NEWTON'S METHOD FOR TRACKING TYPE CONTROL OF THE INSTATIONARY NAVIER-STOKES EQUATIONS

M. HINZE

Fachbereich Mathematik, Technische Universität Berlin, Straße des 17. Juni 136, D-10623 Berlin

E-mail: hinze@math.tu-berlin.de

K. KUNISCH

Institut für Mathematik, Karl-Franzens Universität Graz, Heinrichstr. 36, A-8010 Graz

E-mail: kunisch@kfunigraz.ac.at

Newton's method for optimal open loop optimal control of the instationary Navier-Stokes equation is investigated. It is argued that the a-priori formidable Newton step can be decomposed into nonlinear primal and linear adjoint systems, which is amenable for existing CFL-software. We report a numerical test which demonstrates the feasiblity of the approach. In addition the functional analytic setting of the convergence analysis is presented.

1 The optimal control problem

We consider the optimal control problem

$$\left.\begin{array}{ll} \min_{(y,u)\in W\times U}\ J(y,u) := \frac{1}{2}\int\limits_{Q_o} |y-z|^2\,dxdt + \frac{\alpha}{2}\int\limits_{Q_c} |u|^2\,dxdt & \\ \text{subject to} & \\ \frac{\partial y}{\partial t} + (y\cdot\nabla)y - \nu\Delta y + \nabla p = Bu & \text{in } Q=(0,T)\times\Omega, \\ \text{div } y = 0 & \text{in } Q, \\ y(t,\cdot) = 0 & \text{on } \Sigma = (0,T)\times\partial\Omega, \\ y(0,\cdot) = y_0 & \text{in } \Omega. \end{array}\right\} \quad (1)$$

where $Q_c := \Omega_c \times (0,T)$ and $Q_o := \Omega_o \times (0,T)$, with Ω_c and Ω_o subsets of $\Omega = (0,1)^2$ denoting control and observation volumes, respectively. In this form solving (1) appears at first to be a standard task. However, the formidable size of (1) and the goal of analyzing second order methods necessitate an independent analysis. One of the few contributions focusing on second order methods for optimal control of fluids are given by Ghattas *et al* [3] and Heinkenschloss [4]. These works are restricted to stationary problems, however.

To define the spaces and operators required for the investigation of (1)

we introduce the solenoidal spaces

$$H = \{v \in C_0^\infty(\Omega)^2 \colon \text{ div } v = 0\}^{-|\cdot|_{L^2}}, V = \{v \in C_0^\infty(\Omega)^2 \colon \text{ div } v = 0\}^{-|\cdot|_{H^1}},$$

with the superscripts denoting closures in the respective norms. Further we define

$$W = \{v \in L^2(V) \colon v_t \in L^2(V^*)\} \quad \text{and} \quad Z := L^2(V) \times H,$$

where W is endowed with the norm

$$|v|_W = (|v|^2_{L^2(V)} + |v_t|^2_{L^2(V^*)})^{1/2},$$

and set $\langle \cdot, \cdot \rangle := \langle \cdot, \cdot \rangle_{L^2(V^*), L^2(V)}$, with V^* denoting the dual space of V. Here $L^2(V)$ is an abbreviation for $L^2(0,T;V)$ and similarly $L^2(V^*) = L^2(0,T;V^*)$. Recall that up to a set of measure zero in $(0,T)$ elements $v \in W$ can be identified with elements in $C([0,T];H)$. In (1) further $U = L^2(Q_c)$ denotes the Hilbert space of controls which is identified with its dual U^*. The the cost functional $J \colon L^2(H) \times U \to \mathbb{R}$ is assumed to be bounded from below, weakly lower semi-continuous, twice Fréchet differentiable with locally Lipschitzean second derivative, and radially unbounded in u, i.e. $J(y,u) \to \infty$ as $|u|_U \to \infty$, for every $y \in W$. We define the nonlinear mapping

$$e \colon W \times U \to Z^*$$

by

$$e(y,u) = (\tfrac{\partial y}{\partial t} + (y \cdot \nabla)y - \nu \Delta y - Bu, y(0) - y_0)$$

where $Bu(t,\cdot)$ denotes the Leray-projection of the extension by zero of $u(t,\cdot)$ to the whole of Ω and $y_0 \in H$. In variational form the constraints in (1) can be equivalently expressed as: given $u \in U$ find $y \in W$ such that $y(0) = y_0$ and

$$\langle y_t, v \rangle + \langle (y \cdot \nabla)y, v \rangle + \nu(\nabla y, \nabla v)_{L^2(L^2)} = \langle Bu, v \rangle \ \ \forall\, v \in L^2(V). \quad (2)$$

