Polyhedral Gauß–Seidel Converges

Carsten Gräser and Oliver Sander

March 22, 2010

Abstract

We prove global convergence of an inexact polyhedral Gauß–Seidel method for the minimization of strictly convex functionals that are continuously differentiable on each polyhedron of a polyhedral decomposition of their domains of definition. While being known to be very slow by themselves, such methods are a cornerstone for fast, globally convergent multigrid methods. Our result generalizes the proof of Kornhuber and Krause [7] for differentiable functionals on the Gibbs simplex. Example applications are given that require the generality of our approach.

1 Polyhedral Gauß–Seidel

Consider the minimization problem

$$x^* \in \mathbb{R}^n : \qquad J(x^*) \leq J(y) \qquad \forall y \in \mathbb{R}^n \tag{1}$$

for a functional

$$J : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}.$$

We assume that J and its domain $\operatorname{dom} J := \{x \in \mathbb{R}^n \mid J(x) < \infty\}$ have the following properties:

(A1) J is strictly convex and $\operatorname{dom} J \neq \emptyset$,

(A2) J is coercive, i.e., $J(x) \to \infty$ if $\|x\| \to \infty$,

(A3) J is lower semi-continuous on $\mathbb{R}^n$.

By conditions (A1), (A2), and (A3) the minimization problem (1) has a unique solution [4]. While (A1) already implies continuity of J on the relative interior of $\operatorname{dom} J$ [8], we need a slightly stronger condition.

(A4) J is continuous on $\operatorname{dom} J$ and $\overline{\operatorname{dom} J}$ is a polyhedron.

This continuity condition does only exclude certain functions that degenerate on $\partial(\operatorname{dom} J)$, like the one in the following example. It will be convenient to use the characteristic function $\chi_K : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ of a set $K \subset \mathbb{R}^n$ given by

$$\chi_K(x) := \begin{cases} 0 & \text{if } x \in K, \\ \infty & \text{else.} \end{cases}$$

Example 1. *The functional* $J : \mathbb{R}^2 \to \mathbb{R} \cup \{\infty\}$ *given by*

$$J(x) = \|x\|^2 + \frac{x_2^2}{x_1} + \chi_K(x) \qquad \forall x \neq 0, \qquad J(0) = 0,$$

with $K = \{x \in \mathbb{R}^2 \mid x_1 > 0$ *or* $x = 0\}$ *satisfies (A1)–(A3) but not (A4), since it is not continuous at* $0 \in K = \operatorname{dom} J$.

We now introduce the *Polyhedral Gauß–Seidel Algorithm (PGS)*. Let

$$\mathcal{E} = \{v_1, \dots, v_m\} \subset \mathbb{R}^n$$

be a finite set of search directions, ν an iteration number, and let $x^\nu \in \operatorname{dom} J$. Then one iteration of PGS consists of the following steps:

1. set $x_0^\nu = x^\nu$
2. for $i = 1, \dots, m$ set
$$x_i^\nu = x_{i-1}^\nu + \operatorname*{arg\,min}_{v \in \operatorname{span} v_i} J(x_{i-1}^\nu + v)$$
3. set $x^{\nu+1} = x_m^\nu$.

Note that this is the standard Gauß–Seidel method if $\mathcal{E}$ is the set of Euclidean coordinate vectors.

Remark. *For simplicity we have stated here the exact version of PGS. Later we will also consider an inexact variant, where the minimization in Step 2 only has to be within a given fraction of the true result. Our analysis also covers this more general case.*

While the algorithm does not change if we replace $v \in \mathcal{E}$ by $-v$ we will need directions with the proper sign in many subsequent statements and results. Thus we define

$$\mathcal{E}^\pm := \mathcal{E} \cup -\mathcal{E},$$

and state assumptions on the set $\mathcal{E}^\pm$. We will prove convergence if the set of search directions $\mathcal{E}$ provides descent directions at any point $x \in \operatorname{dom} J \setminus \{x^*\}$, i.e., if $\mathcal{E}$ satisfies:

(A5) For all $x \in \operatorname{dom} J \setminus \{x^*\}$ there is a direction $v \in \mathcal{E}^\pm$ and an $\epsilon > 0$ such that $J(x + \epsilon v) < J(x)$.

While this condition if sufficient for our convergence analysis it can be difficult to verify in practice. We will therefore tighten our assumptions somewhat to allow easier construction of search directions. More specifically, we will also assume that J is piecewise smooth on a partition of $\operatorname{dom} J$ in polyhedra. In the following, $T_\mathcal{C}(x)$ is the tangent cone of a polyhedron $\mathcal{C}$ at x and $\operatorname{cone} M$ is the cone generated by the set M. Both notions are properly defined in the next section. Furthermore, $B_\epsilon(x)$ denotes the open ball with radius ϵ around $x \in \mathbb{R}^n$.

(A6) There is a partitioning of $\overline{\operatorname{dom} J}$ in a finite set of convex closed polyhedra $\mathcal{C}_1, \ldots, \mathcal{C}_N$. Also, there is a set $\mathcal{S} \subset \operatorname{dom} J$ such that for all $i = 1, \ldots, N$ and all $x \in \mathcal{C}_i \cap \mathcal{S}$ there is a subset $\mathcal{E}_{x,i} \subseteq \mathcal{E}^{\pm}$ with

$$T_{\mathcal{C}_i}(x) = \operatorname{cone} \mathcal{E}_{x,i}, \tag{2}$$

and there is an $\epsilon > 0$ and a smooth extension $\bar{J} \in C^1(\overline{B_\epsilon(x)})$ with

$$\bar{J}|_{\overline{B_\epsilon(x)} \cap \mathcal{C}_i \cap \operatorname{dom} J} = J|_{\overline{B_\epsilon(x)} \cap \mathcal{C}_i \cap \operatorname{dom} J}.$$

Furthermore, for all $x \in \operatorname{dom} J \setminus (\mathcal{S} \cup \{x^*\})$ there is a direction $v \in \mathcal{E}^{\pm}$ and an $\epsilon > 0$ such that $J(x + \epsilon v) < J(x)$.

By this assumption J is allowed to be nonsmooth across the intersection of two different $\mathcal{C}_i$ and on the set $\operatorname{dom} J \setminus \mathcal{S}$. Intuitively, $\mathcal{S}$ is the set where J is "smooth enough", and we will show that the existence of descent directions for $x \in \mathcal{S}$ is guaranteed by the condition (2) on the tangent cones. If $\mathcal{S} = \operatorname{dom} J$ this implies that the search directions only need to depend on the partitioning in polyhedra and not on values or derivatives of J. For all $x \notin \mathcal{S}$ additional search directions may have to be provided to ensure convergence.

Example 2. *Let $\alpha \in [0, 1]$. Then the functional $J : \mathbb{R}^2 \to \mathbb{R} \cup \{\infty\}$,*

$$J(x) = \|x\|^2 - (x_1 x_2)^\alpha + \chi_{[0,1]^2}(x),$$

satisfies (A1)–(A4) with $\operatorname{dom} J = [0, 1]^2$. For $\alpha \in \{0, 1\}$ the set $\mathcal{E} = \{e_1, e_2\}$ of the Euclidean coordinate vectors satisfies (A6) with $N = 1$, $\mathcal{C}_1 = \mathcal{S} = \operatorname{dom} J$. If however $\alpha \in (0, 1)$, then $\mathcal{S} = (0, 1]^2$ and for $x \in \operatorname{dom} J \setminus (\mathcal{S} \cup \{0\})$ we even have

$$\lim_{h \searrow 0} \frac{J(x + he_i) - J(x)}{h} = -\infty.$$

However $0 \notin \mathcal{S}$ but neither e_1 nor e_2 is a descent direction there. Hence we need to additionally use, e.g., $e_1 + e_2$ in order to satisfy (A6).

The main result of this article is that the polyhedral Gauß–Seidel algorithm will converge to the unique solution of the minimization problem if the set of search directions $\mathcal{E}$ contains sufficient information about the partitioning of the domain in convex polyhedra $\mathcal{C}_i$ and about points of non-differentiability.

Theorem 1. *Assume that $J : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ has the properties (A1)–(A4) and that the set of search directions $\mathcal{E} = \{v_1, \ldots, v_m\} \subset \mathbb{R}^n$ satisfies (A5) or (A6). Then for all x^0 with $J(x^0) < \infty$ the sequence generated by PGS converges to the unique minimizer x^* of J.*

This article was inspired by a result of Kornhuber and Krause [7], where they proved convergence for a particular J defined on a simplex. The energy fulfilled (A1)–(A4) and was smooth in the interior of the simplex while $\mathcal{E}^{\pm}$ was the set of simplex edges. Unfortunately, their arguments would not carry over to the more general functionals J discussed here.

An early result on constrained Gauß–Seidel methods can be found in the book by Glowinski et al. [5]. There it is shown that the constrained Gauß–Seidel method converges if the functional is strictly convex, differentiable, and if the admissible set is the product of intervals. Tseng [10] generalized this for convex differentiable functionals defined also on products of intervals. By a duality approach the method could also solve strongly convex, not necessarily differentiable minimization problems on polyhedral domains.

