Mathematisches Forschungsinstitut Oberwolfach

Report No. 11/2006

Numerical Techniques for Optimization Problems with PDE Constraints

Organised by
Matthias Heinkenschloss (Houston)
Ronald H.W. Hoppe (Augsburg)
Volker Schulz (Trier)

February 26th – March 4th, 2006

Abstract. Optimization problems with partial differential equation (PDE) constraints arise in many science and engineering applications. Their robust and efficient solution present many mathematical challenges and requires a tight integration of properties and structures of the underlying problem, of fast numerical PDE solvers, and of numerical nonlinear optimization. This workshop brought together experts in the above subdisciplines to exchange the latest research developments, discuss open questions, and to foster interactions between these subdisciplines.

Mathematics Subject Classification (2000): 49-xx, 65-xx.

Introduction by the Organisers

The numerical solution of optimization problems with partial differential equation (PDE) constraints is vital to a growing number of science and engineering applications. The development of robust and efficient algorithms for the solution of these optimization problems presents many challenges that arise out of, e.g., the intricate mathematical structure of these problems, the complicated interactions between numerical methods for PDE and optimization, the large-scale of the optimization problems, and the increasing complexity of applications. To identify and overcome these challenges an integrated approach is needed that builds on a variety of mathematical sub-disciplines, such as theory of PDEs, distributed parameter systems, numerical solution of PDEs, numerical optimization, and numerical linear algebra. This international workshop has brought together some of the leading experts in the fast developing field of optimization problems with PDE constraints to present recent developments in this area as well as to identify open problems and further research needs.

Among the themes of this workshop were the design and analysis of approaches for the solution of PDE constrained optimization problems with additional point-wise constraints on controls and states (the solution of the governing PDE). State constrained problems are particularly challenging because of the low regularity properties of the Lagrange multipliers associated with point-wise constraints on the states. A second theme was the development of adaptive methods for the solution of PDE constrained optimization problems and, more generally, the development of optimization level model reduction techniques for these problems. The goal here is to develop models (through, e.g., mesh adaptation or proper orthogonal model reduction) of the PDE constrained optimization problems that capture the relevant features of the optimization problems with a specified accuracy, but involve as few degrees of freedom as possible and, hence, are computationally less expensive to work with. A third theme was The efficient solution of linear systems arising in optimization algorithms for discretized PDE constrained optimization problems represented another theme. Finally, a number of talks presented advances and challenges in the solution of PDE constrained optimization problems arising in important science and industrial applications.

Numerical Techniques for Optimization Problems with PDE Constraints

Table of Contents

Abstracts

Large-scale Optimization Strategies for Simulated Moving Beds

Lorenz T. Biegler
(joint work with Yoshiaki Kawajiri)

1. Introduction

Chemical separations based on fluid-solid equilibrium are becoming increasingly important for pharmaceutical and fine chemicals industries. Based on a network of chromatographic columns, Simulated Moving Beds (SMB) represent an efficient and effective technology for this task. Mathematical models of SMBs consist of sets of PDEs for mass conservation in solid and fluid phases, along with algebraic consitutive equations and periodic boundary conditions; these models have resisted optimization of realistic processes. This talk presents a full discretization approach to the optimization of SMBs. With PDEs discretized in space and time, the resulting nonlinear program (NLP) is solved with a large-scale Newton-based barrier method. This optimization approach is very efficient and leads to the determination of novel designs and operating strategies. A case study based on the separation of glucose and fructose is presented to demonstrate this approach.

2. Background on Chromatographic Separations

Separation of chemical species by chromatography has been known for over a century. After introducing a small amount of mixture onto a tube (or column) packed with solid adsorbent material followed by a solvent (desorbent) flow, preferential adsorption of each species to the solid causes the mixture to separate into peaks, which are then withdrawn individually. While chromatographic separations are commonplace for lab analysis, high volume separations require the development of continuously operating systems. Such continuous chromatography-based systems have been proposed using the concept of *true moving beds* (TMB). In TMB, the solid and liquid phases move countercurrently, input streams consist of the mixture feed and desorbent, output streams consist of product extract and raffinate, and concentration profiles are established in steady state. However, countercurrent movement of the solid phase is difficult to realize in practice. Instead, TMB behavior can be approximated through a switching of valves for the input and output streams, leading to the concept of *simulated moving beds* (SMB). As shown in Figure 1, the nodal valve positions of the four streams are switched intermittently to simulate movement of the solid. Also, columns between these streams are defined as zones. Note that SMBs are not steady state processes but operate dynamically with periodic boundary conditions, *cyclic steady states*, imposed at each switching time.

Application of SMB processes is growing in the chemical industry. Since their development by UOP in the 60s for aromatic isomers separations, SMBs are now

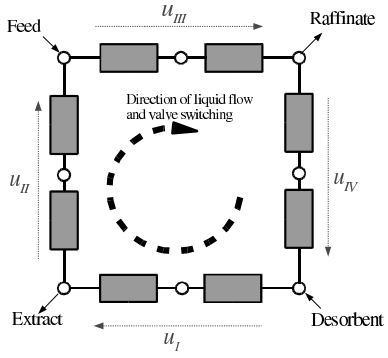

FIGURE 1. Standard SMB process with two columns in each zone

used in the separation of "close-boiling" hydrocarbons, fructose/glucose separation to yield high fructose corn syrup, and enantiomeric species in pharmaceutical processes. Because of their similar molecular properties, most of these species can only be separated by adsorption. Moreover, as more SMB applications are considered, a number of innovative designs have been proposed that extend beyond the structure in Figure 1 (see [1, 2] for a review). However, efficient optimization strategies are needed to evaluate these new designs and demonstrate in a systematic way the advantages offered by each.

3. Optimization Problem Formulation

A typical optimization problem for the design and operation of SMB processes is given by:

$$\max_{u_I(t),u_{II}(t),u_{III}(t),u_{IV}(t),t_{step}} \bar{u}_F := \frac{1}{t_{step}} \int_0^{t_{step}} u_F(t)dt \tag{1}$$

subject to :

$$\epsilon_b \frac{\partial C_i^j(x,t)}{\partial t} + (1-\epsilon_b)\frac{\partial q_i^j(x,t)}{\partial t} + u_m(t)\frac{\partial C_i^j(x,t)}{\partial x} = 0 \tag{2}$$

$$(1-\epsilon_b)\frac{\partial q_i^j(x,t)}{\partial t} = K_{appl\,i}(C_i^j(x,t) - C_i^{j,eq}(x,t)) \tag{3}$$

$$q_i^j(x,t) = f(C_i^{j,eq}(x,t))\ i = 1,\ldots,N_C, j = 1,\ldots,N_{Column} \tag{4}$$

$$u_{IV}(t) + u_D(t) = u_I(t), \quad C_{1,i}^{in}(t)u_I(t) = C_{N_{Column},i}(L,t)u_{IV}(t) \tag{5}$$

$$u_I(t) - u_E(t) = u_{II}(t), \quad u_{II}(t) + u_F(t) = u_{III}(t) \tag{6}$$

$$C_{N_I+N_{II},i}(L,t)u_{II}(t) + C_{F,i}u_F(t) = u_{II}(t)C_{N_I+N_{II}+1,i}^{in}(t) \tag{7}$$

$$u_{III}(t) - u_R(t) = u_{IV}(t), \quad C_{N_I+N_{II}+N_{III},i}(L,t) = C_{N_I+N_{II}+N_{III}+1,n}^{in}(t) \tag{8}$$

$$C_i^j(x,0) = C_i^{j+1}(x,t_{step}),\ j = 1,\ldots,N_{Column}-1 \tag{9}$$

$$q_i^j(x,0) = q_i^{j+1}(x,t_{step}),\ j = 1,\ldots,N_{Column}-1 \tag{10}$$

$$C_i^{N_{Column}}(x,0) = C_i^1(x,t_{step}) \tag{11}$$

$$q_i^{N_{Column}}(x,0) = q_i^1(x,t_{step}) \tag{12}$$

$$C_{E,i}(t) = C_{N_I,i}(L,t) \tag{13}$$

$$(Extract\ Product\ Purity) = \frac{\int_0^{t_{step}} u_E(t)C_{E,k}(t)dt}{\sum_{i=1}^{Nc}\int_0^{t_{step}} u_E(t)C_{E,i}(t)dt} \geq Pur_{min} \tag{14}$$

$$(Extract\ Product\ Recovery) = \frac{\int_0^{t_{step}} u_E(t)C_{E,k}(t)dt}{\int_0^{t_{step}} u_F(t)C_{F,k}(t)dt} \geq Rec_{min} \tag{15}$$

$$u_l \leq u_m(t) \leq u_u,\ m = I, II, III, IV \tag{16}$$

In addition to the objective function (throughput) and inequalities defined for this problem, equations (2)-(4) describe the j^{th} column model with $j = 1,\ldots,N_{Column}$ for species i. Here ϵ_b is the void fraction, $C_i^j(x,t)$ is the concentration in the liquid phase, q_i^j is the concentration in the solid phase, $u_m(t)$ is the superficial liquid velocity in zone m, $C_i^{j,eq}(x,t)$ is the equilibrium concentration in the liquid phase, and $K_{appl\,i}$ is the mass transfer coefficient. Equations (5)-(8) relate the superficial velocities $u_D(t)$, $u_E(t)$, $u_F(t)$, $u_R(t)$ (flow rates divided by cross sectional area of column) of the desorbent, extract, feed, and raffinate, respectively, to the zone velocities, along with their respective concentrations. The cyclic state conditions are given by (9)-(13). Also, t_{step} is the valve switching interval, or step time, Pur_{min} and Rec_{min} are the respective purity and recovery requirements for the desired product k recovered in the extract stream. Finally, u_u and u_l are upper and lower bounds on the zone velocities, respectively.

4. Solution Strategy and Results

Problem (1) -(16) can be addressed by a number of strategies. In particular, the PDEs can be discretized in space, e.g., using the method of lines, and integrated in time. This *single discretization* approach leads to DAE-constrained problem that can be solved with control vector parametrization methods. On the other hand, discretizing (1)-(16) in time and space leads to a fully algebraic system. With the *full discretization* approach, we solve a large NLP with very sparse structure. In both cases, the cyclic steady state constraints (9)-(13) are incorporated within the optimization formulation. As described in [1], the single discretization formulation results from applying central differences in the spatial domain; it is solved with commercial software that couples a DAE- and direct sensitivity solver with a reduced space Sequential Quadratic Programming (SQP) algorithm. The full

discretization formulation uses the same spatial discretization and applies Radau collocation on finite elements in the time dimension. This leads to an NLP that is over 50 times larger than with the single discretization approach. The NLP is modeled in AMPL and solved with IPOPT, a large-scale Newton-based barrier method.

Applied to the separation of glucose and fructose with a product specification of 90 % purity and 80 % recovery of fructose with time-independent zone velocities, both approaches achieve a maximum throughput of $0.521 m/h$; this is almost a 50-fold increase over the base case. For this problem the single discretization approach requires almost 2 CPU hours (2.8 GHz P4), while the full discretization approach is almost 75 times faster. Given the efficiency of the full discretization approach, we have also considered a number of more challenging problems. These include optimization of time-dependent zone velocities (so-called *power feed* operation), which leads to a doubling of the throughput, as well as consideration of nonlinear isotherms. More information on these formulations can be found in [1]. Moreover, in [2] we consider a novel superstructure approach which allows the location of four types of time-dependent streams at all nodes in Figure 1. This optimization formulation embeds several nonconventional processes including VARICOL and variable column zone configurations. For the fructose/glucose separation the superstructure formulation leads to novel design with a further 30 % improvement over the power feed case. Further information on these results can be found in [2].

5. Conclusions and Future Work

SMB processes present a number of challenges for PDE-constrained optimization strategies. We summarize a new full discretization approach for this problem that leads to a large NLP that can be solved efficiently with a Newton-based barrier solver. The approach has been found to be very effective in optimization of several SMB systems. Moreover, fast solution of the NLP allows the optimization of SMB superstructures that lead to improved designs over the standard SMB system.

Future work deals with more challenging optimization problems in addition to throughput maximization. In recent work multi-objective formulations have been considered. Also, previous work has assumed symmetric SMB cycles, where cyclic steady state conditions were written over each time step. This approach is being extended to optimization over the entire SMB cycle with asymmetric operation within the cycle. Finally, we intend to investigate the transient operation of SMBs for startup and transitions to different operating points.

References

[1] Y. Kawajiri and L. T. Biegler, *Optimization Strategies for Simulated Moving Bed and PowerFeed Processes*, AIChE Journal, 52(4):1343-1350, (2006).

[2] Y. Kawajiri and L. T. Biegler, *A Nonlinear Programming Superstructure for Optimal Dynamic Operations of Simulated Moving Bed Processes*, I & EC Research, to appear (2006).

Ill-conditioned State Estimation due to low Observability Measure: Analysis and Numerical Preconditioning

Luise Blank

State estimation is an inverse problem and can be formulated based on regularization techniques as an optimization problem where the model equations act as constraints. Commonly regularization techniques concerning the initial data are applied additionally. However, we show that this is not necessary and avoid otherwise possibly occuring bias. Observability of the model system guarantees the well-posedness of the optimization formulation with respect to the L_2- and $L_\infty-$norm. While well-posedness is a qualitative behaviour, we like to discuss also the quantitive behaviour, namely the conditioning. In this context, we shortly review on observability and introduce a new, to our opinion more appropriate, measure of observability based on the concept of condition numbers. In the linear case we show the context to the well known Gramian matrix, and discuss the quantitative dependence of the observability measure on e.g. the length of the horizon and the eigenvalues of the involved stiffness matrix. Moreover, we derive that the introduced observability measure gives a lower bound on the conditioning of the optimization problem. Consequently, inspite of well-posedness a low observability measure causes bad conditioning.

Either one includes an additional regularization term leading to undesired bias in the solution or one faces in the numerical solution high, problem inherent condition numbers and an undesirable eigenvalue disctribution. Employing iterative solvers this leads in general to a very high number of iterations or it may even happen that iterative solvers cannot be applied without efficient preconditioning. While most developed preconditioners deal with the problem of eigenvalues depending on the frequency, hence depending on the discretization level, only very few deal with problem inherent ill-conditioning. We introduce and suggest a Schur-complement based preconditioner, which can be seen as eigenvalue deflation. While typically for eigenvalue deflation knowledge of an invariant subspace is required the suggested Schur-complement based technique requires the knowledge of a subspace containing an invariant subspace. Choosing the appropriate subspace for the Schur-complement in our application is based on the above analytical considerations and on the Riesz-basis property of wavelets.

To eliminate the dependency of the eigenvalues on the discretization level we apply in addition appropriate scaling of the wavelets as is known. The optimization problem is then solved in a nested iteration process for a hierarchy of subproblems. We conclude with numerical examples, which show the drastic reduction in iteration numbers. In some cases the Schur-complement based preconditioning only enabled us solving the resulting system.

References

[1] L. Blank, *State Estimation analysed as Inverse Problem*, Assessment and Future Directions of Nonlinear Model Predictive Control, Springer Lecture Notes in Control and Information Sciences, eds. Allgöwer, Biegler, Findeisen (2006), to appear.

[2] L. Blank, *State Estimation without Regularizing the Initial Data*, Inverse Problems **20** (2004), 1357–1370.
[3] L. Blank, *Preconditioning via a Schur Complement Method: an Application in State Estimation*, SIAM J. Sci. Comput. **25** (2003), 942–960.
[4] T. Binder and L. Blank and W. Dahmen and W. Marquardt, *Multiscale Concepts for Moving Horizon Optimization*, Grötschel and Krumke and Rambau, Online Optimization of Large Scale Systems, Springer, Berlin, (2001), 341–361.

Bilinear Systems and Quantum Optimal Control Problems

Alfio Borzì
(joint work with Julien Salomon, and Stefan Volkwein)

Nowadays we witness a large growing interest in controlling quantum phenomena in a variety of application systems [4, 6]. Present and perspective applications range from quantum optics and quantum chemistry to semiconductor nanostructures. In the last few years these research areas have received further impetus from the emerging fields of quantum computation and quantum communication, aiming at quantum devices where there is the need to manipulate wavefunctions with highest possible precision.

This high-fidelity quantum-state engineering can only be achieved putting together the most sophisticated experimental and theoretical techniques for control of quantum systems. However, within each field of application one has come up with its own strategies and it is only recently that a common consensus has arisen towards the use of optimal control theory [3, 5].

In the optimal control framework, one starts by defining the optimality criteria in the form of a cost functional. For a desired quantum-state transition, this functional will depend on the final state, the need to suppress population of certain states during the control process, as well as other physically motivated constraints, e.g., limited laser resources. The strategy then is to minimize this cost functional while satisfying the constraints of the underlying dynamic equations governing the evolution of quantum states; e.g., the Schrödinger equation. The calculation of the necessary optimality conditions for this optimization problem results in a system of coupled equations to be solved once.

While we focus on quantum optimal control problems we argue that many of the presented results can be extended to general time-dependent bilinear control problems. Bilinear systems were introduced in the theory of automatic control in the 1960's for electrical engineering applications. They represent a class of nonlinear control strategies with the aim to obtain better system response than possible with linear control. In general, the solution of most bilinear systems poses challenging theoretical and computational problems which are open or have been only partially addressed. This is in particular true for the control of the finite-level quantum mechanical systems discussed in this talk.

It is the purpose of this talk to present a detailed formulation of a class of optimal control problems for finite-level quantum systems [1] and to address their

efficient solution by iterative methods. We prove existence of solutions to the optimal control problems, and investigate first-order necessary optimality conditions and second-order sufficient optimality conditions. We review state-of-the-art monotonic iterative schemes and their convergence properties and use these schemes as benchmark for a new solution procedure that we propose in this paper. This procedure results from combining our extension of a newly proposed nonlinear conjugate gradient method [2] with a cascadic acceleration scheme. Convergence of the proposed nonlinear conjugate gradient method is proved. Numerical results demonstrate robustness and efficiency of the proposed approach.

References

[1] A. Borzì, G. Stadler, and U. Hohenester, *Optimal quantum control in nanostructures: Theory and application to a generic three-level system*, Phys. Rev. A **66** (2002), 053811.

[2] Y.H. Dai and Y. Yuan, *A nonlinear conjugate gradient with a strong global convergence property*, SIAM Journal on Optimization **10** (1999), 177–182.

[3] H.O. Fattorini, *Infinite Dimensional Optimisation and Control Theory*, Cambridge University Press, Cambridge, 1999.

[4] U. Hohenester, *Optical properties of semiconductor nanostructures: Decoherence versus quantum control*, to appear in M. Rieth and W. Schommers, (Eds.), *Handbook of Theoretical and Computational Nanotechnology* (American Scientific Publishers, 2005); cond-mat/0406346.

[5] J.L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer, Berlin, 1971.

[6] H. Rabitz, G. Turinici, and E. Brown, *Control of Molecular Motion: Concepts, Procedures, and Future Prospects*, Ch. 14 in Handbook of Numerical Analysis, Volume X, P. Ciarlet and J. Lions, Eds., Elsevier, Amsterdam, 2003.

Fast Wavelet Schemes for Control Problems Constrained by Elliptic PDEs

Carsten Burstedde

(joint work with Angela Kunoth, and Roland Pabel)

The treatment of many tasks in PDE-constrained optimization involves the solution of linear elliptic subproblems. The areas of applications include heat conduction, diffusion, elasticity and superconductivity. Because of the increased structural complexity compared to the solution of an isolated PDE, highly efficient numerical methods are necessary. A central motivation in this context is to achieve optimal computational complexity, which means that the effort for the numerical solution of the problem should be linear in the amount of unknowns, with preferably small constants. This implies the use of asymptotically optimal preconditioning schemes.

Thus, our first goal in the recent years has been to develop the wavelet framework towards a level of practicability which is comparable with multigrid finite element methods on simple domains in terms of efficiency and the absolute size of the problem to be treated. In the absence of inequality constraints, we believe to have achieved this with the development of a working code which is fully functional for

uniform discretizations of both distributed and Dirichlet boundary control problems. It uses inner-outer inexact conjugate gradient iterations combined with a nested iteration strategy. The wavelet ansatz delivers a preconditioning scheme which is optimal in the sense of uniform bounds, and optimized with respect to the absolute values of the constants involved. On a standard computer, results for up to $6 \cdot 10^6$ unknowns have been obtained (this number does not yet include the various auxiliary and temporary variables), with linear run-time and memory requirements [1, 4, 8, 10].

The second goal is then to make systematic use of the strong analytical qualities of wavelets, and their formulation in an infinite-dimensional function space setting, to provide adaptivity and rigorous error control. A proof for the convergence rates of an adaptive wavelet method for linear elliptic control problems has been obtained in [7]. This approach has been generalized to Dirichlet boundary control [9]. While the central element in these proofs are nested Richardson iterations using operators of bounded (and small) condition numbers, the algorithm implemented in [1] merges the fast conjugate gradient iterations already used in the framework for uniform discretizations with ingredients from optimal adaptive wavelet methods [5]. This implementation permits a quantitative evaluation, and evidence for superlinear convergence has been found in one to three dimensions [1,3]. Numerical experiments show that this adaptive wavelet scheme resolves the control, state and adjoint variables with different index sets adapted to their respective singularities. Also singularities with anisotropic structure are respected by appropriate distributions of coefficients (this is called dimension-adaptivity in the sparse grid context).

The wavelet framework is also well-suited for modeling issues related to the evaluation of Sobolev norms. Since whole ranges of Sobolev scales are numerically accessible, integer and fractional as well as primal and dual norms can be computed in linear time. This is implemented by the construction of appropriate Riesz operators. Care has been taken that this unified scheme delivers results identical to those obtained with standard finite element in the special cases of non-negative integer norms. Other smoothness classes are interpolated equivalently and calibrated by proposing plausible criteria [1, 2].