It is well known, see Temam [6] that for every $u \in Y$ (2) admits a unique solution $y(u) \in W$. Therefore, with respect to existence (1) can equivalently be rewritten as

$$\min \ \hat{J}(u) = J(y(u), u) \ \text{ subject to } \ u \in U, \quad (3)$$

where $y(u) \in W$ satisfies $e(y(u), u) = 0$. It is proved by Abergel *et al* [1] that (3) admits a solution $(y^*, u^*) \in W \times U =: X$. We shall frequently refer to the linearized Navier-Stokes system and the adjoint equations given next:

$$\begin{cases} v_t + (v \cdot \nabla)y + (y \cdot \nabla)v - \nu \Delta v = f \quad \text{in } \Omega \text{ a.e. on } (0,T], \\ v(0) = v_0. \end{cases} \quad (4)$$

and

$$\begin{cases} -w_t + (\nabla y)^t w - (y \cdot \nabla) w - \nu \Delta v = g \quad \text{in } \Omega \text{ a.e. on } [0,T). \\ w(T) = 0. \end{cases} \tag{5}$$

The following proposition is proved by the authors [5]. It is essential for the analysis of Newton and quasi-Newton methods.

Proposition 1.1. Let $x = (y,u) \in W \times U$. Then $e_y(x)\colon W \to Z^*$ is a homeomorphism. Moreover, if the inverse of its adjoint $e_y^{-*}(x)\colon W^* \to Z$ is applied to an element $g \in W^* \cap L^\alpha(V^*)$, where $\alpha \in [1, 4/3]$ then setting $(w, w_0) := e_y^{-*}(x) g \in L^2(V) \times H$ we have $w_t \in L^\alpha(V^*)$, $w(0) = w_0$ and w is the variational solution to (5).

2 Derivatives

To apply Newton's method to (3) we need second order information of the functional $\hat{J}$. This can be obtained by applying the implicit function theorem to the equation $e(y,u) = 0$. The derivatives of the operator e were characterized by the authors [5] which we cite for the convenience of the reader.

Proposition 2.1. The operator $e = (e^1, e^2)\colon X \to Z^*$ is twice continuously differentiable with Lipschitz continuous second derivative. The action of the first two derivatives of e^1 are given by

$$\begin{aligned} \langle e^1_x(x)(w,s), \phi \rangle &= \langle w_t, \phi \rangle + \langle (w \cdot \nabla) y, \phi \rangle + \langle (y \cdot \nabla) w, \phi \rangle \\ &\quad + \nu (\nabla w, \nabla \phi)_{L^2(L^2)} - \langle Bs, \phi \rangle_{L^2(L^2)}, \end{aligned}$$

where $x = (y,u) \in X, (w,s) \in X$ and $\phi \in L^2(V)$, and

$$\begin{aligned} \langle e^1_{xx}(x)(w,s)(v,r), \phi \rangle &= \langle e^1_{yy}(x)(w,v), \phi \rangle = \\ &\quad \langle (w \cdot \nabla) v, \phi \rangle + \langle (v \cdot \nabla) w \phi, v \rangle =: \langle v, H(\phi) w \rangle_{W,W^*}, \end{aligned} \tag{6}$$

where $(v,r) \in X$.

