Newton's Method and Secant Method
for Set-Valued Mappings

Robert Baier[1] and Mirko Hessel-von Molo[2]

[1] University of Bayreuth, Chair of Applied Math., 95440 Bayreuth, Germany,
`robert.baier@uni-bayreuth.de`
[2] University of Paderborn, Chair of Applied Math., 33098 Paderborn, Germany,
`mirkoh@math.uni-paderborn.de`

Abstract. For finding zeros or fixed points of set-valued maps, the fact that the space of convex, compact, nonempty sets of $\mathbb{R}^n$ is not a vector space presents a major disadvantage. Therefore, fixed point iterations or variants of Newton's method, in which the derivative is applied only to a smooth single-valued part of the set-valued map, are often applied for calculations. We will embed the set-valued map with convex, compact images (i.e. by embedding its images) and shift the problem to the Banach space of directed sets. This Banach space extends the arithmetic operations of convex sets and allows to consider the Fréchet-derivative or divided differences of maps that have embedded convex images. For the transformed problem, Newton's method and the secant method in Banach spaces are applied via directed sets. The results can be visualized as usually nonconvex sets in $\mathbb{R}^n$.

Keywords: set-valued Newton's method, set-valued secant method, Gauß-Newton method, directed sets, embedding of convex compact sets

1 Introduction

In this article we illustrate how Newton's method and the similar secant method can be applied to solve two standard problems for set-valued mappings with convex, compact images. Our key method is to embed these problems into a Banach space setting, which allows to reformulate them as zero-finding problems. For references on uses of Newton's method in Banach spaces see e.g. [12,6]. We would like to point out that the ideas presented here differ from the pioneering works (e.g. [10,1]) based on homotopy methods or mutational equations. We start this article with a brief introduction to the Banach space of directed sets. In Sections 2 and 3, we apply this to introduce the Fréchet-derivative as well as Newton's method and the secant method for set-valued maps. We conclude with several numerical examples in Section 4.

1.1 Directed sets

Directed sets provide a way to interpret compact, convex, nonempty sets of $\mathbb{R}^n$, i.e. sets in $\mathcal{C}(\mathbb{R}^n)$, as elements of a Banach space. They were introduced in [3]

(see also references therein), their visualization is studied in [3, Part II]. For $C, D \in \mathcal{C}(\mathbb{R}^n)$ and $M \in \mathbb{R}^{m \times n}$, we write $C + D$ for the *Minkowski sum*, $\lambda \cdot C$ for the *multiplication with non-negative scalars* and $M \cdot C$ for the *linear image*.

Definition 1. Directed intervals resp. sets $\overrightarrow{A} \in \mathcal{D}(\mathbb{R}^n)$ *are defined recursively.*

For $n = 1$, a directed interval is a pair $\overrightarrow{A} = (a_1(l))_{l=\pm 1}$ of two real numbers. (We will also use the notation $\overrightarrow{A} = \overrightarrow{[c,d]}$, where $a_1(-1) = -c$ and $a_1(1) = d$.) In this case, the norm *of $\overrightarrow{A}$ is given by $\|\overrightarrow{A}\|_1 = \max_{l=\pm 1} |a_1(l)|$.*

For $n \geq 2$, a directed set has two components $\overrightarrow{A} = \left(\overrightarrow{A_{n-1}(l)}, a_n(l)\right)_{l \in S_{n-1}}$, given by a map $\overrightarrow{A_{n-1}} : S_{n-1} \to \mathcal{D}(\mathbb{R}^{n-1})$ that is uniformly bounded with respect to $\|\cdot\|_{n-1}$, and a continuous real-valued function $a_n(\cdot)$. The norm *is defined by*

$$\|\overrightarrow{A}\| := \|\overrightarrow{A}\|_n := \max\{\sup_{l \in S_{n-1}} \|\overrightarrow{A_{n-1}(l)}\|_{n-1}, \max_{l \in S_{n-1}} |a_n(l)|\}.$$

