Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

STEFAN VOLKWEIN MARTIN WEISER

Affine Invariant Convergence Analysis for Inexact Augmented Lagrangian-SQP Methods

ZIB-Report 00-56 (Dezember 2000)

Affine Invariant Convergence Analysis for Inexact Augmented Lagrangian-SQP Methods

Stefan Volkwein* Martin Weiser†

4th January 2001

Abstract

An affine invariant convergence analysis for inexact augmented Lagrangian-SQP methods is presented. The theory is used for the construction of an accuracy matching between iteration errors and truncation errors, which arise from the inexact linear system solves. The theoretical investigations are illustrated numerically by an optimal control problem for the Burgers equation.

Keywords: Nonlinear programming, multiplier methods, affine invariant norms, Burgers' equation

MSC (2000): 49M15, 65N99, 90C55

*Karl-Franzens-Universität Graz, Institut für Mathematik, Heinrichstraße 36, A-8010 Graz
e-mail: stefan.volkwein@kfunigraz.ac.at
supported by: Fonds zur Förderung der wissenschaftlichen Forschung under the Special Research Center F003 "Optimization and Control"
†Konrad-Zuse-Zentrum Berlin, Takustraße 7, D-14195 Berlin
e-mail: weiser@zib.de

1 Introduction

This paper is concerned with an optimization problem of the following type:

$$\text{minimize } J(x) \text{ subject to } e(x) = 0, \tag{P}$$

where $J : X \to \mathbb{R}$ and $e : X \to Y$ are sufficiently smooth functions and X, Y are real Hilbert spaces. These types of problems occur, for example, in the optimal control of systems described by partial differential equations. To solve (P) we use the augmented Lagrangian-SQP (sequential quadratic programming) technique as developed in [9]. In this method the differential equation is treated as an equality constraint, which is realized by a Lagrangian term together with a penalty functional. We present an algorithm, which has second-order convergence rate and depends upon a second-order sufficient optimality condition. In comparison with SQP methods the augmented Lagrangian-SQP method has the advantage of a more global behavior, see e.g. [9, 12, 13]. For certain examples we found it to be less sensitive with respect to the starting values, and the region for second-order convergence rate was reached earlier. We shall point out that the penalty term of the augmented Lagrangian functional need not to be implemented but rather that it can be realized by a first-order Lagrangian update.
Augmented Lagrangian-SQP methods applied to problem (P) are essentially Newton type methods applied to the Kuhn–Tucker equations for an augmented optimization problem. Newton methods and their behavior under different linear transformations were studied by several authors, see [3, 4, 5, 6, 8], for instance. In this paper, we combine both lines of work and present an affine invariant setting for analysis and implementation of augmented Lagrangian-SQP methods in Hilbert spaces. An affine invariant convergence theory for inexact augmented Lagrangian-SQP methods is presented. Then the theoretical results are used for the construction of an accuracy matching between iteration errors and truncation errors, which arise from the inexact linear system solves.
The paper is organized as follows. In section 2 the augmented Lagrangian-SQP method is introduced and necessary prerequisites are given. The affine invariance is introduced in section 3. In section 4 an affine invariant convergence result for the augmented Lagrangian-SQP method is presented. Two invariant norms for optimal control problems are analyzed in section 5, and the inexact Lagrangian-SQP method is studied in section 6. In the last section we report on some numerical experiments done for an optimal control problem for the Burgers equation, which is a one-dimensional model for nonlinear convection-diffusion phenomena.

2 The augmented Lagrangian-SQP method

Let us consider the constrained optimal control problem

$$\text{minimize } J(x) \text{ subject to } e(x) = 0, \tag{P}$$

where $J : X \to \mathbb{R}$, $e : X \to Y$ and X, Y are real Hilbert spaces. Throughout we do not distinguish between a functional in the dual space and its Riesz representation in the Hilbert space. The Hilbert space $X \times Y$ is endowed with the Hilbert space product topology and, for brevity, we set $Z = X \times Y$.
Let us present an example for (P) that illustrates our theoretical investigations and that is used for the numerical experiments carried out in section 7. For more details we refer the reader to [14].

Example 2.1. Let Ω denote the interval $(0, 1)$ and set $Q = (0, T) \times \Omega$ for given $T > 0$. We define the space $W(0, T)$ by

$$W(0,T) = \left\{\varphi \in L^2(0,T;H^1(\Omega)) : \ \varphi_t \in L^2(0,T;H^1(\Omega)')\right\},$$

which is a Hilbert space endowed with the common inner product. For controls $u, v \in L^2(0, T)$ the state $y \in W(0, T)$ is given by the weak solution of the unsteady Burgers equation with Robin type boundary conditions, i.e., y satisfies

$$y(0,\cdot) = y_0 \quad \text{in } L^2(\Omega) \tag{2.1a}$$

and

$$\begin{aligned} &\langle y_t(t,\cdot),\varphi\rangle_{(H^1)',H^1} + \sigma_1(t)y(t,1)\varphi(1) - \sigma_0(t)y(t,0)\varphi(0) \\ &+ \int_\Omega \nu y_x(t,\cdot)\varphi' + \big(y(t,\cdot)y_x(t,\cdot) - f(t,\cdot)\big)\varphi\,dx = v(t)\varphi(1) - u(t)\varphi(0) \end{aligned} \tag{2.1b}$$

for all $\varphi \in H^1(\Omega)$ and $t \in (0,T)$ a.e., where $\langle \cdot\,,\cdot\rangle_{(H^1)',H^1}$ denotes the duality pairing between $H^1(\Omega)$ and its dual. We suppose that $f \in L^2(\Omega)$, $y_0 \in L^\infty(\Omega)$, $\sigma_0, \sigma_1 \in L^\infty(0,T)$ and that $\nu > 0$. Recall that $W(0,T)$ is continuously embedded into the space of all continuous functions from $[0,T]$ into $L^2(\Omega)$, denoted by $C([0,T];L^2(\Omega))$, see e.g. [2, p. 473]. Therefore, (2.1a) makes sense. With every controls u, v we associate the cost of tracking type

$$J(y,u,v) = \frac{1}{2}\int_Q |y-z|^2\,dxdt + \frac{1}{2}\int_0^T \alpha\,|u|^2 + \beta\,|v|^2\,dt,$$

where $z \in L^2(Q)$ and $\alpha, \beta > 0$ are fixed. Let $X = W(0,T) \times L^2(0,T) \times L^2(0,T)$, $Y = L^2(0,T;H^1(\Omega)) \times L^2(\Omega)$ and $x = (y,u,v)$. We introduce the bounded operator

$$\tilde{e} : X \to L^2(0,T;H^1(\Omega)'),$$

whose action is defined by

$$\begin{aligned} &\langle \tilde{e}(y,u,v),\lambda\rangle_{L^2(0,T;H^1(\Omega)'),L^2(0,T;H^1(\Omega))} \\ &\quad = \int_0^T \langle y_t(t,\cdot),\lambda(t,\cdot)\rangle_{(H^1)',H^1}\,dt + \int_Q (\nu y_x\lambda_x + yy_x\lambda - f\lambda)\,dxdt \\ &\qquad + \int_0^T \big((\sigma_1 y(\cdot,1) - v)\lambda(\cdot,1) + (\sigma_0 y(\cdot,0) - u)\lambda(\cdot,0)\big)\,dt \end{aligned}$$

for $\lambda \in L^2(0,T;H^1(\Omega))$. Defining $e : X \to Y$ by

$$e(y,u,v) = \big((-\Delta + I)^{-1}\tilde{e}(y,u,v), y(0,\cdot) - y_0\big),$$

where for given $g \in H^1(\Omega)'$ the mapping $(-\Delta + I)^{-1} : H^1(\Omega)' \to H^1(\Omega)$ is the Neumann solution operator associated with

$$\int_\Omega v'\varphi' + v\varphi \, dx = \langle g, \varphi \rangle_{(H^1)',H^1} \text{ for all } \varphi \in H^1(\Omega).$$

Now the optimal control problem can be written in the form (P). ♢

For $c \geq 0$ the augmented Lagrange functional $L_c : Z \to \mathbb{R}$ associated with (P) is defined by

$$L_c(x,\lambda) = J(x) + \langle e(x), \lambda \rangle_Y + \frac{c}{2} \|e(x)\|_Y^2.$$

The following assumption is rather standard for SQP methods in Hilbert spaces.

Assumption 1. *Let $x^* \in X$ be a reference point such that*

a) *J and e are twice continuously Fréchet-differentiable, and the mappings J'' and e'' are Lipschitz-continuous in a neighborhood of x^*,*

b) *the linearization $e'(x^*)$ of the operator e at x^* is surjective,*

c) *there exists a Lagrange multiplier $\lambda^* \in Y$ satisfying the first-order necessary optimality conditions*

$$L_c'(x^*,\lambda^*) = 0, \;\; e(x^*) = 0 \quad \textit{for all } c \geq 0, \tag{2.2}$$

where the Fréchet-derivative with respect to the variable x is denoted by a prime, and

d) *there exists a constant $\kappa > 0$ such that*

$$\langle L_0''(x^*,\lambda^*)\chi, \chi \rangle_X \geq \kappa \|\chi\|_X^2 \quad \textit{for all } \chi \in \ker e'(x^*),$$

where ker *$e'(x^*)$ denotes the kernel or null space of $e'(x^*)$.*

Remark 2.2. In the context of Example 2.1 we write $x^* = (y^*, u^*, v^*)$. It was proved in [14] that Assumption 1 holds provided $\|y^* - z\|_{L^2(Q)}$ is sufficiently small. ♢

The next proposition follows directly from Assumption 1. For a proof we refer to [10] and [11], for instance.

Proposition 2.3. *With Assumption* 1 *holding x^* is a local solution to* (P). *Furthermore, there exists a neighborhood of (x^*,λ^*) such that (x^*,λ^*) is the unique solution of* (2.2) *in this neighborhood.*

The mapping $x \mapsto L_c(x, \lambda^*)$ can be bounded from below by a quadratic function. This fact is referred to as augmentability of L_c and is formulated in the next proposition. For a proof we refer the reader to [9].

Proposition 2.4. *There exist a neighborhood $\hat{U}$ of x^* and a constant $\bar{c} \geq 0$ such that the mapping $x \mapsto L_c''(x, \lambda^*)$ is coercive on the whole space X for all $x \in \hat{U}$ and $c \geq \bar{c}$.*

Remark 2.5. Due to Assumption 1 and Proposition 2.4 there are convex neighborhoods $U(x^*) \subset X$ of x^* and $U(\lambda^*) \subset Y$ of λ^* such that for all $(x, \lambda) \in U = U(x^*) \times U(\lambda^*)$

a) $J(x)$ and $e(x)$ are twice Fréchet-differentiable and their second Fréchet-derivatives are Lipschitz-continuous in $\overline{U(x^*)}$,

b) $e'(x)$ is surjective,

c) $L_0''(x, \lambda)$ is coercive on the kernel of $e'(x)$,

d) the point $z^* = (x^*, \lambda^*)$ is the unique solution to (2.2) in U, and

e) there exist $\tilde{\kappa} > 0$ and $\bar{c} \geq 0$ such that

$$\langle L_c''(x, \lambda)\chi, \chi \rangle_X \geq \tilde{\kappa} \, \|\chi\|_X^2 \quad \text{for all } \chi \in X \text{ and } c \geq \bar{c}. \; \Diamond \tag{2.3}$$

To shorten notation let us introduce the operator

$$F_c(x, \lambda) = \begin{pmatrix} L_c'(x, \lambda) \\ e(x) \end{pmatrix} \quad \text{for all } (x, \lambda) \in U.$$

Then the first-order necessary optimality conditions (2.2) can be expressed as

$$F_c(x^*, \lambda^*) = 0 \quad \text{for all } c \geq 0. \tag{OS}$$

To find x^* numerically we solve (OS) by the Newton method. The Fréchet-derivative of the operator F_c in U is given by

$$\nabla F_c(x, \lambda) = \begin{pmatrix} L_c''(x, \lambda) & e'(x)^\star \\ e'(x) & 0 \end{pmatrix}, \tag{2.4}$$

where $e'(x)^\star : Y \to X$ denotes the adjoint of the operator $e'(x)$.