Minimization problems on polyhedral domains have also been studied intensively in the field of linear programming, where unfortunately it is always assumed that J is linear. We refer the reader to Schrijver [9] for an introduction.

The article proceeds as follows. In Section 2 we briefly review polyhedral sets and list a few basic properties. Section 3 is dedicated to the proof of Theorem 1. In Section 4 we show how sets of search directions can be constructed for actual polyhedral decompositions. Finally, in Section 5 we present a few applications that result in minimization problems of the type considered here.

2 Polyhedral Sets and Their Tangent Cones

Although non-convex polyhedra are sometimes allowed in the literature we stick to the convention that polyhedra are always convex as implied by the following definition.

Definition 1. *A polyhedron is a set*

$$\mathcal{C} = \{w \in \mathbb{R}^n \mid \langle a_i, w\rangle \geq b_i,\ i = 1, \dots, m\},$$

where $a_i \in \mathbb{R}^n$ *and* $b_i \in \mathbb{R}$ *for* $i = 1, \dots, m$.

It is easy to see that polyhedra are closed and convex, but not necessarily bounded. The dimension of a polyhedron $\mathcal{C}$ is the dimension of its affine hull. A face of $\mathcal{C}$ is a set $\mathcal{F}$ such that there exist $a \in \mathbb{R}^n$, $b \in \mathbb{R}$ with $\langle a, x\rangle \geq b$, $\forall x \in \mathcal{C}$ and $\mathcal{F} = \{x \in \mathcal{C} \mid \langle a, x\rangle = b\}$. Faces are again polyhedra. Zero- and one-dimensional faces are called vertices, and edges, respectively [11].

For a polyhedron $\mathcal{C}$ we define its recession cone

$$\operatorname{rec}(\mathcal{C}) = \{y \in \mathbb{R}^n \mid x + ty \in \mathcal{C} \text{ for all } x \in \mathcal{C}, t \geq 0\}.$$

By this definition it is obvious that for any subset $M \subset \mathcal{C}$ we have

$$M + \operatorname{rec}(\mathcal{C}) \subset \mathcal{C}$$

with '=' at least for $M = \mathcal{C}$. The question if there are smaller sets such that equality holds is answered by the following central result about polyhedra, usually named after Minkowski and Weyl.

Theorem 2. *Let* $\mathcal{C}$ *be a polyhedron.*

1. *There is a representation*

$$\mathcal{C} = \left\{ \sum_{i=1}^{k} \lambda_i p_i + \sum_{j=1}^{l} \mu_j d_j \;\middle|\; \lambda_i \geq 0,\ \sum_{i=1}^{k} \lambda_i = 1,\ \mu_j \geq 0 \right\}$$

   with points $p_1, \dots, p_k \in \mathbb{R}^n$ and directions $d_1, \dots, d_l \in \mathbb{R}^n$.

2. *If $\mathcal{C}$ has at least one vertex, then there is a unique minimal such representation, where the p_i are the vertices of $\mathcal{C}$ and the d_j are the infinite edges of $\mathcal{C}$.*

Proof. Schrijver [9, Thm. 8.5] □

Next, we define tangent cones of convex sets.

Definition 2. *Let $K \subset \mathbb{R}^n$ be convex set. The tangent cone $T_K(x)$ of K at x is the set*

$$T_K(x) = \overline{\{v \in \mathbb{R}^n \mid x + \epsilon v \in K \ \text{ for } \epsilon > 0 \text{ small enough}\}}.$$

Tangent cones are cones, i.e., from $v \in T_{\mathcal{C}}(x)$ follows $\lambda v \in T_{\mathcal{C}}(x)$ for all $\lambda \geq 0$. Also, we have always $\mathcal{C} \subset T_{\mathcal{C}}(x) + x$. Finally, tangent cones of polyhedra are polyhedra themselves [8, Thm. 6.46]. The cone generated by a set Y is defined as

$$\operatorname{cone} Y := \left\{ x \in \mathbb{R}^n \;\middle|\; \exists k > 0,\ \lambda_j \geq 0,\ y_j \in Y \text{ such that } x = \sum_{j=1}^{k} \lambda_j y_j \right\}.$$

For polyhedral cones with a vertex we get the following corollary of Theorem 2.

Corollary 1. *Let T be a polyhedral cone with a vertex. Then T is generated by its edges with proper orientation.*

Tangent cones of polyhedra exhibit several useful local properties. As a fundamental tool we find that $\mathcal{C}$ and its tangent cone at x coincide in a neighborhood around x.

Lemma 1. *Let $\mathcal{C}$ be a polyhedron and $x \in \mathcal{C}$. Then there exists an $\epsilon > 0$ such that*

$$(x + T_{\mathcal{C}}(x)) \cap B_\epsilon(x) = \mathcal{C} \cap B_\epsilon(x).$$

Proof. Rockafellar and Wets [8] □

As direct consequence of Lemma 1 we may drop the closure in the definition of $T_K(x)$ if K is a polyhedron.

In general it is not true that the ϵ in Lemma 1 can be chosen uniformly in a neighborhood around some $x_0 \in \mathcal{C}$. As a simple counterexample consider x_0 to be a vertex of $\mathcal{C}$. However for the case of a fixed tangent cone whose apex is moved we get the following local stability result.

Lemma 2. *Let $\mathcal{C}$ be a polyhedron and $x_0 \in \mathcal{C}$. Then there exists an $\epsilon > 0$ such that*

$$(x + T_{\mathcal{C}}(x_0)) \cap B_\epsilon(x) \subseteq \mathcal{C} \cap B_\epsilon(x) \qquad \forall x \in \mathcal{C} \cap B_\epsilon(x_0).$$

Proof. Let $x_0 \in \mathcal{C}$ and select $2\epsilon > 0$ according to Lemma 1 such that

$$(T_{\mathcal{C}}(x_0) + x_0) \cap B_{2\epsilon}(x_0) = \mathcal{C} \cap B_{2\epsilon}(x_0). \tag{3}$$

Then for any $x \in \mathcal{C} \cap B_\epsilon(x_0)$ and $y \in (x + T_{\mathcal{C}}(x_0)) \cap B_\epsilon(x)$ we have

$$y - x,\ x - x_0 \in T_{\mathcal{C}}(x_0) \cap B_\epsilon(0),$$

and thus, using the triangle inequality,

$$y - x_0 \in T_{\mathcal{C}}(x_0) \cap B_{2\epsilon}(0).$$

Adding x_0 and using (3) gives

$$y \in \mathcal{C} \cap B_{2\epsilon}(x_0).$$

Since y was chosen such that $y \in B_\epsilon(x)$ this particularly gives

$$y \in \mathcal{C} \cap B_\epsilon(x),$$

which proves the assertion. □

As a direct corollary of Lemma 2 we find that tangent cones of neighboring points are always bigger if the neighborhood is small enough.

Corollary 2. *Let $\mathcal{C}$ be a polyhedron and $x_0 \in \mathcal{C}$. Then there exists an $\epsilon > 0$ such that*

$$T_{\mathcal{C}}(x_0) \subseteq T_{\mathcal{C}}(x) \qquad \forall x \in \mathcal{C} \cap B_\epsilon(x_0).$$

In general for closed convex sets $\mathcal{C}$ that are not polyhedral neither the assertions of Lemma 1 and Lemma 2 nor the one in Corollary 2 are true. This can easily be checked, e.g., for $\mathcal{C} = \overline{B_1(0)} \subset \mathbb{R}^2$ where $T_{\mathcal{C}}(x) = \{y \in \mathbb{R}^2 \mid \langle y, x \rangle \geq 0\}$ for all $x \in \partial\mathcal{C}$.

The Minkowski–Weyl representation of polyhedra given by Theorem 2 can also be used to characterize the tangent cones in a nice way. This is the key result to constructively describing sets $\mathcal{E}$ of search directions.

Theorem 3. *Let $\mathcal{C}$ be a polyhedron and $x \in \mathcal{C}$. Further let $p_j \in \mathcal{C}, j = 1, \ldots, k$ be points and $d_j \in \operatorname{rec}(\mathcal{C}), j = 1, \ldots, l$ be directions such that*

$$x = \sum_{j=1}^{k} \lambda_j p_j + \sum_{j=1}^{l} \mu_j d_j, \qquad \lambda_j \geq 0, \quad \sum_{j=1}^{k} \lambda_j = 1, \quad \mu_j \geq 0. \tag{4}$$

Then

$$T_{\mathcal{C}}(x) = \sum_{j=1}^{k} \lambda_j T_{\mathcal{C}}(p_j) + \sum_{j=1}^{l} \mu_j \operatorname{cone}\{-d_j\}. \tag{5}$$

Proof. Let $w \in T_{\mathcal{C}}(x)$. By Lemma 1 we get $x + \epsilon w \in \mathcal{C}$ for $\epsilon > 0$ small enough. From $\mathcal{C} \subset p_j + T_{\mathcal{C}}(p_j)$ we know that there are $w_j \in T_{\mathcal{C}}(p_j)$ with

$$x + \epsilon w = p_j + w_j$$

for all $1 \leq j \leq k$. We can now get a representation of ϵw in terms of the w_j by noting that

$$\begin{aligned}
\epsilon w &= \Big(\sum_{j=1}^{k} \lambda_j\Big)(x + \epsilon w) - x = \sum_{j=1}^{k} \Big(\lambda_j (x + \epsilon w)\Big) - x \\
&= \sum_{j=1}^{k} \Big(\lambda_j (p_j + w_j)\Big) - \Big(\sum_{j=1}^{k} \lambda_j p_j + \sum_{j=1}^{l} \mu_j d_j\Big) \\
&= \sum_{j=1}^{k} \lambda_j w_j - \sum_{j=1}^{l} \mu_j d_j.
\end{aligned}$$

Thus we have shown that ϵw, and thus w, is contained in the cone on the right hand side of (5).

