THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Approximation of topology optimization problems using sizing optimization problems

Anton Evgrafov

CHALMERS | GÖTEBORG UNIVERSITY

Department of Mathematics
Chalmers University of Technology and Göteborg University
Göteborg, Sweden 2004

Approximation of topology optimization problems
using sizing optimization problems
Anton Evgrafov
ISBN 91–7291–466–1

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 2148
ISSN 0346–718X

Department of Mathematics
Chalmers University of Technology and Göteborg University
SE-412 96 Göteborg
Sweden
Telephone +46 (0)31–772 1000

Printed in Göteborg, Sweden 2004

Approximation of topology optimization problems using sizing optimization problems

Anton Evgrafov
Department of Mathematics
Chalmers University of Technology and Göteborg University

ABSTRACT

The present work is devoted to approximation techniques for singular extremal problems arising from optimal design problems in structural and fluid mechanics. The thesis consists of an introductory part and four independent papers, which however are united by the common idea of approximation and the related application areas.

In the first half of the thesis we are concerned with finding the optimal topology of truss-like structures. This class of optimal design problems arises when in order to find the optimal truss not only are we allowed to redistribute the material among the structural members (bars), but also to completely remove some parts altering the connectivity (topology) of the structure. The other half of the thesis addresses the question of the optimal topological design of flow domains for Stokes and Navier–Stokes fluids. For flows, optimizing topology means finding the optimal partition of the given design domain into disjoint parts occupied by the fluid and the impenetrable walls, given the in-flow and the out-flow boundaries. In particular, impenetrable walls change the shape and the connectivity of the flow domain.

In the first paper we construct an example demonstrating the singular behaviour of truss topology optimization problems including a linearized global buckling (linear elastic stability) constraint. This singularity phenomenon has not been known before and affects the choice of numerical methods that can be applied to the optimization problem. We propose a simple approximation strategy and establish the convergence of globally optimal solutions to perturbed problems towards globally optimal solutions to the original singular problem.

In the second paper we are concerned with the construction of finer approximating problems that allow us to reconstruct the local behaviour of a general class of singular truss topology optimization problems, namely to approximate stationary points to the limiting problem with sequences of stationary points to the regular approximating problems. We do so on the classic problem of weight minimization under stress constraints for trusses in unilateral contact with rigid obstacles.

In the third paper we extend a design parametrization previously proposed for the topological design of flow domains for Stokes flows to also include the limiting case of porous materials—completely impenetrable walls. We demonstrate that, in general, the resulting design-to-flow mapping is not closed, yet under mild assumptions it is possible to approximate globally optimal minimal-power-dissipation domains using porous materials with diminishing permeability.

In the fourth and last paper we consider the optimal design of flow domains for Navier–Stokes flows. We illustrate the discontinuous behaviour of the design-to-flow mapping caused by the topological changes in the design, and propose "minor" changes to the design parametrization and the equations that allow us to rigorously establish the closedness of the design-to-flow mapping. The existence of optimal solutions as well as the convergence of approximation schemes then easily follows from the closedness result.

LIST OF PUBLICATIONS

This thesis consists of an introductory part and the following papers:

Paper 1 A. Evgrafov, *On globally stable singular topologies*, to appear in Structural and Multidisciplinary Optimization, 2004.

Paper 2 A. Evgrafov and M. Patriksson, *On the convergence of stationary sequences in topology optimization*, submitted to International Journal for Numerical Methods in Engineering, 2004.

Paper 3 A. Evgrafov, *On the limits of porous materials in the topology optimization of Stokes flows*, submitted to Applied Mathematics and Optimization, 2003.

Paper 4 A. Evgrafov, *Topology optimization of slightly compressible fluids*, submitted to Zeitschrift für Angewandte Mathematik und Mechanik, 2004.

CONTENTS

ACKNOWLEDGEMENT

DURING the time spent on writing this dissertation, as well as on many other occasions, I have always been feeling the unflagging support of many people, including

- my supervisor, Prof. Michael Patriksson;
- my wife Elena;
- my friends;
- my parents.

I will ever be indebted to all of you.

Many thanks to the School of Mathematical Sciences for being such a welcoming and friendly place to work, and to the Swedish Research Council for financially supporting me under the grant VR-621-2002-5780.

I dedicate my dissertation to the memory of Joakim Petersson, who untimely passed away on September 30, 2002, at the mere age of 33. This work is to a great extent inspired by his research.

Anton Evgrafov

Göteborg, May 2004

INTRODUCTION AND OVERVIEW

In the present thesis we study approximation techniques for singular extremal problems arising from optimal design problems in structural and fluid mechanics. The purpose of this chapter is to provide an introduction to some of the models and methods coming from bi-level programming, topology optimization, and sensitivity analysis for non-linear programs, as well as to give a summary of the results established in the four appended papers, placing them in a proper perspective.

Mathematical programming with equilibrium constraints

Hierarchical decision-making problems are encountered in a wide variety of domains in the engineering and experimental natural sciences, and in regional planning, management, and economics. These problems are all defined by the presence of two or more objectives with a prescribed order of priority or information. In many applications it is sufficient to consider a sub-class of these problems having two levels, or objectives. We refer to the upper-level as the objective having the highest priority and/or information level; it is defined in terms of an optimization with respect to one set of variables. The lower-level problem, which in the most general case is described by a variational inequality, is then a supplementary problem parameterized by the upper-level variables. These models are known as generalized bi-level programming problems, or mathematical programs with equilibrium constraints (MPEC); see, for example, Luo et al. [LPR96].

Many extremal problems arising from applications in mechanics of solids, structures, and fluids have an inherent bi-level form. The upper-level objective function measures some performance of the system, such as its weight, stiffness, maximal contact force, pressure drop, drag, or power dissipation. This objective function is optimized by selecting design parameters, which in our case will be related to the geometry of the system and the amount of material being used. Further, the upper-level optimization is subject to design constraints, such as limits on the amount of available material, and to behavioural constraints, such as bounds on the displacements and stresses, or buckling/stability constraints. The lower-level problem describes the behaviour of the system given the choices of the design variables, the external forces acting on it, and the boundary conditions.

For linear elastic structures the behaviour is governed by the equilibrium law of minimal potential energy, which determines the values of the state variables (displacements) at the lower level. Equivalently, the equilibrium law can be expressed as a (dual) principle of the minimum of complementary energy, determining the stresses and contact forces.

Similarly, for slow (creep, Stokes) flows the flow velocity is determined by the principle of minimal potential power. In the case of faster flow, the non-linear convection effects must be taken into account in the lower-level problem, leading to the Navier–Stokes equations that (in a weak form) can be formulated as a variational inequality.

Mathematical programs with equilibrium constraints are known to be an especially difficult subclass of non-smooth non-convex NP-hard mathematical problems, which in addition violates standard non-linear programming constraint qualifications [LPR96, Chapter 1]. Furthermore, MPEC problems coming from the topology optimization of structures, solids and fluids typically violate even the novel qualifications, such as the strict complementarity conditions, or strong regularity assumptions, constructed specifically for generalized bi-level programs. Therefore, the use of approximation techniques for the numerical solution of such problems seems inevitable.

Let $\mathbf{x} \in \mathscr{X}$ (respectively, $\mathbf{y} \in \mathscr{Y}$) denote a vector of upper-level, or design (resp. lower-level, or state), variables. Given the performance functional $f : \mathscr{X} \times \mathscr{Y} \to \mathbb{R} \cup \{+\infty\}$, the mathematical program with equilibrium constraints can be stated as follows:

$$\begin{cases} \min\limits_{(\mathbf{x},\mathbf{y})} f(\mathbf{x},\mathbf{y}), \\ \text{s.t. } \ (\mathbf{x},\mathbf{y}) \in S, \\ \qquad \mathbf{y} \in \mathrm{SOL}(\mathbf{x}), \end{cases} \tag{1}$$

where $S \subset \mathscr{X} \times \mathscr{Y}$ is a set representing design, behavioural, and joint constraints, and $\mathrm{SOL} : \mathscr{X} \rightrightarrows \mathscr{Y}$ is a point-to-set mapping defined by the set of solutions to the lower-level parametric optimization, or variational inequality, problem. For example, in the simplest case $\mathrm{SOL}(\mathbf{x})$ may be given by the solution set of the following parametric optimization problem:

$$\begin{cases} \min\limits_{\mathbf{y}} F(\mathbf{x},\mathbf{y}), \\ \text{s.t. } \mathbf{y} \in C, \end{cases} \tag{2}$$

where $F : \mathscr{X} \times \mathscr{Y} \to \mathbb{R} \cup \{+\infty\}$ is a given lower-level objective function (usually convex in the second variable), and $C \subset \mathscr{Y}$ is a set of admissible state vectors (also usually convex). In view of the special structure of the feasible set of (1), one of the central questions in the study of MPEC is the verification of closedness and semi-continuity properties of the point-to-set mapping SOL, which bears close relationships with the verification of the existence of optimal solutions and applicability of optimization algorithms to a given MPEC problem.

For MPEC problems arising in the topology optimization of structures, solids, and fluids, the typical situation is as follows. For design vectors $\mathbf{x}$ found in the *interior* of the design domain, the mapping $\mathbf{x} \rightrightarrows \mathrm{SOL}(\mathbf{x})$ is in fact single-valued and continuous; whereas for designs on the boundary this mapping is at most closed, and may have unbounded, or even empty, images. This behaviour prompts to be exploited algorithmically. Thus, instead of the original MPEC (1) we solve a sequence of approximating problems

$$\begin{cases} \min\limits_{(\mathbf{x},\mathbf{y})} f(\mathbf{x},\mathbf{y}), \\ \text{s.t. } \ (\mathbf{x},\mathbf{y}) \in S_\varepsilon, \\ \qquad \mathbf{y} \in \mathrm{SOL}(\mathbf{x}), \end{cases}$$

where S_ε is a specially constructed approximation of S, avoiding the "bad" points of the design space, and $\varepsilon > 0$ is an approximation parameter eventually tending to zero.

The approximation problems are usually constructed so that the limit points of sequences of globally optimal solutions to the approximating problems are globally optimal in the original, singular problem, as the approximation parameter tends to zero.

Unfortunately, for non-convex, non-smooth, large-scale problems like instances of MPEC problems arising from structural topology optimization one can hardly expect to find globally optimal solutions, and therefore the convergence of globally optimal solutions is a positive result of more theoretical than practical significance. However, the approximation of stationary points for MPEC problems is a complicated task. Firstly, even the practical concept of a stationary point is not so easy to define in this case, because the standard constraint qualifications are violated. Secondly, even more modern MPEC-specific assumptions are violated by the optimal design problems we consider, especially at the boundary of the design domain; since it is the local behaviour of the mapping $\mathbf{x} \rightrightarrows \mathrm{SOL}(\mathbf{x})$ that is important in this case, we need to regularize or approximate the latter point-to-set mapping. Thus, in order to approximate the stationary points of the original, singular MPEC-problem (1) we introduce approximating problems of the form

$$\begin{cases} \min\limits_{(\mathbf{x},\mathbf{y})} f(\mathbf{x},\mathbf{y}), \\ \text{s.t. } (\mathbf{x},\mathbf{y}) \in S_\varepsilon, \\ \qquad \mathbf{y} \in \mathrm{SOL}_\varepsilon(\mathbf{x}), \end{cases}$$

where, as before, $\varepsilon > 0$ is a small approximation parameter, S_ε is an approximation of S, and SOL_ε is a “simple” approximation of SOL. For example, instead of considering *exact* optimal solutions $\mathbf{y}$ in (2) we may require only ε-optimal solutions; other choices are also possible (see [LiM97, FJQ99]).

Truss topology optimization

Truss topology optimization problems play an important role of being model problems in structural optimization owing to their simple and very well developed structure; yet the techniques developed for truss topology optimization problems are applicable to much wider classes of structures than trusses, including frames and finite element discretized models of solids. Systematic research on truss optimization began in the beginning of the previous century with the work of Michell [Mic04], and nowadays this is probably the most advanced area of topology optimization (cf. [BeS03]). However, some computational aspects of truss topology optimization problems still lack a theoretical basis; in the appended papers 1 and 2 we try to resolve some of them. In this section, however, we introduce the necessary truss-specific notation to put the truss topology optimization problems, and the discussion about the approximation techniques, into the general framework of MPEC.

Using the de-facto standard in the field, the *ground structure* approach [DGG64], the *design* of the truss is completely described by prescribing for each bar i, $i = 1,\dots,m$, the amount of material $x_i \geq 0$ allocated to this bar. For convenience we collect all the design variables in a vector $\mathbf{x} = (x_1,\dots,x_m)^t \in \mathbb{R}^m_+$. We introduce an index set of the present (or, active) members in the structure: $\mathscr{I}(\mathbf{x}) = \{\, i = 1,\dots,m \mid x_i > 0 \,\}$. Given a particular design $\mathbf{x} \in \mathbb{R}^m_+$, the equilibrium status of a truss in the presence of rigid obstacles that may come into unilateral frictionless contact with some nodes can be described by specifying

- a pseudo-force s_i (also known as the normalized stress; it is in fact a stress in the bar times its volume) for each bar $i \in \mathscr{I}(\mathbf{x})$ present in the structure. To simplify the notation we collect all values s_i, $i = 1, \ldots, m$, into one vector $\mathbf{s} \in \mathbb{R}^m$, assuming $s_i = 0$ for $i \notin \mathscr{I}(\mathbf{x})$ (since inactive members cannot carry any load);
- a contact force λ_j for each of the potential contact nodes $j = 1, \ldots, r$. These values are collected in a vector $\boldsymbol{\lambda} \in \mathbb{R}^r_+$; and
- a displacement u_k for each of the structural degrees of freedom $k = 1, \ldots, n$. These values are collected in a vector $\mathbf{u} \in \mathbb{R}^n$.

The values of the state variables for a specific design $\mathbf{x}$ are determined using various energy principles. Therefore, we define the complementary energy of the structure as

$$\mathscr{E}(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}) := \frac{1}{2} \sum_{i \in \mathscr{I}(\mathbf{x})} \frac{s_i^2}{E x_i} + \mathbf{g}^t \boldsymbol{\lambda},$$

as well as the linearized strain energy:

$$\Pi(\mathbf{x}, \mathbf{u}) := \frac{1}{2} \mathbf{u}^t \mathsf{K}(\mathbf{x}) \mathbf{u},$$

where $\mathsf{K}(\mathbf{x})$ is the *stiffness matrix* of the structure. The latter matrix is defined as

$$\mathsf{K}(\mathbf{x}) := \sum_{i \in \mathscr{I}(\mathbf{x})} x_i \mathsf{K}_i,$$

where $\mathsf{K}_i = E \mathsf{B}_i^t \mathsf{B}_i$ is the *local stiffness matrix* for the bar $i = 1, \ldots, m$, $\mathsf{B}_i \in \mathbb{R}^{1 \times n}$ is a kinematic transformation matrix for the bar $i = 1, \ldots, m$, and E is the Young modulus of the structural material.

Throughout the thesis, we make the blanket assumption that $\mathsf{K}(\mathbf{x})$ is *positive definite* for every *positive* design $\mathbf{x}$; a necessary and sufficient condition for this property is that $\mathsf{K}(\mathbf{1}^m)$ is positive definite. We do not loose any generality from this assumption, because the positive definiteness can be achieved by starting from an "enough rich" ground structure.

In this notation the equilibrium state of the structure under the external load $\mathbf{f} \in \mathbb{R}^n$ can be characterized using the following primal-dual pair of convex quadratic programming problems:

$$(\mathscr{C})_{\mathbf{x}}(\mathbf{f}) \begin{cases} \min\limits_{(\mathbf{s}, \boldsymbol{\lambda})} \mathscr{E}(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}), \\ \text{s.t. } \mathsf{C}^t \boldsymbol{\lambda} + \sum\limits_{i \in \mathscr{I}(\mathbf{x})} \mathsf{B}_i^t s_i = \mathbf{f}, \\ \qquad\qquad \boldsymbol{\lambda} \geq \mathbf{0}, \end{cases} \qquad (\mathscr{P})_{\mathbf{x}}(\mathbf{f}) \begin{cases} \min\limits_{\mathbf{u}} \Pi(\mathbf{x}, \mathbf{u}) - \mathbf{f}^t \mathbf{u}, \\ \text{s.t. } \mathsf{C}\mathbf{u} \leq \mathbf{g}, \end{cases}$$

where $\mathbf{g} \in \mathbb{R}^r$ is a vector of initial gaps between the contact nodes and rigid obstacles, and $\mathsf{C} \in \mathbb{R}^{r \times n}$ is a kinematic transformation matrix. We have implicitly assumed that the matrix C is *quasi-orthogonal*, that is, that $\mathsf{C}\mathsf{C}^t = \mathsf{I}$. The problem $(\mathscr{C})_{\mathbf{x}}(\mathbf{f})$ is known as the principle of minimum complementary energy, and the problem $(\mathscr{P})_{\mathbf{x}}(\mathbf{f})$ is the principle of minimum potential energy.

Equivalently, the equilibrium problem can be written as a KKT system for the pair

$(\mathscr{C})_{\mathbf{x}}(\mathbf{f})$ and $(\mathscr{P})_{\mathbf{x}}(\mathbf{f})$. Define

$$\mathsf{Q}(\mathbf{x}) := \begin{pmatrix} \mathsf{B}^t & \mathsf{C}^t & 0 \\ 0 & 0 & -\mathsf{C} \\ \mathsf{I} & 0 & -\mathsf{D}(\mathbf{x})\mathsf{B} \end{pmatrix}, \qquad \mathbf{q}(\mathbf{f}) := \begin{pmatrix} -\mathbf{f} \\ \mathbf{g} \\ \mathbf{0} \end{pmatrix},$$

and $Y := \mathbb{R}^m \times \mathbb{R}^r_+ \times \mathbb{R}^n$, where $\mathsf{B} \in \mathbb{R}^{m\times n}$ is the matrix with rows $\mathsf{B}_1,\dots,\mathsf{B}_m$, and $\mathsf{D}(\mathbf{x}) = \operatorname{diag}(\mathbf{x}) \in \mathbb{R}^{m\times m}$. Then, the pair $(\mathbf{s},\boldsymbol{\lambda})$ solves $(\mathscr{C})_{\mathbf{x}}(\mathbf{f})$ and $\mathbf{u}$ solves $(\mathscr{P})_{\mathbf{x}}(\mathbf{f})$ if and only if the vector $\mathbf{y}^* = (\mathbf{s}^t,\boldsymbol{\lambda}^t,\mathbf{u}^t)^t \in Y$ solves the following affine variational inequality problem, denoted $\mathrm{AVI}(\mathbf{q}(\mathbf{f}),\mathsf{Q}(\mathbf{x}),Y)$:

$$[\mathsf{Q}(\mathbf{x})\mathbf{y}^* + \mathbf{q}(\mathbf{f})]^t(\mathbf{y}-\mathbf{y}^*) \geq 0, \quad \text{for all } \mathbf{y} \in Y.$$

Now, one can easily spot the difficulties arising at the boundary of the design domain $\mathbb{R}^m_+$, that is, when some bars are completely removed from the ground structure:

- the objective function of the problem $(\mathscr{C})_{\mathbf{x}}(\mathbf{f})$ is only lower semi-continuous and may assume infinite values on this subset of the design space;
- similarly, the objective function of the dual problem $(\mathscr{P})_{\mathbf{x}}(\mathbf{f})$ is not strongly convex at the boundary of the design domain;
- the affine variational inequality problem $\mathrm{AVI}(\mathbf{q}(\mathbf{f}),\mathsf{Q}(\mathbf{x}),Y)$ may have zero, or multiple, solutions at the boundary.

All these difficulties have immediate implications on the properties of the corresponding multi-mapping $\mathbf{x} \rightrightarrows \mathrm{SOL}(\mathbf{x})$; as a consequence the class of algorithms capable of solving truss topology optimization problems is significantly restricted, for example when compared to the corresponding sizing problems. Thus one is tempted to replace the design domain $\mathbb{R}^m_+$ with the natural approximation $\{\, \mathbf{x} \in \mathbb{R}^m_+ \mid \mathbf{x} \geq \varepsilon \mathbf{1}^m \,\}$, with $\varepsilon \downarrow 0$, thus converting the topology optimization problem into a sequence of *sizing* ones.

For some relatively simple truss topology optimization problems (such as, e.g., compliance minimization problems, possibly with so-called “strong” stress constraints [Ach98]) the strategy we just outlined is sufficient. Such approximations have been rigorously studied for trusses without (Achtziger [Ach98]) and with (Patriksson and Petersson [PaP02]) unilateral constraints.

On the other hand, there are many other classes of topology optimization problems including important mechanical constraints (e.g., stress constraints [SvG68], local buckling constraints [GCY01], and global buckling constraints [Paper 1, this thesis]) where the simple strategy outlined above leads to erroneous results, owing to the complicated singular structure of the feasible set near the points where the truss topology changes. Historically, the study of singularity phenomena for truss topology optimization problems started with problems including stress constraints only. Sved and Ginos [SvG68] observed that such problems may have singular solutions, and the properties of the feasible region were further investigated by Kirsch [Kir90], Cheng and Jiang [ChJ92], and Rozvany and Birker [RoB94]. Cheng and Guo [ChG97] were the first to propose a more sophisticated restriction-relaxation procedure, where not only the lower bounds but also the stress constraints were perturbed. They established the convergence of the optimal values of the perturbed problems to the optimal value of the original problem, while Petersson [Pet01] (using the continuity of certain design-to-state parameterized mappings) has established the convergence of optimal solutions. Since then, the ε-perturbation strategy has been extended by many researchers, in many ways: Duysinx and Bendsøe [DuB98] and Duysinx and Sigmund [DuS98] considered continuum structures; Guo et al. [GCY01] included local buckling constraints into the problem; Patriksson and Petersson [PaP02] generalized the result for trusses including unilateral constraints; Evgrafov et al. [EPP03, EvP03, EvP03a] considered the possibility of introducing stochastic forces; and Evgrafov [Paper 1, this thesis] studied the linearized elastic stability constraint.