Remark 2.6. With Assumptions 1 holding there exists a constant $C > 0$ satisfying

$$\|\nabla F_c(x, \lambda)^{-1}\|_{\mathcal{B}(Z)} \leq C \quad \text{for all } (x, \lambda) \in U \tag{2.5}$$

(see e.g. in [7, p. 114]), where $\mathcal{B}(Z)$ denotes the Banach space of all bounded linear operators on Z. $\Diamond$

Now we formulate the augmented Lagrangian-SQP method.

Algorithm 1. a) *Choose* $(x^0, \lambda^0) \in U$, $c \geq 0$ *and put* $k = 0$.

b) *Set* $\tilde{\lambda}^k = \lambda^k + ce(x^k)$.

c) *Solve for* $(\Delta x, \Delta\lambda)$ *the linear system*

$$\nabla F_0(x^k, \tilde{\lambda}^k) \begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix} = -F_0(x^k, \tilde{\lambda}^k). \tag{2.6}$$

d) *Set* $(x^{k+1}, \lambda^{k+1}) = (x^k + \Delta x, \tilde{\lambda}^k + \Delta\lambda)$, $k = k+1$ *and go back to* b).

Remark 2.7. Since X and Y are Hilbert spaces, (x^{k+1}, λ^{k+1}) can equivalently be obtained from solving the linear system

$$\nabla F_c(x^k, \lambda^k) \begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix} = -F_c(x^k, \lambda^k) \tag{2.7}$$

and setting $(x^{k+1}, \lambda^{k+1}) = (x^k + \Delta x, \lambda^k + \Delta\lambda)$. Equation (2.7) corresponds to a Newton step applied to (OS). This form of the iteration requires the implementation of $e'(x^k)^\star e'(x^k)$, whereas steps b) and c) of Algorithm 1 do not. In case of Example 2.1 this implies an additional solve of the Poisson equation. ♢

3 Affine invariance

Let $\tilde{B} : X \to X$ be an arbitrary isomorphism. We transform the x variable by $x = \tilde{B}y$. Thus, instead of (P) we study the whole class of equivalent transformed minimization problems

$$\text{minimize } J(\tilde{B}y) \text{ subject to } e(\tilde{B}y) = 0 \tag{3.1}$$

with the transformed solutions $\tilde{B}y^* = x^*$. Setting

$$B = \begin{pmatrix} \tilde{B} & 0 \\ 0 & I \end{pmatrix} \quad \text{and} \quad G_c(y, \xi) = B^\star F_c(x, \lambda) \quad \text{with } (x, \lambda) = (\tilde{B}y, \xi),$$

the first-order necessary optimality conditions have the form

$$G_c(y, \xi) = 0 \quad \text{for all } c \geq 0. \tag{$\widehat{\text{OS}}$}$$

Applying Algorithm 1 to ($\widehat{\text{OS}}$) we get an equivalent sequence of transformed iterates.

Theorem 3.1. *Let* $(x^0, \lambda^0) \in U$ *and* $(y^0, \xi^0) = (\tilde{B}^{-1}x^0, \lambda^0)$ *be the starting iterates for Algorithm* 1 *applied to the optimality conditions* (OS) *and* ($\widehat{\text{OS}}$), *respectively. Then both sequences of iterates are well-defined and equivalent in the sense of*

$$(\tilde{B}y^k, \xi^k) = (x^k, \lambda^k) \quad \text{for } k = 0, 1, \dots \,. \tag{3.2}$$

Proof. First note that the Fréchet-derivative of the operator G_c is given by

$$\nabla G_c(y,\xi) = B^\star \nabla F_c(x,\lambda) B \quad \text{with } (x,\lambda) = (\tilde{B}y, \xi). \tag{3.3}$$

To prove (3.2) we use an induction argument. By assumption the identity (3.2) holds for $k = 0$. Now suppose that (3.2) is satisfied for $k \geq 0$. This implies $\tilde{B}y^k = x^k$ and $\xi^k = \lambda^k$. Using step b) of Algorithm 1 it follows that $\tilde{\xi}^k = \xi^k + ce(\tilde{B}y^k) = \tilde{\lambda}^k$. From (3.3),

$$\nabla F_0(x^k, \tilde{\lambda}^k) \begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix} = -F_0(x^k, \tilde{\lambda}^k) \text{ and } \nabla G_0(y^k, \tilde{\xi}^k) \begin{pmatrix} \Delta y \\ \Delta\xi \end{pmatrix} = -G_0(y^k, \tilde{\xi}^k)$$

we conclude that $(\Delta y, \Delta\xi) = (\tilde{B}^{-1}\Delta x, \Delta\lambda)$. Utilizing step d) of Algorithm 1 we get the desired result. □

Remark 3.2. Due to the previous theorem the augmented Lagrangian-SQP method is *invariant under arbitrary transformations* $\tilde{B}$ of the state space X. This nice property should, of course, be inherited by any convergence theory and termination criteria. In section 4 we develop such an invariant theory. ◊

Remark 3.3. The invariance of Newton's method is not limited to transformations of type (3.1). In fact, Newton's method is invariant under arbitrary transformations of domain and image space, i.e., it behaves exactly the same for $AF_c(B\tilde{z}) = 0$ as for $F_c(z) = 0$. Because F_c has a special gradient structure in the optimization context, *meaningful* transformations are coupled due to the chain rule. Meaningful transformations result from transformations of the underlying optimization problem, i.e., transformations of the domain space and the image space of the constraints. Those are of the type

$$\begin{pmatrix} B_1^\star & 0 \\ 0 & B_2^\star \end{pmatrix} F\left(\begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{\lambda} \end{pmatrix} \right).$$

For such general transformations there is no possibility to define a norm in an invariant way, since both the domain and the image space of the constraints are transformed independently: $B_2^\star e(B_1\tilde{x})$. ◊

4 Affine invariant convergence theory

To formulate the convergence theory and termination criteria in terms of an appropriate norm, we use a norm that is invariant under the transformation (3.1).

Definition 4.1. *Let $z \in U$. Then a norm $\|\cdot\|_z : Z \to \mathbb{R}$ is called* affine invariant *for* (OS), *if*

$$\|\nabla F_c(\tilde{z})\Delta z\|_z = \|\nabla G_c(B^{-1}\tilde{z})B^{-1}\Delta z\|_{B^{-1}z} \quad \text{for all } \tilde{z} \in U \text{ and } \Delta z \in Z. \tag{4.1}$$

We call $\{\|\cdot\|_z\}_{z\in U}$ *a* γ-continuous family of invariant norms *for* (OS), *if*

$$\left|\|r\|_{z+\Delta z} - \|r\|_z\right| \leq \gamma \; \|\nabla F_c(z)\Delta z\|_z \|r\|_z \tag{4.2}$$

for every $r, \Delta z \in Z$ *and* $z \in U$ *such that* $z + \Delta z \in U$.

Utilizing affine invariant norms we are able to present an affine invariant convergence theorem for Algorithm 1.

Theorem 4.2. *Assume that there are constants* $\omega \geq 0, \gamma \geq 0$, *and a* γ*-continuous family of affine invariant norms* $\{\|\cdot\|_z\}_{z\in U}$, *such that the operator* ∇F_c *satisfies*

$$\|(\nabla F_c(z + s\Delta z) - \nabla F_c(z))\Delta z\|_{z+\Delta z} \; \leq s\omega\|\nabla F_c(z)\Delta z\|_z^2 \tag{4.3}$$

for $z \in U$ *and* $\Delta z \in Z$ *such that* $z + \Delta z \in U$. *For* $k \in \mathbb{N}$ *let* $h_k = \omega\|F_c(z^k)\|_{z^k}$ *and let*

$$\mathcal{L}(z) = \left\{\zeta \in U : \|F_c(\zeta)\|_\zeta \leq \left(1 + \frac{\gamma}{4}\|F_c(z)\|_z\right)\|F_c(z)\|_z\right\}. \tag{4.4}$$

Suppose that $h_0 < 2$ *and that the level set* $\mathcal{L}(z^0)$ *is closed. Then, the iterates stay in* U *and the residuals converge to zero at a rate of*

$$h_{k+1} \leq \frac{1}{2}\; h_k^2.$$

Additionally, we have

$$\|F_c(z^{k+1})\|_{z^k} \leq \|F_c(z^k)\|_{z^k}. \tag{4.5}$$

Proof. By induction, assume that $\mathcal{L}(z^k)$ is closed and that $h_k < 2$ for $k \geq 0$. Due to Remark 2.5 the neighborhood U is assumed to be convex, so that $z + \eta\Delta z \in U$ for all $\eta \in [0,1]$. From $\nabla F_c(z^k)\Delta z^k = -F_c(z^k)$ we conclude that

$$\begin{aligned} F_c(z^k + \eta\Delta z^k) &= F_c(z^k) + \int_0^\eta \nabla F_c(z^k + s\Delta z^k)\Delta z^k \, ds \\ &= (1-\eta)F_c(z^k) + \int_0^\eta (\nabla F_c(z^k + s\Delta z^k) - \nabla F_c(z^k))\Delta z^k \, ds \end{aligned}$$

for all $\eta \in [0,1]$. Applying (4.2), (4.3), $h_k = \omega\|F_c(z^k)\|_{z^k}$ and $h_k < 2$ we obtain

$$\begin{aligned} &\|F_c(z^k + \eta\Delta z^k)\|_{z^k+\eta\Delta z^k} \\ &\qquad \leq (1-\eta)(1 + \eta\gamma\|F_c(z^k)\|_{z^k})\|F_c(z^k)\|_{z^k} + \int_0^\eta s\omega\|\nabla F_c(z^k)\Delta z^k\|_{z^k}^2 \, ds \\ &\qquad = \left((1-\eta)(1 + \eta\gamma\|F_c(z^k)\|_{z^k}) + \frac{\eta^2 h_k}{2}\right)\|F_c(z^k)\|_{z^k} \\ &\qquad < \left(1 + (\eta - \eta^2)\gamma\|F_c(z^k)\|_{z^k}\right)\|F_c(z^k)\|_{z^k} \\ &\qquad \leq \left(1 + \frac{\gamma}{4}\|F_c(z^k)\|_{z^k}\right)\|F_c(z^k)\|_{z^k}. \end{aligned}$$

If $z^k + \Delta z^k \notin \mathcal{L}(z^k)$, there exists an $\bar{\eta} \in [0,1]$ such that $z^k + \bar{\eta}\Delta z^k \in U \backslash \mathcal{L}(z^k)$, i.e.,

$$\|F_c(z^k + \bar{\eta}\Delta z^k)\|_{z^k+\bar{\eta}\Delta z^k} > \left(1 + \frac{\gamma}{4}\|F_c(z^k)\|_{z^k}\right)\|F_c(z^k)\|_{z^k},$$

which is a contradiction. Hence, $z^{k+1} \in \mathcal{L}(z^k)$ and

$$\|F(z^{k+1})\|_{z^{k+1}} \leq \frac{\omega}{2}\|F_c(z^k)\|_{z^k}^2.$$

Thus, we have $h_{k+1} \leq h_k^2/2$ and $\mathcal{L}(z^{k+1}) \subset \mathcal{L}(z^k)$. Since $\mathcal{L}(z^k)$ is closed, every Cauchy sequence in $\mathcal{L}(z^{k+1})$ converges to a limit point in $\mathcal{L}(z^k)$, which is, by (4.4) and the continuity of the norm, also contained in $\mathcal{L}(z^{k+1})$. Hence, $\mathcal{L}(z^{k+1})$ is closed. □

Remark 4.3. We choose simplicity over sharpness here. The definition of the level set $\mathcal{L}(z)$ can be sharpened somewhat by a more careful estimate of the term

$$(\gamma\|F_c(z^k)\|_{z^k} - 1)\eta + (h_k/2 - \gamma\|F_c(z^k)\|_{z^k})\eta^2. \ \Diamond$$

Theorem 4.2 guarantees that $\lim_{k\to\infty} h_k = 0$. To ensure that $z^k \to z^*$ in Z as $k \to \infty$ we have to require a further property of the invariant norm.