As a space of functions on S_{n-1} with values in a linear space, with the usual pointwise definitions of addition and scalar multiplication, $\mathcal{D}(\mathbb{R}^n)$ is a linear space. With the norm defined above, $\mathcal{D}(\mathbb{R}^n)$ is a Banach space that can be equipped with a partial ordering that represents inclusion of sets (see [3, Part I]). Sets in $\mathcal{C}(\mathbb{R}^n)$ can be *embedded* in this Banach space via $J_n : \mathcal{C}(\mathbb{R}^n) \to \mathcal{D}(\mathbb{R}^n)$ as

$$J_1([c,d]) := \overrightarrow{[c,d]} = (-c, d),$$
$$J_n(C) := \left(J_{n-1}(\Pi_{n-1,l} \cdot Y(l,C)), \delta^*(l,C)\right)_{l \in S_{n-1}} \quad (n \geq 2).$$

Here, $\delta^*(l, C)$ denotes the *support function* of C in direction $l \in S_{n-1}$, $Y(l, C)$ the corresponding *supporting face* (i.e. the maximizers of $\langle l, \cdot \rangle$ on the set C), and $\Pi_{n-1,l}$ is a linear *projection* which maps the orthogonal complement of $\text{span}\{l\}$ into $\mathbb{R}^{n-1}$ and l to the origin in $\mathbb{R}^{n-1}$, see [3, Part I] for details.

Proposition 1 ([3]). *The embedding $J_n : \mathcal{C}(\mathbb{R}^n) \to \mathcal{D}(\mathbb{R}^n)$ is positively linear:*

$$J_n(\lambda \cdot C + \mu \cdot D) = \lambda \cdot J_n(C) + \mu \cdot J_n(D) \qquad (C, D \in \mathcal{C}(\mathbb{R}^n), \ \lambda, \mu \geq 0)$$

In contrast to formulations using pairs of sets (see [3] for references), a directed set (and thus a difference of embedded convex sets) can be visualized as a subset in $\mathbb{R}^n$ which is usually nonconvex. The visualization of the directed set is plotted with arrows (outer normals for embedded convex sets, inner ones for their inverses) on the boundaries (see [3, Part II] for details).

2 Fréchet-Derivative for Set-Valued Maps

As $\mathcal{D}(\mathbb{R}^m)$ forms a Banach space, we use the usual definition of the Fréchet-derivative of a function $\overrightarrow{F} : \mathbb{R}^n \to \mathcal{D}(\mathbb{R}^m)$. To lift a set-valued map $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ with images in $\mathcal{C}(\mathbb{R}^m)$ to the space $\mathcal{D}(\mathbb{R}^m)$, we will use the notation $\overrightarrow{F}(x) = J_m(F(x))$. If $\overrightarrow{F} : \mathbb{R}^n \to \mathcal{D}(\mathbb{R}^m)$ is Fréchet-differentiable, F is called *directed differentiable*. This notion of the derivative generalizes the one of set-valued maps for one variable studied in [8,4].

Lemma 1 (see [2]). *Let $F : \mathbb{R}^n \to \mathcal{C}(\mathbb{R}^m)$ and $w \in \mathbb{R}^n$.*