Despite the clear advantage of approximating the nonsmooth, singular optimization problem with a sequence of smooth and regular ones, all the sizing approximations considered above suffer from a similar difficulty: while the underlying theoretical results are concerned with the approximation of the *globally* optimal solutions, in computational practice it is impossible to solve the nonconvex approximating problems to global optimality. There are also negative results regarding this issue: the ε-perturbation approach may fail to find a globally optimal solution even for topology optimization problems with only 2 design variables [StS01]!

The analysis of the convergence of stationary points to the approximating problems towards stationary points of the limiting (that is, original) problem is difficult; for example, the dependence of the equilibrium state of the structure upon the design near the points where the topology changes is nonsmooth, and even non-Lipschitz continuous. Therefore, Evgrafov and Patriksson [Paper 2, this thesis] have designed an alternative approximation scheme, capable of approximating *both* globally optimal solutions and stationary points through corresponding sequences of global or stationary solutions to approximating problems. To achieve this, we approximate both the design set and the design-to-state mapping $\mathbf{x} \rightrightarrows \mathrm{SOL}(\mathbf{x})$, as we outlined in the previous section. Namely, we consider the following approximating feasible sets:

$$\begin{aligned}\mathscr{F}^{\varepsilon}(\mathbf{f}) := \{\,(\mathbf{x},\mathbf{s},\boldsymbol{\lambda},\mathbf{u}) \in \mathbb{R}^m_+ \times \mathbb{R}^m \times \mathbb{R}^r_+ \times \mathbb{R}^n \mid \mathbf{x} \ge o(\varepsilon)\mathbf{1}^m,\\ \mathscr{E}(\mathbf{x},\mathbf{s},\boldsymbol{\lambda}) + \Pi(\mathbf{x},\mathbf{u}) - \mathbf{f}^t\mathbf{u} \le \varepsilon,\\ \mathsf{B}^t\mathbf{s} + \mathsf{C}^t\boldsymbol{\lambda} = \mathbf{f},\\ \mathsf{C}\mathbf{u} \le \mathbf{g}\,\},\end{aligned}$$

where $o(\varepsilon)$ is a positive function of ε such that $\lim_{\varepsilon\downarrow 0} o(\varepsilon)/\varepsilon = 0$. Of course, the weak duality theorem for convex problems implies that $\mathscr{F}^0(\mathbf{f})$ corresponds to the original design-to-state mapping; for positive values of ε the "state" variables $(\mathbf{s},\boldsymbol{\lambda},\mathbf{u})$ [which in this formulation play a role equal to that of the design variables $\mathbf{x}$, and do not correspond to an equilibrium state of the truss anymore] are required to be primal-dual feasible, but only ε-optimal. We also note that only strictly positive designs are allowed in the approximating sets, so that we isolate the "problematic" points and end up with smooth and more regular problems.

Topology optimization of flow domains

The optimal control of fluid flows has long been receiving considerable attention by engineers and mathematicians, owing to its importance in many applications involving fluid related technology; see, e.g., the recent monographs [Gun03, MoP01]. According to a well-established classification in structural optimization (see [BeS03, page 1]), the absolute majority of works dealing with the optimal design of flow domains fall into the category of shape optimization. (See the bibliographical notes (2) in [BeS03] for classic references in shape optimization.) In the framework of *shape optimization*, the optimization problem formulation can be stated as follows: choose a flow domain out of some family so as to maximize an associated performance functional. The family of domains considered may be as rich as that of all open subsets of a given set satisfying some regularity criterion (see, e.g. [Fei03]), or as poor as the ones obtained from a given domain by locally perturbing some part of the boundary in a Lipschitz manner (cf. [Ton03b, GKM00, GuK98]). Unfortunately, it is typically only the problems in the latter group that can be attacked numerically. On the other hand, *topology optimization* (or, control in coefficients) techniques are known for their flexibility in describing the domains of arbitrary complexity (e.g., the number of connected components need not to be bounded), and at the same time require relatively moderate computational efforts. In particular, one may completely avoid remeshing the domain as the optimization algorithm advances, which eases the integration with existing FEM codes, and simplifies and speeds up sensitivity analysis.

While the field of topology optimization is nowadays very well established for the optimal design of solids and structures, surprisingly little work has been done for the optimal design of fluid domains. Borrvall and Petersson [BoP03] were the first to successfully consider the optimal design of flow domains for minimizing the total power of the incompressible Stokes flows, using inhomogeneous porous materials with a spatially varying Darcy permeability tensor, under a constraint on the total volume of fluid in the control region. Later, this approach has been generalized to include both limiting cases of porous materials, i.e., pure solid and pure flow regions have been allowed to appear in the design domain as a result of the optimization procedure [Paper 3, this thesis]. (We also cite the work of Klarbring et al. [KPTK03], which however study the problem of the optimal design of flow networks, where design and state variables reside in finite-dimensional spaces; in some sense this problem is an analogue of truss design problems if one can carry over the terminology and ideas from the area of optimal design of structures and solids.)

To put the topology optimization of Stokes flows into the framework of MPEC, we need to introduce some fluid-specific notation. Let Ω be a connected bounded domain of $\mathbb{R}^d$, $d \in \{2,3\}$ with a Lipschitz continuous boundary Γ. Borrvall and Petersson [BoP03] proposed to control the Stokes equations in Ω in the following manner: given the prescribed flow velocities $\mathbf{g}$ on the boundary, and forces $\mathbf{f}$ acting in the domain one adjusts the inverse permeability α of the medium occupying Ω, which depends on the control function ρ:

$$\begin{cases} \left.\begin{aligned} -\nu\boldsymbol{\Delta}\mathbf{u} + \alpha(\rho)\mathbf{u} + \boldsymbol{\nabla} p = \mathbf{f}, \\ \operatorname{div}\mathbf{u} = 0 \end{aligned}\right\}, & \quad \text{in } \Omega, \\ \mathbf{u} = \mathbf{g}, & \quad \text{on } \Gamma. \end{cases} \tag{3}$$

In the system (3), $\mathbf{u}$ is the flow velocity, p is the pressure, and ν is the kinematic viscosity. The control domain $\mathscr{H}$ is defined as follows:

$$\mathscr{H} = \{\, \rho \in L^\infty(\Omega) \mid 0 \le \rho \le 1, \text{a.e. in } \Omega, \int_\Omega \rho \le \gamma|\Omega| \,\},$$

where $0 < \gamma < 1$ is the maximal volume fraction that can be occupied by the fluid. Formally, we relate the permeability α^{-1} to ρ using a convex, decreasing, and nonnegative function $\alpha : [0,1] \to \mathbb{R}_+ \cup \{+\infty\}$, defined as

$$\alpha(\rho) = \rho^{-1} - 1.$$

[Thus, $\rho(\mathbf{x}) = 0$ corresponds to zero permeability, or solid regions, which do not permit any flow, while $\rho(\mathbf{x}) = 1$ corresponds to infinite permeability, or 100% flow regions.] In order to introduce the weak formulation of the system (3), we consider the sets of admissible flow velocities:

$$\mathscr{U}_{\text{div}} = \{\, \mathbf{v} \in H^1(\Omega) \mid \mathbf{v} = \mathbf{g} \text{ on } \Gamma,\ \operatorname{div} \mathbf{v} = 0, \text{weakly in } \Omega \,\},$$

and introduce the potential power of the viscous flow through the porous medium:

$$\mathscr{J}(\rho, \mathbf{u}) = \frac{\nu}{2} \int_\Omega \nabla\mathbf{u} \cdot \nabla\mathbf{u} + \frac{1}{2} \int_\Omega \alpha(\rho)\mathbf{u} \cdot \mathbf{u} - \int_\Omega \mathbf{f} \cdot \mathbf{u}.$$

In this notation, we end up with the following lower-level problem corresponding to (2):

$$\begin{cases} \min\limits_{\mathbf{u}} \mathscr{J}(\rho, \mathbf{u}), \\ \text{s.t. } \mathbf{u} \in \mathscr{U}_{\text{div}}. \end{cases} \tag{4}$$

Now we can see that, at least from the approximation viewpoint, the situation in flow topology optimization is similar to the situation in the topology optimization of trusses. Namely, for all designs ρ almost everywhere in Ω satisfying the inequality $\rho \geq \varepsilon$, for some $\varepsilon > 0$, the design-to-flow mapping SOL corresponding to the lower-level problem (4) parameterized by ρ is continuous. However, if we allow solid regions to appear in the domain Ω, the objective function of (4) suddenly becomes discontinuous, and thus induces a non-closed design-to-flow mapping. Therefore, topology optimization problems for Stokes fluids are ill-posed in general; however, one can establish the existence of optimal solutions at least if we take the upper-level objective functional to be equal to the lower-level objective functional, which has numerous applications for fluids [BoP03]. (Minimizing the potential power for Stokes fluids corresponds to minimizing the compliance in linear elasticity.)

Now, if we are interested in modeling faster flows, the non-linear convection effects must be taken into account. In our opinion, the most convenient way to do so is to consider the following fixed-point problem:

$$\mathbf{u} \in \operatorname*{argmin}_{\mathbf{v} \in \mathscr{U}_{\text{div}}} \left\{ \mathscr{J}(\rho, \mathbf{v}) + \int_\Omega (\mathbf{u} \cdot \nabla\mathbf{u}) \cdot \mathbf{v} \right\}. \tag{5}$$

This is the straightforward generalization of the parameterization proposed by Borrvall and Petersson [BoP03] for the Navier–Stokes equations. As the reader may expect, the design-to-flow mapping induced by the problem (5) demonstrates a behaviour that is very similar to the design-to-flow mapping for Stokes flows, except that it may not be single-valued even for positive designs, and that power dissipation is not a lower-semicontinuous function of the design in this case. Thus, regularization of the topology optimization problems for the Navier–Stokes equations is absolutely necessary.

It turns out that if we employ the idea of *filter* [Sig97, SiP98] (which has become quite a standard technique in topology optimization, see [Bou01, BrT01] for the rigorous mathematical treatment) *in addition* to relaxing the incompressibility constraint (which is unique to the topology optimization of fluids) we can establish the continuity of the resulting design-to-flow mapping, and therefore the existence of optimal designs for a great variety of design functionals [Paper 4, this thesis]. Our use of filters significantly differs from the traditional one in the topology optimization of linearly elastic solids, owing to the dissimilar design parameterizations in these two cases.

Summary of the appended papers

Paper 1: We consider a mechanically tractable technique for obtaining robust optimal truss topologies alternatively to stochastic programming, namely, the introduction of global buckling (linear elastic stability) constraint into the optimization problem. This technique has already been considered and interior point algorithms have been proposed to solve this non-convex optimization problem [Koč02, BTJK$^+$00]. We show that the global buckling constraint produces singular feasible sets, which may prevent convergence of standard numerical methods towards optimal solutions; we also propose a simple resolution strategy based on approximation. This manuscript is to appear in *Structural and Multidisciplinary Optimization*, 2004.

Paper 2: All regularization techniques proposed so far for truss topology optimization address the convergence of globally optimal designs, which may not be realistic for this class of large-scale non-convex optimization problems. In this paper we propose an alternative regularization technique, which guarantees the convergence of stationary points to perturbed problems towards stationary points to the original singular problems. Preliminary results have been presented at the 18th International Symposium on Mathematical Programming, 18–22 August 2003, Copenhagen, Denmark. The manuscript was submitted to *International Journal for Numerical Methods in Engineering* in April 2004.

Paper 3: We show that the minimal power dissipation problem for Stokes problems in porous media proposed in [BoP03] can be extended to include solid impenetrable walls. We demonstrate that, in general, the resulting design-to-flow mapping is not closed, yet under mild assumptions it is possible to approximate globally optimal minimal-power-dissipation domains using porous materials with diminishing permeability. The manuscript was submitted to *Applied Mathematics and Optimization* in August 2003.

Paper 4: We show that a straightforward generalization of the design parameterization proposed in [BoP03] for Navier–Stokes flows results in an ill-posed control problem. We propose a regularization technique based on the relaxation of the flow incompressibility requirement and the introduction of a filter into the design parameterization (the latter now regarded as a standard computational technique in topology optimization of linearly elastic continua, see [Sig97, SiP98, Bou01, BrT01]). The manuscript was submitted to *Zeitschrift für Angewandte Mathematik und Mechanik* in February 2004. A condensed version of the paper is accepted for presentation at the 10th AIAA/ISSMO Multidisciplinary Anlysis and Optimization Conference, August 30–September 1, 2004, Albany, NY, USA, and will appear in its proceedings.

References

[Ach98] W. Achtziger. Multiple-load truss topology and sizing optimization: Some properties of minimax compliance. *J. Optim. Theory Appl.*, 98(2):255–280, 1998.

[BeS03] Martin P. Bendsøe and Ole Sigmund. *Topology Optimization: Theory, Methods, and Applications*. Springer-Verlag, Berlin, 2003.

[BoP03] Thomas Borrvall and Joakim Petersson. Topology optimization of fluids in Stokes flow. *Internat. J. Numer. Methods Fluids*, 41(1):77–107, 2003.

[Bou01] Blaise Bourdin. Filters in topology optimization. *Internat. J. Numer. Methods Engrg.*, 50(9):2143–2158, 2001.

[BrT01] Tyler E. Bruns and Daniel A. Tortorelli. Topology optimization of non-linear elastic structures and compliant mechanisms. *Comput. Methods Appl. Mech. Engrg.*, 190(26–27):3443–3459, 2001.

[BTJK+00] Aharon Ben-Tal, Florian Jarre, Michal Kočvara, Arkadi Nemirovski, and Jochem Zowe. Optimal design of trusses under a nonconvex global buckling constraint. *Optim. Eng.*, 1(2):189–213, 2000.

[ChG97] G. Cheng and X. Guo. ε-relaxed approach in structural topology optimization. *Struct. Optim.*, 13:258–266, 1997.

[ChJ92] G. Cheng and Z. Jiang. Study on topology optimization with stress constraints. *Eng. Opt.*, 20:129–148, 1992.

[DGG64] W.S. Dorn, R.E. Gomory, and H.J. Greenberg. Automatic design of optimal structures. *Journal de Mecanique*, 3:25–52, 1964.

[DuB98] P. Duysinx and M. P. Bendsøe. Topology optimization with stress constraints. *Int. J. Numer. Mech. Engrg.*, 43:1453–1478, 1998.

[DuS98] P. Duysinx and O. Sigmund. New developments in handling stress constraints in optimal material distribution. In *Proc. 7th AIAA/USAF/NASA/ISSMO Symposium on Multidiciplinary Analysis and Optimization*, pages 1501–1509, 1998.

[EPP03] Anton Evgrafov, Michael Patriksson, and Joakim Petersson. Stochastic structural topology optimization: Existence of solutions and sensitivity analyses. *ZAMM Z. angew. Math. Mech.*, 83(7):479–492, 2003.

[EvP03a] Anton Evgrafov and Michael Patriksson. Stable relaxations of stochastic stress constrained weight minimization problems. *Struct. Multidisc. Optim.*, 25(3):189–198, 2003.

[EvP03b] Anton Evgrafov and Michael Patriksson. Stochastic structural topology optimization: Discretization and penalty function approach. *Struct. Multidisc. Optim.*, 25(3):174–188, 2003.

[Fei03] Eduard Feireisl. Shape optimization in viscous compressible fluids. *Appl. Math. Optim.*, 47(1):59–78, 2003.

[FJQ99] Francisco Facchinei, Houyuan Jiang, and Liqun Qi. A smoothing method for mathematical programs with equilibrium constraints. *Math. Program.*, 85(1, Ser. A):107–134, 1999.

[GCY01] X. Guo, G. Cheng, and K. Yamazaki. A new approach for the solution of singular optima in truss topology optimization with stress and local buckling constraints. *Struct. and Multidisc. Optim.*, 22:364–372, 2001.

[GKM00] Max D. Gunzburger, Hongchul Kim, and Sandro Manservisi. On a shape control problem for the stationary Navier-Stokes equations. *M2AN Math. Model. Numer. Anal.*, 34(6):1233–1258, 2000.

[GuK98] Max D. Gunzburger and Hongchul Kim. Existence of an optimal solution of a shape control problem for the stationary Navier-Stokes equations. *SIAM J. Control Optim.*, 36(3):895–909, 1998.

[Gun03] Max D. Gunzburger. *Perspectives in flow control and optimization*. Advances in Design and Control. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2003.

[Kir90] U. Kirsch. On singular topologies in optimum structural design. *Struct. Optim.*, 2(3):133–142, 1990.

[Koč02] M. Kočvara. On the modelling and solving of the truss design problem with global stability constraints. *Struct. Multidisc. Optim.*, 23(3):189–203, 2002.

[KPTK03] Anders Klarbring, Joakim Petersson, Bo Torstenfelt, and Matts Karlsson. Topology optimization of flow networks. *Comput. Methods Appl. Mech. Engrg.*, 192(35-36):3909–3932, 2003.

[LiM97] M. B. Lignola and J. Morgan. Stability of regularized bilevel programming problems. *J. Optim. Theory Appl.*, 93(3):575–596, 1997.

[LPR96] Zhi-Quan Luo, Jong-Shi Pang, and Daniel Ralph. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, Cambridge, 1996.

[Mic04] A. G. M. Michell. The limits of economy of material in frame structures. *Phil. Mag.*, 8, 1904.

[MoP01] B. Mohammadi and O. Pironneau. *Applied shape optimization for fluids*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, 2001.

[PaP02] Michael Patriksson and Joakim Petersson. Existence and continuity of optimal solutions to some structural topology optimization problems including unilateral constraints and stochastic loads. *ZAMM Z. Angew. Math. Mech.*, 82(7):435–459, 2002.

[Pet01] Joakim Petersson. On continuity of the design-to-state mappings for trusses with variable topology. *Internat. J. Engrg. Sci.*, 39(10):1119–1141, 2001.

[RoB94] G. I. N. Rozvany and T. Birker. On singular topologies in exact layout optimization. *Struct. Optim.*, 8:228–235, 1994.

[Sig97] Ole Sigmund. On the design of compliant mechanisms using topology optimization. *Mech. Struct. Mach.*, 25(4):493–524, 1997.

[SiP98] Ole Sigmund and Joakim Petersson. Numerical instabilities in topology optimization: A survey on procedures dealing with checkerboards, mesh-dependencies and local minima. *Struct. Multidisc. Optim.*, 16(1):68–75, 1998.

[StS01] M. Stolpe and K. Svanberg. On trajectories of the epsilon-relaxation approach for stress constrained truss topology optimization. *Struct. and Multidisc. Optim.*, 21(2):140–151, 2001.

[SvG68] G. Sved and Z. Ginos. Structural optimization under multiple loading. *Int. J. Mech. Sci.*, 10:803–805, 1968.

[Ton03] Bui An Ton. Optimal shape control problem for the Navier-Stokes equations. *SIAM J. Control Optim.*, 41(6):1733–1747, 2003.

Paper 1

ON GLOBALLY STABLE SINGULAR TOPOLOGIES

Anton Evgrafov*

Abstract

We consider truss topology optimization problems including a global stability constraint, which guarantees a sufficient elastic stability of the optimal structures. The resulting problem is a nonconvex semi-definite program, for which nonconvex interior point methods are known to show the best performance.

We demonstrate that in the framework of topology optimization, the global stability constraint may behave similarly to stress constraints, that is, that some globally optimal solutions are singular and cannot be approximated from the interior of the design domain. This behaviour, which may be called a *global stability singularity phenomenon*, prevents convergence of interior point methods towards globally optimal solutions. We propose a simple perturbation strategy, which restores the regularity of the design domain. Further, to each perturbed problem interior point methods can be applied.

Key words: Topology optimization – global stability – linear buckling – singularity – semi-definite programming

1.1 Introduction

The optimum design of trusses is concerned with the distribution of the available material among structural members (bars) in order to carry a given set of loads as efficiently as possible, subject to mechanical and technological constraints.

In the framework of topology optimization (as opposed to *sizing* optimization), the topology of a truss may change as a result of the optimization process, that is, if a zero amount of material is allocated to some parts; this possibility significantly enlarges the

*Department of Mathematics, Chalmers University of Technology, SE-412 80 Göteborg, Sweden, email: toxa@math.chalmers.se

design space and, at the same time, increases the computational complexity of the problem. The former gives the possibility to obtain optimal designs that perform much better than their "sizing" counterparts [Mic04]; the latter places significant requirements on the topology optimization algorithms.

In most cases, "standard" nonlinear programming algorithms can be applied directly to sizing optimization problems. Therefore, one natural approach to topology optimization is to introduce a small but positive lower bound ε on the bar volumes, converting the problem into a sizing one [Ach98]. Solving the sequence of sizing problems for ε converging to zero produces a sequence of designs, whose limit points one hopes are optimal in the original topology optimization problem.

Unfortunately, some important constraints produce design domains that violate standard nonlinear programming constraint qualifications; in particular, some optimal solutions cannot be reached by sequences of positive feasible designs. The stress singularity phenomenon appearing in topology optimization problems including constraints on the maximal effective stresses in the structural members is probably the most studied in the literature — we mention the works [SvG68, Kir90, ChJ92, RoB94, ChG97, DuB98, DuS98, Pet01, StS01, PaP02, EPP02, EvP03] just to name a few. Recently, local (Euler) buckling constraints were shown to exhibit an even worse singular behaviour, in the sense that singular optimal solutions become disconnected from the rest of the design region [GCY01].

Despite such an attention, the existing strategies for stress constrained problems may fail to discover global optima even for very small problems [StS01]. On the other hand, "real-world" structures may fail not on account of high stresses, but owing to an insufficient elastic stability [TiG61, Koč02]. Rozvany [Roz96] discusses the elastic instability of the solutions to topology optimization problems with stress and local buckling constraints; Cheng et al. [GCY01, Example 4] provide an example of such a globally unstable structure. (The latter reference concludes with a discussion on the inclusion of global stability constraints.)

Unfortunately, to verify the global stability we need to analyse a static equilibrium path, which is defined by the structure loaded from the rest state with a given load, for possible bifurcation points. Being not an easy task even for a given mechanical structure, it is much more difficult to include the global stability restriction into already complicated structural optimization problems. Using a linear buckling model and semi-definite programming techniques, a mechanically viable yet practically solvable model of global stability has been introduced by Kočvara [Koč02] (see also [BTJK$^+$00]). High-performance interior point algorithms are proposed to solve the global stability side constrained topology optimization problems, which makes it possible to solve high-dimensional practical engineering truss design problems.

