TECHNISCHE UNIVERSITÄT
CHEMNITZ

A Conjugate Direction Method for Linear Systems in Banach Spaces

Roland Herzog[*] Winnifried Wollner[†]

October 13, 2016

In this article, the well-known conjugate gradient (CG) method for linear systems in Hilbert spaces is extended to a reflexive Banach space setting. In this setting, the Riesz isomorphism has to be replaced by the duality mapping. Due to the nonlinearity of the duality mapping, the short term recursion and conjugacy of search directions cannot be maintained simultaneously. The well-posedness of the proposed iteration and its global convergence are shown under appropriate conditions. Error bounds and stopping criteria are presented as well. The results extend to a limited-memory variant of the algorithm. The behavior of the method is demonstrated by numerical examples.

Keywords: conjugate direction method; reflexive Banach space; linear operator equation

MSC: 65J10, 65J22, 65F25

1 Introduction

In this paper, we consider an iterative solution algorithm for the linear system

$$A\,x = b \tag{1.1}$$

[*] Technische Universität Chemnitz, Faculty of Mathematics, Professorship Numerical Mathematics (Partial Differential Equations), 09107 Chemnitz, Germany, roland.herzog@mathematik.tu-chemnitz.de, http://www.tu-chemnitz.de/herzog

[†] Technische Universität Darmstadt, Fachbereich Mathematik, Dolivostraße 15, 64293 Darmstadt, Germany, wollner@mathematik.tu-darmstadt.de, http://www.mathematik.tu-darmstadt.de/~wollner

where $A : X \to X^*$ is a bounded linear operator from a reflexive (real) Banach space X into its dual X^*, and $b \in X^*$ holds. It is assumed that X^* is strictly convex. The duality product is denoted by $\langle r, x\rangle$ for $r \in X^*$ and $x \in X$. We further assume A to be self-adjoint, positive and injective, i.e.,

$$\langle A\,x,\, y\rangle = \langle A\,y,\, x\rangle \text{ for all } x, y \in X \quad \text{and} \quad \langle A\,x,\, x\rangle > 0 \text{ for all } x \in X \setminus \{0\}.$$

Notice that under the assumptions above, A cannot be surjective unless the Banach space X is isomorphic to a Hilbert space. In case A is surjective, then necessarily $\langle Ax, x\rangle \geq c_A\, \|x\|_X^2$ holds for some $c_A > 0$; see [Ern and Guermond, 2004, Corollary A.55]. Consequently, X can be equipped with the scalar product $(x, y)_A := \langle Ax,\, y\rangle$, which gives rise to a Hilbert space structure on X with an equivalent norm $\|\cdot\|_A$, cf. [Ern and Guermond, 2004, Proposition A.49]. By the same arguments, we notice that the range of A cannot be a closed subspace of X^* unless X is isomorphic to a Hilbert space.

Hence, we do not assume that A is surjective, and consequently problem (1.1) does not necessarily possess a solution. If a solution $\overline{x}$ of (1.1) exists, then it is necessarily unique due to the injectivity of A. In that case, we are able to show the global strong convergence to zero of the residuals in X^*. The weak convergence of the iterates to $\overline{x}$ can be shown as well.

Finally, we consider the case that A is surjective, which is the generic situation in finite dimensions, where all Banach spaces are isomorphic to a Hilbert space. If A is surjective, then it is boundedly invertible by the Open Mapping Theorem, and hence strong convergence of the iterates in X is obtained. Even then, the use of a conjugate gradient (CG) algorithm in the Hilbert space $(X, \|\cdot\|_A)$ is not a viable option since application of the Riesz isomorphism (as it is needed in each iteration) would require the inversion of A; compare Günnel et al. [2014].

Problem (1.1) is motivated, for instance, by the solution of optimal control problems

$$\text{Minimize} \quad \frac{1}{2}\|S\,q - u_d\|_H^2 + \frac{\gamma}{p}\|q\|_{L^p(\Omega)}^p, \quad q \in L^p(\Omega) \tag{1.2}$$

with $p > 2$ by a Newton-type method. Here $X = L^p(\Omega)$ on some domain $\Omega \subset \mathbb{R}^n$ and $S : X \to H$ denotes a compact linear (control-to-state) mapping into a Hilbert space H, in which the observations of the state variable are made. In this situation, A is defined by

$$\langle A\,q,\, r\rangle = (S\,q, S\,r)_H + (p-1)\,\gamma \int_\Omega |\widetilde{q}|^{p-2}\, q\, r\, \mathrm{d}x, \tag{1.3}$$

where $\widetilde{q} \in X$ is the current linearization point. Such problems arise, for instance, in optimal control with state gradient constraints; see, e.g., Casas and Fernández [1993], Schiela and Wollner [2011]. Notice however, that Newton's method for (1.2) cannot be formulated in $X = L^p(\Omega)$ since $A \in \mathcal{L}(X, X^*)$ is never boundedly invertible, noting that $L^p(\Omega)$ is not isomorphic to a Hilbert space. Nevertheless, A is positive and hence invertible on any finite dimensional subspace of $L^p(\Omega)$, i.e., Newton's method for discretized instances of (1.2) is well-defined. It has been observed in the calculations done for Wollner [2010] that the main difficulty lies in the efficient solution of the Newton systems. The present paper is concerned with exactly that problem in the infinite dimensional Banach space as the limiting case for refined discretization.

Analogously as in the conjugate gradient (CG) method, we introduce the merit function

$$\phi\colon X \to \mathbb{R}, \quad x \mapsto \phi(x) := \frac{1}{2}\langle Ax,\, x\rangle - \langle b,\, x\rangle. \tag{1.4}$$

In order to define directions of steepest descent (the negative gradient in Hilbert space), we require a replacement for the Riesz isomorphism. This role is taken by the duality mapping; see, e.g., Cioranescu [1990]. To be precise, we consider a family of duality mappings $J_s : X \to \mathcal{P}(X^*)$ associated with the family of gauge functions $t \mapsto t^{s-1}$ depending on a parameter $1 < s < \infty$, which are defined by

$$J_s(x) = \{x^* \in X^* : \langle x^*,\, x\rangle = \|x\|_X \|x^*\|_{X^*}, \quad \|x^*\|_{X^*} = \|x\|_X^{s-1}\}.$$

Consequently, the pre-image satisfies

$$\begin{aligned} J_s^{-1}(x^*) &= \{x \in X : x^* \in J_s(x)\} \\ &= \{x \in X : \langle x,\, x^*\rangle = \|x^*\|_{X^*}\|x\|_X, \quad \|x\|_X = \|x^*\|_{X^*}^{s^*-1}\}, \end{aligned}$$

where s^* denotes the index conjugate to s. Notice that since X^* is assumed to be strictly convex, $J_s(x)$ is single-valued; see [Cioranescu, 1990, Chapter II, Corollary 1.5].

From the definition, it is easy to obtain the relations

$$J_r(x) = \|x\|_X^{r-s} J_s(x) \quad \text{and} \quad J_r^{-1}(x^*) = \|J_s^{-1}(x^*)\|_X^{(s-r)/(r-1)} J_s^{-1}(x^*) \tag{1.5}$$

for $1 < r, s < \infty$. Using (1.5) will allow us to show that our algorithm is independent of the choice of the index s in the duality mapping.

A prominent example motivated by (1.3) is the space $X = L^p(\Omega)$, whence $X^* \mathrel{\widehat{=}} L^{p^*}(\Omega)$. In this case,

$$J_p(x) = \operatorname{sgn}(x)\, |x|^{p-1}$$

holds for any $x \in X$.

Let us put our work into perspective. The work closest to ours seems to be Bonesky et al. [2008], where a class of minimization problems is considered which contains (1.3) as a special case. Two iterative methods are proposed, where the updates are performed in X and X^* respectively. However, conjugacy of the search directions is not enforced, which leads to gradient type methods.

In Margotti and Rieder [2015], similar to our motivating example (1.3), a Newton type algorithm is proposed for the solution of a nonlinear equation in Banach spaces. A functional involving the norm of the Newton residual plus a regularization term based on the Bregman distance is approximately minimized. This could in principle be solved by a steepest descent algorithm, utilizing the duality mapping.

In Schöpfer et al. [2006], the authors consider a gradient type method, generalizing the Landweber iteration Landweber [1951] to the Banach space setting, in the dual space X^* for the minimization of residual norm squared for a linear system in Banach spaces. In Schöpfer et al. [2008] this method was improved by sequential subspace optimization, i.e., the consideration of the history of previous directions. Schöpfer and Schuster [2009] improved the efficiency of this method by replacing the Bregman projection onto a set by a sequence of projections onto hyperplanes. In all of the above, conjugacy of the search directions is not enforced.

The rest of the paper is organized as follows. In the following Section 2, we present our conjugate direction method (Algorithm 2.1) for the solution of (1.1). We also prove a number of basic properties about the iterates, which are well known for the CG method in Hilbert spaces. In Section 3, we analyze the global convergence of the proposed algorithm. It is a disadvantage of our method that the amount of storage increases from iteration to iteration. As suggested by one of the reviewers, we therefore study a limited-memory variant in Section 4 with essentially the same properties but a constant amount of storage. Numerical examples are presented in Section 5.

2 A Conjugate Direction Method in Banach Space

To transfer the well known conjugate gradient method for the solution of the system

$$Ax = b \in X^*$$

where X is a Hilbert space to the setting of X being a reflexive Banach space (with strictly convex dual), we propose Algorithm 2.1. The motivation leading to the algorithm is to preserve as many of the properties of the conjugate gradient method in Hilbert space as possible.

