x_{k,\ell}) \overset{(27)}{\le} \frac{\omega_{k,\ell} + C_\omega}{2} R_x(\delta x_{k,\ell}) + C_{f'_x}|f'_x \delta x_{k,\ell}| + |H_x(\delta x_{k,\ell}, \delta x_{k,\ell})|. \tag{53}$$

We insert this relation into (12), the optimality condition for directional minimizers:

$$\begin{aligned}
|f'_{x_k}\delta x_{k,\ell}| = -f'_{x_k}\delta x_{k,\ell} &\overset{(12)}{=} \frac{\omega_{k,\ell}}{2} R_{x_k}(\delta x_{k,\ell}) + H_{x_k}(\delta x_{k,\ell}, \delta x_{k,\ell}) \\
&\overset{(53)}{\ge} \frac{\rho\omega_k - C_\omega}{2} R_{x_k}(\delta x_{k,\ell}) - C_{f'}|f'_x \delta x_{k,\ell}| - |H_{x_k}(\delta x_{k,\ell}, \delta x_{k,\ell})| + H_{x_k}(\delta x_{k,\ell}, \delta x_{k,\ell}) \\
&= \frac{\rho\omega_k - C_\omega}{2} R_{x_k}(\delta x_{k,\ell}) - C_{f'}|f'_x \delta x_{k,\ell}| + 2\min\{H_{x_k}(\delta x_{k,\ell}, \delta x_{k,\ell}), 0\},
\end{aligned}$$

so that we obtain:

$$\frac{(1 + C_{f'})|f'_{x_k}\delta x_{k,\ell}|}{\omega_k R_{x_k}(\delta x_{k,\ell})} \ge \frac{\rho}{2} - \frac{C_\omega}{2\omega_k} - 2\left|\frac{\min\{H_{x_k}(\delta x_{k,\ell}, \delta x_{k,\ell}), 0\}}{\omega_k R_{x_k}(\delta x_{k,\ell})}\right|. \tag{54}$$

Since $\omega_k \to \infty$, the second term on the right hand side of (54) tends to zero. Moreover, by (48), (8), and (51) the same is true for the third term:

$$\begin{aligned}
\lim_{k\in\mathcal{I}\to\infty} \left|\frac{\min\{H_{x_k}(\delta x_{k,\ell}, \delta x_{k,\ell}), 0\}}{\omega_k R_{x_k}(\delta x_{k,\ell})}\right| &\overset{(48)}{\le} \lim_{k\in\mathcal{I}\to\infty} \frac{|\gamma_\ell|\|\delta x_{k,\ell}|^2}{\omega_k R_{x_k}(\delta x_{k,\ell})} \overset{(8)}{\le} \lim_{k\in\mathcal{I}\to\infty} \frac{|\gamma_\ell|}{\omega_k R_{x_k}(\delta x_{k,\ell})^{1/3}} \\
&\overset{(51)}{\le} \lim_{k\in\mathcal{I}\to\infty} \frac{|\gamma_\ell|}{\omega_k^{2/3}(W_0 D_0)^{1/3}} \overset{(42)}{=} 0.
\end{aligned}$$

Hence, in the limit the left hand side of (54) is strictly positive, which implies (52).

Step 3: Let us finally derive our contradiction. Multiplication of (52) with the middle term in (50) implies on the one hand

$$\frac{|f'_{x_k}\delta x_{k,\ell}|}{\|\delta x_{k,\ell}\|} \ge M_0 W_0 > 0 \qquad \forall k \in \mathcal{I} : k \ge k_0. \tag{55}$$

On the other hand we have for $k \in \mathcal{I}, k \geq k_0$:

$$\frac{\|\delta x_{k,\ell}\|}{|\delta x_{k,\ell}|} \geq \frac{|f'_{x_k}\delta x_{k,\ell}|}{\|f'_{x_k}\||\delta x_{k,\ell}|} \overset{(52)}{\geq} M_0 \frac{\omega_k R_{x_k}(\delta x_{k,\ell})}{\|f'_{x_k}\||\delta x_{k,\ell}|}$$
$$\overset{(8)}{\geq} \left(\frac{M_0}{\|f'_{x_k}\|}(\omega_k R_{x_k}(\delta x_{k,\ell}))^{2/3}\right)\omega_k^{1/3} \overset{(46),(51)}{\geq} \frac{M_0}{C_0}(W_0 D_0)^{2/3}\omega_k^{1/3} \overset{(42)}{\to} \infty.$$

and thus

$$\lim_{k\in\mathcal{I}\to\infty} \frac{|\delta x_{k,\ell}|}{\|\delta x_{k,\ell}\|} = 0.$$