However, as we show in this paper, singular optima may also appear in optimization problems with global stability constraints if we consider problems with several loading scenarios, which seems very natural in real-life problems. Therefore, interior point methods, however powerful and modern that they are, applied to such problem instances will produce erroneous results. Using the continuity of design-to-state mappings established by Petersson [Pet01], we show that a simple strategy similar to ε-relaxation method for stress constraints (cf. [ChG97]) can cure the ill-posedness of the feasible region. Further, the interior point method can be applied to the perturbed problems.

1.2 Optimal truss design with a global stability constraint

In this section, we introduce the necessary notation and mechanical principles, and discuss the assumptions on the mechanical structure (utilising a linear buckling model) that naturally lead to the global stability constraint introduced by Kočvara [Koč02] (see also [BTJK$^+$00, SeT00] and references therein). Finally, we state the optimization problem we are going to analyse.

1.2.1 Mechanical equilibrium

Given positions of the nodes, the *design* (and topology in particular) of a truss can be described by prescribing for each bar i, $i = 1,\dots,m$, the amount of material $x_i \geq 0$ allocated to this bar. For convenience we collect all the design variables in a vector $\mathbf{x} = (x_1,\dots,x_m)^t \in \mathbb{R}^m_+$. We introduce an index set of the present (or, active) members in the structure $\mathscr{I}(\mathbf{x}) = \{\, i = 1,\dots,m \mid x_i > 0\,\}$, and denote by $\mathscr{I}^c(\mathbf{x})$ the complement of $\mathscr{I}(\mathbf{x})$ in $\{\,1,\dots,m\,\}$.

For a vector $\mathbf{v} \in \mathbb{R}^n$ and an index set $I = \{\, i_1,\dots,i_{|I|}\,\} \subseteq \{\,1,\dots,n\,\}$, we denote by $\mathbf{v}_I$ the subvector $(v_{i_1},\dots,v_{i_{|I|}})^t$.

Given a particular design $\mathbf{x}$, the equilibrium status of a truss (up to the rigid displacements, which we do not consider) can be described by specifying for each bar $i \in \mathscr{I}(\mathbf{x})$ present in the structure a pseudo-force (also known as the normalised stress) s_i, which is in fact a stress in the bar times its volume. To simplify the notation we collect all values s_i, $i = 1,\dots,m$, into one vector $\mathbf{s}$ of dimension m.

The values of the state variables at equilibrium are determined by the principle of minimum complementary energy; in our case, it is the following quadratic programming problem, parameterised by $\mathbf{x}$:

$$(\mathscr{C})_{\mathbf{x}}(\mathbf{f}) \begin{cases} \min\limits_{\mathbf{s}} \mathscr{E}(\mathbf{x},\mathbf{s}) := \dfrac{1}{2} \sum\limits_{i\in\mathscr{I}(\mathbf{x})} \dfrac{s_i^2}{Ex_i}, \\ \text{s.t.} \begin{cases} \sum\limits_{i\in\mathscr{I}(\mathbf{x})} \mathsf{B}_i^t s_i = \mathbf{f}, \\ \mathbf{s}_{\mathscr{I}^c(\mathbf{x})} = \mathbf{0}, \end{cases} \end{cases}$$

where the data in the problem has the following meaning:
- E is the Young modulus for the structure material;
- $\mathsf{B}_i \in \mathbb{R}^{n\times 1}$ is the kinematic transformation matrix for the bar i;
- $\mathbf{f} \in \mathbb{R}^n$ is the vector of external forces.

We further introduce the vector of nodal displacements $\mathbf{u} \in \mathbb{R}^n$ as a vector of Lagrange multipliers for the force equilibrium constraints in the problem above. Defining the *stiffness matrix* of the structure as

$$\mathsf{K}(\mathbf{x}) := \sum_{i\in\mathscr{I}(\mathbf{x})} x_i \mathsf{K}_i,$$

where $\mathsf{K}_i = E\mathsf{B}_i^t\mathsf{B}_i$ is the *bar stiffness matrix* for the bar i, one can relate the equilibrium displacements directly to the applied force via a system of linear (in $\mathbf{u}$, $\mathbf{f}$) equations: $\mathsf{K}(\mathbf{x})\mathbf{u} = \mathbf{f}$. We, however, avoid this simple, and familiar, formulation, because the matrix

$\mathsf{K}(\mathbf{x})$ is usually not positive definite, unless $\mathbf{x} > \mathbf{0}$ (i.e., when all bars are present in the structure), leading to a non-uniqueness of the equilibrium displacements. On the contrary, the optimal solution to $(\mathscr{C})_{\mathbf{x}}(\mathbf{f})$ is always unique whenever it exists (i.e., if a static equilibrium is possible in the linear model); see [Pet01, Theorem 2.1].

For the rest of the paper, we make the blanket assumption that $\mathsf{K}(\mathbf{x})$ is *positive definite* for every *positive* design $\mathbf{x}$; a necessary and sufficient condition for this property is that $\mathsf{K}(\mathbf{1})$ is positive definite. We do not loose any generality from this assumption, because the positive definiteness can be achieved by starting from an "enough rich" ground structure (see, for example [Ach98, Assumption (A5)]).

1.2.2 Linear buckling model

For the reader's convenience in this subsection we repeat the assumptions on the mechanical structure that lead to the linear buckling model and its representation as a linear matrix inequality; interested reader is referred to Kočvara [Koč02] for more details.

The analysis of the global stability of structural equilibria in its simplest form reduces to the classification of the critical points of a given energy functional as being strict local minimum points (that is, stable points) or not. We also denote by stable points those that are limits of sequences of strict local minima [Koč02].

In the linear model the strain energy in the bar i is related to displacements via

$$W_i^{\text{lin}} = \frac{1}{2}Ex_i(\mathsf{B}_i\mathbf{u})^2, \quad i = 1,\dots,m.$$

The linear strain energy is convex, and, therefore, local maxima or saddle points are impossible in this model, leading to the false conclusion that every equilibrium is stable, if any exists. Therefore, in order to verify the stability of an equilibrium point we must employ a nonlinear model, in which the strain energy takes the form

$$W_i^{\text{nl}} = \frac{1}{2}Ex_i\left[\mathsf{B}_i\mathbf{u} + \frac{1}{2}(\mathsf{B}_i\mathbf{u})^2 + \frac{1}{2}(\mathsf{C}_i\mathbf{u})^2 + \frac{1}{2}(\mathsf{D}_i\mathbf{u})^2\right]^2, \qquad i = 1,\dots,m,$$

where the kinematic transformation matrices $\mathsf{C}_i \in \mathbb{R}^{n\times 1}$ and $\mathsf{D}_i \in \mathbb{R}^{n\times 1}$ account for displacements that are orthogonal to the axial direction of the bar and to each other. (In the two-dimensional model there is of course only one direction that is orthogonal to the axial direction of the bar, whence there will be only one "additional" matrix C_i.)

In order to make the model computationally tractable, yet applicable to a wide class of structures, the following *linear buckling assumptions* are supposed to hold (see [Coo74, BTJK$^+$00, Koč02]):

○ the displacements depend linearly on the load applied for loads less than the critical buckling load;

○ the vector of these linear displacements is orthogonal to the vector of buckling displacements; and

○ the bar axial forces $Ex_i\mathsf{B}_i\mathbf{u}/\ell_i$, $i = 1,\dots,m$ (where ℓ_i is the length of the bar i), remain constant during the deformation caused by buckling.

Under these assumptions, the strain energy can be simplified to the following expression [Koč02]:

$$W_i^{\text{nls}} = \frac{1}{2}\mathbf{u}^t[x_i\mathsf{K}_i + s_i(\mathsf{C}_i^t\mathsf{C}_i + \mathsf{D}_i^t\mathsf{D}_i)]\mathbf{u}, \quad i = 1,\dots,m,$$

where we used the fact that $s_i = Ex_i\mathsf{B}_i\mathbf{u}$ for all bars i. Defining, similarly to the stiffness matrix $\mathsf{K}(\mathbf{x})$, the *geometry matrix*

$$\mathsf{G}(\mathbf{s}) = \sum_{i=1}^m s_i\mathsf{G}_i,$$

where $\mathsf{G}_i = \mathsf{C}_i^t\mathsf{C}_i + \mathsf{D}_i^t\mathsf{D}_i$ is the bar geometry matrix, Kočvara proposes the following stability constraint to be added to the design problem:

$$\mathsf{K}(\mathbf{x}) + \mathsf{G}(\mathbf{s}) \succeq 0.$$

(For two symmetric matrices $\mathsf{M}_1, \mathsf{M}_2 \in \mathbb{R}^{n\times n}$ we write $\mathsf{M}_1 \succeq \mathsf{M}_2$ if and only if the matrix $\mathsf{M}_1 - \mathsf{M}_2$ is positive semi-definite.) The matrix $\mathsf{K}(\mathbf{x}) + \mathsf{G}(\mathbf{s})$ is a Hessian matrix for the simplified nonlinear potential energy functional $\Pi^{\text{nls}}(\mathbf{u}) = \sum_{i=1}^m W_i^{\text{nls}}(\mathbf{u}) - \mathbf{f}^t\mathbf{u}$; it must be positive semi-definite at every local minimum point. We, however, verify this condition at the linear equilibrium state $\mathbf{s}^{\text{lin}}$, motivating it by the fact that, under linear buckling assumptions, $\mathbf{s}^{\text{lin}} \approx \mathbf{s}^{\text{nls}}$ for loads smaller than the buckling load. Arguments for such a stability condition can be found in the papers [BTJK$^+$00, Koč02]; we cite [Koč02, Lemma 1] which asserts that if the stiffness matrix $\mathsf{K}(\mathbf{x})$ is positive definite, then the global stability constraint implies that for any $0 \le \tau < 1$ the load $\tau\mathbf{f}$ is not a classic buckling load for the idealised structure (e.g., see [TiG61]).

1.2.3 Optimization problem

Given N load cases, $\mathbf{f}^1,\dots,\mathbf{f}^N$, we look for a truss of minimal volume that is globally stable as well as stiff w.r.t. each load case. To guarantee the stiffness we require that the inverse quantity $(\mathbf{f}^k)^t\mathbf{u}^k = 2\mathscr{E}(\mathbf{x},\mathbf{s}^k)$, known as the compliance, does not exceed a given amount $c > 0$ for each load case. The problem, which differs from the problem considered by Kočvara [Koč02] only by the fact that we consider several load scenarios, can be formally stated as follows:

$$(\mathscr{P})\begin{cases} \min\limits_{(\mathbf{x},\mathbf{s})} w(\mathbf{x}) := \sum\limits_{i=1}^n x_i, \\ \text{s.t.} \begin{cases} \mathbf{x} \ge \mathbf{0}, \\ \mathscr{E}(\mathbf{x},\mathbf{s}^k) \le 0.5c, \\ \mathsf{K}(\mathbf{x}) + \mathsf{G}(\mathbf{s}^k) \succeq 0, \\ \mathbf{s}^k \text{ solves } (\mathscr{C})_{\mathbf{x}}(\mathbf{f}^k) \end{cases} \quad k = 1,\dots,N, \end{cases}$$

In this form, the problem is an instance of the class of mathematical programs with equilibrium constraints, or bilevel programming problems [LPR96, OKZ98]. For the problems in this class, establishing the existence of solutions is a non-trivial matter; see, for example, [LPR96, Example 1.1.2]. In our case, the equilibrium constraints $(\mathscr{C})_{\mathbf{x}}(\mathbf{f}^k)$ do not, in general, define closed feasible sets in the design space [Pet01, Example 3.1]. However, combined with the energy bound $\mathscr{E}(\mathbf{x},\mathbf{s}^k) \le 0.5c$, the design-to-state mapping $\mathbf{x} \to \mathbf{s}$ is continuous [Pet01, Theorem 3.1], and the corresponding feasible set is closed. Furthermore, the design vector $\mathbf{x} = \alpha\mathbf{1}$ is feasible (stiff and stable) for $\alpha > 0$ large enough. Therefore, the existence of optimal solutions to $(\mathscr{P})$ follows from Weierstrass' theorem.

We also note the presence of the matrix inequality in the formulation of $(\mathscr{P})$; the problem therefore is an instance of semi-definite programming as well. It is customary to solve such problems using interior point techniques; in Section 1.4 we formulate the approximating problems that can be attacked by nonlinear interior point methods. The next section, however, explains why interior point algorithms applied to an equivalent reformulation of the problem $(\mathscr{P})$ as a non-linear semi-definite programming problem, as it has been proposed in [Koč02, BTJK$^+$00], may produce erroneous results.

1.3 A global stability "singularity phenomenon"

Consider the truss structure shown in Fig. 1.1 (a), which consists of 6 bars and has 2 nodes. We consider the two-dimensional case, and thus the structure has $n = 4$ degrees of freedom, which we collect in one vector $\mathbf{u} = (u_x(1), u_y(1), u_x(2), u_y(2))^t$, where $u_v(j)$ is the displacement of the node j along the coordinate axis v. The structure is subject to $N = 2$ load cases: $\mathbf{f}^1 = (0.0, -1.5, 0.0, 0.0)^t$ and $\mathbf{f}^2 = (0.0, 0.0, 0.0, -1.0)^t$. Owing to the vertical symmetry of the ground structure and the load cases, we are interested in symmetric designs only, which allows us to describe the design using only $m = 4$ design variables, collected in one vector $\mathbf{x} = (x_1, x_2, x_3, x_4)^t$ (the correspondence between the design variables and bars is found in Fig. 1.1 (a)). Assuming $E = 1$, the global stiffness and geometry matrices of the structure are, respectively (and approximately)

$$\mathsf{K}(\mathbf{x}) \approx \begin{pmatrix} 7.574\cdot 10^{-3}x_4 & 0.0 & 0.0 & 0.0 \\ 0.0 & x_1 + x_2 + 0.485x_4 & 0.0 & -x_2 \\ 0.0 & 0.0 & 0.1107x_3 & 0.0 \\ 0.0 & -x_2 & 0.0 & x_2 + 1.772x_3 \end{pmatrix},$$

$$\mathsf{G}(\mathbf{s}) \approx \begin{pmatrix} s_1 + s_2 + 0.485s_4 & 0.0 & -s_2 & 0.0 \\ 0.0 & 7.574\cdot 10^{-3}s_4 & 0.0 & 0.0 \\ -s_2 & 0.0 & s_2 + 1.772s_3 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.1107s_3 \end{pmatrix}.$$

Looking at the buckling mode of this structure for the load case 1, which is shown in Fig. 1.1 (b), one can immediately see that the buckling displacements are orthogonal to the linear ones; thus the linear buckling assumptions are most probably verified to some degree of accuracy. To further investigate the linear buckling hypothesis we consider the one-parametric family of loads $\tau\mathbf{f}^1$, $0 \le \tau \le 1$, and plot the graphs of the nodal displacements as functions of τ both for the fully nonlinear strain model and the simplified nonlinear strain model, based on the linear buckling hypothesis. The graphs are shown in Fig. 1.2; we use their similarity as a visual argument in the "proof" of the linear buckling hypothesis. To give some numbers, we note that at $\tau_{\text{buck}} \approx 0.475$ the cosine of the angle between the linear and the buckling displacements is approximately $1.3\cdot 10^{-3}$, and the relative change in the axial forces is 0.85%.

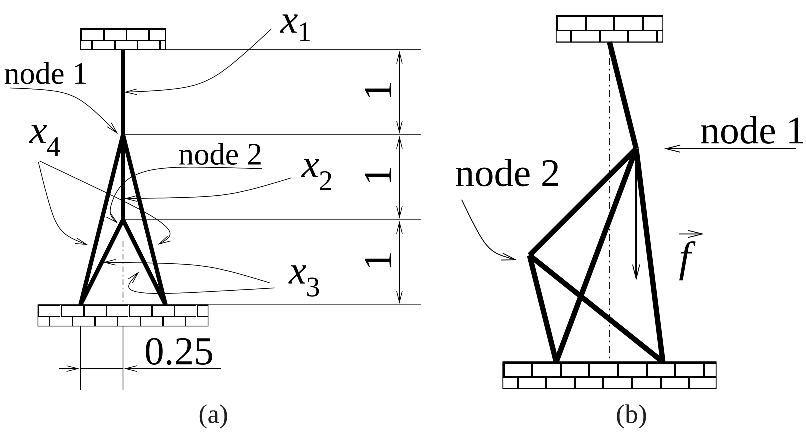
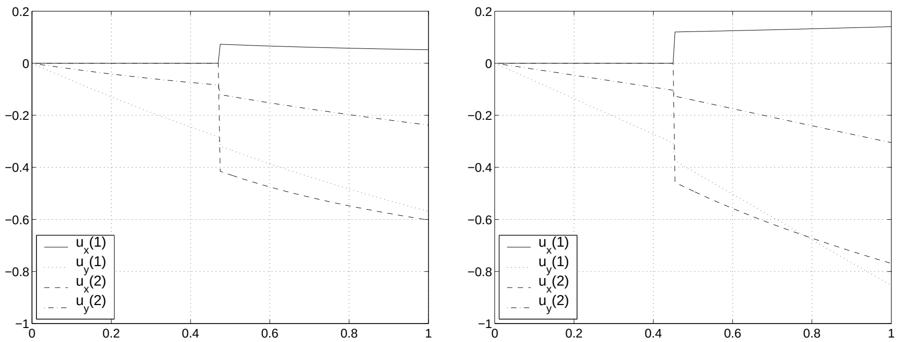


Fig. 1.1: (a) The unstable structure and (b) its buckling mode.

Limiting the compliance to be at most $c = 2$, one can verify analytically that the globally optimal solution to the problem $(\mathscr{P})$ is $\mathbf{x}^* = (1.125, 0.9, 0.0, 0.0)^t$, with optimal weight $w(\mathbf{x}^*) = 2.025$. At the optimal point, not only are the compliance constraints active for both load cases but also the matrix inequality for the first load case; namely, there are three positive eigenvalues of $\mathsf{K}(\mathbf{x}^*) + \mathsf{G}((\mathbf{s}^1)^*)$, where $(\mathbf{s}^1)^*$ is the solution to $(\mathscr{C})_{\mathbf{x}^*}(\mathbf{f}^1)$, and one zero eigenvalue. Solving the state problem $(\mathscr{C})_{\mathbf{x}}(\mathbf{f}^1)$ to obtain a closed form expression for $\mathbf{s}(\mathbf{x})$ and making a first-order Taylor expansion of the smallest eigenvalue of $\mathsf{K}(\mathbf{x}) + \mathsf{G}(\mathbf{s}(\mathbf{x}))$ near $\mathbf{x} = \mathbf{x}^*$ (which, luckily, exists for our example at this point) we get the expression $e(\mathbf{x}) \approx -4.4747x_3 +$ "higher order terms". Thus, for some $\varepsilon > 0$, the globally optimal solution $\mathbf{x}^*$ is separated from every feasible (in particular, stable) design satisfying $\mathbf{x} > \mathbf{0}$, i.e., $\|\mathbf{x} - \mathbf{x}^*\| > \varepsilon$. We see that in the framework of topology optimization the global stability constraint may behave similarly to stress constraints (the singularity phenomenon for the latter was first observed by Sved and Ginos [SvG68]); therefore, we can call this behaviour of the solution to $(\mathscr{P})$ a *global stability singularity phenomenon*.

In Table 1.1, we show the results of the naïve introduction of the positive lower bound $\varepsilon > 0$ into the problem. [We used an SQP-algorithm to solve this non-smooth nonconvex problem in 4 variables, which was first converted to a one-level form by explicitly solving the equilibrium constraint. This is of course unsuitable for realistic problems that will include large semi-definite constraints and therefore require much more cautious algorithmic treatment; we propose one formulation that is suitable for numerical computations in Section 1.4, see problem $(\hat{\mathscr{P}}^\varepsilon)$.] One can clearly see that the sequence of perturbed optimal designs converges to a limit, which is approximately fifteen times heavier than the (unperturbed) optimum! The interior point method applied to an equivalent formulation of this problem instance as a non-linear semi-definite program [that is, to the problem $(\hat{\mathscr{P}}^\varepsilon)$ stated in Section 1.4 for $\varepsilon > 0$, but with the perturbation parameter ε set to zero] would produce similar results, because, as we have shown, it is impossible to approximate the globally optimal solution from the interior of the design domain.

We further note that the semi-definite approximation method proposed by Ben-



(a) (b)

Fig. 1.2: Nodal displacements calculated (a) for fully nonlinear and (b) simplified nonlinear strain models [$\mathbf{x} = (1.125, 0.9, 1, 1)^t$].

ε	$\mathbf{x}^*$	$w(\mathbf{x}^*)$
1	$(26.993, 2.656, 1.0, 1.0)^t$	33.649
10^{-1}	$(30.08, 0.326, 0.1, 0.1)^t$	30.810
10^{-3}	$(30.13, 0.326, 0.1, 1\cdot 10^{-2})^t$	30.654
10^{-7}	$(30.126, 0.326, 0.1, 1\cdot 10^{-7})^t$	30.652
0	$(1.125, 0.9, 0.0, 0.0)^t$	2.025

Tab. 1.1: Results for the naïve ε-perturbation approach.

Tal et al. [BTJK$^+$00] cannot reach the optimal solution either. Indeed, let $\mathscr{Y} \subset \mathbb{R}^m \times \mathbb{R}^k$ be the set formed by the linear matrix inequalities, such that $\mathrm{Proj}_{\mathbb{R}^m}(\mathscr{Y}) = \{\, \mathbf{x} \in \mathbb{R}^m \mid \exists \mathbf{z} \in \mathbb{R}^k : (\mathbf{x}, \mathbf{z}) \in \mathscr{Y} \,\} \subset \mathscr{X}$, where $\mathscr{X}$ is the set of feasible designs for the problem $(\mathscr{P})$. Then, $\mathrm{Proj}_{\mathbb{R}^m}(\mathrm{int}(\mathscr{Y})) \subset \mathrm{int}(\mathscr{X})$, and an interior point method applied to the problem having $\mathscr{Y}$ as a feasible set would not converge towards a point $\mathbf{y}^* = (\mathbf{x}^*, \mathbf{z}^*)$, for some $\mathbf{z}^* \in \mathbb{R}^k$, because $\mathbf{x}^* \notin \mathrm{cl}(\mathrm{int}(\mathscr{X}))$.