The most striking difference is the lack of a Riesz isomorphism $J : X \to X^*$ between X and its dual X^*. This is compensated by the use of the duality mapping, which is, by our assumptions, single-valued. As we will show in Proposition 2.7, the iterates produced by Algorithm 2.1 are independent of the index s in the duality mapping. Therefore, we simply write J instead of J_s.

In contrast to the Riesz isomorphism, however, the duality mapping is *nonlinear*. Therefore, the well-known identity

$$\begin{aligned}\operatorname{span}\{d^0, \dots, d^k\} &= \mathcal{K}_{k+1}(J^{-1}A; J^{-1}r^0)\\ &:= \operatorname{span}\{J^{-1}r^0, (J^{-1}A)J^{-1}r^0, \dots, (J^{-1}A)^k J^{-1}r^0\}\end{aligned}$$

between the space of search directions d^j, $j = 0, \dots, k$, and the Krylov subspace $\mathcal{K}_{k+1}$, as it holds for the Hilbert space CG method, is no longer valid. Consequently, not all properties of CG can be preserved. We favor here to preserve the A-conjugacy, i.e., $\langle Ad^i, d^j \rangle = 0$ for $i \neq j$, of the search directions, at the expense of losing short recurrences. Note that A-conjugacy can be obtained analogously to the Hilbert space setting by a Gram-Schmidt-like procedure. The resulting algorithm reads as follows.

Algorithm 2.1 (Conjugate Direction Method in Banach Space)**.**

```
1:  Set $r^0 := b - Ax^0 \in X^*$
2:  Set $d^0 := J^{-1}(r^0) \in X$
3:  Set $k := 0$
4:  while not converged do
5:      Set $\alpha^k := \dfrac{\langle r^k, d^k \rangle}{\langle Ad^k, d^k \rangle}$
6:      Set $x^{k+1} := x^k + \alpha^k d^k$
7:      Set $r^{k+1} := r^k - \alpha^k A d^k$
8:      Set $\beta_i^{k+1} := \dfrac{\langle Ad^i, J^{-1}(r^{k+1}) \rangle}{\langle Ad^i, d^i \rangle}$    for $i = 0, \ldots, k$
9:      Set $d^{k+1} := J^{-1}(r^{k+1}) - \sum_{i=0}^{k} \beta_i^{k+1} d^i$
10:     Set $k := k + 1$
11: end while
```

We refer to the quantity $r^k \in X^*$ as the residual. By induction, it is obvious that $r^k = b - Ax^k$ holds.

Remark 2.2.

1. Similarly as the conjugate gradient method in Hilbert spaces, Algorithm 2.1 requires one application of A plus one application of the inverse duality mapping J^{-1} per iteration. The latter takes the role of a nonlinear preconditioner.
2. In contrast to the conjugate gradient method, however, we lose short recurrences, and therefore Algorithm 2.1 requires the storage of all search directions d^i, as well as the storage of their images Ad^i. This drawback can be overcome by a limited-memory variant, whose properties are studied in Section 4.
3. When X is a Hilbert space, then J^{-1} is identical to (a scalar multiple of) the (linear) Riesz map, and Algorithm 2.1 coincides with the conjugate gradient method. Full orthogonalization is not necessary then, i.e., the sum in step 9 can be truncated to the last term ($i = k$). Previous search directions d^i as well as their images Ad^i need not be stored.

Recall from (1.4) the merit function $\phi(x) := \frac{1}{2}\langle Ax, x \rangle - \langle b, x \rangle$. In this section, we do not assume that (1.1) is solvable. However, in case that (1.1) does possess a (unique) solution $\overline{x}$, a straightforward calculation shows that also

$$\phi(x) = \frac{1}{2}\|x - \overline{x}\|_A^2 - \frac{1}{2}\|\overline{x}\|_A^2 \tag{2.1}$$

holds. (We use here the symbol $\|x\|_A := \langle Ax, x \rangle^{1/2}$ although A does not necessarily define a norm.)

From Algorithm 2.1, it is easy to see that $x^k = \overline{x}$ implies $x^{k+1} = \overline{x}$ as well. Hence we assume in the following lemma that this is not the case.

Lemma 2.3. The iterates of Algorithm 2.1 satisfy the following properties as long as $x^k \neq \overline{x}$.

1. The directions d^j are A-conjugate, i.e.,
$$\langle Ad^j,\, d^k\rangle = 0 \quad \text{for all } 0 \le j < k.$$
2. The residual r^{k+1} satisfies
$$\langle r^{k+1},\, d^j\rangle = 0 \quad \text{for all } 0 \le j \le k.$$
3. The search directions satisfy $d^k \neq 0$, and $\dim \operatorname{span}\{d^0, \dots d^k\} = k+1$.
4. The residuals and search directions satisfy the following relations:
$$\langle r^k,\, d^k\rangle = \|r^k\|_{X^*}^{p^*} \quad \text{and} \quad \|r^k\|_{X^*} \le \|d^k\|_X^{p-1} \quad \text{for all } k \ge 0,$$
where p^* is the conjugate index to the gauge parameter p chosen for the duality mapping.
5. The number α^k is the exact minimizer of the uniformly convex quadratic polynomial
$$\alpha \mapsto \phi(x^k + \alpha\, d^k).$$
6. The vector x^{k+1} satisfies
$$x^{k+1} = \phi(x^k + \alpha^k\, d^k) = \operatorname*{arg\,min}_{d \in \operatorname{span}\{d^0, \dots, d^k\}} \phi(x^0 + d)$$

Proof:

1. The proof proceeds by induction over k. For $k = 0$ nothing has to be shown since the admissible index range for j is empty. Assuming that the assertion is true for some k, consider for $j < k+1$,
$$\begin{aligned}\langle Ad^j,\, d^{k+1}\rangle &= \langle Ad^j,\, J^{-1}(r^{k+1})\rangle - \sum_{i=0}^{k} \beta_i^{k+1} \langle Ad^j,\, d^i\rangle \\ &= \langle Ad^j,\, J^{-1}(r^{k+1})\rangle - \beta_j^{k+1} \langle Ad^j,\, d^j\rangle \\ &= \langle Ad^j,\, J^{-1}(r^{k+1})\rangle - \frac{\langle Ad^j,\, J^{-1}(r^{k+1})\rangle}{\langle Ad^j,\, d^j\rangle} \langle Ad^j,\, d^j\rangle = 0.\end{aligned}$$
2. The proof is again by induction over k. For $k = 0$, we have
$$\langle r^1,\, d^0\rangle = \langle r^0 - \alpha^0 Ad^0,\, d^0\rangle = \langle r^0,\, d^0\rangle - \frac{\langle r^0,\, d^0\rangle}{\langle Ad^0,\, d^0\rangle} \langle Ad^0,\, d^0\rangle = 0.$$

Assuming the assertion to hold for some k, we have in case $j \le k-1$

$$\langle r^{k+1}, d^j \rangle = \langle r^k - \alpha^k A d^k, d^j \rangle = 0$$

using the A-conjugacy of d^j and d^k. In case $j = k$, we have

$$\langle r^{k+1}, d^k \rangle = \langle r^k - \alpha^k A d^k, d^k \rangle = \langle r^k, d^k \rangle - \frac{\langle r^k, d^k \rangle}{\langle A d^k, d^k \rangle} \langle A d^k, d^k \rangle = 0.$$

3. The assertion follows, again, by induction. For $k = 0$, suppose $x^0 \neq \overline{x}$. Then, certainly, $r^0 \neq 0$ and by definition of $d^0 = J^{-1}(r^0)$, or $r^0 = J(d^0)$, since

$$\begin{aligned}\langle r^0, d^0 \rangle &= \langle r_0, J^{-1}(r_0) \rangle \\ &= \|r_0\|_{X^*}^{s^*} \neq 0\end{aligned}$$

and thus $d^0 \neq 0$. Here s is the index of the duality mapping $J = J_s$, and s^* is its conjugate. (Notice that the choice of s is irrelevant for the argument.)

Assuming that the assertion is true for some k, we can show the assertion for $k+1$. Again, certainly, $r^{k+1} \neq 0$ holds whenever $x^{k+1} \neq \overline{x}$. As for $k = 0$, this implies $J^{-1}(r^{k+1}) \neq 0$ and, more importantly, $\langle r^{k+1}, J^{-1}(r^{k+1}) \rangle \neq 0$. Using part (2) of the present lemma, we calculate

$$\begin{aligned}\langle r^{k+1}, d^{k+1} \rangle &= \langle r^{k+1}, J^{-1}(r^{k+1}) \rangle - \sum_{i=0}^{k} \beta_i^{k+1} \langle r^{k+1}, d^i \rangle \\ &= \langle r^{k+1}, J^{-1}(r^{k+1}) \rangle \neq 0.\end{aligned}$$

Hence $d^{k+1} \neq 0$ holds as claimed.

To show the linear independence of $d^0, \dots, d^k$, consider

$$0 = \sum_{i=0}^{k} \sigma_i \, d^i$$

with $\sigma_i \in \mathbb{R}$. Testing this equation with $A d^j$ and using the A-conjugacy of the directions, we obtain

$$0 = \sum_{i=0}^{k} \sigma_i \, \langle A d^i, d^j \rangle = \sigma_j \, \langle A d^j, d^j \rangle$$

for any $j = 0, \dots, k$. The positivity of A implies $\sigma_j = 0$, which proves the assertion.