Corollary 4.4. *If, in addition to the assumptions of Theorem* 4.2, *there exists a constant* $\tilde{C} > 0$ *such that*

$$\|\zeta\|_Z \leq \tilde{C}\,\|\nabla F_c(z)\zeta\|_z \quad \textit{for all } \zeta \in Z \textit{ and } z \in U,$$

then the iterates converge to the solution $z^* = (x^*, \lambda^*)$ *of* (OS).

Proof. By assumption and Theorem 4.2 we have

$$\|\Delta z^k\|_Z \leq \tilde{C}\,\|F_c(z^k)\|_{z^k} \leq \tilde{C}\left(\frac{h_0}{2}\right)^k \|F_c(z^0)\|_{z^0}.$$

Thus, $\{z^k\}_{k\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{L}(z^0) \subset U$. Since $\mathcal{L}(z^0)$ is closed, the claim follows by Remark 2.5-d). □

For actual implementation of Algorithm 1 we need a *convergence monitor* indicating whether or not the assumptions of Theorem 4.2 may be violated, and a *termination criterion* deciding whether or not the desired accuracy has been achieved.
From (4.5), a new iterate z^{k+1} is accepted, whenever

$$\|F_c(z^{k+1})\|_{z^k} < \|F_c(z^k)\|_{z^k}. \tag{4.6}$$

Otherwise, the assumptions of Theorem 4.2 are violated and the iteration is considered as to be non-convergent. The use of the norm $\|\cdot\|_{z^k}$ for both the old and the new iterate permits an efficient implementation. Since in many cases the norm $\|F_c(z^{k+1})\|_{z^k}$ is defined in terms of $\overline{\Delta z^{k+1}} = \nabla F_c(z^k)^{-1}F_c(z^{k+1})$, the derivative

need not be evaluated at the new iterate. If a factorization of $\nabla F_c(z^k)$ is available via a direct solver, it can be reused at negligible cost even if the convergence test fails. If an iterative solver is used, $\overline{\Delta z^{k+1}}$ in general provides a good starting point for computing Δz^{k+1}, such that the additional cost introduced by the convergence monitor is minor.
The SQP iteration will be terminated with a solution z^{k+1} as soon as

$$\|F_c(z^{k+1})\|_{z^k} \leq \text{TOL } \|F_c(z^0)\|_{z^0}$$

with a user specified tolerance TOL. Again, the use of the norm $\|\cdot\|_{z^k}$ allows an efficient implementation.

5 Invariant norms for optimization problems

What remains to be done is the construction of a γ-continuous family of invariant norms. In this section we introduce two different norms.

5.1 First invariant norm

The first norm takes advantage of the parameter c in the augmented Lagrangian. As we mentioned in Remark 2.5, there exists a $\bar{c} \geq 0$ such that $L''_c(z)$ is coercive on X for all $z \in U$ and $c \geq \bar{c}$. Hence, the operator $L''_c(z)^{-1}$ belongs to $\mathcal{B}(Z)$ for all $c \geq \bar{c}$.
Let us introduce the operator $S_c : U \to \mathcal{B}(Z)$ by

$$S_c(z) = \begin{pmatrix} L''_c(z) & 0 \\ 0 & I \end{pmatrix} \quad \text{for all } z \in U \text{ and } c \geq 0. \tag{5.1}$$

Since $L''_c(z)$ is self-adjoint for all $z \in U$, $S_c(z)$ is self-adjoint as well. Due to (2.3) the operator $S_c(z)$ is coercive for all $z \in U$ and $c \geq \bar{c}$. Thus, for all $z \in U$

$$\|S_c^{1/2}(z) \cdot \| = \sqrt{\langle S_c(z) \cdot , \cdot \rangle_Z} \tag{5.2}$$

is a norm on Z for $c \geq \bar{c}$.

Proposition 5.1. *Let $c \geq \bar{c}$. Then, for every $z \in U$ the mapping*

$$\|r\|_z = \|S_c(z)^{1/2} \nabla F_c(z)^{-1} r\| \quad \textit{for } r \in Z \tag{5.3}$$

defines an affine invariant norm for (2.2).

Proof. Let $z \in U$ be arbitrary. Since $\|S_c^{1/2}(z) \cdot \|$ defines a norm on Z for $c \geq \bar{c}$ and $\nabla F_c(z)$ is continuously invertible by Remark 2.6, it follows that $\|\cdot\|_z$ is a norm on Z. Now we prove the invariance property (4.1). Let $\tilde{L}_c$ denote the augmented Lagrangian associated with the transformed problem (3.1). Then we have $\tilde{L}''_c(\zeta) = \tilde{B}^\star L''_c(z)\tilde{B}$ for $z = B\zeta \in U$. Hence, setting $\tilde{S}_c(\zeta) = B^\star S_c(z) B$ we get

$$\|r\|_\zeta = \|\tilde{S}_c(\zeta)^{1/2} \nabla G_c(\zeta)^{-1} r\| \quad \text{for } r \in Z.$$

From (3.3) we conclude that

$$\nabla F_c(z)^{-1}\nabla F_c(\tilde z) = B\nabla G_c(\zeta)^{-1}\nabla G_c(\tilde\zeta)B^{-1} \tag{5.4}$$

with $z = B\zeta, \tilde z = B\tilde\zeta \in U$. Using (5.3) and (5.4) we obtain

$$\begin{aligned}\|\nabla F_c(\tilde z)\delta z\|_z &= \|S_c(z)^{1/2}\nabla F_c(z)^{-1}\nabla F_c(\tilde z)\delta z\| \\ &= \|S_c(z)^{1/2}B\nabla G_c(\zeta)^{-1}\nabla G_c(\tilde\zeta)B^{-1}\delta z\| \\ &= \|(B^\star S_c(z)B)^{1/2}\nabla G_c(\zeta)^{-1}\nabla G_c(\tilde\zeta)B^{-1}\delta z\| \\ &== \|\nabla G_c(\tilde\zeta)B^{-1}\delta z\|_\zeta,\end{aligned}$$

which gives the claim. □

In order to show the γ-continuity (4.2) required for Theorem 4.2, we need the following lemma.

Lemma 5.2. *Suppose that $c \geq \bar c$ and that there exists a constant $\omega \geq 0$ such that*

$$\|(\nabla F_c(z+\delta z) - \nabla F_c(z))\zeta\|_{z+\delta z} \leq \omega\ \|\nabla F_c(z)\delta z\|_z \|\nabla F_c(z)\zeta\|_z \tag{5.5}$$

for all $\zeta \in Z$, $z \in U$ and $\delta z \in Z$ such that $z + \delta z \in U$. Then we have

$$\|S_c(z+\delta z)^{1/2}\zeta\| \leq \sqrt{1+\omega(1+C_e)\ \|\nabla F_c(x,\lambda)\delta z\|_z}\ \|S_c(z)^{1/2}\zeta\|,$$

where

$$C_e = \sup\left\{\frac{\|e'(x)\xi\|_Y^2}{\langle L_c''(x,\lambda)\xi,\xi\rangle_X} : (x,\lambda) \in \overline{U}, \xi \in X\backslash\{0\}\right\} > 0.$$

Proof. Let $\zeta = (\zeta_1,\zeta_2)^\mathsf{T} \in Z$ and $z \in U$. From (5.1) and (5.2) we infer

$$\begin{aligned}\|S_c(z+\delta z)^{1/2}\zeta\|^2 &= \langle S_c(z+\delta z)\zeta,\zeta\rangle_Z \\ &= \langle S_c(z)\zeta,\zeta\rangle_Z + \langle (S_c(z+\delta z) - S_c(z))\zeta,\zeta\rangle_Z \\ &\leq \|S_c(z)^{1/2}\zeta\|^2 + \langle (L_c''(z+\delta z) - L_c''(z))\zeta_1,\zeta_1\rangle_X.\end{aligned} \tag{5.6}$$

By assumption $S_c(z)$ is continuously invertible. Utilizing the Lipschitz assumption (5.5) the second additive term on the right-hand side can be estimated as

$$
\begin{aligned}
&\langle (L_c''(z+\delta z) - L_c''(z))\zeta_1, \zeta_1\rangle_X \\
&\quad = \langle (\nabla F_c(z+\delta z) - \nabla F_c(z))(\zeta_1, 0)^\mathsf{T}, (\zeta_1, 0)^\mathsf{T}\rangle_Z \\
&\quad = \langle \nabla F_c(z) S_c(z)^{-1} S_c(z) \nabla F_c(z)^{-1} (\nabla F_c(z+\delta z) - \nabla F_c(z))(\zeta_1, 0)^\mathsf{T}, (\zeta_1, 0)^\mathsf{T}\rangle_Z \\
&\quad = \langle S_c(z) \nabla F_c(z)^{-1} (\nabla F_c(z+\delta z) - \nabla F_c(z))(\zeta_1, 0)^\mathsf{T}, S_c(z)^{-1} \nabla F_c(z) (\zeta_1, 0)^\mathsf{T}\rangle_Z \\
&\quad \le \|S_c(z)^{1/2} \nabla F_c(z)^{-1} (\nabla F_c(z+\delta z) - \nabla F_c(z))(\zeta_1, 0)^\mathsf{T}\| \\
&\qquad\quad \cdot \|S_c(z)^{-1/2} \nabla F_c(z) (\zeta_1, 0)^\mathsf{T}\| \\
&\quad = \|(\nabla F_c(z+\delta z) - \nabla F_c(z))(\zeta_1, 0)^\mathsf{T}\|_z \|S_c(z)^{-1/2} \nabla F_c(z)(\zeta_1, 0)^\mathsf{T}\| \\
&\quad \le \omega \ \|\nabla F_c(z)\delta z\|_z \ \|\nabla F_c(z)(\zeta_1, 0)^\mathsf{T}\|_z \ \|S_c(z)^{-1/2} \nabla F_c(z)(\zeta_1, 0)^\mathsf{T}\| \\
&\quad \le \omega \ \|\nabla F_c(z)\delta z\|_z \ \|S_c(z)^{1/2}\zeta\| \ \|S_c(z)^{-1/2} \nabla F_c(z)(\zeta_1, 0)^\mathsf{T}\|.
\end{aligned}
$$