To show that also '$\supset$' holds in (5), let first $w \in \lambda_j T_{\mathcal{C}}(p_j)$ for some j. Using Lemma 1 we find $p_j + \epsilon w \in \mathcal{C}$ for $\epsilon > 0$ small enough. Hence $x + \lambda_j \epsilon w$ can be represented similarly to x in (4) with p_j replaced by $p_j + \epsilon w$, i.e.,

$$x + \lambda_j \epsilon w \in \mathcal{C} + \mathrm{rec}(\mathcal{C}) = \mathcal{C}.$$

This implies that both $\lambda_j \epsilon w$ and w are contained in $T_{\mathcal{C}}(x)$ and thus $\lambda_j T_{\mathcal{C}}(p_j) \subset T_{\mathcal{C}}(x)$. Now consider $-\mu_j d_j$ for some j. Then we have

$$x - \frac{\mu_j}{2} d_j \in \mathcal{C} + \mathrm{rec}(\mathcal{C}) = \mathcal{C},$$

which implies $-(\mu_j/2)d_j \in T_{\mathcal{C}}(x)$. Hence the cone spanned by all such w and $-\mu_j d_j$ is also contained in $T_{\mathcal{C}}(x)$. $\square$

Note that (5) is equivalent to

$$T_{\mathcal{C}}(x) = \sum_{j:\lambda_j > 0} T_{\mathcal{C}}(p_j) + \sum_{j:\mu_j > 0} \mathrm{cone}\{-d_j\}.$$

Combining Theorems 2 and 3 we see that any tangent cone $T_{\mathcal{C}}(x)$ is spanned by a selection from a finite number of cones, namely

$$T_{\mathcal{C}}(p_1), \ldots T_{\mathcal{C}}(p_k), \quad \text{and} \quad \mathrm{cone}\{-d_1\}, \ldots, \mathrm{cone}\{-d_l\},$$

with fixed p_j, d_j selected according to Theorem 2.

3 Convergence of Polyhedral Gauß–Seidel

The convergence proof for the PGS algorithm and its inexact version uses the same building blocks as the proof for the special case in [7]:

1. monotonicity of intermediate iterates with respect to J,
2. continuity of the minimization operator, $x \mapsto \arg\min_{\alpha \in \mathbb{R}} J(x + \alpha d)$ (in [7] this was only assumed),
3. the fact that fixed points of the PGS method are minimizers.

While the monotonicity is obvious since the algorithm is based on successive minimization Properties 2 and 3 are not trivial in the presented general setting. We will prove them in the following two subsections.

3.1 Continuity of the Minimization Operator

In contrast to the case where the nonsmooth part of J decomposes into search directions, i.e.,

$$J(x) = J_0(x) + \sum_{i=1}^{n} J_i(x_i) \tag{6}$$

with continuously differentiable J_0 and scalar functions J_i (as in, e.g., [6]), the proof of the continuity of $x \mapsto \arg\min_{\alpha \in \mathbb{R}} J(x + \alpha d)$ can no longer be based on standard arguments for the Lipschitz continuity of the solution operator to scalar variational inequalities. Instead, we give a direct proof.

Lemma 3. *Let $J : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ satisfy (A1)–(A4). Then the set $\overline{\operatorname{dom} J} \setminus \operatorname{dom} J$ is closed.*

Proof. Assume that $A := \overline{\operatorname{dom} J} \setminus \operatorname{dom} J$ is not closed. Then there is a sequence $x^\nu \in A$ such that $x^\nu \to x \in \overline{A} \setminus A \subset (\overline{\operatorname{dom} J} \cap \operatorname{dom} J) = \operatorname{dom} J$.

For each x^ν there must then also be a sequence $x_k^\nu \in \operatorname{dom} J$ with $x_k^\nu \to x^\nu$ for $k \to \infty$. Using lower semi-continuity of J we find that

$$\varliminf_{k \to \infty} J(x_k^\nu) \geq J(x^\nu) = \infty.$$

Hence for each ν there must be a k_ν such that $\|x^\nu - x_{k_\nu}^\nu\| \leq \|x - x^\nu\|$ and $J(x_{k_\nu}^\nu) \geq \frac{1}{\nu}$ which implies

$$\varliminf_{\nu \to \infty} J(x_{k_\nu}^\nu) = \infty.$$

On the other hand from $\|x - x_{k_\nu}^\nu\| \leq 2\|x - x^\nu\|$ we get $x_{k_\nu}^\nu \to x$ and thus by continuity $J(x_{k_\nu}^\nu) \to J(x) \neq \infty$. ∎

Lemma 4. *Let $J : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ satisfy (A1)–(A4). Then, for any fixed subspace V of $\mathbb{R}^n$ the mapping*

$$\begin{aligned} \alpha(\cdot, V) \quad &: \quad \operatorname{dom} J \to V \\ \alpha(x, V) \quad &= \quad \operatorname*{arg\,min}_{v \in V} J(x+v) \end{aligned}$$

is continuous.

Proof. Since the assertion is obvious for $V = \{0\}$ we only consider $V \neq \{0\}$. Assume that $\alpha(\cdot, V)$ is not continuous at some $x \in \operatorname{dom} J$ and set $y = x + \alpha(x, V) \in \operatorname{dom} J$. Then there is a sequence

$$x^\nu \in \operatorname{dom} J, \qquad x^\nu \to x$$

such that $y^\nu = x^\nu + \alpha(x^\nu, V) \in \operatorname{dom} J$ satisfies $\|y^\nu - y\| \geq \delta$ for some $\delta > 0$. Since J is continuous, $J(x^\nu)$ is bounded and by definition of α we have

$$J(y^\nu) \leq J(x^\nu) \leq C \tag{7}$$

for some constant C. Hence the sequence y^ν is bounded and there exists a subsequence of y^ν converging to some $\bar{y}$. Lower semi-continuity of J and (7) directly imply $J(\bar{y}) \leq C$ and thus $\bar{y} \in \operatorname{dom} J$. For simplicity we relabel and call this convergent subsequence y^ν from now on.

By Lemma 2 applied to $\mathcal{C} := \overline{\operatorname{dom} J}$ there is an $\epsilon > 0$ such that

$$(z + T_{\mathcal{C}}(\bar{y})) \cap B_\epsilon(z) \subset \mathcal{C} \qquad \forall z \in \mathcal{C} \cap B_\epsilon(\bar{y}).$$

Using this and convergence of y^ν we find that there is a $\nu_0 \in \mathbb{N}$ such that

$$(y^\nu + T_{\mathcal{C}}(\bar{y})) \cap B_\epsilon(y^\nu) \subset \mathcal{C} \qquad \nu > \nu_0. \tag{8}$$

Now consider $\lambda(y - \bar{y})$ with

$$\lambda := \min\left\{ \frac{\epsilon}{2\|y - \bar{y}\|}, \frac{1}{2} \right\} \in (0, 1).$$

With this we have $y^\nu + \lambda(y - \bar{y}) \in B_\epsilon(y^\nu)$ and by $y, \bar{y} \in \mathcal{C}$ we also have $y - \bar{y} \in T_{\mathcal{C}}(\bar{y})$. Using (8) this implies

$$y^\nu + \lambda(y - \bar{y}) \in \mathcal{C} = \overline{\operatorname{dom} J}.$$

Strict convexity of J, $\lambda \in (0, 1)$, and the minimality of y also give

$$\begin{aligned} J(\bar{y} + \lambda(y - \bar{y})) &< (1 - \lambda) J(\bar{y}) + \lambda J(y) \\ &< (1 - \lambda) J(\bar{y}) + \lambda J(\bar{y}) = J(\bar{y}) < \infty, \end{aligned} \tag{9}$$

and hence $\bar{y} + \lambda(y - \bar{y}) \in \operatorname{dom} J$. Thus for large enough ν we even have

$$y^\nu + \lambda(y - \bar{y}) \in \operatorname{dom} J,$$

otherwise Lemma 3 would imply $\bar{y} + \lambda(y - \bar{y}) \in \overline{\operatorname{dom} J} \setminus \operatorname{dom} J$. Now we can use the continuity of J on $\operatorname{dom} J$ and the inequality

$$J(y^\nu) \leq J(y^\nu + \lambda(y - \bar{y}))$$

which, together with (9), gives the contradiction

$$J(\bar{y}) \leq J(\bar{y} + \lambda(y - \bar{y})) < J(\bar{y}).$$

□

For the case $\dim V = 1$, Lemma 4 reduces to the following result, where, in an abuse of notation, $\alpha(\cdot, \cdot)$ now returns a scalar instead of a vector.