As a consequence of variations of standard estimates for the Navier-Stokes and the linearized Navier-Stokes equations and the implicit function theorem the first derivative of the mapping $u \to y(u)$ at u in direction δu is given by

$$y'(u)\delta u = -e_y^{-1}(x) e_u(x) \delta u, \tag{7}$$

where $x = (y(u), u)$. By the chain rule we thus obtain

$$\langle \hat{J}'(u), \delta u \rangle_U = \langle J_u(x) - e_u^*(x) e_y^{-*}(x) J_y(x), \delta u \rangle_U.$$

Introducing the variable

$$\lambda = -e_y^{-*}(x) J_y(x) \in Z \tag{8}$$

we obtain the Riesz representation for the first derivative of $u \to \hat{J}(u)$:

$$\hat{J}'(u) = J_u(x) + e_u^* \lambda. \tag{9}$$

Here $\lambda = (\lambda^1, \lambda^0) \in Z$, $\lambda_t^1 \in L^{4/3}(V^*)$, $\lambda^1 \in C(H)$ and λ^1 is the variational solution of

$$\begin{cases} -\lambda_t^1 + (\nabla y)^t \lambda^1 - (y \cdot \nabla)\lambda^1 - \nu \Delta \lambda^1 = -J_y(x) \\ \lambda^1(T) = 0, \end{cases} \tag{10}$$

where the first equation holds in $L^{4/3}(V^*) \cap W^*$.
The computation of the representation of the second derivative of $\hat{J}''$ of $\hat{J}$ is more involved. Introducing the Lagrangian $L\colon X \times Z \to \mathbb{R}$

$$L(x, \lambda) = J(x) + \langle e(x), \lambda \rangle_{Z^*, Z} \tag{11}$$

and the matrix operators

$$T(x) = \begin{pmatrix} -e_y^{-1}(x) e_u(x) \\ Id_U \end{pmatrix} \in \mathcal{L}(U, X) \tag{12}$$

we observe that the second derivative of $\hat{J}$ can be expressed as

$$\hat{J}''(u) = T^*(x) L_{xx}(x, \lambda) T(x), \tag{13}$$

where $x = (y(u), u)$ and the second derivative of L with respect to x is given by

$$L_{xx}(x, \lambda) = \begin{pmatrix} J_{yy}(x) + \langle e_{yy}^1(x)(\cdot, \cdot),\, \lambda^1 \rangle & 0 \\ 0 & J_{uu}(x) \end{pmatrix} \in \mathcal{L}(X, X^*).$$

3 The algorithm

This section contains a description of Newton's method to solve (3). Throughout u^* denotes a (local) solution to (3). For the sake of reference we specify the Newton algorithm.

Algorithm 3.1. (Newton Algorithm).

1. Choose $u^0 \in N(u^*)$, set $k = 0$.
2. Do until convergence
   - i) solve $\hat{J}''(u^k)\delta u^k = -\hat{J}'(u^k)$,
   - ii) update $u^{k+1} = u^k + \delta u^k$,
   - iii) set $k = k + 1$.

Let us consider the linear system in 2. i). Its dimension is that of the control space U. From the characterization of the Hessian $\hat{J}''(u^k)$ we conclude that its evaluation requires as many solutions to the linearized Navier–Stokes equation (7) with appropriate right hand sides as is the dimension of U. If U is infinite dimensional then an appropriate discretization must be carried out. Let us assume now that the dimension of U is large so that direct evaluation of $\hat{J}''(u^k)$ is not feasible. As a consequence 2. i) will be solved iteratively, e. g. by a conjugate gradient technique. We shall then refer to 2. i) as the "inner " loop as opposed to the do–loop in 2. which is the "outer" loop of the Newton algorithm. The inner loop at iteration level k of the outer loop requires to

(i) evaluate $\hat{J}'(u^k)$, i. e. given u^k compute $y(u^k)$ from (2) and λ^1 from (10) with $x = (y(u^k), u^k)$,

(ii) iteratively evaluate the action of $\hat{J}''(u^k)$ on δ_j^k, the j–th iterate of the inner loop on the k–th level of the outer loop.

The iterate $q = \hat{J}''(u^k)\delta_j^k$ can be evaluated by successively applying the steps

a) solve (4) with $v_0 = 0$ and $f = B\delta_j^k$ in $L^2(V^*)$ for $v \in W$ where $y = y(u^k)$,

b) evaluate $J_{yy}(x)v + \langle e_{yy}^1(x)(v,\cdot), \lambda^1\rangle$,

c) solve (5) with $g = J_{yy}(x)v + \langle e_{yy}^1(x)(v,\cdot), \lambda^1\rangle$ in W^* for $w \in L^2(V)$,

d) and finally set $q := J_{uu}\delta u + B^*w$.