Note that

$$
\begin{aligned}
&\|S_c(z)^{-1/2} \nabla F_c(z)(\zeta_1, 0)^\mathsf{T}\|^2 \\
&\quad = \langle \nabla F_c(z)(\zeta_1, 0)^\mathsf{T}, S_c(z)^{-1} \nabla F_c(z)(\zeta_1, 0)^\mathsf{T}\rangle_Z = \langle L_c''(z)\zeta_1, \zeta_1\rangle_X + \|e'(x)\zeta_1\|_Y^2 \\
&\quad \le (1 + C_e) \ \langle L_c''(z)\zeta_1, \zeta_1\rangle_X = (1 + C_e) \ \|S_c(z)^{1/2}(\zeta_1, 0)^\mathsf{T}\|^2.
\end{aligned}
$$

This implies

$$
\langle (L_c''(z+\delta z) - L_c''(z))\zeta_1, \zeta_1\rangle_X \le \omega(1 + C_e) \ \|\nabla F_c(z)\delta z\|_z \ \|S_c(z)^{1/2}\zeta\|^2. \tag{5.7}
$$

Inserting (5.7) into (5.6) the claim follows. □

Proposition 5.3. *Let all hypotheses of Lemma* 5.2 *be satisfied. Then* $\{\|\cdot\|_z\}_{z\in U}$ *is a* $\omega(3 + C_e)/2$*-continuous family of invariant norms with*

$$
\|\zeta\|_Z \le \frac{1}{\sqrt{\tilde\kappa}} \ \|\nabla F_c(z)\zeta\|_z \tag{5.8}
$$

for all $\zeta \in Z$ *and* $z \in U$*, where* $\tilde\kappa > 0$ *was introduced in* (2.3).

Proof. From (5.3) it follows that

$$
\begin{aligned}
\|r\|_{z+\delta z} \le \ &\|S_c(z+\delta z)^{1/2} \nabla F_c(z)^{-1} r\| \\
&+ \|S_c(z+\delta z)^{1/2} (\nabla F_c(z+\delta z)^{-1} - \nabla F_c(z)^{-1}) r\|.
\end{aligned}
$$

We estimate the additive terms on the right-hand side separately. Using Lemma 5.2 we find

$$
\|S_c(z+\delta z)^{1/2} \nabla F_c(z)^{-1} r\| \le \sqrt{1 + \omega(1 + C_e)\|\nabla F_c(z)\delta z\|_z} \|r\|_z.
$$

Applying (5.3) and (5.5) we obtain

$$\begin{aligned}
&\|S_c(z+\delta z)^{1/2}(\nabla F_c(z+\delta z)^{-1} - \nabla F_c(z)^{-1})r\| \\
&\quad = \|S_c(z+\delta z)^{1/2}\nabla F_c(z+\delta z)^{-1}(\nabla F_c(z) - \nabla F_c(z+\delta z))\nabla F_c(z)^{-1}r\| \\
&\quad = \|(\nabla F_c(z) - \nabla F_c(z+\delta z))\nabla F_c(z)^{-1}r\|_{z+\delta z} \le \omega\ \|\nabla F_c(z)\delta z\|_z \|r\|_z.
\end{aligned}$$

Hence,

$$\|r\|_{z+\delta z} \le \Big(1 + \frac{\omega}{2}(3 + C_e)\|\nabla F_c(z)\delta z\|_z\Big)\|r\|_z$$

and it follows that $\{\|\cdot\|_z\}_{z\in U}$ is a $\omega(3+C_e)/2$-continuous family of invariant norms. Finally, from

$$\begin{aligned}
\|\nabla F_c(z)\zeta\|_z^2 &= \langle S_c(z)\nabla F_c(z)^{-1}\nabla F_c(z)\zeta, \nabla F_c(z)^{-1}\nabla F_c(z)\zeta\rangle_Z \\
&= \langle S_c(z)\zeta,\zeta\rangle_Z \ge \tilde{\kappa}\ \|\zeta\|_Z^2
\end{aligned}$$

we infer (5.8). □

5.2 Second invariant norm

In section 5.1 we introduced an invariant norm provided the augmentation parameter in Algorithm 1 satisfies $c \ge \bar{c}$. But in many applications the constant $\bar{c}$ is not explicitly known. Thus, $L_c''(x,\lambda)^{-1}$ need not to be bounded for $c \in [0,\bar{c})$, so that $S_c(x,\lambda)$ given by (5.1) might be singular. To overcome this difficulties we define a second invariant norm that is based on a splitting $X = \ker e'(x) \oplus \bar{X}$, such that at least the coercivity of $L_0''(x,\lambda)$ on $\ker e'(x)$ can be utilized. For that purpose let us introduce the bounded linear operator $T_c(x,\lambda) : \ker e'(x) \times Y \times Y \to Z$ by

$$T_c(x,\lambda) = \begin{pmatrix} L_c''(x,\lambda) & e'(x)^\star & 0 \\ e'(x) & 0 & I \end{pmatrix} \quad \text{for } (x,\lambda) \in U \text{ and } c \ge 0.$$

Lemma 5.4. *For every $(x,\lambda) \in U$ and $c \ge 0$ the operator $T_c(x,\lambda)$ is an isomorphism.*

Proof. Let $r = (r_1, r_2)^\mathsf{T} \in Z$ be arbitrary. Then the equation $T_c(x,\lambda)\zeta = r$ for $\zeta = (\zeta_1,\zeta_2,\zeta_3)^\mathsf{T} \in \ker e'(x) \times Y \times Y$ is equivalent with

$$\nabla F_c(x,\lambda)\begin{pmatrix}\zeta_1\\ \zeta_2\end{pmatrix} = \begin{pmatrix} r_1 \\ 0\end{pmatrix} \quad \text{and} \quad \zeta_3 = r_2. \tag{5.9}$$

Due to Remark 2.6 the operator $\nabla F_c(x,\lambda)$ is continuously invertible for all $(x,\lambda) \in U$ and $c \ge 0$. Thus, ζ is uniquely determined by (5.9), and the claim follows. □

We define the bounded linear operator $R_c(x,\lambda) : \ker e'(x) \times Y \times Y \to Z \times Y$ as

$$R_c(x,\lambda) = \begin{pmatrix} L_c''(x,\lambda) & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{pmatrix} \quad \text{for } (x,\lambda) \in U \text{ and } c \ge 0. \tag{5.10}$$

Note that $R_c(x,\lambda)$ is coercive and self-adjoint. Next we introduce the invariant norm

$$\|r\|_z = \sqrt{\langle R_c(z)T_c(z)^{-1}r, T_c(z)^{-1}r\rangle_{Z\times Y}} \quad \text{for } z \in U \text{ and } r \in Z. \tag{5.11}$$

To shorten notation, we write $\|r\|_z = \|R_c(z)^{1/2}T_c(z)^{-1}r\|$.

Proposition 5.5. *For every $z \in U$ the mapping given by* (5.11) *is an affine invariant norm for* (OS). *Moreover, there exists a constant $\bar{C} > 0$ such that*

$$\|\zeta\|_Z \le \bar{C}\ \|\nabla F_c(z)\zeta\|_z \quad \textit{for all } \zeta \in Z \textit{ and } z \in U. \tag{5.12}$$

Proof. Let $z \in U$ be arbitrary. Since $R_c(z)$ is coercive and $T_c(z)$ is continuously invertible, it follows that $\|\cdot\|_z$ defines a norm on Z. Now we prove the invariance property (4.1). For $(x,\lambda) = (\tilde{B}y,\xi) \in U$ we have

$$\begin{pmatrix} \tilde{B}^\star L_c''(y,\xi)\tilde{B} & \tilde{B}^\star e'(y)^\star & 0 \\ e'(y)B & 0 & I \end{pmatrix} = B^\star T_c(x,\lambda)\begin{pmatrix} B & 0 \\ 0 & I \end{pmatrix}. \tag{5.13}$$

Utilizing (3.3), (5.11) and (5.13) the invariance property follows. Finally, setting $\xi = T_c(z)^{-1}\nabla F_c(z)\zeta$ we conclude

$$\begin{aligned} \|\nabla F_c(z)\zeta\|_z^2 &= \langle R_c(z)T_c(z)^{-1}\nabla F_c(z)\zeta, T_c(z)^{-1}\nabla F_c(z)\zeta\rangle_{\ker e'(x)\times Y\times Y} \\ &= \langle R_c(z)\xi,\xi\rangle_{\ker e'(x)\times Y\times Y} \ge c_1\ \|\xi\|^2_{\ker e'(x)\times Y\times Y} \end{aligned}$$

for some $c_1 > 0$ from the coercivity of $R_c(z)$. Since both $\nabla F_c(z)$ and $T_c(z)$ are isomorphisms by Remark 2.6 and Lemma 5.4, there is some $c_2 > 0$ such that

$$\|\xi\|_{\ker e'(x)\times Y\times Y} \ge c_2\ \|\zeta\|_Z,$$

and (5.12) follows. □

The following proposition guarantees that $\{\|\cdot\|_z\}_{z\in U}$ is a γ-continuous family of invariant norms for (OS).

Proposition 5.6. *Suppose that there exists a constant $\omega \ge 0$ such that*

$$\|(\nabla F_c(z+\delta z) - \nabla F_c(z))\zeta\|_{z+\delta z} \le \omega\ \|\nabla F_c(z)\delta z\|_z \|\nabla F_c(z)\zeta\|_z \tag{5.14}$$

for all $\zeta \in Z$, $z \in U$ and $\delta z \in Z$ such that $z + \delta z \in U$. Then we have

$$\|r\|_{z+\delta z} \le \left(1 + \frac{3\omega}{2}\ \|\nabla F_c(z)\delta z\|_z\right)\|r\|_z.$$

For the proof of the previous proposition, we will use the following lemmas.

Lemma 5.7. *With the assumption of Proposition* 5.6 *holding and $z = (x,\lambda)$ it follows that*

$$\|R_c(z+\delta z)^{1/2}\zeta\| \le \sqrt{1 + \omega\|\nabla F_c(z)\delta z\|_z}\|R_c(z)^{1/2}\zeta\|$$

for all $\zeta \in \ker e'(x) \times Y \times Y$ and $c \ge 0$.