(i) If $F(x) = U$, then $\overrightarrow{F}'(x)(w) = J_m(\{0\})$.
(ii) If $F(x) = r(x) \cdot G(x) + h(x)$ with $r : \mathbb{R}^n \to [0, \infty)$, $h : \mathbb{R}^n \to \mathbb{R}^m$, $G : \mathbb{R}^n \to \mathcal{C}(\mathbb{R}^m)$ which are (directed) Fréchet-differentiable, then $\overrightarrow{F}'(x)(w) = r'(x) \cdot w \cdot \overrightarrow{G}(x) + r(x) \cdot \overrightarrow{G}'(x)(w) + J_m(\{h'(x) \cdot w\})$.
Especially, if $F(x) = r(x) \cdot U$ with $U \in \mathcal{C}(\mathbb{R}^m)$, then $\overrightarrow{F}'(x)(w) = r'(x) \cdot w \cdot \overrightarrow{U}$.
(iii) If $\overrightarrow{F}(x) = r(x) \cdot \overrightarrow{G}(x) + \overrightarrow{H}(x)$ and $r : \mathbb{R}^n \to \mathbb{R}$, $\overrightarrow{G}, \overrightarrow{H} : \mathbb{R}^n \to \mathcal{D}(\mathbb{R}^m)$ are Fréchet-differentiable, then $\overrightarrow{F}'(x)(w) = r'(x) \cdot w \cdot \overrightarrow{G}(x) + r(x) \cdot \overrightarrow{G}'(x)(w) + \overrightarrow{H}'(x)(w)$.

Proof. (i) A constant map in a Banach space has 0 as Fréchet-differential.
(ii) follows immediately from (iii) with $\overrightarrow{G}(x) = J_m(G(x))$, $J_m(r(x) \cdot G(x)) = r(x) \cdot J_m(G(x))$ and $\overrightarrow{H}(x) = J_m(\{h(x)\})$. Since $-J_m(\{h(x)\}) = J_m(\{-h(x)\})$ we can directly show from the definition that $\overrightarrow{H}'(x)(w) = J_m(\{h'(x) \cdot w\})$.
(iii) Clearly, the sum and the product of two functions are Fréchet-differentiable, if the operands have this property. $\qquad\square$

We now study explicit formulas for the Fréchet-derivative.

Proposition 2. *Let $F : \mathbb{R}^n \to \mathcal{C}(\mathbb{R}^m)$ and $w, x \in \mathbb{R}^n$.*
(i) If $m = 1$ and $F(x) = [g(x), h(x)]$ with two differentiable functions $g, h : \mathbb{R}^n \to \mathbb{R}$ and $g(x) \le h(x)$ for all $x \in \mathbb{R}^n$, then $\overrightarrow{F}$ is differentiable with

$$\overrightarrow{F}'(x)(w) = \overrightarrow{[g'(x) \cdot w, h'(x) \cdot w]}.$$

(ii) If $m \ge 2$, the Fréchet-differential of $\overrightarrow{F}(x) = (\overrightarrow{F}_{m-1}(l; x), f_m(l; x))_{l \in S_{m-1}}$ is

$$\overrightarrow{F}'(x)(w) = (\overrightarrow{F}'_{m-1}(l; x)(w), f'_m(l; x) \cdot w)_{l \in S_{m-1}}.$$

Proof. (i) apply Lemma 1(ii) to $F(x) = g(x) + (h(x) - g(x)) \cdot [-1, 1]$
(ii) for a proof see [2, Proposition 3.6] $\qquad\square$

3 Newton's Method and Secant Method

For the sake of brevity, we refer to the literature for convergence results about iterative methods in Banach spaces. For Newton's method see e.g. [12,5,6], for the secant method e.g. [9] and for Gauß-Newton method e.g. [7].

3.1 Newton's Method for Directed Sets

Problem 1. Let $F : \mathbb{R}^n \to \mathcal{C}(\mathbb{R}^m)$ be given. We are looking for a solution $\hat{x} \in \mathbb{R}^n$ of $0 \in F(\hat{x})$. Using the partial order "$\le$" on $\mathcal{D}(\mathbb{R}^m)$ and the embedded map $\overrightarrow{F}$, this problem can be transformed into $J_m(\{0\}) \le \overrightarrow{F}(\hat{x})$.

To apply Newton's iteration for directed sets to Problem 1 we choose a starting point $x^0 \in \mathbb{R}^n$ and iteratively obtain $x^{k+1} = x^k + d^k$, where $d^k \in \mathbb{R}^n$

has to solve the linear inequality

$$-\overrightarrow{F}(x^k) \leq \overrightarrow{F}'(x^k)(d^k)\,. \tag{1}$$

Checking this inequality essentially means checking the positivity of a function, which is easily realized algorithmically.

