GAMMAS 2025, Vol. 7, No. 1, p. 14–30 DOI: 10.14464/gammas.v7i1.810

International Association
of Applied Mathematics and Mechanics
– Archive for Students –

Global Convergence of Semismooth Newton Methods for Quadratic Problems

Hannah Rickmann[a,⋆], Evelyn Herberg[a], Roland Herzog[a]

[a] Interdisciplinary Center for Scientific Computing, Heidelberg University, 69120 Heidelberg, Germany

received 27.12.2024, accepted 26.03.2025, published 06.05.2025

⋆ corresponding author: hannah.rickmann@iwr.uni-heidelberg.de

Abstract: *This paper investigates the semismooth Newton method and its application to solving constrained quadratic optimization problems. We begin by reviewing various generalized concepts of differentiability, such as Clarke's generalized Jacobian, slanting functions, and Newton differentiability, providing a comparative analysis to clarify their differences. We then establish the equivalence between the semismooth Newton method (SSN) and the primal-dual active set algorithm (PDASA), and key global convergence results are summarized. Our study focuses on the cycle behavior of small-dimensional quadratic problems. For the two-dimensional case, we prove global convergence for arbitrary quadratic problems, demonstrating the robustness of the method in this setting. A necessary condition for cycles of certain lengths is derived and used to identify possible cycling patterns, for problems in three dimensions. While only two of the cycling patterns were observed in randomly generated examples, the other remain a theoretical possibility, suggesting further exploration of the method's behavior.*

Keywords: semismooth Newton method, primal-dual active set method, global convergence, cycling behavior

1 Introduction

Optimization of quadratic problems is a fundamental problem in many scientific and industrial fields. These problems arise in various domains, particularly in optimal control problems, as regularization terms in machine learning and as subproblems in Sequential Quadratic Programming (SQP) methods. Let us consider the finite-dimensional constrained quadratic program

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad & f(x) = \frac{1}{2}(x, Ax) - (b, x), \\ \text{subject to} \quad & x \le u, \end{aligned} \tag{QP}$$

where $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$ and $u \in (\mathbb{R}\cup\{\infty\})^n$. By $(\cdot,\cdot)$ we denote the Euclidean inner product in $\mathbb{R}^n$. Problem (QP) is a constrained quadratic problem with unilateral bounds in finite dimensions. We note that considering only unilateral instead of bilateral bounds is a restriction. Further, we assume that A is positive definite and symmetric; therefore, (QP) has a unique solution x^*. The KKT conditions [17, Theorem 12.1] imply, that the corresponding Lagrange multiplier μ^* is also unique since the linear independence constraint qualification (LICQ) holds. For convex quadratic problems, we can write the optimality system of necessary and sufficient conditions in complementarity form [17, Chapters 16.4 and 16.6] as

$$\begin{aligned} & Ax + \mu = b, \\ & u \ge x, \quad \mu \ge 0, \quad (\mu, u - x) = 0. \end{aligned} \tag{1.1}$$

These conditions constitute a linear complementarity problem (LCP), a well-studied class of problems in math-

ematical programming [8]. Problem (1.1) can be equivalently reformulated as a nonsmooth root-finding problem using the max-complementarity function:

$$F(x,\mu) = \begin{pmatrix} Ax+\mu-b \\ \mu - \max(0, \mu + c\,(x-u)) \end{pmatrix} = 0, \tag{1.2}$$

for any $c > 0$ [23]. The well-known semismooth Newton method is a powerful tool to solve such problems. It is also applicable to more general complementarity problems, regardless of their connection to quadratic optimization problems.

Semismooth functions and the corresponding Newton methods were studied in [18] and [19], which demonstrate local superlinear convergence similar to the classical Newton method for differentiable operators. However, their analysis is limited to finite dimensions, which is not suitable, e. g., for applications to optimal control problems or other infinite-dimensional constrained optimization problems.

For infinite-dimensional spaces, the concepts of slanting functions [5] and Newton differentiability [10] were devised. Concurrently, an alternative approach was developed in [3], [2], specifically in the context of infinite-dimensional quadratic problems. The authors introduced a primal-dual active set strategy, which was later found to be interpretable as a semismooth Newton method [9]. There exists extensive research on the application of the primal-dual active set strategy for optimal control problems; see, e. g. [11, 12]. Other methods to solve optimal control problems with bound constraints on the controls include interior-point methods [25] and gradient projection methods [14].

One of the objectives of the present paper is to give an overview about different concepts frequently called "semismoothness" and to clarify the differences between these concepts in Section 2. Second, we explicitly state the semismooth Newton method in Section 3 along with some known global convergence theorems. Finally, we present new findings on convergence of the method for dimensions $n = 2$ and $n = 3$, and provide an abstract cycle condition in Section 4.

2 Differentiability Concepts

The optimality system (1.2) has the form of a root-finding problem, which is commonly solved by Newton's method. To address the lack of differentiability in the classical sense, we review generalized derivatives [7] and the concept of semismooth functions, which was first introduced in [16] and extended in [19], initially in finite dimensions.

To add to our historical overview, we revisit related concepts in infinite-dimensional spaces, specifically slanting functions [5] and Newton differentiability [10]. In the literature, the term "semismooth" is used inconsistently, and *all* of the above concepts are often subsumed under the term "semismoothness". We aim to give an overview and clarify the differences between these concepts. In Fig. A.1 we provide a timeline of concepts and corresponding papers.

2.1 Generalized Derivatives

From now, suppose that X is a Banach space, $D \subseteq X$ is an open subset, and $f\colon D \to \mathbb{R}$ is a Lipschitz continuous function. Clarke in [7] defines *generalized directional derivatives*

$$f^\circ(x;d) = \limsup_{y \to x,\, t \searrow 0} \frac{f(y+td) - f(y)}{t}$$

and *generalized subdifferentials*

$$\partial f(x) = \{\zeta \in X^* \mid (\zeta, d) \le f^\circ(x,d) \text{ for all } d \in X\}.$$

Here X^* denotes the dual space of X. Observe that this definition is more general than the classical directional derivative, as it allows for a limit superior instead of a limit and $y \neq x$. Further, the concept was initially referred to as "generalized gradient" but we now denote it as "generalized subdifferential" for consistency.

In finite-dimensional spaces such as $X = \mathbb{R}^n$, Rademacher's theorem states that Lipschitz continuity on an open subset implies differentiability at almost all points of that open subset; see [6, Theorem 2.5.1]. Denoting the set of points at which f lacks differentiability as Ω_f, we can characterize the generalized subdifferential as

$$\partial f(x) = \operatorname{conv}\{\lim \nabla f(x^{(k)}) \mid x^{(k)} \to x,\ x^{(k)} \notin S \cup \Omega_f\}.$$

where $S \subseteq \mathbb{R}^n$ is an arbitrary Lebesgue null set and conv denotes the convex hull.

For vector valued functions $F\colon \mathbb{R}^n \to \mathbb{R}^m$, we define the *generalized Jacobian* [6, Definition 2.6.1] of F at $x \in \mathbb{R}^n$ as

$$\partial F(x) = \operatorname{conv}\{\lim JF(x^{(k)}) \mid x^{(k)} \to x,\ x^{(k)} \notin \Omega_F\},$$

i. e., as the convex hull of all $m \times n$ matrices Z that are derived as the limit of a sequence in the form $JF(x^{(k)})$ where $x^{(k)} \to x$, $x^{(k)} \notin \Omega_F$, and JF represents the conventional $m \times n$ Jacobian matrix of partial derivatives.

2.2 Semismooth Functions

For finite-dimensional spaces, the notion of semismooth functions has been initially introduced for functionals [16] and later extended to general maps [19].

The concept also exists in infinite dimensions; see Section 2.3.

Definition 2.1 ([19, p. 355]). *A function $F\colon \mathbb{R}^n \to \mathbb{R}^m$ is called* semismooth *at $x \in \mathbb{R}$ if F is Lipschitz continuous in a neighborhood of x and*

$$\lim_{\substack{d^{(k)} \to d,\, t^{(k)} \searrow 0 \\ G^{(k)} \in \partial F(x + t^{(k)} d^{(k)})}} G^{(k)} d^{(k)}$$

exists for all $d \in \mathbb{R}^n$.

The limit in this definition is understood in a way that for all sequences $d^{(k)} \to d$ and $t^{(k)} \searrow 0$ we are free to choose one generalized Jacobian $G^{(k)} \in \partial F\big(x + t^{(k)} d^{(k)}\big)$ for every $k \in \mathbb{N}$. Semismooth functions satisfy many useful properties [16]. For instance, convex functions as well as continuously differentiable functions are always semismooth, and the composition of semismooth functions yields again a semismooth function.

For a function $F\colon \mathbb{R}^n \to \mathbb{R}^m$ that is Lipschitz continuous in a neighborhood of x, semismoothness at x is equivalent to directional differentiability at x together with

$$\max_{G \in \partial F(x+h)} \|F(x+h) - F(x) - Gh\| = o(\|h\|), \tag{2.1}$$

as $h \to 0$ [23, Proposition 2.3].

Example 2.2 ([9, p. 5]). *The generalized subdifferential of the maximum function $x \mapsto \max(0, x)$ at $x \in \mathbb{R}$ can be computed as*

$$\partial \max(0,x) = \begin{cases} 1 & \text{if } x > 0, \\ [0,1] & \text{if } x = 0, \\ 0 & \text{if } x < 0. \end{cases}$$

The semismoothness of the max-function can easily be shown with (2.1). From this it follows that (1.2) is a semismooth problem, consisting of a composition of semismooth functions. In the subsequent analysis, we consistently adopt

$$G(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x \le 0, \end{cases} \tag{2.2}$$

as a specific representative of the set-valued generalized subdifferential of the maximum function at 0.

Example 2.3 ([16, p. 962]). *The locally Lipschitz continuous function*

$$f\colon \mathbb{R} \to \mathbb{R}, \quad f(x) = \begin{cases} x^2 \sin(\frac{1}{x}) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0 \end{cases}$$

is not semismooth at $x = 0$. On $\mathbb{R}_{>0}$ and $\mathbb{R}_{<0}$, the function is continuously differentiable, and we have $\partial f(x) = \{2x \sin(\frac{1}{x}) - \cos(\frac{1}{x})\}$ for $x \neq 0$. Conversely, for $x = 0$ and $d = 1$ we obtain

$$\begin{aligned} &\operatorname{conv}\left\{ \lim_{t^{(k)} \searrow 0} Gd \;\middle|\; G \in \partial F(x + t^{(k)} d) \right\} \\ &= \operatorname{conv}\left\{ \lim_{t^{(k)} \searrow 0} G \;\middle|\; G \in \partial F(t^{(k)}) \right\} \\ &= \operatorname{conv}\left\{ \lim_{t^{(k)} \searrow 0} 2t^{(k)} \sin\Big(\frac{1}{t^{(k)}}\Big) - \cos\Big(\frac{1}{t^{(k)}}\Big) \right\} \\ &= [-1, 1]. \end{aligned}$$

This indicates that the requirement for a unique limit in Definition 2.1 is not satisfied.