Proof. Let $z = (x, \lambda) \in U$ and $\zeta = (\zeta_1, \zeta_2, \zeta_3)^\top \in \ker e'(x) \times Y \times Y$. Using (5.10) and (5.11) we obtain

$$\|R_c(z+\delta z)^{1/2}\zeta\|^2 \le \|R_c(z)^{1/2}\zeta\|^2 + \langle (L_c''(z+\delta z) - L_c''(z))\zeta_1, \zeta_1\rangle_X. \tag{5.15}$$

For all $c \ge 0$ the operator $R_c(z)$ is continuously invertible. Furthermore, $R_c(z)$ is self-adjoint. Thus, applying (5.14) and

$$\nabla F_c(z)(\zeta_1, 0)^\top = T(z)^\star(\zeta_1, 0)^\top = R_c(z)(\zeta_1, 0, 0)^\top$$

the second additive term on the right-hand side of (5.15) can be estimated as

$$\begin{aligned}
&\langle (L_c''(z+\delta z) - L_c''(z))\zeta_1, \zeta_1\rangle_X \\
&= \langle (\nabla F_c(z+\delta z) - \nabla F_c(z))(\zeta_1, 0)^\top, (\zeta_1, 0)^\top\rangle_Z \\
&= \langle T_c(z) R_c(z)^{-1} R_c(z) T_c(z)^{-1} (\nabla F_c(z+\delta z) - \nabla F_c(z))(\zeta_1, 0)^\top, (\zeta_1, 0)^\top\rangle_Z \\
&= \langle R_c(z) T_c(z)^{-1} (\nabla F_c(z+\delta z) - \nabla F_c(z))(\zeta_1, 0)^\top, R_c(z)^{-1} T_c(z)^\star (\zeta_1, 0)^\top\rangle_{Z\times Y} \\
&\le \|R_c(z)^{1/2} T_c(z)^{-1} (\nabla F_c(z+\delta z) - \nabla F_c(z))(\zeta_1, 0)^\top\| \; \|R_c(z)^{-1/2} T_c(z)^\star (\zeta_1, 0)^\top\| \\
&\le \omega \|\nabla F_c(z)\delta z\|_z \|\nabla F_c(z)(\zeta_1, 0)^\top\|_z \|R_c(z)^{1/2}(\zeta_1, 0, 0)^\top\| \\
&\le \omega \|\nabla F_c(z)\delta z\|_z \|R_c(z)^{1/2}\zeta\|^2.
\end{aligned}$$

Inserting this bound in (5.15) the claim follows. □

Lemma 5.8. *Let the assumptions of Theorem* 5.6 *are satisfied. Then*

$$\|(T_c(z+\delta z) - T(z))T_c(z)^{-1} r\|_{z+\delta z} \le \omega \|\nabla F_c(z)\delta z\|_z \|r\|_z \quad \textit{for all } r \in Z.$$

Proof. For arbitrary $r = (r_1, r_2)^\top \in Z$ we set $\zeta = (\zeta_1, \zeta_2, \zeta_3)^\top = T_c(z)^{-1} r$. Using (5.9) and (5.14) we estimate

$$\begin{aligned}
&\|(T_c(z+\delta z) - T_c(z))T_c(z)^{-1} r\|_{z+\delta z} \\
&= \|(\nabla F_c(z+\delta z) - \nabla F_c(z))(\zeta_1, \zeta_2)^\top\|_{z+\delta z} \le \omega \|\nabla F_c(z)\delta z\|_z \|\nabla F_c(z)(\zeta_1, \zeta_2)^\top\|_z \\
&= \omega \|\nabla F_c(z)\delta z\|_z \|(r_1, 0)\|_z \le \omega \|\nabla F_c(z)\delta z\|_z \|r\|_z
\end{aligned}$$

so that the claim follows. □

Proof of Proposition 5.6. Let $z, z+\delta z \in U$. Utilizing (5.11), Lemmas 5.7 and 5.8 we find

$$\begin{aligned}
\|r\|_{z+\delta z} &= \|R_c(z+\delta z)^{1/2} T_c(z+\delta z)^{-1} r\| \\
&\le \|R_c(z+\delta z)^{1/2} T_c(z)^{-1} r\| + \|R_c(z+\delta z)^{1/2}(T_c(z+\delta z)^{-1} - T_c(z)^{-1})r\| \\
&\le \sqrt{1 + \omega \|\nabla F_c(z)\delta z\|_z}\, \|r\|_z + \|(T_c(z) - T_c(z+\delta z))T_c(z)^{-1} r\|_{z+\delta z} \\
&\le \sqrt{1 + \omega \|\nabla F_c(z)\delta z\|_z}\, \|r\|_z + \omega \|\nabla F_c(z)\delta z\|_z \|r\|_z \\
&\le \Big(1 + \frac{3\omega}{2}\, \|\nabla F_c(z)\delta z\|_z\Big) \|r\|_z.
\end{aligned}$$

Hence, $\{\|\cdot\|_z\}_{z\in U}$ is a $3\omega/2$-continuous family of invariant norms. □

Remark 5.9. Note that the Lipschitz constant of the second norm does not involve C_e and hence is independent of the choice of c. In contrast, choosing c too small may lead to a large Lipschitz constant of the first norm and thus can affect the algorithm.

5.3 Connection to the optimization problem

When solving optimization problems of type (P), feasibility $e(x) = 0$ and optimality are the relevant quantities. This is well reflected by the proposed norms $\|\cdot\|_z$. Let $z = (x, \lambda)$ and $\Delta z = (\Delta x, \Delta\lambda)^\mathsf{T} = -\nabla F_c(z)^{-1} F_c(z)$. Using Taylor's theorem (see [15, p. 148]) and the continuity of L_0'', we obtain for the first norm

$$\begin{aligned}
\|F_c(z)\|_z^2 &= \langle S_c(z)\Delta z, \Delta z\rangle_Z \\
&= \langle L_c''(z)\Delta x, \Delta x\rangle_X + \|\Delta\lambda\|_Y^2 \\
&= \langle L_0''(z)\Delta x, \Delta x\rangle_X + c\|e'(x)\Delta x\|_Y^2 + \|\Delta\lambda\|_Y^2 \\
&= \langle L_0''(z^*)\Delta x, \Delta x\rangle_X + o(\|z^* - z\|_Z^2) + c\|e'(x)\Delta x\|_Y^2 + \|\Delta\lambda\|_Y^2 \\
&= 2(L_0(z) - L_0(z^*)) + o(\|z^* - z\|_Z^2) + c\|e(x)\|_Y^2 + \|\Delta\lambda\|_Y^2 \\
&= 2\big(J(x) - J(x^*) - \langle\lambda, e(x)\rangle_Y\big) + c\|e(x)\|_Y^2 + \|\Delta\lambda\|_Y^2 + o(\|z^* - z\|_Z^2).
\end{aligned}$$

The second norm is based on the partitioning $F_c(x, \lambda) = (L_c'(x, \lambda), e(x))^\mathsf{T}$ and correspondingly on a splitting of the Newton correction into a *optimizing* direction $\nabla F_c(x, \lambda)(\zeta_1, \zeta_2)^\mathsf{T} = -(L_c'(x, \lambda), 0)^\mathsf{T}$ tangential to the constraints manifold and a *feasibility* direction $\nabla F_c(x, \lambda)(\xi_1, \xi_2)^\mathsf{T} = -(0, e(x))^\mathsf{T}$. Since $e'(x)\zeta_1 = 0$, we have for $z = (x, \lambda)$

$$\begin{aligned}
\|F_c(z)\|_z^2 &= \langle L_c''(z)\zeta_1, \zeta_1\rangle_X + \|\zeta_2\|_Y^2 + \|e(x)\|_Y^2 \\
&= \langle L_0''(z)\zeta_1, \zeta_1\rangle_X + \|\zeta_2\|_Y^2 + \|e(x)\|_Y^2 \\
&= \langle L_0''(z^*)\zeta_1, \zeta_1\rangle_X + o(\|z^* - z\|_Z^2) + \|\zeta_2\|_Y^2 + \|e(x)\|_Y^2 \\
&= 2(L_0(z) - L_0(z^*)) + \|\zeta_2\|_Y^2 + \|e(x)\|_Y^2 + o(\|z^* - z\|_Z^2) \\
&= 2\big(J(x) - J(x^*) - \langle\lambda, e(x)\rangle_Y\big) + \|e(x)\|_Y^2 + \|\zeta_2\|_Y^2 + o(\|z^* - z\|_Z^2).
\end{aligned}$$

Recall that $\Delta\lambda = \zeta_2 + \xi_2$. Thus, in the proximity of the solution, both affine invariant norms measure the quantities we are interested in when solving optimization problems, in addition to the error in the Lagrange multiplier and the optimizing direction's Lagrange multiplier component, respectively.

6 Inexact augmented Lagrangian-SQP methods

Taking discretization errors or truncation errors resulting from iterative solution of linear systems into account, we have to consider inexact Newton methods, where an inner residual remains:

$$\nabla F_c(z^k)\delta z^k = -F_c(z^k) + r^k, \qquad (6.1)$$
$$z^{k+1} = z^k + \delta z^k.$$

With slightly stronger assumptions and a suitable control of the inner residual, a similar convergence theory can be established as in section 4. Note that exact affine invariance is preserved only in case the inner iteration is affine invariant, too.

Theorem 6.1. *Assume there are constants* $\omega \geq 0$, $\gamma \geq 0$, *and a* γ*-continuous family of affine invariant norms* $\{\|\cdot\|_z\}_{z\in U}$, *such that the operator* ∇F_c *satisfies*

$$\|(\nabla F_c(z+s\delta z) - \nabla F_c(z))\delta z\|_{z+\delta z} \leq s\omega\|\nabla F_c(z)\delta z\|_z^2 \qquad (6.2)$$

for $s \in [0,1]$, $z \in U$, *and* $\delta z \in Z$ *such that* $z + \delta z \in U$. *Let* $h_k = \omega\|F_c(z^k)\|_{z^k}$ *and*

$$\mathcal{L}(z) = \Big\{\zeta \in U : \|F_c(\zeta)\|_\zeta \leq \Big(1 + \frac{\gamma}{2}\|F_c(z)\|_z\Big)\|F_c(z)\|_z\Big\}.$$

Suppose that $h_0 < 2$ *and that the level set* $\mathcal{L}(z^0)$ *is closed. If the inner residual* r^k *resulting from the inexact solution of the Newton correction* (6.1) *is bounded by*

$$\|r^k\|_{z^k} \leq \delta_k \; \|F_c(z^k)\|_{z^k}, \qquad (6.3)$$

with

$$(1+\delta_k)^2\frac{h_k}{2} + \big(1 + \gamma(1+\delta_k)\|F_c(z^k)\|_{z^k}\big)\delta_k \leq \Theta_k \leq \bar{\Theta} < 1, \qquad (6.4)$$

then the iterates stay in U *and* h_k *converges to zero as* $k \to \infty$ *at a rate of*

$$h_{k+1} \leq \Theta_k h_k. \qquad (6.5)$$

Additionally,

$$\|F_c(z^{k+1})\|_{z^k} \leq \Big(\delta_k + \frac{1}{2}(1+\delta_k)^2 h_k\Big)\|F_c(z^k)\|_{z^k}. \qquad (6.6)$$

Proof. Analogously to the proof of Theorem 4.2, one obtains

$$F_c(z^k + \eta\delta z^k) = (1-\eta)F_c(z^k) + \eta r^k + \int_0^\eta (\nabla F_c(z^k + s\delta z^k) - F_c(z^k))\delta z^k \, ds$$

for all $\eta \in [0,1]$. From (6.3) we have

$$\gamma\|\nabla F_c(z^k)\delta z^k\|_{z^k} = \gamma\|F_c(z^k) - r^k\|_{z^k} \leq \gamma(1+\delta_k)\|F_c(z^k)\|_{z^k}. \qquad (6.7)$$