Remark 1. In many publications, e.g. in the study of necessary optimality conditions in non-smooth optimization or optimal control problems, the set-valued map F is given in the form $F(x) = g(x) + H(x)$ with a smooth function $g : \mathbb{R}^n \to \mathbb{R}^m$ and an u.s.c. map $H : \mathbb{R}^n \Rightarrow \mathbb{R}^m$ (representing subdifferentials or normal cones in the mentioned applications), see e.g. [5,6]. The set-valued derivative of F is avoided and an inclusion step of the form

$$0 \in g(x^k) + g'(x^k) \cdot d^k + H(x^k + d^k)$$

is studied. While in this case the computation of d^k typically requires a nonlinear inclusion to be solved, for this splitting of F (with Fréchet-differentiable $\overrightarrow{H}$) Newton's method as suggested above leads to an inclusion problem linearly in d^k, since Newton's method in (1) demands a solution of

$$-J_m(\{g(x^k)\}) - \overrightarrow{H}(x^k) \leq J_m(\{g'(x^k) \cdot d^k\}) + \overrightarrow{H}'(x^k)(d^k)\,.$$

Problem 2. Let $G, H : \mathbb{R}^n \to \mathcal{C}(\mathbb{R}^m)$ be given. We are looking for a solution $\hat{x} \in \mathbb{R}^n$ of $G(\hat{x}) = H(\hat{x})$. Transformed to $\mathcal{D}(\mathbb{R}^m)$, this reads $\overrightarrow{G}(\hat{x}) = \overrightarrow{H}(\hat{x})$.

Using $\overrightarrow{F} = \overrightarrow{G} - \overrightarrow{H}$, Newton's iteration with directed sets for this problem proceeds similarly to the one for Problem 1, with the difference that the updates d^k have to solve the linear equation

$$\overrightarrow{F}'(x^k)(d^k) = -\overrightarrow{F}(x^k)\,. \tag{2}$$

3.2 Secant Method Based on Directed Sets

To apply the secant method as studied e.g. in [11,9], we need a notion of divided differences. Set-valued divided differences in one variable are studied e.g. in [8,4]. Following [11], we introduce multivariate divided differences for directed sets.

Definition 2. *Let $\overrightarrow{F} : \mathbb{R}^n \to \mathcal{D}(\mathbb{R}^m)$ and $u, v, w \in \mathbb{R}^n$ be given such that $u_j \neq v_j$ for $j = 1, \ldots n$. We define*

$$\overrightarrow{F}[u,v]^j := \tfrac{1}{u_j - v_j} \cdot \left(\overrightarrow{F}(u_1, \ldots, u_{j-1}, u_j, v_{j+1}, \ldots, v_n) \right.$$
$$\left. -\overrightarrow{F}(u_1, \ldots, u_{j-1}, v_j, v_{j+1}, \ldots, v_n) \right) \in \mathcal{D}(\mathbb{R}^m)\,,$$
$$\overrightarrow{F}[u,v](w) := \sum_{j=1}^n w_j \cdot \overrightarrow{F}[u,v]^j \in \mathcal{D}(\mathbb{R}^m)\,.$$

In the limit for $u \to v$, this definition yields the Fréchet-differential $\overrightarrow{F}'(u)$.