To solve root-finding problems as (1.2), we can use a generalized Newton method [19]. At an iterate $x^{(k)}$, we use the update rule

$$x^{(k+1)} = x^{(k)} - (G^{(k)})^{-1} F(x^{(k)}),$$

where $G^{(k)}$ is a nonsingular element of the generalized Jacobian $\partial F(x^{(k)})$. This is called the semismooth Newton (SSN) method.

Algorithm 1 Semismooth Newton method

Input: $x^{(0)} \in \mathbb{R}^n$ and set $k := 0$.

1: **while** stopping criterion is not met **do**
2: Choose $G^{(k)} \in \partial F(x^{(k)})$
3: Determine the Newton update

$$G^{(k)} \delta x = -F(x^{(k)}).$$

4: Set $x^{(k+1)} = x^{(k)} + \delta x$.
5: Set $k := k + 1$.
6: **end while**

The algorithm is usually terminated by achieving a residual of sufficiently small norm. One can show that *locally*, the SSN method is well-defined and superlinearly convergent [19, Theorem 3.2].

2.3 Slanting Functions and Newton Differentiability

To extend the concept of semismooth functions to infinite-dimensional spaces, we review slanting functions [5] and Newton differentiability [10]. Throughout this section, we consider Banach spaces X and Y and an open subset $D \subseteq X$. Moreover, $\mathscr{L}(X, Y)$ denotes the space of bounded linear operators $X \to Y$.

Definition 2.4 ([5, Definition 2.1 and 2.3]). *A function $F\colon D \to Y$ is said to be* slantly differentiable *at $x \in D$ if*

there exists a mapping $G\colon D \to \mathscr{L}(X,Y)$ *and* $\varepsilon > 0$ *such that* $\{G(x+h) \mid \|h\| < \varepsilon\}$ *is bounded in the operator norm and*

$$\lim_{h\to 0} \frac{\|F(x+h) - F(x) - G(x+h)h\|}{\|h\|} = 0.$$

The function G *is called* slanting function *for* F *at* x. *Further, we call the set*

$$\partial_S F(x) := \{\lim G(x^{(k)}) \mid x^{(k)} \to x,\ x^{(k)} \in D\}$$

the slant derivative *of* F *associated with* G *at* $x \in D$.

Definition 2.5 ([5, Definition 2.2]). *A function* $F\colon D \to Y$ *is said to be* slantly differentiable *in an open domain* $U \subseteq D$ *if there exists a mapping* $G\colon D \to \mathscr{L}(X,Y)$, *such that* G *is a slanting function for* F *at every* $x \in U$. *In this case,* G *is called* slanting function *for* F *in* U.

This notion of differentiability now enables us to extend the concept of semismoothness to infinite-dimensional spaces.

Definition 2.6 ([5, Definition 3.2]). *A function* $F\colon D \to Y$ *is called* semismooth *at* $x \in D$ *if there is a slanting function* G *for* F *in a neighborhood* $\mathcal{N}(x)$ *of* x *such that* G *and the associated slant derivative satisfy the following two conditions:*

(*i*) $\lim\limits_{t\searrow 0} G(x+th)h$ *exists for every* $h \in X$ *and*

$$\lim_{\|h\|\to 0} \frac{\|\lim_{t\searrow 0} G(x+th)h - G(x+h)h\|}{\|h\|} = 0,$$

(*ii*) *For all* $V \in \partial_S F(x+h)$, *we have*

$$\|G(x+h)h - Vh\| = o(\|h\|).$$

This alternative definition of semismooth functions aligns with Definition 2.1 in finite-dimensional spaces [5, Theorem 3.3]. We also mention the related definitions of semismoothness in [23] and [24, Definition 3.1].

Let us comment on some of the important properties of slanting functions [5], which contributes to a better understanding of the concept.

(*i*) Due to the point-wise definition of slanting functions, a function F can be slantly differentiable at every point of D, but still there may be no common slanting function of F at all points of D [5, Remark (1)]. For instance, if F is Fréchet differentiable at x, we define $G(y) := F'(x)$ as a constant function for all $y \in D$. Hence, G constitutes a slanting function for F at the point x, although it generally does not fulfill this role at other points within D. Conversely, if F is continuously differentiable in D and we set $G(y) = F'(y)$ for all $y \in D$, then G satisfies Definition 2.5 and it is a slanting function for F in D.

(*ii*) Not all continuous functions are slantly differentiable [5, Remark (8)]. For instance, consider the real-valued function

$$f(x) = \begin{cases} \sqrt{x} & \text{if } x \geq 0, \\ -\sqrt{-x} & \text{if } x < 0. \end{cases}$$

Then as $h \searrow 0$ we have

$$\frac{f(h) - f(0) - G(h)h}{h} = \frac{\sqrt{h} - G(h)h}{h} = \frac{1}{\sqrt{h}} - G(h).$$

Since $1/\sqrt{h} \to \infty$ as $h \searrow 0$, it follows that there is no uniformly bounded function G such that the right hand side goes to zero. In this example, f is continuous but not slantly differentiable at zero.

(*iii*) A function $F\colon D \to Y$ is slantly differentiable at x if and only if there exists a neighborhood of x such that F is Lipschitz continuous in this neighborhood [5, Theorem 2.6].

An adaptation of Definition 2.5 was published in [10] three years after the introduction of slant differentiability. It involved a relaxation of the boundedness condition on $\{G(x+h)\}$ for slant differentiability, introducing a new concept known as Newton differentiability.

Definition 2.7 ([10, Definition 1.1]). *A mapping* $F\colon D \to Y$ *is called* Newton-differentiable *or* generalized differentiable *on the open subset* $U \subseteq D$ *if there exists a family of generalized derivatives* $G\colon U \to \mathscr{L}(X,Y)$, *such that*

$$\lim_{h\to 0} \frac{\|F(x+h) - F(x) - G(x+h)h\|}{\|h\|} = 0$$

holds for every $x \in U$.

In general, G is not unique and the definition of Newton differentiability is not restrictive regarding the type of generalized derivative that is being used. By choosing a single-valued selection of Clarke's Jacobian as generalized derivative G, (2.1) implies that every semismooth function is also Newton differentiable. Therefore, Newton differentiability extends the notion of differentiability even further than semismooth functions: F is not required to be locally Lipschitz continuous, the generalized derivative is not restricted to be Clarke's generalized Jacobian and the existence of a directional derivative is not necessary. Furthermore, if F is semismooth, every single-valued selection of Clarke's Jacobian serves as the generalized derivative G in Definition 2.7.

Example 2.8. *In Example 2.2 we demonstrated that the maximum function* $x \mapsto \max(0,x)$ *for* $x \in \mathbb{R}$ *is semismooth and that a possible representative of the set-valued*

generalized subdifferential is

$$G(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x \le 0. \end{cases}$$

This immediately implies that the maximum function is Newton differentiable. Because $G(x) \in \{0,1\}$, the operator family $\{G(x)\}$ in Definition 2.4 is uniformly bounded. Thus, G is also a slanting function for the maximum function for all $x \in \mathbb{R}$.

Example 2.9. *In Example 2.3 we argue that*

$$f : \mathbb{R} \to \mathbb{R}, \quad f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0 \end{cases}$$

is not semismooth at zero. But, f is slantly differentiable at zero with any slanting function G that satisfies $\lim_{h\to 0} G(h) = 0$ ([5, Remark (6)]). Then we have

$$\lim_{h\to 0} \frac{h^2 \sin\left(\frac{1}{h}\right) - G(h)h}{h} = \lim_{h\to 0} \left[h \sin\left(\frac{1}{h}\right) - G(h) \right] = 0.$$

This example illustrates that both slant differentiability and Newton differentiability are indeed extensions of semismoothness.

In the academic literature, the term "semismooth" is used inconsistently. Various concepts, including slant differentiability and Newton differentiability, are sometimes referred to as "semismoothness". Although the differences might be minor, it is essential to evaluate their significance for the specific application. Importantly, even though Newton differentiability is a weaker condition compared to slant differentiability or semismoothness, a successful local convergence analysis of the Newton method remains feasible.

Theorem 2.10 (Local superlinear convergence; [9, Theorem 1.1])**.** *Suppose that x^* is a solution to $F(x) = 0$ and that F is Newton differentiable in an open neighborhood $\mathcal{N}(x)$ containing x^* with slanting function $G(x)$. If $G(x)$ is nonsingular for all $x \in \mathcal{N}(x)$ and $\{\|G(x)^{-1}\| \mid x \in \mathcal{N}(x)\}$ is bounded, then the Newton iteration*

$$x^{(k+1)} = x^{(k)} - G(x^{(k)})^{-1} F(x^{(k)}),$$

converges superlinearly to x^*, provided that $\|x^{(0)} - x^*\|$ is sufficiently small.

Note that in the original formulation of this theorem in [9, Theorem 1.1], the authors use the term "slantly differentiable" instead of "Newton differentiable", while referring to the concept that we denote as Newton differentiability.

Proof. The proof follows the arguments of the proof in [9, Theorem 1.1] and we have adapted the phrasing used there. Let $B(x^*, r) \subseteq \mathcal{N}(x)$ denote a ball of radius r centered at x^*. Since $\|G(x)^{-1}\|$ is bounded in U, we find a positive constant M such that $\|G(x)^{-1}\| \le M$ in $B(x^*, r)$. Let $\eta \in (0,1]$ be arbitrary. Using Definition 2.7 of Newton differentiability, there exists $\rho \in (0, r)$ such that

$$\|F(x^* + h) - F(x^*) - G(x^* + h)h\| < \frac{\eta}{M}\|h\| \tag{2.3}$$

holds for all $\|h\| < \rho$. Assuming $\|x^{(k)} - x^*\| < \rho$ and using the above equation with $h = x^{(k)} - x^*$, the Newton iterates satisfy

$$\begin{aligned} &\|x^{(k+1)} - x^*\| \\ &\quad = \|x^{(k)} - G(x^{(k)})^{-1} F(x^{(k)}) - x^*\| \\ &\quad = \|-G(x^{(k)})^{-1}\left[F(x^{(k)}) - G(x^{(k)})(x^{(k)} - x^*)\right]\| \\ &\quad \le \|G(x^{(k)})^{-1}\| \, \|F(x^{(k)}) - G(x^{(k)})(x^{(k)} - x^*)\| \\ &\quad \le \underbrace{\|G(x^{(k)})^{-1}\|}_{\le M} \|F(x^{(k)}) - \underbrace{F(x^*)}_{=0} - G(x^{(k)})(x^{(k)} - x^*)\| \\ &\quad \le M \frac{\eta}{M} \|x^{(k)} - x^*\| = \eta \|x^{(k)} - x^*\|. \end{aligned}$$

Consequently, if we choose $x^{(0)}$ such that $\|x^{(0)} - x^*\| < \rho$, all iterates are well-defined and satisfy $\|x^{(k)} - x^*\| < \rho$. Since η is arbitrarily chosen, $x^{(k)} \to x^*$ converges superlinearly. □

Note that with the definition of Newton differentiability, the proof of superlinear convergence is short and concise. The reason for this is that the definition presupposes the upper bound, which is a crucial and fundamental aspect of the proof.