Corollary 3. *Let $J : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ satisfy (A1)–(A4). Then, for any fixed direction $v \in \mathbb{R}^n \setminus \{0\}$ the mapping*

$$\begin{aligned} \alpha(\cdot, v) \quad &: \quad \operatorname{dom} J \to \mathbb{R} \\ \alpha(x, v) \quad &= \quad \operatorname*{arg\,min}_{\hat{\alpha} \in \mathbb{R}} J(x + \hat{\alpha} v) \end{aligned}$$

is continuous.

Note that the assumption that $\overline{\operatorname{dom} J}$ is a polyhedron is not only needed in order to allow for a partitioning of $\overline{\operatorname{dom} J}$ in polyhedra. Indeed, $\alpha(\cdot, v)$ need not be continuous if $\operatorname{dom} J$ has a curved boundary.

Example 3. *Let $K \subset \mathbb{R}^3$ be the cone $K := \{x \in \mathbb{R}^3 \mid x_1^2 + x_2^2 \leq x_3\}$. Then the functional $J : \mathbb{R}^3 \to \mathbb{R} \cup \{\infty\}$,*

$$J(x) = \|x\|^2 + \chi_K(x),$$

satisfies (A1)–(A4) except that $\overline{\operatorname{dom} J} = K$ is not a polyhedron. For $x = (1, 0, 1)$ and $\alpha(\cdot, x)$ as defined in Corollary 3 we have $\alpha(x, x) = -1$ but for any $y \neq x$ with $y_1^2 + y_2^2 = y_3 = 1$ we get $\alpha(y, x) = 0$, since

$$J(y + \epsilon x) > J(y) \qquad \text{and} \qquad y - \epsilon x \notin K \qquad \forall \epsilon > 0.$$

Thus $\alpha(\cdot, x)$ is not continuous in x.

3.2 Fixed Points are Minimizers

The second essential ingredient for the convergence proof of the Polyhedral Gauß–Seidel method is the fact that fixed points of the algorithm are solutions. If J decomposes according to (6) and the search directions form a basis, this can be shown by suitably combining the one-dimensional variational inequalities that determine the fixed point property. The result is an n-dimensional variational inequality which shows that the fixed point is indeed a minimum of the functional [7]. In our more general case this is not possible.

While (A5) directly implies that fixed points are minimizers showing this for (A6) is more complicated. The proof of the following lemma circumvents the non-smoothness by localizing near the fixed point such that the problem looks like a smooth problem there.

Lemma 5. *Assume that $J : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ has the properties (A1)–(A4) and that the set of search directions $\mathcal{E} = \{v_1, \dots, v_m\} \subset \mathbb{R}^n$ satisfies (A6). Then $\mathcal{E}$ satisfies (A5), i.e., for all $x \in \operatorname{dom} J \setminus \{x^*\}$ there is a direction $v \in \mathcal{E}^\pm$ and an $\epsilon > 0$ such that $J(x + \epsilon v) < J(x)$.*

Proof. Consider $x \in \mathcal{S} \setminus \{x^*\}$. Then there is a $\mathcal{C}_i$ such that

$$x \in \mathcal{C}_i \qquad \text{and} \qquad (x, x^*) \cap \mathcal{C}_i \neq \emptyset,$$

where we have used (x, x^*) to denote the open line segment from x to x^*. By (A6) there is $\mathcal{E}_{x,i} \subset \mathcal{E}^\pm$ such that $T_{\mathcal{C}_i}(x) = \operatorname{cone} \mathcal{E}_{x,i}$. For $w := x^* - x \in T_{\mathcal{C}_i}(x)$ this guarantees that there are $d_1, \dots, d_k \in \mathcal{E}_{x,i}$ and $\lambda_1, \dots, \lambda_k \geq 0$ such that

$$w = \sum_{j=1}^{k} \lambda_j d_j. \tag{10}$$

Now let $\bar{J} \in C^1(\overline{B_\epsilon(x)})$ be the smooth extension of J on $B_\epsilon(x)$ for some $\epsilon > 0$ provided by (A6). Then w is a descent direction for $\bar{J}$ and we can use the representation (10) to get

$$0 > \frac{\partial \bar{J}}{\partial w}(x) = \big\langle \nabla \bar{J}(x), w \big\rangle = \sum_{j=1}^{k} \lambda_j \big\langle \nabla \bar{J}(x), d_j \big\rangle = \sum_{j=1}^{k} \lambda_j \frac{\partial \bar{J}}{\partial d_j}(x).$$

Since all λ_j are nonnegative there must be one descent direction d_j satisfying

$$\lim_{h \searrow 0} \frac{\bar{J}(x + h d_j) - \bar{J}(x)}{h} = \frac{\partial \bar{J}}{\partial d_j}(x) < 0$$

and hence $\bar{J}(x + h d_j) < \bar{J}(x) = J(x)$ for small enough $h > 0$. In order to show $J(x + h d_j) < J(x)$ note that by Lemma 1 we have

$$x + h d_j \in \mathcal{C}_i \subset \overline{\operatorname{dom} J}$$

for small enough $h > 0$. Now Lemma 3 implies $x + h d_j \in \operatorname{dom} J$ for small enough $h > 0$ (otherwise we would have $x \in \overline{\operatorname{dom} J} \setminus \operatorname{dom} J$). ☐

Lemma 6. *Assume that $J : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ has the properties (A1)–(A4), that the set of search directions $\mathcal{E} = \{v_1, \dots, v_m\} \subset \mathbb{R}^n$ satisfies (A5), and let $x \in \operatorname{dom} J$ such that*

$$\alpha(x, v_i) = 0 \qquad \forall i = 1, \dots, m.$$

Then x is the unique minimizer x^* of J.

Proof. If $x \neq x^*$ assumption (A5) provides $J(x + \epsilon v_i) < J(x)$ for some $\epsilon > 0$ and $v_i \in \mathcal{E}^\pm$ and thus $\alpha(x, v_i) \neq 0$. ☐

3.3 Convergence

We are now ready to prove the convergence of the Polyhedral Gauß–Seidel Algorithm. Since in practical applications the exact minimizers in the search directions are in general not easily available we prove convergence for an inexact version that only requires to have at least a fixed portion of the exact correction. Convergence of the exact PGS method is then a special case of this result.

Additionally to the prerequisites of the PGS algorithm let $\epsilon \in [0,1]$ be arbitrary but fixed. Then using $\alpha(\cdot,\cdot)$ as defined in Corollary 3 one iteration of inexact PGS(ϵ) consists of the following steps:

1. set $x_0^\nu = x^\nu$
2. for $i = 1, \dots, m$ select
$$x_i^\nu \in x_{i-1}^\nu + [\epsilon, 1]\alpha(x_{i-1}^\nu, v_i)v_i$$
3. set $x^{\nu+1} = x_m^\nu$.

Theorem 4. *Assume that $J : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ has the properties (A1)–(A4) and that the set of search directions $\mathcal{E} = \{v_1, \dots, v_m\} \subset \mathbb{R}^n$ satisfies (A5) or (A6). Furthermore let $\epsilon \in (0,1]$ be arbitrary but fixed. Then for all x^0 with $J(x^0) < \infty$ the sequence generated by PGS(ϵ) converges to the unique minimizer x^* of J.*

Proof. By definition the iterates satisfy $J(x^{\nu+1}) \leq J(x^\nu)$. Since J is coercive this implies that x^ν is bounded and has a convergent subsequence.

Now consider any such convergent subsequence $x^{\nu_k} \to x_0$. Then by monotonicity and continuity of J we have $x_0 \in \operatorname{dom} J$. Note that by convexity $J(x + t\alpha(x,v)v)$ is monotonically decreasing for all x, v and for $t \nearrow 1$. For

$$M_i(x) := x + \epsilon\alpha(x, v_i)v_i$$

this implies

$$J(x^{\nu_{k+1}}) \leq J(x_i^{\nu_k}) \leq J(M_i(x_{i-1}^{\nu_k})) \leq J(x_{i-1}^{\nu_k}) \leq J(x^{\nu_k}) \tag{11}$$

and guarantees that all intermediate iterates also stay in $\operatorname{dom} J$.

Obviously we have $x_0^{\nu_k} = x^{\nu_k} \to x_0$. Now assume that $x_{i-1}^{\nu_k} \to x_0$ holds true for some $i \in \{1, \dots, m\}$. Then taking the limit in (11) gives

$$J(M_i(x_{i-1}^{\nu_k})) \to J(x_0)$$

while continuity of M_i (following from Corollary 3) gives

$$J(M_i(x_{i-1}^{\nu_k})) \to J(M_i(x_0)).$$

Hence we have $J(x_0) = J(M_i(x_0))$ which, by strict convexity, implies $M_i(x_0) = x_0$ and $\alpha(x_0, v_i) = 0$. Thus

$$x_{i-1}^{\nu_k} + \rho_i^{\nu_k}\alpha(x_{i-1}^{\nu_k}, v_i)v_i \to x_0$$

holds true for any bounded sequence $\rho_i^{\nu_k}$ which in particular provides $x_i^{\nu_k} \to x_0$.