We set $\chi_k = \gamma(1+\delta_k)\|F_c(z^k)\|_{z^k}$. Using (4.2), (6.2), (6.3), (6.4), and (6.7) we find

$$\begin{aligned}
&\|F_c(z^k+\eta\delta z^k)\|_{z^k+\eta\delta z^k}\\
&\le (1-\eta)\|F_c(z^k)\|_{z^k+\eta\delta z^k} + \eta\|r^k\|_{z^k+\eta\delta z^k} + \int_0^\eta s\omega\|\nabla F(z^k)\delta z^k\|_{z^k}^2\,ds\\
&\le (1+\gamma\eta\|\nabla F_c(z^k)\delta z^k\|_{z^k})\Big((1-\eta)\|F_c(z^k)\|_{z^k} + \eta\|r^k\|_{z^k}\Big)\\
&\quad + \frac{\eta^2}{2}(1+\delta_k)^2 h_k\|F_c(z^k)\|_{z^k}\\
&\le (1+\eta\chi_k)\Big((1-\eta)\|F_c(z^k)\|_{z^k} + \eta\delta_k\|F_c(z^k)\|_{z^k}\Big) + \frac{\eta^2}{2}(1+\delta_k)^2 h_k\|F_c(z^k)\|_{z^k}\\
&= \Big((1-\eta)(1+\eta\chi_k) + (1+\eta\chi_k)\eta\delta_k + \frac{\eta^2}{2}(1+\delta_k)^2 h_k\Big)\|F_c(z^k)\|_{z^k}\\
&\le ((1-\eta)(1+\eta\chi_k) + \Theta_k\eta)\|F_c(z^k)\|_{z^k}.
\end{aligned}$$

A short computation yields

$$(1-\eta)(1+\eta\chi_k) + \Theta_k\eta \le 1 + \frac{\chi_k}{4} \le 1 + \frac{\gamma}{2}\|F_c(z^k)\|_{z^k}$$

and thus

$$\|F_c(z^k+\eta\delta z^k)\|_{z^k+\eta\delta z^k} \le \Big(1+\frac{\gamma}{2}\|F_c(z^k)\|_{z^k}\Big)\|F_c(z^k)\|_{z^k}.$$

If $z^{k+1}\notin U$, then there is some $\eta^*\in[0,1]$ such that $z^k+\eta\delta z^k\in U$ for $\eta\in[0,\eta^*]$ and $z^k+\eta^*\delta z^k\notin\mathcal{L}(z^k)$, i.e. $\|F_c(z^k+\eta^*\delta z^k)\|_{z^k+\eta\delta z^k} > (1+\gamma\|F_c(z^k)\|_{z^k}/2)\|F_c(z^k)\|_{z^k}$, which is a contradiction. Thus, $z^{k+1}\in U$. Furthermore,

$$\|F_c(z^{k+1})\|_{z^{k+1}} \le \Theta_k\|F_c(z^k)\|_{z^k}$$

and therefore $\mathcal{L}(z^{k+1})\subset\mathcal{L}(z^k)$ is closed. □

The next corollary follows analogously as Corollary 4.4.

Corollary 6.2. *If, in addition to the assumptions of Theorem* 6.1*, there exists a constant* $\hat{C}>0$ *such that*

$$\|\zeta\|_Z \le \hat{C}\|\nabla F_c(z)\zeta\|_z$$

for all $\zeta\in Z$ *and* $z\in U$*, then the iterates converge to the solution* $z^*=(x^*,\lambda^*)$ *of* (OS).

For actual implementation of an inexact Newton method following Theorem 6.1 we need to satisfy the accuracy requirement (6.4). Thus, we do not only need an error estimator for the inner iteration computing δ_k, but also easily computable estimates

$[\omega]$ and $[\gamma]$ for the Lipschitz constants ω and γ in case no suitable theoretical values can be derived. Using (6.6) and $h_k = \omega\|F_c(z^k)\|_z$, we construct the estimator

$$[\omega]_k = \frac{2}{(1+\delta_k)^2\|F_c(z^k)\|_{z^k}}\left(\frac{\|F_c(z^{k+1})\|_{z^k}}{\|F_c(z^k)\|_{z^k}} - \delta_k\right) \le \omega$$

in case $\|F_c(z^{k+1})\|_{z^k} > \delta_k\|F_c(z^k)\|_{z^k}$. Due to possible cancellation of significant digits, the estimate can only be expected to be reliable if $\|F_c(z^{k+1})\|_{z^k} \gg \delta_k\|F_c(z^k)\|_{z^k}$, i.e., in the beginning of the iteration when in fact the nonlinearity limits the convergence speed. The degradation of the estimate when the inaccuracy starts to limit the convergence rate can be observed in the numerical examples in Section 7. By (4.2) we have

$$\|F_c(z^{k+1})\|_{z^{k+1}} \le \left(1+\gamma\|\nabla F_c(z^k)\delta z^k\|_{z^k}\right)\|F_c(z^{k+1})\|_{z^k}$$

and thus we define

$$[\gamma]_k = \frac{1}{\|\nabla F_c(z^k)\delta z^k\|_{z^k}}\left|\frac{\|F_c(z^{k+1})\|_{z^{k+1}}}{\|F_c(z^{k+1})\|_{z^k}} - 1\right| \le \gamma.$$

Furthermore an estimate of h_k is required for the computation of δ_k from (6.4). Unfortunately, when defining $[h_k] = [\omega]_k\|F_c(z^k)\|_{z^k}$, the evaluation of the norm $\|F_c(z^k)\|_{z^k}$ in general requires the yet unknown Newton correction δz^k and therefore δ_k. Using the Lipschitz continuity (4.2) of the norm, h_k can be substituted instead by

$$[h_k] = \|F_c(z^k)\|_{z^{k-1}}\left(1-[\gamma]_k\|\nabla F_c(z^{k-1})\delta z^{k-1}\|_{z^{k-1}}\right)[\omega]_k. \tag{6.8}$$

Together the estimates can be used in the actually implementable accuracy requirement

$$(1+\delta_k)^2\frac{[h_k]}{2} + (1+(1+\delta_k)[\gamma]_k\|F_c(z^k)\|_{z^k})\delta_k \le \Theta_k.$$

Remark 6.3. If an inner iteration is used for approximately solving the Newton equation (6.1) which provides the *orthogonality relation* $(\delta z^k, \Delta z^k - \delta z^k)_{z^k} = 0$ in a scalar product $(\cdot,\cdot)_{z^k}$ that induces the affine invariant norm, the estimates can be tightened by substituting $(1+\delta_k)^2$ by $1+\delta_k^2$. Furthermore, the norm $\|\Delta z^k\|_{z^k}$ of the exact Newton correction is computationally available, which permits the construction of algorithms that are robust even for large inaccuracies δ_k. The application of a conjugate gradient method that is confined to the null space of the linearized constraints [1] to augmented Lagrangian-SQP methods can be the focus of future research. $\Diamond$

7 Numerical experiments

This section is devoted to present numerical tests for Example 2.1 that illustrate the theoretical investigations of the previous sections. To solve (P) we apply the so-called "optimize-then-discretize" approach: we compute an approximate solution by discretizing Algorithm 1, i.e., by discretizing the associated system (2.6). In the context of Example 2.1 we have $x^k = (y^k, u^k, v^k)$, $\delta x = (\delta y, \delta u, \delta v) \in W(0,T) \times L^2(0,T) \times L^2(0,T)$. To reduce the size of the system we take advantage of a relationship between the SQP steps δu, δv for the controls and the SQP step $\delta\lambda$ for the Lagrange multiplier. In fact, from

$$
\begin{aligned}
\frac{\partial^2 L_0}{\partial u^2}(x^k, \tilde{\lambda}^k)\delta u + \frac{\partial e}{\partial u}(x^k)^\star \delta\lambda &= -\frac{\partial L_0}{\partial u}(x^k, \tilde{\lambda}^k), \\
\frac{\partial^2 L_0}{\partial v^2}(x^k, \tilde{\lambda}^k)\delta v + \frac{\partial e}{\partial v}(x^k)^\star \delta\lambda &= -\frac{\partial L_0}{\partial v}(x^k, \tilde{\lambda}^k)
\end{aligned}
$$

we infer that

$$
\begin{aligned}
\delta u &= -\frac{1}{\alpha}\left(\tilde{\lambda}^k(\cdot,0) - \lambda^k(\cdot,0) + \delta\lambda(\cdot,0)\right) \quad \text{in } (0,T), \\
\delta v &= \frac{1}{\beta}\left(\tilde{\lambda}^k(\cdot,1) - \lambda^k(\cdot,1) + \delta\lambda(\cdot,1)\right) \quad \text{in } (0,T),
\end{aligned}
\tag{7.1}
$$

where $\tilde{\lambda}^k = \lambda^k + ce(x^k)$ by step b) of Algorithm 1. Inserting (7.1) into (2.6) we obtain a system only in the unknowns $(\delta y, \delta\lambda)$. Note that the second Fréchet-derivative of the Lagrangian is given by

$$
\langle L_0''(x^k, \tilde{\lambda}^k)\zeta, \xi\rangle_X = \int_Q \zeta_1\xi_1\big(1 + 2\tilde{\lambda}^k\big)\,dx + \int_0^T \alpha\zeta_2\xi_2 + \beta\zeta_3\xi_3\,dt
$$

for $\zeta = (\zeta_1, \zeta_2, \zeta_3)$, $\xi = (\xi_1, \xi_2, \xi_3) \in X$. The solution $(\delta y, \delta u, \delta v, \delta\lambda)$ of (2.6) is computed as follows: First we solve

$$
\begin{aligned}
y_t - \nu y_{xx} + (y^k y)_x &= -e^k && \text{in } Q, \\
\nu y_x(\cdot,0) + \sigma_0 y(\cdot,0) + \frac{\lambda(\cdot,0)}{\alpha} &= \frac{1}{\alpha}\left(\lambda^k(\cdot,0) - \tilde{\lambda}^k(\cdot,0)\right) && \text{in } (0,T), \\
\nu y_x(\cdot,1) + \sigma_1 y(\cdot,1) - \frac{\lambda(\cdot,1)}{\beta} &= \frac{1}{\beta}\left(\tilde{\lambda}^k(\cdot,1) - \lambda^k(\cdot,1)\right) && \text{in } (0,T), \\
y(0,\cdot) &= 0 && \text{in } \Omega, \\
(1 - \tilde{\lambda}^k_x)y - \lambda_t - \nu\lambda_{xx} - y^k\lambda_x &= y^k - z && \text{in } Q, \\
\nu\lambda_x(\cdot,0) + (y(\cdot,0) + \sigma_0)\lambda(\cdot,0) &= 0 && \text{in } (0,T), \\
\nu\lambda_x(\cdot,1) + (y(\cdot,1) + \sigma_1)\lambda(\cdot,1) &= 0 && \text{in } (0,T), \\
\lambda(T,\cdot) &= 0 && \text{in } \Omega,
\end{aligned}
\tag{7.2}
$$

where $e^k = y^k_t - \nu y^k_{xx} + y^k y^k_x - f$, and set $\delta y = y$ and $\delta\lambda = \lambda$. Then we obtain δu and δv from (7.1). For more details we refer the reader to [14].

For the time integration we use the backward Euler scheme while the spatial variable is approximated by piecewise linear finite elements. The programs are written in MATLAB, version 5.3, executed on a Pentium III 550 MHz personal computer.