The third goal is characterized by future plans. Nonlinear constraints are expected to be tractable by SQP approaches, which reduce to a sequence of linear control problems of the type covered above. An adaptive analysis and implementation can be based on [6]. Moreover, inequality constraints are considered, where bounds on the control are more directly accessible than those on the state. Concerning the class of primal-dual active set methods, adaptive wavelets are in principle capable of resolving boundary submanifolds between active and inactive sets. However, pointwise bounds interfere with the fundamental multilevel characterization based on difference spaces. The class of barrier and interior point methods is known to have conditioning issues in the limit of zero penalization, where the norm equivalences available in the wavelet framework may be extendable in a way which matches the problem formulation via barrier terms. Currently,

both approaches are being investigated, since it is not clear which benefits most from the theoretical background available in the wavelet setting.

References

[1] C. Burstedde, *Fast Optimised Wavelet Methods for Control Problems Constrained by Elliptic PDEs*, Dissertation, Institut für Angewandte Mathematik, Universität Bonn, 2005.

[2] C. Burstedde, *On the numerical evaluation of general Sobolev norms*, Preprint submitted for publication, 2006.

[3] C. Burstedde and A. Kunoth, *Variants of adaptive wavelet methods for elliptic PDEs*, Manuscript in preparation, 2006.

[4] C. Burstedde and A. Kunoth, *Fast iterative solution of elliptic control problems in wavelet discretization*, Preprint No. 127, SFB 611, Universität Bonn, Dec. 2003. Revised and accepted for publication in J. Comp. Appl. Math., August, 2005.

[5] A. Cohen, W. Dahmen, and R. DeVore, *Adaptive wavelet methods for elliptic operator equations – convergence rates*, Math. Comp., **70**:27–75, 2001.

[6] A. Cohen, W. Dahmen, and R. DeVore,*Adaptive wavelet schemes for nonlinear variational problems*, SIAM J. Numer. Anal., **41**(5):1785–1823, 2003.

[7] W. Dahmen and A. Kunoth, *Adaptive wavelet methods for linear-quadratic elliptic control problems: convergence rates*, SIAM J. Contr. Optim., **43**(5):1640–1675, 2005.

[8] A. Kunoth, *Fast iterative solution of saddle point problems in optimal control based on wavelets*, Comput. Optim. Appl., **22**:225–259, 2002.

[9] A. Kunoth, *Adaptive wavelet schemes for an elliptic control problem with Dirichlet boundary control*, Numer. Algor., **39**(1-3):199–220, 2005.

[10] R. Pabel, *Wavelet methods for a PDE constrained control problem with Dirichlet boundary control*, Diploma thesis, Institut für Angewandte Mathematik, Universität Bonn, 2005, handed in and being reviewed.

Simulation of an Optimal Design for Maximal Torsion Stiffness with a Converging Adaptive FEM

Carsten Carstensen

(joint work with Sören Bartels)

1. Overview

The adaptive finite element method (AFEM) is considered for a problem in topology optimisation. The optimal design problem for maximal torsion stiffness of an infinite bar of given geometry $\Omega \times \mathbb{R}$ aims to describe the unknown distribution of two materials of prescribed amounts. The relaxation of the associated energy densities leads to a degenerated convex minimization problem which caused some attention by [12, 13, 16, 17, 19, 20, 24, 25] amongst others.

The problem is recast in Section 2 and leads to some convexified minimization problem of the form

$$E(v) := \int_\Omega W(Dv)\,dx + \int_\Omega v\,dx \quad \text{for all } v \in V := H^1_0(\Omega)$$

Notice that the right hand side is constant an whence there is no data oscillation. The edge-oriented AFEM therefore consists of a loop: On each level ℓ, given V_ℓ, do

$$\text{SOLVE} \Rightarrow \text{ESTIMATE} \Rightarrow \text{MARK} \Rightarrow \text{REFINE}$$

to generate $V_{\ell+1}$ based on $u_\ell := \arg\min E(V_\ell)$, f, V_ℓ, a posteriori error estimators.

Section 3 analyses the convergence of the algorithm in the spirit of [5, 8, 15, 22, 23, 26, 27]

2. Mathematical Modelling

Given a simply-connected bounded 2D Lipschitz domain $\Omega \subset \mathbb{R}^2$ consider the infinite cylinder $\Omega \times \mathbb{R}$ under torsion. The dsign Task is to maximise the torsion stiffness for an optimal composition of the prescribed section Ω with two materials of reciprocal shear stiffness $0 < \mu_1 < \mu_2 < \infty$ located at Ω_1 and Ω_2:

$$\Omega = \Omega_1 \cup \Omega_2 \quad \text{and} \quad \Omega_1 \cap \Omega_2 = \emptyset.$$

Side restriction with given $0 < \Theta < 1$ on the area $|\cdot|$:

$$|\Omega_1| = \Theta\,|\Omega| \quad \text{and} \quad |\Omega_2| = (1-\Theta)|\Omega|.$$

Design optimal shear modulus $\mu(x)^{-1}$ distribution ($\mu_\Theta := \Theta\,\mu_1 + (1-\Theta)\mu_2$)

$$\mu \in \mathcal{M}_\Theta := \{\mu \in L^\infty(\Omega) : \mu = \mu_{1,2} \text{ a.e. in } \Omega,\ \int_\Omega \mu(x)\,dx = \mu_\Theta\,|\Omega|\}$$

and then define $\Omega_j := \{x \in \Omega : \mu(x) = \mu_j\}$ up to sets of measure zero.
The reciprocal torsion T^{-1} is also given by a minimisation problem for the 2D stress vector $\sigma = (\sigma_1, \sigma_2)$ in the section Ω,

$$T^{-1} = \inf_{\sigma \in \Sigma} E/m^2$$

with the elastic energy E and the resulting 2D moment m

$$E = \frac{1}{2}\int_\Omega \mu(x)|\sigma(x)|^2\,dx \quad \text{and} \quad m = \int_\Omega (x_2\sigma_1 - x_1\sigma_2)\,dx$$

The stress field $\sigma \in L^2(\Omega;\mathbb{R}^2)$ satisfies

$$\mathrm{div}\sigma = 0 \text{ in } \Omega \quad \text{and} \quad \sigma \cdot \nu = 0 \text{ along } \partial\Omega$$

Hence there exists $u \in V := H^1_0(\Omega)$ s.t.

$$\sigma = (-\partial u/\partial x_2, \partial u/\partial x_1).$$

This leads to some reformulations in terms of $v \in V$,

$$m = -\int_\Omega (x_2\partial u/\partial x_2 + x_1\partial u/\partial x_1)^2\,dx = 2\int_\Omega u\,dx,$$

and so to the representation

$$8T^{-1} = \inf_{v \in V} \frac{\int_\Omega \mu|Dv|^2\,dx}{(\int_\Omega v\,dx)^2}.$$

Some minimiser $u \in V$ is given as the weak solution of

$$-\mathrm{div}(\mu Du) = 1 \text{ in } \Omega$$

and this unique u attains the minimum in

$$-\frac{T}{16} = \min_{v\in V}\Big(\frac{1}{2}\int_\Omega \mu|Dv|^2\,dx - \int_\Omega v\,dx\Big).$$

Select some minimiser $u \in V$ as the weak solution of

$$-\mathrm{div}(\mu Du) = 1 \text{ in } \Omega$$

and this unique u attains the minimum in

$$-\frac{T}{16} = \min_{v\in V}\Big(\frac{1}{2}\int_\Omega \mu|Dv|^2\,dx - \int_\Omega v\,dx\Big).$$

For $\mu \in \mathcal{M}$ and $v \in V := H_0^1(\Omega)$ set

$$F(\mu, v) := \frac{1}{2}\int_\Omega \mu|Dv|^2\,dx - \int_\Omega v\,dx.$$

Then, the optimal design problem for the maximal torsion stiffness reads:

$$(M) \quad \inf_{\mu\in\mathcal{M}_\Theta}\inf_{v\in V} F(\mu, v).$$

Solutions exists in terms of Young measures and can be recovered from:

$$(S) \quad \sup_{\lambda\in\mathbb{R}}\inf_{v\in V} G(\lambda, v)$$

with

$$G(\lambda, v) = \int_\Omega g_\lambda(|Dv|)\,dx + \int_\Omega v\,dx + \lambda\mu_\Theta|\Omega|$$

One Theorem by Kohn&Strang shows (M) $\Leftrightarrow$ (S).
Since g_λ is nonconvex the infimum may in fact be not attained. Infimal value computable with *degenerated convex minimisation problem* for

$$E(v) := \int_\Omega W(Dv)\,dx + \int_\Omega v\,dx \quad \text{for all } v \in V := H_0^1(\Omega)$$

where g_λ is replaced by its convex hull $W(\cdot) := g_\lambda^{**}(|\cdot|)$

$$g_\lambda^{**}(t) := \begin{cases} \mu_2(t^2/2 - \lambda) & \text{für } t \le t_1, \\ \sqrt{t_1 t_2 \mu_1 \mu_2}\,t - \lambda(\mu_1 + \mu_2) & \text{for } t_1 \le t \le t_2, \\ \mu_1(t^2/2 - \lambda) & \text{für } t_2 \le t. \end{cases}$$

($t_1 := \sqrt{2\lambda\mu_1/\mu_2}$ and $t_2\mu_1 := \mu_2 t_1$)

3. Convergence of AFEM

The numerical analysis of this degenerated minimization problem is delicate for possibly multiple primal variables u but unique derivatives $\sigma := DW(\nabla u)$. A refined a posteriori error estimate still suffers from the reliability-efficiency gap. However, it motivates a simple edge-based adaptive mesh-refining algorithm (AFEM) that is not a priori guaranteed to refine everywhere. Its convergence is based on energy estimates and some refined convexity control. Numerical experiments illustrate even optimal convergence rates of the proposed AFEM.
This section ends with a few remarks on the convergence proof. This is based on some Refined convexity control in the sense of

$$|DW(A) - DW(B)|^2 \lesssim W(A) - W(B) - DW(A) \cdot (B - A)$$

An application with $A = Du_\ell$ and $B = Du$ yields

$$\|\sigma - \sigma_\ell\|L^2(\Omega)^2 + \delta_\ell \lesssim R_\ell(u - u_\ell)$$

with

$$\delta_\ell := E(u_\ell) - E(u) \quad \text{and} \quad R_\ell(v) := \int_\Omega (v + \sigma_\ell \cdot Dv)dx$$

In the next steo oone proves a reliabile estimate of $\|R_\ell\|_{V^*}$ as for linear problems

$$\|R_\ell\|^2_{V^*} \lesssim \sum_{E\in E_\ell} \eta_E^2$$

The bulk criterion in the step MARK leads (as for linear problems)

$$\|R_\ell\|^2_{V^*} \lesssim \sum_{E\in M_\ell} \eta_E^2$$

Discrete local efficiency with inner node property for $E \in M_\ell$ as in the work [3,27] allows the proof of

$$\eta_E^2 \lesssim \|\sigma_{\ell+1} - \sigma_\ell\|L^2(\omega_E)^2$$

A second application of refined convexity control leads to

$$\|\sigma_{\ell+1} - \sigma_\ell\|L^2(\Omega)^2 \lesssim \delta_\ell - \delta_{\ell+1}$$

which is actually known from [12, 13]. Altogether one deduces

$$\|\sigma - \sigma_\ell\|L^2(\Omega)^4 + \delta_\ell^2 \lesssim \delta_\ell - \delta_{\ell+1}$$

The preceeding arguments allow a proof of the following theorem:
Theorem: $\exists \kappa < 1$ s.t.

$$\kappa \|\sigma - \sigma_\ell\|L^2(\Omega)^4 + \delta_{\ell+1} \le (1 - \kappa\delta_\ell)\delta_\ell.$$

More details and numericla examples shall appear in [2].

References

[1] Allaire, G. (2002) *Shape optimization by the homogenization method.* Springer-Verlag, New York.

[2] Bartels, S. and Carstensen, C. (2006). *AFEM for an optimal design problem.*

[3] Bochetto.

[4] Brenner, S.C. and Carstensen, C. (2004) *Finite Element Methods.* Chapter 4 in Encyclopedia of Computational Mechanics, John Wiley and Sons, 2004.

[5] Binev, P., Dahmen, W. and DeVore, R. (2004). *Adaptive finite element methods with convergence rates.* Numer. Math., **97**, 2, 219–268.

[6] Carstensen, C. (2004). *Some remarks on the history and future of averaging techniques in a posteriori finite element error analysis.* ZAMM Z. Angew. Math. Mech., **84**, 1, 3–21.

[7] Carstensen, C. (2004). *Convergence of adaptive finite element methods in computational mechanics.* Proceedings of the Sixth World Congress on Computational Mechanics, September 2004.

[8] Carstensen, C. and Hoppe, R.W.H. (2005). *Error Reduction and Convergence for an AMFEM, ANCFEM, 2D eddy current...* 2005ff

[9] Carstensen, C. and Jochimsen, K. (2003). *Adaptive finite element methods for microstructures? Numerical experiments for a 2-well benchmark.* Computing, **71**, 2, 175–204.

[10] Carstensen, C. and Müller, S. (2002). *Local stress regularity in scalar nonconvex variational problems.* SIAM J. Math. Anal., **34**, 2, 495–509.

[11] Carstensen, C. and Orlando, A. and Valdman, J. (2005). *A convergent adaptive finite element method for the primal problem of elastoplasticity.* Preprint 2005-12.

[12] Carstensen, C. and Plecháč, P. (1997). *Numerical solution of the scalar double-well problem allowing microstructure.* Math. Comp., **66**, 219, 997–1026.

[13] Carstensen, C. and Plecháč, P. (2000). *Numerical analysis of compatible phase transitions in elastic solids.* SIAM J. Numer. Anal., **37**, 6, 2061–2081.

[14] Cherkaev, A. (2000) *Variational methods for structural optimization* Springer-Verlag, New York.

[15] Dörfler, W. (1996). *A convergent adaptive algorithm for Poisson's equation.* SIAM J. Numer. Anal., **33**, 3, 1106–1124.

[16] French, D. (1989). *On the convergence of finite element approximations of relaxed variational problem.* IMA Preprint 503, Minneapolis 1989.

[17] Goodman, J. and Kohn, R.V. and Reyna, L. (1986). *Numerical study of a relaxed variational problem from optimal design.* Comput. Methods Appl. Mech. Engrg., **57**, 1, 107–127.

[18] Hiriart-Urruty, J.-B. (1998). *Optimisation et Analyse Convexe.* Presses Universitaires de France, Paris.

[19] Kohn, R. and Strang, G. (1986). *Optimal design and relaxation of variational problems I, II, III.* Comm. Pure Appl. Math. **39** (1986) pp. 113-137, 139-182, 353-377.

[20] Kawohl, B. and Stara, J. and Wittum, G. (1991). *Analysis and numerical studies of a shape design problem.* Archive for Rational Mechanics, 114, 349-363 (1991).

[21] Morin, P. and Nochetto, R.H. and Siebert, K.G. (2000). *Data oscillation and convergence of adaptive FEM.* SIAM J. Numer. Anal., **38**, 2, 466–488.

[22] Morin, P. and Nochetto, R.H. and Siebert, K.G. (2003). *Local problems on stars: a posteriori error estimators, convergence, and performance.* Math. Comp., **72**, 243, 1067–1097.

[23] Morin, P. and Nochetto, R.H. and Siebert, K.G. (2003). *Convergence of adaptive finite element methods.* SIAM Review **44** (2003) 631–658.

[24] Murat, F. and Tartar, L. (1985). *Optimality conditions and homogenization,* in: *Nonlinear Variational Problems.* Eds.: A. Marino, L. Modica, S. Spagnolo and M. Degiovanni, Pitman Research Notes in Math. **127** (1985) pp. 1-8.

[25] Murat, F. and Tartar, L. (1985). *Calcul des variations et homogenization,* in: *Les méthodes de l'homogeneisation: théorie et applications en physique.* Eds.: D. Bergman et al. Collection de la Direction des Études et Recherches d'Electricité de France **57** (1985) pp. 319-369.

[26] Stevenson, R. (2005) *Optimality of AFEM.* Preprint 2005.
[27] Veeser, A. (2002). *Convergent adaptive finite elements for the nonlinear Laplacian.* Numer. Math., **92**, 4, 743–770.
[28] Zeidler, E. (1988). *Nonlinear functional analysis and its applications. IV Applications to Mathematical Physics.* Springer-Verlag, New York.

Optimal Motions of Multibody Systems in Resistive Media

Felix L. Chernousko

Various non-conventional principles of motion of multibody systems in resistive media are considered. These principles do not imply using wheels, legs, tracks, screws, propellers, jets, and other devices and are based on relative motions of some parts of the moving bodies in the presence of external resistance forces. Namely, two kinds of motions are investigated:

— snake-like locomotion of multilink systems;

— progressive movement of a body caused by internal motions of a mass inside this body.

Both types of motion are possible only in the presence of external resistance forces such as Coulomb's dry friction or viscous resistance.

For the snake-like locomotion, multilink mechanisms equipped with actuators installed at the joints are considered. The mechanism can move along a horizontal plane in the presence of Coulomb's dry friction between the mechanism and the plane. The locomotion is a result of periodic twisting of the mechanism at its joints controlled by the actuators. For two-link and three-link systems, these periodic motions consist of alternating slow and fast phases. Multilink mechanisms with more than four links can perform smooth wave-like locomotion.

Optimal geometrical and mechanical parameters of the multilink mechanisms as well as optimal parameters of their periodic motions are obtained that maximize the average speed of locomotion. The optimization is carried out numerically. It is shown that the optimization leads to a considerable gain in the locomotion speed. The results of optimization are compared with the experimental data. The results of computer simulation, animation and videofilms are presented.

Another kind of motion is caused by a periodic displacement of a certain mass inside the moving body. Here, various types of external resistance forces are examined: Coulomb's dry friction, linear and quadratic resistance forces depending on the velocity of the body. Both symmetric (isotropic) and asymmetric (anisotropic) cases of resistance are considered; in the latter case, the resistance force for the forward and backward motions of the body are different.

Optimal relative periodic motions of the internal mass are obtained that result in the maximum average speed of the moving body. The analytical and numerical results of the optimization are presented. Experiments confirm the obtained theoretical results. Videofilms of experiments are demonstrated.

The obtained results are of interest with respect to mobile robotic systems, especially, for mini-robots that can move inside tubes or in liquids.

References

[1] F.L. Chernousko, *The motion of a multilink system along a horizontal plane*, Journal of Applied Mathematics and Mechanics **64**, No. 1, (2000), 5–15.

[2] F.L. Chernousko, *The wavelike motion of a multilink system on a horizontal plane*, Journal of Applied Mathematics and Mechanics **64**, No. 4, (2000), 497–508.

[3] A.S. Smyshlayev, F.L. Chernousko. *Optimization of the motion of multilink robots on a horizontal plane*, Journal of Computer and Systems Science International **40**, No. 2, (2001), 340–348.

[4] F.L. Chernousko. *The optimum rectilinear motion of a two-mass system*, Journal of Applied Mathematics and Mechanics **66**, No. 1, (2002), 1–7.

[5] F.L. Chernousko. *Snake-like locomotions of multilink mechanisms*, Journal of Vibration and Control **9**, No. 1-2, (2003), 237–256.

[6] F.L. Chernousko. *On the motion of a body containing a movable internal mass*, Doklady Physics **50**, No. 11, (2005), 593–597.

Shape Optimization for Elliptic PDE: Coercivity, Convergence and Stability of Numerical Methods

Karsten Eppler

(joint work with Helmut Harbrecht)

We consider the numerical solution of the following simple class of problems. For a sufficiently large closed "security set" $D \subset \mathbb{R}^n$ we consider the set of all simply connected (more general: with a fixed finite number of connected components) and bounded domains $\Omega \subset \operatorname{int} D$ with sufficiently regular boundary $\Gamma = \partial\Omega$, $\Gamma \in C^{2,\alpha}$, for example. Then, for a given elliptic second order operator L with associated boundary operator B, an optimization problem is investigated as follows

$$\begin{aligned} J(\Omega; u_\Omega) &= \int_\Omega j(x, u_\Omega, \nabla_x u_\Omega)dx \to \min, \quad \text{subject to} \\ Lu &= f \text{ in } \Omega, \qquad Bu = g \text{ on } \Gamma, \\ \int_\Omega h_i(x)dx &\le c_i, \quad i = 1(1)n, \; \int_\Gamma h_j d\sigma \le c_j, \quad j = 1(1)m. \end{aligned}$$

Here, the data f, g, h_i and j are defined on D and $D \times \mathbb{R}^{n+1}$, respectively. Typical domain and boundary integral constraints in applications are the Volume $V(\Omega)$ and the perimeter $P(\Omega)$ of the domain. Contrary to classical control problems, the domains itself, resp. their boundaries Γ now serve as the optimization variable(s).

Based on a related shape calculus [1, 16], the aim is to develop efficient optimization algorithms for the solution of these problems with particular emphasis on second order methods. Complete boundary integral representations for the shape gradient and the shape Hessian provides uptdate rules directly for the boundary [2, 3]. In some specific cases, integral equation methods using wavelet compression techniques turn out to be a powerful tool [15]. In particular, applications have been investigated from planar elasticity [4, 5], from exterior electromagnetic

shaping of liquid metals [6,8], from shape identification in electrical impedance tomography [9,10] as well as from free boundary problems in electrochemical processing [11]. Furthermore, efficient and accurate numerical algorithms are developed by using either BEM and/or FEM/BEM-coupling tools for compactly supported integral objectives with applications in fluid dynamics and magnetostatics [12,13].

Moreover, accuracy and efficiency of computations in shape optimization were substantially increased during the last years for several other approaches as well by using advanced numerical methods. Nevertheless, despite of improved accuracy, essentially different behaviour of optimization algorithms was observed for formally (almost) similar classes of problems. We will illustrate this for the abovementioned examples. It is shown in the talk, that the different behaviour of numerical algorithms heavily depends on the nature of the underlying shape problem, characterized by the coercivity properties of the shape Hessian at stationary domains. Finally, a concept is discussed for proving convergence of optimal solutions of finite dimensional auxiliary problems to the optimal shape of the original problem in case of *well-posed problems* [14].

References

[1] M. Delfour and J.-P. Zolesio, *Shapes and Geometries*, SIAM, Philadelphia, 2001.

[2] K. Eppler, *Boundary integral representations of second derivatives in shape optimization*, Discussiones Mathematicae (Differential Inclusion Control and Optimization), **20**:63–78, 2000.