By induction we have especially shown that $\alpha(x_0, v_i) = 0$ for all $i = \{1, \ldots, m\}$. Thus Lemma 6 and Lemma 5 provide $x_0 = x^*$. Since this is true for any convergent subsequence and there is at least one convergent subsequence we have shown $x^\nu \to x^*$. □

Our main result Theorem 1 directly follows from Theorem 4 with $\epsilon = 1$.

Remark 1. *Note that the result still holds if Step 3 is replaced by*

3'. select $x^{\nu+1}$ such that $J(x^{\nu+1}) \leq J(x_m^\nu)$.

Hence an additional minimization step can be inserted after each Gauß–Seidel sweep. This is important for the convergence of multigrid algorithms, where this minimization step is the coarse grid correction.

4 Construction of Search Directions

In the previous section we have proved that the PGS algorithm converges if any tangent cone $T_{\mathcal{C}_i}(x)$ of the partition $\operatorname{dom} J = \mathcal{C}_1 \cup \cdots \cup \mathcal{C}_N$ for $x \in \operatorname{dom} J$ is generated by a subset of the set $\mathcal{E}$ of search directions and if $\mathcal{E}$ contains descent directions for all points $x \in \operatorname{dom} J \setminus \mathcal{S}$ where J is not piecewise smooth. While the construction of descent directions for $x \in \operatorname{dom} J \setminus \mathcal{S}$ is in general dependent on the values and derivatives of J we will show in this section how to construct sets $\mathcal{E}$ generating the tangent cones in a systematic way.

In the following we will restrict our attention to a single polyhedron $\mathcal{C}$ and construct a set of search directions $\mathcal{E}_\mathcal{C}$ for $\mathcal{C}$. The set $\mathcal{E}$ of search directions for an entire polyhedral decomposition can then be constructed by taking the union of all sets of search directions $\mathcal{E}_{\mathcal{C}_i}$ for $\mathcal{C}_i$. Note that in many cases this union will eliminate many redundant directions. This is important because the time-complexity of a single Gauß–Seidel iteration is $O(|\mathcal{E}|)$. One particular noteworthy case is when the $\mathcal{C}_i$ form the cells of a hyperplane arrangement (cf. Section 5.3).

4.1 Spanning Sets for Polyhedra

For any polyhedron $\mathcal{C}$ the first part of Theorem 2 provides a decomposition of $\mathcal{C}$ into points and directions. If such a decomposition can be found in practice, then Theorem 3 can be evoked to construct search directions from a finite number of tangent cones.

Lemma 7. *Let*

$$\mathcal{C} = \left\{ \sum_{i=1}^{k} \lambda_i p_i + \sum_{j=1}^{l} \mu_j d_j \;\middle|\; \lambda_i \geq 0,\ \sum_{i=1}^{k} \lambda_i = 1,\ \mu_j \geq 0 \right\}$$

be a representation of $\mathcal{C}$ *with points* $p_1, \dots, p_k \in \mathbb{R}^n$ *and directions* $d_1, \dots, d_l \in \mathbb{R}^n$*, and for each* $i = 1, \dots, k$ *let* $\mathcal{E}_{T_\mathcal{C}(p_i)}$ *be a finite set such that* $T_\mathcal{C}(p_i) = \operatorname{cone} \mathcal{E}_{T_\mathcal{C}(p_i)}$*. Then for all* $x \in \mathcal{C}$ *there is a subset* $\mathcal{E}_x \subset \mathcal{E}$ *of*

$$\mathcal{E} := \bigcup_{i=1}^{k} \mathcal{E}_{T_\mathcal{C}(p_i)} \cup \bigcup_{j=1}^{l} \{-d_j\}$$

such that $T_\mathcal{C}(x) = \operatorname{cone} \mathcal{E}_x$*.*

Proof. Let $x \in \mathcal{C}$ be given by (4). Then Theorem 3 proves the assertion since

$$\begin{aligned} T_\mathcal{C}(x) &= \sum_{j=1}^{k} \lambda_j T_\mathcal{C}(p_j) + \sum_{j=1}^{l} \mu_j \operatorname{cone}\{-d_j\} \\ &= \operatorname{cone}\Bigg(\underbrace{\bigcup_{i:\lambda_i>0} \mathcal{E}_{T_\mathcal{C}(p_i)} \cup \{-d_j \mid \mu_j > 0\}}_{=:\mathcal{E}_x}\Bigg). \end{aligned}$$

□

In order to give an explicit construction for the search directions let

$$\operatorname{lineal}(\mathcal{C}) := \{y \in \mathbb{R}^n \mid x + ty \in \mathcal{C} \quad \text{for all } x \in \mathcal{C}, t \in \mathbb{R}\}$$

be the lineality space of $\mathcal{C}$. Picking a complementary subspace U to $\operatorname{lineal}(\mathcal{C})$ we can decompose $\mathcal{C}$ as the Minkowski sum

$$\mathcal{C} = \operatorname{lineal}(\mathcal{C}) + (\mathcal{C} \cap U), \tag{12}$$

where $\mathcal{C} \cap U$ is pointed, i.e., it has lineality space $\{0\}$ [11].

Lemma 8. *Let* $\mathcal{C}$ *be a polyhedron,* $\mathcal{E}_{\operatorname{lineal}(\mathcal{C})}$ *a basis of* $\operatorname{lineal}(\mathcal{C})$*, and* $\mathcal{E}_{\mathcal{C}\cap U}$ *the set of edge vectors of* $\mathcal{C} \cap U$ *for a subspace* U *complementary to* $\operatorname{lineal}(\mathcal{C})$*. Then for all* $x \in \mathcal{C}$ *there is a subset* $\mathcal{E}_x \subset (\mathcal{E}_{\operatorname{lineal}(\mathcal{C})} \cup \mathcal{E}_{\mathcal{C}\cap U})^{\pm}$ *such that* $T_\mathcal{C}(x) = \operatorname{cone} \mathcal{E}_x$*.*

Proof. Since $\mathcal{C} \cap U$ is pointed it has at least one vertex (see [9]). Thus the second part of Theorem 2 gives a representation of $\mathcal{C} \cap U$ where $p_1, \dots, p_k$ are the vertices of $\mathcal{C} \cap U$. Then each $T_{\mathcal{C}\cap U}(p_j)$ is a cone with a vertex and by Corollary 1 generated by the edge vectors adjacent to p_j with proper orientation. Hence Lemma 7 implies that for each $z \in \mathcal{C} \cap U$ there is $\mathcal{E}_z \subset \mathcal{E}_{\mathcal{C}\cap U}$ with $T_{\mathcal{C}\cap U}(z) = \operatorname{cone} \mathcal{E}_z$.

Now let $x \in \mathcal{C}$. Then there are $y \in \operatorname{lineal}(\mathcal{C})$ and $z \in \mathcal{C} \cap U$ with $x = y + z$. For this representation Exercise 6.44 in [8] implies

$$T_\mathcal{C}(x) = \overline{T_{\operatorname{lineal}(\mathcal{C})}(y) + T_{\mathcal{C}\cap U}(z)}$$

and hence for some $\mathcal{E}_z \subset \mathcal{E}_{\mathcal{C}\cap U}$

$$\begin{aligned} T_\mathcal{C}(x) &= \overline{\operatorname{lineal}(\mathcal{C}) + \operatorname{cone} \mathcal{E}_z} \\ &= \overline{\operatorname{cone}\left(\mathcal{E}^{\pm}_{\operatorname{lineal}(\mathcal{C})} \cup \mathcal{E}_z\right)} = \operatorname{cone}\left(\mathcal{E}^{\pm}_{\operatorname{lineal}(\mathcal{C})} \cup \mathcal{E}_z\right). \end{aligned}$$

□

For the special case that $\mathcal{C}$ is a polyhedron with at least one vertex we have $\operatorname{lineal}(\mathcal{C}) = \{0\}$ and $U = \mathbb{R}^n$. Then Lemma 8 reduces to:

Corollary 4. *Let $\mathcal{C}$ be a polyhedron with at least one vertex and $\mathcal{E}$ be the set of edge vectors of $\mathcal{C}$. Then for all $x \in \mathcal{C}$ there is a subset $\mathcal{E}_x \subset \mathcal{E}^{\pm}$ such that $T_{\mathcal{C}}(x) = \operatorname{cone} \mathcal{E}_x$.*

Corollary 4 covers the main assumption on search directions in (A6). If $\mathcal{S} = \operatorname{dom} J$ it implies that the set of edge vectors satisfies (A6). Otherwise further directions for $x \notin \mathcal{S}$ mav be necessary.