4. By the proof of part (3), we have

$$\langle r^k, d^k \rangle = \langle r^k, J^{-1}(r^k) \rangle.$$

By the definition of J^{-1}, this implies

$$\langle r^k, d^k \rangle = \|r^k\|_{X^*}^{p^*}.$$

and thus

$$\|r^k\|_{X^*} \le \|d^k\|_X^{p-1}.$$

5. We calculate

$$\phi(x^k + \alpha\, d^k) = \frac{\alpha^2}{2}\langle Ad^k, d^k\rangle + \alpha\langle Ax^k - b, d^k\rangle + \frac{1}{2}\langle Ax^k, x^k\rangle - \langle b, x^k\rangle.$$

Due to the positivity and injectivity of A and part (3) of the present lemma, we know $\langle Ad^k, d^k\rangle > 0$. Hence ϕ is a strictly convex quadratic polynomial in α. Consequently, its unique minimizer is characterized by the condition

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}\alpha}\phi(x^k + \alpha\, d^k) \\ &= \alpha\langle Ad^k, d^k\rangle + \langle Ax^k - b, d^k\rangle \\ &= \alpha\langle Ad^k, d^k\rangle - \langle r^k, d^k\rangle, \end{aligned}$$

which holds iff $\alpha = \alpha^k = \frac{\langle r^k, d^k\rangle}{\langle Ad^k, d^k\rangle}$.

6. Minimizing ϕ over $x^0 + \operatorname{span}\{d^0, \ldots, d^k\}$ is equivalent to minimizing

$$\sigma = (\sigma_0, \ldots, \sigma_k) \mapsto \phi\big(x^0 + \sum_{i=0}^{k} \sigma_i\, d^i\big)$$

over $\mathbb{R}^{k+1}$. Analogously as in part (5), this function is strictly convex, and its unique minimizer is characterized by the conditions

$$0 = \langle Ax^0 + \sum_{i=0}^{k} \sigma_i\, A\, d^i - b, d^j\rangle \quad \text{for all } j = 0, \ldots, k.$$

Using the A-conjugacy twice, this is equivalent to

$$\begin{aligned} 0 &= \langle Ax^0 + \sigma_j\, Ad^j - b, d^j\rangle \\ &= \langle Ax^0 + \sum_{i=0}^{j-1} \alpha^i Ad^i + \sigma_j\, Ad^j - b, d^j\rangle \\ &= \langle Ax^j + \sigma_j\, Ad^j - b, d^j\rangle \\ &= \sigma_j\,\langle Ad^j, d^j\rangle - \langle r^j, d^j\rangle. \end{aligned}$$

holding for all $j = 0, \ldots, k$. Comparing this with the formula for α^j shows $\sigma_j = \alpha^j$, which implies the assertion.

□

Remark 2.4. It follows from part (5) and the properties of A that $\langle Ad^k, d^k\rangle \neq 0$ holds unless $x^k = \bar{x}$. Consequently, α^k and β_i^k are well defined, and Algorithm 2.1 either produces an infinite sequences of well-defined iterates, or stops with $x^k = \bar{x}$.

Corollary 2.5. When X is finite dimensional of dimension n, then $\bar{x}$ exists and Algorithm 2.1 converges in at most n iterations.

Proof: This property follows directly from Lemma 2.3, parts (3) and (6). □

Remark 2.6. In practical realizations of the algorithm it is convenient to stop the iteration once the residual of the equation has become sufficiently small. As in the CG method, we can access the norm of the residual directly, noticing that by part (4) of Lemma 2.3, we can compute the norm of the residual by

$$\|r^k\|_{X^*} = \langle r^k, d^k \rangle^{(p-1)/p}, \tag{2.2}$$

where p is the gauge parameter chosen for the duality mapping. The quantity $\langle r^k, d^k \rangle$ is readily available, hence stopping the iteration once

$$\|r^k\|_{X^*} \le \text{TOL} \tag{2.3}$$

for some prescribed tolerance TOL provides an economic stopping criterion.

As was pointed out in (1.5), duality maps with different gauging parameters return results which are scalar multiples of each other. This does not affect the outcome of the one-dimensional minimization. Consequently, it is not surprising that the iterates x^k of Algorithm 2.1 do not depend on the choice of the gauging parameter in the duality map. This is made precise in the following proposition.

Proposition 2.7. Consider the iterates x^k, r^k, and d^k of Algorithm 2.1 when the duality map J_s is used, and denote by $\widehat{x}^k$, $\widehat{r}^k$ and $\widehat{d}^k$ the corresponding iterates when the duality map J_r is used. If $x^0 = \widehat{x}^0$, then $x^k = \widehat{x}^k$ and $r^k = \widehat{r}^k$ holds for all k, and $\widehat{d}^k = \delta^k \, d^k$ with $\delta^k = \|J_s^{-1}(r^k)\|_X^{(s-r)/(r-1)}$.

Proof: The proof proceeds by induction over k. The claim is true for $k = 0$ since $x^0 = \widehat{x}^0$, whence $r^0 = \widehat{r}^0$ holds. Consequently,

$$d^0 = J_s^{-1}(r^0) \quad \text{and} \quad \widehat{d}^0 = J_r^{-1}(\widehat{r}^0) = J_r^{-1}(r^0) = \delta^0 \, d^0$$

holds with $\delta^0 = \|J_s^{-1}(r^0)\|_X^{(s-r)/(r-1)}$; see (1.5). Now consider the step from k to $k+1$ in Algorithm 2.1. We have

$$\widehat{\alpha}^k = \frac{\langle \widehat{r}^k, \widehat{d}^k \rangle}{\langle A\widehat{d}^k, \widehat{d}^k \rangle} = \frac{1}{\delta^k} \frac{\langle r^k, d^k \rangle}{\langle Ad^k, d^k \rangle} = \frac{\alpha^k}{\delta^k}$$

and therefore

$$\widehat{x}^{k+1} = \widehat{x}^k + \widehat{\alpha}^k \widehat{d}^k = x^k + \frac{\alpha^k}{\delta^k} \delta^k \, d^k = x^k + \alpha^k \, d^k = x^{k+1}$$

as well as $\widehat{r}^{k+1} = r^{k+1}$. We next consider, for $0 \le i \le k$, the expression

$$\begin{aligned}\widehat{\beta}_i^{k+1} &= \frac{\langle A\widehat{d}^i, J_r^{-1}(r^{k+1})\rangle}{\langle A\widehat{d}^i, \widehat{d}^i\rangle} = \frac{1}{\delta^i}\frac{\langle Ad^i, J_s^{-1}(r^{k+1})\rangle}{\langle Ad^i, d^i\rangle}\|J_s^{-1}(r^{k+1})\|^{(s-r)/(r-1)}\\ &= \frac{1}{\delta^i}\,\beta_i^{k+1}\,\|J_s^{-1}(r^{k+1})\|^{(s-r)/(r-1)},\end{aligned}$$

which follows by (1.5) and the induction hypothesis. Therefore, we obtain

$$\begin{aligned}\widehat{d}^{k+1} &= J_r^{-1}(r^{k+1}) - \sum_{i=0}^{k}\widehat{\beta}_i^{k+1}\widehat{d}^i\\ &= \|J_s^{-1}(r^{k+1})\|^{(s-r)/(r-1)}\Big[J_s^{-1}(r^{k+1}) - \sum_{i=0}^{k}\frac{1}{\delta^i}\beta_i^{k+1}\,\delta^i\,d^i\Big]\\ &= \|J_s^{-1}(r^{k+1})\|^{(s-r)/(r-1)}d^{k+1},\end{aligned}$$

which concludes the proof. □

By Proposition 2.7, it is justified that we denote the duality map in Algorithm 2.1 simply by J.

3 Convergence Analysis

We now analyze the convergence of our algorithm. Due to Remark 2.4, we only need to consider the case that Algorithm 2.1 produces an infinite sequence of iterates and does not stop prematurely. We assume this throughout this section, since otherwise the statements are trivially true.

Lemma 3.1. Let the merit function ϕ given in (1.4) be bounded from below along the sequence of iterates x^k generated by Algorithm 2.1. Then the sequences r^k and d^k satisfy

$$\frac{\langle r^k, d^k\rangle^2}{\langle Ad^k, d^k\rangle} \to 0 \quad \text{as } k \to \infty,$$

and r^k converges strongly to zero in X^*.