[3] K. Eppler, *Optimal shape design for elliptic equations via BIE-methods*, Journal of Applied Mathematics and Computer Science, **10**:487–516, 2000.

[4] K. Eppler and H. Harbrecht, *Numerical solution of elliptic shape optimization problems using wavelet-based BEM*, Optim. Methods Softw., **18**:105-123, 2003.

[5] K. Eppler and H. Harbrecht, *2nd Order Shape Optimization using Wavelet BEM*, Preprint 06-2003, TU Berlin, 2003 to appear in Optim. Methods Softw.

[6] K. Eppler and H. Harbrecht, *Exterior Electromagnetic Shaping using Wavelet BEM*, Math. Meth. Appl. Sci., **28**:387–405 (2005).

[7] K. Eppler and H. Harbrecht, *Second order Lagrange multiplier approximation for constrained shape optimization problems*, in *Control and Boundary Analysis*, (Lecture Notes in Pure and Applied Mathematics), edited by J. Cagnol et. al., Marcel Dekker, Boca Raton, **240**:107–118, 2005.

[8] K. Eppler and H. Harbrecht, *Fast wavelet BEM for 3d electromagnetic shaping*, Appl. Numer. Math., **54**:537–554 (2005).

[9] K. Eppler and H. Harbrecht, *A regularized Newton method in electrical impedance tomography using shape Hessian information*, Control & Cybernetics **34**:203–225, 2005.

[10] K. Eppler and H. Harbrecht, *Shape optimization in 3D electrical impedance tomography*, WIAS-Preprint 963, WIAS Berlin, 2004, to appear in Proceeding of IFIP WG 7.2 conference.

[11] K. Eppler and H. Harbrecht, *Efficient treatment of stationary free boundary problems*, WIAS-Preprint 965, WIAS Berlin, 2004, submitted to Appl. Numer. Math.

[12] K. Eppler and H. Harbrecht, *Compact gradient tracking in shape optimization*, WIAS-Preprint 1054, WIAS Berlin, 2005, submitted to Comput. Optim. App.

[13] K. Eppler and H. Harbrecht, *Coupling of FEM and BEM in shape optimization*, WIAS-Preprint 1029, WIAS Berlin, 2005, to appear in Numer. Math.

[14] K. Eppler, H. Harbrecht, and R. Schneider, *On Convergence in Elliptic Shape Optimization*, WIAS-Preprint 1016, WIAS Berlin, 2005, submitted to SIAM J. Control Optim.

[15] H. Harbrecht, and R. Schneider, *Biorthogonal wavelet bases for the boundary element method*, Math. Nachr. 269/270:167–188, 2004.
[16] J. Sokolowski and J.-P. Zolesio, *Introduction to Shape Optimization*, Springer, Berlin, 1992.

Numerical Shape Optimization for Aircraft Design - Status and Future Challenges

Norbert Kroll and Nicolas R. Gauger

This presentation is a survey lecture about the activities in aircraft design at the Institute of Aerodynamics and Flow Technology of DLR and in particular about adjoint based methods.

There are many ingredients required to establish an efficient and flexible numerical optimization capability for aircraft design. These include suitable techniques for geometry parametrization, handling of geometrical and physical constraints, meshing and mesh deformation as well as mesh adaptation methods, efficient and accurate flow solvers as well as a flexible tool-set containing both deterministic and non-deterministic optimization strategies [1].

Over the last years, numerical shape optimization is one of the major issues of the Institute of Aerodynamics and Flow Technology at DLR. The presentation will focus on both algorithmic developments as well as realistic applications. Several applications including cruise and high-lift flight configurations will be shown.

One key activity is the derivation and implementation of continuous [2] and discrete [3] adjoint approaches for the DLR flow simulation software MEGAFLOW [4] based on the solution of the compressible Euler and Navier-Stokes equations. There are the MEGAFLOW solvers FLOWer, coded in Fortran and working on structured meshes, as well as TAU, coded in C and working on hybrid meshes. Initial activities are launched for the automated generation of discrete adjoint solvers by the use of AD tools. For the FLOWer code the AD tool TAF [5] is used and ADOL-C [6] for the TAU code.

Current status and future needs in algorithmic development for large scale optimization problems will be discussed. One of the future challenges in aircraft design addresses multi-objective and multi-disciplinary optimization. In this context, a coupled aero-structure adjoint formulation for efficient multi-disciplinary wing optimization and its application for maximum aircraft range will be presented [7]. As further future challenges we identify the efficient calculation of Pareto fronts in multi-objective design, as well as the efficient evaluation of numerical uncertainties and robust design cases.

Acknowledgements. The authors acknowledge the contributions of their colleagues at DLR, namely J. Brezillon, O. Brodersen, R. Dwight, A. Fazzolari, T. Gerhold, U. Herrmann, W. Khier, J. Raddatz, T. Schwarz, D. Vollmer, M. Widhalm and J. Wild. Furthermore, the authors acknowledge the contributions of R. Giering and T. Kaminski from FastOpt, F. Deister and W. Haase from EADS-M as well as A. Walther from TU Dresden.

References

[1] N. Kroll, N. Gauger, *Ongoing activities in shape optimization within the German Project MEGADESIGN*, Proceedings of ECCOMAS 2004, Jyväskylä, Finland, 2004.

[2] J. Brezillon and N. Gauger, *2D and 3D Aerodynamic Shape Optimization using the Adjoint Approach*, Aerospace Science and Technology, Vol. 8, No. 8, pp. 715-727, 2004.

[3] J. Brezillon, R. Dwight, *Discrete adjoint of the Navier-Stokes equations for aerodynamic shape optimization*, Proceedings of EUROGEN05, Munich, 2005.

[4] N. Kroll, C.C. Rossow, D. Schwamborn, K. Becker and G. Heller, *MEGAFLOW - a Numerical Flow Simulation Tool for Transport Aircraft Design*, ICAS 2002-1.10.5, 23rd International Congress of Aeronautical Sciences, Toronto, 2002.

[5] R. Giering, T. Kaminski, T. Slawig, *Applying TAF to a Navier-Stokes solver that simulates an Euler flow around an airfoil*, Future Generation Computer Systems 21(8),Elsevier 2005.

[6] A. Griewank, D. Juedes, H. Mitev, J. Utke, O. Vogel, A. Walther, *ADOL-C: A Package for the Automatic Differentiation of Algorithms Written in C/C++*, Technical University of Dresden, Institute of Scientific Computing and Institute of Geometry, 1999.

[7] A. Fazzolari, N. Gauger, J. Brezillon, *An aero-structure adjoint formulation for efficient multidisciplinary wing optimization*, Proceedings of EUROGEN05, Munich, 2005.

A Posteriori Estimates for Optimal Control Problems with Control Constraints

Alexandra Gaevskaya

(joint work with Ronald H.W. Hoppe, and Sergey I. Repin)

In this talk, we present two approaches to a posteriori analysis of optimal control problems associated with elliptic type partial differential equations with control constraints. The analysis is performed for distributed and boundary control problems of the following form:

Problem 1. Given $\psi \in L_\infty(\Omega)$, $y^d \in L_2(\Omega)$, $u^d \in L_2(\Omega)$, $f \in L_2(\Omega)$, and $a \in \mathbb{R}_+$, consider the distributed control problem

$$
\begin{aligned}
&\text{minimize } J(y(u),u) := \frac{1}{2}\,\|y-y^d\|^2 + \frac{a}{2}\,\|u-u^d\|^2 \\
&\text{over } (y,u) \in Y := H_0^1(\Omega)\times L^2(\Omega)\,, \\
&\text{subject to } -\Delta\, y = u+f \quad \text{a.e. in } \Omega\,, \\
&\qquad u \in K := \{v \in L^2(\Omega) \mid v \le \psi \text{ a.e. in } \Omega\}\,.
\end{aligned}
$$

Problem 2. Given $\psi_1, \psi_2 \in L_\infty(\Gamma_N)$, $y^d \in L_2(\Omega)$, $u^d \in L_2(\Gamma_N)$, $f \in L_2(\Omega)$, and $a \in \mathbb{R}_+$, consider the boundary control problem

$$\begin{aligned}
&\text{minimize } J(y(u),u) := \frac{1}{2}\|y-y^d\|^2 + \frac{a}{2}\|u-u^d\|^2_{L_2(\Gamma_N)} \\
&\text{over } (y,u) \in Y := H^1_{0,\Gamma_D}(\Omega) \times L^2(\Gamma_N)\ , \\
&\text{subject to } -\Delta\ y\ =\ f \quad \text{a.e. in } \Omega\ , \\
&\qquad\qquad \frac{\partial y}{\partial n}\ =\ u \quad \text{on} \quad \Gamma_N\ , \\
&\qquad u \in K\ :=\ \{v \in L^2(\Gamma_N)\ |\psi_1 \le\ v \le \psi_2 \text{ a.e. on } \Gamma_N\}\ .
\end{aligned}$$

The first approach is based on residual-type a posteriori error estimators and incorporates data oscillations. In the framework of this approach, we construct an adaptive finite element method that consists of successive loops of the sequence

$$\text{SOLVE} \to \text{ESTIMATE} \to \text{MARK} \to \text{REFINE}$$

Here, the step SOLVE stands for the numerical solution of the finite element discretized problem, whereas the following three steps include the implementation of the residual-type estimator (ESTIMATE), appropriate marking strategy (MARK), and refinement process (REFINE). Up to data oscillations, the residual-type error estimator does provide an upper and a lower bound for the errors in the state, the co-state, the control, and the co-control (see, e.g., [1–3]). In [1], we provide convergence of the adaptive scheme based on the residual-type estimator in terms of guaranteed error reduction in the state, the co-state, the control, and the co-control for the case of distributed control problem.

The second approach involves the so-called functional type a posteriori error estimates that provide sharp upper bounds for the error with respect to any feasible approximation of the state (see, e.g., [4], [5]). Using functional type estimates, we obtain directly computable upper bounds for the cost functionals of the respective optimal control problems (the majorants). It is proved that a numerical strategy based upon using the majorants produces sequences of control and state functions which provide a value of the cost functional as close to the optimal value as it is required. Moreover, the respective sequences of control and state functions tend to the desired solution of the original problem (see [6]). We further note that the majorants can be used to find guaranteed and easily computable upper bounds for the cost functional when the optimization problem is solved by known methods.

At the end, we present results of numerical experiments that illustrate the performance of both approaches.

References

[1] A. Gaevskaya, R.H.W.Hoppe, Y. Iliash, and M. Kieweg, *Convergence analysis of an adaptive finite element method for distributed control problems with control constraints*, to appear in Proc. Conf. Optimal Control for PDEs, Oberwolfach, April 17-13, 2005 (G. Leugering et al.: eds.), Birkhaeuser, Basel.

[2] M. Hintermueller, R.H.W. Hoppe, Y. Iliash, and M. Kieweg, *An a posteriori error analysis of adaptive finite element methods for distributed elliptic control problems with control constraints*, to appear in ESAIM.
[3] A. Gaevskaya, R.H.W.Hoppe, Y. Iliash, and M. Kieweg, *A posteriori error analysis of adaptive finite element methods for control constrained distributed and boundary control problems*, to appear in Proc. Int. Conf. on Scientific Computing, Moscow, September 15-16, 2005 (O. Pironneau et al.; eds).
[4] S. Repin, *A posteriori error estimation for nonlinear variational problems by duality theory*, Zapiski Nauchn. Semin. POMI,**243** (1997), 201–214.
[5] S. Repin, *A posteriori error estimation for variational problems with uniformly convex functionals*, Math. Comp., 69 (2000), 230, 481-500.
[6] A. Gaevskaya, R.H.W. Hoppe, and S. Repin. *A Posteriori Estimates for Cost Functionals of Optimal Control Problems*, to appear in Proc. ENUMATH 2005 Conf., Santiago de Compostela, July 18-22, 2005.

Efficient Methods for Aerodynamic Shape Optimization

Ilia Gherman
(joint work with Subhendu Hazra, Volker Schulz, Joël Brezillon and Nicolas Gauger)

We consider the general optimization problem

$$\begin{aligned} &\min \quad J(u,q) \\ &\text{subject to} \quad c(u,q)=0 \end{aligned} \tag{1}$$

where u is the state variable, q the design variable, and J the cost function. The constraint function c is such that $\frac{\partial c}{\partial u} =: C_u$ is invertible. Our main application field is geometric design of an aircraft. Hence, in our applications, c represents the flow-equations (e.g., Euler-flow). The cost function J is usually the drag, and the designs q are some parameterization of the geometry.

Our aim is to construct optimization methods based on already existing flow-solvers such that the overall optimization effort is equal to just several simulations. In [2, 3], a *one-shot* method was presented and efficiently applied to drag minimization of an RAE 2822 airfoil. The one-shot method is based on the ideas of reduced SQP methods. The overall optimization effort is of factor 4 compared to one simulation.

An important issue within the one-shot algorithm is the approximation of the reduced Hessian B. In [1,2], we investigate different choices of the reduced Hessian approximations. Our conclusion is that the approximation of the reduced Hessian should be consistent with the solution of the forward and the adjoint equations. In [3], we use two approximations of the reduced Hessian:

(B1) $B = \beta I$, where β is a constant;
(B2) $B = \beta_k I$, where β_k depends on the gradient information from the last iteration step.

In problem (1), there are no additional state constraints. In practical applications, drag minimization without additional constraints can lead to a solution

with very low lift- and pitching moment values. Such a solution is physically not meaningful. Hence, the aerodynamic constraints have to be stated explicitly. We now consider the problem

$$\begin{aligned} &\min && J(u,q) \\ &\text{subject to} && c(u,q) = 0 \\ &&& \ell(u,q) \geq 0, \end{aligned} \tag{2}$$

where $\ell(u,q) \geq 0$ represents the scalar constraint of maintaining constant lift [5]. The one-shot algorithm is extended to handle the additional constraint. Here, an additional adjoint equation has to be solved in order to compute reduced gradient with respect to lift. Similar to the "unconstrained" version, the one-shot method for the constrained problem is based on *partially* reduced SQP methods. A further extention is the use of L-BFGS updates with m stages for the approximation of the reduced Hessian. The choice of B as in (B2) above corresponds now to a L-BFGS update of the reduced Hessian with $m = 0$ stages.

Applying our method to the constrained drag minimization of an RAE 2822 airfoil, after design-update step we need to perform a *back projection* step for the design variables in order to fulfill the lift constraint [5]. This is due to the curvature in the constant-lift manifold.

The overall optimization effort for the constrained case is of factor 7 compared to one simulation. Very similar results are achieved imposing pitching moment constraint. Also, the "constrained" version of one-shot can be easily extended to handle both lift and pitching moment constraints [5].

The use of L-BFGS updates with $m > 0$ stages reduces the number of optimization iterations compared to other choices of the approxmations of the reduced Hessian.

The developed one-shot method was further applied to a 3D case [4]. There, the wing and the body of a Supersonic Commertial Transport aircraft were optimized with additional constraint of constant lift. The optimization effort is about 8 simulation runs.

Comparing the performance of the one-shot method in applications to existing black-box steepest descent methods we see a considerable reduction of the overall computational effort.

Acknowledgments. This research has been supported under the MEGADESIGN Project by the German Federal Ministry of Economics and Technology (BMWi).

References

[1] I. Gherman and V. Schulz, *Preconditioning of One-Shot Pseudo-Timestepping Methods for Shape Optimization*, PAMM Vol. **5**(1), 741–742, 2005.

[2] S. B. Hazra and V. Schulz, *Simultaneous Pseudo-timestepping for PDE-model based optimization problems*, Bit Numerical Mathematics, Vol. **44**(3), 457–472, 2004.

[3] S. B. Hazra, V. Schulz, J. Brezillon, and N. Gauger, *Aerodynamic shape optimization using simultaneous pseudo-timestepping*, Journal of Computational Physics, Vol. **204**(1), 46–64, 2005.

[4] S. B. Hazra, V. Schulz, and J. Brezillon, *Pseudo-Time Stepping for 3D Aerodynamic Shape Optimization*, Forschungsbericht Nr. 05 – 2, Universität Trier, Mathematik/Informatik, 2005.

[5] S. B. Hazra and V. Schulz, *Simultaneous Pseudo-Timestepping for Aerodynamic Shape Optimization Problems with State Constraints*, Journal of Numerical Mathematics (submitted).

Iterative Solvers for PDE Constrained Optimization

Philip E. Gill

(joint work with Randolph E. Bank, Anders Forsgren, and Joshua Griffin)

The "all-in-one" approach for PDE-constrained optimization involves the solution of a finite-dimensional optimization problem in the finite-difference, finite-element or finite-volume variables. This approach has the benefit that the PDE needs to be solved accurately only in the final stages of the optimization. In addition, the resulting optimization problem is *smooth*, with both first and second derivatives available at little cost. These benefits come at the price of an extremely large finite-dimensional optimization problem—possibly with millions of variables and constraints. It is vital, therefore, that the optimization method exploits the PDE structure by using parallel adaptive multilevel PDE techniques.

Here we focus on general nonconvex PDE-constrained problems. The discretized problem is written in the form

$$\min_{x\in\Re^n} f(x) \;\text{ subject to }\; c(x)=0, \;\; l\le x\le u,$$

where x denotes the vector of primal variables, c denotes the vector of nonlinear constraint functions. The Jacobian of $c(x)$ has the form $J = (\,J_1 \;\; J_2\,)$, where J_1 is an invertible sparse matrix with PDE structure. The Lagrange multipliers (i.e., dual variables) associated with the constraints $c(x) = 0$ and $l \le x \le u$ are denoted by λ and z respectively. A range of algorithms and software exist for *general-purpose* nonconvex optimization. In many cases, these algorithms are guaranteed to converge to a local minimizer of the problem with few assumptions on the problem. These general-purpose algorithms are based on a combination of sophisticated line-search, trust-region and filter techniques (see, e.g., [2, 3, 12]). The challenge is to develop methods that have the same reliability on the huge structured problems that arise from PDE discretizations.

If there are no inequality constraints we may solve a sequence of unconstrained minimization problems, each parameterized by an approximate multiplier vector λ_e and a scalar penalty parameter $1/\mu$. The objective function is the *primal-dual augmented Lagrangian*:

$$\mathcal{M}(x,\lambda) = f(x) - c(x)^T\lambda_e + \frac{1}{2\mu}\|c(x)\|_2^2 + \frac{1}{2\mu}\|c(x)+\mu(\lambda-\lambda_e)\|_2^2.$$

Unlike the conventional augmented Lagrangian, this function is minimized with respect to both x and λ. If μ is sufficiently small and λ_e is the optimal multiplier vector λ^*, then the primal-dual solution is an unconstrained minimizer of $\mathcal{M}(x,\lambda)$.

(We allow $\lambda_e \not\to \lambda^*$ for regularization purposes, but then require that $\mu \to 0$.) When inequality constraints are present we include the upper and lower bounds in the objective as a primal-dual barrier term. This gives an unconstrained objective function $\mathcal{M}(x, \lambda, z)$ that now involves x and the dual variables (λ, z) associated with both the equality and inequality constraints (see [4]).

It can be shown that if Newton's method is used to minimize $\mathcal{M}(x, \lambda, z)$, the Newton equations are equivalent to a KKT saddle-point system $Bx = b$ of the form

$$\begin{pmatrix} H & -A^T \\ A & D \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \tag{1}$$

where x_1, x_2 denote generic vectors of variables, H is the Lagrangian Hessian, and A and b depend on the derivatives of f and c. The matrix D is positive definite but *inherently ill-conditioned*, i.e., $D_{ii} = O(\mu)$ and $1/D_{jj} = O(\mu)$ for some i and j. The function $\mathcal{M}$ is not generally convex but it may be minimized by line-search and trust-region Newton methods based on finding an approximate solution of linear systems with the same structure as in (1).

The approximate solver must satisfy three requirements. First, it must treat the inherent ill-conditioning in D. Second, the solver must be able to detect if the underlying Newton system is not positive definite. This is required by methods that generate iterates converging to points satisfying the second-order necessary conditions for optimality. If the underlying Newton system is positive definite, then $H + AD^{-1}A^T$ is positive definite, and the matrix B of (1) is said to have *correct inertia*. If B does not have correct inertia, then the solution of (1) is of no interest and an alternative system is solved. The third requirement is that the solver generates iterates with increasing norm. This *monotonicity property* is a vital ingredient of many trust-search and trust-region methods based on Krylov-based iterative solvers (see, e.g., [13, 14]).

A method satisfying all three requirements may be defined by observing that (1) is equivalent to the *doubly-augmented system*:

$$\begin{pmatrix} H + 2A^TD^{-1}A & A^T \\ A & D \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 + 2A^TD^{-1}b_2 \\ b_2 \end{pmatrix}.$$

This system is positive definite if B has correct inertia and exhibits the monotonicity property when solved using a preconditioned conjugate-gradient method. The ill-conditioning is treated using *constraint preconditioning* (see, e.g., [9–11]). As preconditioner we use the matrix

$$P = \begin{pmatrix} M + 2A^TD^{-1}A & A^T \\ A & D \end{pmatrix}, \tag{2}$$

where M is some simple positive-definite matrix that approximates H. The equations $Pv = r$ used to apply the preconditioner are solved by exploiting the equivalence of the systems:

$$\begin{pmatrix} M + 2A^TD^{-1}A & A^T \\ A & D \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}, \tag{3}$$

$$\begin{pmatrix} M & -A^T \\ A & D \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} r_1 - 2A^TD^{-1}r_2 \\ r_2 \end{pmatrix}, \quad \text{and} \tag{4}$$

$$(M + A^TD^{-1}A)v_1 = r_1 - A^TD^{-1}r_2, \;\; v_2 = D^{-1}(r_2 - Av_1). \tag{5}$$

This equivalence allows us to compute the solution of (3) by solving either (4) or (5). In [5] it is shown that the preconditioner (2) provably removes the inherent ill-conditioning in B.