Via Lemma 2.2 we conclude that the normalized sequence $\delta x_{k,\ell}/\|\delta x_{k,\ell}\|$ converges to 0 weakly in X. Since $f'_{x_k} \to f'_*$ strongly, we obtain a contradiction to (55):

$$\lim_{k\in\mathcal{I}\to\infty} \frac{|f'_{x_k}\delta x_{k,\ell}|}{\|\delta x_{k,\ell}\|} = f'_* 0 = 0.$$

This is due to a standard result in functional analysis which states that the duality product is continuous with respect to strong convergence in the dual space and weak convergence in the primal space. □

If x_k does not converge, we can still show convergence properties for f'_{x_k}, following the standard pattern that continuity of f'_{x_k} yields subsequential convergence of f'_{x_k} to 0, while uniform continuity yields convergence of the whole sequence.

Theorem 4.4. *Let f be continuously Fréchet differentiable. Assume that $f(x_k)$ is bounded from below and H_x satisfies the Gårding inequality* (7)*. Further, assume that R_x satisfies* (8)*. Then*

$$\liminf_{k\to\infty} \|f'_{x_k}\| = 0. \tag{56}$$

If x_k converges, or if f' is uniformly continuous, then

$$\lim_{k\to\infty} \|f'_{x_k}\| = 0.$$

Proof. If x_k converges, then $\|f'_{x_k}\| \to 0$ by Theorem 4.3.

Otherwise, for the purpose of contradiction we assume that $\|f'_{x_k}\|$ is bounded away from zero. Then x_k is a Cauchy sequence in X by (44), and hence convergent to a limit point x_* by completeness of X. This is a contradiction our premise that x_k does not converge and hence (56) must hold.

It remains to assert that $\mathcal{L}^\nu$ is finite for any $\nu > 0$ if f'_x is uniformly continuous. For this we use a standard trick (cf. e.g. [4, Thm 6.4.6]), exploiting uniform continuity of the function $x \to f'_x$. For any index $n \in \mathcal{L}^\nu$ choose the first index $k(n) \in \mathbb{N} \setminus \mathcal{L}^{\nu/2}$ that satisfies $k(n) > n$. Then $\{j : n \leq j < k(n)\} \subset \mathcal{L}^{\nu/2}$ and by (44)

$$\lim_{n\to\infty} \|x_n - x_{k(n)}\| \leq \lim_{n\to\infty} \sum_{j=n}^{j<k(n)} \|\delta x_j\| = 0$$

and thus, if $\mathcal{L}^\nu$ was infinite, $\|f'_{x_n} - f'_{x_{k(n)}}\| \to 0$. However, eventually

$$\|f'_{x_n}\| \geq \nu,$$
$$\|f'_{x_{k(n)}}\| \leq \nu/2.$$

Hence, we have a contradiction and $\mathcal{L}^\nu$ must be finite. This argument holds for every $\nu > 0$ and thus implies $\lim_{k\to\infty} \|f'_{x_k}\| = 0$. □

4.2 Local convergence

Next we consider local convergence of our method towards a local minimizer x_*. We will first show under some convexity assumptions that a computed sequence converges to x_* if is started close enough. Then we will show under additional smoothness assumptions that our globalization scheme does not interfere with any method to compute search directions and finally we will show local superlinear convergence if directional minimizers along inexact Newton steps are used as trial steps.