In conclusion, it can be stated that the various differentiability concepts presented in this section are closely connected. We provide a timeline in Fig. A.1 of the mentioned concepts and corresponding papers. The original definition of semismooth functions in finite-dimensional spaces [16], [19] relies on Clarke's generalized derivatives [6]. Chen, Nashed, Qi extended this concept to spaces of infinite dimensions by introducing slanting functions [5]. Clarke's concept of generalized derivatives can serve as slanting functions in finite-dimensional spaces.

Newton differentiability is a minor modification of slant differentiability, which eliminates specific restrictions on the slanting function, thus providing a more general framework. Importantly, we argue that, by selecting Clarke generalized derivatives as the generalized derivative in the Newton differentiability framework, it can be demonstrated that all semismooth functions are Newton differentiable.

3 Semismooth Newton Method

The generalized or semismooth Newton method is applicable to semismooth and Newton differentiable problems, such as (1.2). In the following we explicitly compute the update steps of the semismooth Newton method for the root-finding problem (1.2). Consistent with the generalized subdifferential of the maximum function (2.2), we define active and inactive index sets

$$\begin{aligned}\mathscr{A} &= \{1 \le i \le n \,|\, \mu_i + c\,(x_i - u_i) > 0\},\\ \mathscr{I} &= \{1 \le i \le n \,|\, \mu_i + c\,(x_i - u_i) \le 0\},\end{aligned}$$

for (1.2). On these active and inactive sets, (1.2) simplifies to

$$\begin{aligned}&Ax + \mu - b = 0\\ &\begin{cases}\mu_i = 0 & \text{if } i \in \mathscr{I}\\ c\,(x_i - u_i) = 0 & \text{if } i \in \mathscr{A}.\end{cases}\end{aligned}$$

Therefore, calculating the generalized subdifferential of $F(x,\mu)$ based on the active and inactive sets, we obtain the Newton update $(\delta x, \delta\mu)$ solving the system

$$\begin{aligned}&\begin{pmatrix} A_{\mathscr{I}\mathscr{I}} & A_{\mathscr{I}\mathscr{A}} & \mathrm{id}_{\mathscr{I}\mathscr{I}} & 0\\ A_{\mathscr{A}\mathscr{I}} & A_{\mathscr{A}\mathscr{A}} & 0 & \mathrm{id}_{\mathscr{A}\mathscr{A}}\\ 0 & 0 & \mathrm{id}_{\mathscr{I}\mathscr{I}} & 0\\ 0 & -c\,\mathrm{id}_{\mathscr{A}\mathscr{A}} & 0 & 0\end{pmatrix}\begin{pmatrix}\delta x_{\mathscr{I}}\\ \delta x_{\mathscr{A}}\\ \delta\mu_{\mathscr{I}}\\ \delta\mu_{\mathscr{A}}\end{pmatrix}\\ &\quad = -\begin{pmatrix}(Ax+\mu-b)_{\mathscr{I}}\\ (Ax+\mu-b)_{\mathscr{A}}\\ \mu_{\mathscr{I}}\\ c\,(u-x)_{\mathscr{A}}\end{pmatrix}.\end{aligned}$$

Here, we rearrange x, μ, A, b, u according to the disjoint index sets $\mathscr{A}$ and $\mathscr{I}$. For the matrix A this leads to the block matrix representation

$$\begin{pmatrix} A_{\mathscr{I}\mathscr{I}} & A_{\mathscr{I}\mathscr{A}}\\ A_{\mathscr{A}\mathscr{I}} & A_{\mathscr{A}\mathscr{A}}\end{pmatrix}.$$

Note that, for instance, we use the notation $A_{\mathscr{I}\mathscr{A}}$ to denote the submatrix of A formed by selecting the rows from $\mathscr{I}$ and columns from $\mathscr{A}$. The Newton update is used to calculate the next iterate $(x+\delta x, \mu+\delta\mu)$. This update can be equivalently interpreted as the following primal-dual active set algorithm (PDASA), compare [9, Chapter 2] for a derivation of the update step.

Algorithm 2 PDASA for upper bound constraints

Input: $x^{(0)}, \mu^{(0)}$

1: Set $k := 0$
2: **while** stopping criterion is not met **do**
3: Set

$$\begin{aligned}\mathscr{A}^{(k)} &= \{1 \le i \le n \,|\, \mu_i^{(k)} + c\,(x_i^{(k)} - u_i) > 0\}\\ \mathscr{I}^{(k)} &= \{1 \le i \le n \,|\, \mu_i^{(k)} + c\,(x_i^{(k)} - u_i) \le 0\}\end{aligned}$$

4: Determine $x^{(k+1)}, \mu^{(k+1)}$ from

$$\begin{aligned}&Ax^{(k+1)} + \mu^{(k+1)} = b,\\ &\mu^{(k+1)} = 0 \text{ on } \mathscr{I}^{(k)}, \quad x^{(k+1)} = u \text{ on } \mathscr{A}^{(k)}.\end{aligned}$$

5: Set $k := k+1$.
6: **end while**

Note that the iterations of this algorithm do not depend on $c > 0$ for $k \ge 1$ [15], since, for every index $i \in \{1,\dots,n\}$, we have $\mu_i^{(k)} = 0$ or $x_i^{(k)} - u_i = 0$. Further, we can state some useful properties of the iterates of the algorithm [13, p. 194f.]:

$$\begin{aligned}&A_{\mathscr{A}^{(k)}\mathscr{A}^{(k)}}\delta x_{\mathscr{A}^{(k)}} + A_{\mathscr{A}^{(k)}\mathscr{I}^{(k)}}\delta x_{\mathscr{I}^{(k)}} + \delta\mu_{\mathscr{A}^{(k)}} = 0,\\ &A_{\mathscr{I}^{(k)}\mathscr{I}^{(k)}}\delta x_{\mathscr{I}^{(k)}} + A_{\mathscr{I}^{(k)}\mathscr{A}^{(k)}}\delta x_{\mathscr{A}^{(k)}} + \delta\mu_{\mathscr{I}^{(k)}} = 0,\end{aligned} \tag{3.1a}$$

$$\begin{aligned}&x^{(k)} - u \ge 0 \quad \text{and} \quad \mu^{(k)} \ge 0 \quad \text{on } \mathscr{A}^{(k)},\\ &x^{(k)} - u \le 0 \quad \text{and} \quad \mu^{(k)} \le 0 \quad \text{on } \mathscr{I}^{(k)},\end{aligned} \tag{3.1b}$$

$$\begin{aligned}&\delta x = u - x^{(k)} \le 0 \quad \text{on } \mathscr{A}^{(k)},\\ &\delta\mu = -\mu^{(k)} \quad \ge 0 \quad \text{on } \mathscr{I}^{(k)}.\end{aligned} \tag{3.1c}$$

Further, if $\mathscr{A}^{(k)} = \mathscr{A}^{(k+1)}$, then we have found the solution $(x^{(k+1)}, \mu^{(k+1)}) = (x^*, \mu^*)$ for the root-finding problem (1.2) [13, Remark 7.1.1]. In the following we review necessary and sufficient conditions for this to happen.

From the equivalence of the SSN method and the PDAS algorithm, it directly follows that the PDAS algorithm converges locally superlinearly to the solution of the root-finding problem (1.2) [9, Theorem 3.1]. The global convergence theory is not that straightforward.

Recall that $A \in \mathbb{R}^{n\times n}$ is called an *M-matrix* if A is nonsingular, $(a_{ij}) \le 0$ for $i \ne j$, and $A^{-1} \ge 0$ entry-wise. This is the case, for instance, when (QP) arises from standard discretizations of an obstacle problem [13, Chapter 4.7.4].

For M-matrices, Algorithm 2 exhibits global convergence in a particular manner. This is shown in the following theorem, which we state here for symmetric matrices A.

Theorem 3.1 ([13, Theorem 7.4])**.** *Suppose that A is a symmetric M-matrix. Then $x^{(k)} \to x^*$, $\mu^{(k)} \to \mu^*$ for arbitrary initial data $(x^{(0)}, \mu^{(0)})$. Moreover, $x^* \le x^{(k+1)} \le x^{(k)}$*

for all $k \geq 1$, $x^{(k)} \leq u$ for all $k \geq 2$, and there exists k_0 such that $\mu^{(k)} \geq 0$ for all $k \geq k_0$.

As shown in the following theorem, global convergence can also be shown for a broader class of matrices. Recall that $A \in \mathbb{R}^{n\times n}$ is called a *P-matrix* if all its principal minors are positive [13, p. 197]. Note that every M-matrix and every positive definite matrix is a P-matrix [4, p. 271].

Theorem 3.2 ([13, Theorem 7.5.]). *Suppose that A is a symmetric P-matrix of size $n \times n$. Moreover, suppose that for every disjoint union $\mathscr{A} \sqcup \mathscr{I} = \{1,\dots,n\}$ we have*

$$\left\| \left((A_{\mathscr{I}\mathscr{I}})^{-1} A_{\mathscr{I}\mathscr{A}}\right)_+ \right\|_1 < 1 \quad \text{and} \quad \sum_{i\in\mathscr{I}} \left[(A_{\mathscr{I}\mathscr{I}})^{-1} x_{\mathscr{I}}\right]_i > 0$$

for all $x_{\mathscr{I}} \geq 0$ with $x_{\mathscr{I}} \neq 0$, then $\lim_{k\to\infty} x^{(k)} = x^*$. Here, $(\cdot)_+$ denotes the positive part of every component in a matrix. When $\mathscr{I} = \emptyset$, then both conditions are void. When $\mathscr{A} = \emptyset$, the first condition is void.