Finally, we note that the problems including free vibration constraints (see [Koč02, Section 6]), which is sometimes substituted for the problems with the global stability constraint, are convex and, therefore, do not exhibit any singularities. Thus, our example further illustrates significant differences between the two problems.

1.4 ε-perturbation approach

The instance of the problem $(\mathscr{P})$ given in the previous section clearly demonstrates the need to relax the stability constraint in order to be able to use the interior point machinery. We employ an idea similar to the ε-relaxation of the stress constraints, proposed by Cheng and Guo [ChG97] (see also [Pet01]). Let $o : \mathbb{R}_{++} \to \mathbb{R}_{++}$ be a function such that $\lim_{\varepsilon \to +0} o(\varepsilon)/\varepsilon = 0$. We introduce the positive lower bound $o(\varepsilon)$ on the design variables, thus *restricting* the feasible set, while at the same time *relaxing* the stability constraint by adding a positive definite matrix $\varepsilon\mathsf{I}$:

$$(\mathscr{P}^\varepsilon)\begin{cases}\min\limits_{(\mathbf{x},\mathbf{s})} w(\mathbf{x}) \\ \text{s.t.} \begin{cases}\mathbf{x} \geq o(\varepsilon)\mathbf{1}, \\ \mathscr{E}(\mathbf{x},\mathbf{s}^k) \leq 0.5c, \\ \mathsf{K}(\mathbf{x}) + \mathsf{G}(\mathbf{s}^k) + \varepsilon\mathsf{I} \succeq 0, \\ \mathbf{s}^k \text{ solves } (\mathscr{C})_{\mathbf{x}}(\mathbf{f}^k)\end{cases}\Bigg\} \quad k = 1,\dots,N.\end{cases}$$

Using the locally directionally Lipschitz dependence of the state variables on the design [Pet01, Theorem 3.3], we can easily prove the convergence of a sequence of globally optimal solutions to $\{(\mathscr{P}^\varepsilon)\}$ to globally optimal solutions to $(\mathscr{P})$ as $\varepsilon \to +0$.

Theorem 1.1. *Let $\{\varepsilon_i\}$ be a positive sequence, converging to zero, and further let $\{\mathbf{x}_i\}$ be a sequence of designs that are globally optimal in $\{(\mathscr{P}^{\varepsilon_i})\}$. Then, every limit point $\hat{\mathbf{x}}$ of $\{\mathbf{x}_i\}$ (and there exists at least one) is a design that is globally optimal in $(\mathscr{P})$.*

Proof. As we have already mentioned, for sufficiently large $\alpha > 0$ the design $\alpha\mathbf{1}$ is feasible in $(\mathscr{P})$; thus it is also feasible in $(\mathscr{P}^{\varepsilon_i})$ for all i large enough. Therefore, eventually, the subsequence $\{\mathbf{x}_i\}$ lies in the compact set $\{\, \mathbf{x} \in \mathbb{R}^m_+ \mid w(\mathbf{x}) \leq w(\alpha\mathbf{1})\,\}$, which implies the existence of limit points for the original sequence.

Without any loss of generality, we assume that the original sequence $\{\mathbf{x}_i\}$ converges to $\hat{\mathbf{x}}$. Let $\{\mathbf{s}_i^k\}$, $k = 1,\dots,N$, be the corresponding sequence of state vectors. Owing to the uniform energy bound $\mathscr{E}(\mathbf{x}_i,\mathbf{s}_i^k) \leq 0.5c$, for each $k = 1,\dots,N$, we have that $\{\mathbf{s}_i^k\} \to \hat{\mathbf{s}}^k$, where $\hat{\mathbf{s}}^k$ moreover solves $(\mathscr{C})_{\hat{\mathbf{x}}}(\mathbf{f}^k)$ (cf. [Pet01, Theorem 3.1]). The lower semicontinuity of $\mathscr{E}$ (cf. [Pet01, Lemma 2.1]) together with the continuity of the global stability constraint, implies that $\hat{\mathbf{x}}$ is feasible in $(\mathscr{P})$. Thus, we have proved the inequality

$$\mathrm{val}(\mathscr{P}) \leq w(\hat{\mathbf{x}}) \leq \liminf_{i\to\infty} \mathrm{val}(\mathscr{P}^{\varepsilon_i}).$$

On the other hand, let $\mathbf{x}^*$ be a design that is optimal in $(\mathscr{P})$, and consider a sequence of positive designs $\{\hat{\mathbf{x}}_i\} := \{\mathbf{x}^* + o(\varepsilon_i)\mathbf{1}\}$. Then, owing to [Pet01, Theorem 3.3], there is a constant $C > 0$ such that for the distance between the corresponding state vectors the following inequality holds: $\|\hat{\mathbf{s}}_i^k - (\mathbf{s}^*)^k\| \leq Co(\varepsilon_i)$, $k = 1,\dots,N$. Given the additive structure of the stiffness and geometry matrices this implies

$$\begin{aligned}\mathsf{K}(\hat{\mathbf{x}}_i) + \mathsf{G}(\hat{\mathbf{s}}_i^k) + \varepsilon_i\mathsf{I} \succeq \mathsf{K}(\mathbf{x}^*) &+ \mathsf{G}((\mathbf{s}^*)^k) \\ &+ \mathsf{K}(o(\varepsilon_i)\mathbf{1}) - \mathsf{G}(Co(\varepsilon_i)\mathbf{1}) + \varepsilon_i\mathsf{I} \succeq 0, \\ &\qquad k = 1,\dots,N,\end{aligned}$$

owing to the global stability of $\mathbf{x}^*$ and the properties of $o(\cdot)$. Thus we have proved the reverse inequality:

$$\limsup_{i\to\infty} \mathrm{val}(\mathscr{P}^{\varepsilon_i}) \geq \lim_{i\to\infty} w(\hat{\mathbf{x}}_i) = \mathrm{val}(\mathscr{P}),$$

which concludes the proof. □

Each problem $(\mathscr{P}^\varepsilon)$ is a sizing optimization problem including matrix inequality constraints. In the following proposition, we show that the feasible design set of such problems is regular, as opposed to the singular feasible set of the original problem $(\mathscr{P})$.

ε	$\mathbf{x}^*$	$w(\mathbf{x}^*)$
1	$(1.0, 1.0, 1.0, 1.0)^t$	6
10^{-1}	$(1.08, 0.795, 2.31 \cdot 10^{-2}, 1.0 \cdot 10^{-2})^t$	1.943
10^{-3}	$(1.121, 0.890, 2.24 \cdot 10^{-3}, 1.0 \cdot 10^{-4})^t$	2.015
10^{-4}	$(1.1246, 0.899, 2.235 \cdot 10^{-4}, 1.0 \cdot 10^{-6})^t$	2.024
10^{-5}	$(1.125, 0.9, 2.235 \cdot 10^{-5}, 1.0 \cdot 10^{-8})^t$	2.025
0	$(1.125, 0.9, 0.0, 0.0)^t$	2.025

Tab. 1.2: Results for the ε-perturbation scheme.

Proposition 1.2. *Every design* $\mathbf{x}$ *that is feasible in* $(\mathscr{P}^\varepsilon)$, *can be approximated by a sequence of* strictly *feasible points, that is, for which the inequality constraints are strictly satisfied.*

Proof. The sequence of designs $\{\alpha_k \mathbf{x}\}$, where $\{\alpha_k\} \downarrow 1$, satisfies the requirements of the claim. □

Owing to the inequality $\mathbf{x} \geq o(\varepsilon)\mathbf{1} > \mathbf{0}$, the stiffness matrix $\mathsf{K}(\mathbf{x})$ is positive definite. Therefore, the equation $\mathsf{K}(\mathbf{x})\mathbf{u}^k = \mathbf{f}^k$ is uniquely solvable, and the constraint $2\mathscr{E}(\mathbf{x}, \mathbf{s}^k) = (\mathbf{f}^k)^t \mathbf{u}^k \leq c$ can be equivalently written as the linear matrix inequality constraint after an application of the Schur Complement Theorem (this is a standard technique in semi-definite programming):

$$\hat{\mathsf{K}}^k(\mathbf{x}) := \begin{pmatrix} c & (\mathbf{f}^k)^t \\ \mathbf{f}^k & \mathsf{K}(\mathbf{x}) \end{pmatrix} \succeq 0.$$

Furthermore, denoting the unique equilibrium displacements by $\mathbf{u}^k(\mathbf{x})$, one can write the unique solution to $(\mathscr{C})_{\mathbf{x}}(\mathbf{f}^k)$ as a function of the design by using the following expression: $s_i^k(\mathbf{x}) = E x_i \mathsf{B}_i \mathbf{u}^k(\mathbf{x})$. We further define the matrix $\hat{\mathsf{G}}^k(\mathbf{x}) := \mathsf{G}(\mathbf{s}^k(\mathbf{x}))$ to write the *nested* formulation of the problem $(\mathscr{P}^\varepsilon)$, which includes design variables only and is very similar to the one introduced by Kočvara [Koč02]:

$$(\hat{\mathscr{P}}^\varepsilon) \begin{cases} \min_{\mathbf{x}} w(\mathbf{x}) \\ \text{s.t.} \begin{cases} \mathbf{x} \geq o(\varepsilon)\mathbf{1}, \\ \hat{\mathsf{K}}^k(\mathbf{x}) \succeq 0, \\ \mathsf{K}(\mathbf{x}) + \hat{\mathsf{G}}^k(\mathbf{x}) + \varepsilon \mathsf{I} \succeq 0 \end{cases} \left.\vphantom{\begin{matrix} a \\ b \end{matrix}}\right\} \quad k = 1, \ldots, N. \end{cases}$$

This formulation contains only simple design and matrix inequality constraints. Furthermore, Proposition 1.2 guarantees that every feasible point can be approximated as a sequence of strictly feasible points. Therefore, we can apply a nonlinear interior point method (e.g., see [WSV00, Jar00]) to solve this problem.

One can of course argue that as ε goes to zero, there might be "fewer and fewer" interior points around globally optimal solutions. Implementations of interior point methods, however, take special precautions to the numerical ill-posedness appearing as the iterates approach the boundary (cf. [FGW02] and references therein). Therefore, the numerical problems appearing as "boundary approaches" the current point (i.e., as the perturbation parameter ε decreases) will not prevent convergence of the method.

In Table 1.2, we summarise the results of the ε-perturbation scheme applied to our numerical example. We have chosen $o(\varepsilon) = \varepsilon^2$; an SQP algorithm has been used for the numerical solution. [The comments made about the use of an SQP-solver for solving the problem $(\mathscr{P})$ of course apply to the present situation as well.] One can see that the sequence of perturbed optimal designs converges to the singular global optimum, as Theorem 1.1 predicts.

1.5 Discussion

1.5.1 Stress constraints

In the same way that the stress (and local buckling) constraints alone do not guarantee global stability, globally stable designs might include overstressed bars (cf. [SeT00]) which significantly reduce the life-time of the structure. Therefore, there might be an engineering interest in including stress constraints into the structural problem formulation $(\mathscr{P})$.

Let $\sigma_i > 0$ be the maximal allowable stress in the bar i, $i = 1,\dots,m$. The stress constraints in our notation then take the form $|s_i| \le \sigma_i x_i$, $i = 1,\dots,m$. The easiest way to add stress constraints in our problem is via a penalty function (see [EvP03]):

$$g(\mathbf{x},\mathbf{s}) = \sum_{i\in\mathscr{I}(\mathbf{x})} \frac{[|s_i| - \sigma_i x_i]_+^2}{x_i}.$$

In this way, we only need to change the objective function of the problem $(\hat{\mathscr{P}}^\varepsilon)$ to $w(\mathbf{x}) + \mu(\varepsilon)g(\mathbf{x},\mathbf{s}(\mathbf{x}))$, where $\mu : \mathbb{R}_{++} \to \mathbb{R}_{++}$ is a penalty parameter. The speed at which $\mu(\varepsilon)$ grows must be “synchronised” with the speed at which $o(\varepsilon)$ converges to zero (see [EvP03] for details).

Stress constraints, however, significantly contribute to the nonconvexity of the resulting nested formulation of the problem. Therefore, instead of using the nested formulation together with nonlinear interior point algorithms, one can exploit the convexity of $(\mathbf{x},\mathbf{s}) \to g(\mathbf{x},\mathbf{s})$ as well as the linearity of $(\mathbf{x},\mathbf{s}) \to \mathsf{K}(\mathbf{x}) + \mathsf{G}(\mathbf{x},\mathbf{s})$ by using a semi-definite approximation approach, as proposed by Ben-Tal et al. [BTJK$^+$00] for the original problem $(\mathscr{P})$ without stress constraints. The ε-perturbation of the global stability constraint as well as the treatment of the stress constraints via a penalty function will guarantee the approximability of the globally optimal solutions from the interior of the feasible domain.

1.5.2 Global vs. local optimality

Since each of the problems $(\mathscr{P}^\varepsilon)$ is nonconvex, it is still possible to construct numerical examples demonstrating the non-convergence of the ε-perturbation approach in practice. Such examples are based on the local nature of the nonlinear interior point methods (see [StS01] for examples of non-convergence for stress-constrained problems solved using the ε-relaxation approach of Cheng and Guo [ChG97]); they do not contradict Theorem 1.1, which makes an assertion about the sequence of *global* solutions. Nevertheless, we believe that the results of this paper contribute to the deeper understanding of the problems including a global stability constraint, as well as to the construction of more efficient algorithms for this practically important class of problems.

A convergence analysis of sequences of local minima and stationary points to various sizing approximations of topology optimization problems is one of the topics of our current research.

Acknowledgement: The author is grateful to Michael Patriksson for his careful reading of the manuscript and suggestions that improved the presentation. He also thanks two anonymous reviewers for their valuable comments. This research is supported by the Swedish Research Council (grant 621-2002-5780).

References

[Ach98] W. Achtziger. Multiple-load truss topology and sizing optimization: Some properties of minimax compliance. *J. Optim. Theory Appl.*, 98(2):255–280, 1998.

[BTJK+00] Aharon Ben-Tal, Florian Jarre, Michal Kočvara, Arkadi Nemirovski, and Jochem Zowe. Optimal design of trusses under a nonconvex global buckling constraint. *Optim. Eng.*, 1(2):189–213, 2000.

[ChG97] G. Cheng and X. Guo. ε-relaxed approach in structural topology optimization. *Struct. Optim.*, 13:258–266, 1997.

[ChJ92] G. Cheng and Z. Jiang. Study on topology optimization with stress constraints. *Eng. Opt.*, 20:129–148, 1992.

[Coo74] R. D. Cook. *Concepts and Applications in Finite Element Analysis.* J. Wiley & Sons: New York, 1974.

[DuB98] P. Duysinx and M. P. Bendsøe. Topology optimization with stress constraints. *Int. J. Numer. Mech. Engrg.*, 43:1453–1478, 1998.

[DuS98] P. Duysinx and O. Sigmund. New developments in handling stress constraints in optimal material distribution. In *Proc. 7th AIAA/USAF/NASA/ISSMO Symposium on Multidiciplinary Analysis and Optimization*, pages 1501–1509, 1998.

[EPP02] Anton Evgrafov, Michael Patriksson, and Joakim Petersson. On stochastic structural topology optimizaion. In C. C. Baniotopoulos, editor, *Nonsmooth/nonconvex mechanics with applications in engineering*, pages 97–104, Department of Civil Engineering, Faculty of Engineering, Aristotle University of Thessaloniki, Thessaloniki, Greece, july 2002.

[EvP03] Anton Evgrafov and Michael Patriksson. Stochastic structural topology optimization: Discretization and penalty function approach. *Struct. Multidisc. Optim.*, 25(3):174–188, 2003.

[FGW02] Anders Forsgren, Philip E. Gill, and Margaret H. Wright. Interior methods for nonlinear optimization. *SIAM Review*, 44(4):525–597, 2002.

[GCY01] X. Guo, G. Cheng, and K. Yamazaki. A new approach for the solution of singular optima in truss topology optimization with stress and local buckling constraints. *Struct. and Multidisc. Optim.*, 22:364–372, 2001.

[Jar00] Florian Jarre. An interior method for nonconvex semidefinite programs. *Optim. Eng.*, 1(4):347–372, 2000.

[Kir90] U. Kirsch. On singular topologies in optimum structural design. *Struct. Optim.*, 2(3):133–142, 1990.

[Koč02] M. Kočvara. On the modelling and solving of the truss design problem with global stability constraints. *Struct. Multidisc. Optim.*, 23(3):189–203, 2002.

[LPR96] Zhi-Quan Luo, Jong-Shi Pang, and Daniel Ralph. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, Cambridge, 1996.

[Mic04] A. G. M. Michell. The limits of economy of material in frame structures. *Phil. Mag.*, 8, 1904.

[OKZ98] Jiři Outrata, Michal Kočvara, and Jochem Zowe. *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*. Kluwer Academic Publishers, Dordrecht, 1998.

[PaP02] Michael Patriksson and Joakim Petersson. Existence and continuity of optimal solutions to some structural topology optimization problems including unilateral constraints and stochastic loads. *ZAMM Z. Angew. Math. Mech.*, 82(7):435–459, 2002.

[Pet01] Joakim Petersson. On continuity of the design-to-state mappings for trusses with variable topology. *Internat. J. Engrg. Sci.*, 39(10):1119–1141, 2001.

[RoB94] G. I. N. Rozvany and T. Birker. On singular topologies in exact layout optimization. *Struct. Optim.*, 8:228–235, 1994.

[Roz96] G. I. N. Rozvany. Difficulties in truss topology optimization with stress, local buckling and system stability constraints. *Struct. and Multidisc. Optim.*, 11:213–217, 1996.

[SeT00] R. Sedaghati and B. Tabarrok. Optimum design of truss structures undergoing large deflections subject to a system stability constraint. *Int. J. Numer. Meth. Engng*, 48:421–434, 2000.

[StS01] M. Stolpe and K. Svanberg. On trajectories of the epsilon-relaxation approach for stress constrained truss topology optimization. *Struct. and Multidisc. Optim.*, 21(2):140–151, 2001.

[SvG68] G. Sved and Z. Ginos. Structural optimization under multiple loading. *Int. J. Mech. Sci.*, 10:803–805, 1968.

[TiG61] S. P. Timoshenko and J. M. Gere. *Theory of elastic stability*. New York: McGraw-Hill, 1961.

[WSV00] Henry Wolkowicz, Romesh Saigal, and Lieven Vandenberghe, editors. *Handbook of semidefinite programming*, volume 27 of *International Series in Operations Research & Management Science*. Kluwer Academic Publishers, Boston, MA, 2000.

Paper 2

ON THE CONVERGENCE OF STATIONARY SEQUENCES IN TOPOLOGY OPTIMIZATION

Anton Evgrafov* and Michael Patriksson†

Abstract

We consider structural topology optimization problems including unilateral constraints arising from non-penetration conditions in contact mechanics. The resulting non-convex non-smooth problems are instances of mathematical programs with equilibrium constraints (MPEC), or bi-level programs. Applying nested (implicit programming) algorithms to this class of problems is problematic owing to the singularity of the feasible set. We propose a perturbation strategy combining the relaxation of the equilibrium constraint with the restriction of the design domain to its regular part only. This strategy allows us to attack the problem numerically using standard nonlinear programming algorithms.

We rigorously study the optimality conditions for the original singular problem as well as the convergence of stationary points and globally optimal solutions to approximating problems towards respectively stationary points and globally optimal solutions to the original problem. A limited numerical benchmarking of the algorithm is performed.

Keywords: topology optimization, ε-perturbation, local optimality, stress singularity, MPEC, smoothing

2.1 Introduction

THE optimum design of trusses is concerned with the distribution of the available material among structural members (bars) in order to carry a given set of loads as efficiently

*Department of Mathematics, Chalmers University of Technology, SE-412 96 Göteborg, Sweden, e-mail: toxa@math.chalmers.se

†Department of Mathematics, Chalmers University of Technology, SE-412 96 Göteborg, Sweden, e-mail: mipat@math.chalmers.se

as possible, subject to mechanical and technological constraints. The distinguishing feature of structural optimization problems is the presence of the complicating *equilibrium* constraint, relating *design* variables (i.e., those controlling the material distribution) with *state* variables (e.g., nodal displacements and stresses in the structural members). Verbally, the relation between the two sets of variables can be formulated as follows: the state variables solve a parametric optimization problem with design variables as parameters. Therefore, the problem belongs to a class of difficult optimization problems known as mathematical programs with equilibrium constraints (MPEC), or generalized bi-level programming problems.

In the framework of topology optimization (as opposed to *sizing* optimization), the topology of a truss may change as a result of the optimization process, that is, if a zero amount of material is allocated to some parts; this possibility significantly enlarges the design space and, at the same time, increases the computational complexity of the problem. The former implies the possibility to obtain optimal designs that perform much better than their “sizing” counterparts [Mic04]; the latter places significant requirements on algorithms for solving topology optimization problems.

In most cases, “standard” algorithms for differentiable nonlinear programming problems can be applied directly to sizing optimization problems. Therefore, one natural approach to topology optimization is to introduce a small but positive lower bound ε on the bar volumes, thus converting the problem into a sizing one. Solving a sequence of sizing problems for ε converging to zero produces a sequence of designs, whose limit points one hopes are optimal in the original topology optimization problem.

Unfortunately, some important constraints produce design domains that violate standard nonlinear programming constraint qualifications; in particular, some optimal solutions cannot be reached as limits of any sequence of strictly positive feasible designs. The stress *singularity* phenomenon appearing in topology optimization problems with constraints on the maximal effective stresses in the structural members is probably the one most studied and the one that has attracted the most recent interest—we mention the work in [SvG68, Kir90, ChJ92, RoB94, ChG97, DuB98, DuS98, Pet01, StS01, PaP02, EPP03, EvP03], just to name a few references. Similarly, local (Euler) buckling constraints [GCY01], and global (system) stability constraints [Evg04] are known to exhibit a singular behaviour.