Unlike standard constraint preconditioners it is not necessary to solve the preconditioning equations $Pv = r$ exactly. In particular, a second-level preconditioned iterative method may be used to find approximate solutions of $Pv = r$. The requirements for the second-level solver are less restrictive because it is unnecessary to enforce monotonicity on the iterates or detect incorrect inertia of the preconditioner equations. For this reason, any conventional parallel multigrid and or multigraph preconditioner may be applied to the system $Pv = r$ (see, e.g., [1, 7–9]). The system $Pv = r$ is ill-conditioned, but since approximate solutions of the preconditioning equations need not satisfy the monotonicity property, the ill-conditioning may be eliminated using an appropriate diagonal similarity transformation (see, e.g., [6, p. 576]).

References

[1] G. Biros and O. Ghattas, Inexactness issues in the Lagrange-Newton-Krylov-Schur method for PDE-constrained optimization, In *Large-Scale PDE-Constrained Optimization (Santa Fe, NM, 2001)*, volume 30 of *Lect. Notes Comput. Sci. Eng.*, pages 93–114. Springer, Berlin, 2003.

[2] A. R. Conn, N. I. M. Gould, and Ph. L. Toint, *Trust-Region Methods*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000.

[3] R. Fletcher and S. Leyffer, Nonlinear programming without a penalty function, *Math. Program.*, 91(2, Ser. A):239–269, 2002.

[4] A. Forsgren and P. E. Gill, Primal-dual interior methods for nonconvex nonlinear programming, *SIAM J. Optim.*, 8:1132–1152, 1998.

[5] A. Forsgren, P. E. Gill, and J. D. Griffin, Iterative methods for systems arising in interior-point methods, Numerical Analysis Report NA 04-2, University of California, San Diego, 2004.

[6] A. Forsgren, P. E. Gill, and M. H. Wright, Interior methods for nonlinear optimization, *SIAM Rev.*, 44:525–597, 2002.

[7] O. Ghattas and J.-H. Bark, Large-scale SQP methods for optimization of Navier-Stokes flows, In *Large-Scale Optimization with Applications, Part II (Minneapolis, MN, 1995)*, pages 247–270. Springer, New York, 1997.

[8] O. Ghattas and C. E. Orozco, A parallel reduced Hessian SQP method for shape optimization, In *Multidisciplinary Design Optimization (Hampton, VA, 1995)*, pages 133–152. SIAM, Philadelphia, PA, 1997.

[9] C. Keller, N. I. M. Gould, and A. J. Wathen, Constraint preconditioning for indefinite linear systems, *SIAM J. Matrix Anal. Appl.*, 21(4):1300–1317 (electronic), 2000.

[10] D. Loghin and A. J. Wathen, Analysis of preconditioners for saddle-point problems, *SIAM J. Sci. Comput.*, 25(6):2029–2049 (electronic), 2004.
[11] M. F. Murphy, G. H. Golub, and A. J. Wathen, A note on preconditioning for indefinite linear systems, *SIAM J. Sci. Comput.*, 21(6):1969–1972 (electronic), 2000.
[12] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer-Verlag, New York, 1999.
[13] Ph. L. Toint, Towards an efficient sparsity exploiting Newton method for minimization, In I. S. Duff, editor, *Sparse Matrices and Their Uses*, pages 57–88, London and New York, 1981. Academic Press.
[14] T. Steihaug, The conjugate gradient method and trust regions in large scale optimization, *SIAM J. Numer. Anal.*, 20:626–637, 1983.

Nested Checkpointing in Optimal Control of Evolutions

Andreas Griewank
(joint work with Julia Sternberg)

We consider a discretized evolution with m-dimensional control over l time-steps on a state space of dimension n. Given an initial state we wish to minimize an objective defined as a function of the final state, by first or second order methods. The computation of total objective gradients with respect to the time-variant control vector requires a forward simulation of the evolution followed by a 'reverse', or 'adjoint' sweep. On nonlinear problems the reverse sweep is defined in terms of the forward trajectory, which is usually stored in a data structure of size $l \times n$. If we wish to apply a second order method of Newton type a third sweep becomes necessary and the storage requirement grows by the factor n to the order $l \times n^2$. To avoid these potentially prohibitive memory requirements we propose to replace the full storage/retrieval scheme by checkpointing strategies that repeatedly reexecutes the first forward and second reverse sweep in pieces. It is shown that the resulting operations count and storage requirement grow for the first order gradient and second order Newton method only by the factors $log(l)$ and $log^2(l)$, respectively. We report numerical results from an application to laser hardening of steel.

References

[1] A. Griewank, *Achieving logarithmic growth of temporal and spatial complexity in reverse automatic differentiation*, Optimization Methods and Software, **1(1)** (1992), 35–54.
[2] A. Griewank, *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*, Number 19 in Frontiers in Appl. Math. SIAM, Philadelphia, 2000.
[3] J. Sternberg, *Reduction of storage requirement by checkpointing for time-dependent optimal control problems*, PhD Thesis, Institute of Scientific Computing, TU Dresden, 2005.
[4] J. Sternberg and A. Griewank, *Reduction of storage requirement by checkpointing for time-dependent optimal control problems in ODEs*, Automatic Differentiation: Applications, Theory, and Implementations, editors M. Bücker and G. Corliss and P. Hovland and U. Naumann and B. Norris, Springer, (2006), 99–110.
[5] J. Sternberg and A. Griewank, *Nested memory-reduced procedure for solution of optimal control problem using Pantoja/Riccati method*, in preparation.

Hierarchical Matrices

Wolfgang Hackbusch

1. Introduction

1.1. **Dense Large-Scale Matrices.** Partial differential equations lead to large systems of equations. In particular, the boundary element method (cf. [31]) produces fully populated (dense) matrices. For large n, it is not possible to store all entries of a dense $n \times n$-matrix. Therefore, one needs data-sparse representations, which reduce the storage requirement to $\mathcal{O}(n \log^q n)$ for some $q \geq 0$ (and introduce an additional approximation error). The second requirement is that matrices represented in this format can easily multiplied by a vector. Such techniques are the panel clustering method (cf. [29]), which includes the multipole method as a spacial case, and compression techniques in the case of wavelets bases (cf. [12]).

The new feature of hierarchical matrices is the fact that they support *all* matrix operations, i.e., matrix addition, matrix multiplication, matrix inversion and the LU decomposition. This is also of interest for the sparse matrices from FEM, since the inverse matrix, which in general is completely dense, can be computed. The mentioned matrix operations are performed only approximately.

1.2. **Construction of Hierarchical Matrices.** The main idea in the construction of hierarchical matrices is the combination of two steps.

Step I: The matrix is partitioned into certain blocks defining the partition $\mathcal{P}$. For this purpose, for a given index set I "cluster tree" $T(I)$ and a block cluster tree are constructed which define the underlying hierarchical structure. A so-called admissibility condition is used to select the blocks of partition $\mathcal{P}$, typically contains small blocks along the diagonal and to large blocks far from the diagonal.

Step II: We fill all blocks in the partitioning $\mathcal{P}$ from above by rank-k matrices, i.e., the matrix block $M|_b = (M_{ij})_{i\in\tau, j\in\sigma}$ (for $b = \tau \times \sigma \in \mathcal{P}$) must satisfy $rank(M|_b) \leq k$. Each of these matrix blocks is represented by two matrices $A_b \in \mathbb{R}^{\#\tau\times k}$, $B_b \in \mathbb{R}^{\#\sigma\times k}$ so that

$$M|_b = A_b B_b^\top .$$

The storage costs of A_b, B_b are $(\#\tau + \#\sigma)\, k$ instead of $\#\tau * \#\sigma$ for the naive method. The resulting definition of hierarchical matrices of local rank k (more precisely $\leq k$) is

$$\mathcal{H}(k, \mathcal{P}) := \left\{ M \in \mathbb{R}^{I\times I} : rank(M|_b) \leq k \text{ for all } b \in \mathcal{P} \right\} .$$

A similar construction is possible for rectangular matrix from $\mathbb{R}^{I\times J}$.

Introductory papers on hierarchical matrices are [23], [8], [26], [14], [24]. Details of the construction are explained in [18].

1.3. **Accuracy and Costs.** The stiffness matrices from the boundary element method as well as the inverse of the finite element stiffness matrix (for elliptic boundary value problems) have the following property: Let A be the exact matrix. The singular values of $A|_b$ for $b \in \mathcal{P}$ (not of the whole matrix A) decay exponentially. This allows to approximate $A|_b$ up to an error ε with a rank-k matrix where $k = \mathcal{O}(\log^q \frac{1}{\varepsilon})$ where often q is the spatial dimension. Hence, A can be well approximated by some $A_{\mathcal{H}} \in \mathcal{H}(k, \mathcal{P})$.

In the BEM case, this follows easily from the smoothness of the fundamental solution. Concerning the inverse FEM stiffness matrix, this result is proved in [2] and holds even for nonsmooth coefficients.

Concerning the costs the following asymptotics hold for $k \ll n = \#I$.

- The *storage* of hierarchical matrices from $\mathcal{H}(k, \mathcal{P})$ requires $\mathcal{O}(nk \log n)$ units.
- The *matrix-vector multiplication* $(A \in \mathcal{H}(k, \mathcal{P}), x \in \mathbb{R}^I \mapsto Ax)$ is exact and requires $\mathcal{O}(nk \log n)$ arithmetical operations.
- The *matrix-matrix addition* $(A, B \in \mathcal{H}(k, \mathcal{P}) \mapsto C \approx A + B)$ computes the approximate sum in $\mathcal{O}(nk^2 \log n)$ operations.
- The *matrix-matrix multiplication* $(A, B \in \mathcal{H}(k, \mathcal{P}) \mapsto C \approx A * B)$ computes the approximate product in $\mathcal{O}(nk^2 \log^2 n)$ operations.
- The same cost estimate holds for the *matrix inversion* $(A \in \mathcal{H}(k, \mathcal{P}) \mapsto C \approx A^{-1})$.

The details and, e.g., characterisations of the constants involved in $\mathcal{O}(\cdot)$ are to be found in [18].

2. Application Fields

BEM matrices: The first goal in the boundary element method (BEM) is to generate a data-sparse approximation A' to the dense system matrix A in order to reduce the storage requirements. The error should be comparable with the already existing discretisation error. This can be achieved with hierarchical matrices of the local rank $k = \mathcal{O}(\log n)$. For the solution phase, one constructs a rougher approximation A'' by a further data compression step (cf. [17]) and approximates the LU factors which are used to build a fast iteration. The observation is that the solution phase is much faster than the generation phase (cf. [1]).

FEM preconditioning: Since any rough approximation B to A^{-1} is a good preconditioner, one can either compute the inverse of A in $\mathcal{H}(k, \mathcal{P})$ or the LU decomposition. The most effective LU approach is described in [21], [32].

Solution operator and Schur complements: The elliptic solution operator S (mapping the right-hand side and boundary data into the solution) is given by the inverse matrix, which is available in $\mathcal{H}(k, \mathcal{P})$. In optimisation calculations, often Schur complements of the form ASB appear, where e.g. B maps the control into the right-hand side and/or the boundary data and A extracts data from the solution. Such products are computable in $\mathcal{H}(k, \mathcal{P})$. Schur complement also appear during the solution of saddle point problems.

Matrix equations: The Lyapunov equation $AX + XA = C$ or the nonlinear Riccati equation $A^\top X + XB - XFX + G = 0$ appear in control problems. Traditional methods try to solve for the n^2 unknown entries of X, which in the best case costs $\mathcal{O}(n^2)$ operations. Representing X in $\mathcal{H}(k, \mathcal{P})$, one can solve these equations in almost linear time, provided that the coefficient matrix A arises from an elliptic operator (as in control problems with a state governed by an elliptic boundary value problem). Details can be found in [15], [19] and [20].

Matrix functions: For parabolic problems the matrix exponential function $\exp(-tA)$ is of interest, where A is the positive definite discretisation of an elliptic operator. The method proposed and analysed in [13] represents $\exp(-tA)$ by a Cauchy integral. Replacing the integral by a numerical quadrature, we are able to compute $\exp(-tA)$ with accuracy ε with a cost of order $\mathcal{O}(n \log^p \frac{1}{\varepsilon} \log^q n)$. Another interesting function is the sign-function $\operatorname{sign}(A)$ (see [20]).

Problems in high spatial dimensions: Related techniques be applied to problems in high spatial dimensions when Kronecker products of matrices can be used. An example is given in [16]: The discrete Laplace operator A in $[0,1]^d$ corresponding to $n = 1024$ nodal points in each direction and $d = 2048$ is an $N \times N$-matrix of size $N = 1024^{2048} \approx 1.2 \times 10^{6165}$. Nevertheless the inverse can be computed with high accuracy in 5 minutes (see [16] for older results).

$\mathcal{H}^2$-Matrices: Besides the hierarchy in the cluster tree, one can install a second hierarchy. Instead of the general rank-k matrices for the matrix blocks one requires the matrix blocks $M|_b$ ($b = \tau \times \sigma$) to belong to a fixed tensor space $V_\tau \times V_\sigma$ (i.e., $M|_b = A_b B_b^\top$ with $range(A_b) \subset V_\tau$ and $range(B_b) \subset V_\sigma$). In addition one needs a compatibility condition for V_τ and $V_{\tau'}$ when τ' is the son of τ in the cluster tree. The arising $\mathcal{H}^2$-matrices lead to lower cost in two aspects. First, the log-factor can be avoided, i.e. storage and cost are of order $O(n)$. Second, the constants in the $O(\ldots)$ estimate becomes smaller. The $\mathcal{H}^2$-technique has first been presented in [28] and then worked out by S. Börm [3], [4], [5], [6], [9], [25].

References

[1] M. Bebendorf, Hierarchical LU decomposition based preconditioners for BEM, *Computing 74* (2004), 225–247.

[2] M. Bebendorf and W. Hackbusch, Existence of $\mathcal{H}$-matrix approximants to the inverse FE-matrix of elliptic operators with L^∞-coefficients *Numer. Math. 95* (2003), 1-28.

[3] S. Börm, $\mathcal{H}^2$-matrices - multilevel methods for the approximation of integral operators, *Comput. Vis. Sci. 7* (2004), 173–181.

[4] S. Börm, Approximation of integral operators by $\mathcal{H}^2$-matrices with adaptive bases, *Computing 74* (2005), 249–271.

[5] S. Börm, $\mathcal{H}^2$-matrix arithmetics in linear complexity, Computing (to appear in 2006).

[6] S. Börm, Data-sparse approximation of non-local operators by $\mathcal{H}^2$-matrices, Preprint 44, Max-Planck-Institut Mathematik in den Naturwissenschaften, Leipzig, 2005.

[7] S. Börm and L. Grasedyck, Hybrid cross approximation of integral operators, *Numer. Math. 101*, (2005), 251–271.

[8] S. Börm, L. Grasedyck, and W. Hackbusch, Introduction to hierarchical matrices with applications, *Eng. Anal. Bound. Elem. 27* (2003), 405-422.

[9] S. Börm and W. Hackbusch, Data-sparse approximation by adaptive $\mathcal{H}^2$-matrices, *Computing 69* (2002), 1–35.

[10] S. Börm and J. Ostrowski, Fast evaluation of boundary integral operators arising from an eddy current problem, *J. Comput. Phys. 193* (2003), 67–85.
[11] S. Börm and S.A. Sauter, BEM with linear complexity for the classical boundary integral operators, *Math. Comp. 74* (2005), 1139–1177.
[12] W. Dahmen, S. Prössdorf, and R. Schneider, Wavelet approximation methods for pseudodifferential equations II: Matrix compression and fast solution, *Adv. Comput. Math. 1* (1993), 259-335.
[13] I.P. Gavrilyuk, W. Hackbusch, and B.N. Khoromskij, $\mathcal{H}$-Matrix approximation for the operator exponential with applications, *Numer. Math. 92* (2002), 83-111.
[14] L. Grasedyck, *Theorie und Anwendungen Hierarchischer Matrizen*, Dissertation, Universität zu Kiel, 2001.
[15] L. Grasedyck, Existence of a low rank or $\mathcal{H}$-matrix approximant to the solution of a Sylvester equation, *Numer. Linear Algebra Appl. 11* (2004), 371–389.
[16] L. Grasedyck, Existence and computation of a low Kronecker-rank approximant to the solution of a tensor system with tensor right-hand side, *Computing, 72* (2004), 247–266.
[17] L. Grasedyck, Adaptive recompression of $\mathcal{H}$-matrices for BEM, Computing 74:205–223, 2005.
[18] L. Grasedyck and W. Hackbusch, Construction and arithmetics of $\mathcal{H}$-matrices, *Computing 70* (2003), 295-334.
[19] L. Grasedyck and W. Hackbusch. A multigrid method to solve large scale Sylvester equations, Report 48, Max-Planck-Institut für Mathematik, Leipzig, 2004.
[20] L. Grasedyck, W. Hackbusch, and B.N. Khoromskij, Solution of large scale algebraic matrix Riccati equations by use of hierarchical matrices, *Computing 70* (2003), 121-165.
[21] L. Grasedyck, R. Kriemann, and S. Le Borne, Parallel black box domain decomposition based $\mathcal{H}$-LU preconditioning, Preprint 115, Max-Planck-Institut Mathematik in den Naturwissenschaften, Leipzig, 2005.
[22] L. Grasedyck and S. Le Borne, $\mathcal{H}$-matrix preconditioners in convection-dominated problems, *SIAM J. Matr. Anal. Appl.* (to appear).
[23] W. Hackbusch, A sparse matrix arithmetic based on $\mathcal{H}$-matrices. Part I: Introduction to $\mathcal{H}$-matrices, *Computing 62* (1999), 89-108.
[24] W. Hackbusch, *Panel clustering techniques and hierarchical matrices for BEM and FEM.* Chapter 20 in E. Stein, R. de Borst, and T.J.R. Hughes, editors, Encyclopedia of computational mechanics, Wiley, Chichester, 2004.
[25] W. Hackbusch and S. Börm: $\mathcal{H}^2$-Matrix approximation of integral operators by interpolation, *Appl. Numer. Math. 43* (2002), 129-143.
[26] W. Hackbusch, L. Grasedyck, and S. Börm, An introduction to hierarchical matrices, *Math. Bohem. 127* (2002), 229–241.
[27] W. Hackbusch and B.N. Khoromskij, A sparse $\mathcal{H}$-matrix arithmetic. Part II: Application to multi-dimensional problems, *Computing 64* (2000), 21-47.
[28] W. Hackbusch, B.N. Khoromskij, and S.A. Sauter, On $\mathcal{H}^2$-matrices. In H.-J. Bungartz, R.H.W. Hoppe, and C. Zenger, editors, *Lectures on Applied Mathematics*, pages 9–29. Springer-Verlag, Berlin, 2000.
[29] W. Hackbusch and Z.P. Nowak, On the fast matrix multiplication in the boundary element method by panel clustering, *Numer. Math. 54* (1989), 463–491.
[30] R. Kriemann, Parallel $\mathcal{H}$-matrix arithmetics on shared memory systems, *Computing 74* (2005), 273-297.
[31] S. Sauter and C. Schwab, *Randelementmethoden*, Teubner, Stuttgart, 2004.
[32] S. Le Borne, L. Grasedyck, and R. Kriemann, *Domain-decomposition based $\mathcal{H}$-LU preconditioners*, 16th International Conference of Domain Decomposition Methods, ddm.org, 2005.
[33] C.H. Wolters, L. Grasedyck, and W. Hackbusch, Efficient computation of lead field bases and influence matrix for the FEM-based EEG and MEG inverse problem - part I: complexity considerations, *Inverse Problems 20* (2004), 1099–1116.

Feasible and Non-Interior Path-following in constrained minimization with low multiplier regularity

Michael Hintermüller

The efficient numerical solution of inequality constrained minimization problems in function space with low Lagrange multiplier regularity is still a significant challenge. Prototype problems include optimal control problems with pointwise state constraints, boundary control problems where the control has to satisfy pointwise constraints on the boundary, or classes of control problems with "solutions très faibles" (in the sense of Lions). The common feature is low regularity of the Lagrange multipliers associated with the pointwise constraints. This has an immediate effect when characterizing optimality of solutions. Indeed, in the presence of inequality constraints first order optimality conditions typically involve a so-called *complementarity system* (see, e.g., [4]) which in turn is influenced by the regularity of the Lagrange multiplier associated with the inequality constraint. To be specific, let us assume that $x_1 \in X_1$ has to satisfy the pointwise (almost everywhere) constraint

$$x_1 \leq \psi, \tag{1}$$

where X_1 denotes a Hilbert space continuously embedded into $L^2(\omega)$ and $\omega \subset \mathbb{R}^m$ is a bounded domain. Further $\psi \in L^2(\omega)$, and $\leq$ represents the natural ordering in $L^2(\omega)$. Let $\lambda \in X_1^*$ denote the Lagrange multiplier associated with the constraint (1), where X_1^* is the topological dual space of X_1. First order optimality characterizations include the complementarity condition

$$x_1 \leq \psi, \quad \langle \lambda, x_1 - \xi \rangle_{X_1^*, X_1} \geq 0 \quad \text{for all } \xi \leq \psi,\ \xi \in X_1, \tag{2}$$

at an optimal solution x_1 with associated Lagrange multiplier λ. Without additional regularity, system (2) does not admit a pointwise interpretation, which is frequently crucial for numerical algorithms.

In fact, solution techniques and their (local) convergence behavior often hinge on the multiplier regularity. Classical active set methods, for instance, require a pointwise (almost everywhere) interpretation of λ for the active set estimation. In the case of pointwise constraints, techniques like the projected gradient methods will not work without modification since the sum of the iteration variable and the gradient of the objective, which coincides with the negative multiplier, is needed for the update. Since they have different regularity properties this is not feasible in general. An analogous comment applies for projected Newton techniques. Recently it was found that semismooth Newton methods are highly efficient in solving certain classes of constrained optimization problems in function space [1–5]. These methods rely on a pointwise almost everywhere interpretation of the complementarity system (2) and smoothing properties of the control-to-adjoint-state mapping. In fact, the pointwise interpretation allows to express (2) equivalently as

$$\lambda - \max(0, \lambda + c(x_1 - \psi)) = 0, \tag{3}$$

for some arbitrarily fixed $c > 0$, and the smoothing of the control-to-adjoint-state operator typically implies that the mapping

$$\theta : x_1 \mapsto \lambda(x_1) + c(x_1 - \psi)$$

can be considered as $\theta : X_1 \to L^q(\omega)$ with $q > 2$. The norm gap between $L^q(\omega)$ and the space $L^2(\omega)$, in which the inequality (1) is posed, is crucial in proving generalized differentiability of

$$x_1 \mapsto \max(0, \theta(x_1));$$

and in arguing well-definedness and locally superlinear convergence of the generalized (semismooth) Newton method for solving the underlying nonsmooth first order optimality system; see [4] for details. Again, the low multiplier regularity may prevent the pointwise interpretation and/or the smoothing of the control-to-adjoint-state mapping.