Proof: Let $\overline{\phi}$ be a lower bound on $\phi(\overline{x})$. A direct calculation shows

$$\phi(x^{k+1}) - \phi(x^k) = -\frac{1}{2}\frac{\langle r^k, d^k\rangle^2}{\langle Ad^k, d^k\rangle}$$

and we obtain

$$\overline{\phi} - \phi(x^0) \le \phi(x^{k+1}) - \phi(x^0) = -\frac{1}{2}\sum_{i=0}^{k}\frac{\langle r^i, d^i\rangle^2}{\langle Ad^i, d^i\rangle} < 0.$$

The summability implies convergence

$$\frac{\langle Ad^k, d^k\rangle}{\langle r^k, d^k\rangle^2} \to 0.$$

Consequently there exists a sequence of real numbers $c_k \to 0$ such that

$$\langle r^k, d^k\rangle^2 \le c_k \langle Ad^k, d^k\rangle.$$

Now, using the definition of d^k and β_i^k together with the A-conjugacy, see part (1) of Lemma 2.3, we obtain

$$\begin{aligned}
\langle Ad^k, d^k\rangle &= \langle Ad^k, J^{-1}(r^k) - \sum_{i=0}^{k-1} \beta_i^k d^i\rangle \\
&= \langle Ad^k, J^{-1}(r^k)\rangle \\
&= \langle AJ^{-1}(r^k), J^{-1}(r^k)\rangle - \sum_{i=0}^{k-1} \beta_i^k \langle Ad^i, J^{-1}(r^k)\rangle \\
&= \langle AJ^{-1}(r^k), J^{-1}(r^k)\rangle - \sum_{i=0}^{k-1} \frac{\langle Ad^i, J^{-1}(r^k)\rangle^2}{\langle Ad^i, d^i\rangle} \\
&\le \langle AJ^{-1}(r^k), J^{-1}(r^k)\rangle \\
&\le \|A\|_{\mathcal{L}(X,X^*)} \|J^{-1}(r^k)\|_X^2 \\
&= \|A\|_{\mathcal{L}(X,X^*)} \|r^k\|_{X^*}^{2s^*-2}.
\end{aligned} \tag{3.1}$$

Now, by part (4) of Lemma 2.3, it holds

$$\|r^k\|^{2s^*} = \langle r^k, d^k\rangle^2 \le c_k \langle Ad^k, d^k\rangle \le c_k \|A\|_{\mathcal{L}(X,X^*)} \|r^k\|_{X^*}^{2s^*-2}$$

and the strong convergence of the residuals is shown. □

From this lemma, we can deduce convergence of the residuals as it is shown in the following.

Corollary 3.2. Suppose that $b \in \operatorname{range}(A)$, and that $\bar{x}$ is the unique solution to (1.1). Then for any initial guess $x^0 \in X$, the residuals r^k generated by Algorithm 2.1 converge strongly to zero in X^*. In addition, if the sequence x^k has a weak accumulation point $\widehat{x}$, then $\widehat{x} = \bar{x}$.

Proof: Using the alternative form (2.1), we see that ϕ is bounded from below by $\phi(\bar{x})$, and Lemma 3.1 gives the claimed strong convergence of the residuals. Now, if any subsequence x^k converges weakly to some $\widehat{x}$, then we obtain by weak continuity of A that

$$b - A\widehat{x} \leftharpoondown b - Ax^k = r^k \to 0$$

and hence $\widehat{x}$ must coincide with $\bar{x}$. □

Corollary 3.3. If A is surjective in addition to the assumptions of Corollary 3.2, then x^k converges strongly in X to $\overline{x}$.

Proof: If A is surjective, and hence possesses a continuous inverse A^{-1}, then the strong convergence of x^k follows from

$$x^k - \overline{x} = x^k - A^{-1}b = -A^{-1}r^k \to 0.$$

□

As we have remarked earlier, if the range of A is closed, then X is in fact a Hilbert space. Hence the above corollary is useful only in settings where the standard CG method is, in principle applicable as well. However, if the range of A is not closed, then it is also not surjective and thus the generic situation will be that $b \notin \operatorname{range} A$, but $b \in \operatorname{cl}\operatorname{range} A$ meaning that there exists a sequence of values $y^k \in X$ such that $b = \lim_{k\to\infty} Ay^k$. If $b \notin \operatorname{range} A$, then y^k cannot have a (weak) accumulation point in X, and consequently $\|y^k\|_X \to \infty$. It may still be true that $\|y^k\|_A$ remains bounded; which is a typical situation if A is a compact operator. Indeed in this situation it is still possible to derive the convergence of the residuals, as shown in the following corollary.

Corollary 3.4. Suppose that there exists a sequence $y^k \in X$ satisfying $\|y^k\|_A \le c < \infty$ for all $k \ge 0$ and

$$b - Ay^k \rightharpoonup 0 \qquad \text{in } X^*$$

as $k \to \infty$. Then for any initial guess $x^0 \in X$, the residuals r^k generated by Algorithm 2.1 converge strongly to zero in X^*.

Proof: Let $x \in X$ be arbitrary. We notice that by definition of y^k,

$$\begin{aligned}\phi(x) &= \frac{1}{2}\langle Ax,\, x\rangle - \lim_{k\to\infty}\langle Ay^k,\, x\rangle \\ &= \lim_{k\to\infty} \frac{1}{2}\|x - y^k\|_A^2 - \frac{1}{2}\|y^k\|_A^2 \\ &\ge -\lim_{k\to\infty}\frac{1}{2}\|y^k\|_A^2 \\ &> -\infty.\end{aligned}$$

Lemma 3.1 proves the assertion. □

In order to obtain quantitative convergence results, we need to assume the surjectivity (bounded invertibility) of A. As was mentioned in the introduction, A endows X with a Hilbert space structure in this case. Utilizing this scalar product, one could employ a standard CG or steepest descent algorithm, both of which would converge in one step, but would require the inversion of A during the formation of the steepest descent direction, so this is not a viable option.

Therefore, we would like to investigate the speed of convergence of Algorithm 2.1. Due to the lack of equivalence between the typical Krylov spaces and the space spanned by the search directions, we cannot expect to obtain the same convergence result available for the CG method in Hilbert spaces, whose proof is based upon this relation and the linearity of J.

The next best result would be a speed of convergence comparable to the steepest descent method in Hilbert spaces. To obtain such a result, the well known Kantorovich inequality is required; see Kantorovič [1948] or the translation Kantorovich [1952]. To the best of our knowledge, no generalization of this inequality to the Banach space setting is available. The Hilbert space analysis relies on the existence of a (Schauder) basis consisting of eigenvectors of A, a concept not applicable in our setting. Therefore, we extend the concept of the numerical range of the operator to our setting. However, the reader should be aware that this extension is not the one typically considered, see, e.g., [Trefethen and Embree, 2005, page 172 et.seq.], where the operator is taken to be a mapping from X into X.

Lemma 3.5. Suppose that A is surjective, i.e., boundedly invertible. Denote by $0 < \hat\lambda, \Lambda < \infty$ the bounds on the numerical ranges of A^{-1} and A, i.e.,

$$\hat\lambda^{-1} = \sup_{r \in X^* \setminus \{0\}} \frac{\langle r, A^{-1} r\rangle}{\|r\|^2_{X^*}}, \quad \Lambda = \sup_{x \in X \setminus \{0\}} \frac{\langle Ax, x\rangle}{\|x\|^2_X}.$$

Then the iterates generated by Algorithm 2.1 satisfy

$$\frac{\langle r^k, d^k\rangle^2}{\langle Ad^k, d^k\rangle\langle r^k, A^{-1}r^k\rangle} \geq \frac{\hat\lambda}{\Lambda}$$

Proof: As in the proof of Lemma 3.1, see (3.1), we utilize

$$\langle Ad^k, d^k\rangle \leq \langle AJ^{-1}(r^k), J^{-1}(r^k)\rangle.$$

We conclude that

$$\begin{aligned}\langle Ad^k, d^k\rangle\langle A^{-1}r^k, r^k\rangle &\leq \langle AJ^{-1}(r^k), J^{-1}(r^k)\rangle\langle A^{-1}r^k, r^k\rangle \\ &\leq \Lambda\hat\lambda^{-1}\|J^{-1}(r^k)\|^2_X\|r^k\|^2_{X^*} \\ &= \Lambda\hat\lambda^{-1}\langle r^k, J^{-1}(r^k)\rangle^2,\end{aligned}$$

where the last equality follows by definition of J^{-1}. Noticing that r^k is conjugate to all previous directions, see part (2) of Lemma 2.3, we obtain

$$\begin{aligned}\langle r^k, J^{-1}(r^k)\rangle &= \langle r^k, J^{-1}(r^k)\rangle - \sum_{i=0}^{k-1}\beta_i^k\langle r^k, d^i\rangle \\ &= \langle r^k, d^k\rangle,\end{aligned}$$

which concludes the proof. □

Theorem 3.6. Under the assumptions of Lemma 3.5, the iterates generated by Algorithm 2.1 satisfy

$$\phi(x^{k+1}) - \phi(\bar{x}) \le \Big(1 - \frac{\hat{\lambda}}{\Lambda}\Big)\big(\phi(x^k) - \phi(\bar{x})\big),$$

where ϕ is given in (1.4).

Moreover, let $0 < \lambda < \infty$ be the lower bound on the numerical range of A, i.e.,

$$\lambda = \inf_{x \in X} \frac{\langle Ax, x\rangle}{\|x\|_X^2},$$

then

$$\|x^k - \bar{x}\|_X \le \sqrt{\frac{\Lambda}{\lambda}} \Big(1 - \frac{\hat{\lambda}}{\Lambda}\Big)^{k/2} \|x^0 - \bar{x}\|_X.$$

Proof: Since A is surjective and positive, we have $\langle Ax,\, x\rangle \ge c_A \, \|x\|_X^2$ as mentioned in the introduction, and hence $\lambda \ge c_A > 0$ holds.