At first glance, the conditions of this theorem might not seem intuitive, and it is not immediately clear which matrices will satisfy them. However, Theorem 3.1 identifies a class of matrices that meet these conditions. Specifically, for any M-matrix A, we have $(A_{\mathscr{I}\mathscr{I}})^{-1} \geq 0$ and $(A_{\mathscr{I}\mathscr{I}})^{-1} A_{\mathscr{I}\mathscr{A}} \leq 0$ for every possible choice of $\mathscr{A}$ and $\mathscr{I}$ [4, p. 134]. These properties are already utilized in the proof of Theorem 3.1. Additionally, they imply that $\|((A_{\mathscr{I}\mathscr{I}})^{-1} A_{\mathscr{I}\mathscr{A}})_+\|_1 < 1$ and $\sum_{i\in\mathscr{I}} [(A_{\mathscr{I}\mathscr{I}})^{-1} x_{\mathscr{I}}]_i > 0$ for $x_{\mathscr{I}} \geq 0$ with $x_{\mathscr{I}} \neq 0$. As mentioned earlier, every M-matrix is also a P-matrix, allowing us to apply Theorem 3.2.

Hintermüller, Ito, Kunisch [9] have also demonstrated a perturbation result, which asserts the global convergence when A is a small perturbation of an M-matrix. This is of particular interest in numerical implementations, since the properties of M-matrices, such as the non-positivity of off-diagonal elements, are not stable under small perturbations.

Theorem 3.3 ([13, Theorem 7.6]). *Suppose that M is an M-matrix, K is arbitrary and $A = M + K$ is symmetric. If $\|K\|$ is sufficiently small, the root-finding problem (1.2) admits a unique solution (x^*, μ^*), Algorithm 2 is well-defined and $\lim_{k\to\infty} x^{(k)} = x^*$.*

A precise characterization of "sufficiently small" can be found in the proof of [13, Theorem 7.6]. In particular, the smallness of $\|K\|$ is determined by the parameters ρ and σ defined in that proof, which quantify the perturbation bounds ensuring the theorem's conclusions hold.

For completeness, we briefly mention that there are also convergence results for bilaterally constrained problems. That is, we add the constraints $x \geq \ell$ for some $-\infty < \ell < u$ to (QP). In this case, convergence results are more complex to express and significantly less intuitive to comprehend. For some results on the convergence of the PDASA for this case, we refer the reader to [12].

4 Cycle Analysis

In this section, we aim to further explore the global convergence behavior of the PDASA. In particular, we seek to go beyond the conditions of Theorems 3.1 to 3.3 and investigate the convergence behavior for problems that do *not* satisfy these conditions. To this end, we consider the following standing assumptions for $A \in \mathbb{R}^{n\times n}$ in (QP).

Assumption 4.1 (standing assumptions).

(*i*) *A is symmetric and positive definite. The latter is equivalent to all leading principal minors of A being positive.*

(*ii*) *A is not an M-matrix. Therefore, there either exist indices $i \neq j$ such that*

$$a_{ij} > 0 \quad \text{or} \quad A^{-1} \ngeq 0 \text{ entry-wise.}$$

(*iii*) *There exists a disjoint union $\mathscr{A} \sqcup \mathscr{I} = \{1,\dots,n\}$ such that*

$$\left\| \left((A_{\mathscr{I}\mathscr{I}})^{-1} A_{\mathscr{I}\mathscr{A}}\right)_+ \right\|_1 \geq 1 \tag{4.1}$$

with $\mathscr{I}$ and $\mathscr{A}$ nonempty, or

$$\sum_{i\in\mathscr{I}} \left[(A_{\mathscr{I}\mathscr{I}})^{-1} x_{\mathscr{I}}\right]_i \leq 0 \tag{4.2}$$

for some $x_{\mathscr{I}} \geq 0$ with $x_{\mathscr{I}} \neq 0$ and $\mathscr{I}$ nonempty.

(*iv*) *For any M-matrix M and $K := A - M$, the matrix K is not sufficiently small in the sense of Theorem 3.3.*

For matrices that satisfy these three conditions, the global convergence of PDASA for the unilaterally constrained (QP) remains undetermined, as none of the theorems from Section 3 is applicable.

This section is principally inspired by the author's implementation of the PDASA, accessible on GitHub [20]. We have implemented Algorithm 2, choosing $c = 1$, and have conducted multiple experiments with random input data for A, b, and u. Our objective has been to identify patterns in problem instances where convergence was not achieved. Since there exist only finitely many possible partitions of the index set $\{1,\dots,n\}$, the

algorithm will either exhibit convergence or encounter a cyclic pattern, disrupting convergence. Thus, the examination of the algorithm's convergence is equivalent to examining the potential cycling patterns that may occur for different problem instances. We consider only very low-dimensional situations, where we can fully characterize the behavior of the algorithm.

It is important to realize that each iteration of the algorithm is completely determined by the current active and inactive sets. The active indices of the primal variable are constrained to the upper bound, while the inactive indices of the dual variable are set to zero. The remaining variables, $x_{\mathscr{I}}$ and $\mu_{\mathscr{A}}$, are determined such that they satisfy the linear system $Ax+\mu = b$. To analyze the iterative process of the algorithm, it is sufficient to examine the active and inactive sets at each iteration.

For $n = 1$, convergence is always achieved within a maximum of two iterations. This directly follows from the fact that we have only two possible active sets ($\emptyset$, $\{1\}$) and that identical active/inactive sets in consecutive iterations indicate that a solution has been found.

For $n = 2$, we have observed global convergence of PDASA for all problem instances. A proof of this result is provided in Section 4.1.

For $n = 3$, we have noticed that the cycles disrupting convergence in numerical examples are precisely of two distinct types. We provide an analysis and mathematical validation of the observed characteristics in Section 4.3.

4.1 Two-Dimensional Case

In this section we consider the two-dimensional case, $n = 2$. We demonstrate in Theorem 4.3 below that in this scenario, the PDASA exhibits convergence for any given initial iterate and any set of problem data. The problem data satisfies

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \quad u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix},$$
$$A^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

In the two-dimensional case presently considered, we explicitly restate the conditions from Assumptions (i) and (ii), as these two are specifically required for subsequent steps:

(i) A is symmetric and positive definite if and only if

$$a_{12} = a_{21}, \quad a_{11} > 0, \quad a_{11}a_{22} - a_{12}^2 > 0.$$

(ii) A is not an M-matrix if and only if $a_{12} > 0$.

In the following analysis, we set $c = 1$ to simplify the calculations. As noted earlier, the choice of $c > 0$ does not affect the algorithm's iterates beyond the first iteration. Therefore, this analysis remains valid for the general case. As mentioned before, it suffices to know the previous active set to characterize the update step in the PDASA. We need to distinguish three cases.

1. $\mathscr{A}^{(k)} = \{1,2\}$: The subsequent iterate is given by $x^{(k+1)} = u$ and $\mu^{(k+1)} = b - Ax^{(k+1)} = b - Au$. The following active set will therefore be determined by evaluating the signs of

$$\mu^{(k+1)} + x^{(k+1)} - u = \begin{pmatrix} b_1 - a_{11}u_1 - a_{12}u_2 \\ b_2 - a_{22}u_2 - a_{12}u_1 \end{pmatrix}. \tag{4.3}$$

If, for instance, $b_1 - a_{11}u_1 - a_{12}u_2 > 0$, then $1 \in \mathscr{A}^{(k+1)}$, whereas $b_1 - a_{11}u_1 - a_{12}u_2 \le 0$ would imply $1 \in \mathscr{I}^{(k+1)}$.

2. $\mathscr{A}^{(k)} = \{i\}$ and $\mathscr{I}^{(k)} = \{j\}$: In this case, we have $x_i^{(k+1)} = u_i$ and $\mu_j^{(k+1)} = 0$. Solving the linear equation system leads to

$$x_j^{(k+1)} = \frac{b_j - a_{ij}u_i}{a_{jj}},$$
$$\mu_i^{(k+1)} = b_i - a_{ii}u_i - a_{ij}\frac{b_j - a_{ij}u_i}{a_{jj}},$$

and hence

$$\mu^{(k+1)} + x^{(k+1)} - u = \begin{pmatrix} b_i - a_{ii}u_i - a_{ij}\frac{b_j - a_{ij}u_i}{a_{jj}} \\ \frac{b_j - a_{ij}u_i}{a_{jj}} - u_j \end{pmatrix}. \tag{4.4}$$

3. $\mathscr{A}^{(k)} = \emptyset$: Finally, we have $\mu^{(k+1)} = 0$ and $x^{(k+1)} = A^{-1}b$ and hence

$$\mu^{(k+1)} + x^{(k+1)} - u = \begin{pmatrix} \frac{b_2a_{12} - b_1a_{22}}{a_{12}^2 - a_{11}a_{22}} - u_1 \\ \frac{b_1a_{12} - b_2a_{11}}{a_{12}^2 - a_{11}a_{22}} - u_2 \end{pmatrix}. \tag{4.5}$$

With the three equations (4.3), (4.4) and (4.5), we can fully describe the behavior of PDASA in problem dimension $n = 2$.

Example 4.2. *The purpose of this example is to present a case where PDASA converges globally but that is not covered by the known convergence results Theorems 3.1 to 3.3. We set*

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}, \quad b = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \quad u = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

as problem data and observe the active set behavior illustrated in Fig. 4.1. In this and all other figures, each circle represents one particular active set. The arrows indicate transitions to the subsequent active set. The circle colored in green represents the active set identifying

the unique solution of the problem. Once the algorithm reaches this active set, it remains there and can be terminated. The first active set is determined by the initial iterate $(x^{(0)}, \mu^{(0)})$.

In this example, A satisfies Assumption 4.1. In particular, A is symmetric and positive definite. Further, it is not an M-matrix and Theorems 3.1 to 3.3 are not applicable. Still, the PDASA converges for any initial guess. In Theorem 4.3 we show that this is not a coincidence in case $n = 2$.

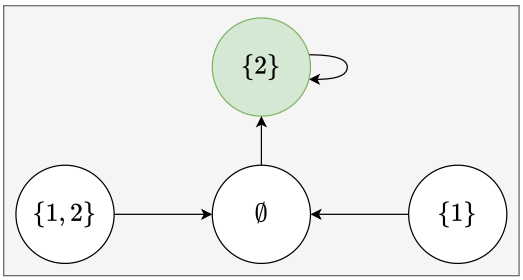

Figure 4.1 – Active set behavior for Example 4.2.

Theorem 4.3. *Suppose that* $A \in \mathbb{R}^{2\times 2}$ *is symmetric and positive definite,* $b \in \mathbb{R}^2$, *and* $u \in \mathbb{R}^2$. *Then PDASA converges for any initial guess* $x^{(0)}$.

To prove this theorem, the following lemma exhibits three rules that necessarily hold for all possible sequences of active and inactive sets.