This approach generalizes the divided differences in one variable and satisfies the characteristic equation

$$\overrightarrow{F}[u,v](u-v) = \overrightarrow{F}(u) - \overrightarrow{F}(v) \quad (u,v \in \mathbb{R}^n). \tag{3}$$

To apply the *secant method* to Problem 1, one chooses two starting points $x^{-1}, x^0 \in \mathbb{R}^n$ and iteratively computes $x^{k+1} = x^k + d^k$, where d^k solves

$$-\overrightarrow{F}(x^k) \le \overrightarrow{F}[x^k, x^{k-1}](d^k). \tag{4}$$

To treat Problem 2 with the secant method, we again set $\overrightarrow{F} = \overrightarrow{G} - \overrightarrow{H}$, choose initial points x^{-1} and x^0 and compute $x^{k+1} = x^k + d^k$, where d^k solves

$$\overrightarrow{F}[x^k, x^{k-1}](d^k) = -\overrightarrow{F}(x^k). \tag{5}$$

An advantage of (5) is that the derivative $\overrightarrow{F}'$ is avoided and replaced by differences of function values. It is obvious that the iterations (2) and (5) generalize their pointwise analogues.

4 Examples

In this section we present numerical examples for the procedures proposed above.

Example 1. Let $U_1, \ldots, U_M \in \mathcal{C}(\mathbb{R}^n)$ be such that the vectors $J_n(U_i)$ are linearly independent in $\mathcal{D}(\mathbb{R}^n)$. Consider an unknown convex combination $U_0 = \sum_{i=1}^M \lambda_i \cdot U_i$ with coefficients $\lambda_i \ge 0$, $i = 1, \ldots, M$ and $\sum_{i=1}^M \lambda_i = 1$. In order to apply Newton's method for Problem 2 to compute the vector $\lambda \in \mathbb{R}^M$, we define $G(x) = \sum_{i=1}^M x_i \cdot U_i$ for $x \in \mathbb{R}^M$ and $H(x) = U_0$, and set $\overrightarrow{F}(x) = \sum_{i=1}^M x_i \cdot J_n(U_i) - J_n(U_0)$. By Proposition 1

$$\overrightarrow{F}(x) = \sum_{i=1}^M (x_i - \lambda_i) \cdot J_n(U_i) \quad \text{and} \quad \overrightarrow{F}'(x)(w) = \sum_{i=1}^M w_i \cdot J_n(U_i),$$

so that the Newton step results in

$$\sum_{i=1}^M d_i^k \cdot J_n(U_i) = -\sum_{i=1}^M (x_i^k - \lambda_i) \cdot J_n(U_i) \text{ and } \sum_{i=1}^M (x_i^k - \lambda_i + d_i^k) \cdot J_n(U_i) = 0.$$

From the assumption of linear independence we obtain $d_i^k - \lambda_i + x_i^k = 0$ for $i = 1, \ldots, M$, and hence $d^k = -x^k + \lambda$. This means that already the first iteration step leads to the solution $x^1 = \hat{x} = \lambda$. We obtain the same picture for the secant method. Since

$$\overrightarrow{F}[x^k, x^{k-1}]^i = \frac{1}{x_i^k - x_i^{k-1}} \cdot \left(x_i^k \cdot J_n(U_i) - x_i^{k-1} \cdot J_n(U_i) \right) = J_n(U_i),$$

$$\overrightarrow{F}[x^k, x^{k-1}](d^k) = \sum_{i=1}^M d_i^k \cdot \overrightarrow{F}[x^k, x^{k-1}]^i = \sum_{i=1}^M d_i^k \cdot J_n(U_i),$$

the left-hand side of (5) equals the one of the Newton step (2).

From now on, we consider Problem 2.

Example 2. Let $\alpha, \beta \in \mathbb{R}$ and consider the functions $g(v) = \alpha\|v\|_2^2 v + \beta v$, $h(v) = \|v\|_2 v$ for $v \in \mathbb{R}^2$. We consider Problem 2 with $G, H : \mathbb{R} \to \mathcal{C}(\mathbb{R}^2)$, $G(x) = g(x \cdot B_1(0))$, $H(x) = h(x \cdot B_1(0))$. A simple calculation shows that $g(x \cdot B_1(0)) = (\alpha x^3 + \beta x) \cdot B_1(0)$ and $h(x \cdot B_1(0)) = x^2 \cdot B_1(0)$. The solutions for $\alpha = -1$, $\beta = \frac{3}{4}$ are $\widehat{x} \in \{0, -\frac{3}{2}, \frac{1}{2}\}$. To eliminate the trival solution $\widehat{x} = 0$, we divide the leading term by x and obtain $\overrightarrow{F}(x) = (\alpha x^2 + \beta - x) \cdot J_2(B_1(0))$. For this, one has