The proof follows the standard argumentation of the convergence proof for the steepest descent method. As in the proof of Lemma 3.1, we use

$$\phi(x^{k+1}) = \phi(x^k) - \frac{1}{2}\frac{\langle r^k,\, d^k\rangle^2}{\langle Ad^k,\, d^k\rangle}$$

so that we can represent the decay in the function values as follows,

$$\phi(x^{k+1}) - \phi(\bar{x}) = \phi(x^k) - \phi(\bar{x}) - \frac{1}{2}\frac{\langle r^k,\, d^k\rangle^2}{\langle Ad^k,\, d^k\rangle}.$$

For the first two terms on the right hand side, we calculate

$$\begin{aligned}
\phi(x^k) - \phi(\bar{x}) &= \frac{1}{2}\langle A(x^k - \bar{x}),\, x^k - \bar{x}\rangle \\
&= \frac{1}{2}\langle A(x^k - \bar{x}),\, A^{-1}A(x^k - \bar{x})\rangle \\
&= \frac{1}{2}\langle r^k,\, A^{-1}r^k\rangle.
\end{aligned}$$

Combining the previous two equations, we conclude

$$\begin{aligned}
\phi(x^{k+1}) - \phi(\bar{x}) &= \left(1 - \frac{\langle r^k,\, d^k\rangle^2}{\langle Ad^k,\, d^k\rangle\langle r^k,\, A^{-1}r^k\rangle}\right)\big(\phi(x^k) - \phi(\bar{x})\big) \\
&\le \Big(1 - \frac{\hat{\lambda}}{\Lambda}\Big)\big(\phi(x^k) - \phi(\bar{x})\big)
\end{aligned}$$

due to Lemma 3.5.

To prove the second claim, we use (2.1) and the first part of the theorem to get

$$\begin{aligned}\frac{\lambda}{2}\|x^k-\overline{x}\|_X^2 &\le \frac{1}{2}\langle A(x^k-\overline{x}), x^k-\overline{x}\rangle\\ &= \phi(x^k)-\phi(\overline{x})\\ &\le \left(1-\frac{\hat{\lambda}}{\Lambda}\right)^k(\phi(x^0)-\phi(\overline{x}))\\ &= \frac{1}{2}\left(1-\frac{\hat{\lambda}}{\Lambda}\right)^k\langle A(x^0-\overline{x}), x^0-\overline{x}\rangle\\ &\le \frac{\Lambda}{2}\left(1-\frac{\hat{\lambda}}{\Lambda}\right)^k\|x^0-\overline{x}\|_X^2,\end{aligned}$$

from where the assertion follows. □

4 Limited-Memory Variant

As was mentioned in Remark 2.2, one drawback of Algorithm 2.1 is the required amount of storage, which increases from iteration to iteration. We therefore consider in this section the following limited-memory variant.

Algorithm 4.1 (Limited-Memory Conjugate Direction Method in Banach Space).

```
1:  Set $r^0 := b - Ax^0 \in X^*$
2:  Set $d^0 := J^{-1}(r^0) \in X$
3:  Set $k := 0$
4:  while not converged do
5:      Set $\alpha^k := \frac{\langle r^k, d^k\rangle}{\langle Ad^k, d^k\rangle}$
6:      Set $x^{k+1} := x^k + \alpha^k d^k$
7:      Set $r^{k+1} := r^k - \alpha^k Ad^k$
8:      Set $\beta_i^{k+1} := \frac{\langle Ad^i, J^{-1}(r^{k+1})\rangle}{\langle Ad^i, d^i\rangle}$   for $i = M_L(k), \dots, k$
9:      Set $d^{k+1} := J^{-1}(r^{k+1}) - \sum_{i=M_L(k)}^{k} \beta_i^{k+1} d^i$
10:     Set $k := k + 1$
11: end while
```

Notice that the only difference to Algorithm 2.1 lies in the range of previous search directions d^i against which the subsequent direction d^{k+1} is being 'orthogonalized'; see steps 8 and 9. The quantity $M_L(k)$ is defined as

$$M_L(k) := \max\{0, k-L+1\},$$

and consequently $M_{L+1}(k) = \max\{0, k - L\}$ holds. The parameter $L \geq 0$ denotes the size of the memory. Quantities d^i and Ad^i with $i < M_L(k)$ are no longer needed and can be dropped. For $L = 0$, the index range in steps 8 and 9 is empty and we obtain a steepest descent-like method, while $L = \infty$ amounts to an unlimited amount of memory and the method coincides with Algorithm 2.1.

Let us point out which of the results of Section 2 and 3 continue to hold or need to be modified. The modifications of the proofs for the respective properties of Algorithm 2.1 are obvious and are therefore not repeated.

Similarly as in Lemma 2.3, we get

$$\langle Ad^j, d^k \rangle = 0 \quad \text{for all } M_{L+1}(k) \leq j < k$$

and

$$\langle r^{k+1}, d^j \rangle = 0 \quad \text{for all } M_{L+1}(k) \leq j \leq k.$$

As long as $x^k \neq \overline{x}$ holds, the search directions continue to satisfy $d^k \neq 0$, but we only get

$$\dim \operatorname{span}\{d^k, \ldots d^{k+L}\} = L + 1$$

and consequently $\dim \operatorname{span}\{d^0, \ldots d^k\} \geq \min\{k+1, L+1\}$. Indeed, the occurrence of linearly dependent (in fact, recurring) search directions is well known for the steepest descent method ($L = 0$) in Hilbert spaces. The properties

$$\langle r^k, d^k \rangle = \|r^k\|_{X^*}^{p^*} \quad \text{and} \quad \|r^k\|_{X^*} \leq \|d^k\|_X^{p-1},$$

which are essential for the evaluation of the stopping criterion, see (2.2) and (2.3), remain valid. Property (6) of Lemma 2.3 needs to be modified as follows:

$$x^{k+1} = \phi(x^k + \alpha^k d^k) = \operatorname*{arg\,min}_{d \in \operatorname{span}\{d^{M_{L+1}(k)}, \ldots, d^k\}} \phi(x^{M_{L+1}(k)} + d).$$

Clearly, the finite termination property (Corollary 2.5) no longer holds in general unless the memory window happens to be sufficiently large. As already mentioned, the residual norm can still be computed from (2.2). Independence of the method of the gauging parameter (Proposition 2.7) remains valid as well.

It is interesting that the convergence results of Section 3 continue to hold without modification. One may therefore argue that the extra storage and algebraic manipulations necessary to incorporate previous search directions do not seem to pay off. In practice, however, we observe that the storage of at least a handful of directions has a significant impact on the convergence history. However, an improvement of the steepest-descent like convergence result (Theorem 3.6) which reflects these experimental observations remains a topic for future research.

5 Numerical Experiments

In this section, we present a number of numerical experiments. The stopping criterion is $\|r^k\|_{X^*} \le 10^{-8}$ in each case. In Section 5.1 and 5.2, we consider the original method (Algorithm 2.1) with full orthogonalization. In Section 5.3 we study the impact of limited memory (Algorithm 4.1).

5.1 Experiments for Algorithm 2.1 in $X = \ell^p$

We consider the sequence space $X = \ell^p$ with $p \in (1, \infty)$ and define the operator $A \in \mathcal{L}(X, X^*)$ by

$$\left(A\,x\right)_n = \frac{x_n}{n} \quad \text{for } x \in \ell^p \text{ and } n \in \mathbb{N},$$

i.e., A is represented by an infinite diagonal matrix with entries $1/n$. Notice that A has dense range, since any element $b \in \ell^{p^*}$ can be approximated by the sequence $A\,x^k$ with $x^k_n = b_n/n$ for $n \le k$ and $x^k_n = 0$ otherwise, Consequently, we only need to consider two cases, $b \in \operatorname{range}(A)$ and $b \in \operatorname{cl}\operatorname{range}(A) \setminus \operatorname{range}(A)$.

Experiment 5.1 ($b \in \operatorname{range}(A)$)**.** We set $p = 10$ and use $b_n = n^{-1.2}$, which belongs to $\ell^{p^*} = \ell^{10/9}$. Since $\overline{x}_n = n^{-0.2}$ belongs to ℓ^p and satisfies $A\,\overline{x} = b$, b belongs to $A(\ell^p)$. Notice that $\overline{x} \notin \ell^2$ and thus $A\,x = b$ is not solvable in ℓ^2.

Figure 5.1 shows the convergence behavior of Algorithm 2.1 starting at initial guess $x^0 = 0$. As we showed in Proposition 2.7 and confirmed computationally, the gauging parameter has no influence on the iterates of the method. In our experiments, we used $s = 2$. Since ℓ^p is infinite dimensional, we consider the corresponding problem for finite sequences of length $N \in \{10^3, 10^4, 10^5\}$.

For comparison, we also calculated the results of the standard CG method applied to the respective finite dimensional problems in ℓ^2. To achieve a fair comparison, the CG method was emulated by Algorithm 2.1 using as preconditioner J^{-1} the identity map, i.e., even full orthogonalization is used which mitigates round-off error. Figure 5.1 shows that our new method performs much better than standard CG, and the difference becomes more pronounced as the problem size grows. This has to be expected since the limiting, infinite dimensional, problem has no solution in ℓ^2.

Experiment 5.2 ($b \notin \operatorname{range}(A)$)**.** We set again $p = 10$ but use $b_n = 1/n$, which belongs to $\ell^{p^*} = \ell^{10/9}$. However, the only solution to $A\,x = b$ would be $\overline{x}_n = 1$, which is not in ℓ^p. Hence $A\,x = b$ is not solvable in ℓ^p.