Sizing approximations, studied in the cited papers, are all concerned with approximations of the *globally* optimal solutions. In computational practice, however, it is impossible to solve the non-convex approximating problems to global optimality. Since most numerical nonlinear optimization algorithms can only find *stationary points* of the approximating sizing problems, in this paper we study the limit points of such sequences. We show that they are indeed stationary (in some sense) in the limiting (that is, original) topology optimization problem as well.

2.1.1 Equilibrium problem

We consider a truss with m bars and n degrees of freedom. There are r designated nodes of the truss that may come into frictionless unilateral contact with rigid obstacles.

Given positions of the nodes, the *design* (and topology in particular) of a truss can be described by prescribing for each bar i, $i = 1,\dots,m$, the amount of material $x_i \geq 0$

allocated to this bar. For convenience we collect all the design variables in a vector $\mathbf{x} = (x_1,\dots,x_m)^t \in \mathbb{R}^m_+$. We introduce an index set of the present (or, active) members in the structure $\mathscr{I}(\mathbf{x}) = \{\, i = 1,\dots,m \mid x_i > 0\,\}$, and denote by $\mathscr{I}^c(\mathbf{x})$ the complement of $\mathscr{I}(\mathbf{x})$ in $\{\,1,\dots,m\,\}$.

Given a particular design $\mathbf{x}$, the equilibrium status of a truss can be described by specifying

- a pseudo-force s_i (also known as the normalized stress, which is in fact a stress in the bar times its volume) for each bar $i \in \mathscr{I}(\mathbf{x})$ present in the structure . To simplify the notation we collect all values s_i, $i = 1,\dots,m$, into one vector $\mathbf{s} \in \mathbb{R}^m$, assuming $s_i = 0$ for $i \notin \mathscr{I}(\mathbf{x})$;
- a contact force λ_j for each of the potential contact nodes $j = 1,\dots,r$. These values are collected in a vector $\boldsymbol{\lambda} \in \mathbb{R}^r_+$; and
- a displacement u_k for each of the structural degrees of freedom $k = 1,\dots,n$. These values are collected in a vector $\mathbf{u} \in \mathbb{R}^n$.

The triple $(\mathbf{s},\boldsymbol{\lambda},\mathbf{u})$ will be referred to as *state* variables.

For a vector $\mathbf{v} \in \mathbb{R}^q$, and an index set $I = \{\, i_1,\dots,i_{|I|}\,\} \subseteq \{\,1,\dots,q\,\}$, we denote by $\mathbf{v}_I$ the subvector $(v_{i_1},\dots,v_{i_{|I|}})^t$.

The values of the state variables for a specific design $\mathbf{x}$ are determined using various energy principles. Therefore, we define the complementary energy of the structure as

$$\mathscr{E}(\mathbf{x},\mathbf{s},\boldsymbol{\lambda}) := \frac{1}{2}\sum_{i\in\mathscr{I}(\mathbf{x})}\frac{s_i^2}{Ex_i} + \mathbf{g}^t\boldsymbol{\lambda},$$

as well as the linearized strain energy:

$$\Pi(\mathbf{x},\mathbf{u}) := \frac{1}{2}\mathbf{u}^t\mathsf{K}(\mathbf{x})\mathbf{u},$$

where $\mathsf{K}(\mathbf{x})$ is *stiffness matrix* of the structure. The latter matrix is defined as

$$\mathsf{K}(\mathbf{x}) := \sum_{i\in\mathscr{I}(\mathbf{x})} x_i\mathsf{K}_i,$$

where $\mathsf{K}_i = E\mathsf{B}_i^t\mathsf{B}_i$ is the *local stiffness matrix* for the bar $i = 1,\dots,m$, $\mathsf{B}_i \in \mathbb{R}^{1\times n}$ is a kinematic transformation matrix for the bar $i = 1,\dots,m$, and E is the Young modulus of the structural material.

For the rest of the paper, we make the blanket assumption that $\mathsf{K}(\mathbf{x})$ is *positive definite* for every *positive* design $\mathbf{x}$; a necessary and sufficient condition for this property is that $\mathsf{K}(\mathbf{1}^m)$ is positive definite. We do not loose any generality from this assumption, because the positive definiteness can be achieved by starting from an “enough rich” ground structure.

In these notations the equilibrium state of the structure under the external load $\mathbf{f} \in \mathbb{R}^n$ can be characterized using a primal-dual pair of convex quadratic programming problems:

$$(\mathscr{C})_{\mathbf{x}}(\mathbf{f}) \begin{cases} \min\limits_{(\mathbf{s},\boldsymbol{\lambda})} \mathscr{E}(\mathbf{x},\mathbf{s},\boldsymbol{\lambda}), \\ \text{s.t. } \mathsf{C}^t\boldsymbol{\lambda} + \sum\limits_{i\in\mathscr{I}(\mathbf{x})} \mathsf{B}_i^t s_i = \mathbf{f}, \\ \qquad\qquad \boldsymbol{\lambda} \geq \mathbf{0}, \end{cases} \qquad (\mathscr{P})_{\mathbf{x}}(\mathbf{f}) \begin{cases} \min\limits_{\mathbf{u}} \Pi(\mathbf{x},\mathbf{u}) - \mathbf{f}^t\mathbf{u}, \\ \text{s.t. } \mathsf{C}\mathbf{u} \leq \mathbf{g}, \end{cases}$$

where $\mathbf{g} \in \mathbb{R}^r$ is a vector of gaps between the contact nodes and rigid obstacles, and $\mathsf{C} \in \mathbb{R}^{r\times n}$ is a kinematic transformation matrix. We have implicitly assumed that the matrix C is *quasi-orthogonal*, that is, that $\mathsf{C}\mathsf{C}^t = \mathsf{I}$. The problem $(\mathscr{C})_{\mathbf{x}}(\mathbf{f})$ is known as the principle of minimum complementary energy, and the problem $(\mathscr{P})_{\mathbf{x}}(\mathbf{f})$ is the principle of minimum potential energy.

Equivalently, the equilibrium problem can be written as a KKT system for the pair $(\mathscr{C})_{\mathbf{x}}(\mathbf{f})$ and $(\mathscr{P})_{\mathbf{x}}(\mathbf{f})$. Define

$$\mathsf{Q}(\mathbf{x}) := \begin{pmatrix} \mathsf{B}^t & \mathsf{C}^t & 0 \\ 0 & 0 & -\mathsf{C} \\ \mathsf{I} & 0 & -\mathsf{D}(\mathbf{x})\mathsf{B} \end{pmatrix}, \qquad \mathbf{q}(\mathbf{f}) := \begin{pmatrix} -\mathbf{f} \\ \mathbf{g} \\ \mathbf{0} \end{pmatrix},$$

and $Y := \mathbb{R}^m \times \mathbb{R}^r_+ \times \mathbb{R}^n$, where $\mathsf{B} \in \mathbb{R}^{m\times n}$ is the matrix with rows $\mathsf{B}_1, \dots, \mathsf{B}_m$, and $\mathsf{D}(\mathbf{x}) = \operatorname{diag}(\mathbf{x}) \in \mathbb{R}^{m\times m}$. Then, the pair $(\mathbf{s}, \boldsymbol{\lambda})$ solves $(\mathscr{C})_{\mathbf{x}}(\mathbf{f})$ and $\mathbf{u}$ solves $(\mathscr{P})_{\mathbf{x}}(\mathbf{f})$ if and only if the vector $\mathbf{y}^* = (\mathbf{s}^t, \boldsymbol{\lambda}^t, \mathbf{u}^t)^t \in Y$ solves the affine variational inequality problem $\mathrm{AVI}(\mathbf{q}(\mathbf{f}), \mathsf{Q}(\mathbf{x}), Y)$ [see, e.g., [FaP03] for the definition]:

$$[\mathsf{Q}(\mathbf{x})\mathbf{y}^* + \mathbf{q}(\mathbf{f})]^t(\mathbf{y} - \mathbf{y}^*) \geq 0, \quad \text{for all } \mathbf{y} \in Y.$$

Either of the equilibrium problem formulations $(\mathscr{C})_{\mathbf{x}}(\mathbf{f})$, $(\mathscr{P})_{\mathbf{x}}(\mathbf{f})$, or $\mathrm{AVI}(\mathbf{q}(\mathbf{f}), \mathsf{Q}(\mathbf{x}), Y)$ has its advantages and disadvantages. For example, the problem $(\mathscr{C})_{\mathbf{x}}(\mathbf{f})$ possesses at most one optimal solution for every design $\mathbf{x} \in \mathbb{R}^m_+$; at the same time, the objective function $\mathscr{E}$ is only lower semicontinuous (and may be infinite) for some $\mathbf{x} \in \partial\mathbb{R}^m_+$. Both problems have been studied by Patriksson and Petersson [PaP02], and we summarize some of their results below.

Proposition 2.1. (i) *The multi-mapping* $\mathbf{x} \rightrightarrows \operatorname{argmin}(\mathscr{C})_{\mathbf{x}}(\mathbf{f})$ *is at most single-valued for every* $\mathbf{x} \in \mathbb{R}^m_+$, *and is single-valued for every* $\mathbf{x} > \mathbf{0}^m$. *Moreover, this mapping is locally directionally Lipschitz for every* $\mathbf{x} \in \mathbb{R}^m_+$.

Furthermore, the mapping $\mathbf{x} \rightrightarrows \operatorname{argmin}(\mathscr{C})_{\mathbf{x}}(\mathbf{f}) \cap \{ (\mathbf{s}, \boldsymbol{\lambda}) \in \mathbb{R}^m \times \mathbb{R}^r_+ \mid \mathscr{E}(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}) \leq C \}$ *is continuous for every constant* $C > 0$.
(ii) *The multi-mapping* $\mathbf{x} \rightrightarrows \operatorname{argmin}(\mathscr{P})_{\mathbf{x}}(\mathbf{f})$ *is single-valued for every* $\mathbf{x} > \mathbf{0}^m$. *Moreover, this mapping is locally directionally* (*upper*) *Lipschitz for every* $\mathbf{x} \in \mathbb{R}^m_+$.
(iii) *The multi-mapping* $\mathbf{x} \rightrightarrows \mathrm{SOL}(\mathbf{q}(\mathbf{f}), \mathsf{Q}(\mathbf{x}), Y)$ *is closed and locally directionally* (*upper*) *Lipschitz for every* $\mathbf{x} \in \mathbb{R}^m_+$.

Our ultimate goal in this paper is to establish stationary conditions that must be verified by limit points of certain sequences of positive designs. We cannot use the equilibrium formulation given by the problem $(\mathscr{C})_{\mathbf{x}}(\mathbf{f})$ for this purpose, because its objective violates such a basic condition for sensitivity analysis as continuity. Neither the problem $(\mathscr{P})_{\mathbf{x}}(\mathbf{f})$ is suitable for us, because the design-to-state mapping it induces is not closed. Therefore, we will use the primal-dual characterization of the equilibrium given by $\mathrm{AVI}(\mathbf{q}(\mathbf{f}), \mathsf{Q}(\mathbf{x}), Y)$ in the sequel.

We close the subsection by defining the feasible set generated by the equilibrium constraint:

$$\mathscr{F}(\mathbf{f}) := \{ (\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}, \mathbf{u}) \subset \mathbb{R}^m_+ \times \mathbb{R}^m \times \mathbb{R}^r_+ \times \mathbb{R}^n \mid (\mathbf{s}, \boldsymbol{\lambda}, \mathbf{u}) \in \mathrm{SOL}(\mathbf{q}(\mathbf{f}), \mathsf{Q}(\mathbf{x}), Y) \}. \quad (2.1)$$

2.1.2 Weight minimization problem

We use a stress constrained weight minimization problem of a truss subject to unilateral frictionless contact with some rigid obstacles as a representative of the difficult structural optimization problems. To skip one index and simplify the notation we consider a single load case only; this does not affect the applicability of our results to multiple load cases in any way.

The weight minimization problem can be written as follows:

$$(\mathscr{W})\begin{cases} \displaystyle\min_{(\mathbf{x},\mathbf{s},\boldsymbol{\lambda},\mathbf{u})} w(\mathbf{x}) := \sum_{i=1}^{m} x_i, \\ \text{s.t. } (\mathbf{x},\mathbf{s},\boldsymbol{\lambda},\mathbf{u}) \in \mathscr{F}(\mathbf{f}), \\ \qquad \underline{\sigma}_i x_i \le s_i \le \overline{\sigma}_i x_i, \quad i = 1,\dots,m, \end{cases}$$

where $\underline{\sigma}_i \le 0$ and $\overline{\sigma}_i \ge 0$ are the stress bounds in compression and tension for the bar $i = 1,\dots,m$, and $\mathscr{F}(\mathbf{f})$ is given by (2.1).

The results of the present paper are of course applicable to a wider class of problems than $(\mathscr{W})$. For example, more general objective functions can be considered as long as they are reasonably regular [differentiable, or Lipschitz continuous w.r.t. $(\mathbf{x},\mathbf{s},\boldsymbol{\lambda},\mathbf{u})$]; additional constraints may be considered [such as bounds on admissible displacements, local buckling constraints, or global stability constraints]. However, to keep the notation simple we do not discuss such straightforward generalizations in detail.

2.2 Previous results

2.2.1 ε-perturbation of Cheng and Guo and variations

The so-called ε-perturbation of structural topology optimization problems, or approximation with a sequence of *sizing* optimization problems, has become a classic topic. Convergence results of this type allow one, at least in principle, to compute optimal solutions to structural topology optimization problems by solving a sequence of smooth non-convex approximating problems. Such approximations do not suffer from many numerical difficulties possessed by the original model problem $(\mathscr{W})$, so that efficient solvers are readily available.

For some truss topology optimization problems (such as, e.g., compliance minimization, possibly with so-called "strong" stress constraints [Ach98]) the naïve replacement of the lower bound 0 on design variables with a small positive value $\varepsilon > 0$ tending to zero (whence the name—ε-perturbation) is sufficient. Such a strategy has been rigorously studied for trusses, without (Achtziger [Ach98]) and with (Patriksson and Petersson [PaP02]) unilateral constraints.

On the other hand, there are many other classes of topology optimization problems including important mechanical constraints (e.g., stress constraints [SvG68], local buckling constraints [GCY01], and global buckling constraints [Evg04]) where the simple strategy outlined above leads to erroneous results, owing to the complicated singular structure of the design domain near the points where the truss topology changes. Historically, the study of singularity phenomena for truss topology optimization problems started with problems including stress constraints only. Sved and Ginos [SvG68] observed that such problems may have singular solutions, and the properties of the feasible region were further investigated by Kirsch [Kir90], Cheng and Jiang [ChJ92], and Rozvany and Birker [RoB94]. Cheng and Guo [ChG97] were the first to propose a more sophisticated restriction-relaxation procedure, where not only lower bounds but also stress constraints were perturbed. They established the convergence of the optimal values of the perturbed problems to the optimal value of the original problem, while Petersson [Pet01] (using the continuity of certain design–to–state parameterized mappings) has established the convergence of optimal solutions. Since then, the ε-perturbation has been extended by many authors in many ways: Duysinx and Bendsøe [DuB98] and Duysinx and Sigmund [DuS98] considered continuum structures; Guo et al. [GCY01] included local buckling constraints into the problem; Patriksson and Petersson [PaP02] generalized the result for trusses including unilateral constraints; Evgrafov et al. [EPP03] considered the possibility of stochastic forces; and Evgrafov [Evg04] studied the linearized elastic stability constraint.

Despite the clear advantage of approximating the nonsmooth, singular optimization problem with a sequence of smooth and regular ones, all the sizing approximations considered above suffer from the same difficulty. While the underlying theoretical results are concerned with the approximation of the *globally* optimal solutions, in computational practice it is impossible to solve the non-convex approximating problems to global optimality. There are also negative results regarding this issue: the ε-perturbation approach may fail to find a globally optimal solution even for topology optimization problems with only 2 design variables (see [StS01])!

The analysis of the convergence of stationary points to the approximating problems towards stationary points of the limiting (that is, original) problem is difficult; for example, the dependence of the equilibrium state of the structure upon the design near the points where the topology changes is nonsmooth, and even non-Lipschitz continuous.

In constructing a new ε-perturbation we try to address these above issues, concentrating on the convergence of *both* globally optimal solutions and stationary points towards the respective limits.

2.2.2 The extended formulation of Stolpe and Svanberg

Recently, Stolpe and Svanberg [StS03] proposed an alternative method for the solution of the truss topology optimization problems including stress and local buckling constraints, which is based on the Karush–Kuhn–Tucker (KKT) formulation of the equilibrium constraint. In this formulation the state variables are treated equally to the design variables, and artificial lower bounds on the design are unnecessary. In the absence of unilateral constraints, the formulation is suitable for any SQP algorithm, and for some numerical examples Stolpe and Svanberg report that such an algorithm has a better performance than an ε-perturbation based approach. Later, a branch-and-cut algorithm based on this formulation has been developed [Sto03, Paper D]; furthermore, Achtziger [Ach03] has made the conjecture that every globally optimal solution to a topology optimization problem including stress and local buckling constraints (but not including unilateral constraints) is a KKT point in the extended formulation.

Unfortunately, the KKT formulation of the lower level equilibrium problem for trusses with unilateral constraints includes complementarity conditions, which are known to violate standard nonlinear programming constraint qualifications. Therefore, the extended formulation cannot be used directly to solve topology optimization problems for trusses in contact with rigid obstacles, or including tensile-only members (ropes or cables).

We therefore propose a new approximation scheme, which allows for the violation of the lower-level equilibrium conditions, and thus does not include the complicating complementarity constraints.

2.3 A smoothing method for a general MPEC problem

Among iterative algorithms for MPEC problems, our ε-perturbation method is special in that it combines relaxation (of the equilibrium conditions) and restrification (of the design space). Most iterative algorithms for general MPEC problems belong to the relaxation category, wherein constraints are penalized or complementarity conditions are smoothed. In the latter category, the method of Facchinei et al. [FJQ99] has relations to ours that are interesting to explore, in order to analyze the strength of our convergence results. Due to the stronger regularity properties of the problem considered in [FJQ99], their convergence results are shown to be stronger; we then seek to explain why local, iterative methods for our problem are unlikely to yield better convergence characteristics than those that we reach in this paper.

2.3.1 The problem

Consider the problem to

$$(\mathscr{M})\quad \begin{cases} \min\limits_{(\mathbf{x},\mathbf{y})} f(\mathbf{x},\mathbf{y}), \\ \text{s.t.} \begin{cases} \mathbf{x} \in X, \\ \mathbf{y}\,\text{solves}\,\mathrm{VI}(\mathbf{F}(\mathbf{x},\cdot),C(\mathbf{x})), \end{cases} \end{cases}$$

where $f : \mathbb{R}^{n+m} \mapsto \mathbb{R}$ is continuously differentiable, $X \subset \mathbb{R}^n$ is nonempty and compact, and, for each $\mathbf{x} \in X$ and for a continuously differentiable function $\mathbf{F} : \mathbb{R}^{n+m} \mapsto \mathbb{R}^m$, $\mathrm{VI}(\mathbf{F}(\mathbf{x},\cdot),C(\mathbf{x}))$ denotes the variational inequality defined by the pair $(\mathbf{F}(\mathbf{x},\cdot),C(\mathbf{x}))$,

$$\mathbf{y} \in C(\mathbf{x}); \qquad \mathbf{F}(\mathbf{x},\mathbf{y})^t(\mathbf{z}-\mathbf{y}) \geq 0, \quad \mathbf{z} \in C(\mathbf{x}),$$

where

$$C(\mathbf{x}) = \{\, \mathbf{y} \in \mathbb{R}^m \mid g_i(\mathbf{x},\mathbf{y}) \geq 0,\ i = 1,\ldots,\ell \,\},$$

$\mathbf{g} : \mathbb{R}^{n+m} \mapsto \mathbb{R}^\ell$ being twice continuously differentiable and concave in the second argument.

For the lower-level VI, we assume that $C(\mathbf{x}) \neq \emptyset$ for all $\mathbf{x}$ in an open set A containing X, that $C(\mathbf{x})$ is uniformly compact on A (with $C(\mathbf{x}) \subset B$ for some open bounded set $B \subset \mathbb{R}^m$), that $\mathbf{F}(\mathbf{x},\cdot)$ is uniformly strongly monotone on B for all $\mathbf{x} \in A$, and that for every pair $(\mathbf{x},\mathbf{y})$ for which $\mathbf{x} \in X$ and $\mathbf{y}$ solves $\mathrm{VI}(\mathbf{F}(\mathbf{x},\cdot),C(\mathbf{x}))$, the partial gradients $\nabla_{\mathbf{y}}\, g_i(\mathbf{x},\mathbf{y})$,

$i \in \mathscr{I}(\mathbf{x},\mathbf{y}) := \{\, i = 1,\dots,\ell \mid g_i(\mathbf{x},\mathbf{y}) = 0 \,\}$, are linearly independent (that is, the linear independence CQ, LICQ).

By these assumptions, it is clear that $\mathrm{VI}(\mathbf{F}(\mathbf{x},\cdot), C(\mathbf{x}))$ has a unique solution for each $\mathbf{x} \in X$. By the concavity of $\mathbf{g}$ and the LICQ, each lower-level VI is equivalent to the existence of a (unique) multiplier vector $\boldsymbol{\lambda} \in \mathbb{R}^\ell$ such that

$$\mathbf{F}(\mathbf{x},\mathbf{y}) - \nabla_{\mathbf{y}} \mathbf{g}(\mathbf{x},\mathbf{y})\boldsymbol{\lambda} = \mathbf{0}^m, \tag{2.2a}$$

$$\mathbf{0}^\ell \leq \mathbf{g}(\mathbf{x},\mathbf{y}) \perp \boldsymbol{\lambda} \geq \mathbf{0}^\ell \tag{2.2b}$$

holds. (That these KKT conditions are necessary follows by the LICQ; sufficiency follows from concavity.) We can therefore replace the lower-level problem $\mathrm{VI}(\mathbf{F}(\mathbf{x},\cdot), C(\mathbf{x}))$ in the problem $(\mathscr{M})$ by (2.2). This non-smooth reformulation has been utilized in the development of iterative algorithms. Since it, however, does not satisfy any CQ, due to the presence of the complementarity conditions, it is tempting to consider perturbations of the KKT system. Let

$$\mathbf{H}_0(\mathbf{x},\mathbf{y},\mathbf{z},\boldsymbol{\lambda}) := \begin{pmatrix} \mathbf{F}(\mathbf{x},\mathbf{y}) - \nabla_{\mathbf{y}} \mathbf{g}(\mathbf{x},\mathbf{y}) \\ \mathbf{g}(\mathbf{x},\mathbf{y}) - \mathbf{z} \\ -2\min(\mathbf{z},\boldsymbol{\lambda}) \end{pmatrix}, \qquad (\mathbf{x},\mathbf{y},\mathbf{z},\boldsymbol{\lambda}) \in \mathbb{R}^{n+m+2\ell}.$$

The KKT system (2.2) is equivalent to the statement that $\mathbf{H}_0(\mathbf{x},\mathbf{y},\mathbf{z},\boldsymbol{\lambda})\mathbf{0}^{m+2\ell}$. We therefore write

$$(\mathscr{P}) \begin{cases} \min\limits_{(\mathbf{x},\mathbf{y})} f(\mathbf{x},\mathbf{y}), \\ \text{s.t.} \begin{cases} \mathbf{x} \in X, \\ \mathbf{H}_0(\mathbf{x},\mathbf{y},\mathbf{z},\boldsymbol{\lambda}) = \mathbf{0}^{m+2\ell}, \end{cases} \end{cases}$$

which is an equivalent, non-smooth, restatement of $(\mathscr{M})$, in the sense that the two problems share global as well as local optimal solutions in $\mathbf{x}$ (cf. [FJQ99, Proposition 1]).