Let us start with some auxiliary estimates, which capture the effect of positive curvature of H_x along a directional minimizer. These estimates do not rely on a fraction of Cauchy decrease condition:

Lemma 4.5. *Let δv be a directional minimizer and*

$$\gamma_{\delta v} := \frac{H_x(\delta v, \delta v)}{\|\delta v\|^2} \geq 0.$$

Then we have the following estimates:

$$m_x^\omega(\delta v) \leq -\frac{\gamma_{\delta v}}{2}\|\delta v\|^2 \tag{57}$$

$$\gamma_{\delta v}\|\delta v\| \leq \|f'_x\|. \tag{58}$$

Proof. Equation (57) directly follows from (14), taking into account positivity of R_x.

Equation (12) yields

$$\gamma_{\delta v}\|\delta v\|^2 \leq \gamma_{\delta v}\|\delta v\|^2 + \frac{\omega}{2}R_x(\delta v) = H_x(\delta v, \delta v) + \frac{\omega}{2}R_x(\delta v) = -f'_x\delta v \leq \|f'_x\|\|\delta v\|$$

and thus (58). □

4.2.1 Convergence to local minmizers

Our basic theoretical framework comprises the following assumptions, which we impose throughout the whole section. For fast local convergence we will later impose further smoothness assumptions.

Assumption 4.6. Let $x_* \in X$ be a local minimizer, and assume that there exists a neighborhood U of x_* with the following properties:

(i) The assumptions of Theorem 4.4(i) on global convergence hold in U.

(ii) For $\varepsilon > 0$ define the local level sets

$$L_\varepsilon := \{x \in U : f(x) \leq f(x_*) + \varepsilon\} \subset U.$$

Assume that these sets form a neighborhood base of x_*, i.e., each neighborhood of x_* contains one of these level sets (and hence all with smaller ε). This implies that x_* is a local minimizer. The converse is not true, in general.

(iii) We have the estimate

$$\exists \alpha < \infty \ : \ f(x) - f(x_*) \leq \alpha\|f'_x\|\|x - x_*\| \quad \forall x \in U.$$

This holds with $\alpha = 1$, if f is convex in U, and implies, together with (ii) that x_* is an isolated critical point.

(iv) The ellipticity assumption (11) for H_x holds in U:

$$\exists \gamma > 0 : \quad \gamma \|\delta x\|^2 \leq H_x(\delta x, \delta x) \quad \forall x \in X, \forall \delta x \in X$$

If f is twice differentiable and $H_x = f''_x$, then this implies convexity of f in U and thus (iii).

It follows from continuity of f that the interior of L_ε is non-empty, and (ii) implies via differentiability of f that $f'_{x_*} = 0$. Alternatively to (iii) we could assume continuous invertibility of the mapping $x \to f'_x$.

First we show that if our algorithm comes close to a local minimizer with the above properties, then it will converge towards this minimizer.

Lemma 4.7. *If Assumption 4.6 holds, then there exists $\varepsilon_0 > 0$ such that if $x \in L_\varepsilon$, and δx is an acceptable directional minimizer then $x + \delta x \in L_\varepsilon$ for all $0 < \varepsilon < \varepsilon_0$.*

Proof. By Assumption 4.6(ii) we can choose for any neighborhood $V \subset U$ of x_* an $\varepsilon > 0$, such that $L_\varepsilon \subset V$. Recall that H_x is uniformly elliptic on U and thus on V with a constant $\gamma > 0$. By continuity of f'_x we can in turn choose V, such that $\|f'_x\| \leq \gamma^{-1}\nu$ for every $x \in V$, for every given $\nu > 0$. It follows by (58) that $\|\delta x\| \leq \nu$ for every acceptable directional minimizer, and thus $x + \delta x \in U$, as long as V and ν have been chosen sufficiently small, and $x \in L_\varepsilon \subset V$. Thus, we conclude by the descent property that $x + \delta x \in L_\varepsilon \subset V$, again. □

Proposition 4.8. *Suppose that Assumption 4.6 holds. If the sequence of iterates, generated by our algorithm comes sufficiently close to x_*, then it converges to x_*.*