4.2 Energies with a Block Structure

Many problems obtained by discretized partial differential equations exhibit a block structure where the nonsmooth part of the energy decouples into local blocks that are only coupled globally by a smooth energy.

In the following we assume $n = n_1 + \cdots + n_k$ and identify $\mathbb{R}^n$ with $\mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_k}$ as well as $(x_1, \dots, x_n) \in \mathbb{R}^n$ with $(x_{(1)}, \dots, x_{(k)}) \in \Pi_{i=1}^k \mathbb{R}^{n_i}$, where $x_{(i)} \in \mathbb{R}^{n_i}$ denotes the i-th block of $x \in \mathbb{R}^n$. Introducing the restriction operator

$$R_i : \mathbb{R}^n \to \mathbb{R}^{n_i}, \qquad R_i x := x_{(i)}$$

we find that $R_i^T : \mathbb{R}^{n_i} \to \mathbb{R}^n$ is the extension of vectors in $\mathbb{R}^{n_i}$ to $\mathbb{R}^n$ by zero.

We consider functionals $J : \mathbb{R}^n = \Pi_{i=1}^k \mathbb{R}^{n_i} \to \mathbb{R} \cup \{\infty\}$ of the form

$$J(x) = J_0(x) + \sum_{i=1}^{k} J_i(x_{(i)}).$$

The topic of this section is the construction of search directions v for J from search directions $v_{(i)}$ for the J_i. For $i > 0$ a descent direction $v_{(i)}$ of J_i satisfies $\partial J_i / \partial v_{(i)} < 0$. However this does in general not imply $\partial J / \partial v < 0$ for $v = R_i^T v_{(i)}$. In order to exclude this problem we introduce the following general assumption which is stronger then (A6). Lemma 9 then shows that if the individual block functionals J_i satisfy (A6'), so does the entire functional J.

(A6') J and $\mathcal{E}$ satisfy (A6) and for all $x \in \operatorname{dom} J \setminus (\mathcal{S} \cup \{x^*\})$ there is a direction $v \in \mathcal{E}^{\pm}$ such that

$$\lim_{h \searrow 0} \frac{J(x + hv) - J(x)}{h} = -\infty.$$

Lemma 9. *Assume that $J : \mathbb{R}^n = \Pi_{i=1}^k \mathbb{R}^{n_i} \to \mathbb{R} \cup \{\infty\}$ has the properties (A1)–(A4) and that it is given by*

$$J(x) = J_0(x) + \sum_{i=1}^{k} J_i(x_{(i)})$$

for a continuous differentiable convex function $J_0 : \mathbb{R}^n \to \mathbb{R}$ and convex functions $J_i : \mathbb{R}^{n_i} \to \mathbb{R} \cup \{\infty\}$. Let the search directions $\mathcal{E}_i \subset \mathbb{R}^{n_i}$ for J_i satisfy (A6') for all $i = 1, \dots, k$. Then (A6') is also satisfied by the search directions $\mathcal{E}$ for J given by

$$\mathcal{E} := \bigcup_{i=1}^{k} R_i^T \mathcal{E}_i.$$

Proof. First we note that $\operatorname{dom} J = \Pi_{i=1}^k (\operatorname{dom} J_i)$. A partitioning of $\overline{\operatorname{dom} J}$ in the sense of (A6) is given by

$$\{\mathcal{C}_1, \dots, \mathcal{C}_N\} = \{\mathcal{C} \subset \mathbb{R}^n \mid \mathcal{C} = \Pi_{i=1}^k \mathcal{C}_{(i),j_i} \text{ for } 1 \le j_i \le N_i\}$$

where $\mathcal{C}_{(i),1}, \dots, \mathcal{C}_{(i),N_i}$ is the partitioning of $\overline{\operatorname{dom} J_i}$ from (A6). Analogously define $\mathcal{S} = \Pi_{i=1}^k \mathcal{S}_i$ using the smoothness domains $\mathcal{S}_i$ of J_i.

Let $x \in \mathcal{C}_j \cap \mathcal{S}$ for some $j \in \{1, \dots, N\}$. Then for each $i \in \{1, \dots, k\}$ we have $x_{(i)} \in \mathcal{C}_{(i),j_i} \cap \mathcal{S}_i$ for some $j_i \in \{1, \dots, N_i\}$. Hence there are $\epsilon_i > 0$ and smooth extensions $\bar{J}_i \in C^1(\overline{B_{\epsilon_i}(x_{(i)})})$ such that

$$\bar{J}(x) := J_0(x) + \sum_{i=1}^{k} \bar{J}_i(x_{(i)})$$

is continuously differentiable on $\Pi_{i=1}^k \overline{B_{\epsilon_i}(x_{(i)})}$ which contains $\overline{B_\epsilon(x)}$ for some $\epsilon > 0$. By Theorem 6.9, Corollary 6.29, and Proposition 6.41 in [8]; and (A6) for J_i we have

$$\begin{aligned} T_{\mathcal{C}_j}(x) &= \Pi_{i=1}^k T_{\mathcal{C}_{(i),j_i}}(x_{(i)}) \\ &= \Pi_{i=1}^k \operatorname{cone} \mathcal{E}_{x_{(i)},j_i} = \operatorname{cone} \underbrace{\bigcup_{i=1}^{k} R_i^T \mathcal{E}_{x_{(i)},j_i}}_{=:\mathcal{E}_{x,j} \subset \mathcal{E}}. \end{aligned}$$

Now let $x \in \operatorname{dom} J \setminus (\mathcal{S} \cup \{x^*\})$. Then there is at least one i such that $x_{(i)} \notin \mathcal{S}_i$. Hence there is $v \in \mathcal{E}_i$ with

$$\lim_{h \searrow 0} \frac{J(x + hR_i^T v) - J(x)}{h} = \frac{\partial J_0(x)}{\partial R_i^T v} + \lim_{h \searrow 0} \frac{J_i(x_{(i)} + hv) - J_i(x_{(i)})}{h} = -\infty.$$

□

Lemma 9 allows to construct search directions $\mathcal{E}$ for the PGS method for J from local directions for the separate blocks. If directions in $R_i^T \mathcal{E}_i$ are processed before directions in $R_j^T \mathcal{E}_j$, $i < j$, one global PGS step is equivalent to subsequent local PGS steps for all blocks of J keeping all other blocks fixed at a time.

5 Applications

In this last section we give three example applications that all lead to minimization problems of the type considered in this article. All three of them arise as discretizations of certain partial differential equations. We do not present actual numerical experiments, because Gauß–Seidel methods for PDE problems are known to converge only slowly anyways. We see their purpose primarily as being the basis for fast nonlinear multigrid methods for polyhedral minimization problems, which we will cover in a subsequent paper.

5.1 Bound Constraints

Let $A \in \mathbb{R}^{n\times n}$ by symmetric and positive definite, $f \in \mathbb{R}^n$ and consider the minimization problem (1) for the functional

$$J(x) = \frac{1}{2}\langle Ax, x\rangle - \langle f, x\rangle + \chi_K(x) \qquad \text{with } K = \prod_{i=1}^{n}[a_i, b_i], \tag{13}$$

where $a_i, b_i \in \mathbb{R}$ for all $1 \leq i \leq n$.

Remark. *This problem arises for example from the discretization of an obstacle problem obtained if the Laplace equation $-\Delta u = f$ is complemented by inequality constraints*

$$\underline{\psi}(x) \leq u(x) \leq \overline{\psi}(x) \quad a.e.$$

for given continuous functions $\underline{\psi}, \overline{\psi} : \Omega \to \mathbb{R}$, $\underline{\psi} \leq \overline{\psi}$ pointwise. Using the scalar projection $P_{[a,b]} : \mathbb{R} \to [a, b]$ such an obstacle problem can be written as

$$P_{[\underline{\psi}-u,\overline{\psi}-u]}\big(-\Delta u - f\big) = 0 \quad \text{on } \Omega, \qquad u|_{\partial\Omega} = 0.$$

The functional J is C^∞ on $\operatorname{dom} J = K$, strictly convex and coercive—in other words it fulfills (A1)–(A3). The admissible set K is a polyhedron, because it can be described as the intersection of the half-spaces

$$K = \big\{x \in \mathbb{R}^n \;\big|\; \langle e_i, x\rangle \geq a_i, \langle -e_i, x\rangle \geq -b_i,\ i = 1, \ldots, n\big\},$$

the e_i being the Euclidean coordinate directions. Hence (A4) is also fulfilled. By Theorems 1 and Corollary 4 the Polyhedral Gauß–Seidel method converges for the functional (13) if each edge of K is parallel to a search direction. The edges of K are the closed segments

$$\big[(s_1, \ldots, s_{i-1}, a_i, s_{i+1}, \ldots, s_n), (s_1, \ldots, s_{i-1}, b_i, s_{i+1}, \ldots, s_n)\big] \subset \mathbb{R}^n,$$

for $i = 1, \ldots, n$ and $s_j \in \{a_j, b_j\}$. These edges are parallel to the Euclidean coordinate vectors e_i. Hence we recover the well-known result that the Gauß–Seidel method converges for a convex quadratic functional with bound constraints only if the search directions include the coordinate vectors [5].