Lemma 4.4. *For* $A \in \mathbb{R}^{2\times 2}$ *satisfying Assumption 4.1, indices* $i, j \in \{1,2\}$, $i \neq j$, *and iteration indices* $k, k' \in \mathbb{N}$, $k \neq k'$, *the following holds:*

(*i*) *if* $\mathscr{A}^{(k)} = \{1,2\}$ *and* $i \notin \mathscr{A}^{(k+1)}$, *then, for* $\mathscr{A}^{(k')} = \{j\}$, *it follows that* $i \notin \mathscr{A}^{(k'+1)}$.

(*ii*) *if* $\mathscr{A}^{(k)} = \{1,2\}$ *and* $\mathscr{A}^{(k+1)} = \{i\}$, *then, for* $\mathscr{A}^{(k')} = \{i\}$, *it follows that* $i \in \mathscr{A}^{(k'+1)}$.

(*iii*) *if* $\mathscr{A}^{(k)} = \emptyset$ *and* $i \in \mathscr{A}^{(k+1)}$, *then, for* $\mathscr{A}^{(k')} = \{i\}$, *it follows that* $i \in \mathscr{A}^{(k'+1)}$.

Proof. For the first rule, we observe (using (4.3)) that the conditions $\mathscr{A}^{(k)} = \{1,2\}$ and $i \notin \mathscr{A}^{(k+1)}$ are equivalent to

$$b_i - a_{ii}u_i - a_{ij}u_j \leq 0.$$

Furthermore we characterize $\mathscr{A}^{(k')} = \{j\}$ and $i \in \mathscr{A}^{(k'+1)}$ based on (4.4) with

$$\frac{b_i - a_{ij}u_j}{a_{ii}} - u_i > 0.$$

Since A is positive definite, we have $a_{ii} > 0$. Consequently, this results in a contradiction and $i \notin \mathscr{A}^{(k'+1)}$.

For the second rule, having $\mathscr{A}^{(k)} = \{1,2\}$ and $\mathscr{A}^{(k+1)} = \{i\}$ can (using (4.3)) be equivalently expressed as

$$b_i - a_{ii}u_i - a_{ij}u_j > 0, \tag{4.6a}$$
$$b_j - a_{jj}u_j - a_{ij}u_i \leq 0. \tag{4.6b}$$

Additionally, for $\mathscr{A}^{(k')} = \{i\}$ and $i \notin \mathscr{A}^{(k'+1)}$ we write (using (4.4))

$$b_i - a_{ii}u_i - a_{ij}\frac{b_j - a_{ij}u_i}{a_{jj}} \leq 0.$$

Combining this with (4.6a), we derive

$$\begin{aligned} 0 &> a_{ij}u_j - a_{ij}\frac{b_j - a_{ij}u_i}{a_{jj}} \\ &= a_{ij}\frac{a_{jj}u_j + a_{ij}u_i - b_j}{a_{jj}}, \end{aligned}$$

which contradicts (4.6b), because according to Assumption 4.1 we have $a_{ij} > 0$ and $a_{jj} > 0$. This establishes the second rule.

For the third rule, we characterize $\mathscr{A}^{(k)} = \emptyset$ and $i \in \mathscr{A}^{(k+1)}$ (using (4.5)) by

$$\frac{b_j a_{ij} - b_i a_{jj}}{a_{ij}^2 - a_{ii}a_{jj}} - u_i > 0,$$

and it follows that $b_j a_{ij} + u_i(a_{ii}a_{jj} - a_{ij}^2) < b_i a_{jj}$. This is because A satisfies Assumption 4.1, thus, $\det(A) = a_{ii}a_{jj} - a_{ij}^2 > 0$. Moreover, $\mathscr{A}^{(k')} = \{i\}$ and $i \notin \mathscr{A}^{(k'+1)}$ is equivalent to (using (4.4))

$$b_i - a_{ii}u_i - a_{ij}\frac{b_j - a_{ij}u_i}{a_{jj}} \leq 0.$$

Since $a_{jj} > 0$ by Assumption 4.1, this is equivalent to $b_j a_{ij} + u_i(a_{ii}a_{jj} - a_{ij}^2) \geq b_i a_{jj}$. This is a contradiction, hence the third rule follows. □

Next, we rule out two specific active set cycles by means of a contradiction that satisfy the necessary conditions of Lemma 4.4. This narrows down the number of possible active set cycles further.

Lemma 4.5. *Suppose that* $A \in \mathbb{R}^{2\times 2}$ *satisfies Assumption 4.1. The following cycles cannot occur in the PDASA.*

(*i*) $\{i\} \to \{j\} \to \{i\}$,

(*ii*) $\emptyset \to \{1,2\} \to \emptyset$.

Here i *and* j *denote indices in* $\{1,2\}$ *with* $i \neq j$.

Proof. We characterize the cycle $\{i\} \to \{j\} \to \{i\}$ using (4.4), resulting in the inequalities

$$b_i - a_{ii}u_i - a_{ij}\frac{b_j - a_{ij}u_i}{a_{jj}} \le 0, \tag{4.7a}$$

$$\frac{b_j - a_{ij}u_i}{a_{jj}} - u_j > 0, \tag{4.7b}$$

$$b_j - a_{jj}u_j - a_{ij}\frac{b_i - a_{ij}u_j}{a_{ii}} \le 0, \tag{4.7c}$$

$$\frac{b_i - a_{ij}u_j}{a_{ii}} - u_i > 0. \tag{4.7d}$$

The first inequality (4.7a) is equivalent to

$$a_{ii}u_i + \frac{a_{ij}}{a_{jj}}b_j - \frac{a_{ij}^2}{a_{jj}}u_i \ge b_i$$

and (4.7c) is equivalent to (using $a_{ii} > 0$)

$$\frac{a_{ii}b_j + (a_{ij}^2 - a_{jj}a_{ii})u_j}{a_{ij}} \le b_i.$$

Together this implies

$$\frac{a_{ii}a_{jj} - a_{ij}^2}{a_{ij}a_{jj}}b_j \le \frac{a_{ii}a_{jj} - a_{ij}^2}{a_{jj}}u_i + \frac{a_{ii}a_{jj} - a_{ij}^2}{a_{ij}}u_j.$$

Combining this with (4.7b), this leads to the contradiction

$$\begin{aligned} a_{ij}u_i + a_{jj}u_j < b_j &\le \frac{a_{ij}a_{jj}}{a_{jj}}u_i + \frac{a_{ij}a_{jj}}{a_{ij}}u_j \\ &= a_{ij}u_i + a_{jj}u_j. \end{aligned}$$

The cycle $\emptyset \to \{1,2\} \to \emptyset$ can be described using (4.5) and (4.3), resulting in the inequalities

$$\frac{b_2a_{12} - b_1a_{22}}{a_{12}^2 - a_{11}a_{22}} - u_1 > 0, \tag{4.8a}$$

$$\frac{b_1a_{12} - b_2a_{11}}{a_{12}^2 - a_{11}a_{22}} - u_2 > 0, \tag{4.8b}$$

$$b_1 - a_{11}u_1 - a_{12}u_2 \le 0, \tag{4.8c}$$

$$b_2 - a_{22}u_2 - a_{12}u_1 \le 0. \tag{4.8d}$$

The first inequality (4.8a) is equivalent to

$$\frac{a_{22}}{a_{12}}b_1 - \frac{a_{11}a_{22} - a_{12}^2}{a_{12}}u_1 > b_2$$

and (4.8b) is equivalent to

$$\frac{a_{12}}{a_{11}}b_1 + \frac{a_{11}a_{22} - a_{12}^2}{a_{11}}u_2 < b_2.$$

Together this implies

$$\frac{a_{11}a_{22} - a_{12}^2}{a_{12}a_{11}}b_1 > \frac{a_{11}a_{22} - a_{12}^2}{a_{12}}u_1 + \frac{a_{11}a_{22} - a_{12}^2}{a_{11}}u_2.$$

Combining this with (4.8c) leads to the contradiction

$$\begin{aligned} a_{11}u_1 + a_{12}u_2 \ge b_1 &> \frac{a_{12}a_{11}}{a_{12}}u_1 + \frac{a_{12}a_{11}}{a_{11}}u_2 \\ &= a_{11}u_1 + a_{12}u_2. \end{aligned}$$

□

The proof of Theorem 4.3 is now a direct consequence of the previous lemmas. To see this, we list all combinatorially possible cycles. Letting $i, j \in \{1,2\}$ and $i \ne j$, the following is a complete list of cycles:

$\emptyset \to \{i\} \to \emptyset$	*(iii)*
$\emptyset \to \{i\} \to \{j\} \to \emptyset$	*(iii)*
$\emptyset \to \{i\} \to \{j\} \to \{1,2\} \to \emptyset$	*(i)*
$\emptyset \to \{i\} \to \{1,2\} \to \emptyset$	*(i)*
$\emptyset \to \{i\} \to \{1,2\} \to \{j\} \to \emptyset$	*(ii)*
$\emptyset \to \{1,2\} \to \emptyset$	
$\emptyset \to \{1,2\} \to \{i\} \to \emptyset$	*(ii)*
$\emptyset \to \{1,2\} \to \{i\} \to \{j\} \to \emptyset$	*(i)*
$\{i\} \to \{j\} \to \{i\}$	
$\{i\} \to \{j\} \to \{1,2\} \to \{i\}$	*(i)*
$\{i\} \to \{1,2\} \to \{i\}$	*(i)*

The number in parentheses behind each cycle indicates which of the three rules from Lemma 4.4 shows that this cycle is actually impossible. The remaining two cycles are ruled out by Lemma 4.5. In summary, we have systematically excluded all potential cyclic behaviors under the condition that A satisfies Assumption 4.1, thereby ensuring convergence of the PDASA to the optimal solution also in cases not covered by the known convergence Theorems 3.1 and 3.2.

Note that in the proofs of Lemma 4.4 and Lemma 4.5, we heavily used the positive definiteness of A. In case that A is negative definite, the update step of the PDASA remains well-defined, as the linear equation in the update step can be uniquely solved. Consider the following example:

$$A = \begin{pmatrix} -2 & 0 \\ 0 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad u = \begin{pmatrix} -2 \\ -2 \end{pmatrix},$$

$$x^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \mu^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This setup presents a strictly concave problem and it is thus an unbounded minimization problem. The PDASA generates the cyclic active set sequence

$$\{1,2\} \to \emptyset \to \{1,2\},$$

which was ruled out by Lemma 4.5 for positive definite matrices.

4.2 Abstract Cycle Analysis

In [1, Lemma 4.4], Ben Gharbia, Gilbert established a necessary condition for the presence of a cycle of length 3 in the PDASA. Here we extend their result to cycles of arbitrary length.