Figure 5.2 shows the convergence behavior of Algorithm 2.1 starting at initial guess $x^0 = 0$. As in the previous example, we consider the corresponding problem for finite

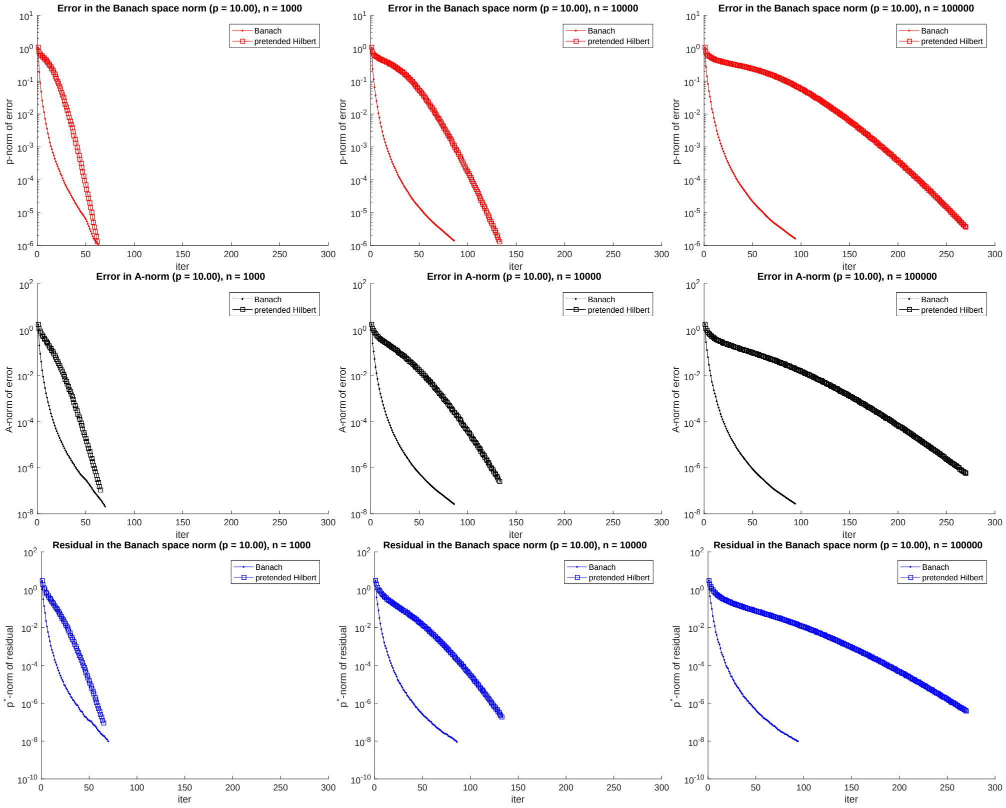

Figure 5.1: The first row shows the evolution of $\|x^k - \overline{x}\|_{\ell^p}$ for the proposed conjugate direction method (Algorithm 2.1) applied to Experiment 5.1, compared to the CG method, which is applicable only in the discretized setting. The problem size $N \in \{10^3, 10^4, 10^5\}$ is increasing from left to right. The second and third rows show corresponding plots for the error $\|x^k - \overline{x}\|_A$ and the residual $\|r^k\|_{\ell^{p^*}}$, respectively.

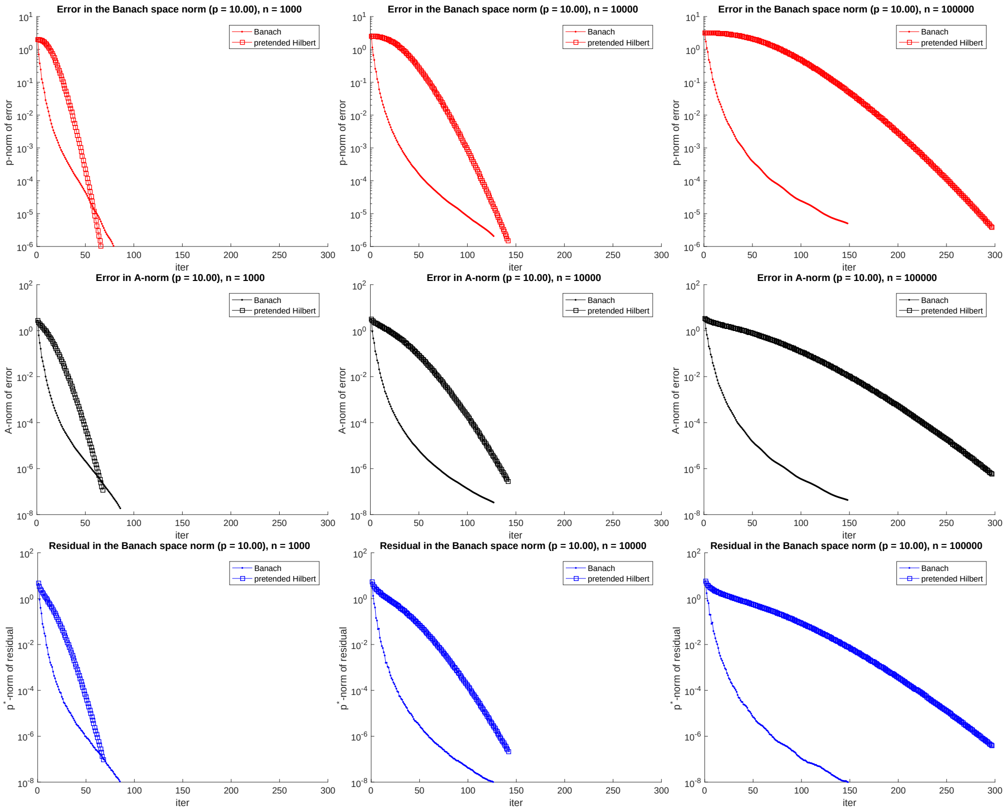

Figure 5.2: The first row shows the evolution of $\|x^k - \overline{x}\|_{\ell^p}$ for the proposed conjugate direction method (Algorithm 2.1) applied to Experiment 5.2, compared to the CG method, which is applicable only in the discretized setting. The problem size $N \in \{10^3, 10^4, 10^5\}$ is increasing from left to right. The second and third rows show corresponding plots for the error $\|x^k - \overline{x}\|_A$ and the residual $\|r^k\|_{\ell^{p^*}}$, respectively.

sequences of length $N \in \{10^3, 10^4, 10^5\}$. For each finite dimensional space, the corresponding solution is $\overline{x}_n = 1$. Despite the degeneracy of the problem, our method shows only a mild dependence on the discretization parameter, in contrast to standard CG.

As expected, the convergence of Algorithm 2.1 to the finite dimensional solution now exhibits a dependence on the problem size. Nevertheless, Figure 5.2 shows that our new method performs much better than standard CG, and the difference becomes more pronounced as the problem size grows.

Both Experiment 5.1 and Experiment 5.2 indicate the usefulness of considering algorithms explicitly for the Banach space setting.

5.2 Experiments for Algorithm 2.1 in $X = L^p(\Omega)$

The experiments in this section are based on the optimal control model problem (1.2) with $X = L^p(\Omega)$ for some $p > 2$ and observation space $H = L^2(\Omega)$. The solution operator $S : X \to H$ is defined by the unique solution $u = S\,q$ of the partial differential equation (PDE) $-\triangle u + u = q$ in Ω with homogeneous Neumann boundary conditions. We focus on the solution of the equation $A\,\overline{q} = b$ at a given point $\widetilde{q} \in L^p(\Omega)$, where A is given by (1.3). In the sequel, we describe a number of experiments to investigate the behavior of our conjugate direction method. As in the previous section, we compare it to the classical conjugate gradient method, which is applicable of course to any discretized version of the problem but disregards the underlying mapping properties of $A : X \to X^*$.

All experiments are set up on the unit circle $\Omega \subset \mathbb{R}^2$. For the application of the solution operator S, we discretize the PDE state u as well as the control q using a standard piecewise linear, continuous finite element basis on various levels of uniform refinement. The implementation was done with MATLAB's PDE toolbox (R2015b). For the evaluation of the L^p-norm, we utilize the following nodal quadrature formula

$$\|q\|^p_{L^p(\Omega)} \approx \sum_{T \in \mathcal{T}_h} \frac{|T|}{3} \sum_{i=1}^{3} q(v_{T,i})^p$$

on the triangulation $\mathcal{T}_h$, where $v_{T,i}$ represents the vertices of the triangle $T \in \mathcal{T}_h$. The corresponding duality map for $s = 2$ is based upon this quadrature formula and it corresponds to a weighted ℓ^p duality mapping for the coefficient vector representing q.

Experiment 5.3. In this experiment, we use $p = 20$ and $\gamma = 10^2$. As our point of linearization, we use $\widetilde{q} = \left(x_1^2 + x_2^2\right)^{-1/(p+\varepsilon)}$ for some small $\varepsilon > 0$, which belongs to $X = L^p(\Omega)$. In our calculations, we used $\varepsilon = 0.1$. In this situation, $A : X \to X^*$ is positive since

$$\langle A\,q,\, q\rangle = (S\,q, S\,q)_H + (p-1)\,\gamma \int_\Omega |\widetilde{q}|^{p-2}\, q^2 \,\mathrm{d}x > 0 \quad \text{for } q \in L^p(\Omega),\ q \neq 0$$

holds. The right hand side is chosen as $b = A\,\overline{q} \in L^{p^*}(\Omega)$ for some known function $\overline{q}$ so that we can measure the error of the iterates. In this experiment, we choose $\overline{q} = \widetilde{q}$.

The numerical results for Experiment 5.3 with initial guess $q^0 = 0$ are displayed in Figure 5.3. In the first row, we show the evolution of the error $\|\overline{q} - q^k\|_{L^p(\Omega)}$ for the proposed conjugate direction method (Algorithm 2.1), compared to the CG method. The CG method is set up in the Hilbert space $L^2(\Omega)$, which of course does not reflect the true mapping property of A since A does not map $L^2(\Omega)$ into its dual. We therefore expect the performance of CG to degrade on finer meshes. As we did previously, the CG method is emulated by Algorithm 2.1 using as preconditioner J^{-1} the linear Riesz operator, i.e., full orthogonalization is used. As before, we used the gauging parameter $s = 2$.