$$\overrightarrow{F}'(x) = (2\alpha x - 1) \cdot \overrightarrow{B} \quad \text{with } \overrightarrow{B} = J_2(B_1(0)) \text{ and}$$

$$\overrightarrow{F}[u, v] = \frac{1}{u - v} \cdot \left((\alpha u^2 - u + \beta) - (\alpha v^2 - v + \beta)\right) \cdot \overrightarrow{B} = (\alpha(u + v) - 1) \cdot \overrightarrow{B} \,.$$

For the Newton step we thus get $d^k = -\frac{\alpha(x^k)^2 - x^k + \beta}{2\alpha x^k - 1}$, while the iteration for the secant method results in $d^k = -\frac{\alpha(x^k)^2 - x^k + \beta}{\alpha(x^k + x^{k-1}) - 1}$. In Fig. 1(a) the iterates $\overrightarrow{X}^k := x^k \cdot J_2(B_1(0))$ of the secant method are depicted for the starting values $x^{-1} = 0$, $x^0 = -\frac{3}{4}$. Here, the iterates $k = 0, 2, 3$ are inverses of embedded sets (inner normals), whereas the other iterates are embedded convex sets (outer normals). For $k = -1$, the embedded origin is visualized in the plot.

In Example 3, we solve Problem 2 using the Gauß-Newton method in [7] and the Gauß-secant method for two-dimensional directed sets. Similar to the approximation of convex sets with finitely many supporting hyperplanes, we approximate a set-valued map $\overrightarrow{F} : \mathbb{R}^n \to \mathcal{D}(\mathbb{R}^2)$ by choosing a finite number of unit vectors $l^\mu \in S_1$, $\mu = 1, \ldots, M$ and evaluating $\overrightarrow{F}(x) = (f_1(x; \eta; l))_{\eta=\pm 1}, f_2(x; l))_{l \in S_1}$ in these directions, leading to a discretized map $\overrightarrow{F}_M : \mathbb{R}^n \to \mathbb{R}^{3M}$ defined by

$$\left(\overrightarrow{F}_M(x)\right)_i = \begin{cases} f_1(x; -1; l^i) & \text{if } i = 1, \ldots, M, \\ f_1(x; \ 1; l^{i-M}) & \text{if } i = M+1, \ldots, 2M, \\ f_2(x; l^{i-2M}) & \text{if } i = 2M+1, \ldots, 3M. \end{cases}$$

Recalling the ideas behind the Gauß-Newton method, we replace $\overrightarrow{F}_M(x) = 0$ by the minimization problem $\min_{x \in \mathbb{R}^n} \|\overrightarrow{F}_M(x)\|_2^2$ which is solved iteratively by minimizing $\|\overrightarrow{F}_M(x^k) + \overrightarrow{F}'_M(x^k)(d^k)\|_2^2$ or $\|\overrightarrow{F}_M(x^k) + \overrightarrow{F}_M[x^k, x^{k-1}](d^k)\|_2^2$, respectively. This leads to the normal equations

$$\overrightarrow{F}'_M(x^k)^\top \overrightarrow{F}'_M(x^k)(d^k) = -\overrightarrow{F}'_M(x^k)^\top \overrightarrow{F}_M(x^k) \quad \text{and} \tag{6}$$

$$\overrightarrow{F}[x^k, x^{k-1}]^\top \overrightarrow{F}[x^k, x^{k-1}](d^k) = -\overrightarrow{F}[x^k, x^{k-1}]^\top \overrightarrow{F}_M(x^k) \tag{7}$$