2.3.2 A smooth approximation

Facchinei et al. [FJQ99] consider a smooth reformulation of the problem $(\mathscr{P})$, as follows. We introduce the function $\phi : \mathbb{R}^2 \mapsto \mathbb{R}$ by

$$\phi_\mu(a,b) := \sqrt{(a-b)^2 + 4\mu^2} - (a+b), \qquad (a,b) \in \mathbb{R}^2.$$

For this function, we have that ([FJQ99, Proposition 2])

$$\phi_\mu(a,b) = 0 \qquad \Longleftrightarrow \qquad a \geq 0, b \geq 0, ab = \mu^2.$$

For $\mu = 0$, $\phi_\mu(a,b) = -2\min(a,b)$; for $\mu \neq 0$, ϕ_μ is in C^∞; and for every pair (a,b), $\lim_{\mu \to 0} \phi_\mu(a,b) = -2\min(a,b)$. The function ϕ_μ therefore serves as a smooth perturbation of the min function. We consider replacing the operator $\mathbf{H}_0$ in the problem $(\mathscr{P})$ above with the smooth operator $\mathbf{H}_\mu$, defined by

$$\mathbf{H}_\mu(\mathbf{x},\mathbf{y},\mathbf{z},\boldsymbol{\lambda}) := \begin{pmatrix} \mathbf{F}(\mathbf{x},\mathbf{y}) - \nabla_{\mathbf{y}} \mathbf{g}(\mathbf{x},\mathbf{y}) \\ \mathbf{g}(\mathbf{x},\mathbf{y}) - \mathbf{z} \\ \boldsymbol{\Phi}_\mu(\mathbf{z},\boldsymbol{\lambda}) \end{pmatrix}, \qquad (\mathbf{x},\mathbf{y},\mathbf{z},\boldsymbol{\lambda}) \in \mathbb{R}^{n+m+2\ell},$$

where

$$\boldsymbol{\Phi}(\mathbf{z},\boldsymbol{\lambda}) := (\phi(z_1,\lambda_1),\ldots,\phi(z_\ell,\lambda_\ell)),$$

thus defining the smoothing problem

$$(\mathscr{P}_\mu)\begin{cases}\min\limits_{(\mathbf{x},\mathbf{y})} f(\mathbf{x},\mathbf{y}),\\ \text{s.t.}\begin{cases}\mathbf{x}\in X,\\ \mathbf{H}_\mu(\mathbf{x},\mathbf{y},\mathbf{z},\boldsymbol{\lambda}) = \mathbf{0}^{m+2\ell}.\end{cases}\end{cases}$$

While $(\mathscr{P}_0)$ coincides with the non-smooth problem $(\mathscr{M})$, the problem $(\mathscr{P}_\mu)$ for $\mu \neq 0$ is a smooth optimization problem. We denote the feasible set to $(\mathscr{P}_\mu)$ by $\mathscr{F}_\mu \subset \mathbb{R}^{n+m+2\ell}$. The function $\mathbf{H}_\mu$ has favourable properties: it is not only locally Lipschitz continuous for every μ but also regular (in the sense that the directional derivative exists in all directions and equals the Clarke derivative; cf. [FJQ99, Lemma 1]), and its generalized Jacobian with respect to $(\mathbf{y},\mathbf{z},\boldsymbol{\lambda})$ is non-singular for every μ and feasible point of the problem $(\mathscr{P}_\mu)$, cf. [FJQ99, Proposition 3]. Further, for every $\overline{\mathbf{x}} \in X$ and $\mu \in \mathbb{R}$ there exists a unique point in $\mathscr{F}_\mu$ such that its $\mathbf{x}$-part equals $\overline{\mathbf{x}}$, and this vector,

$$\overline{\mathbf{w}}_\mu := (\overline{\mathbf{x}},\mathbf{y}_\mu(\overline{\mathbf{x}}),\mathbf{y}_\mu(\overline{\mathbf{x}}),\boldsymbol{\lambda}_\mu(\overline{\mathbf{x}})), \tag{2.3}$$

is continuous in μ. Based on these properties, Facchinei et al. [FJQ99, Theorem 1] establish that the feasible sets $\mathscr{F}_\mu$ of the problem $(\mathscr{P}_\mu)$ are non-empty and uniformly compact; this is crucial, because then by the continuity of f, the problems $(\mathscr{P}_\mu)$ have optimal solutions.

2.3.3 Optimality conditions

We develop the optimality conditions of the problem $(\mathscr{P}_\mu)$. The first-order optimality conditions for the problem $(\mathscr{P}_\mu)$ can be written as follows: with $\mathbf{L}(\mathbf{x},\mathbf{y},\boldsymbol{\lambda}) := \mathbf{F}(\mathbf{x},\mathbf{y}) - \nabla_{\mathbf{y}}\mathbf{g}(\mathbf{x},\mathbf{y})\boldsymbol{\lambda}$, if $(\mathbf{x},\mathbf{y},\mathbf{z},\boldsymbol{\lambda})$ is a locally optimal solution to $(\mathscr{P}_\mu)$, then there exist vectors $(\theta,\rho,\sigma) \in \mathbb{R}^{m+2\ell}$ and $\mathbf{s} \in N_X(\mathbf{x}) \times \mathbb{R}^{m+2\ell}$ such that (cf. [FJQ99, Theorem 2])

$$\mathbf{0}^{n+m} \in \nabla f(\mathbf{x},\mathbf{y}) + \nabla\mathbf{L}(\mathbf{x},\mathbf{y},\boldsymbol{\lambda})\theta + \nabla(\mathbf{g}(\mathbf{x},\mathbf{y})-\mathbf{z})\rho + \sum_{i=1}^{\ell}\partial\phi_\mu(z_i,\lambda_i)\sigma_i + M\|(1,\theta,\rho,\sigma)\|\mathbf{s}, \tag{2.4}$$

where M is a Lipschitz constant for $(f,\mathbf{H}_\mu)$ around $(\mathbf{x},\mathbf{y},\mathbf{z},\boldsymbol{\lambda})$. Since the multiplier for $\nabla f(\mathbf{x},\mathbf{y})$ is non-zero (it then equals 1, without any loss of generality), this condition is stronger than the Fritz–John conditions, and is in fact the KKT conditions for the problem. While one may then refer to this condition for $\mu = 0$ as the KKT conditions for the MPEC problem $(\mathscr{M})$, Facchinei et al. [FJQ99] refer to it is as *strong C-stationarity* (SCS).

2.3.4 Global convergence

A global version of the smoothing algorithm is immediate: with an arbitrary choice of starting point $\mathbf{w}^0 := (\mathbf{x}^0,\mathbf{y}^0,\mathbf{z}^0,\boldsymbol{\lambda}^0)$, and $\{\mu_\tau\}$ being any sequence of non-zero numbers with $\lim_{\tau\to\infty}\mu_\tau = 0$, we define the sequence $\{\mathbf{w}^\tau\}$ to be given by globally optimal solutions $\mathbf{w}^\tau := (\mathbf{x}^\tau,\mathbf{y}^\tau,\mathbf{z}^\tau,\boldsymbol{\lambda}^\tau)$ to the problems $(\mathscr{P}_{\mu_\tau})$.

For this algorithm, it is not difficult to establish (cf. [FJQ99, Theorem 3]) that the sequence $\{\mathbf{w}^\tau\}$ is bounded and that every limit point is a globally optimal solution to $(\mathscr{P})$; it follows from the uniform compactness of the feasible sets $\mathscr{F}_{\mu_\tau}$ and the continuity of the trajectory defined by $\bar{\mathbf{w}}_\mu$ in (2.3).

2.3.5 Convergence to stationary points

A more practical algorithm is obtained by replacing, in the algorithm above, global optimality in $(\mathscr{P}_{\mu_\tau})$ of the vector $\mathbf{w}^\tau$ by stationarity in the sense of the KKT system (2.4). For this algorithm, it is shown in [FJQ99, Theorem 4] that the sequence $\{\mathbf{w}^\tau\}$ of KKT points in $(\mathscr{P}_{\mu_\tau})$ is bounded and every limit point is an SCS point in $(\mathscr{P})$. A crucial part of the proof is the continuity property of any sequence of KKT points in the problem $(\mathscr{P}_\mu)$ as μ tends to zero (cf. [FJQ99, Proposition 4]). The proof of the convergence result also establishes the important result that the sequence $\{(\theta^\tau, \rho^\tau, \sigma^\tau)\}$ of KKT multipliers is bounded. This is a crucial part of any analysis of the stationarity property of a limit point.

A yet more practical algorithm is also devised, in which the sequence $\{\mathbf{w}^\tau\}$ of vectors is allowed to be defined by near-feasible and approximate KKT points. In other words, in each iteration τ, the distance from the vector $\mathbf{w}^\tau$ to the feasible set $\mathscr{F}_\mu$ of the problem $(\mathscr{P}_{\mu_\tau})$ is bounded by $\varepsilon_\tau > 0$, and the Euclidean length of the vector defining the right-hand side of the inclusion (2.4) is also bounded above by this value. Theorem 5 in [FJQ99] then states that if $\{\varepsilon_\tau\} \downarrow 0$ as $\{\mu_\tau\} \to 0$, then the sequence $\{\mathbf{w}^\tau\}$ of approximate KKT points is bounded and every limit point is, again, a SCS point in $(\mathscr{P})$.

The latter algorithm was coded and tested in [FJQ99] on some small and medium-size MPEC problems; each problem $(\mathscr{P}_{\mu_\tau})$ was then solved by utilizing an SQP algorithm. They report that it compares favourably with, for example, the implicit programming algorithms proposed in [Out94, OuZ95].

2.4 A new smoothing approach to topology optimization

2.4.1 Motivation

The smoothing algorithm described in Section 2.3 may unfortunately not be applied to truss topology optimization problems, out of which $(\mathscr{W})$ is a typical example. The latter problem violates several assumptions that are vital for the smoothing algorithm of Facchinei et al. [FJQ99], the most important being the lack of the uniform strong monotonicity by the lower-level problem $\mathrm{AVI}(\mathbf{q}(\mathbf{f}), \mathrm{Q}(\mathbf{x}), Y)$. In addition, some of the variables (that is, $\mathbf{u}$) may not be uniformly bounded, and upper-level joint constraints (such as stress constraints) are essential in the problem $(\mathscr{W})$.

In order to overcome the difficulties outlined we introduce an alternative perturbation scheme for solving stress constrained weight minimization problems for trusses including unilateral constraints. It resembles the ε-perturbation approach of Cheng and Guo [ChG97] (cf. Section 2.2.1) by the fact that we introduce positive lower bounds on the design variables, thus restricting the design domain. There are important differences, however: instead of relaxing the technological constraints (e.g., stress constraints in the original paper [ChG97]) we relax the equilibrium constraint; to accomplish this, we formulate the optimization problem using *both* the design and the state variables, similarly to the extended formulation of Stolpe and Svanberg [StS03] (cf. Section 2.2.2).

2.4.2 Relaxed equilibrium problem

Formally, fix an arbitrary $\varepsilon \geq 0$ and consider the following perturbation of the feasible set $\mathscr{F}(\mathbf{f})$ (cf. (2.1)):

$$\begin{aligned}\mathscr{F}^{\varepsilon}(\mathbf{f}) := \{\, (\mathbf{x},\mathbf{s},\boldsymbol{\lambda},\mathbf{u}) \in \mathbb{R}^m_+ \times \mathbb{R}^m \times \mathbb{R}^r_+ \times \mathbb{R}^n \mid \mathbf{x} \geq o(\varepsilon)\mathbf{1}^m,\\ \mathscr{E}(\mathbf{x},\mathbf{s},\boldsymbol{\lambda}) + \Pi(\mathbf{x},\mathbf{u}) - \mathbf{f}^t\mathbf{u} \leq \varepsilon\\ \mathsf{B}^t\mathbf{s} + \mathsf{C}^t\boldsymbol{\lambda} = \mathbf{f},\\ \mathsf{C}\mathbf{u} \leq \mathbf{g}\,\},\end{aligned}$$

where $o(\varepsilon)$ is a positive function of ε such that $\lim_{\varepsilon\downarrow 0} o(\varepsilon)/\varepsilon = 0$. Of course, the weak duality theorem for convex problems implies that $\mathscr{F}^0(\mathbf{f}) = \mathscr{F}(\mathbf{f})$; for positive values of ε the "state" variables $(\mathbf{s},\boldsymbol{\lambda},\mathbf{u})$ (which in the extended formulation play a role equal to that of the design variables $\mathbf{x}$, and do not correspond to an equilibrium state of the truss anymore) are required to be primal-dual feasible, but only ε-optimal.

From the theoretical point of view, allowing for ε-optimal solutions to the lower-level equilibrium problem means that we "regularize" the bi-level programming problem $(\mathscr{W})$, in the sense defined by [LiM97]; this will allow us to obtain the convergence of both globally optimal solutions and stationary points (see below). At least of equal importance is the practical interpretation of the method, where the relaxation parameter ε comes from the approximate numerical solution of the equilibrium problem (e.g., using the existing finite element software).

Finally, for every $\varepsilon > 0$ we consider the following *perturbed* version of the stress constrained weight minimization problem:

$$(\mathscr{W}^{\varepsilon}) \quad \begin{cases} \min\limits_{(\mathbf{x},\mathbf{s},\boldsymbol{\lambda},\mathbf{u})} w(\mathbf{x}), \\ \text{s.t. } (\mathbf{x},\mathbf{s},\boldsymbol{\lambda},\mathbf{u}) \in \mathscr{F}^{\varepsilon}(\mathbf{f}), \\ \qquad \underline{\sigma}_i x_i \leq s_i \leq \overline{\sigma}_i x_i, \quad i = 1,\dots,m. \end{cases}$$

Remark 2.2. In the multiple-load case, the problem $(\mathscr{W}^{\varepsilon})$ will have several constraints of the form $(\mathbf{x},\mathbf{s}^k,\boldsymbol{\lambda}^k,\mathbf{u}^k) \in \mathscr{F}^{\varepsilon}(\mathbf{f}^k)$, where $\mathbf{f}^k$ is a vector of external forces corresponding to the load case k, and the triple $(\mathbf{s}^k,\boldsymbol{\lambda}^k,\mathbf{u}^k) \in \mathbb{R}^m \times \mathbb{R}^r_+ \times \mathbb{R}^n$ represents the "state" variables for the load case k, $k = 1,\dots,\ell$.

In the rest of the section, we study the theoretical properties of the point-to-set mapping $\varepsilon \rightrightarrows \mathscr{F}^{\varepsilon}(\mathbf{f})$ which will allow us to establish the convergence of globally optimal solutions as well as stationary points as $\varepsilon \downarrow 0$.

2.4.3 Properties of $\varepsilon \rightrightarrows \mathscr{F}^{\varepsilon}(\mathbf{f})$

In this section we show that the point-to-set mapping $\varepsilon \rightrightarrows \mathscr{F}^{\varepsilon}(\mathbf{f})$ enjoys most of the nice properties one can expect from a point-to-set mapping: under some mild conditions it has compact (although, unfortunately, non-convex) images, and is closed and lower semicontinuous at zero [AuF90, Chapter 1]. Furthermore, in Proposition 2.7 we demonstrate a continuity of the design-to-force "sub-mapping" $\mathbf{x} \rightrightarrows (\mathbf{s}, \boldsymbol{\lambda})$ (see Proposition 2.7 for the formal definition), the property originally established for the unperturbed feasible set $\mathscr{F}$ by Petersson [Pet01] for trusses without unilateral constraints, and later generalized by Patriksson and Petersson [PaP02].

We formulate the results as a sequence of short propositions.

Proposition 2.3 (Closed images). *For each $\varepsilon \geq 0$ the set $\mathscr{F}^\varepsilon(\mathbf{f})$ is closed.*

Proof. The claim follows easily from the lower semicontinuity of $\mathscr{E}(\cdot,\cdot,\cdot)$ (cf. [PaP02, Lemma 3.2]) together with the continuity of the other functions defining $\mathscr{F}^\varepsilon(\mathbf{f})$, $\varepsilon \geq 0$. □

Proposition 2.4 (Lower semicontinuity). *The multi-function $\varepsilon \rightrightarrows \mathscr{F}^\varepsilon(\mathbf{f})$ is lower semicontinuous at zero.*

Proof. Let $(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}, \mathbf{u}) \in \mathscr{F}(\mathbf{f})$. Then, $\{\, (\mathbf{x} + o(\varepsilon)\mathbf{1}^m, \mathbf{s}, \boldsymbol{\lambda}, \mathbf{u})\,\} \in \mathscr{F}^\varepsilon(\mathbf{f})$ for all enough small $\varepsilon > 0$, where $\mathbf{1}^m = (1, \dots, 1)^t \in \mathbb{R}^m_+$. □

Remark 2.5. The same construction establishes the lower semicontinuity of the multi-functions $\varepsilon \to (\mathscr{F}^\varepsilon \cap K)$, where (independent of ε) the closed set K may represent stress, stiffness, or global stability constraints, or any combination thereof.

We stress that the classic ε-perturbation of Cheng and Guo [ChG97] results in a l.s.c. mapping including design variables *only*; i.e., there might be some displacement vectors corresponding to the limiting design that cannot be approximated with the displacements corresponding to positive designs.

Proposition 2.6 (Closedness). *The multi-function $\varepsilon \rightrightarrows \mathscr{F}^\varepsilon(\mathbf{f})$ is closed at zero.*

Proof. The claim follows from the lower semicontinuity of $\mathscr{E}(\cdot,\cdot,\cdot)$ (cf. [PaP02, Lemma 3.2]) together with the continuity of the other functions, defining the sets $\mathscr{F}^\varepsilon(\mathbf{f})$, $\varepsilon \geq 0$. □

Proposition 2.7 (Continuity of the design–to–force mapping). *Let $\{\varepsilon_k\}$ be a positive sequence, converging to zero. Assume that $(\mathbf{x}^k, \mathbf{s}^k, \boldsymbol{\lambda}^k, \mathbf{u}^k) \in \mathscr{F}^{\varepsilon_k}(\mathbf{f})$, and that $\{\mathbf{x}^k\} \to \mathbf{x}$. Suppose further that for each $k = 1, 2, \dots$, $i = 1, \dots, m$, the stress constraints $\underline{\sigma}_i x_i^k \leq s_i^k \leq \overline{\sigma}_i x_i^k$ constraints are satisfied. Then, $\{(\mathbf{s}^k, \boldsymbol{\lambda}^k)\} \to (\mathbf{s}, \boldsymbol{\lambda})$, this limit vector solves $(\mathscr{C})_{\mathbf{x}}(\mathbf{f})$, and there is a vector $\mathbf{u}$ solving $(\mathscr{P})_{\mathbf{x}}(\mathbf{f})$. [In particular, $(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}, \mathbf{u}) \in \mathscr{F}$.]*

Proof. The additional stress constraints imply the uniform boundedness of the sequence of complementary energies $\{\mathscr{E}(\mathbf{x}^k, \mathbf{s}^k, \boldsymbol{\lambda}^k)\}$, as has been established in [PaP02]. Therefore, the sequence $\{(\mathbf{s}^k, \boldsymbol{\lambda}^k)\}$ is bounded, owing to the coercivity of $\mathscr{E}$, which is locally uniform with respect to the design. Let $(\mathbf{s}, \boldsymbol{\lambda})$ be an arbitrary limit point of this sequence. The lower semicontinuity of $\mathscr{E}$ and the uniform boundedness of energies yield that

$$\mathscr{E}(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}) \leq \liminf_{k \to \infty} \mathscr{E}(\mathbf{x}^k, \mathbf{s}^k, \boldsymbol{\lambda}^k) < \infty.$$

Therefore, the problem $(\mathscr{C})_{\mathbf{x}}(\mathbf{f})$ is feasible and thus possesses a unique optimal solution (cf. [PaP02, Theorem 2.1]).

Let now $(\tilde{\mathbf{s}},\tilde{\boldsymbol{\lambda}})$ be an arbitrary force distribution that is feasible in $(\mathscr{C})_{\mathbf{x}}(\mathbf{f})$. Then, from the ε_k-optimality of $(\mathbf{s}^k,\boldsymbol{\lambda}^k)$ and feasibility of $(\tilde{\mathbf{s}},\tilde{\boldsymbol{\lambda}})$ in $(\mathscr{C})_{\mathbf{x}^k}(\mathbf{f})$ it follows that

$$\mathscr{E}(\mathbf{x},\mathbf{s},\boldsymbol{\lambda}) \le \liminf_{k\to\infty} \mathscr{E}(\mathbf{x}^k,\mathbf{s}^k,\boldsymbol{\lambda}^k) \le \lim_{k\to\infty} \mathscr{E}(\mathbf{x}^k,\tilde{\mathbf{s}},\tilde{\boldsymbol{\lambda}}) + \varepsilon_k = \mathscr{E}(\mathbf{x},\tilde{\mathbf{s}},\tilde{\boldsymbol{\lambda}}),$$

where the equality follows from the continuity of $\mathscr{E}(\cdot,\tilde{\mathbf{s}},\tilde{\boldsymbol{\lambda}})$ (cf. [PaP02, Lemma 3.2]). Therefore, $(\mathbf{s},\boldsymbol{\lambda})$ is optimal in $(\mathscr{C})_{\mathbf{x}}(\mathbf{f})$. It follows that $(\mathbf{s},\boldsymbol{\lambda})$ must be the only limit point of the sequence $\{(\mathbf{s}^k,\boldsymbol{\lambda}^k)\}$.