Theorem 4.6 (Necessary condition for m-cycle)**.** *Suppose that $A \in \mathbb{R}^{n\times n}$ is symmetric, positive definite and that the PDASA produces a cycle by visiting $m \geq 2$ pairwise distinct points $x^{(1)} \to \ldots \to x^{(m)}$ and $x^{(m+1)} = x^{(1)}$. Then, for each $k \in \{1,\ldots,m\}$ the set*

$$\left(\mathscr{A}^{(k-1)} \cap \mathscr{I}^{(k)} \cap \mathscr{I}^{(k+1)}\right) \cup \left(\mathscr{I}^{(k-1)} \cap \mathscr{A}^{(k)} \cap \mathscr{A}^{(k+1)}\right),$$

is nonempty. We recall that $\mathscr{A}^{(k)}$ denotes the active set at iterate $x^{(k)}$, with the conventions $\mathscr{A}^{(0)} := \mathscr{A}^{(m)}$ and $\mathscr{A}^{(m+1)} := \mathscr{A}^{(1)}$.

Proof. The proof follows the structure of the proof of [1, Lemma 4.4], replacing the cycle length 3 with an arbitrary cycle length m.

We observe that for this cycle the update rules of the algorithm imply

$$x^{(2)}_{\mathscr{A}^{(1)}} = u_{\mathscr{A}^{(1)}}, \quad \mu^{(2)}_{\mathscr{I}^{(1)}} = 0, \quad x^{(1)}_{\mathscr{A}^{(m)}} = u_{\mathscr{A}^{(m)}}, \quad \mu^{(1)}_{\mathscr{I}^{(m)}} = 0.$$

Now we calculate

$$\begin{aligned} x^{(2)} - x^{(1)} &= \begin{pmatrix} (x^{(2)} - x^{(1)})_{\mathscr{A}^{(1)}\cap\mathscr{A}^{(2)}\cap\mathscr{A}^{(m)}} \\ (x^{(2)} - x^{(1)})_{\mathscr{A}^{(1)}\cap\mathscr{A}^{(2)}\cap\mathscr{I}^{(m)}} \\ (x^{(2)} - x^{(1)})_{\mathscr{A}^{(1)}\cap\mathscr{I}^{(2)}\cap\mathscr{A}^{(m)}} \\ (x^{(2)} - x^{(1)})_{\mathscr{A}^{(1)}\cap\mathscr{I}^{(2)}\cap\mathscr{I}^{(m)}} \\ (x^{(2)} - x^{(1)})_{\mathscr{I}^{(1)}\cap\mathscr{A}^{(2)}\cap\mathscr{A}^{(m)}} \\ (x^{(2)} - x^{(1)})_{\mathscr{I}^{(1)}\cap\mathscr{A}^{(2)}\cap\mathscr{I}^{(m)}} \\ (x^{(2)} - x^{(1)})_{\mathscr{I}^{(1)}\cap\mathscr{I}^{(2)}\cap\mathscr{A}^{(m)}} \\ (x^{(2)} - x^{(1)})_{\mathscr{I}^{(1)}\cap\mathscr{I}^{(2)}\cap\mathscr{I}^{(m)}} \end{pmatrix} \\ &= \begin{pmatrix} (u-u)_{\mathscr{A}^{(1)}\cap\mathscr{A}^{(2)}\cap\mathscr{A}^{(m)}} \\ (u-x^{(1)})_{\mathscr{A}^{(1)}\cap\mathscr{A}^{(2)}\cap\mathscr{I}^{(m)}} \\ (u-u)_{\mathscr{A}^{(1)}\cap\mathscr{I}^{(2)}\cap\mathscr{A}^{(m)}} \\ (u-x^{(1)})_{\mathscr{A}^{(1)}\cap\mathscr{I}^{(2)}\cap\mathscr{I}^{(m)}} \\ (x^{(2)}-u)_{\mathscr{I}^{(1)}\cap\mathscr{A}^{(2)}\cap\mathscr{A}^{(m)}} \\ (x^{(2)} - x^{(1)})_{\mathscr{I}^{(1)}\cap\mathscr{A}^{(2)}\cap\mathscr{I}^{(m)}} \\ (x^{(2)}-u)_{\mathscr{I}^{(1)}\cap\mathscr{I}^{(2)}\cap\mathscr{A}^{(m)}} \\ (x^{(2)} - x^{(1)})_{\mathscr{I}^{(1)}\cap\mathscr{I}^{(2)}\cap\mathscr{I}^{(m)}} \end{pmatrix} \begin{matrix} = 0 \\ \leq 0 \\ = 0 \\ \leq 0 \\ \geq 0 \\ \\ \leq 0 \\ \\ \end{matrix} \end{aligned}$$

where the sign of the entries follows from (3.1b). Further, we calculate

$$\begin{aligned} A(x^{(2)} - x^{(1)}) &= \begin{pmatrix} (Ax^{(2)} - Ax^{(1)})_{\mathscr{A}^{(1)}\cap\mathscr{A}^{(2)}\cap\mathscr{A}^{(m)}} \\ (Ax^{(2)} - Ax^{(1)})_{\mathscr{A}^{(1)}\cap\mathscr{A}^{(2)}\cap\mathscr{I}^{(m)}} \\ (Ax^{(2)} - Ax^{(1)})_{\mathscr{A}^{(1)}\cap\mathscr{I}^{(2)}\cap\mathscr{A}^{(m)}} \\ (Ax^{(2)} - Ax^{(1)})_{\mathscr{A}^{(1)}\cap\mathscr{I}^{(2)}\cap\mathscr{I}^{(m)}} \\ (Ax^{(2)} - Ax^{(1)})_{\mathscr{I}^{(1)}\cap\mathscr{A}^{(2)}\cap\mathscr{A}^{(m)}} \\ (Ax^{(2)} - Ax^{(1)})_{\mathscr{I}^{(1)}\cap\mathscr{A}^{(2)}\cap\mathscr{I}^{(m)}} \\ (Ax^{(2)} - Ax^{(1)})_{\mathscr{I}^{(1)}\cap\mathscr{I}^{(2)}\cap\mathscr{A}^{(m)}} \\ (Ax^{(2)} - Ax^{(1)})_{\mathscr{I}^{(1)}\cap\mathscr{I}^{(2)}\cap\mathscr{I}^{(m)}} \end{pmatrix} \\ &= \begin{pmatrix} (Ax^{(2)} - Ax^{(1)})_{\mathscr{A}^{(1)}\cap\mathscr{A}^{(2)}\cap\mathscr{A}^{(m)}} \\ (Ax^{(2)} - b)_{\mathscr{A}^{(1)}\cap\mathscr{A}^{(2)}\cap\mathscr{I}^{(m)}} \\ (Ax^{(2)} - Ax^{(1)})_{\mathscr{A}^{(1)}\cap\mathscr{I}^{(2)}\cap\mathscr{A}^{(m)}} \\ (Ax^{(2)} - b)_{\mathscr{A}^{(1)}\cap\mathscr{I}^{(2)}\cap\mathscr{I}^{(m)}} \\ (b - Ax^{(1)})_{\mathscr{I}^{(1)}\cap\mathscr{A}^{(2)}\cap\mathscr{A}^{(m)}} \\ (b-b)_{\mathscr{I}^{(1)}\cap\mathscr{A}^{(2)}\cap\mathscr{I}^{(m)}} \\ (b - Ax^{(1)})_{\mathscr{I}^{(1)}\cap\mathscr{I}^{(2)}\cap\mathscr{A}^{(m)}} \\ (b-b)_{\mathscr{I}^{(1)}\cap\mathscr{I}^{(2)}\cap\mathscr{I}^{(m)}} \end{pmatrix} \\ &= \begin{pmatrix} (Ax^{(2)} - Ax^{(1)})_{\mathscr{A}^{(1)}\cap\mathscr{A}^{(2)}\cap\mathscr{A}^{(m)}} \\ -\mu^{(2)}_{\mathscr{A}^{(1)}\cap\mathscr{A}^{(2)}\cap\mathscr{I}^{(m)}} \\ (Ax^{(2)} - Ax^{(1)})_{\mathscr{A}^{(1)}\cap\mathscr{I}^{(2)}\cap\mathscr{A}^{(m)}} \\ -\mu^{(2)}_{\mathscr{A}^{(1)}\cap\mathscr{I}^{(2)}\cap\mathscr{I}^{(m)}} \\ \mu^{(1)}_{\mathscr{I}^{(1)}\cap\mathscr{A}^{(2)}\cap\mathscr{A}^{(m)}} \\ (b-b)_{\mathscr{I}^{(1)}\cap\mathscr{A}^{(2)}\cap\mathscr{I}^{(m)}} \\ \mu^{(1)}_{\mathscr{I}^{(1)}\cap\mathscr{I}^{(2)}\cap\mathscr{A}^{(m)}} \\ (b-b)_{\mathscr{I}^{(1)}\cap\mathscr{I}^{(2)}\cap\mathscr{I}^{(m)}} \end{pmatrix} \begin{matrix} \leq 0 \\ \\ \geq 0 \\ \leq 0 \\ = 0 \\ \leq 0 \\ = 0 \end{matrix} \end{aligned}$$

The positive definiteness of A implies

$$(x^{(2)} - x^{(1)})^{\mathrm{T}} A (x^{(2)} - x^{(1)}) > 0.$$

By contrast, $\mathscr{I}^{(1)} \cap \mathscr{I}^{(2)} \cap \mathscr{A}^{(m)} \cup \mathscr{A}^{(1)} \cap \mathscr{A}^{(2)} \cap \mathscr{I}^{(m)}$ being empty implies $(x^{(2)} - x^{(1)})^{\mathrm{T}} A (x^{(2)} - x^{(1)}) \leq 0$ and we would have $x^{(2)} = x^{(1)}$, which contradicts the assumption of cycle length $m \geq 2$. The same is observed when starting with $x^{(2)}$ and $x^{(3)}$, $x^{(3)}$ and $x^{(4)}$, and so on. Thus, by cyclically permuting the indices, we obtain that the set

$$\left(\mathscr{A}^{(k-1)} \cap \mathscr{I}^{(k)} \cap \mathscr{I}^{(k+1)}\right) \cup \left(\mathscr{I}^{(k-1)} \cap \mathscr{A}^{(k)} \cap \mathscr{A}^{(k+1)}\right),$$

is nonempty for $k \in \{1,2,\ldots,m\}$. □

By considering this necessary condition for 2-cycles in more detail, we can show:

Corollary 4.7 ([1, Lemma 4.3])**.** *Suppose that A is symmetric and positive definite, then the PDASA does not form 2-cycles.*

Proof. By similar calculations as in the proof of Theorem 4.6, we can show that

$$(x^{(2)} - x^{(1)})^{\mathrm{T}} A (x^{(2)} - x^{(1)}) \leq 0.$$

Since A is positive definite, this implies $x^{(2)} = x^{(1)}$, which is a contradiction. □

The results from Theorem 4.6 and Corollary 4.7 present an alternative proof of Theorem 4.3 [1, Proposition 4.5], i. e., global convergence of the PDASA for two-dimensional problems.