as iterative steps. The (i, j)-th entry of the matrix of divided differences is given for $j = 1, 2$ by the expressions

$$\frac{f_1(x_1^k, x_2^{k+j-2}; -1; l^i) - f_1(x_1^{k+j-2}, x_2^{k-1}; -1; l^i)}{x_j^k - x_j^{k-1}} \quad \text{for } i = 1, \ldots, M,$$

$$\frac{f_1(x_1^k, x_2^{k+j-2}; 1; l^i) - f_1(x_1^{k+j-2}, x_2^{k-1}; 1; l^i)}{x_j^k - x_j^{k-1}} \quad \text{for } i = M+1, \ldots, 2M,$$

$$\frac{f_2(x_1^k, x_2^{k+j-2}); l^{i-2M}) - f_2(x_1^{k+j-2}, x_2^{k-1}; l^{i-2M})}{x_j^k - x_j^{k-1}} \quad \text{for } i = 2M+1, \dots, 3M,$$

which are much simpler to compute than the corresponding partial derivatives.

In the next nonlinear test example we are looking for parameters in the determining matrix of an ellipsoid centered in the origin which coincides with the unit ball.

Example 3. For $x \in \mathbb{R}^2$ we consider $G(x) = A(x) \cdot B_1(0)$ with $A(x) = \begin{pmatrix} x_1 & 0 \\ 0 & x_2 \end{pmatrix}$, $H(x) = B_1(0)$ in Problem 2. In this example, it is rather complicated to evaluate the derivative of $\overrightarrow{F}(x) = \overrightarrow{G}(x) - \overrightarrow{H}(x)$ analytically. Clearly, $\hat{x} = (1,1)^\top$ is one solution and $\overrightarrow{F}'(x) = (\overrightarrow{G}'_1(x;l), g_2(x;l))_{l \in S_1}$ with

$$f_2(x;l) = g_2(x;l) - 1 = \delta^*(l, G(x)) - 1 = \sqrt{l_1^2 x_1^2 + l_2^2 x_2^2} - 1,$$

$$\frac{\partial f_2}{\partial x}(x;l) = \frac{1}{\sqrt{l_1^2 x_1^2 + l_2^2 x_2^2}} \begin{pmatrix} l_1^2 x_1 \\ l_2^2 x_2 \end{pmatrix}^\top.$$

The derivative of $\overrightarrow{F}'_1(x;l)$ within the Gauß-Newton method is complicated, since it involves the projection to $\mathbb{R}$, whereas the calculation of $\overrightarrow{F}[u,v](w)$ in the Gauß-secant method only requires the weighted sum of differences of function values. Table 1 demonstrates that the Gauß-secant method converges rather rapidly, but slower than the Gauß-Newton method. In Fig. 1(b) the iterates $\overrightarrow{X}^k := J_2(A(x^k) \cdot B_1(0))$ of the secant method are depicted for the starting values $x^{-1} = (4,3)^\top$, $x^0 = (3,2)^\top$. Depending on the starting values, convergence to one of the solutions $(\pm 1, \pm 1)^\top$ can be observed.

To sketch the outline of further research, we mention [2]. In this publication, Newton's method for set-valued maps is used to approximate invariant sets, i.e. to find sets P for which $g(P) = P$. This task is similar to Problem 2, if the set is parametrized e.g. as $P(x) = x \cdot B_1(0)$ (see Example 2) and $G(x) = g(P(x))$, $H(x) = h(P(x))$, $h(x) = x$, but considerably more complicated, as the map $\overrightarrow{G}$ depends on a directed set. This leads to a zero-finding problems of the form $0 = \overrightarrow{F}(\overrightarrow{X}) = J_n(g(V_n(\overrightarrow{X}))) - \overrightarrow{X}$, the solution of which requires much more advanced Banach space techniques than in the present paper.