Proof. By Lemma 4.7 the sequence, generated by our algorithm remains in L_ε, as long as one iterate comes sufficiently close to x_*. Thus, $\|x_k - x_*\|$ remains bounded. Theorem 4.4 implies $\|f'_{x_{k_j}}\| \to 0$, at least for a subsequence x_{k_j}, and thus

$$f(x_{k_j}) - f(x_*) \leq \alpha \|f'_{x_{k_j}}\| \|x_{k_j} - x_*\| \to 0.$$

So, for each $\varepsilon > 0$, $x_{k_j} \in L_\varepsilon$, eventually. Since x_k does not leave level sets by Lemma 4.7, the same holds for the whole sequence. Since the level sets form a neighborhood base of x_*, we conclude that $x_k \to x_*$. □

4.2.2 Asymptotic behaviour of the globalization scheme

Next, we will study conditions under which the effect of globalization vanishes close to x_*. We do this by comparing the actually computed step δx, some directional minimizer of the model function m^ω_x with a step Δx in the same direction computed for $\omega = 0$, i.e., the minimizer of

$$q_x(v) = f(x) + f'_x v + \frac{1}{2} H_x(v, v) = f(x) + m^0_x(v)$$

on $\operatorname{span}\{\delta x\}$. Close to x_* the Hessian H_x is elliptic by assumption, so that Δx is well defined.

Considering a sequence $x_k \to x_*$ and corresponding sequences ω_k and δx_k, generated by our algorithm, we will show in the following that the quotients

$$\lambda_k := \frac{\|\delta x_k\|}{\|\Delta x_k\|} \leq 1$$

tend to 1. Note that by definition of Δx_k and δx_k we have $\delta x_k = \lambda_k \Delta x_k$.

For the following we will only need a slightly weaker version of the upper bound of (8):

$$x_k \to x_*, v_k \to 0 \quad \text{implies} \quad \lim_{k\to\infty} \frac{R_{x_k}(v_k)}{\|v_k\|^2} = 0. \tag{59}$$

Lemma 4.9. *Let x_k be any sequence of iterates with accepted steps δx_k, such that H_{x_k} are uniformly elliptic. Then*

$$\lim_{k\to\infty} \frac{\omega_k R_{x_k}(\delta x_k)}{\|\delta x_k\|^2} = 0 \quad \Rightarrow \quad \lim_{k\to\infty} \lambda_k = 1.$$

Proof. To show the above equivalence we insert δx_k and Δx_k into (12) and set

$$\gamma_k := \frac{H_{x_k}(\delta x_k, \delta x_k)}{\|\delta x_k\|^2} = \frac{H_{x_k}(\Delta x_k, \Delta x_k)}{\|\Delta x_k\|^2}.$$

We obtain from (12) (with $\omega = 0$ for Δx):

$$\begin{aligned} \|\delta x_k\| \left(\frac{\omega_k}{2} \frac{R_{x_k}(\delta x_k)}{\|\delta x_k\|^2} + \gamma_k \right) &\overset{(12)}{=} |f'_{x_k}\delta x_k| / \|\delta x_k\| \\ &= |f'_{x_k}\Delta x_k| / \|\Delta x_k\| \underset{\omega=0}{\overset{(12)}{=}} \|\Delta x_k\| \gamma_k \end{aligned}$$

By assumption, the sequence γ_k is positive and bounded away from 0 and thus we obtain by division

$$1 \geq \lambda_k = \frac{\|\delta x_k\|}{\|\Delta x_k\|} = \frac{\gamma_k}{\frac{\omega_k}{2} \frac{R_{x_k}(\delta x_k)}{\|\delta x_k\|^2} + \gamma_k}$$

The right hand side tends to 1, if $\frac{\omega_k R_{x_k}(\delta x_k)}{\|\delta x_k\|^2} \to 0$. □

The following result is an immediate consequence:

Corollary 4.10. *Let x_k be a converging sequence, such that H_{x_k} are uniformly elliptic, and suppose that* (59) *holds. If ω_k is bounded, then* $\lim_{k\to\infty} \lambda_k = 1$.

To show boundedness of ω_k we consider the acceptance indicators η_k as defined in (24) and show that they tend to 1 asymptotically if the quadratic model is really a second order approximation of f in the sense of (10):

$$\lim_{k\to\infty} \frac{w_{x_k}(\delta x_k)}{\|\delta x_k\|^2} = 0.$$

It can be shown that such a condition holds, if f is twice continuously differentiable in a neighborhood of x_* and $H_x = f''_x$.