The first row of Figure 5.3 shows the evolution of $\|\overline{q} - q^k\|_{L^p(\Omega)}$ for the proposed conjugate direction method (Algorithm 2.1) applied to Experiment 5.3, compared to the CG method, which is applicable only in the discretized setting. The mesh level and thus problem size is increasing from left to right. The second and third rows show corresponding plots for the error $\|\overline{q} - q^k\|_A$ and the residual $\|r^k\|_{L^{p^*}(\Omega)}$, respectively.

We observe, as expected, a mesh dependent behavior of both our method and standard CG. For a large range of iteration numbers, our method performs significantly better than standard CG with respect to the three different convergence measures. Although finally the CG method seems to exhibit a better slope, the break-even point moves to higher and higher iteration numbers as the problem size grows.

5.3 Impact of Limited Memory (Algorithm 4.1)

In this section we briefly study the impact of limited memory in the generation of search directions, i.e., we compare the convergence behavior of Algorithm 4.1 for various memory size parameters $L \geq 0$ with Algorithm 2.1, which corresponds to $L = \infty$. For completeness, we repeat all Experiments 5.1, 5.2 and 5.3.

For the well-posed problem in Experiment 5.1, see Figure 5.4, we observe numerically that the steepest-descent like method ($L = 0$) performs poorly and that the incorporation of a number of previous search directions clearly pays off. On the other hand, already $L = 3$ directions seem to suffice to achieve a convergence behavior comparable to the unlimited memory case. This observation is true for all three convergence measures considered.

The results are, generally speaking, the same for Experiment 5.2 although here, interestingly the method with $L = 3$ seems to outperform the one with $L = \infty$ in the long run for sufficiently large problems; see Figure 5.5.

In Experiment 5.3, the method with few search directions $L \in \{0, 1, 2\}$ eventually outperforms the method using $L = \infty$. Hence in this example, using too large memory does not pay off.

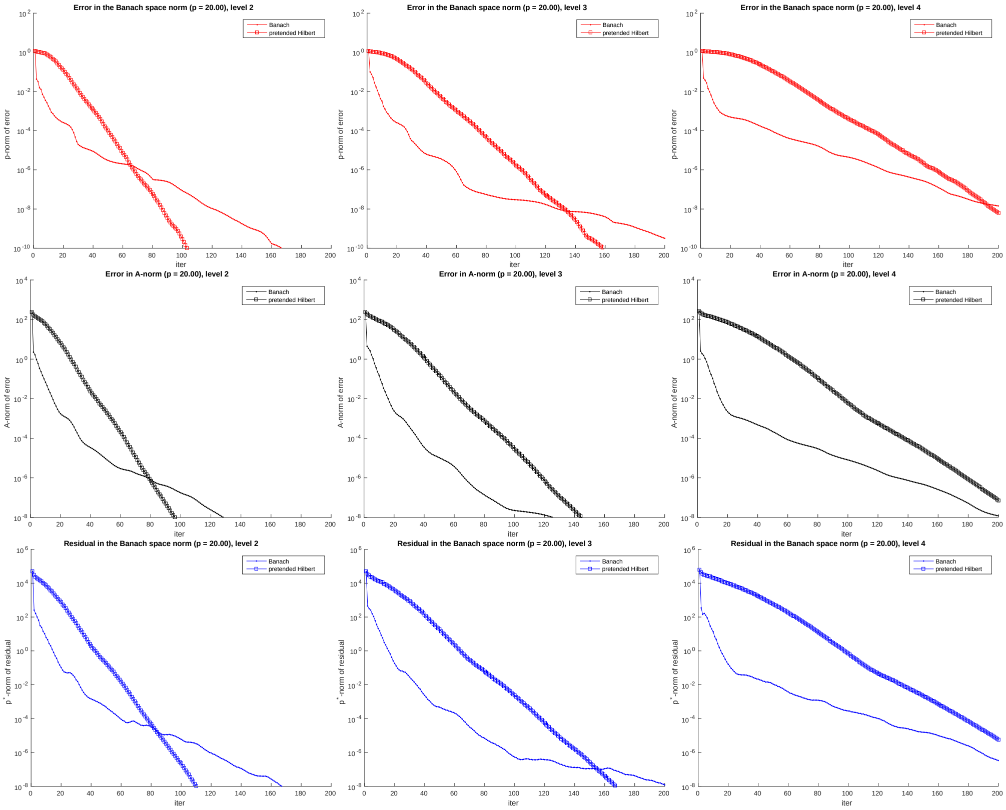

Figure 5.3: The first row shows the evolution of $\|\bar{q} - q^k\|_{L^p(\Omega)}$ for the proposed conjugate direction method (Algorithm 2.1) applied to Experiment 5.3, compared to the CG method, which is applicable only in the discretized setting. The mesh level and thus problem size is increasing from left to right. The second and third rows show corresponding plots for the error $\|\bar{q} - q^k\|_A$ and the residual $\|r^k\|_{L^{p^*}(\Omega)}$, respectively.

Let us mention that in all three experiments, the dominant cost per iteration is the evaluation of A and the application of the inverse duality mapping. Even for large iteration numbers, we found the algebraic operations necessary to incorporate many previous search directions negligible; see steps 8 and 9 in Algorithm 2.1. Clearly, an important benefit of using the limited-memory variant is the aspect of constant storage, especially for large-scale problems. This has to be weighted against the alteration of the convergence history, which may degrade or even improve compared to the full-memory variant, depending on the problem. This issue deserves further investigation.

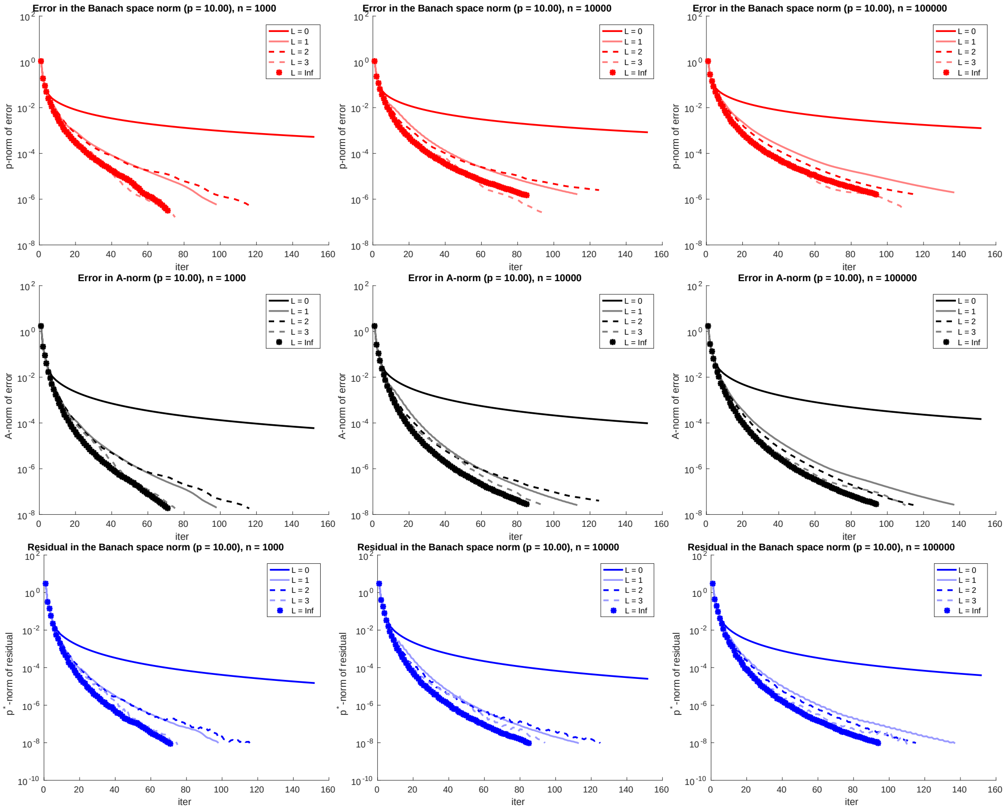

Figure 5.4: The first row shows the evolution of $\|x^k - \overline{x}\|_{\ell^p}$ for the proposed conjugate direction method (Algorithm 2.1, $L = \infty$) applied to Experiment 5.1, compared to the limited-memory variant (Algorithm 4.1) for memory sizes $L \in \{0, 1, 2, 3\}$. The problem size $N \in \{10^3, 10^4, 10^5\}$ is increasing from left to right. The second and third rows show corresponding plots for the error $\|x^k - \overline{x}\|_A$ and the residual $\|r^k\|_{\ell^{p^*}}$, respectively.