5.2 Vector-Valued Allen–Cahn Equations with a Logarithmic Potential

In this section we consider the vector-valued Allen–Cahn equation [2, 6] as a model for phase transitions in a mixture of p phases. Let

$$G = \Big\{\lambda \in \mathbb{R}^p \;\big|\; \lambda_i \geq 0, i = 1, \dots, p, \; \sum_{i=1}^{p} \lambda_i = 1\Big\}$$

be the $p-1$-dimensional Gibbs simplex. It is a $p-1$-dimensional bounded polyhedron embedded in $\mathbb{R}^p$. For a domain $\Omega \subset \mathbb{R}^d$ we are looking for a time-dependent field of order parameters

$$u(\cdot, \cdot) \; : \; \Omega \times [0, T) \to G$$

solving

$$\epsilon u_t = \epsilon \Delta u - \frac{1}{\epsilon} P(\nabla \Psi_\theta)(u), \qquad u(0) = u_0, \tag{14}$$

and subject to natural boundary conditions. In (14), ϵ is a parameter, P is the orthogonal projection onto $G_0 := \{v \in \mathbb{R}^p \mid \sum_i v_i = 0\}$, and the potential $\Psi_\theta : G \to \mathbb{R}$ is given by

$$\Psi_\theta(u) = \Phi_\theta(u) + \theta_c \frac{N}{2} \sum_{i=1}^{p} u_i (Cu)_i, \qquad \Phi_\theta(u) = \theta \sum_{i=1}^{p} u_i \ln(u_i),$$

where $\theta > 0$ is the temperature, $\theta_c > 0$ is a critical temperature, and C is a symmetric matrix. The potential Φ_θ and thus Ψ_θ can be naturally extended to the limiting case $\theta = 0$, which is the obstacle potential

$$\Phi_0(u) = \chi_G(u).$$

For $\theta = 0$ the derivative $\nabla \Psi_\theta$ in (14) becomes the subdifferential $\partial \Psi_\theta$ and the equation (14) becomes an inclusion.

We discretize (14) in space using finite elements and in time using the semi-implicit scheme also used by Kornhuber and Krause [7] and obtain discrete problems

$$u^k \in S_G^h \quad a(u^k, v - u^k) + \frac{\tau}{\epsilon^2} (\Phi_\theta(v), 1)_h - \frac{\tau}{\epsilon^2} (\Phi_\theta(u^k), 1)_h \geq l^k(v - u^k) \qquad \forall v \in S_G^h$$

in the constraint linear finite element space

$$S_G^h := \{v \in (S^h)^p \mid v(x) \in G \; \forall x \in \Omega\},$$

with the $(H^1(\Omega))^p$-elliptic bilinear form

$$a(v, w) := (v, w)_h + \tau (\nabla v, \nabla w),$$

$(\cdot,\cdot)_h$ the lumped scalar product, and

$$l^k(v) := \big((I-\theta_c N\frac{\tau}{\epsilon^2}C)u^{k-1}, v\big)_h.$$

For each time k this is equivalent to a minimization problem for the functional

$$\begin{aligned} J &: (\mathbb{R}^p)^n \to \mathbb{R}\cup\{\infty\} \\ J(x) &= \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle + \sum_{i=1}^n \big(\omega_i \Phi_\theta(x_{(i)}) + \chi_G(x_{(i)})\big). \end{aligned} \tag{15}$$

Here x is the coefficient vector of u^k for the nodal basis $\lambda_1,\dots,\lambda_n$ of S^h, $x_{(i)} = R_i x \in \mathbb{R}^p$ is the i-th block of x as introduced in Section 4, and ω_i is the weight $\omega_i = (\lambda_i, 1)_h$.

Lemma 10. *The functional J fulfills the assumptions (A1)–(A4).*

Proof. The quadratic part of J is obviously strictly convex, coercive, and continuous. Noting that $\psi(t) := t\ln(t)$, $\psi(0) := 0$ has the properties

$$\lim_{t\searrow 0}\psi(t) = 0, \qquad \psi''(t) = t^{-1} > 0 \quad \forall t\in(0,1),$$

we find that ψ is continuous and convex on $[0,1]$ and thus the nonlinear part of J is continuous and convex on the polyhedron $\operatorname{dom} J = (G)^n$. ☐

Lemma 11. *Let $\mathcal{E}_G$ be the set of edge vectors of G and $\theta \geq 0$. Then for each $i = 1,\dots,n$ the set $\mathcal{E}_G$ satisfies (A6') for $J_i : \mathbb{R}^p \to \mathbb{R}\cup\{\infty\}$ with*

$$J_i(y) = \omega_i \Phi_\theta(y) + \chi_G(y).$$

Proof. First we note that since Φ_θ is finite on G we have $\operatorname{dom} J = \overline{\operatorname{dom} J} = G$ and the trivial partitioning $N = 1$, $\mathcal{C}_1 = G$. Then for all $y \in G$ Corollary 4 guarantees the existence of a subset $\mathcal{E}_y \subset \mathcal{E}_G^{\pm}$ with $T_{\mathcal{C}_1}(y) = \operatorname{cone}\mathcal{E}_y$.

Now we examine the smoothness of J_i on the partitioning. For $\theta = 0$ we have $J_i \equiv 0$ on G. Hence we can select $\mathcal{S} = G$ and extend J_i trivially to a smooth function on any $\overline{B_\epsilon(y)}$ for $y \in \mathcal{S}$ and $\epsilon > 0$. Since $\operatorname{dom} J \setminus \mathcal{S} = \emptyset$ we have shown (A6') for $\theta = 0$.

For $\theta > 0$ we select the relative interior $\mathcal{S} := G \cap (\mathbb{R}^+)^p$. Then we have $J_i = \omega_i\Phi_\theta$ on $\mathcal{S}$. Hence for any $y \in \mathcal{S}$ a smooth extension of J_i to $\overline{B_\epsilon(y)}$ is given by $\omega_i\Phi_\theta$ if $\epsilon > 0$ is small enough. Now let $y \in G\setminus\mathcal{S}$. Then there is an index j_1 with $y_{j_1} = 0$ and an index j_2 with $y_{j_2} > 0$. Consider the edge $v = e_{j_1} - e_{j_2} \in \mathcal{E}_G^{\pm}$ from vertex j_2 to vertex j_1 of G. The directional derivative of J_i at y in the direction of v is

$$\lim_{h\searrow 0}\frac{J_i(y+hv) - J_i(u)}{h} = \theta\lim_{h\searrow 0}\ln(h) - \theta(1+\ln(y_{j_2})) = -\infty.$$

Thus we have shown (A6') for J_i. ☐

Now Lemma 10 and Lemma 11 together with Lemma 9 allow to invoke Theorem 1 and to obtain the following convergence result.

Corollary 5. *Let $\mathcal{E}_G$ be the set of edge vectors of G. Then for all $\theta \geq 0$ the polyhedral Gauß–Seidel method with $\mathcal{E} = \bigcup_{i=1}^n R_i^T \mathcal{E}_G$ converges to the unique minimum of the Allen–Cahn functional* (15).

Hence Theorem 4.1 from Kornhuber and Krause [7] turns out to be a special case of our convergence theory.

5.3 Discontinuous Galerkin Methods

Let Ω be a domain in $\mathbb{R}^d$ and $H^1(\Omega)$ the space of scalar first-order Sobolev functions on Ω. We consider a minimization problem for the functional

$$\mathcal{J}(v) = \frac{1}{2}a(v,v) - b(v) + \int_\Omega \Phi(v(\xi))\, d\xi$$

in $H^1(\Omega)$ and subject to suitable boundary conditions. Here, $a(\cdot,\cdot)$ is a symmetric, coercive, and continuous bilinear form, $b(\cdot)$ is a linear form, and $\Phi : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ is convex, coercive, lower semi-continuous, proper, and once continuously differentiable everywhere in its domain except at a finite number of points $\kappa_1, \dots, \kappa_m$. If, additionally, Φ satisfies certain growth conditions, it follows that the overall functional $\mathcal{J}$ is strictly convex, coercive, and lower semi-continuous [6], and hence there exists a unique minimizer.