4.3 Three-Dimensional Case

In this section, we exploit the necessary condition for m-cycles as stated in Theorem 4.6 to investigate the possible cycle behaviors in case $n = 3$.

For a comprehensive analysis of all possible cycles, it is necessary to examine cycles of length $m \leq 7$, as there are $2^n = 8$ possible active sets, one of which is optimal. Notice that any cycle that can occur must have length $m > 2$ as established in Corollary 4.7. However, in numerical examples we have typically observed that $m \leq n$, which gives us reason to believe that cycles of greater length may also lead to contradictions. Proving this property is an open question.

Consequently, in the following, we limit our analysis to $m = 3$, since extending our analysis to include $m \in \{4,5,6,7\}$ is beyond the scope of this work due to the substantial complexity involved and the lack of an effective methodology to handle them.

We begin with two examples that show that cyclic behavior can indeed occur.

Example 4.8 (First Cycle Type)**.** *The data*

$$A = \begin{pmatrix} 11 & -4 & 9 \\ -4 & 10 & -7 \\ 9 & -7 & 9 \end{pmatrix}, \quad b = \begin{pmatrix} -6 \\ 9 \\ -3 \end{pmatrix}, \quad u = \begin{pmatrix} -4 \\ 8 \\ 7 \end{pmatrix},$$

leads to the active set behavior illustrated in Fig. 4.2. Recall that the initial active set is exclusively determined by the initial iterate $(x^{(0)}, \mu^{(0)})$. Based on this initial active set, the algorithm will either converge or enter a cycle from which it cannot escape.

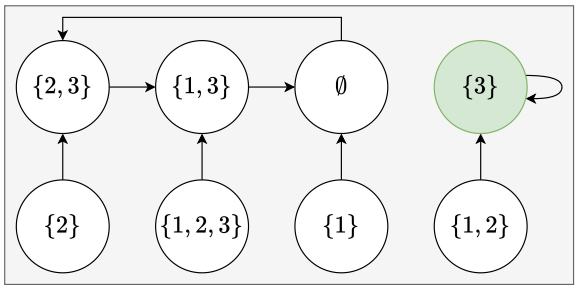

Figure 4.2 – Active set behavior for Example 4.8.

Example 4.9 (Second Cycle Type)**.** *The data*

$$A = \begin{pmatrix} 9 & -1 & -7 \\ -1 & 11 & 14 \\ -7 & 14 & 22 \end{pmatrix}, \quad b = \begin{pmatrix} -1 \\ 5 \\ 2 \end{pmatrix}, \quad u = \begin{pmatrix} -6 \\ 0 \\ 0 \end{pmatrix},$$

leads to the active set behavior illustrated in Fig. 4.3. Note that this is complementary to the active sets in Example 4.8 in the sense that the roles of active and inactive sets are interchanged.

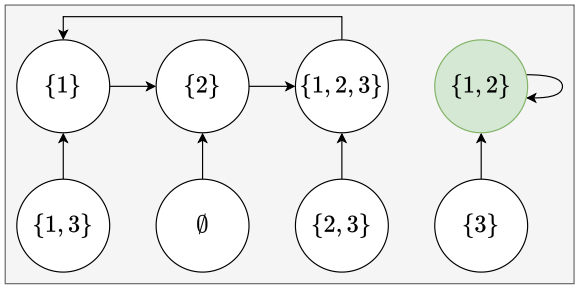

Figure 4.3 – Active set behavior for Example 4.9.

A variation of the upper bound shows that the cycle behavior depends on the upper bound. We emphasize departure from the original data u in bold face. For example,

$$u' = \begin{pmatrix} -6 \\ \mathbf{5} \\ \mathbf{-4} \end{pmatrix}$$

leads to the same cycle pattern as u. By contrast,

$$u'' = \begin{pmatrix} -6 \\ \mathbf{14} \\ 0 \end{pmatrix} \quad \text{and} \quad u''' = \begin{pmatrix} -6 \\ \mathbf{-1} \\ 0 \end{pmatrix}$$

lead to the active set behaviors shown in Figs. 4.4 and 4.5. In case of u'', all bounds are inactive.

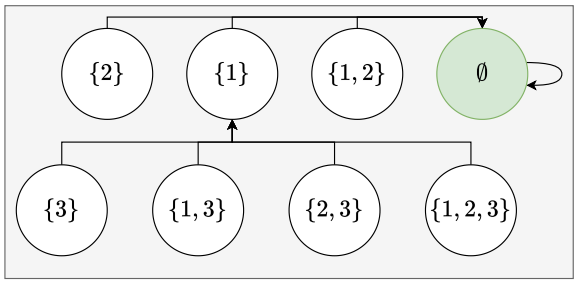

Figure 4.4 – Active set behavior for Example 4.9 with u''.

Numerical investigations with a large set of randomly generated problem data A, b, u suggest that Figs. 4.2 and 4.3 show the only possible cycle patterns in case $n = 3$, modulo permutations of $\{1,2,3\}$. The natural question arises as to why we observe precisely these

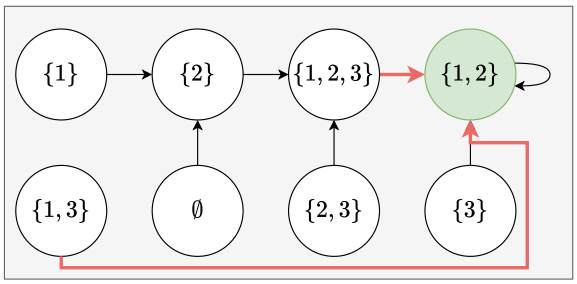

Figure 4.5 – Active set behavior for Example 4.9 with u'''.

two cycle types and whether we can find a theoretical upper bound on the number of such cycle types.

First, we test every potential active set cycle configuration with a cycle length of 3. In total, there exist 8 distinct possible active sets. Thus, there are $\binom{8}{3} = 56$ distinct combinations of three active sets that can be part of a cycle. A tedious but straightforward analysis reveals that the majority of these configurations are inconsistent with the necessary condition provided by Theorem 4.6. The only remaining feasible cycle types can be written as

$$\begin{aligned}
&\{i,j\} \to \{i,k\} \to \emptyset \to \{i,j\} && \text{(first type)}\\
&\{k\} \to \{j\} \to \{i,j,k\} \to \{k\} && \text{(second type)}\\
&\{k\} \to \{j\} \to \{i\} \to \{k\} && \text{(third type)}\\
&\{i,j\} \to \{i,k\} \to \{j,k\} \to \{i,j\} && \text{(fourth type)}
\end{aligned}$$

where $i, j, k \in \{1,2,3\}$ are three pairwise distinct indices. The first and second cycle types were observed to occur in Examples 4.8 and 4.9. It can be proved that the fourth cycle type can not occur:

Lemma 4.10. *Suppose that $A \in \mathbb{R}^{3\times 3}$ is symmetric and positive definite. Then, the active set sequence $\{i,j\} \to \{i,k\} \to \{j,k\} \to \{i,j\}$ is impossible.*

Proof. We start by listing the characterizing inequalities: For $\{i,j\} \to \{i,k\}$ we have

$$K := \frac{b_k - a_{ik}u_i - a_{jk}u_j}{a_{kk}} - u_k \geq 0, \quad (4.9a)$$

$$b_i - a_{ii}u_i - a_{ij}u_j - a_{ik}\frac{b_k - a_{ik}u_i - a_{jk}u_j}{a_{kk}} \geq 0, \quad (4.9b)$$

$$b_j - a_{ij}u_i - a_{jj}u_j - a_{jk}\frac{b_k - a_{ik}u_i - a_{jk}u_j}{a_{kk}} < 0. \quad (4.9c)$$

For $\{i,k\} \to \{j,k\}$ we have

$$J := \frac{b_j - a_{jk}u_k - a_{ij}u_i}{a_{jj}} - u_j \geq 0, \quad (4.10a)$$

$$b_k - a_{kk}u_k - a_{ik}u_i - a_{jk}\frac{b_j - a_{jk}u_k - a_{ij}u_i}{a_{jj}} \geq 0, \quad (4.10b)$$

$$b_i - a_{ik}u_k - a_{ii}u_i - a_{ij}\frac{b_j - a_{jk}u_k - a_{ij}u_i}{a_{jj}} < 0. \quad (4.10c)$$

For $\{j,k\} \to \{i,j\}$ we have

$$I := \frac{b_i - a_{ij}u_j - a_{ik}u_k}{a_{ii}} - u_i \geq 0, \quad (4.11a)$$

$$b_j - a_{jj}u_j - a_{jk}u_k - a_{ij}\frac{b_i - a_{ij}u_j - a_{ik}u_k}{a_{ii}} \geq 0, \quad (4.11b)$$

$$b_k - a_{jk}u_j - a_{kk}u_k - a_{ik}\frac{b_i - a_{ij}u_j - a_{ik}u_k}{a_{ii}} < 0. \quad (4.11c)$$

Since A is positive definite, it has positive diagonal entries, so (4.9a) implies

$$b_k - a_{ik}u_i - a_{jk}u_j - a_{kk}u_k \geq 0.$$

Combining this with (4.11c) we get

$$a_{ik}\frac{b_i - a_{ij}u_j - a_{ik}u_k}{a_{ii}} > a_{ik}u_i.$$

If $a_{ik} \leq 0$, this would result in a contradiction. Hence we must have $a_{ik} > 0$. By similar reasoning, we deduce $a_{ij} > 0$ and $a_{jk} > 0$. Combining (4.9b) and (4.10c), we obtain

$$a_{ij}J > a_{ik}K,$$

and similarly

$$a_{jk}K > a_{ij}I \quad \text{and} \quad a_{ik}I > a_{jk}J,$$

using (4.9c), (4.11b), (4.10b) and (4.11c). Combining these results with the fact that a_{ij}, a_{jk}, a_{ik}, I, J, $K \geq 0$, we conclude

$$a_{ij}J > a_{ik}K > \frac{a_{ij}}{a_{jk}}a_{ik}I > \frac{a_{ij}}{a_{jk}}a_{jk}J = a_{ij}J,$$

which is a contradiction and completes the proof. □

Despite thorough investigation, we were unable to find examples exhibiting the third cycle pattern. It is noteworthy that the first and second cycle types are complementary, as are the third and fourth cycle types. Further investigation along these lines might contribute to a deeper understanding of the PDASA also for values of $n > 3$.

5 Conclusion

This paper summarizes various notions of generalized derivatives, including Clarke's generalized Jacobian, slanting functions, and Newton differentiability. We state the equivalence of SSN applied to (QP) and the PDASA together with known global convergence results.