Run 7.1 (Neumann control). In the first example we choose $T = 1$, $\nu = 0.1$, $\sigma_0 = \sigma_1 = 0$, $f = 0$, and

$$y_0 = \begin{cases} 1 & \text{in } (0, 0.5], \\ 0 & \text{otherwise.} \end{cases}$$

The grid is given by

$$x_i = \frac{i}{50} \quad \text{for } i = 0, \ldots, 50 \quad \text{and} \quad t_j = \frac{jT}{50} \quad \text{for } j = 0, \ldots, 50.$$

To solve (2.1) for $u = v = 0$ we apply the Newton method at each time step. The algorithm needs one second CPU time. The value of the cost functional is 0.083. Now we turn to the optimal control problem. We choose $\alpha = \beta = 0.01$, and the desired state is $z(t, \cdot) = y_0$ for $t \in (0, T)$. In view of the choice of z and the nonlinear convection term yy_x in (2.1b) we can interprete this problem as determining u in such a way that it counteracts the uncontrolled dynamics which smoothes the discontinuity at $x = 0.5$ and transports it to the left as t increases. The discretization of (7.2) leads to an indefinite system

$$H^k \begin{pmatrix} \delta y \\ \delta \lambda \end{pmatrix} = r^k \quad \text{with} \quad H^k = \begin{pmatrix} A^k & (B^k)^\top \\ B^k & C^k \end{pmatrix}. \tag{7.3}$$

As starting values for Algorithm 1 we take $y^0 = 0$, $u^0 = v^0 = 0$ and $\lambda^0 = 0$.

(i) First we solve (7.3) by an LU-factorization (MATLAB routine `lu`) so that the theory of section 4 applies. According to section 4 we stop the SQP iteration if

$$\|F_c(z^{k+1})\|_{z^k} \le 10^{-3} \cdot \|F_c(z^0)\|_{z^0}. \tag{7.4}$$

In the case if $\|F_c(z^0)\|_{z^0}$ is very large, the factor 10^{-3} on the right-hand side of (7.4) might be too big. To avoid this situation Algorithm 1 is terminated if (7.4) and in addition

$$\|F_c(z^{k+1})\|_{z^k} < 10^{-3}$$

hold. The augmented Lagrangian-SQP method stops after four iterations. The CPU times for different values of c can be found in Table 7.7. Let us mention that for $c = 0.1$ the algorithm needs 102.7 seconds and for $c = 1$ we observe divergence of Algorithm 1. As it was proved in [12] the set of admissible starting values reduces whenever c enlarges. The value of the cost functional is 0.041. In Figure 7.1 the residuum $t \mapsto \|y(t, \cdot) - z(t, \cdot)\|_{L^2(\Omega)}$ for the solution of (2.1) for $u = v = 0$ as well as for the optimal state is plotted. Furthermore, the optimal controls are presented. The decay of $\|F_c(z^{k+1})\|_{z^k}$, $k = 0, \ldots, 3$, for the first invariant norm given by (5.3) and for different values of c is shown in Table 7.1. Recall that the invariant norm is only defined for

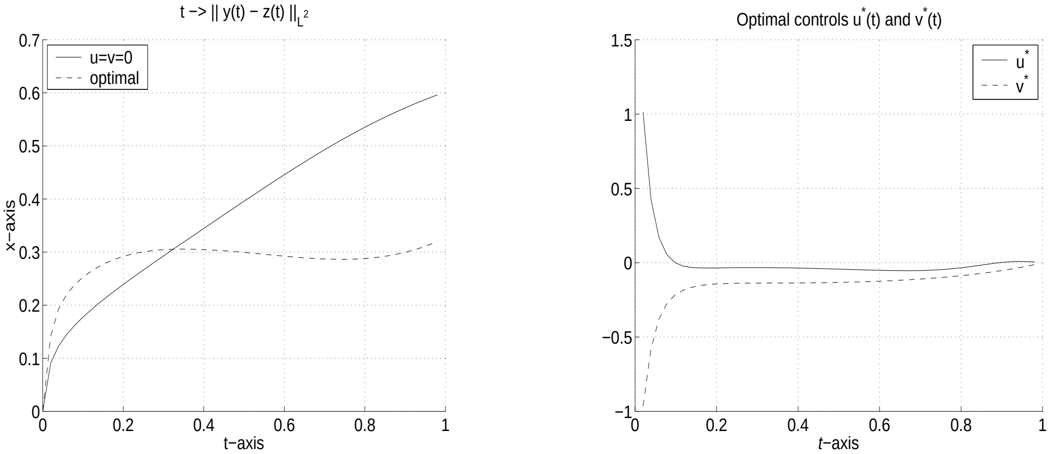

Figure 7.1: Run 7.1: residuum $t \mapsto \|y(t,\cdot) - z(t,\cdot)\|_{L^2(\Omega)}$ and optimal controls.

	$c = 0$	$c = 10^{-3}$	$c = 10^{-2}$
$\|F_c(z^0)\|_{z^0}$	4.636278	4.630344	4.642807
$\|F_c(z^1)\|_{z^0}$	1.635481	1.625800	1.581022
$\|F_c(z^2)\|_{z^1}$	0.210650	0.202490	0.184842
$\|F_c(z^3)\|_{z^2}$	0.003625	0.003234	0.002663
$\|F_c(z^4)\|_{z^3}$	0.000002	0.000001	0.000001

Table 7.1: Run 7.1-**(i)**: decay of $\|F_c(z^{k+1})\|_{z^k}$ for the first norm.

	$c = 0$	$c = 10^{-3}$	$c = 10^{-2}$
$[\kappa]_0$	0.020	0.020	0.020
$[\kappa]_1$	0.019	0.019	0.020
$[\kappa]_2$	0.004	0.023	0.024
$[\kappa]_3$	0.021	0.022	0.025

Table 7.2: Run 7.1-**(i)**: values of $[\kappa]_k$ for different c.

$\|F_0(z^0)\|_{z^0}$	$\|F_0(z^1)\|_{z^0}$	$\|F_0(z^2)\|_{z^1}$	$\|F_0(z^3)\|_{z^2}$	$\|F_0(z^4)\|_{z^3}$
26.77865	4.91492	0.63812	0.0105	0.00002

Table 7.3: Run 7.1-**(i)**: decay of $\|F_c(z^{k+1})\|_{z^k}$ for the second norm.

$c \geq \bar{c}$. Unfortunately, the constant $\bar{c} \geq 0$ is unknown. We proceed as follows: Choose a fixed value for c and compute

$$[\kappa]_k = \frac{\langle L''_c(x^k, \lambda^k)\delta x, \delta x\rangle_X}{\|\delta x\|_X^2}$$

in each level of the SQP iteration. Whenever $[\kappa]_k$ is greater than zero, we have coercivity in the direction of the SQP step. Otherwise, c needs to be increased. In Table 7.2 we present the values for $[\kappa]_k$. We observed numerically that $[\kappa]_k$ is positive for $k = 0, \ldots, 3$. Moreover, $[\kappa]_k$ increased if c increased.

Next we tested the second norm introduced in (5.11) for $c = 0$. Again, the augmented Lagrangian-SQP method stops after four iterations and needs 97.4 seconds CPU time. Thus, both invariant norms lead to a similar performance of Algorithm 1. The decay of $\|F_c(z^{k+1})\|_{z^k}$ can be found in Table 7.3.

(ii) Now we solve (7.3) by an inexact generalized minimum residual (GMRES) method (MATLAB routine `gmres`). As a preconditioner for the GMRES method we took an incomplete LU-factorization of the matrix

$$D = \begin{pmatrix} I & P^\top \\ P & 0 \end{pmatrix} \tag{7.5}$$

by utilizing the MATLAB function `luinc(`D`,1e-05)`. Here, the matrix P is the discretization of the heat operator $y_t - \nu y_{xx}$ with the homogeneous Robin boundary conditions $\nu y_x(\cdot, 0) + \sigma_0 y(\cdot, 0) = \nu y_x(\cdot, 1) + \sigma_1 y(\cdot, 1) = 0$ in $(0, T)$. The same preconditioner is used for all Newton steps. We chose $\Theta_k = 0.6$ for all k. In section 6 we introduced estimators for the constants ω and γ, denoted by $[\omega]_k$ and $[\gamma]_k$, respectively. Using $[\omega]_k$ and $[\gamma]_k$ we computed $[h_k]$ by (6.8). To compute $\|F_c(z^0)\|_{z^0}$ and $\|F_c(z^1)\|_{z^0}$ we take $\delta_0 = 10^{-10}$ Then we determine δ_k as follows:

	$c=0$	$c=10^{-3}$	$c=10^{-2}$
$\|F_c(z^0)\|_{z^0}$	4.63628	4.63030	4.64281
$\|F_c(z^1)\|_{z^1}$	1.25125	1.24517	1.65165
$\|F_c(z^2)\|_{z^2}$	0.24235	0.23416	1.45285
$\|F_c(z^3)\|_{z^3}$	0.04235	0.04071	0.40544
$\|F_c(z^4)\|_{z^4}$	0.00525	0.00479	0.05616
$\|F_c(z^5)\|_{z^5}$	0.00110	0.00099	0.00595
$\|F_c(z^6)\|_{z^6}$	0.00028	—	0.00113
$\|F_c(z^7)\|_{z^7}$	—	—	0.00008

Table 7.4: Run 7.1-**(ii)**: decay of $\|F_c(z^k)\|_{z^k}$ for the first norm with $\Theta_k = 0.6$.

$\delta_k = \Theta_k$;

while $(1+\delta_k)^2 \frac{[h_k]}{2} + (1 + (1+\delta_k)[\gamma]_k \|F_c(z^k)\|_{z^k})\delta_k > \Theta_k$ **do**

$\delta_k = \frac{\delta_k}{2}$;

end while;

for $k \geq 1$. Since by Propositions 5.3 and 5.6 γ is of the same order of magnitude as ω, we set $[\gamma]_0 = [\omega]_0$. The decay of $\|F(z^k)\|_{z^k}$ is presented in Table 7.4. Note that $h_{k+1}/h_k = \|F(z^{k+1})\|_{z^{k+1}}/\|F(z^k)\|_{z^k}$ holds. Thus we can check estimate (6.5) without knowing the Lipschitz constant ω. For our example it turns out that (6.5) is satisfied numerically during the SQP iteration, see Table 7.4. Algorithm 1 stops after at most six iterations. In particular, for $c = 10^{-3}$ the augmented Lagrangian-SQP method has the best performance. In Tables 7.5 and 7.6 the values of the estimators are presented. As we mentioned in section 6 the value of $[\omega]_k$ can only be reliable if $\|F_c(z^{k+1})\|_{z^k} \gg \delta_k \ \|F_c(z^k)\|_{z^k}$ holds. Introducing

$$G_c^k = \left| \|F_c(z^{k+1})\|_{z^k} - \delta_k \ \|F_c(z^k)\|_{z^k} \right|, \quad k \geq 0 \text{ and } c \geq 0,$$

we find $G_0^k \geq 0.065$ for $k = 0, \ldots, 3$, but $G_0^4 \approx 0.0075$ and $G_0^5 \approx 0.0039$. Therefore, for $c = 0$ the values $[\omega]_4$ and, in particular, $[\omega]_5$ indicate that the inaccuracy starts to limit the convergence rate. In Table 7.7 the CPU times for the first norm are presented. It turns out that the performance of the inexact method does not change significantly for different values of Θ_k. The first norm leads to a better performance of the inexact method. Compared to part **(i)** the CPU time is reduced by about 33% if one takes the first norm. In case of the second norm the reduction is only about 9%, compare Table 7.8. Finally we test the inexact method utilizing a variable Θ_k. We choose $\Theta_0 = 0.9$ and $\Theta_k = \Theta_{k-1}/2$ for $k \geq 1$. It turns out that the inexact method does not speed up significantly for the first norm, but in case of the second norm the algorithm needs 81.8 seconds, i.e., the CPU time of the exact method is reduced by about 16%.