We consider the following problem class: Let X_1, X_2 and W be real Hilbert spaces with $X_1 \hookrightarrow L^2(\omega) \hookrightarrow X_1^*$, where X_1^* denotes the dual of X_1 and ω a bounded domain in $\mathbb{R}^m$. Further set $X = X_1 \times X_2$ and let $x = (x_1, x_2)$ denote a generic element in X. Let $E \in \mathcal{L}(X, W)$, $f \in W$, and $\psi \in X_1$. Further let $J : X \to \mathbb{R}$ denote a quadratic functional satisfying

$$\langle J'(x) - J'(y), x - y\rangle_{X^*,X} \geq \alpha |x - y|_X^2 \tag{4}$$

whenever $E(x - y) = 0$ for $x, y \in X$, for some $\alpha > 0$. Here $\langle \cdot, \cdot \rangle_{X^*,X}$, at times denoted by $\langle \cdot, \cdot \rangle$, stands for the duality pairing between X and X^*. We set

$$C := \begin{pmatrix} E \\ (I, 0) \end{pmatrix} : X \to W \times X_1$$

and assume that C is surjective. The problem under consideration is

$$\text{(P)} \qquad \begin{aligned} &\text{minimize } J(x) \text{ over } x \in X \\ &\text{subject to } Ex = f, \\ &\qquad\qquad x_1 \leq \psi, \end{aligned}$$

where $\leq$ denotes the ordering in $L^2(\omega)$. Note that (P) covers state constrained optimal control problems for which it is known that the Lagrange multiplier associated with the pointwise inequality constraint is a Borel measure only. In order to have a regular approximation of this Lagrange multiplier we consider the regularized problems

$$(P_\gamma) \qquad \begin{aligned} &\text{minimize } J(x, \gamma) := J(x) + \frac{1}{2\gamma} \int_\omega |\left(\bar\lambda + \gamma(x_1 - \psi)\right)^+|^2 dw \text{ over } x \in X \\ &\text{subject to } Ex = f, \end{aligned}$$

where $\gamma > 0$ represents a relaxation (or regularization) parameter and $\bar\lambda \in L^2(\omega)$ is an optional shift-parameter. Further, $(\cdot)^+ = \max(\cdot, 0)$. Let x_γ denote its solution with $p_\gamma \in W^*$ the corresponding adjoint state and

$$\lambda_\gamma = (\bar\lambda + \gamma(x_{1,\gamma} - \psi))^+ \in L^2(\omega).$$

The regularization parameter γ induces a *primal-dual path*

$$\mathcal{C} = \{(x_\gamma, p_\gamma, \lambda_\gamma) \in X \times W^* \times X_1^* : \gamma \in (0, \infty)\}.$$

We show that the path is bounded, that it converges to a solution of the original problem, and that it is locally Lipschitz continuous. Further, under a strict complementarity assumption, it is (strongly) differentiable w.r.t. γ.

We also consider the *primal-dual path-value functional*

$$\gamma \mapsto V(\gamma) = J(x_\gamma) + \frac{1}{2\gamma} \int_\omega |(\bar\lambda + \gamma(x_{1,\gamma} - \psi))^+|^2 dw$$

defined on $(0, \infty)$. It is differentiable with the first derivative satisfying

$$\dot V(\gamma) = -\frac{1}{2\gamma^2} \int_\omega |(\bar\lambda + \gamma(x_{1,\gamma} - \psi))^+|^2 + \frac{1}{\gamma} \int_\omega (\bar\lambda + \gamma(x_{1,\gamma} - \psi))^+ (x_{1,\gamma} - \psi).$$

Further it can be shown, again under a strict complementarity assumption, that it is twice differentiable w.r.t. γ. For $\bar\lambda = 0$ it follows that for $\gamma \geq 0$

$$\dot V(\gamma) > 0 \quad \text{and} \quad \ddot V(\gamma) < 0$$

whenever the unconstrained solution does not solve (P). For certain problem classes (like obstacle problems) it can be shown that there exists $\bar\lambda \geq 0$ such that x_γ is feasible for (P). In this case we infer

$$\dot V(\gamma) < 0 \quad \text{and} \quad \ddot V(\gamma) > 0$$

whenever the unconstrained solution does not solve (P). These properties are utilized for defining model functions of $V(\gamma)$ which are then the basis for an efficient update strategy of γ in an algorithmic framework. Further, for every $\gamma > 0$ the corresponding regularized problem is efficiently solved by a class of semismooth Newton methods, or equivalently primal-dual active set strategies, which converge locally superlinearly in function space.

In our numerical practice the overall solution algorithm exhibits a mesh-independent behavior and compares favorably with primal-dual path-following interior point methods.

For proof details and a report on extensive numerical tests including a comparison with interior point methods and the blending with a nested iteration concept for speed-up we refer to [6].

References

[1] X. Chen, Z. Nashed, and L. Qi. Smoothing methods and semismooth methods for nondifferentiable operator equations. *SIAM J. Numer. Anal.*, 38(4):1200–1216 (electronic), 2000.

[2] J.-C. de los Reyes and K. Kunisch. A semi-smooth Newton method for control constrained optimal control of the Navier Stokes equations. *Nonlinear Analysis*. to appear.

[3] M. Hintermüller and M. Hinze. A SQP-semismooth Newton-type algorithm applied to control of the instationary Navier-Stokes system subject to control constraints. Technical Report TR03-11, Department of Comput. and Appl. Math., Rice University, Houston, Texas, USA, 2003. submitted to SIAM J. Optimization.

[4] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semismooth Newton method. *SIAM J. Optim.*, 13(3):865–888, 2003.

[5] M. Hintermüller, V. A. Kovtunenko, and K. Kunisch. Generalized Newton methods for crack problems with nonpenetration condition. *Numer. Methods Partial Differential Equations*, 21(3):586–610, 2005.
[6] M. Hintermüller and K. Kunisch. Feasible and non-interior path-following in constrained minimization with low multiplier regularity. Report 01-05, Department of Mathematics and Scientific Computing, University of Graz, Heinrichstraße 36, A-8010 Graz, Austria, October 2005.

Convergence of a Finite Element Approximation to a State Constrained Elliptic Control Problem

Michael Hinze
(joint work with Klaus Deckelnick)

In this talk we discuss a finite element discretization of a control problem with pointwise state constraints. Let $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$ be a bounded, convex domain with a smooth boundary. For a given function $u \in L^2(\Omega)$ we denote by $y = \mathcal{G}(u)$ the solution of the Neumann problem

$$\begin{aligned} -\Delta y + y &= u \quad \text{in } \Omega, \\ \partial_\nu y &= 0 \quad \text{on } \partial\Omega. \end{aligned}$$

Here ν denotes the outward pointing unit normal to $\partial\Omega$. It is well known that $y \in H^2(\Omega)$ and

$$\|y\|_{H^2} \leq C\|u\|_{L^2}. \tag{1}$$

We now consider the following control problem

$$\begin{aligned} &\min_{u \in L^2(\Omega)} J(u) = \frac{1}{2}\int_\Omega |y - y_0|^2 + \frac{\alpha}{2}\int_\Omega |u - u_0|^2 \\ &\text{subject to } y = \mathcal{G}(u) \text{ and } y(x) \leq b(x) \text{ in } \Omega. \end{aligned} \tag{2}$$

Here, $\alpha > 0$ and $y_0, u_0 \in H^1(\Omega)$ as well as $b \in W^{2,\infty}(\Omega)$ are given functions. We denote by $\mathcal{M}(\bar{\Omega})$ the space of Radon measures which is defined as the dual space of $C^0(\bar{\Omega})$ and endowed with the norm

$$\|\mu\|_{\mathcal{M}(\bar{\Omega})} = \sup_{f \in C^0(\bar{\Omega}), |f| \leq 1} \int_{\bar{\Omega}} f d\mu.$$

The analysis of (2) is well understood for the problem under consideration. Since the state constraints form a convex set and the cost functional is quadratic it is not difficult to establish the existence of a unique solution $u \in L^2(\Omega)$ to this problem. Moreover,

Theorem 1. *A function $u \in L^2(\Omega)$ is a solution of (2) if and only if there exist $\mu \in \mathcal{M}(\bar{\Omega})$ and $p \in L^2(\Omega)$ such that with $y = \mathcal{G}(u)$ there holds*

$$(3) \int_\Omega p(-\Delta v + v) = \int_\Omega (y - y_0)v + \int_{\bar{\Omega}} v d\mu \qquad \forall v \in H^2(\Omega) \text{ with } \partial_\nu v = 0 \text{ on } \partial\Omega$$

$$(4) \qquad p + \alpha(u - u_0) = 0 \qquad \text{a.e. in } \Omega$$

$$(5) \qquad \mu \geq 0,\ y(x) \leq b(x) \text{ a.e. in } \Omega, \text{ and} \qquad \int_{\bar{\Omega}} (b - y) d\mu = 0.$$

The study of (2) is complicated by the presence of the measure μ on the right hand side of (3). As a consequence, the solution p of this problem is no longer in $H^1(\Omega)$ but only in $W^{1,s}(\Omega)$ for all $1 \leq s < \frac{d}{d-1}$. This fact also accounts for the form of the weak formulation (3).

In this talk we develop a finite element approximation of problem (2). The underlying idea consists in approximating the cost functional J by a sequence of functionals J_h where h is a mesh parameter related to a sequence of triangulations. The definition of J_h involves the approximation of the state equation by linear finite elements and enforces constraints on the state in the nodes of the triangulation.

We shall prove that the minima u_h of J_h converge in L^2 to the minimum u of J as $h \to 0$ and that the states y_h convergence strongly in H^1 as well as uniformly to the corresponding optimal state y. Furthermore, we shall prove that the associated discrete adjoint states p_h converge strongly in L^2 to p, and that the discrete multipliers μ_h converge weakly in $\mathcal{M}(\bar{\Omega})$ to μ. Finally we carry out an error analysis in the two– and three–dimensional case and obtain

$$\|u - u_h\|_{L^2},\ \|y - y_h\|_{H^1} = O(h^{2-\frac{d}{2}-\epsilon})\ (h \to 0)\ (\epsilon > 0 \text{ arbitrary}).$$

The complete analysis is presented in [1].

References

[1] K. Deckelnick and M. Hinze, *Convergence of a finite element approximation to a state constrained elliptic control problem*, Preprint MATH–NM–01–2006, Institut fr Numerische Mathematik, TU Dresden (2006).

Modelling and Identification of Hysteresis in PDEs

Barbara Kaltenbacher

Memory effects such as rate independent hysteresis appear in a number of physical phenomena modelled by PDEs. We here focus on the classical example of magnetics as well as a model problem in piezoelectricity, the first one leading to a parabolic PDE,

$$(1) \qquad \mu_0 \vec{u}_t + \mathbf{P}[\vec{u}]_t + \nabla \times \ \nabla \times \vec{u} = f\,,$$

the second to a hyperbolic one

$$(2) \qquad \rho u_{tt} = \mathbf{P}[u_x]_x\,.$$

In both cases, hysteresis can be modelled in a very general way by Preisach operators $\mathbf{P}$ appearing in place of coefficients within the respective differential operator. An important task in material characterization is the identification of the operator $\mathbf{P}$ in (1) or (2), respectively, from overposed boundary data.

Hysteresis Operators. The most simple example of a hysteresis operator is a hysteretic relay with up- and down- switching thresholds α and β, respectively, given by

$$\mathbf{R}_{\beta,\alpha}[v](t) = \begin{cases} +1 & \text{if } v(t) > \alpha \text{ or } (w(t_i) = +1 \wedge v(t) > \beta \\ -1 & \text{if } v(t) < \beta \text{ or } (w(t_i) = -1 \wedge v(t) < \alpha \end{cases} \qquad t \in [t_i, t_{i+1}]\,,$$

where $t_0, t_1, \ldots, t_N$ a monotonicity partition of the time interval, i.e., chosen such that v is monotone on $[t_i, t_{i+1}]$, $i = 0, \ldots, N-1$. Another example is the mechanical play, described by the recursion

$$\mathbf{F}_r[v](t) = w(t) = \max\{v(t) - r, \min\{v(t) + r, w(t_i)\}\} \quad t \in [t_i, t_{i+1}]\,. \tag{3}$$

These two simple examples serve as building blocks in a very general hysteresis model, the Preisach operator, that is a weighted superposition of elementary relays:

$$\begin{aligned} \mathbf{P}[v](t) &= \iint_{\alpha,\beta\in S} \wp(\beta,\alpha)\mathbf{R}_{\beta,\alpha}[v](t)\, d(\alpha,\beta) \\ &= \iint_{\alpha,\beta\in S^+(t)} \wp(\beta,\alpha)\, d(\alpha,\beta) - \iint_{\alpha,\beta\in S^-(t)} \wp(\beta,\alpha)\, d(\alpha,\beta) \\ &= w_0 + 2\int_0^\infty \int_0^{\mathbf{F}_r[v](t)} \wp(s-r, s+r)\, ds\, dr \end{aligned} \tag{4}$$

where the latter identity holds due to the fact that the interface between the sets $S^+(t)$, $S^-(t)$ of up and down switched relays moves according to the rules of the mechanical play. It is essential for the efficient numerical simulation of hysteretic processes (especially in the context of PDEs, where each space point has its own memory), that Preisach operators forget certain passages in the past according to deletion rules (monotone deletion, Madelung rule, wipe out). Therewith, to compute the output at a some time instance $\bar{t}$, not the whole input $(v(t))_{t\in[0,\bar{t}]}$ is required, but only a (short) string of values $v(t_i)_{i\in\{1,\ldots,N\}}$ that have "survived" deletion. The actual computation of the output can then efficiently be done via a finite sum

$$\mathbf{P}[v](t) = \iint_{\alpha,\beta\in S} \wp(\beta,\alpha)\mathbf{R}_{\beta,\alpha}[v](t)\, d(\alpha,\beta) = \mathbf{P}[v_0] + \sum_{i=1}^{N} \mathbf{E}(v(t_i), v(t_{i+1}))$$

with a precomputed Everett (or shape) function $\mathbf{E}$

$$\mathbf{E}(v_*, v^*) = 2\iint_{v_*\le\beta\le\alpha\le v^*} \wp(\beta,\alpha)\, d(\alpha,\beta) \text{ if } v_* < v^* \quad \mathbf{E}(-v^*, -v_*) = \mathbf{E}(v_*, v^*)\,.$$

Under certain conditions on the weight function $\wp$ or the shape function $\mathbf{E}$, the operator $\mathbf{P}$ is Lipschitz continuous from $C(0,T)$ into itself, and has some coercivity and convexity properties. However, $\mathbf{P}$ is not differentiable, even for smooth $\wp$ and

E, and **P** is not a monotone operator. Therefore, well-posedness proofs for (1), (2) use techniques different from the usual arguments for nonlinear evolution equations with monotone operators (cf. e.g. Zeidler 1998), see [1], [3], and the references therein, especially Visintin 1994, Krejčí, 1996.

Hysteresis Identification. To discuss some numerical methods for the reconstruction of $\wp$ in PDEs from additional boundary measurements, we concentrate on a hyperbolic model problem

$$(*)\begin{cases} \text{PDE}: & \rho u_{tt} - (\mathbf{P}[u_x])_x = 0 \quad x \in [0,L],\ t \in [0,T] \\ \text{boundary conditions}: & u(0,t) = f_0(t) \qquad u(L,t) = f_L(t) \\ \text{initial conditions}: & u(x,0) = u_0(x) \qquad u_t(x,0) = u_1(x) \end{cases}$$

$$(**)\ \text{measurements}: \quad (\mathbf{P}[u_x])(L,t) = g(t)$$

Since **P** appears not only in the PDE but also in the measurement boundary condition, a most straightforward approach for recovering **P** is the following *alternating iteration*: In each step

- solve PDE (*) for u
- solve measurements (**) for **P**.

This amounts to a fixed point method whose convergence at least theoretical imposes restrictions on the magnitude of the excitation. To avoid this, we apply *Newton type methods* to (**) considered as an operator equation $F(\mathbf{P}) = g$, where the forward operator F contains solution of (*). In doing so, we have to take into account the fact that **P** is not differentiable in a classical sense. Hence, we propose to choose one of the following strategies:

i) Approximate **P** by a superposition operator **S**
ii) Semismooth Newton: Find slanting function **G** for **P**
iii) Broyden's method

ad i) The main loop of a hysteresis operator can be approximated by its centerline or alternatively by the so-called the commutation curve (i.e., the virgin curve and its prolongation into the lower half plane). Either of these two curves can be explicitely expressed via the shape function **E** and used to define a superposition operator **S**. Replacing $\mathbf{P}'$ (which does not exist in a classical sense) in the definition of a Newton step by $\mathbf{S}'$, we arrive at an iterative identification method for **P**.

ad ii) The slope approximation derived from the centerline or commutation curve may be poor in minor loops, that are usually much flatter than the main loop. Moreover, superposition operator approximations of the Preisach operator neglect its memory. To avoid these two drawbacks, we also study a generalized differentiability concept, namely slant differentiability (cf., e.g., Chen, Nashed and Qi 1997, Hintermüller, Ito, and Kunisch 2002, M. Ulbrich 2003). Using the representation of the Preisach operator via the mechanical play (4), the recursion (3), as well as the slant derivative of the max and the min function together with

the chain rule, we arrive at a slanting function $\mathbf{G}$ (i.e., "$\mathbf{P}'$") as follows:

$$(\mathbf{G}[v]h)(t_m) = \sum_{i=0}^{m} h(t_i) \int_0^\infty \wp(w_m^r - r, w_m^r + r) \mathbf{I}_{R^i[v]}(r)\, dr$$

where $w_i^r = \mathbf{F}_r[v](t_i)$ and $R^i[v] = \{r \in \mathbb{R}^+ \mid w_i^r \neq w_{i-1}^r \wedge w_i^r = w_{i+1}^r = \cdots = w_j^r\}$. for v piecewise linear with breakpoints at t_m, $m = 1, \ldots, M$. However, the efficient implementation of this generalized derivative in a semismooth Newton method for hysteresis identification is still an open problem, due to the arising integral as well as the computationally costly case distinctions in the definition of the sets R^i.

ad iii) Broyden's method, that approximates $F'(\mathbf{P})$ according to a rank one update formula, has the advantage of only involving values of F. For well-posed problems, convergence of Broyden's method even for operators that are only Lipschitz continuous has been established by Griewank, 1987. In the here relevant ill-posed situation, these results can not directly be carried over; Still a regularizing effect due to the finite rank update procedure can be observed and made rigorous under certain conditions (cf. BK 1998).

In a series of numerical tests we investigated convergence of the above sketched methods both with exact and with noisy data, for different types of hysteresis operators (with narrow and with broad main loop, respectively, as well as incorporating saturation), cf. [2].

References

[1] M. Brokate, J. Sprekels, *Hysteresis and Phase Transitions*, Springer, New York, 1996.
[2] B. Kaltenbacher, M. Kaltenbacher, *Modelling and iterative identification of hysteresis via Preisach operators in PDEs*, LSE-report 2005/04, University of Erlangen, 2005.
[3] B. Kaltenbacher, in preparation.

Model Reduction for Nonlinear Least Squares

C. Tim Kelley
(joint work with Dan Sorensen, Jill Reese, and Corey Winton)

1. Introduction

We propose a model reduction method for moderately sized nonlinear least squares problems for which the evaluation of the residual requires the solution of a partial differential equation on a large grid. Using POD [1] as a model, we manufacture snapshots of the solution of the PDE and use the singular value decomposition to build a low-dimensional Galerkin basis to reduce the dimension of the PDE solve.

At this stage the research is speculative, and we offer only simple one dimensional computational studies to support our assertions.

Suppose we have a vector $p \in R^N$ of N design parameters, and seek to fit M data points to minimize

$$f(p) = \frac{1}{2} R(p)^T R(p) = \frac{1}{2} \sum_{i=1}^{M} (S_i(p) - d_i)^2$$

where $d \in R^M$ are the data, and $S : R^N \to R^M$ is the model.

The target problems will have large residuals, and the least squares problems themselves are very poorly scaled. We can remedy the scaling problems with a logarithmic change of variables.

One example is for $S_i(p) = u(p)(x_i)$, where u is the solution of a PDE, and x_i is a spatial data point. In any case, we will assume that in order to evaluate S you have to solve a *linear, for now,* discretized PDE

$$L(p)w = g \tag{1}$$

on a spatial mesh with M_x points. We assume that

$$S(p) = S(w(p)) \tag{2}$$

depends on p through w.

We follow the POD paradigm, and see to generate a sequence of solutions $\{w_k\}_{k=1}^K$ for various values of p. We let W be the $M_x \times K$ matrix with the u_k's as its columns. We let

$$U \Sigma V^T = W$$

be the SVD of W, and take the first M_S columns of U to be our reduced basis.

The next step is to change the way the least squares residual is computed. We define a coarse mesh model S^c as follows. We begin by solving the reduced PDE

$$U^T L(p) U w^c = U^T p, \tag{3}$$

which we will assume is a well-posed problem. Notice the notation, $w^c \in R^{M_K}$ is the vector of coefficients of the solution $Uw^c \in R^{M_x}$ of the reduced PDE. We've reduced the number of unknowns in the PDE from M_x to M_k in this way.

Then we simply define

$$S^c(p) = S(Uw^c(p)), \tag{4}$$

and

$$R^c(p) = S^c(p) - d. \tag{5}$$

The reduced problem is to minimize

$$\frac{1}{2} \sum_{i=1}^{M} (S_i^c(p) - d_i)^2.$$

2. Example

These problems are badly scaled. We remedy this with a logarithmic change of variables. We solve the optimization problem with a Gauss-Newton pseudo-transient continuation scheme [2].