We discretize the problem with a Discontinuous Galerkin (DG) method. For this we introduce a grid $\mathcal{G}$ on Ω consisting of n elements and define the finite element space

$$V_{\mathrm{DG}}^r = \Big\{ v_h \in L^2(\Omega) \;\Big|\; v_h|_T \text{ is a polynomial of order } r \text{ for all elements } T \in \mathcal{G} \Big\}.$$

Since we want to obtain a discrete convex minimization problem we have to use a DG method that preserves the symmetry of $a(\cdot,\cdot)$. Discretization of the forms $a(\cdot,\cdot)$ and $b(\cdot)$ by such DG methods is well known from the literature [1]. Here we concentrate on the nonlinear term

$$\phi(v) := \int_\Omega \Phi(v(\xi))\, d\xi.$$

Let

$$\Lambda := \big\{ \lambda_{(j),i} \in V_{\mathrm{DG}}^r \;\big|\; \lambda_{(j),i} = \tilde{\lambda}_i \circ F_j \big\}$$

be a basis of V_{DG}^r, where the $\tilde{\lambda}_i$, $i = 1, \dots, p$ are a set of shape functions on the reference element and F_j, $j = 1, \dots, n$ is the affine mapping from element j onto the reference element. We obtain an algebraic functional

$$\begin{aligned} J \quad &: \quad \mathbb{R}^{pn} \to \mathbb{R} \cup \{\infty\} \\ J(x) \quad &= \quad \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle + \mathcal{I}_\Omega\, \Phi\Big(\sum_{j=1}^n \sum_{i=1}^p x_{(j),i} \lambda_{(j),i}(\xi) \Big)\, d\xi, \end{aligned}$$

where A is a symmetric positive-definite matrix, $x_{(j),i}$ is the coefficient corresponding to basis function $\lambda_{(j),i}$, and $\mathcal{I}_\Omega \dots d\xi$ is an approximation of $\int_\Omega \dots d\xi$ by numerical quadrature.

Since the support of each basis function from Λ is restricted to a single element we can write J as a sum

$$J(x) = J_0(x) + \sum_{j=1}^{n} J_j(x_{(j)})$$

with $J_0 := \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle$ and

$$J_j(x_{(j)}) := \mathcal{I}_{T_j} \Phi\Big(\sum_{i=1}^{p} x_{(j),i}\lambda_{(j),i}(\xi)\Big)\, d\xi.$$

Choosing a quadrature rule consisting of K points q_k in the reference element and corresponding weights w_k we get

$$J_j(x_{(j)}) = \sum_{k=1}^{K} w_k |\det F_j^{-1}(q_k)| \Phi\Big(\sum_{i=1}^{p} x_{(j),i}\lambda_{(j),i}(F_j^{-1}(q_k))\Big).$$

We simplify the notation by defining $\hat{w}_{j,k} := w_k|\det F_j^{-1}(q_k)|$ and introducing $B \in \mathbb{R}^{K\times p}$ with rows B_k and $B_{k,i} = \tilde{\lambda}_i(q_k) = \lambda_{(j),i}(F_j^{-1}(q_k))$ such that

$$J_j(x_{(j)}) = \sum_{k=1}^{K} \hat{w}_{j,k}\Phi(B_k x_{(j)}).$$

Minimization problems for the discrete functional J can now be treated by the Polyhedral Gauß–Seidel method.

Lemma 12. *The DG functional J has the properties (A1)–(A4).*

Proof. Since A is symmetric and positive definite the functional J_0 has the properties (A1)–(A4). For each $j = 1, \dots, n$, J_j is convex and hence $J = J_0 + \sum_{i=1}^{n} J_j$ is strictly convex and coercive. This shows (A1) and (A2).

To see (A3) and (A4), we define for every quadrature point q_k and every element T_j the function

$$\Phi_{j,k} : \mathbb{R}^{pn} \to \mathbb{R} \cup \{\infty\}, \qquad \Phi_{j,k}(x) := \Phi(B_k R_j x)$$

using the restriction operator R_j from Section 4.2. Since Φ is lower semi-continuous it must also be continuous on its (one-dimensional) domain. Hence the same it true for $\Phi_{j,k}$ and the weighted sum J_j. The set $\operatorname{dom} \Phi_{j,k}$ is a polyhedron and so is

$$\operatorname{dom} J = \bigcap_{j=1}^{n} \bigcap_{k=1}^{K} \operatorname{dom} \Phi_{j,k}.$$

□

It remains to be shown that search directions having the property (A6) can be found. We again concentrate on a single element. The corresponding property for the entire problem can then be established by Lemma 9.

Consider an element j. The set of all $y \in \mathbb{R}^p$ where $\Phi_k(\cdot) := \Phi(B_k \cdot)$ is not continuously differentiable forms a finite set of parallel hyperplanes given by

$$\mathcal{H}_{k,l} := \{y \in \mathbb{R}^p \mid \langle B_k, y\rangle = \kappa_l\}, \qquad l = 1, \dots, m,$$

where we now write $\langle B_k, y\rangle$ instead of $B_k y$ to be consistent with the notation introduced in Section 2. As a sum of functions is differentiable if all addends are, the regions of differentiability of J_j are the cells of a hyperplane arrangement $\mathcal{A}_j$ in $\mathbb{R}^p$, generated by the $\mathcal{H}_{k,l}$. These cells form a decomposition of $\operatorname{dom} J_j$ into convex polyhedra. From the smoothness of Φ it follows that J_j will be smooth on each cell of the arrangement and piecewise smooth on the lower-dimensional faces. Hence the arrangement $\mathcal{A}_j$ is a decomposition of $\operatorname{dom} J_j$ as mandated by (A6).

Finally, to obtain a convergent polyhedral Gauß–Seidel method we need a set of search directions $\mathcal{E}$ such that subsets of $\mathcal{E}^{\pm}$ span all tangent cones of all polyhedra $\mathcal{C}_i \in \mathcal{A}$ at all $x \in \mathcal{C}_i$. To compute such directions from the B_k, $k = 1, \dots, K$ and κ_l, $l = 1, \dots, m$, note that it is possible to compute the entire arrangement structure in $O((mK)^p)$ time using an algorithm of Edelsbrunner et al. [3]. This may seem like a lot; it is, however, optimal. Also the arrangement structure depends only on the element type, the DG basis, the quadrature rule and the non-differentiability of Φ. Hence for a given Φ and a given global DG space it can be computed once and for all. On the other hand, in most cases the entire arrangement structure is more information than necessary, because many faces in an arrangement are parallel and hence spanned by the same directions. More efficient algorithms may possibly be found exploiting the special structure of the problem. We leave this task for a future publication and close with an example showing the easiest possible case.

Example 4. *Let $\tilde{\lambda}_i$, $i = 1, 2, 3$ be the first-order Lagrangian shape functions on a triangle, the quadrature rule consisting of a single point $(\frac{1}{3}, \frac{1}{3})$ and $\Phi : \mathbb{R} \to \mathbb{R}$ continuous everywhere and C^∞ everywhere except at $\kappa_1 = 1$. Then $J_j : \mathbb{R}^3 \mapsto \mathbb{R}$ and $\operatorname{dom} J_j = \mathbb{R}^3$. The arrangement $\mathcal{A}$ consists of the single hyperplane*

$$\langle B_1, x\rangle = \kappa_1 = 1, \qquad B_1 = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$$

and the two open half-spaces $\langle B_1, x\rangle < 1$ and $\langle B_1, x\rangle > 1$. Define the two closed polyhedra $\mathcal{C}_1 := \{x \in \mathbb{R}^3 \mid \langle B_1, x\rangle \leq 1\}$ and $\mathcal{C}_2 := \{x \in \mathbb{R}^3 \mid \langle B_1, x\rangle \geq 1\}$. There are four tangent cones, namely $T_{\mathcal{C}_1}(x), T_{\mathcal{C}_2}(x)$ with $\langle B_1, x\rangle = 1$; $T_{\mathcal{C}_1}(x)$ for $\langle B_1, x\rangle < 1$, and $T_{\mathcal{C}_2}(x)$ for $\langle B_1, x\rangle > 1$. The ones for $\mathcal{C}_1$ are spanned by

$$v_1 = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right), \qquad v_2 = (1, -1, 0), \qquad v_3 = (0, 1, -1),$$

and the ones for $\mathcal{C}_2$ by $-v_1$, v_2, and v_3. Hence $\mathcal{E} = \{v_1, v_2, v_3\}$ is a suitable set of search directions.

References

[1] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39(5):1749–1779, 2002.

[2] L. Bronsard and F. Reitich. On three-phase boundary motion and the singular limit of a vector-valued Ginzburg-Landau equation. *Arch. Rational Mech. Anal.*, 124:355–379, 1993.

[3] H. Edelsbrunner, J. O'Rourke, and R. Seidel. Constructing arrangements of lines and hyperplanes with applications. *SIAM J. Comput.*, 15(2):341–363, 1986.

[4] I. Ekeland and R. Temam. *Convex Analysis*. North-Holland, 1976.

[5] R. Glowinski, J. L. Lions, and R. Trémolières. *Numerical Analysis of Variational Inequalities*. North-Holland, 1981.

[6] R. Kornhuber. *Adaptive Monotone Multigrid Methods for Nonlinear Variational Problems*. Teubner, Stuttgart, 1. edition, 1997.

[7] R. Kornhuber and R. Krause. Robust multigrid methods for vector-valued Allen–Cahn equations with logarithmic free energy. *Comp. Visual. Sci.*, 9: 103–116, 2006.

[8] R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis*. Springer Verlag, 1998.

[9] A. Schrijver. *Theory of Linear and Integer Programming*. John Wiley & Sons, Chichester, 1986.

[10] P. Tseng. Decomposition algorithm for convex differentiable minimization. *J. Optim. Theory Appl.*, 70(1):109–135, 1991.

[11] G. M. Ziegler. *Lectures on Polytopes*. Springer Verlag, 1994.