	$c=0$	$c=10^{-3}$	$c=10^{-2}$
$[\omega]_0$	0.15	0.15	0.15
$[\omega]_1$	0.08	0.07	0.01
$[\omega]_2$	0.06	0.05	0.38
$[\omega]_3$	0.17	0.24	0.18
$[\omega]_4$	3.53	3.38	0.13
$[\omega]_5$	22.00	—	1.62
$[\omega]_6$	—	—	6.11

Table 7.5: Run 7.1-(**ii**): values of $[\omega]_k$ for $\Theta_k = 0.6$.

	$c=0$	$c=10^{-3}$	$c=10^{-2}$
$[\gamma]_0$	0.15	0.15	0.15
$[\gamma]_1$	0.05	0.05	0.01
$[\gamma]_2$	0.06	0.06	0.08
$[\gamma]_3$	0.02	0.01	0.09
$[\gamma]_4$	0.01	0.07	0.17
$[\gamma]_5$	0.02	—	1.89
$[\gamma]_6$	—	—	20.15

Table 7.6: Run 7.1-(**ii**): values of $[\gamma]_k$ for $\Theta_k = 0.6$.

	$c=0$	$c=10^{-3}$	$c=10^{-2}$
exact	97.5	96.8	96.9
inexact, $\Theta_k = 0.3$	63.5	64.1	64.0
inexact, $\Theta_k = 0.4$	61.8	61.9	63.1
inexact, $\Theta_k = 0.5$	61.7	62.8	63.5
inexact, $\Theta_k = 0.6$	64.1	62.0	65.7
inexact, $\Theta_k = 0.9$	61.6	64.7	68.9

Table 7.7: Run 7.1-(**ii**): CPU times in seconds for the first norm.

	first norm	second norm
exact	97.5	97.4
inexact, $\Theta_k = 0.3$	63.5	86.9
inexact, $\Theta_k = 0.4$	61.8	91.5
inexact, $\Theta_k = 0.5$	61.7	90.0
inexact, $\Theta_k = 0.6$	64.1	87.8
inexact, $\Theta_k = 0.9$	61.6	86.3

Table 7.8: Run 7.1-(**ii**): CPU times in seconds for both norms and $c = 0$.

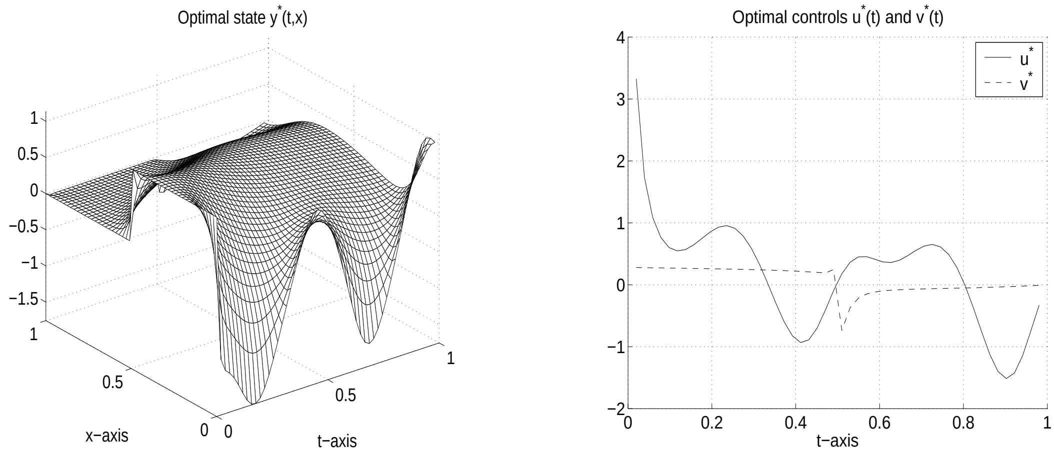

Figure 7.2: Run 7.2: optimal state and controls.

	$c = 0$	$c = 10^{-3}$	$c = 10^{-2}$
$\|F_c(z^0)\|_{z^0}$	3.11799	3.12494	3.15978
$\|F_c(z^1)\|_{z^0}$	1.25420	1.29953	1.75698
$\|F_c(z^2)\|_{z^1}$	0.18289	0.18768	0.26507
$\|F_c(z^3)\|_{z^2}$	0.01361	0.00849	0.01200
$\|F_c(z^4)\|_{z^3}$	0.00009	0.00003	0.00006

Table 7.9: Run 7.2-**(i)**: decay of $\|F_c(z^{k+1})\|_{z^k}$ for different c.

Run 7.2 (Robin control). We choose $T = 1$, $\nu = 0.05$, $\sigma_0(t) = \sin(4\pi t)$, $f = 0$, $\alpha = \beta = 0.01$,

$$\sigma_1 = \begin{cases} -10 & \text{in } \left(0, \dfrac{T}{2}\right), \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad y_0 = \begin{cases} 1 & \text{in } \left(0, \dfrac{1}{2}\right), \\ 0 & \text{otherwise.} \end{cases}$$

The desired state was taken to be $z(t, \cdot) = y_0 \cos(4\pi t)$ for $t \in [0, T]$.

(i) First we again solve (7.3) by an LU-factorization. We take the same starting values and stopping criteria as in Run 7.1. The augmented Lagrangian-SQP method stops after four iteration and needs 105 seconds CPU time. The discrete optimal solution is plotted in Figure 7.2. From Table 7.9 it follows that (4.6) is satisfied numerically. Let us mention that $[\kappa]_0, \dots, [\kappa]_3$ are positive for $c \in \{0, 10^{-3}, 10^{-2}\}$. For the needed CPU times we refer to Tables 7.11 and 7.12.

(ii) Now we solve (7.3) by an inexact GMRES method. As a preconditioner we take the same as in Run 7.1. We choose $\Theta_k = 0.5$ for all k. The decay of $\|F(z^k)\|_{z^k}$ is presented in Table 7.10. As in part **(i)** we find that $[\kappa]_k > 0$

	$c=0$	$c=10^{-3}$	$c=10^{-2}$
$\Vert F_c(z^0)\Vert_{z^0}$	3.1180	3.1249	3.1598
$\Vert F_c(z^1)\Vert_{z^1}$	1.2917	1.2812	1.2248
$\Vert F_c(z^2)\Vert_{z^2}$	0.3094	0.2894	0.2206
$\Vert F_c(z^3)\Vert_{z^3}$	0.0368	0.0348	0.0533
$\Vert F_c(z^4)\Vert_{z^4}$	0.0062	0.0045	0.0051
$\Vert F_c(z^5)\Vert_{z^5}$	0.0012	0.0007	0.0018
$\Vert F_c(z^6)\Vert_{z^6}$	0.0002	—	0.0006

Table 7.10: Run 7.2-**(ii)**: decay of $\Vert F_c(z^k)\Vert_{z^k}$ for $\Theta_k = 0.5$.

	$c=0$	$c=10^{-3}$	$c=10^{-2}$
exact	105.1	105.7	105.7
inexact, $\Theta_k = 0.5$	50.4	49.3	51.6
inexact, $\Theta_k = 0.6$	51.3	51.3	54.0
inexact, $\Theta_k = 0.75$	53.2	50.7	55.2
inexact, $\Theta_k = 0.9$	52.8	49.9	55.8

Table 7.11: Run 7.2-**(ii)**: CPU times in seconds for the first norm.

for all test runs. The needed CPU times are shown in Table 7.11. As we can see, the inexact augmented Lagrangian-SQP method with GMRES is much faster than the exact one using the LU-factorization. For the first norm the CPU time is reduced by about 50%, and for the second norm by about 28%. Moreover, for our example the best choice for c is $c = 10^{-3}$. For smaller values of Θ_k the method does not speed up significantly. As in Run 7.1 we test the inexact method utilizing a variable Θ_k. Again we choose $\Theta_0 = 0.9$ and $\Theta_k = \Theta_{k-1}/2$ for $k \geq 1$. As in Run 7.1 the inexact method does not speed up significantly for the first norm, but in case of the second norm the algorithm needs 68.5 seconds, i.e., the CPU time of the exact method is reduced by about 35%.

	first norm	second norm
exact	105.1	105.5
inexact, $\Theta_k = 0.5$	50.4	75.3
inexact, $\Theta_k = 0.6$	51.0	75.4
inexact, $\Theta_k = 0.75$	53.2	77.2
inexact, $\Theta_k = 0.9$	52.8	77.5

Table 7.12: Run 7.2-**(ii)**: CPU times in seconds for both norms and $c = 0$.

References

[1] D. Braess, P. Deuflhard, and K. Lipnikov. A Subspace Cascadic Multigrid Method for Mortar Elements, Preprint SC 99-07, Konrad-Zuse-Zentrum Berlin, 1999.

[2] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology. Volume 5: Evolution Problems I.* Springer-Verlag, Berlin, 1992.

[3] P. Deuflhard. *Newton Methods for Nonlinear Problems. Affine Invariance Properties and Adaptive Algorithms.* In preparation.

[4] P. Deuflhard and G. Heindl. Affine invariant convergence theorems for Newton's method and extensions to related methods. *SIAM J. Numer. Anal.*, 16(1):1–10, 1979.

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

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

[7] V. Girault and P.-A. Raviart. *Finite Element Methods for Navier-Stokes Equations. Theory and Algorithms.* Springer-Verlag, Berlin, 1980.

[8] A. Hohmann. *Inexact Gauss Newton Methods for Parameter Dependent Nonlinear Problems.* PhD thesis, Freie Universität Berlin, 1994.

[9] K. Ito and K. Kunisch. Augmented Lagrangian-SQP-methods in Hilbert spaces and application to control in the coefficient problems. *SIAM J. on Optimization*, 6(1):96–125, February 1996.

[10] D. G. Luenberger. *Optimization by Vector Space Methods.* John Wiley & Sons, Inc., New York, 1969.

[11] H. Maurer and J. Zowe. First and second order necessary and sufficient optimality conditions for infinite-dimensional programming problems. *Math. Programming*, 16:98–110, 1979.

[12] S. Volkwein. Mesh-independence for an augmented Lagrangian-SQP method in Hilbert spaces. *SIAM J. Control and Optimization*, 38(3):767–785, 2000.

[13] S. Volkwein. Mesh-independence of Lagrange-SQP methods with Lipschitz-continuous Lagrange multiplier updates. Submitted, 2000.

[14] S. Volkwein. Second-order conditions for boundary control problems of the Burgers equation. Submitted, 2000.

[15] E. Zeidler, *Nonlinear Functional Analysis and its Applications I.* Springer-Verlag, New York, 1986.