The existence of at least one dual optimal solution $\mathbf{u}$ to $(\mathscr{P})_{\mathbf{x}}(\mathbf{f})$ follows. □

Proposition 2.8 (Compact images). *For every* $\varepsilon > 0$ *and every constant* $M > 0$ *the set* $\{(\mathbf{x},\mathbf{s},\boldsymbol{\lambda},\mathbf{u}) \in \mathscr{F}^{\varepsilon} \mid \|\mathbf{x}\| \le M\}$ *is compact.*

Proof. The function $\mathscr{E}(\mathbf{x},\mathbf{s},\boldsymbol{\lambda}) + \Pi(\mathbf{x},\mathbf{u}) - \mathbf{f}^t\mathbf{u}$ is continuous as well as coercive in $(\mathbf{s},\boldsymbol{\lambda},\mathbf{u})$, uniformly in $\mathbf{x}$ for all $\mathbf{x} \ge o(\varepsilon)\mathbf{1}^m$, with $\|\mathbf{x}\| \le M$. □

In the subsections that follow we apply the continuity results we have just established to show that the ε-perturbed problems can indeed be used as approximating problems for small ε, both if we are interested in globally optimal solutions and stationary points.

2.4.4 Regularity of $(\mathscr{W}^{\varepsilon})$

To be of practical use, every approximating problem $(\mathscr{W}^{\varepsilon})$ should be easier to solve than the original problem $(\mathscr{W})$. Clearly, the functions defining the constraints of $(\mathscr{W}^{\varepsilon})$ are continuously differentiable on some neighbourhood of the feasible set $\mathscr{F}^{\varepsilon}$ for every $\varepsilon > 0$; therefore, the smooth Fritz–John conditions must hold at optimal points. The following (purely academic) example shows that the feasible sets of the optimization problems $(\mathscr{W}^{\varepsilon})$ do not in general verify MFCQ, and therefore we cannot expect the KKT conditions to be satisfied at every point of local minimum. On the other hand, in Proposition 2.10 we show that under rather mild additional conditions MFCQ is verified, so that standard nonlinear programming algorithms can be used to find locally optimal solutions of $(\mathscr{W}^{\varepsilon})$.

Example 2.9. Consider a simple 1-bar structure shown in Figure 2.1 that is made of (academic) material with the Young modulus $E = 1$. Let $f = 3$, $g = 2$, $\varepsilon = 1$, $\overline{\sigma} = (2-\sqrt{2})$, and consider the point of global minimum $(x,s,\lambda,u) = (1, 2-\sqrt{2}, 1+\sqrt{2}, 2)$. At this feasible in $(\mathscr{W}^{\varepsilon})$ point the active constraints are:

$$\begin{cases} s+\lambda = f, \\ -x \le -\varepsilon^2, \\ u \le g, \\ s \le \overline{\sigma}x, \\ \dfrac{s^2}{2x} + g\lambda + \dfrac{1}{2}u^2x - fu \le \varepsilon \end{cases} \Leftrightarrow \begin{cases} s+\lambda = 3, \\ -x \le -1, \\ u \le 2, \\ s \le (2-\sqrt{2})x, \\ \dfrac{s^2}{2x} + 2\lambda + \dfrac{1}{2}u^2x - 3u \le 1. \end{cases}$$

It is easy to verify that there is no direction $\mathbf{d} \in \mathbb{R}^4$ such that

$$\begin{cases} \begin{pmatrix} 0 & 1 & 1 & 0 \end{pmatrix} \mathbf{d} = 0, \\ \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -\overline{\sigma} & 1 & 0 & 0 \end{pmatrix} \mathbf{d} \leq \mathbf{0}^3, \\ \begin{pmatrix} -\frac{s^2}{2x^2} + \frac{u^2}{2} & \frac{s}{x} & g & xu - f \end{pmatrix} \mathbf{d} < 0, \end{cases}$$

so that MFCQ is violated at (x, s, λ, u).

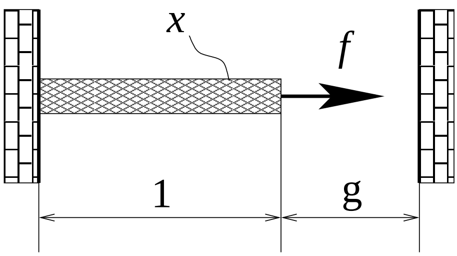


Fig. 2.1: 1-bar truss structure.

While MFCQ is violated at the point of global minimum in the Example 2.9, this does not prevent the KKT conditions to hold at this point, because the more basic Abadie's CQ is still verified. While for realistic trusses the latter CQ is close to impossible to verify, the following result resolves the problem of verifying a CQ in most practical situations.

Proposition 2.10. *Let* $(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}, \mathbf{u})$ *be a point of local minimum for* $(\mathscr{W}^\varepsilon)$, $\varepsilon > 0$. *Suppose that any of the following conditions are verified:*
(i) $\mathscr{E}(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}) + \Pi(\mathbf{x}, \mathbf{u}) - \mathbf{f}^t\mathbf{u} < \varepsilon$*;*
(ii) $r = 0$, *that is, no rigid obstacles are present;*
(iii) $\mathbf{u}$ *is not the equilibrium displacement corresponding to* $\mathbf{x}$.
Then, Abadie's CQ hold at $(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}, \mathbf{u})$. *In particular, the KKT-conditions for* $(\mathscr{W}^\varepsilon)$ *hold at this point.*

Proof. Suppose that (i) holds. Then the relaxed equilibrium constraint is passive, and the feasible set of the problem $(\mathscr{W}^\varepsilon)$ is locally around $(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}, \mathbf{u})$ defined by affine constraints only, which guarantees Abadie's CQ.

Alternatively, assume that there are no rigid obstacles, i.e., (ii) holds. Consider the direction $\mathbf{d} = (\alpha\mathbf{x}, \mathbf{0}^m, \mathbf{0}^r, -\beta\mathbf{u})$, where $\alpha > 0$, $\beta \geq 0$ are parameters to be determined. This direction is *feasible* with respect to all linear constraints of $(\mathscr{W}^\varepsilon)$. Furthermore, an easy calculation shows that

$$\nabla[\mathscr{E}(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}) - \Pi(\mathbf{x}, \mathbf{u}) - \mathbf{f}^t\mathbf{u}]^t\mathbf{d} = -\alpha[\mathscr{E}(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}) - \Pi(\mathbf{x}, \mathbf{u})] - \beta[2\Pi(\mathbf{x}, \mathbf{u}) - \mathbf{f}^t\mathbf{u}] < 0,$$

for some $\alpha > 0$, $\beta \geq 0$, owing to the inequality

$$0 < \varepsilon = [\mathscr{E}(\mathbf{x},\mathbf{s},\boldsymbol{\lambda}) - \Pi(\mathbf{x},\mathbf{u})] + [2\Pi(\mathbf{x},\mathbf{u}) - \mathbf{f}^t\mathbf{u}].$$

Thus, the MFCQ is verified, implying Abadie's CQ.

At last, assume that (iii) is verified. Similarly to the case (ii) we can show that the direction $\mathbf{d} = (\alpha\mathbf{x}, \mathbf{0}^m, \mathbf{0}^r, \beta[\mathbf{u}(\mathbf{x}) - \mathbf{u}])$ satisfies the requirements of MFCQ for some $\alpha > 0$, $\beta \geq 0$, where $\mathbf{u}(\mathbf{x})$ is the equilibrium displacement, corresponding to $\mathbf{x}$. □

Naturally, all three assumptions of Proposition 2.10 are violated by Example 2.9.

It is interesting to note that topology optimization problems for trusses without unilateral constraints are always qualified in the sense of Mangasarian–Fromowitz; it is probably even more interesting to see that the violation of MFCQ may happen even for "nice" feasible points that verify a strict complementarity assumption for MPEC problems (like the point considered in Example 2.9).

2.4.5 Optimality conditions for $(\mathscr{W})$

Motivated by the description of the feasible sets of the approximating problems $(\mathscr{W}^\varepsilon)$, $\varepsilon > 0$, in terms of differentiable inequalities which lead to at least Fritz–John necessary optimality conditions (see Example 2.9 and Proposition 2.10), we may use the same description with $\varepsilon = 0$ in order to develop non-smooth necessary optimality conditions for $(\mathscr{W})$. The biggest difficulty we encounter is the loss of continuity (not to mention differentiability) of the complementary energy function $\mathscr{E}$. Indeed, if we look at the constraint involving $\mathscr{E}$:

$$\mathscr{E}(\mathbf{x},\mathbf{s},\boldsymbol{\lambda}) + \Pi(\mathbf{x},\mathbf{u}) - \mathbf{f}^t\mathbf{u} \leq 0, \tag{2.5}$$

we note that the function on the left-hand side of the inequality is neither Lipschitz continuous nor convex, and therefore the classic subdifferentials of such functions are not defined. On the other hand, we may use the structure of this function: it is continuously differentiable everywhere except when $\mathbf{x} \in \partial\mathbb{R}^m_+$, and it is a sum of convex and Lipschitz continuous functions. Therefore, the notion of *limiting subdifferential* ∂_a is well defined for such functions (see [Mor76]). In particular, it holds that

$$\partial_a[\mathscr{E}(\mathbf{x},\mathbf{s},\boldsymbol{\lambda}) + \Pi(\mathbf{x},\mathbf{u}) - \mathbf{f}^t\mathbf{u}] = \partial_a\mathscr{E}(\mathbf{x},\mathbf{s},\boldsymbol{\lambda}) + \nabla[\Pi(\mathbf{x},\mathbf{u}) - \mathbf{f}^t\mathbf{u}].$$

As a result, we obtain the following non-smooth Fritz–John type optimality conditions.

Proposition 2.11. *Let $(\mathbf{x},\mathbf{s},\boldsymbol{\lambda},\mathbf{u})$ be a point of local minimum for $(\mathscr{W})$. To simplify notation we write all inequality and equality constraints of $(\mathscr{W})$, except the relaxed equilibrium, constraint in the form:*

$$\mathsf{A}_i(\mathbf{x}^t,\mathbf{s}^t,\boldsymbol{\lambda}^t,\mathbf{u}^t)^t \leq \mathbf{b}_i,$$
$$\mathsf{A}_e(\mathbf{x}^t,\mathbf{s}^t,\boldsymbol{\lambda}^t,\mathbf{u}^t)^t = \mathbf{b}_e,$$

where $\mathsf{A}_i \in \mathbb{R}^{N_i\times(m+m+r+n)}$, $\mathsf{A}_e \in \mathbb{R}^{N_e\times(m+m+r+n)}$, $\mathbf{b}_i \in \mathbb{R}^{N_i}$, and $\mathbf{b}_e \in \mathbb{R}^{N_e}$ are matrices and vectors of appropriate sizes. Then, the non-smooth Fritz–John optimality conditions *hold at* $(\mathbf{x},\mathbf{s},\boldsymbol{\lambda},\mathbf{u})$*, that is, there are vectors* $\mu_i \in \mathbb{R}^{N_i}_+$, $\mu_e \in \mathbb{R}^{N_e}$*, and numbers* $\mu_0, \mu \in \mathbb{R}_+$ *not all equal to zero such that:*

$$\begin{aligned} \mathbf{0}^{m+m+r+n} &\in \mu_0 \nabla w(\mathbf{x}) + \mathsf{A}_i^t \mu_i + \mathsf{A}_e^t \mu_e + \mu[\partial_a \mathscr{E}(\mathbf{x},\mathbf{s},\boldsymbol{\lambda}) + \nabla(\Pi(\mathbf{x},\mathbf{u}) - \mathbf{f}^t\mathbf{u}))], \quad \textit{and} \\ 0 &= \mu_i^t[\mathsf{A}_i(\mathbf{x}^t,\mathbf{s}^t,\boldsymbol{\lambda}^t,\mathbf{u}^t)^t - \mathbf{b}_i]. \end{aligned} \tag{2.6}$$

In general, we cannot expect the KKT conditions to be satisfied at every point of local minimum, because the problem $(\mathscr{W})$ is usually much less regular than its approximation $(\mathscr{W}^\varepsilon)$, $\varepsilon > 0$, and even the latter problem may violate the standard nonlinear programming constraint qualifications (see Example 2.9). In fact, in Problem 2, Subsection 2.5.2, we obtained a locally optimal solution that satisfies the system (2.6) only with $\mu_0 = 0$. It is sad to note that this example does not contain any contact conditions, and the optimal solution we obtained is non-singular (in particular, no bars were removed), yet it is only a FJ point in our formulation. On the positive side, at least if unilateral condition are absent, the conditions (2.6) imply the fulfillment of the KKT conditions for a related optimization problem that has a clear engineering interpretation. Namely, the stationary point obtained is a KKT point for a "semi-fixed topology" optimization problem, in which the given subset of the bars is removed from the ground structure; formally, the following result holds.

Proposition 2.12. *Assume that the unilateral constraints are absent and that the point* $(\hat{\mathbf{x}},\hat{\mathbf{s}},\hat{\mathbf{u}}) \in \mathbb{R}^m_+ \times \mathbb{R}^m \times \mathbb{R}^n$ *satisfies the FJ optimality conditions* (2.6)*. Let* $\hat{\mathscr{I}} = \{\, i = 1,\dots,m \mid \hat{x}_i = 0\,\}$*. Then, the point* $(\hat{\mathbf{x}},\hat{\mathbf{s}},\hat{\mathbf{u}})$ *is a KKT-point for the following problem:*

$$(\hat{\mathscr{W}}) \begin{cases} \min\limits_{(\mathbf{x},\mathbf{s},\mathbf{u})} w(\mathbf{x}) \\ \textit{s.t. } \mathsf{B}^t\mathbf{s} = \mathbf{f}, \\ \qquad \mathscr{E}(\mathbf{x},\mathbf{s},\boldsymbol{\lambda}) + \Pi(\mathbf{x},\mathbf{u}) - \mathbf{f}^t\mathbf{u} = 0, \\ \qquad \underline{\sigma}_i x_i \le s_i \le \overline{\sigma}_i x_i, \quad i \in \{\,1,\dots,m\,\} \setminus \hat{\mathscr{I}}, \\ \qquad x_i = s_i = 0, \quad i \in \hat{\mathscr{I}}. \end{cases}$$

Proof. Clearly the point $(\hat{\mathbf{x}},\hat{\mathbf{s}},\hat{\mathbf{u}})$ is feasible in the problem $(\hat{\mathscr{W}})$. Furthermore, it is easy to check that the feasible set of the problem $(\hat{\mathscr{W}})$ verifies a Mangasarian–Fromowitz type constraint qualification at $(\hat{\mathbf{x}},\hat{\mathbf{s}},\hat{\mathbf{u}})$ [one can, for example, take the direction $\mathbf{d} = (\hat{\mathbf{x}},\mathbf{0},-\mathbf{u}) \in \mathbb{R}^m_+ \times \mathbb{R}^m \times \mathbb{R}^n$ to verify that], and thus the FJ conditions (2.6) [that can be viewed as FJ conditions for $(\hat{\mathscr{W}})$] also imply the KKT conditions. □

There are of course other approaches to optimality conditions for MPEC. For example, [OKZ98, Theorem 7.2] establishes non-smooth KKT-type conditions for a problem rather similar to $(\mathscr{W})$. However, the strong regularity condition on the lower-level problem assumed in [OKZ98, Theorem 7.2] is violated by our problem, because the displacements $\mathbf{u}$ are in general not uniquely determined for designs $\mathbf{x} \in \partial\mathbb{R}^m_+$.

2.4.6 Global convergence

Convergence of globally optimal solutions to relaxed weight minimization problems with stress constraints $(\mathscr{W}^{\varepsilon})$ towards globally optimal solutions to the limiting problem $(\mathscr{W})$ as $\varepsilon \downarrow 0$ follows easily, given the results of the previous subsections.

Proposition 2.13. *Consider a positive sequence $\{\varepsilon_k\}$ converging to zero. Let $\{(\mathbf{x}_{\varepsilon_k}, \mathbf{s}_{\varepsilon_k}, \boldsymbol{\lambda}_{\varepsilon_k}, \mathbf{u}_{\varepsilon_k})\}$ be a corresponding sequence of globally optimal solutions to $\{(\mathscr{W}^{\varepsilon_k})\}$. Then, an arbitrary limit point of this sequence is a globally optimal solution to the limiting problem $(\mathscr{W})$.*

Proof. That globally optimal solutions to the sequence of problems $\{(\mathscr{W}^{\varepsilon_k})\}$ exist follows by the coercivity of the objective w.r.t. the design variables, Proposition 2.8, and Weierstrass' Theorem).

Without any loss of generality, assume that $\lim_{k\to+\infty}(\mathbf{x}_{\varepsilon_k}, \mathbf{s}_{\varepsilon_k}, \boldsymbol{\lambda}_{\varepsilon_k}, \mathbf{u}_{\varepsilon_k}) = (\widetilde{\mathbf{x}}, \widetilde{\mathbf{s}}, \widetilde{\boldsymbol{\lambda}}, \widetilde{\mathbf{u}})$. Then, owing to Proposition 2.6, the point $(\widetilde{\mathbf{x}}, \widetilde{\mathbf{s}}, \widetilde{\boldsymbol{\lambda}}, \widetilde{\mathbf{u}})$ is feasible in $(\mathscr{W})$. Together with Remark 2.5 and the continuity of the objective functional this proves the claim. □

In general, the displacement component $\{\mathbf{u}_k\}$ of the sequence of global optimal solutions we study in Proposition 2.13 need not to have any limit points. However, we may use the fact that our objective function is independent of the displacements and utilize Proposition 2.7 to establish the following result.

Proposition 2.14. *Consider a positive sequence $\{\varepsilon_k\}$ converging to zero. Let $\{(\mathbf{x}_{\varepsilon_k}, \mathbf{s}_{\varepsilon_k}, \boldsymbol{\lambda}_{\varepsilon_k}, \mathbf{u}_{\varepsilon_k})\}$ be a corresponding sequence of globally optimal solutions to $\{(\mathscr{W}^{\varepsilon_k})\}$. Then, an arbitrary limit point $(\mathbf{x}_0, \mathbf{s}_0, \boldsymbol{\lambda}_0)$ of the sequence $\{(\mathbf{x}_{\varepsilon_k}, \mathbf{s}_{\varepsilon_k}, \boldsymbol{\lambda}_{\varepsilon_k})\}$* (*and there is at least one*) *corresponds to some globally optimal solution $(\mathbf{x}_0, \mathbf{s}_0, \boldsymbol{\lambda}_0, \mathbf{u}_0)$ to the limiting problem $(\mathscr{W})$.*

Proof. Similar to the proof of Proposition 2.13, but uses Proposition 2.7 instead of Proposition 2.6. □

2.4.7 Convergence of stationary points

The main result of this paper, Theorem 2.15, uses the fact that stress constraints are imposed. Furthermore, we need to make an assumption that the sequence of displacements $\{\mathbf{u}^{\varepsilon}\}$ produced by the smoothing procedure is bounded as $\varepsilon \downarrow 0$. We cannot guarantee the latter property without imposing explicit bounds on the displacements; however, our computational experience with the smoothing approach we introduce in this paper confirms that convergence of displacements takes place in practice. In any case, Proposition 2.4 asserts that it is at least possible to approximate every equilibrium state using the relaxation approach we propose; this is contrary to traditional ε-relaxation, where some equilibrium displacements cannot be approximated.

Theorem 2.15. *Consider a positive sequence $\{\varepsilon_k\}$ converging to zero. Let $\{(\mathbf{x}_{\varepsilon_k}, \mathbf{s}_{\varepsilon_k}, \boldsymbol{\lambda}_{\varepsilon_k}, \mathbf{u}_{\varepsilon_k})\}$ be a sequence of KKT-points to $\{(\mathscr{W}^{\varepsilon_k})\}$. Then, every limit point of this sequence is feasible in the limiting problem $(\mathscr{W})$, and in addition it verifies the non-smooth FJ stationarity conditions* (2.6).

Proof. Without loss of generality we assume that $\{(\mathbf{x}_{\varepsilon_k}, \mathbf{s}_{\varepsilon_k}, \boldsymbol{\lambda}_{\varepsilon_k}, \mathbf{u}_{\varepsilon_k})\} \to (\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}, \mathbf{u})$ as k converges to infinity. Owing to Proposition 2.6, the point $(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}, \mathbf{u})$ is feasible in $(\mathscr{W})$.

The stress constraints imply that the gradients $\nabla\mathscr{E}(\mathbf{x}_{\varepsilon_k}, \mathbf{s}_{\varepsilon_k}, \boldsymbol{\lambda}_{\varepsilon_k})$ are uniformly bounded for $k = 1, 2, \dots$ Therefore, the sequence $\{\nabla\mathscr{E}(\mathbf{x}_{\varepsilon_k}, \mathbf{s}_{\varepsilon_k}, \boldsymbol{\lambda}_{\varepsilon_k})\}$ has at least one limit point that by definition is a member of $\partial_a\mathscr{E}(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda})$. It is now an easy exercise to verify that the point $(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}, \mathbf{u})$ satisfies the system (2.6). □

Again, the optimality conditions we obtain in Theorem 2.15 are rather weak, but we cannot expect more from points of local minima for $(\mathscr{W})$ in general; see the discussion in Section 2.4.5.