References

1. Aubin, J.P.: Mutational and Morphological Analysis. Tools for Shape Evolution and Morphogenesis. Systems & Control: Foundations & Applications, Birkhäuser Boston Inc., Boston, MA (1999)
2. Baier, R., Dellnitz, M., Hessel-von Molo, M., Kevrekidis, I.G., Sertl, S.: The computation of invariant sets via Newton's method (May 2010), submitted, 21 pages.
3. Baier, R., Farkhi, E.: Differences of Convex Compact Sets in the Space of Directed Sets. Part I: The Space of Directed Sets. Set-Valued Anal. 9(3), 217–245 (2001), Part II: Visualization of Directed Sets. Same volume, 247–252.

8

4. Baier, R., Perria, G.: Set-valued Hermite interpolation. J. Approx. Theory (2011), doi: 10.1016/j.jat.2010.11.004, 24 pages.
5. Dontchev, A.L., Hager, W.W., Veliov, V.M.: Uniform convergence and mesh independence of Newton's method for discretized variational problems. SIAM J. Control Optim. 39(3), 961–980 (electronic) (2000)
6. Dontchev, A.L., Veliov, V.M.: Metric regularity under approximations. Control Cybernet. 38(4B), 1283–1303 (2009)
7. Li, C., Zhang, W.H., Jin, X.Q.: Convergence and uniqueness properties of Gauss-Newton's method. Comput. Math. Appl. 47(6-7), 1057–1067 (2004)
8. Perria, G.: Set-valued interpolation. Bayreuth. Math. Schr. 79, 154 pp. (2007)
9. Potra, F.A.: An error analysis for the secant method. Numer. Math. 38(3), 427–445 (1981/82)
10. Saint-Pierre, P.: Newton and other continuation methods for multivalued inclusions. Set-Valued Anal. 3(2), 143–156 (1995)
11. Schmidt, J.W.: Eine Übertragung der Regula Falsi auf Gleichungen in Banachräumen. II. Nichtlineare Gleichungssysteme. Z. Angew. Math. Mech. 43(3), 97–110 (1963)
12. Yamamoto, T.: A method for finding sharp error bounds for Newton's method under the Kantorovich assumptions. Numer. Math. 49(2–3), 203–220 (1986)

(a) Example 2   (b) Example 3
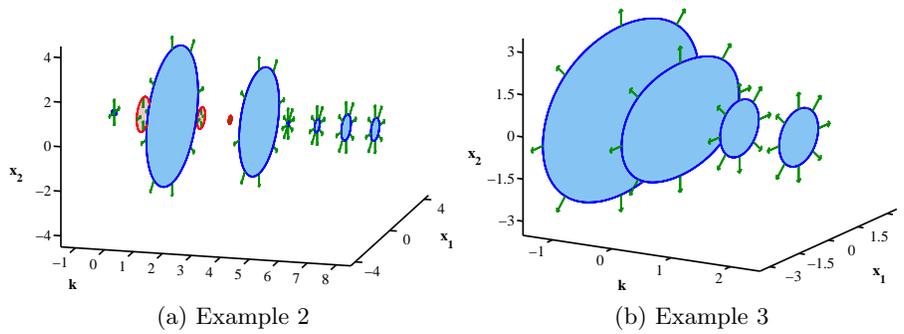
Fig. 1. Iterates of the secant method for two examples

k	x_1^k	x_2^k	$\frac{1}{2}\|\overrightarrow{F}_M(x^k)\|_2^2$	x_1^k	x_2^k	$\frac{1}{2}\|\overrightarrow{F}_M(x^k)\|_2^2$
-1				4.00000	3.00000	149.2847951
0	3.00000000	2.00000000	68.997425636149	3.00000	2.00000	68.9974256
1	1.01002370	1.00647268	0.001644629887	0.98748	0.98710	0.0032343
2	1.00000078	1.00000078	0.000000000012	0.99942	0.99933	0.0000081

Table 1. Iterates x^k of the Gauß-Newton and the secant method for Example 3