Proposition 4.11. *Suppose that $x_k \to x_*$ and assume that the second order approximation error estimate* (10) *holds. Then, independently of the choice of $\omega_k \geq 0$ we conclude for η_k, defined in* (24)*:*

$$\liminf_{k\to\infty} \eta_k \geq 1$$

for any corresponding sequence of directional minimizers δv_k.

Proof. Since, by assumption $x_k \to x_*$, we also have $\|f'_{x_k}\| \to 0$ and thus $\|\delta v_k\| \to 0$ by (58). Thus, by (10) we conclude

$$\lim_{k\to\infty} \frac{w_{x_k}(\delta v_k)}{\|\delta v_k\|^2} = 0, \quad \text{while by (57) we have} \quad \frac{m_{x_k}^{\omega_k}(\delta v_k)}{\|\delta v_k\|^2} \le -\frac{\gamma}{2}.$$

Thus, taken together, we obtain

$$\lim_{k\to\infty} \frac{w_{x_k}(\delta v_k)}{m_{x_k}^{\omega_k}(\delta v_k)} = 0.$$

Hence, by definition (2) (recall that $m_{x_k}^{\omega_k}(\delta v_k) < 0$)

$$\begin{aligned}\liminf_{k\to\infty} \eta_k &= \liminf_{k\to\infty} \frac{f(x_k+\delta v_k) - f(x_k)}{m_{x_k}^{\omega_k}(\delta v_k)} = \liminf_{k\to\infty} \frac{m_{x_k}^{\omega_k}(\delta v_k) - \frac{\omega_k}{6} R_{x_k}(\delta v_k) + w_{x_k}(\delta v_k)}{m_{x_k}^{\omega_k}(\delta v_k)} \\ &\ge \lim_{k\to\infty} \left(1 + \frac{w_{x_k}(\delta v_k)}{m_{x_k}^{\omega_k}(\delta v_k)}\right) = 1.\end{aligned}$$

□

Theorem 4.12. *In addition to Assumption 4.6 suppose that* (59) *and* (10) *hold in* U *along* x_k *generated by our algorithm. If* x_k *comes sufficiently close to* x_* *then* $x_k \to x_*$, ω_k *is bounded and* $\lambda_k \to 1$.

Moreover, eventually, $\iota_{\mathrm{mod}} = 0$ *and all calls of subroutine "CompAccStep" terminate after one iteration.*

Proof. By Theorem 4.8 we conclude that $x_k \to x_*$ and by (58) $\|\delta v_k\| \to 0$ for any directional minimizer δv_k of $m_{x_k}^{\omega_k}$. In particular the quasi-Cauchy steps δx_k^C and the accepted steps δx_k tend to 0 in $\|\cdot\|$-norm.

By Proposition 4.11 eventually every trial step is accepted with some $\eta_k > \overline{\eta}$. Hence, subroutine "CompAccStep" terminates at the first step and by our algorithmic restriction (30) ω_k is not increased anymore so that it follows that ω_k is bounded above. This and $\|\delta x_k^C\| \to 0$ implies via (59) that $\lim_{k\to\infty} \frac{\omega_k R_{x_k}(\delta x_k^C)}{\|\delta x_k^C\|^2} = 0$ so that $\iota_{\mathrm{mod}} = 0$, eventually.

Finally, Lemma 4.9, taking into account boundedness of ω_k and $\|\delta x_k\| \to 0$ yields $\lambda_k \to 1$. □

4.2.3 Fast local convergence along Newton directions

As an illustration of this result consider the case, where δx is computed from a Newton direction Δx^N in case that $H_x = f''_x$ is elliptic:

$$\Delta x^N \in \operatorname{argmin} q_x \quad \Leftrightarrow \quad f'_x v + H_x(\Delta x^N, v) = 0 \quad \forall v \in X.$$

In the following, we denote by $\|v\|_{H_x} := H_x(v,v)^{1/2}$ the energy norm. Under our assumptions, we have equivalence of norms:

$$\exists \gamma > 0, \Gamma < \infty : \quad \gamma\|v\|^2 \le \|v\|_{H_x}^2 \le \Gamma\|v\|^2.$$

It is well known that the sequence, generated by these steps converges locally superlinearly to x_* as long as f is twice continuously differentiable in a neighbourhood of x_*. Let us denote by δx^N the directional minimizer of m_x^ω in Newton direction.