We then investigate the global convergence behavior of the PDASA for quadratic problems with symmetric positive definite matrices and upper bound constraints. For problems in dimension $n = 2$, we provide a new proof of global convergence. In case $n = 3$, we identify three theoretically possible cycles of length 3, two of which have been confirmed to occur in randomly generated examples.

During this research, several limitations were identified that could be addressed in future work. First, while the possibility of longer cycles in the case of $n = 3$ could not be ruled out, it has not been observed in practice. This suggests that a deeper understanding of the complementarity of problems concerning their active and inactive set behavior could be advantageous.

Overall, the global convergence theory even for problems with one-sided constraints remains incomplete. While known results provide sufficient conditions depending only on properties of the matrix A, the actual convergence behavior also depends on b and u.

CRediT Author Statement

Hannah Rickmann: Software, Validation, Formal analysis, Investigation, Writing - Original Draft, Visualization
Evelyn Herberg: Conceptualization, Methodology, Writing - Review and Editing, Supervision, Project administration
Roland Herzog: Conceptualization, Methodology, Writing - Review and Editing, Supervision, Project administration

A Timeline of Differentiability Concepts

For an illustration of the historical development of differentiability concepts, we provide a timeline of concepts and corresponding papers in Fig. A.1.

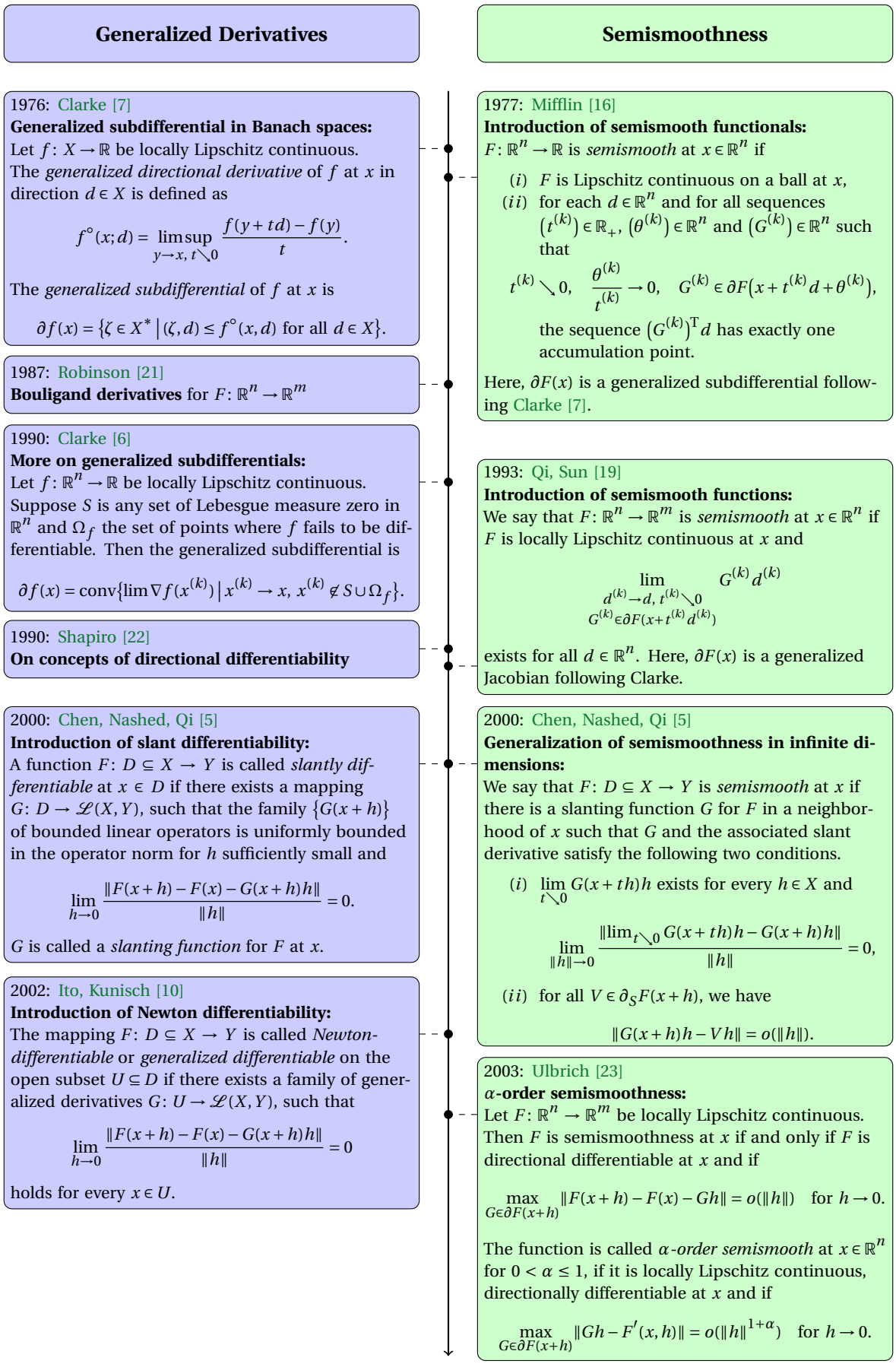

Figure A.1 – Timeline of differentiability concepts.

References

[1] I. Ben Gharbia; J. C. Gilbert. *Nonconvergence of the plain Newton-min algorithm for linear complementarity problems with a P-matrix. The full report*. 2011. HAL: hal-04148354.

[2] M. Bergounioux; M. Haddou; M. Hintermüller; K. Kunisch. "A comparison of a Moreau-Yosida-based active set strategy and interior point methods for constrained optimal control problems". *SIAM Journal on Optimization* 11.2 (2000), pp. 495–521. DOI: 10.1137/s1052623498343131.

[3] M. Bergounioux; K. Ito; K. Kunisch. "Primal-dual strategy for constrained optimal control problems". *SIAM Journal on Control and Optimization* 37.4 (1999), pp. 1176–1194. DOI: 10.1137/s0363012997328609.

[4] A. Berman; R. J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. Philadelphia: SIAM, 1994. DOI: 10.1137/1.9781611971262.

[5] X. Chen; Z. Nashed; L. Qi. "Smoothing methods and semismooth methods for nondifferentiable operator equations". *SIAM Journal on Numerical Analysis* 38 (2000), pp. 1200–1216. DOI: 10.1137/S0036142999356719.

[6] F. H. Clarke. *Optimization and Nonsmooth Analysis*. 2nd ed. Vol. 5. Classics in Applied Mathematics. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 1990. DOI: 10.1137/1.9781611971309.

[7] F. H. Clarke. "A new approach to Lagrange multipliers". *Mathematics of Operations Research* 1.2 (1976), pp. 165–174. DOI: 10.1287/moor.1.2.165.

[8] R. W. Cottle; J.-S. Pang; R. E. Stone. *The Linear Complementarity Problem*. Society for Industrial and Applied Mathematics, 2009. DOI: 10.1137/1.9780898719000. eprint: https://epubs.siam.org/doi/pdf/10.1137/1.9780898719000. URL: https://epubs.siam.org/doi/abs/10.1137/1.9780898719000.

[9] M. Hintermüller; K. Ito; K. Kunisch. "The primal-dual active set strategy as a semismooth Newton method". *SIAM Journal on Optimization* 13.3 (2002), pp. 865–888. DOI: 10.1137/s1052623401383558.

[10] K. Ito; K. Kunisch. "Semi-smooth Newton methods for variational inequalities of the first kind". *RAIRO Modélisation Mathématique et Analyse Numérique* 37 (2002), pp. 41–62. DOI: 10.1051/m2an:2003021.

[11] K. Ito; K. Kunisch. "Semi-smooth Newton methods for state-constrained optimal control problems". *Systems and Control Letters* 50 (2003), pp. 221–228. DOI: 10.1016/S0167-6911(03)00156-7.

[12] K. Ito; K. Kunisch. "The primal-dual active set method for nonlinear optimal control problems with bilateral constraints". *SIAM Journal on Control and Optimization* 43.1 (2004), pp. 357–376. DOI: 10.1137/s0363012902411015.

[13] K. Ito; K. Kunisch. *Lagrange Multiplier Approach to Variational Problems and Applications*. Vol. 15. Advances in Design and Control. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 2008. DOI: 10.1137/1.9780898718614.

[14] C. T. Kelley; E. W. Sachs. "Multilevel algorithms for constrained compact fixed point problems". *SIAM Journal on Scientific Computing* 15.3 (1994), pp. 645–667. DOI: 10.1137/0915042.

[15] K. Kunisch; A. Rösch. "Primal-dual active set strategy for a general class of constrained optimal control problems". *SIAM Journal on Optimization* 13.2 (2002), pp. 321–334. DOI: 10.1137/s1052623499358008.

[16] R. Mifflin. "Semismooth and semiconvex functions in constrained optimization". *SIAM Journal on Control and Optimization* 15.6 (1977), pp. 959–972. DOI: 10.1137/0315061.

[17] J. Nocedal; S. J. Wright. *Numerical Optimization*. 2nd ed. New York: Springer, 2006. DOI: 10.1007/978-0-387-40065-5.

[18] L. Q. Qi. "Convergence analysis of some algorithms for solving nonsmooth equations". *Mathematics of Operations Research* 18.1 (1993), pp. 227–244. DOI: 10.1287/moor.18.1.227.

[19] L. Qi; J. Sun. "A nonsmooth version of Newton's method". *Mathematical Programming* 58.1–3 (1993), pp. 353–367. DOI: 10.1007/bf01581275.

[20] H. Rickmann. *Primal-Dual Active Set Algorithm for Quadratic Optimization Problems*. GitHub. 2024. URL: https://github.com/HannahRickmann/SSN.

[21] S. M. Robinson. “Local structure of feasible sets in nonlinear programming, Part III: stability and sensitivity”. *Nonlinear Analysis and Optimization*. Springer Berlin Heidelberg, 1987, pp. 45–66. DOI: 10.1007/bfb0121154.

[22] A. Shapiro. “On concepts of directional differentiability”. *Journal of Optimization Theory and Applications* 66.3 (1990), pp. 477–487. DOI: 10.1007/bf00940933.

[23] M. Ulbrich. “Semismooth Newton methods for operator equations in function spaces”. *SIAM Journal on Control and Optimization* 13.3 (2003), pp. 805–842. DOI: 10.1137/s1052623400371569.

[24] M. Ulbrich. *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*. Vol. 11. MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, 2011. DOI: 10.1137/1.9781611970692.

[25] M. Ulbrich; S. Ulbrich. “Superlinear convergence of affine-scaling interior-point Newton methods for infinite-dimensional nonlinear problems with pointwise bounds”. *SIAM Journal on Control and Optimization* 38.6 (2000), pp. 1938–1984. DOI: 10.1137/s0363012997325915.