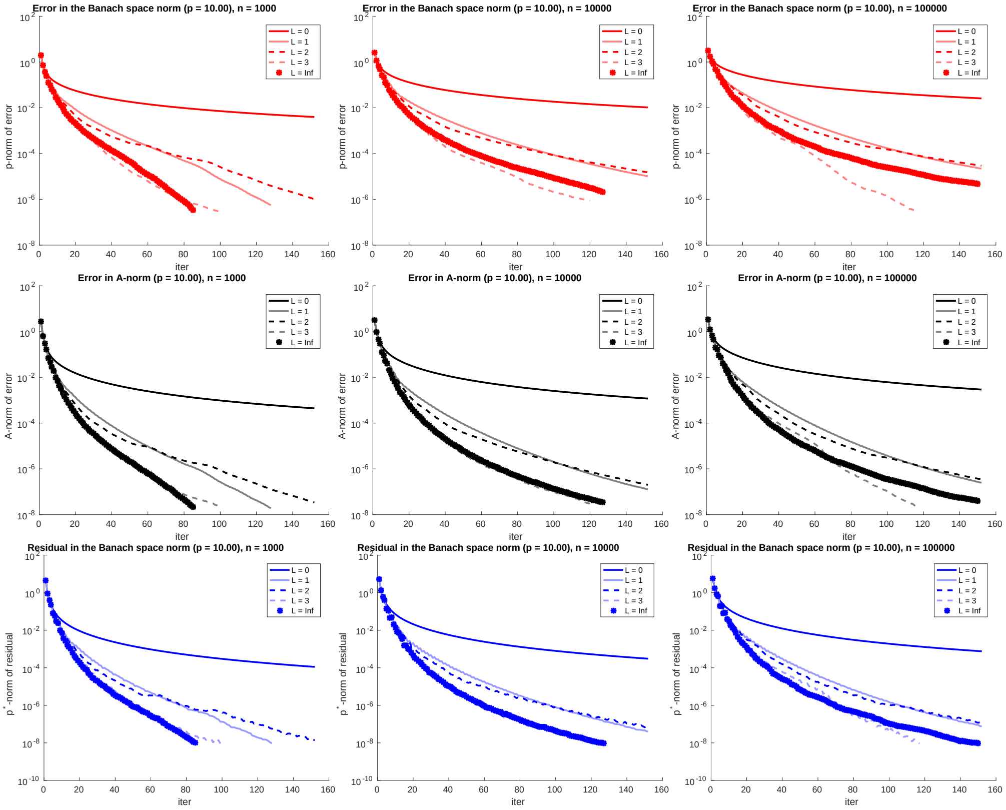

Figure 5.5: The first row shows the evolution of $\|x^k - \overline{x}\|_{\ell^p}$ for the proposed conjugate direction method (Algorithm 2.1, $L = \infty$) applied to Experiment 5.2, compared to the limited-memory variant (Algorithm 4.1) for memory sizes $L \in \{0, 1, 2, 3\}$. The problem size $N \in \{10^3, 10^4, 10^5\}$ is increasing from left to right. The second and third rows show corresponding plots for the error $\|x^k - \overline{x}\|_A$ and the residual $\|r^k\|_{\ell^{p^*}}$, respectively.

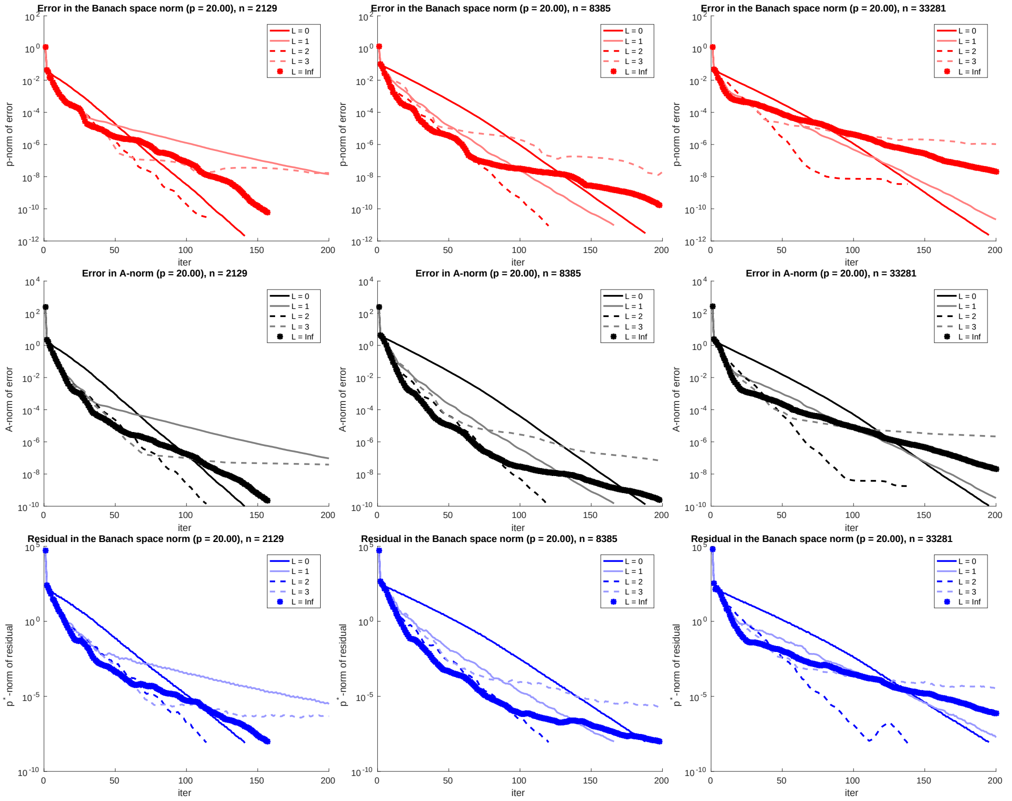

Figure 5.6: The first row shows the evolution of $\|\overline{q} - q^k\|_{L^p(\Omega)}$ for the proposed conjugate direction method (Algorithm 2.1, $L = \infty$) applied to Experiment 5.3, compared to the limited-memory variant (Algorithm 4.1) for memory sizes $L \in \{0, 1, 2, 3\}$. The mesh level and thus problem size is increasing from left to right. The second and third rows show corresponding plots for the error $\|\overline{q} - q^k\|_A$ and the residual $\|r^k\|_{L^{p^*}(\Omega)}$, respectively.

Acknowledgments

The authors would like to thank two anonymous referees for their valuable suggestions, which helped improve the presentation. We are particularly thankful to one referee who suggested to consider the limited-memory version of Algorithm 2.1, which led to the addition of Sections 4 and 5.3.

References

T. Bonesky, K. S. Kazimierski, P. Maass, F. Schöpfer, and T. Schuster. Minimization of Tikhonov functionals in Banach spaces. *Abstract and Applied Analysis*, pages Art. ID 192679, 19, 2008. ISSN 1085-3375. doi: 10.1155/2008/192679.

E. Casas and L.A. Fernández. Optimal control of semilinear elliptic equations with pointwise constraints on the gradient of the state. *Applied Mathematics and Optimization. An International Journal with Applications to Stochastics*, 27(1):35–56, 1993. doi: 10.1007/BF01182597.

I. Cioranescu. *Geometry of Banach spaces, duality mappings and nonlinear problems*, volume 62 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1990. ISBN 0-7923-0910-3. doi: 10.1007/978-94-009-2121-4.

A. Ern and J.-L. Guermond. *Theory and Practice of Finite Elements*. Springer, Berlin, 2004.

A. Günnel, R. Herzog, and E. Sachs. A note on preconditioners and scalar products in Krylov subspace methods for self-adjoint problems in Hilbert space. *Electronic Transactions on Numerical Analysis*, 41:13–20, 2014.

L. V. Kantorovič. Functional analysis and applied mathematics. *Uspekhi Mat. Nauk (N.S.)*, 3(6(28)):89–185, 1948. ISSN 0042-1316.

L. V. Kantorovich. *Functional analysis and applied mathematics*. NBS Rep. 1509. U. S. Department of Commerce, National Bureau of Standards, Los Angeles, Calif., 1952. Translated by C. D. Benster.

L. Landweber. An iteration formula for Fredholm integral equations of the first kind. *American Journal of Mathematics*, 73(3):615–624, 1951. doi: 10.2307/2372313.

F. Margotti and A. Rieder. An inexact Newton regularization in Banach spaces based on the nonstationary iterated Tikhonov method. *Journal of Inverse and Ill-Posed Problems*, 23(4):373–392, 2015. ISSN 0928-0219. doi: 10.1515/jiip-2014-0035.

A. Schiela and W. Wollner. Barrier methods for optimal control problems with convex nonlinear gradient state constraints. *SIAM Journal on Optimization*, 21(1):269–286, 2011.

F. Schöpfer and T. Schuster. Fast regularizing sequential subspace optimization in Banach spaces. *Inverse Problems. An International Journal on the Theory and Practice of Inverse Problems, Inverse Methods and Computerized Inversion of Data*, 25(1):015013, 22, 2009. ISSN 0266-5611. doi: 10.1088/0266-5611/25/1/015013.

F. Schöpfer, A. K. Louis, and T. Schuster. Nonlinear iterative methods for linear ill-posed problems in Banach spaces. *Inverse Problems. An International Journal on the Theory and Practice of Inverse Problems, Inverse Methods and Computerized Inversion of Data*, 22(1):311–329, 2006. ISSN 0266-5611. doi: 10.1088/0266-5611/22/1/017.

F. Schöpfer, T. Schuster, and A. K. Louis. Metric and Bregman projections onto affine subspaces and their computation via sequential subspace optimization methods. *Journal of Inverse and Ill-Posed Problems*, 16(5):479–506, 2008. ISSN 0928-0219. doi: 10.1515/JIIP.2008.026.

L. N. Trefethen and M. Embree. *Spectra and pseudospectra*. Princeton University Press, Princeton, NJ, 2005. The behavior of nonnormal matrices and operators.

W. Wollner. A posteriori error estimates for a finite element discretization of interior point methods for an elliptic optimization problem with state constraints. *Computational Optimization and Applications. An International Journal*, 47(1):133–159, 2010. ISSN 0926-6003. doi: 10.1007/s10589-008-9209-2.