Lemma 4.13. *δx^N satisfies the fraction of Cauchy decrease condition* (21) *if*

$$\beta \le 1 - \frac{\omega}{3}\frac{R_x(\Delta x^N)}{\|\Delta x^N\|_{H_x}^2}. \tag{60}$$

Proof. We compute, using that δx^N and Δx^N are directional minimizers of m_x^ω and m_x^0:

$$\begin{aligned} m_x^\omega(\delta x^N) &\le m_x^\omega(\Delta x^N) = m_x^0(\Delta x^N) + \frac{\omega}{6}R_x(\Delta x^N) \\ &= -\frac{1}{2}\|\Delta x^N\|_{H_x}^2 + \frac{\omega}{6}R_x(\Delta x^N) = -\frac{1}{2}\left[1 - \frac{\omega}{3}\frac{R_x(\Delta x^N)}{\|\Delta x^N\|_{H_x}^2}\right]\|\Delta x^N\|_{H_x}^2. \end{aligned}$$

Observing that the term in square brackets is greater or equal β by (60) we can continue to compute:

$$m_x^\omega(\delta x^N) \le -\frac{1}{2}\beta\|\Delta x^N\|_{H_x}^2 = \beta m_x^0(\Delta x^N) = \beta \inf m_x^0 \le \beta \inf m_x^\omega \le \beta m_x^\omega(\delta x^C).$$

□

In the following we consider for a sequence x_k the Newton steps Δx_k^N computed at x_k and corresponding directional minimizers δx_k^N of $m_{x_k}^{\omega_k}$.

Theorem 4.14. *Suppose that the conditions of Theorem 4.12 hold and assume that f is twice continuously differentiable. Assume that $\underline{\beta} < 1$ in Condition 3.5.*

Then, if x_k comes sufficiently close to x_*, eventually all δx_k^N are acceptable, so that $x_{k+1} = x_k + \delta x_k^N$, and the sequence x_k converges locally superlinearly to x_*.

Proof. By boundedness of ω_k, equivalence of the norms $\|\cdot\|$ and $\|\cdot\|_{H_x}$, and (59) we obtain that the right hand side of (60) tends to 1 and is thus larger than $\underline{\beta}$, eventually. Thus, eventually, δx_k^N is acceptable in terms of Condition 3.5 (recall that eventually $\iota_{\text{mod}} = 0$ by Theorem 4.12), and also in terms of Condition 3.7 by Proposition 4.11. Hence $x_{k+1} = x_k + \delta x_k^N$. Now we compute

$$\begin{aligned} \frac{\|x_k + \delta x_k^N - x_*\|_{H_x}}{\|x_k - x_*\|_{H_x}} &\le \frac{\|x_k + \Delta x_k^N - x_*\| + \|\delta x_k^N - \Delta x_k^N\|_{H_x}}{\|x_k - x_*\|_{H_x}} \\ &= \frac{\|x_k + \Delta x_k^N - x_*\|_{H_x}}{\|x_k - x_*\|_{H_x}} + \frac{(1-\lambda_k)\|\Delta x_k^N\|_{H_x}}{\|x_k - x_*\|_{H_x}}. \end{aligned}$$

The first term of the right hand side vanishes asymptotically due to local superlinear convergence of Newton's method, which also implies $\frac{\|\Delta x_k^N\|_{H_x}}{\|x_k - x_*\|_{H_x}} \to 1$. Then the second term vanishes asymptotically due to $\lambda_k \to 1$ by Theorem 4.12 and thus

$$\lim_{k\to\infty} \frac{\|x_{k+1} - x_*\|_{H_x}}{\|x_k - x_*\|_{H_x}} = 0.$$

By induction we conclude superlinear convergence of x_k to x_*, also w.r.t $\|\cdot\|$ by equivalence of norms. □

References

[1] R.A. Adams. *Sobolev Spaces.* Academic Press, 1975.