It can be shown that the functional appearing in b) is an element of $W^* \cap L^{4/3}(V^*)$, and hence (5) is meaningful in the sense of admitting a solution w in $L^2(V)$ with $w_t \in L^{4/3}(V^*)$. Summarizing, for the outer iteration of the Newton method one Navier–Stokes solve for $y(u^k)$ and one linearized Navier–Stokes solve for $\lambda(u^k)$ are required. For the inner loop one forward (–in time) as well as one backwards linearized Navier–Stokes solve per iteration is necessary.

Concerning initialization we observe that if initial guesses $(y_0, u_0) \in W \times U$ are available (with y_0 not necessarily $y(u_0)$) then alternatively to the initialization in Algorithm 3.1 this information can be used advantageously to compute the adjoint variable λ^1 required for the initial guess for the right hand side of the linear system as well as to carry out steps a) - c) for the evaluation of the Hessian. There is no necessity to recompute $y(u_0)$ from u_0.

To prove local quadratic convergence of the method we need to impose a smallness assumption on the difference $|y(u^*) - z|_{L^2(H)}$. It guarantees second order sufficient optimality of the local minimizer u^*. The proof of the following theorem can then be given by standard arguments.

Theorem 3.1. If $|y(u^*) - z|_{L^2(H)}$ is sufficiently small then there exist a neighbourhood $\mathcal{U}(u^*)$ such that for every $u^0 \in \mathcal{U}(u^*)$ the iterates $\{u^n\}_{n\in\mathbb{N}}$ of Newton's Algorithm 3.1 converge quadratically to u^*.

It is worth noting that by choosing the initial approximation to the Hessian appropriately local super-linear convergence for quasi-Newton methods in our context can also be proved. This can be deduced from the special structure of the Hessian $\hat{J}''(u)$ in (13) which can be decomposed into the sum of a compact part and $J_{uu}(y(u), u)$ [5].

4 Numerical Example

Here we present numerical results for the control of driven cavity flow with control problem given by (3). The discretization of the Navier-Stokes equations, its linearization and adjoint was carried out by using parts of the code developed by Bänsch [2], which is based on Taylor-Hood finite elements for the spatial discretization. As time step size we took $\delta t = .00625$, which resulted in 160 grid points for the time grid, and 545 pressure and 2113 velocity nodes for the spatial discretization. This results in a total number of unknowns (primal-, adjoint-, and control variables) of the optimization problem of order $2.2{*}10^6$. The time horizon could still be increased or the mesh size decreased by utilizing reduced storage techniques at the expense of additional cpu-time, but we shall not pursue this aspect here. All computations were performed on a DEC-ALPHA^{TM}station 500.

We present the results for $\Omega_c = \Omega_o = (0,1)^2$, $\alpha = 10^{-2}$ and Re=400. Due to the size of the discretized control space we have to apply the inexact version of Newton's method. As iterative solver for 2.i) in Algorithm 3.1 we utilize the conjugate gradient algorithm. We therefore can only expect super-linear convergence of Newton's method which in fact is confirmed in Table 1. A comparison of Newton's method to the gradient and conjugate gradient algorithm with step-size control for this example is presented by the authors [5] and confirmes the superiority of Newton's method over the gradient algorithm. In Figure 2 the evolution of the cost functional and the control cost as a function of time are documented. It can be observed that graphically there is no significant change after the third iteration of the method. These comments hold for quite a wide range of values for α. In Fig. 1 the target flow together with the uncontrolled flow, the controlled flow and the control action at the end of the time interval are presented. Note that different observation and control volumes result in smaller control and observation volumes as in our example, and thus the primal and adjoint equations are numerically simpler to solve. We refer Hinze *et al* [5] for a more detailed discussion of this topic.