We consider the boundary value problem

$$-(k(p)w')' = g,\ w(0) = w_0, w(1) = w_1, \tag{6}$$

where $k(p)$ is a piecewise constant function taking values $\{p_j\}_{j=1}^N$. We assume that

$$k(p) \in [10^{-6}, 10^{-2}],$$

and $w_0, w_1 \in [10^5, 10^7]$. We'll use

$$w_0 = 10^5 \text{ and } w_1 = 10^6$$

in this example.

The exact data are

$$k^e(p) = p_{i+1}^e \text{ for } x_i^e < x < x_{i+1}^e, 0 \le i \le 4,$$

where the exact interface points are $x_i^e = i/4$. The exact conductivities are

$$p^e = (10^{-4}, 10^{-3}, 10^{-2}, 10^{-5})^T.$$

I'll let $u^e \in R^{M_x}$ be the solution of (6). We will generate the data from u^e and perturbations of u^e.

We will try to approximate p for fixed choices of the interfaces.

In all cases $m_x = 2^{10} - 1$. We discretize with standard finite elements with $m_x = 2^{10} - 1$ interior nodes.

We generate the data by solving

$$A(p^{true})u^{true} = b(p^{true})$$

and then letting the data be

$$d_j = u_{j_k}^{true}$$

for $1 \le j \le M$, where the sequence $\{z_k\} = \{x_{j_k}\}$ is a subset of the grid.

To make sure that p stays within sensible limits, we imposed simple bound constraints

$$10^{-6} \le p_i \le 10^{-2},$$

so $-6 \le q_i \le -2$.

To generate the reduced basis we take a few Euler steps for $p' = -R'(p)^T R(p)$ to build the snapshots.

The algorithm we used in the example was

Begin with an initial iterate p
for i = 1,..., isweep **do**
  Build the reduced basis.
  Take lots of iterations on f^c (reduced model).
  Take 3 iterations on f (full model).
end for
Take 10 iterations on f.

We consider one test case. We perturb the data by 1% with a random vector, and change the interfaces. The interface for the unknown conductivity is now $x_i^e = i/3$. So the number and locations of the interfaces have changed. In Table 1 we tabulate the errors in $\log(p)$ and the residuals as functions of the number of $isweep$, the number of times we rebuild the reduced basis. $isweep = 0$ is the result of using the full model. The table illustrates that $isweep = 1$ is a reasonable choice.

TABLE 1. Case 3 results

isweep	0	1	2	3	4
f	3.16e-02	3.16e-02	3.16e-02	3.16e-02	1.11e-01
$E(\log(p))$	0.00e+00	1.09e-07	1.16e-07	1.29e-07	3.92e-01

REFERENCES

[1] K. Karhunen, *Zur Spektraltheorie stochastischer Prozesse*, Ann. Acad. Sci. Fennacie, 37 (1946).
[2] C. T. Kelley, L. Qi, L.-Z. Liao, J. P. Reese, and C. Winton, *Pseudo-transient continuation and optimization*, Submitted for publication, 2006.

Optmization with Partial Differential Equations on Graphs

GÜNTER LEUGERING

We consider partial differential equations defined on networked domains, such as graphs or more generally multi-linked domains. In particular, we consider elliptic, hyperbolic and Petrowski-type systems on such networks subject to controls at the boundary or in the coefficients. Optimal control or optimization problems in general for such systems appear frequently in modern technologies on various scales, such as traffic-, water-, gas-networks and flexible structures on the macro-scale, micro-mechanical structures, blood-flow, neuronal networks on the micro-scale, photonic crystals, nano-tubes on the meso- and nano-scale. The first problem to address is the proper modeling of the PDE-systems on networks. While there are still interesting and challenging questions in this respect, in particular for traffic-networks, there is meanwhile an extensive literature available. Spectral analysis, optimization, optimal control and the optimization of networks all by themselves are, however, a major focus of current research. In the talk presented we consider

on periodic networks as being used in nano-tubes and photonic crystals. One considers Maxwell's equations, or wave-equations in locally 1-d situation, as well as Schrödinger equtations with varying potentials. Bilinear controllability is a major concern from a mathematical point of view. Optimal control of such systems is the subject of Quantum Control. In case of very fine structures the question of homogenization comes into play, therefore we discuss an average model residing in a 2-d or 3-d domain which is the limit in the sense of two-scale convergence for a special product topology applied to Kuratowski's set-convergence. See [2] The homogenization procedure pursued here uses supports of singular measures as being representative for the edges of the graph in a typical cell. The novelty lies with inclusion of optimization into the procedure. The point of view taken here is the direct method of calculus of variations. Thus, we do not homogenize optimality systems, but rather the optimization problem itself. We also discuss the role of the Bloch-Floquet theory and Bloch-expansions in order to diagonalize the operators on the ϵ-scale.The second major focus of the talk is devoted to the concept of topological gradients developed by J. Sokolwoski. We present a new concept for graphs [1]. In this concept one mimics the situation for 2-d domains, in that one considers the possibility of replacing a multiple node with high edge degree by a cycle of new nodes having smaller edge degree. Thus one 'digs a hole' of 'size' ρ into the graph. An energy functional is considered parametrized by ρ, and an asymptotic analysis is given which gives rise to a proper definition of a topological gradient. This will make it possible to systematically treat shape and topology optimization problems on networks.

References

[1] G. Leugering and J. Sokolowski, *Topological gradients for problems on graphs*, preprint 2006.
[2] G.Leugering and P. Kogut, *Homogenization of optimal control problems on periodic domains*, preprint 2006.

On Two Applications of PDE Constrained Optimization

Hans Josef Pesch
(joint work with Verena Petzet, Kati Sternberg, and Kurt Chudej)

Two challenging application problems of PDE constrained optimization are investigated and presented with their associated numerical results.

First, a parameter optimization problem for the optimal positioning and dimensioning of laser beams in multi-beam laser welding techniques is investigated. The objective function to be minimized is a quantification of the likelihood of hot crack initiation which constitutes a severe problem in laser welding technology. The optimization problem is subject to several constraints. One is given in form of a boundary value problem for an elliptic PDE, the others are inequality constraints including also a state variable inequality constraint. The solution of the PDE can here be obtained semi-analytically, i. e. parts of the solution formulae in form of an absolutely convergent series must be computed numerically, here by evaluating

modified Bessel functions. Experimental results show that the computed optimal design of the multi-beam laser configuration indeed avoids the arising of hot cracks.

The second problem, deals with molten carbonate fuel cells. Such fuel cells allow an efficient and environmentally friendly energy production by converting the chemical energy contained in the fuel gas. Their dynamical behavior can be described by a large scale system of 2D quasilinear partial differential algebraic equations (PDAEs) of dimension 28. They are of mixed parabolic-hyperbolic type with integral terms in the right hand side and initial and nonlinear boundary conditions, the latter governed by a system of ordinary differential algebraic equations. Numerical results of simulations including of optimal control of the fuel cell through certain boundary conditions in case of almost time optimal load changes are presented.

Acknowledgement: These research projects have been supported by the Bayerische Forschungsstiftung and the German Federal Ministry of Education and Research.

References

[1] K. Chudej, *Index Analysis for Singular PDE Models of Fuel Cells*, in: A. Di Bucchianico, R. M. M. Mattheij, M. A. Peletier (Eds.), Progress in Industrial Mathematics at ECMI 2004. — Berlin: Springer, 212–216, 2006.

[2] K. Chudej, P. Heidebrecht, V. Petzet, S. Scherdel, K. Schittkowski, H. J. Pesch, and K. Sundmacher, *Index Analysis and Numerical Solution of a Large Scale Nonlinear PDAE System Describing the Dynamical Behaviour of Molten Carbonate Fuel Cells*, Zeitschrift für Angewandte Mathematik und Mechanik **85** (2005), 132–140.

[3] K. Chudej, K. Sternberg, P. Heidebrecht, H. J. Pesch, and J. Rang, *On Molten Carbonate Fuel Cells: 2D PDAE Models, Index Analysis, and Numerical Simulation*, submitted for publication.

[4] K. Chudej, K. Sternberg, and H. J. Pesch, *Numerical Simulation of Molten Carbonate Fuel Cells*, in: E. P. Hofer, E. Reithmeier (Eds.), Modeling and Control of Autonomous Decision Support Based Systems. — Aachen: Shaker Verlag, 227-234, 2005.

[5] P. Heidebrecht, *Modelling, Analysis and Optimisation of a Molten Carbonate Fuel Cell with Direct Internal Reforming (DIR-MCFC).* Fortschritt-Berichte VDI, Reihe 3, Nr. 826, Düsseldorf, 2005.

[6] P. Heidebrecht and K. Sundmacher, *Molten carbonate fuel cell (MCFC) with internal reforming: model-based analysis of cell dynamics*, Chemical Engineering Science **58**, 2003, 1029–1036.

[7] P. Heidebrecht and K. Sundmacher, *Dynamic Modeling and Simulation of a Countercurrent Molten Carbonate Fuel Cell (MCFC) with Internal Reforming*, Fuel Cells **3–4**, 2002, 166–180.

[8] V. A. Karkhin, H. J. Pesch, A. S. Ilin, A. A. Prikhodovsky, V. V. Plochikhine, M. V. Makhutin, and H.-W. Zoch, *Effects of Latent Heat of Fusion on Thermal Processes During Welding*, in: H. Cerjak, H. K. D. H. Bhadeshia, E. Kozeschnik (Eds.): Mathematical Modelling of Weld Phenomena 7. — Graz: Verlag der Technischen Universität Graz, 39–62, 2005

[9] V. A. Karkhin, H. J. Pesch, A. S. Ilin, A. A. Prikhodovsky, V. V. Plochikhine, M. V. Makhutin, and H.-W. Zoch, *Effects of Latent Heat of Fusion on Thermal Processes in Laser Welding of Alumium Alloys*, Science and Technology of Welding and Joining 10, No. 1 (2005), 1–7.

[10] H. J. Pesch, V. Petzet, C. Büskens, V. A. Karkhin, M. V. Makhutin, A. A. Prikhodovsky, and V. V. Plochikhin, *Numerical Optimization as a Key Tool for the Improvement of Advanced*

Multi-Beam Laser Welding Techniques, in: A. Bode, F. Durst (Eds.), High Performance Computing in Science and Engineering. — Berlin: Springer, 153–166, 2005.

[11] H. J. Pesch, K. Sternberg, and K. Chudej, *Towards the Numerical Solution of a Large Scale PDAE Constrained Optimization Problem Arising in Molten Carbonate Fuel Cell Modeling*, in: G. Di Pillo, M. Roma (Eds.), Large Scale Nonlinear Optimization. — New York: Springer, 2006, 243–253.

[12] J. Rang and K. Chudej, *A perturbation index for a singular PDE model of a fuel cell*, Mathematik-Bericht Nr. 2004/6, Institut für Mathematik, Technische Universität Clausthal, Germany, 2004.

Adaptive Solution of PDE-Constrained Optimal Control Problems

Rolf Rannacher

(joint work with Roland Becker, Boris Vexler, Malte Braack, and Dominik Meidner)

This invited survey lecture discusses various aspects of mesh adaptation in solving PDE-constrained optimal control problems. Within the framework of a Galerkin discretization mesh adaptation can be based on a posteriori error estimation following the general concept of the 'Dual Weighted Residual' method ('DWR' method) developed in [3] and [1]. By the Lagrangian formalism the optimization problem is reformulated as a saddle-point boundary value problem, the so-called KKT system, which is discretized by a Galerkin finite element method. The accuracy of the discretization is controlled by residual-based a posteriori error estimates. This paves the way toward model reduction in PDE-constrained optimal control.

We consider an optimal control problem (OCP) in the abstract setting

$$J := J(u,q) \to \min, \qquad A(u,q) = 0 \tag{1}$$

where u is the *state* and q the *control*, which are required to satisfy a stationary or nonstationary PDE system. In solving this problem numerically, one is confronted with the following fundamental questions:

- *What is the appropriate notion of admissibility* of states $u = u(q)$?
  - Discretization introduces perturbation of state equation.
  - Since accuracy in the solution of PDEs is expensive, the work should be reduced by adaptive discretization.
  - Accuracy requirements should observe intrinsic problem sensitivities and are a modeling issue.
- *How to measure* admissibility?
  - In PDEs the choice of error measures is a delicate matter.
  - Error estimates should not rely on unknown or worst-case oriented stability properties of the underlying PDE model.

The DWR method offers a universal approach for dimension reduction which is based on computationally generated sensitivity information and can be applied for general PDE systems and any Galerkin-type discretization in space or space-time.

Adaptivity in optimal control may aim at error control w.r.t. the cost functional,

$$|J_h - J| \le \eta(u_h, q_h, \lambda_h) := \sum\nolimits_{K\in\mathbb{T}_h} \eta_K, \tag{2}$$

$J_h := J(u_h, q_h)$, and mesh adaptation driven by the local error indicators η_K,

$$\eta_K \gg \delta \;\Rightarrow\; \text{refine}, \quad \eta_K \approx \delta \;\Rightarrow\; \text{keep}, \quad \eta_K \ll \delta \;\Rightarrow\; \text{coarsen}$$

Other relevant error measures may be given by certain norms of state $\|u - u_h\|$ and control $\|q - q_h\|$.

Functional-oriented a posteriori error estimates can be derived by the general approach of the DWR method. The frame is the Galerkin discretization of the abstract optimal control problem (1) in subspaces $V_h \times Q_h \subset V \times Q$,

$$J(u_h, q_h) = \min! \qquad a(u_h, q_h)(\psi_h) = 0. \tag{3}$$

Using the Lagrangian functional $L(u, q, \lambda) := J(u, q) - a(u, q)(\lambda)$, with the adjoint variable λ, the discrete first-order optimality condition (KKT system) reads

$$\left\{ \begin{array}{l} J'_u(u_h, q_h)(\phi_h) - a'_u(u_h, q_h)(\phi_h, \lambda_h) \\ J'_q(u_h, q_h)(\chi_h) - a'_q(u_h, q_h)(\chi_h, \lambda_h) \\ -a(u_h, q_h)(\psi_h) \end{array} \right\} = 0 \quad \forall\{\phi_h, \chi_h, \psi_h\}. \tag{4}$$

with the corresponding residuals

$$\begin{aligned} \rho^*(\lambda_h)(\cdot) &:= J'_u(u_h, q_h)(\cdot) - a'_u(u_h, q_h)(\cdot, \lambda_h) \\ \rho^q(q_h)(\cdot) &:= J'_q(u_h, q_h)(\cdot) - a'_q(u_h, q_h)(\cdot, \lambda_h) \\ \rho(u_h)(\cdot) &:= -a(u_h)(\cdot) \end{aligned}$$

Then, adopting the natural concept of error control w.r.t. the cost functional, an abstract argument yields the following a posteriori error representation

$$J_h - J = \tfrac{1}{2}\underbrace{\rho^*(\lambda_h)(u - I_h u)}_{\text{dual residual}} + \tfrac{1}{2}\underbrace{\rho^q(q_h)(q - I_h q)}_{\text{control residual}} + \tfrac{1}{2}\underbrace{\rho(u_h)(\lambda - I_h\lambda)}_{\text{primal residual}} + R_h, \tag{5}$$

for arbitrary $I_h u, I_h\lambda \in V_h$ and $I_h q \in Q_h$. The remainder R_h is cubic in the errors $e^u := u - u_h$, $e^q := q - q_h$, and $e^\lambda := \lambda - \lambda_h$.

For prototypical examples, such as Neumann boundary control in the stationary semi-linear diffusion problem (for details see [5])

$$-\Delta u + s(u) = f \quad \text{in } \Omega \subset \mathbb{R}^2, \quad \partial_n u_{|\Gamma_N} = 0, \quad \partial_n u_{|\Gamma_C} = q, \tag{6}$$

with the cost functional ($\bar{u} \equiv 1,\ \alpha \ge 0$)

$$J(u, q) = \tfrac{1}{2}\|u - \bar{u}\|^2_{\Gamma_O} + \tfrac{1}{2}\alpha\|q\|^2_{\Gamma_C} \;\to\; \min! \tag{7}$$

the error representation (5) can be converted into an error estimate of the form

$$|J_h - J| \;\le\; \eta_\omega := \sum_{K\in\mathbb{T}_h} \underbrace{\{\, \rho^u_K\, \omega^\lambda_K + \rho^\lambda_K\, \omega^u_K + \rho^q_K\, \omega^q_K \,\}}_{\eta_K \text{ (refinement indicators)}}, \tag{8}$$

Here, the residual terms and weights are given in the usual form by $\rho^u_K := \|f + \Delta u_h - s(u_h)\|_K + h_K^{-1/2}\|[\partial_n u_h]\|_{\partial K}$ and $\omega^\lambda_K := \|\lambda - I_h\lambda\|_K + h_K^{1/2}\|\lambda - I_h\lambda\|_{\partial K}$, respectively, and similarly for the other terms. The weights may be approximated

by local post-processing, e.g., simply by patch-wise higher-order interpolation, $(\lambda - I_h\lambda)_{|K} \approx (I_{2h}^{(2)}\lambda_h - \lambda_h)_{|K}$. The mesh adaptation follows the 'error balancing' concept,

$$\eta := \sum\nolimits_{K\in\mathbb{T}_h} \eta_K, \quad \eta_K \approx TOL/N, \quad N := \#\{K \in \mathbb{T}_h\} \quad \Rightarrow \quad \eta \approx TOL.$$

The 'weighted' error estimate (8) is compared with 'energy norm-type' error estimators of the form

$$\eta_E := c_I \Big(\sum\nolimits_{K\in\mathbb{T}_h} h_K^2 \left\{ (\rho_K^u)^2 + (\rho_K^\lambda)^2 + (\rho_K^q)^2 \right\} \Big)^{1/2}, \tag{9}$$

which have been given by Liu/Yan [11, 12], Dahmen/Kunoth [8, 10], and Hoppe et al. [9]. It has to be noted that both estimates, (8) and (9), do not work well for mesh adaptation in parameter estimation due to the inherent ill-conditioning of such problems. In this case the DWR method needs to be applied in a more sophisticated way, in order to give error bounds for the control, $\|q - q_h\|$.

The superiority of mesh adaptation by the DWR method over other residual-based approaches is demonstrated by examples from optimal flow control (drag minimization in viscous flow; see Becker [2]) and parameter estimation as well as model calibration in reactive flows (determination of Arrhenius coefficients and diffusion models from measurements; see Becker/Vexler [7]). Other applications of the DWR method to various kinds of optimal control problems can be found in the survey article [4]. Current work concerns the use of the DWR method in optimal control of nonstationary problems, for control and state constraints, in estimating of *distributed* parameters, in fluid-structure interaction, in optimal experimental design, and last but not least in software development [6].

References

[1] W. Bangerth and R. Rannacher, *Adaptive Finite Element Methods for Differential Equations*, Lectures in Mathematics, ETH Zürich, Birkhäuser, Basel 2003.

[2] R. Becker, *Mesh adaptation for stationary flow control*, J. Math. Fluid Mech. 3, 317-341 (2001).

[3] R. Becker and R. Rannacher, *An optimal control approach to error estimation and mesh adaptation in finite element methods*, Acta Numerica 2000 (A. Iserles, ed.), pp. 1-102, Cambridge University Press, 2001.

[4] R. Becker, D. Meidner, R. Rannacher, and B. Vexler: *Adaptive finite element methods for PDE-constrained optimal control problems*, Reactive Flows, Diffusion and Transport (R. Rannacher, ed.), Springer, to appear 2006.

[5] R. Becker, H. Kapp, and R. Rannacher, *Adaptive finite element methods for optimal control of partial differential equations: basic concepts*, SIAM J. Optim. Control 39, 113-132 (2000).

[6] R. Becker, D. Meidner, and B. Vexler, *RoDoBo: A C++ library for optimization with stationary and nonstationary PDEs*, Institute of Applied Mathematics, University of Heidelberg, URL http://www.rodobo.uni-hd.de, 2005.

[7] R. Becker and B. Vexler, *A Posteriori error estimation for finite element discretization of parameter identification problems*, Numer. Math. 96, 435–459 (2004).

[8] W. Dahmen and A. Kunoth, *Adaptive wavelet methods for linear-quadratic elliptic control problems*, SIAM J. Contr. Optim. 43, 1640–1675 (2005).

[9] R. Hoppe, *Adaptive finite element methods for optimally controlled elliptic problems with control constraints*, Lecture at Workshop on PDE Constrained Optimization, Center for International Mathematics, Tomar, Portugal, July 26–29,2005.
[10] A. Kunoth, *Adaptive wavelet schemes for an elliptic control problem with Dirichlet boundary control*, Numer. Algor. 39, 199-220 (2005).
[11] W. Liu and N. Yan, *A posteriori error estimates for some model boundary control problems*, J. Comput. Appl- Math. 120, 159–173 (2000).
[12] W. Liu and N. Yan, *A posteriori error estimates for control problems governed by parabolic equations*, Numer. Math. 93, 497–521 (2003).

A virtual control concept for optimal control problems with pointwise state constraints

Arnd Rösch
(joint work with Svetlana Cherednichenko, and Klaus Krumbiegel)

We investigate an optimal control problem with pointwise control and state constraints

$$(P)\quad \begin{cases} \min J(y,u) := & \frac{1}{2}\|y-y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} & & \\ \quad \text{subject to} & (Ay)(x) = u(x) & & \text{in } \Omega \\ & y = 0 & & \text{on } \Gamma \\ & 0 \le u(x) \le b & \text{a.e.} & \text{in } \Omega \\ & y(x) \ge y_c(x) & \text{a.e.} & \text{in } \Omega' \end{cases}$$

where A stands for an elliptic operator. Problems of this type contain theoretical and numerical difficulties. One possibility to overcome these problems is a Lavrentiev type regularization introduced in Meyer, Rösch, and Tröltzsch [1]. Here, the state constraint $y = y_c$ is replaced by a mixed control-state constraint $y + \varepsilon u \le y_c$. In [1], the strong convergence of regularized solutions to the solution of the original problem has been shown.