[2] Coralia Cartis, Nicholas Gould, and Philippe Toint. Adaptive cubic regularisation methods for unconstrained optimization. part i: motivation, convergence and numerical results. *Mathematical Programming*, pages 1–51, 2009. 10.1007/s10107-009-0286-5.

[3] Coralia Cartis, Nicholas Gould, and Philippe Toint. Adaptive cubic regularisation methods for unconstrained optimization. part ii: worst-case function- and derivative-evaluation complexity. *Mathematical Programming*, pages 1–25, 2010. 10.1007/s10107-009-0337-y.

[4] A.R. Conn, N.I.M. Gould, and P.L. Toint. *Trust-Region Methods.* SIAM, 2000.

[5] B. Dacorogna. *Direct Methods in the Calculus of Variations.* Springer, 2^{nd} edition, 2008.

[6] P. Deuflhard and M. Weiser. Local inexact Newton multilevel FEM for nonlinear elliptic problems. In M.-O. Bristeau, G. Etgen, W. Fitzgibbon, J.-L. Lions, J. Periaux, and M. Wheeler, editors, *Computational science for the 21st century*, pages 129–138. Wiley, 1997.

[7] P. Deuflhard and M. Weiser. Global inexact Newton multilevel FEM for nonlinear elliptic problems. In W. Hackbusch and G. Wittum, editors, *Multigrid Methods V*, Lecture Notes in Computational Science and Engineering, pages 71–89. Springer, 1998.

[8] I. Ekeland and R. Témam. *Convex Analysis and Variational Problems.* Number 28 in Classics in Applied Mathematics. SIAM, 1999.

[9] A. Griewank. The modification of Newton's method for unconstrained optimization by bounding cubic terms. Technical Report NA/12, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, 1981.

[10] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE constraints*, volume 23 of *Mathematical Modelling: Theory and Applications.* Springer, New York, 2009.

[11] C. T. Kelley and E. W. Sachs. A trust region method for parabolic boundary control problems. *SIAM J. Optim.*, 9(4):1064–1081, 1999. Dedicated to John E. Dennis, Jr., on his 60th birthday.

[12] Olga A. Ladyzhenskaya and Nina N. Ural′tseva. *Linear and quasilinear elliptic equations.* Translated from the Russian by Scripta Technica, Inc. Translation editor: Leon Ehrenpreis. Academic Press, New York-London, 1968.

[13] J. L. Lions. *Optimal control of systems governed by partial differential equations*, volume 170 of *Grundlehren der mathematischen Wissenschaften.* Springer-Verlag, 1971.

[14] E. Schechter. *Handbook of Analysis and its Foundations.* Academic Press, 1997.

[15] Ph. L. Toint. Global convergence of a class of trust-region methods for nonconvex minimization in Hilbert space. *IMA J. Numer. Anal.*, 8(2):231–252, 1988.

[16] Philippe L. Toint. Nonlinear stepsize control, trust regions and regularizations for unconstrained optimization. *Optim. Methods Softw.*, 28(1):82–95, 2013.

[17] Fredi Tröltzsch. *Optimal control of partial differential equations*, volume 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. Theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels.

[18] M. Ulbrich. *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*. MOS-SIAM Series on Optimization 11. SIAM, 2011.

[19] M. Ulbrich, S. Ulbrich, and M. Heinkenschloss. Global convergence of trust-region interior-point algorithms for infinite-dimensional nonconvex minimization subject to pointwise bounds. *SIAM J. Control Optim.*, 37(3):731–764, 1999.

[20] M. Weiser, P. Deuflhard, and B. Erdmann. Affine conjugate adaptive newton methods for nonlinear elastomechanics. *Opt. Meth. Softw.*, 22(3):413–431, 2007.

[21] E. Zeidler. *Nonlinear Functional Analysis and its Applications*, volume III. Springer, New York, 1985.

[22] E. Zeidler. *Nonlinear Functional Analysis and its Applications*, volume II/A. Springer, New York, 1990.

[23] E. Zeidler. *Nonlinear Functional Analysis and its Applications*, volume II/B. Springer, New York, 1990.