Table 1. Performance of Newton's method for Example 1

Iteration	CG-steps	$\frac{\lvert\hat{J}'(u)\rvert}{\lvert\hat{J}'(u^0)\rvert}$	$\frac{\lvert\delta u^k\rvert_U}{\lvert\delta u^{k-1}\rvert_U}$	$\hat{J}(u^k)$
1	-	1.e0	-	1.196202e-2
2	13	3.358825e-1	1.	3.226486e-3
3	11	5.058497e-2	0.492	1.617913e-3
4	18	8.249029e-3	0.422	1.482032e-3
5	17	1.409278e-4	0.079	1.480533e-3
6	19	4.686819e-6	0.032	1.480534e-3

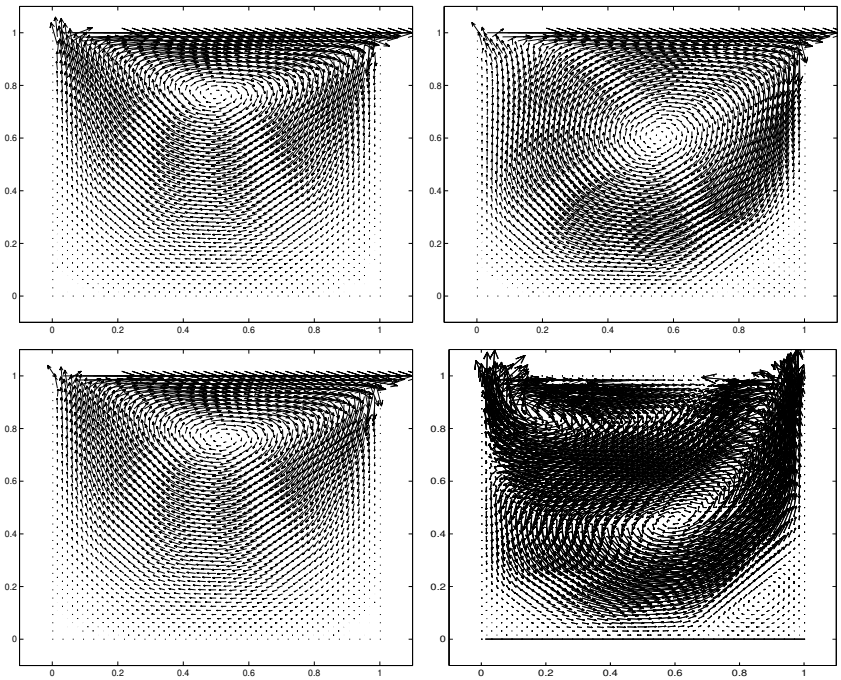

Figure 1. top left: target flow, top right:uncontrolled flow, bottom left: controlled flow, bottom right: control force

References

1. F. Abergel and R. Temam, Theoret. Comput. Fluid Dynamics **1** (1990), 303–325.

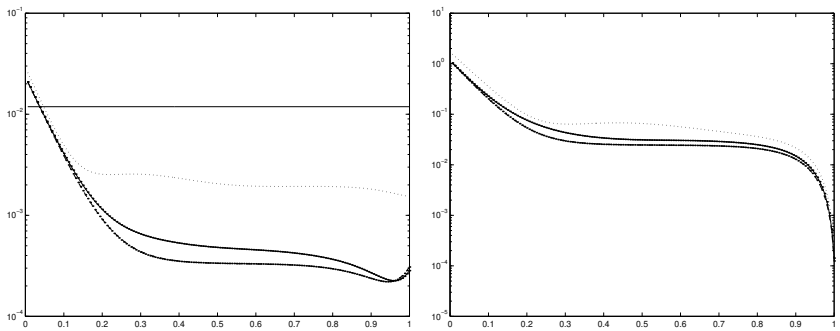

Figure 2. Re=400, $\alpha = 10^{-2}$: Evolution of cost functional (left) and control cost (right) for different iteration levels and relative accuracy = 1.d-5

2. E. Bänsch, J. Comp. Math. **36**, 1991
3. O. Ghattas and J.J. Bark, Journal of Computational Physics **136** (1997), 231–244.
4. M. Heinkenschloss, in Optimal Control: Theory, Algorithms, and Applications. Kluwer Academic Publishers B.V., 1998
5. M. Hinze and K. Kunisch, Karl-Franzens Universität Graz, Institut für Mathematik, Spezialforschungsbereich F003, Bericht 165, 1999
6. R. Temam: Navier-Stokes Equations, North Holland, Amsterdam, 1979.