In this paper, we extend these results. Under a Slater type assumption, we prove an estimate for the regularization error of the form

$$\nu\|\bar u - \bar u_\varepsilon\|^2_{L^2(\Omega)} + \|\bar y - \bar y_\varepsilon\|^2_{L^2(\Omega)} \le \varepsilon C.$$

Moreover, we investigate the case of perturbed data

$$\begin{aligned} \|y_d - y_d^\sigma\|_{L^2(\Omega)} &\le \sigma_d, \\ \|y_c - y_c^\sigma\|_{L^\infty(\Omega')} &\le \sigma_c. \end{aligned}$$

In this case, the stability estimate

$$\nu\|\bar u - \bar u_\varepsilon^\sigma\|^2_{L^2(\Omega)} + \frac{1}{2}\|\bar y - \bar y_\varepsilon^\sigma\|^2_{L^2(\Omega)} \le C_1\varepsilon + C_2\sigma_c + \frac{1}{2}\sigma_d^2$$

is obtained. For more details see [2].

However, this approach needs a specific situation: The state constraints have to be required where the control is acting. Clearly, the situation changes completely if we consider the boundary control problem

$$(P')\begin{cases} \text{minimize} & J(y,u) := \dfrac{1}{2}\|y-y_d\|^2_{L^2(\Omega)} + \dfrac{\nu}{2}\|u\|^2_{L^2(\Gamma)} & \\ \text{subject to} & Ay + y = 0 & \text{in } \Omega \\ & \partial_{n_A} y = u & \text{on } \Gamma \\ & u_a \le u(x) \le u_b & \text{a.e. on } \Gamma \\ & y(x) \ge y_c(x) & \text{a.e. in } \Omega, \end{cases}$$

Here, we will introduce a new virtual control concept by

$$(P'_\varepsilon)\begin{cases} \min & J_\varepsilon(y,u,v) := \dfrac{1}{2}\|y-y_d\|^2_{L^2(\Omega)} + \dfrac{\nu}{2}\|u\|^2_{L^2(\Gamma)} + \dfrac{f(\varepsilon)}{2}\|v\|^2_{L^2(\Omega)} & \\ \text{subject to} & Ay + y = g(\varepsilon)v & \text{in } \Omega \\ & \partial_{n_A} y = u & \text{on } \Gamma \\ & u_a \le u(x) \le u_b & \text{a.e. on } \Gamma \\ & y(x) \ge y_c(x) - h(\varepsilon)v(x) & \text{a.e. in } \Omega \\ & 0 \le v(x) \le v_b & \text{a.e. in } \Omega \end{cases}$$

with a virtual control v. The functions f, g, and h are suitable chosen continuous functions. Under a general assumption, we prove an estimate for the regularization error of the form

$$\nu\|\bar{u}-\bar{u}_\varepsilon\|^2_{L^2(\Gamma)} + \|\bar{y}-\bar{y}_\varepsilon\|^2_{L^2(\Omega)} \le C_1\frac{\kappa(\varepsilon)}{\kappa(\varepsilon)+\tau} + C_2\frac{(g(\varepsilon))^2}{f(\varepsilon)}$$

with

$$\kappa(\varepsilon) = bh(\varepsilon) + \frac{Cg(\varepsilon)}{\sqrt{f(\varepsilon)}}.$$

Acknowledgements. This work is supported by FWF P18090-N12.

References

[1] C. Meyer, A. Rösch, and F. Tröltzsch, *Optimal control of PDEs with regularized pointwise state constraints*, Computational Optimization and Applications, accepted for publication.

[2] A. Rösch and S. Cherednichenko, *Error estimates for the regularization of optimal control problems with pointwise control and state constraints*, Zeitschrift für Analysis und ihre Anwendungen, submitted.

Numerical Methods for Shape Optimization of Variational Inequalities

Jan Sokolowski
(joint work with Piotr Fulmanski, Antoine Laurain, Jean-Francois Scheid, and Antoni Zochowski)

Shape gradients and topological derivatives are used for numerical methods of levelset type used for numerical solution of energy maximization for a class of unilateral problems. We combine the recent results on topological differentiability of solutions to variational inequalities [14] with the technique of level set [15]. The numerical method is devised for scalar problem, however the extension to the elasticity boundary value problems with unilateral conditions is straightforward.

Topology and shape optimization. The shape sensitivity analysis is combined with the topological derivatives in order to perform the shape and topology optimization. The examples of applications include Signorini problem and the frictionless contact problem in elasticity. The directional differentiability of shape functionals with respect to smooth perturbations of the boundary of geometrical domain is already proved in the monograph [9]. Recently, the form of topological derivatives for variational inequalities is established in the paper [14].

Numerical method. We propose a numerical method which combines the level set approach with the shape and topological derivatives in order to solve a class of shape optimization problems for variational inequalities. This is a joint work with Piotr Fulmanski (Lodz), Antoine Laurain and Jean-Francois Scheid (Nancy). The first numerical results confirm that the method is convergent to optimal shape - for maximization of the energy in the case of Signoroni problem.

Asymptotic analysis. Theoretical justification of procedures for computing topological derivatives for variational inequalities is given in [14]. The references on the subject of the topology and shape optimizations are listed below. In the general case of elliptic boundary value problems, the topological derivatives are derived using the outer and inner approximations of singularly perturbed elliptic problems. In the particular case of unilateral problems, the outer approximation is only constructed.

References

[1] J. Cea, S. Garreau, Ph. Guillaume, and M. Masmoudi, *The shape and topological optimizations connection*, Comput. Methods Appl. Mech. Eng. **188**, No. 4(2000), pp. 713-726.
[2] H.A. Eschenauer, V.V. Kobelev, and A. Schumacher, *Bubble method for topology and shape optimization of structures* Struct. Optimiz. **8**(1994) pp. 42-51.
[3] T. Lewinski and J. Sokołowski, *Energy change due to appearing of cavities in elastic solids*, *Int. J. Solids & Structures* **40**(2003), 1765-1803.
[4] W. G. Mazja, S. A. Nazarov, and B. A. Plamenevskii, *Asymptotic theory of elliptic boundary value problems in singularly perturbed domains* Vol. 1, 2, Basel: Birkhäuser Verlag, 2000.
[5] S. A. Nazarov and B. A. Plamenevsky, *Elliptic Problems in Domains with Piecewise Smooth Boundaries*, De Gruyter Exposition in Mathematics 13, Walter de Gruyter, 1994.

[6] S. A. Nazarov and J. Sokołowski, *Asymptotic analysis of shape functionals*, Journal de Mathématiques pures et appliquées, 82-2(2003), 125-196.
[7] S. A. Nazarov and J. Sokołowski, *Selfadjoint extensions for elasticity system in application to shape optimization*, Bulletin of the Polish Academy of Sciences – Mathematics, 52 (2004), no. 3, 237–248.
[8] M. Rao and J. Sokołowski, *Tangent sets in Banach spaces and applications to variational inequalities*, Les prépublications de l'Institut Élie Cartan 42/2000.
[9] J. Sokołowski and J-P. Zolesio, *Introduction to Shape Optimization. Shape Sensitivity Analysis*, Springer Verlag, 1992.
[10] J. Sokołowski and A. Żochowski, *On topological derivative in shape optimisation*, INRIA-Lorraine, Rapport de Recherche No. 3170, 1997.
[11] J. Sokołowski and A.Żochowski, *On topological derivative in shape optimization*, SIAM Journal on Control and Optimization. **37**, Number 4 (1999), pp. 1251–1272.
[12] J. Sokołowski and A. Żochowski, *Topological derivatives of shape functionals for elasticity systems* Mechanics of Structures and Machines 29(2001), pp. 333-351.
[13] J. Sokołowski and A. Żochowski, *Optimality conditions for simultaneous topology and shape optimization*, SIAM Journal on Control and Optimization. **42**, Number 4 (2003), pp. 1198–1221.
[14] J. Sokołowski and A. Żochowski, *Modelling of topological derivatives for contact problems*, Numerische Mathematik, **102**, No. 1 (2005), pp. 145-179.
[15] P. Fulmanski, A. Laurain, J.-F. Scheid, and J. Sokołowski, *Shape optimization for Signorini problem by a levelset method*, in preparation.

On Regularization of Elliptic Control Problems With Pointwise state Constraints

Fredi Tröltzsch
(joint work with Peter Deuflhard, Christian Meyer, Uwe Prüfert, Anton Schiela, and Martin Weiser)

Consider the elliptic optimal control problem with pointwise constraints on the control and the state,

$$\min J(y,u) = \frac{1}{2}\,\|y-y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\,\|u\|^2_{L^2(Q)} \tag{1}$$

subject to

$$-\Delta y + y = u \qquad \text{in } \Omega \tag{2}$$

with homogeneous Neumann or Dirichlet boundary condition. Moreover, pointwise constraints are imposed on the state $y \in C(\bar\Omega)$,

$$y_a \le y(x) \le y_b \qquad \forall x \in \bar\Omega. \tag{3}$$

The following data are given: A bounded domain $\Omega \subset \mathbb{R}^N$, $N \in \{2,3\}$ with $C^{0,1}$-boundary Γ, a function $y_d \in L^2(\Omega)$, $\nu > 0$, and bounds $y_a < y_b$. For all $u \in L^2(\Omega)$, a unique state $y = y(u)$ exists in $Y = H^1(\Omega) \cap C(\bar\Omega)$. Moreover, there exists a unique optimal control $\bar u$ with associated state $\bar y$.

The theory of necessary optimality conditions shows that the Lagrange multipliers for the state constraints associated with the optimal state $\bar y$ are regular Borel measures, cf. Casas [4], for instance. We also refer to the discussion of

the structure of the multipliers by Bergounioux and Kunisch [2]. In view of these results, the numerical solution of such problems is more complicated than for control-constrained problems so that different regularization methods have been investigated.

Ito and Kunisch [5] and Bergounioux and Kunisch [3] have suggested a Moreau-Yosida based regularization technique that is related to a penalization of the state constraints. The resulting unconstrained problems can be efficiently solved by a primal-dual active set strategy or, in the nonlinear case, by a semismooth Newton method.

Later, inspired by the theory of problems with bottleneck constraints, Meyer et al. [7] applied a Lavrentiev type regularization to a problem with mixed control-state constraints of the type

$$\lambda\, u(x) + y(x) \geq y_a(x), \qquad u(x) \geq 0 \quad \text{a.e. in } \Omega.$$

Here, the corresponding Lagrange multipliers are functions of L^p-spaces, cf. [10]. The proof of this regularity result is fairly technical.

It is much easier in the case of the constraints $y_a \leq \lambda u + y \leq y_b$ ($\lambda > 0$) which can be transformed to a control-constrained problem by defining a new control $v := \lambda u + y$. This idea, that has already been mentioned in [10], was applied to an associated semilinear elliptic problem by Meyer and Tröltzsch [8]. After regularization, a primal-dual active set method can be directly applied. Numerical tests showed that the performance of this method is comparable to the one by Bergounioux and Kunisch mentioned above. The main strength of the Lavrentiev regularization is that the structure of a state-constrained problem is preserved and the Lagrange multipliers are regular. This facilitates the analysis of some numerical methods in function space.

Prüfert et al. [9] investigated a primal-dual interior point method for an elliptic problem with unilateral and regularized state constraint. Here, the regularization permits to set up the method in function space and to work out an associated convergence analysis. For instance, the existence of the central path and convergence rates can be shown. The main idea of this interior point method consists of adding a logarithmic barrier function to the objective (1) and to solve the problem

$$\min \Big\{ J(y,u) - \mu \int_\Omega \big(\ln(\lambda u + y - y_a) + \ln(y_b - \lambda u - y) \big)\, dx \Big\} \tag{4}$$

subject to the equation (2) and $y_a < y(x) < y_b$.

Here, $\mu > 0$ is a positive path parameter tending to zero, while $\lambda > 0$ is a fixed small Lavrentiev regularization parameter. Though the introduction of the log-barrier functional (4) is a regularization itself, the additional Lavrentiev regularization has several positive effects. First of all, it can be shown that, for all fixed $\mu > 0$, the problem (4), (2) has a unique solution (u_μ, y_μ) such that $\lambda u_\mu + y_\mu$ has a distance $\varepsilon(\mu) > 0$ to the boundaries y_a, y_b. In this way, the solution stays in the interior of the admissible set. Moreover, the mapping $\mu \mapsto (u_\mu, y_\mu)$ is

continuously differentiable from $(0, \infty)$ to suitable spaces and it holds

$$\|(u_\mu, y_\mu) - (\bar{u}, \bar{y})\|_{L^\infty(\Omega)\times C(\bar{\Omega})} \leq c\sqrt{\mu}$$

as $\mu \downarrow 0$. Altogether, this shows the existence of the central path and the existence of its limit. The main tool behind the proof is the transformation of the problem (4), (2) with mixed control-state constraints to a problem with control constraints by introducing the auxiliary control $v := \lambda u + y$.

The performance of active set and interior point methods was studied detailed by Bergounioux et al. [1] for discretized versions of elliptic problems. Inspired by their work, Meyer et al. [6] compared an active set and an interior point method for regularized bilateral state constraints in function space. Here, the Lavrentiev regularization turns out to be very useful. While the interior point method can be directly extended to function spaces, this is not clear for active set strategies. Without regularization, measures appear as Lagrange multipliers. This causes specific difficulties, since primal-dual active set strategies are based on a pointwise evaluation of controls and Lagrange multipliers.

We should also point out that the regularization improves the performance of the numerical methods. The condition numbers of the associated system matrices increased considerably with decreasing parameter λ. As a natural consequence, iteration numbers of the methods increase as well.

References

[1] M. Bergounioux, M. Haddou, M. Hintermüller, and K. Kunisch, *A comparison of a Moreau-Yosida-based active set strategy and interior point methods for constrained optimal control problems*, SIAM J. Optimization, **11** (2000), 495–521.

[2] M. Bergounioux and K. Kunisch, *On the structure of the Lagrange multiplier for state-constrained optimal control problems*, Systems and Control Letters, **48** (2002), 16–176.

[3] M. Bergounioux and K. Kunisch, *Primal-dual active set strategy for state-constrained optimal control problems*, Computational Optimization and Applications, **22** (2002), 193–224.

[4] E. Casas, *Control of an elliptic problem with pointwise state constraints*, SIAM J. Control and Optimization, **4** (1986), 1309–1322.

[5] K. Ito and K. Kunisch, *Augmented Lagrangian methods for nonsmooth, convex optimization in Hilbert spaces, Nonlinear Analysis, Theory, Methods and Applications*, **41** (2000), 591–616.

[6] C. Meyer, U. Prüfert, and F. Tröltzsch, *On two numerical methods for state-constrained elliptic control problems*, Technical report, Institut für Mathematik, Technische Universität Berlin, **5** (2005), submitted.

[7] C. Meyer, A. Rösch, and F. Tröltzsch, *Optimal control of PDEs with regularized pointwise state constraints*, Technical report, Institut für Mathematik, Technische Universität Berlin, **14** (2003), accepted by Computational Optimization and Applications.

[8] C. Meyer and F. Tröltzsch, *On an elliptic optimal control problem with pointwise mixed control-state constraints*, in A. Seeger (Ed): Recent Advances in Optimization. Proceedings of the 12th French-German-Spanish Conference on Optimization held in Avignon, September 20-24, 2004, Lectures Notes in Economics and Mathematical Systems, **563**, (2006), 187– 204.

[9] U. Prüfert, F. Tröltzsch, and M. Weiser, *The convergence of an interior point method for an elliptic control problem with mixed control-state constraints*, Technical report, Institut für Mathematik, Technische Universität Berlin, submitted.

[10] F. Tröltzsch, *Regular Lagrange multipliers for control problems with mixed pointwise control-state constraints*, SIAM J. on Optimization, **15(2)**(2005), 616–634.

Numerical solution of regularized optimal control problems with state constraints

Michael Ulbrich

We consider optimal control problems with (regularized) state constraints. For elliptic and parabolic optimal control problems with Lavrentiev-regularized state constraints we show how the problem can be converted to a control constrained problem. It is proved that the semismooth Newton approach can be applied to this problem and that the resulting method converges superlinearly. Then, we consider elliptic optimal control problems with (regularized) state constraints and with bounds on the control. We handle the control bounds by barrier functions and show rates of convergence for the Lavrentiev-regularized solutions to the solution of the state constrained problem. A semismooth reformulation of the optimality system and a sutable smoothing step for the controls are presented. Again, a superlinearly convergent nonsmooth Newton method can be applied. The efficiency of the presented approaches is documented by numerical results.

Multilevel preconditioning of interior point and semismooth Newton methods for parabolic optimal control problems with bound constraints

Stefan Ulbrich

In this talk we consider the numerical solution of optimal control problems governed by parabolic PDEs with control constraints. We discuss the construction of multigrid preconditioners for the iterative solution of the linear Newton systems arising in interior point and semismooth Newton methods.

We briefly introduce semismooth Newton and interior point methods for control constrained parabolic optimal control problems and state recent convergence results [2–6].

We show that for both optimization methods the underlying Newton system can be reduced to a coupled parabolic system. We propose three reductions of this system to symmetric systems, which are positive definite for convex problems and allow the detection of negative curvature for nonconvex problems. We now propose two strategies to solve these systems iteratively: The first is to apply a preconditioned CG-method to one of the reduced symmetric systems. Here, the preconditioner is obtained by applying a parallel multigrid method to the equivalent coupled parabolic system. The multigrid solver augments previous approaches [1] with time-domain decomposition techniques for the smoother, which leads to a highly parallel algorithm. The smoother uses a forward-backward sweep of collective Gauss-Seidel steps in parallel on small time slabs (we use 10 time

steps in each slab). Our second strategy is to use a simplified preconditioner for the reduced symmetric system that is based on an approximate factorization of the system in a forward and a backward parabolic problem. We show that this preconditioner yields in the case without bound constraints a condition number of at most 2 if the preconditioner is computed exactly. In practice we replace exact solves by a $V(1,1)$ multigrid cycle as approximate solver in the preconditioner. Finally, we show numerical results for both strategies, which show the efficiency and robustness of our approach.

References

[1] A. Borzi, *Multigrid methods for parabolic distributed optimal control problems*, J. Comp. Appl. Math, 157 (2003), 365-382.

[2] M. Hintermüller, K. Ito, and K. Kunisch, *The primal-dual active set strategy as a semi-smooth newton method*, SIAM J. Optim., 13 (2003), pp. 865–888.

[3] F. Prüfert, U. Tröltsch, and M. Weiser, *The convergence of an interior point method for an elliptic control problem with mixed control-state constraints*, Tech. Report 36-2004, Inst. of Math., TU Berlin, 2005.

[4] A. Schiela and M. Weiser, *Superlinear convergence of the Control Reduced Interior Point Method for PDE Constrained Optimization*, ZIB Report 05-15, 2005.

[5] M. Ulbrich, *Semismooth Newton methods for operator equations in function spaces*, SIAM J. Optim., 13 (2003), pp. 805–841.

[6] M. Ulbrich and S. Ulbrich, *Primal-dual interior point methods for PDE-constrained optimization*, Technical report, Department of Mathematics, TU Darmstadt, 2006.

[7] M. Weiser, *Interior point methods in function space*, SIAM J. Control Optim., 44 (2005), pp. 1766–1786.

New Aspects in Proper Orthogonal Decomposition Model Reduction

Stefan Volkwein

(joint work with Michael Hinze, Martin Kahlbacher, and Karl Kunisch)

Proper orthogonal decomposition (POD) is a powerful technique for model reduction of linear and non-linear systems. It is based on a Galerkin type discretization with basis elements created from the system itself.

In the first part (see [5]) error estimates for Galerkin POD methods for linear and semi-linear elliptic, parameter-dependent systems are proved. The resulting error bounds depend on the number of POD basis functions and on the parameter grid that is used to generate the snapshots and to compute the POD basis.

In the second part (see [4]) we investigate POD discretizations of abstract linear-quadratic optimal control problems with control constraints. We apply the discrete technique developed in [3] and prove error estimates for the corresponding discrete controls, where we combine error estimates for the state and the adjoint system from [7, 8]. The main result of the present work can be formulated as follows.

Main result: Let $\bar{u}$ denote the solution of the linear-quadratic optimal control problem, and $\bar{u}^\ell$ its POD approximation using ℓ POD basis functions for the Galerkin ansatz. Then

$$\bar{u}^\ell - \bar{u} \sim \bar{p}^\ell - \bar{p},$$

where $\bar{p} = \bar{p}(\bar{u})$ and $\bar{p}^\ell = p^\ell(\bar{u})$ denote the corresponding solutions of the continuous and discrete adjoint systems, respectively, associated to the same control $\bar{u}$. To the authors knowledge, the presented POD error estimates are the first ones for discrete controls computed by a POD Galerkin scheme.

In third part (see [6]) we focus on a model reduction technique for open loop optimal control problems of the form

$$(\mathbf{P}) \qquad \min J(y,u) \quad \text{s.t.} \quad \begin{cases} y_t(t) = F(y(t)) + B(u(t)) & \text{for } t \in (0,T], \\ y(0) = y_0. \end{cases}$$

Here J denotes a cost-functional depending on the state y and the control u, with y and u depending on time $t \in [0,T]$ and the spatial variable x from the domain $\Omega \subset \mathbb{R}^d$ with $d \in \mathbb{N}$. The differential equation constraint in $(\mathbf{P})$ will be considered in weak form in a separable Hilbert space V of functions defined on Ω. The precise conditions on the possibly nonlinear operator $F : V \to V^*$ and the control operator $B : U \to V^*$ will not be given in this report. Here U denotes the control space. We utilize a Gelfand triple $V \subset H \subset V^*$, with V compactly and densely embedded in the real Hilbert space H, and V^* denoting the dual space of V. Further y_0 stands for the given initial condition.

While significant advances have been made over the last decade in efficient solving of $(\mathbf{P})$ good reasons remain for applying model reduction techniques to the partial differential equation and solving the optimal control problem for the reduced system. Such a procedure can become necessary in the case a large systems of partial differential equations or in the context of real time requirements, for example. It provides a possibility for obtaining closed loop solutions based on solving the Hamilton-Jacobi-Bellman equation which is otherwise computationally infeasible even for very coarse spatial discretizations.

To obtain low dimensional models of complex high dimensional systems many different approaches, including balanced truncation and reduced basis methods, were proposed. Here we focus on model reduction based on proper orthogonal decomposition (POD). It is based on a Galerkin technique where the basis functions $\{\psi_i\}_{i=1}^\ell$ are chosen as the solutions to

$$(1) \qquad \begin{cases} \min\limits_{\{\psi_i\}_{i=1}^\ell} \displaystyle\int_0^T \Big\| y(t,\cdot) - \sum_{i=1}^{\ell} \langle y(t,\cdot), \psi_i \rangle_X \, \psi_i \Big\|_X^2 \, \mathrm{d}t \\ \text{subject to } \langle \psi_i, \psi_j \rangle_X = \delta_{i,j} \quad \text{for } 1 \le i,j \le \ell, \end{cases}$$

with X a Hilbert space satisfying $V \subset X \subset H$. The basis defined by (1) is given by the eigenfunctions corresponding to the ℓ largest eigenvalues λ_i of the following eigenvalue problem:

$$(2a) \qquad \mathcal{R}\psi_i := \int_0^T \langle y(t), \psi_i \rangle_X \, y(t) \, \mathrm{d}t = \lambda_i \psi_i \qquad \text{for } 1 \le i \le \ell,$$

$$(2b) \qquad \langle \psi_j, \psi_i \rangle_X = \delta_{ij} \qquad \text{for } 1 \le i,j \le \ell.$$

Let us denote $X^\ell = \operatorname{span}\{\psi_i\}_{i=1}^\ell$ endowed with the norm induced by X. In (2) the function y denotes a solution to the dynamical system in $(\mathbf{P})$ computed at

a reference control u. Hence the Galerkin subspace X^ℓ depends on the control as well. Typical choices for X are $X = H$ and $X = V$. Our assumptions will guarantee that $X^\ell \subset V$. Let $P^\ell : H \to X^\ell$ denote the orthogonal projection with respect to the H norm. Since $X^\ell \subset V$ the projection P^ℓ can be extended to a bounded linear mapping from $V^* \to X^\ell \subset V^*$ satisfying $\langle P^\ell f - f, \phi \rangle ngle_{V^*,V} = 0$ for all $\phi \in X^\ell$ and $f \in V^*$. In the POD-approach to $(\mathbf{P})$ the dynamical system is replaced by a Galerkin projection on the POD-subspace X^ℓ. This results in

$$(\mathbf{P}^\ell) \qquad \min J(y^\ell, u) \quad \text{s.t.} \quad \begin{cases} y_t^\ell = P^\ell \left(F(y^\ell(t)) + B(u(t)) \right) & \text{for } t \in (0, T], \\ y^\ell(0) = P^\ell y_0. \end{cases}$$

The discretization of the control variable is a different issue that is not the focus of the present research. In Section 2 the control space U will be chosen to be finite dimensional. Note that the projection in $(\mathbf{P}^\ell)$ depends on the state y and hence on the reference control u at which the eigenvalue problem (2) is solved for the basis $\{\psi_i\}_{i=1}^\ell$. This may deter from one of the main advantages of the POD-approach for model reduction, which consists in the fact that unlike typical finite element basis functions the elements of the POD-basis reflect the dynamics of the system. In optimal control this feature gets lost if the dynamics of the state corresponding to the reference control is significantly different from that of the trajectory corresponding to the optimal control of $(\mathbf{P})$ or the POD-Galerkin approximation $(\mathbf{P}^\ell)$. To eliminate this weakness of the conventional approach we propose to consider

$$(\mathbf{P}^\ell_{OS\text{-}POD}) \qquad \begin{cases} \min J(y^\ell, u) \text{ subject to} \\ \\ y_t^\ell(t) = P^\ell(u)\, F(y^\ell(t)) + P^\ell(u)\, B(u(t)), \text{ for } t \in (0, T], \\ y^\ell(0) = P^\ell(u)\, y_0, \\ \\ y_t(t) = F(y(t)) + B(u(t)), \text{ for } t \in (0, T], \\ y(0) = y_0, \\ \\ \mathcal{R}(y(u))\psi_i = \lambda_i \psi_i \text{ for } 1 \le i \le \ell, \\ \langle \psi_j, \psi_i \rangle_X = \delta_{ij} \text{ for } 1 \le i, j \le \ell, \end{cases}$$

where we now indicate the dependence of the projection $(\mathbf{P}^\ell)$ and the correlation operator $\mathcal{R}$ on the control u. The first three lines in $(\mathbf{P}^\ell_{OS\text{-}POD})$ coincide with $(\mathbf{P}^\ell)$, the next two are the infinite dynamical system and the last two represent the eigenvalue problem characterizing the POD basis. For the optimal solution $(y^{\ell *}, y^*, \psi_i^*, \lambda_i^*, u^*)$ the problem formulation $(\mathbf{P}^\ell_{OS\text{-}POD})$ has the property that the associated POD-reduced system is computed from the trajectory corresponding to the optimal control u^* and thus, differently from $(\mathbf{P}^\ell)$, the problem of unmodelled is removed. $(\mathbf{P}^\ell_{OS\text{-}POD})$ can be considered as an optimization problem in the variables $(y^\ell, y, \psi_i, \lambda_i, u)$ or, alternatively it can be looked upon in the reduced sense with u the only independent variable, and $y^\ell, y, \psi_i, \lambda_i$ dependent variables

defined by the equations in $(\mathbf{P}^{\ell}_{OS\text{-}POD})$. In either case $(\mathbf{P}^{\ell}_{OS\text{-}POD})$ is more complicated than the original problem and we thus need to justify its scope. We shall show that the optimality system for $(\mathbf{P}^{\ell}_{OS\text{-}POD})$ involves two adjoint equations. The adjoint equation for the finite dimensional system is the common one for the POD-approach. It contains the linearization of the cost J with respect to the state-variable y^{ℓ} as forcing function. The second adjoint equation results from the infinite dimensional system and contains as forcing term information of the linearization of the correlation operator $\mathcal{R}$. The gradient of the reduced functional $u \mapsto J(y^{\ell}(u), u)$ with y^{ℓ} the solution to the POD system in $(\mathbf{P}^{\ell}_{OS\text{-}POD})$ can be expressed in terms of the solutions to these two adjoints. This can be utilized for practical realizations of $(\mathbf{P}^{\ell}_{OS\text{-}POD})$. It suggests a splitting of the variables z into $z_1 = (y^{\ell}, u)$ and $z_2 = (y, \psi_i, \lambda_i)$. Minimizing $J(y^{\ell}, u)$ with respect to the former for fixed z_2 results in the common POD-optimization problem for which first order, gradient based, or second order methods can be used. Minimization with respect to the second set of variables requires one additional forward and one adjoint sweep of the full system if gradient iterations are used. Thus for practical realization of $(\mathbf{P}^{\ell}_{OS\text{-}POD})$ we suggest a splitting scheme where minimization with respect to the z_1 is done more accurately than with respect to z_2. The benefit for this extra work is that the POD basis is updated in the direction of the minimum of J.

In the above discussion we presented OS-POD by means of the spatial correlation operator $\mathcal{R}$. There is an equivalent formulation based on a temporal correlation operator $\mathcal{K}$. For numerical realizations one can choose the formulation involving $\mathcal{R}$, respectively $\mathcal{K}$, depending on which of the two results in a descretized eigenvalue problem of smaller dimension.

To overcome the problem of unmodelled dynamics in the POD-basis it was proposed earlier [2, 9] to update the basis after a solution to $(\mathbf{P}^{\ell})$, to recompute the solution to $(\mathbf{P}^{\ell})$ with this new basis, and to possibly iterate. In [1] this updating procedure was combined with a trust region strategy which determines whether at the end of an optimization step an update of the POD-basis should be performed. The main difference between the procedure just described and $(\mathbf{P}^{\ell}_{OS\text{-}POD})$ lies in the fact that the former updates the basis according to the optimal control obtained from $(\mathbf{P}^{\ell}_{OS\text{-}POD})$ whereas the update of the POD-basis following $(\mathbf{P}^{\ell}_{OS\text{-}POD})$ respects the goal of minimizing J. On a computational level, updating the basis at the end of a $(\mathbf{P}^{\ell}_{OS\text{-}POD})$ solution step requires one forward solve of the full system, whereas a gradient step for $(\mathbf{P}^{\ell})$ for z_2 requires one forward and one backwards solve of the full system.

References

[1] E. Arian, M. Fahl, and E. Sachs, *Trust-region proper orthogonal decomposition for flow control*, Preprint.

[2] K. Afanasiev and M. Hinze, *Adaptive control of a wake flow using proper orthogonal decomposition*, Lecture Notes in Pure and Applied Mathematics, Marcel Dekker, **216**:317–332, 2001.

[3] M. Hinze, *A variational discretization concept in control constrained optimization: the linear-quadratic case*, Computational Optimization and Applications, **30**:45-61, 2005.
[4] M. Hinze and S. Volkwein, *Error estimates for abstract linear-quadratic optimal control problems using proper orthogonal decomposition*, Submitted, 2005.
[5] M. Kahlbacher and S. Volkwein, *Galerkin proper orthogonal decomposition methods for parameter dependent elliptic systems*, Submitted, 2006.
[6] K. Kunisch and S. Volkwein, *Proper Orthogonal Decomposition for Optimality Systems*, Submitted, 2006.
[7] K. Kunisch and S. Volkwein, *Galerkin proper orthogonal decomposition methods for parabolic problems*, Numerische Mathematik, 90:117-148, 2001.
[8] K. Kunisch, and S. Volkwein, *Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics*, SIAM J. Numer. Anal., **40**:492-515, 2002.
[9] S.S. Ravindran, *Adaptive reduced-order controllers for a thermal flow system using proper orthogonal decomposition*, SIAM S. Sci. Comput., **23**:1924–1942, 2002.

Interior Point Aspects in PDE Optimization

Martin Weiser

We consider three aspects of interior point methods applied to PDE constrained optimization problems. These are control reduction and parametric sensitivities in control constrained problems and self-concordance of integrated barrier functions for state constrained semi-infinite programs.

Control Reduction. Using the logarithmic barrier function for a model problem of the tracking type

$$\min_{u\in L_\infty, y\in H_0^1} \frac{1}{2}\|y-y_d\|_{L_2}^2 + \frac{\alpha}{2}\|u\|_{L_2}^2$$

subject to an elliptic state equation $-\Delta y = u$ and box constraints $-1 \le u \le 1$ on the control, the primal interior point formulation consists of the relaxed necessary conditions

$$\begin{aligned} y - y_d - \Delta\lambda &= 0 \\ \alpha u - \lambda - \frac{\mu}{1-u} + \frac{\mu}{1+u} &= 0 \\ -\Delta y - u &= 0. \end{aligned} \tag{1}$$

Transfering an idea of Hinze [2], the optimality condition (1) can be used to pointwisely eliminate the control $u = u(\lambda,\mu)$, giving the reduced system

$$\begin{aligned} y - y_d - \Delta\lambda &= 0 \\ -\Delta y - u(\lambda,\mu) &= 0. \end{aligned}$$

A similar idea has been suggested by Rösch [5] as a postprocessing step for discretized problems and by Ulbrich and Ulbrich [9] as a projection step inside affine scaling methods.

There are two advantages of this approach: (i) Since only the state and the adjoint state have to be discretized, the discretization error estimates for linear and quadratic finite elements are not impeded by the kinks occuring in the control

at the boundary of the active set. Thus, relatively high accuracy can be obtained on rather coarse meshes [10]. In particular, the active set can be represented with subgrid resolution. (ii) Since the control enters only in an integrated way, integral norms are sufficient for a convergence analysis. Assuming strong strict complementarity, superlinear convergence of the control reduced interior point method can be shown [7].

Parametric Sensitivities. Whenever the problem setting depends on some parameter p, the question arises how the parametric sensitivity $u_p(\mu)$of the interior point approximation $u(\mu)$ for some $\mu > 0$ is related to the parametric sensitivity u_p^* of the continuous solution u^*. Again assuming the strong strict complementarity, it can be shown [1] that

$$\|u(\mu) - u^*\|_{L_q} \leq c\mu^{(1+q)/2q} \quad \text{and} \|u_p(\mu) - u_p^*\|_{L_q} \leq c\mu^{1/2q}.$$

These convergence rates are also observed in numerical examples. The reason for the loss of convergence rate of the sensitivities is that, the sensitivity of the exact solution usually exhibits a discontinuity at the boundary of the active set, whereas the sensitivity of the interior point approximation is continuous. This explains in particular, that in the L_∞-norm, no convergence can be shown at all.

Semi-Infinite Programs. Many applications involve distributed state constraints but only a finite number of controls. Writing the state y as a function $y(u)$ of the control u, the optimal control problem falls in the class of semi-infinite optimization problems. Several attempts have been done at extending interior point methods from finite dimensional (non)linear programming to the semi-infinite case by simply substituting the sum over logarithmic barrier functions by an integral [4, 6, 8]. The results have been widely disappointing, in particular, no satisfactory convergence theory could be established. One reason might be that no integrated barrier function can lead to a ϑ-selfconcordant barrier function for general semi-infinite programs [3]. As a consequence, integrated barrier approaches were mostly dropped in favor of local reduction methods.

In PDE constrained optimization, however, local reduction methods are less attractive, since (i) a discretization is needed for the PDE anyway and (ii) local reduction methods usually exhibit a relatively small domain of convergence. Thus, integrated barrier methods are algorithmically attractive in this context. In fact, for a restricted setting, in a neighborhood of a generic solution with finitely many touch points with strictly positive definite Hessian of the state, θ-selfconcordance can be shown for the integrated barrier function

$$\int_\Omega y(x)^{-d/2}\, dx \quad \text{for } \Omega \subset \mathbb{R}^d.$$

This result suggests that integrated barrier approaches should be efficient in most practical problems, provided a suitable barrier function is used.

References

[1] R. Griesse, and M. Weiser, *On the interplay between interior point approximation and parametric sensitivities in optimal control*, Report 05-44, Zuse Institute Berlin, 2005.

[2] M. Hinze, *A variational discretization concept in control constrained optimization: the linear-quadratic case*, Comput. Optim. Appl., **30**:45–63, 2005.

[3] F. Jarre, *Comparing two interior-point approaches for semi-infinite programs*, Technical report, Universität Trier, April 1999.

[4] Z.-Q. Luo, C. Roos, and T. Terlaky, *Complexity analysis of logarithmic barrier decomposition methods for semi-infinite linear programming*, Appl. Numer. Math., **29**(3):379–394, 1999.

[5] A. Rösch, *Error estimates for linear-quadratic control problems with control constraints*, Optim. Meth. Softw., **21**(1):121–134, 2006.

[6] U. Schättler, *An interior-point method for semi-infinite programming problems*, Ann. Oper. Res., **62**:277–301, 1996.

[7] A. Schiela and M. Weiser, *Superlinear convergence of the control reduced interior point method for pde constrained optimization*, Report 05-15, Zuse Institute Berlin, 2005.

[8] L. Tunçel and M.J. Todd, *Asymptotic behavior of interior-point methods: a view from semi-infinite programming*, Math. Oper. Res., **21**(2):354–381, 1996.

[9] M. Ulbrich and S. Ulbrich, *Superlinear convergence of affine-scaling interior-point Newton methods for infinite-dimensional nonlinear problems with pointwise bounds*, SIAM J. Control Optim., **38**(6):1938–1984, 2000.

[10] M. Weiser, T. Gänzler, and A. Schiela, *Control reduced primal interior point methods*, Report 04-38, ZIB, 2004.

Reporter: Volker Schulz

Participants

Prof. Dr. Randolph E. Bank
Dept. of Mathematics
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0112
USA

Prof. Lorenz T. Biegler
Chemical Engineering Department
Carnegie Mellon University
5000, Forbes Avenue
Pittsburgh PA 15213
USA

Dr. Luise Blank
Naturwissenschaftliche Fakultät I
Mathematik
Universität Regensburg
93040 Regensburg

Dr. Bart van Bloemen Waanders
Sandia National Laboratories
Optimization/Uncertainty
Est Dept. (9211)
P.O. Box 5800, MS–0847
Albuquerque NM 87185-0847
USA

Prof. Dr. Dr. h.c. Hans Georg Bock
Interdisziplinäres Zentrum
für Wissenschaftliches Rechnen
Universität Heidelberg
Im Neuenheimer Feld 368
69120 Heidelberg

Dr. Alfio Borzi
Institut für Mathematik
Karl-Franzens-Universität Graz
Heinrichstr. 36
A-8010 Graz

Prof. Dr. Dr. h.c. Roland Bulirsch
Zentrum Mathematik
Technische Universität München
Boltzmannstr. 3
85747 Garching bei München

Prof. Dr. John Allen Burns
ICAM
Virginia Tech
Blacksburg, VA 24061-0531
USA

Dr. Carsten Burstedde
Institut für Angewandte Mathematik
Universität Bonn
Wegelerstr. 6
53115 Bonn

Prof. Dr. Carsten Carstensen
Institut für Mathematik
Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin

Prof. Dr. Felix L. Chernousko
Institute for Problems in
Mechanics
Russian Academy of Sciences
Prospekt Vernadskogo 101-1
Moscow 119 526
RUSSIA

Dr. Karsten Eppler
Weierstraß-Institut für
Angewandte Analysis und Stochastik
im Forschungsverbund Berlin e.V.
Mohrenstr. 39
10117 Berlin

Alexandra Gaevskaya
Institut für Mathematik der
Universität Augsburg
Universitätsstr. 14
86159 Augsburg

Prof. Dr. Nicolas R. Gauger
HU Berlin und DLR Braunschweig
AS-BS, Numerical Methods
Lilienthalplatz 7
38108 Braunschweig

Ilia Gherman
Fachbereich IV - Mathematik
Numerik - Optimierung, partielle
Differentialgleichungen
Universität Trier
54286 Trier

Prof. Dr. Philip E. Gill
Dept. of Mathematics
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0112
USA

Prof. Dr. Roland Glowinski
Department of Mathematics
University of Houston
651 P. G. Hoffman Hall
Houston TX 77204-3008
USA

Prof. Dr. Andreas Griewank
Institut für Mathematik
Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin

Prof. Dr. Wolfgang Hackbusch
Max-Planck-Institut für Mathematik
in den Naturwissenschaften
Inselstr. 22 - 26
04103 Leipzig

Prof. Dr. Matthias Heinkenschloss
CAAM Dept. MS 134
Rice University
6100 S. Main Street
Houston, TX 77005-1892
USA

Prof. Dr. Vincent Heuveline
Institut f. Angewandte Mathematik
Universität Karlsruhe (TH)
-Rechenzentrum-
Zirkel 2
76128 Karlsruhe

Prof. Dr. Michael Hintermüller
Institut für Mathematik
Karl-Franzens-Universität Graz
Heinrichstr. 36
A-8010 Graz

Prof. Dr. Michael Hinze
Institut für Numerische Mathematik
TU Dresden
Willersbau C 318
01062 Dresden

Prof. Dr. Ronald H.W. Hoppe
Lehrstuhl für Angewandte
Mathematik I
Universität Augsburg
Universitätsstr. 14
86159 Augsburg

Prof. Dr. Barbara Kaltenbacher
Lehrstuhl für Sensorik
Friedrich-Alexander-Universität
Erlangen
Paul-Gordan-Str. 3/5
91052 Erlangen

Prof. C.Tim Kelley
Department of Mathematics
North Carolina State University
Campus Box 8205
Raleigh, NC 27695-8205
USA

Prof. Dr. Karl Kunisch
Institut für Mathematik und
wissenschaftliches Rechnen
Karl-Franzens-Universität Graz
Heinrichstr. 36
A-8010 Graz

Prof. Dr. Ulrich Langer
Institut für Numerische Mathematik
Johannes Kepler Universität Linz
Altenbergerstr.69
A-4040 Linz

Prof. Dr. Günter Leugering
Lehrst. f. Angewandte Mathematik II
Universität Erlangen-Nürnberg
Martensstr. 3
91058 Erlangen

Christopher Linsenmann
Department of Mathematics
University of Houston
Houston TX 77204-3008
USA

Prof. Dr. Hans Josef Pesch
Lehrstuhl für
Ingenieurmathematik - FAN
Universitätsstraße 30
95447 Bayreuth

Prof. Dr. Rolf Rannacher
Institut für Angewandte Mathematik
Universität Heidelberg
Im Neuenheimer Feld 294
69120 Heidelberg

PD Dr. Arnd Rösch
Johann Radon Institute
Austrian Academy of Sciences
Altenberger Straße 69
A-4040 Linz

Prof. Dr. Ekkehard Sachs
Abteilung Mathematik
Fachbereich IV
Universität Trier
54286 Trier

Prof. Dr. Volker Schulz
Fachbereich IV - Mathematik
Numerik - Optimierung, partielle
Differentialgleichungen
Universität Trier
54286 Trier

Prof. Dr. Jan Sokolowski
Departement de Mathematiques
Universite de Nancy I
BP 239
F-54506 Vandoeuvre-les-Nancy Cedex

Prof. Dr. Fredi Tröltzsch
Institut für Mathematik
Technische Universität Berlin
Sekr. MA 4-5
Strasse des 17. Juni 136
10623 Berlin

Prof. Dr. Michael Ulbrich
Fachbereich Mathematik
Universität Hamburg
Bundesstr. 55
20146 Hamburg

Prof. Dr. Stefan Ulbrich
Fachbereich Mathematik
AG 10
TU Darmstadt
Schloßgartenstr. 7
64289 Darmstadt

Dr. Stefan Volkwein
Institut für Mathematik und
wissenschaftliches Rechnen
Karl-Franzens-Universität Graz
Heinrichstr. 36
A-8010 Graz

Dr. Martin Weiser
Konrad-Zuse-Zentrum für
Informationstechnik Berlin (ZIB)
Takustr. 7
14195 Berlin

Prof. Dr. Gabriel Wittum
IWR Technische Simulation
Universität Heidelberg
Im Neuenheimer Feld 368
69120 Heidelberg