2.5 Numerical experiments

While a substantial amount of theoretical studies of topology optimization problems for trusses including unilateral frictionless contact has been carried out (see, e.g., [BTKNZ99, PaP02, EPP03]), surprisingly little numerical experience has been reported. Therefore we use a comprehensive numerical study of Stolpe [Sto03] (who was interested in finding globally optimal solutions using a branch-and-cut algorithm) as a rich and authoritative source of benchmark problems, unfortunately however for trusses without contact. We also compare our algorithm against a few tests of "classic" MPEC algorithms (implicit programming-based algorithm, IMPA, [LPR96, Section 6.3], and penalty interior point algorithm, PIPA, [LPR96, Section 6.1]), MMA [Sva87] (see also [Sva02]), as well as the smoothing algorithm [FJQ99], made by Hilding [HKP99, Hil00]. Unfortunately, the latter studies are not concerned with *topology* optimization (i.e., a strictly positive bound on the bar volumes is imposed) and stress constraints are not included.

Below we present some preliminary numerical experience with an academic implementation of our approximation method.

2.5.1 Implementation issues

A sequence of smooth optimization subproblems $\{(\mathscr{W}^{\varepsilon_k})\}$ has been solved using the SQP-solver SNOPT [GMS02]. The optimal solution obtained at step k was used as a starting point for the step $k+1$. We used a simple update rule for ε: $\varepsilon_{k+1} = \gamma\varepsilon_k$, where $\gamma \in [0.25, 0.75]$.

The biggest computational difficulty we have noted is that the projected gradient of the potential energy with respect to displacements is close to zero for all points feasible in $(\mathscr{W}^{\varepsilon})$ when ε is small, resulting in rather slow progress of the optimization procedure based on the first order information only. The use of second order information in this case seems essential for improving the performance.

Another problem is that the complementary energy has a rather unusual scaling when the design variables $\mathbf{x}$ are close to the boundary $\partial\mathbb{R}^m_+$. While we obtained satisfactory results with automatic scaling in SNOPT, a specific scaling of the relaxed equilibrium constraint may be necessary for more robust convergence of the algorithm.

2.5.2 Numerical results: topology optimization, contact-less case

A number of "classic" weight minimization problems for trusses without frictionless contact but including stress, and possibly local buckling constraints and bounds on displacements have been solved to global optimality by Stolpe [Sto03]. We benchmark our relaxation algorithm against the results reported in the cited paper and find that in many cases our local algorithm is capable of finding globally optimal solutions. We keep the problem numbers assigned by Stolpe [Sto03] and report the results we obtained on a subset of these problems in Tables 2.1 (only stress and/or displacement constraints) and 2.2 (stress and local buckling constraints).

Since we use a local algorithm to solve non-convex optimization problems, starting the optimization procedure from different starting points may result in obtaining different optimal solutions. We started the algorithm from the design obtained by uniformly distributing structural material among bars, and calculating the corresponding equilibrium forces/displacements.

Some comments are in order. In problems 24–26 the number of bars in the structure is $m = 10$, but the volumes of 4 of them are fixed, which leaves us only 6 design variables. In addition, these are the only problems with displacement constraints, and the optimal weight we report differs from the known globally optimal solution despite the small value of the relaxation parameter ε we used. The reason for such a behaviour is that the potential energy $\Pi(\mathbf{x},\mathbf{u}) - \mathbf{f}^t\mathbf{u}$ becomes rather insensitive to some components of the displacements for designs $\mathbf{x}$ that are close to the boundary $\partial\mathbb{R}^m_+$. In problems 24–26 this allows the optimization procedure to choose displacements that are reasonably far from the equilibrium displacements (compared to the size of the relaxation parameter ε) but are feasible with respect to the imposed bounds on the displacements. (Recall, that Proposition 2.7 does not guarantee the convergence of the displacements as designs converge.) This may or may not be a problem in practice, depending on how stringent the displacement constraints are, if present. In particular, we guarantee the convergence of forces, and always keep the stress (and local buckling) constraints satisfied, which means that the structure will not suffer from destructive stresses. (Even though stress constraints are imposed not on the "equilibrium" stresses, stress bounds are usually chosen far from the point where plastic deformation occurs.) In any case, our algorithm successfully finds the optimal topology, which is of major importance in many applications.

In problem 17 our algorithm indeed finds a better solution to the classic 25-bar truss problem stated in [ScF74] than the one reported in [Sto03, ScF74]. The reason for this small victory of a local optimization algorithm over a global one is that the branch-and-cut method developed in [Sto03] may be applied only to problems with bounds imposed on all variables involved. In the original formulation of the problem 17 taken from [ScF74] there are no upper bounds on the volumes of the bars, and the optimal weight of the truss we obtained for the *original* formulation is 510.157. On the other hand, Stolpe [Sto03] imposes artificial bounds on the design variables for the branch-and-cut method to function, which leads to a globally optimal solution with the weight 545.264; in fact, the newly introduced bounds are inactive at the latter solution but owing to the non-convexity of the problem they cannot be safely removed without changing the optimal solution. The last comment about the problem 17 is that in the original formulation there are only 2 load

Problem	m	n	k	w_{our}	$w_{[\text{Sto03}]}$	source
2	5	4	2	39.9856	33.5000	[ChG97]
5	4	2	2	185.597	185.667	[Hob96]
9	10	8	1	4896.95	4898.31	[ScF74]
11	10	8	1	1583.99	1584.00	[ScF74]
13	10	8	1	4425.16	4426.52	[ScF74]
15	10	8	1	1655.99	1656.00	[ScF74]
*17	25	18	5	510.157	545.264	[ScF74]
18	10	8	1	1583.99	1584.00	[ChG97]
23	5	4	1	24.0000	24.0000	[ChJ92]
24	10(6)	8	1	18199.4	18211.8	[Kir90]
25	10(6)	8	1	20021.8	20035.3	[Kir90]
26	10(6)	8	1	22799.7	22817.3	[Kir90]
27	10	8	1	1979.99	1980.00	[GCY01]
28	5	4	2	79.9713	79.9716	[GCY01]

Tab. 2.1: Results of numerical experiments: weight minimization under stress and/or displacement constraints

Problem	m	n	k	w_{our}	$w_{[\text{Sto03}]}$	source
5	4	2	2	408.312	408.628	[Hob96]
27	10	8	1	8553.44	8553.44	[GCY01]
28	5	4	2	105.831	105.831	[GCY01]

Tab. 2.2: Results of numerical experiments: weight minimization under stress and local buckling constraints

scenarios and many linear constraints on the design variables related to the required symmetry of the truss. Instead, we consider all design variables to be independent and obtain a symmetric solution by introducing additional load cases.

2.5.3 Numerical results: sizing optimization of trusses in contact

Hilding et al. [HKP99] (see also Hilding [Hil00]) were interested in minimizing the maximal contact force, that is, to achieve as uniform contact pressures as possible. The formal problem statement can be written as follows:

$$(\Lambda) \begin{cases} \displaystyle\min_{(\mathbf{x},\mathbf{s},\boldsymbol{\lambda},\mathbf{u},\lambda_{\max})} \lambda_{\max}, \\ \text{s.t. } (\mathbf{x},\mathbf{s},\boldsymbol{\lambda},\mathbf{u}) \in \mathscr{F}(\mathbf{f}), \\ \qquad \underline{x}_i \le x_i \le \overline{x}_i, \quad i = 1,\dots,m, \\ \qquad w(\mathbf{x}) \le \overline{w}, \\ \qquad \lambda_\ell \le \lambda_{\max}, \quad \ell = 1,\dots,r, \end{cases}$$

where $\underline{x}_i$, $\overline{x}_i$, $\overline{w}$ are given positive numbers, $i = 1, \dots, m$. In general, allowing lower bounds on the design variables to be zero results in an ill-posed optimization problem, unless bounds on the compliance of the structure or stress constraints are added (see [PaP02]).

The problem (Λ) is thus not a topology optimization problem and does not suffer from the difficulties outlined in Section 2.4.1; in particular, the smoothing method of Facchinei [FJQ99] outlined in Section 2.3 is directly applicable to this problem (see [Hil00], where smoothing was used for "the heuristic avoiding of local minima") and we use it as one of the benchmarks for our new smoothing algorithm.

On some instances of the problem (Λ) Hilding et al. [HKP99] also implemented and tested some classic MPEC algorithms (IMPA [LPR96, Section 6.3] and PIPA [LPR96, Section 6.1]) on the family of structures shown in Figure 2.2. Also, they tested on (Λ) a very popular method in the structural optimization community: the method of moving asymptotes, MMA, [Sva87], even though it is not guaranteed to work on this problem.

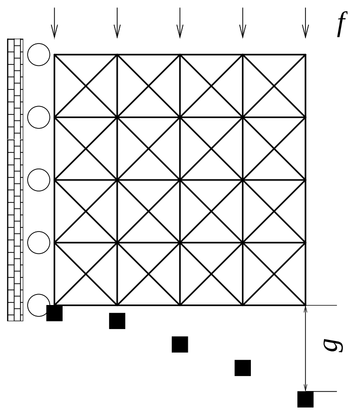


Fig. 2.2: Test problem found in [Hil00]. 5×5 case is shown.

We apply SNOPT to the following relaxation of the problem (Λ):

$$(\Lambda^\varepsilon) \begin{cases} \min\limits_{(\mathbf{x},\mathbf{s},\boldsymbol{\lambda},\mathbf{u},\lambda_{\max})} \lambda_{\max}, \\ \text{s.t. } (\mathbf{x},\mathbf{s},\boldsymbol{\lambda},\mathbf{u}) \in \mathscr{F}^\varepsilon(\mathbf{f}), \\ \qquad \underline{x}_i \leq x_i \leq \overline{x}_i, \quad i = 1,\dots,m, \\ \qquad w(\mathbf{x}) \leq \overline{w}, \\ \qquad \lambda_\ell \leq \lambda_{\max}, \quad \ell = 1,\dots,r, \end{cases}$$

where $\varepsilon > 0$ is a relaxation parameter. We report the results we obtained for trusses of different sizes (see Figure 2.2) in Table 2.3 along with the results found in [HKP99, Hil00].

We report the size of the structure, the number of bars (design variables) and the optimal values obtained by PIPA and IMPA/MMA as reported in [HKP99] (the two latter algorithms are reported to produce the same optimal values); the optimal values produced by IMPA and MMA as applied to the smoothed MPEC using the methodology introduced

Truss size	m	PIPA	IMPA/MMA	Smooth.+IMPA	Smooth.+MMA	New smoothing
3×3	58	-	3.0	1.0	1.0	1.0
4×4	113	-	2.0	1.0	1.0	1.0
5×5	190	1.67	2.5	1.0	1.25	1.07
10×10	875	2.5	5.0	-	-	2.0

Tab. 2.3: Results of numerical experiments: contact force minimization.

in [FJQ99], as reported in [Hil00]; and the optimal values obtained using our new smoothing procedure. The “-” sign in the table columns means that the corresponding algorithm has not been applied to a given problem instance.

One can see that our algorithm favourably competes with classic MPEC algorithms on these tests. As we already mentioned, general MPEC algorithms cannot be applied to truss optimization problems if we remove strictly positive lower bounds on the design variables, i.e., consider topology optimization problems.

2.6 Conclusions and further research

In this paper we proposed a new algorithm for solving MPEC problems arising from the topology optimization of trusses with unilateral contact conditions. The algorithm is based on the approximation of topology optimization problems with sizing-type problems, where in addition we relax the equilibrium constraint. We studied the convergence of global optimal solutions and stationary points to approximating problems towards, respectively, globally optimal solutions and stationary points to the original, singular problem. We have also performed some numerical testing of the proposed method.

Many open problems remain. On the numerical side, we need a better implementation (probably utilizing second order information); also, a much more thorough numerical testing should be done, especially for trusses with unilateral contact. However, in our opinion, the most challenging task is to improve the optimality conditions we obtained in this paper. To do that, the comparative analysis of modern KKT-type optimality conditions for general MPEC problems (see, e.g., [FlK02a, FlK02b, FlK02c]) and the FJ-type optimality conditions we obtained needs to be performed. We hope to address these questions in our future research.

Acknowledgement: This research is supported by the Swedish Research Council (grant 621-2002-5780).

References

[Ach98] W. Achtziger. Multiple-load truss topology and sizing optimization: Some properties of minimax compliance. *J. Optim. Theory Appl.*, 98(2):255–280, 1998.

[Ach03] W. Achtziger. On the optimality conditions and primal-dual methods for the detection of singular optima. In Carlo Cinquini, Marco Rovati, Paolo Venini, and Roberto Nascimbene, editors, *Short Papers of The Fifth World Congress of Structural and Multidisciplinary Optimization*, pages 1–2, 2003.

[AuF90] Jean-Pierre Aubin and Hélène Frankowska. *Set-Valued Analysis*. Birkhäuser Boston Inc., Boston, MA, 1990.

[BTKNZ99] A. Ben-Tal, M. Kočvara, A. Nemirovski, and J. Zowe. Free material design via semidefinite programming: the multiload case with contact conditions. *SIAM J. Optim.*, 9(4):813–832, 1999.

[ChG97] G. Cheng and X. Guo. ε-relaxed approach in structural topology optimization. *Struct. Optim.*, 13:258–266, 1997.

[ChJ92] G. Cheng and Z. Jiang. Study on topology optimization with stress constraints. *Eng. Opt.*, 20:129–148, 1992.

[DuB98] P. Duysinx and M. P. Bendsøe. Topology optimization with stress constraints. *Int. J. Numer. Mech. Engrg.*, 43:1453–1478, 1998.

[DuS98] P. Duysinx and O. Sigmund. New developments in handling stress constraints in optimal material distribution. In *Proc. 7th AIAA/USAF/NASA/ISSMO Symposium on Multidiciplinary Analysis and Optimization*, pages 1501–1509, 1998.

[EPP03] Anton Evgrafov, Michael Patriksson, and Joakim Petersson. Stochastic structural topology optimization: Existence of solutions and sensitivity analyses. *ZAMM Z. angew. Math. Mech.*, 83(7):479–492, 2003.

[Evg04] Anton Evgrafov. On globally stable singular topologies. To appear in *Struct. Multidisc. Optim.*, 2004.

[EvP03] Anton Evgrafov and Michael Patriksson. Stochastic structural topology optimization: Discretization and penalty function approach. *Struct. Multidisc. Optim.*, 25(3):174–188, 2003.

[FaP03] Francisco Facchinei and Jong-Shi Pang. *Finite-dimensional variational inequalities and complementarity problems. Vol. I*. Springer Series in Operations Research. Springer-Verlag, New York, 2003.

[FJQ99] Francisco Facchinei, Houyuan Jiang, and Liqun Qi. A smoothing method for mathematical programs with equilibrium constraints. *Math. Program.*, 85(1, Ser. A):107–134, 1999.

[FlK02a] Michael L. Flegel and Christian Kanzow. An Abadie-type constraint qualification for mathematical problems with equilibrium constraints. Preprint, University of Würzburg, Institute of Applied Mathematics and Statistics, November 2002.

[FlK02b] Michael L. Flegel and Christian Kanzow. On the Guignard constraint qualification for mathematical problems with equilibrium constraints. Preprint, University of Würzburg, Institute of Applied Mathematics and Statistics, October 2002.

[FlK02c] Michael L. Flegel and Christian Kanzow. Optimality conditions for mathematical programs with equilibrium constraints: Fritz John and Abadie-type approaches. Preprint, University of Würzburg, Institute of Applied Mathematics and Statistics, May 2002.

[GCY01] X. Guo, G. Cheng, and K. Yamazaki. A new approach for the solution of singular optima in truss topology optimization with stress and local buckling constraints. *Struct. and Multidisc. Optim.*, 22:364–372, 2001.

[GMS02] Philip E. Gill, Walter Murray, and Michael A. Saunders. SNOPT: an SQP algorithm for large-scale constrained optimization. *SIAM J. Optim.*, 12(4):979–1006, 2002.

[Hil00] Daniel Hilding. A heuristic smoothing procedure for avoiding local optima in optimization of structures subject to unilateral constraints. *Struct. Multidisc. Optim.*, 20(1):29–36, Aug 2000.

[HKP99] Daniel Hilding, Anders Klarbring, and Jong-Shi Pang. Minimization of maximum unilateral force. *Comput. Methods Appl. Mech. Engrg.*, 177(3-4):215–234, 1999.

[Hob96] A. S. Hoback. Optimization of singular problems. *Struct. Optim.*, 12:93–97, 1996.

[Kir90] U. Kirsch. On singular topologies in optimum structural design. *Struct. Optim.*, 2(3):133–142, 1990.

[LiM97] M. B. Lignola and J. Morgan. Stability of regularized bilevel programming problems. *J. Optim. Theory Appl.*, 93(3):575–596, 1997.

[LPR96] Zhi-Quan Luo, Jong-Shi Pang, and Daniel Ralph. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, Cambridge, 1996.

[Mic04] A. G. M. Michell. The limits of economy of material in frame structures. *Phil. Mag.*, 8, 1904.

[Mor76] B. Sh. Mordukhovich. Maximum principle in the problem of time optimal response with nonsmooth constraints. *Prikl. Mat. Meh.*, 40(6):1014–1023, 1976.

[OKZ98] Jiři Outrata, Michal Kočvara, and Jochem Zowe. *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*. Kluwer Academic Publishers, Dordrecht, 1998.

[Out94] J. V. Outrata. On optimization problems with variational inequality constraints. *SIAM Journal on Optimization*, 4:340–357, 1994.

[OuZ95] J. Outrata and J. Zowe. A numerical approach to optimization problems with variational inequality constraints. *Mathematical Programming*, 68:105–130, 1995.

[PaP02] Michael Patriksson and Joakim Petersson. Existence and continuity of optimal solutions to some structural topology optimization problems including unilateral constraints and stochastic loads. *ZAMM Z. Angew. Math. Mech.*, 82(7):435–459, 2002.

[Pet01] Joakim Petersson. On continuity of the design-to-state mappings for trusses with variable topology. *Internat. J. Engrg. Sci.*, 39(10):1119–1141, 2001.

[RoB94] G. I. N. Rozvany and T. Birker. On singular topologies in exact layout optimization. *Struct. Optim.*, 8:228–235, 1994.

[ScF74] L. A. Schmit and B. Farshi. Some approximation consepts for structural synthesis. *AIAA Journal*, 12(5):692–699, 1974.

[Sto03] M. Stolpe. *On Models and Methods for Global Optimization of Structural Topology*. PhD thesis, Optimization and Systems Theory, Department of Mathematics, Royal Institute of Technology, Stockholm, Sweden, 2003.

[StS01] M. Stolpe and K. Svanberg. On trajectories of the epsilon-relaxation approach for stress constrained truss topology optimization. *Struct. and Multidisc. Optim.*, 21(2):140–151, 2001.

[StS03] M. Stolpe and K. Svanberg. A note on stress-constrained truss topology optimization. *Struct. and Multidisc. Optim.*, 25(1):62–64, 2003.

[Sva87] Krister Svanberg. The method of moving asymptotes—a new method for structural optimization. *Internat. J. Numer. Methods Engrg.*, 24(2):359–373, 1987.

[Sva02] Krister Svanberg. A class of globally convergent optimization methods based on conservative convex separable approximations. *SIAM J. Optim.*, 12(2):555–573, 2002.

[SvG68] G. Sved and Z. Ginos. Structural optimization under multiple loading. *Int. J. Mech. Sci.*, 10:803–805, 1968.

Paper 3

ON THE LIMITS OF POROUS MATERIALS IN THE TOPOLOGY OPTIMIZATION OF STOKES FLOWS

Anton Evgrafov*

Abstract

We consider a problem concerning the distribution of a solid material in a given bounded control volume with the goal to minimize the potential power of the Stokes flow with given velocities at the boundary through the material-free part of the domain. We also study the relaxed problem of the optimal distribution of the porous material with a spatially varying Darcy permeability tensor, where the governing equations are known as the Darcy-Stokes, or Brinkman, equations. We show that the introduction of the requirement of zero power dissipation due to the flow through the porous material into the relaxed problem results in it becoming a well-posed mathematical problem, which admits optimal solutions that have extreme permeability properties (i.e., assume only zero or infinite permeability); thus, they are also optimal in the original (non-relaxed) problem.

Two numerical techniques are presented for the solution of the constrained problem. One is based on a sequence of optimal Brinkman flows with increasing viscosities, from the mathematical point of view nothing but the exterior penalty approach applied to the problem. Another technique is more special, and is based on the "sizing" approximation of the problem using a mix of two different porous materials with high and low permeabilities, respectively.

This paper thus complements the study of Borrvall and Petersson [Internat. J. Numer. Methods Fluids, vol. 41, no. 1, pp. 77–107, 2003], where only sizing optimization problems are treated.

Keywords. Topology optimization, Fluid mechanics, Stokes flow.

AMS subject classification. 49J20, 49J45, 76D55, 62K05.

---

*Department of Mathematics, Chalmers University of Technology, SE-412 96 Göteborg, Sweden, e-mail: toxa@math.chalmers.se

3.1 Introduction

WHILE topology optimization of structures in (very) rough terms can be described as the science of introducing holes in the structure so as to improve the structural performance, in the vast majority of the literature on the subject, especially computationally-oriented one, the appearence of holes is *precluded* from the very beginning by the requirement that the minimal structural dimension is positive at every point.

The reason for introducing such a constraint is two-fold. From the numerical point of view, the FEM-stiffness matrix of the governing differential equation is guaranteed to be positive definite in this case, resulting in stable numerical procedures. However, more importantly, allowing some structural parts to disappear we often end up with an optimization problem having a non-closed feasible set and, as a result, lacking optimal solutions.

In topology optimization of solids and structures the classic problem of minimizing the structural compliance is known to possess optimal solutions, if we allow microstructures to be used in the optimal structure (cf. [BeS03, Appendix 5.2]). At the same time, if we are interested in a pure solid–void design, free of microstructures, the same problem lacks optimal solutions. Since the "grey" optimal solutions (the ones involving micostructures, as opposed to "black–white" pure solid–void solutions) are usually difficult to interpret and to manufacture, various restriction or regularization methods are considered in order to reduce the amount of the "microstructural material" in the optimal structure; see the bibliographical notes (8) in [BeS03]. The pure void parts, the very heart of the topology optimization, are not allowed to appear in such methods and are usually modelled by a very compliant material. However, the limits of optimal designs as the properties of the compliant substitute approach those of void are not investigated.

In the case of topology optimization of truss structures, the question of the continuity of the optimal solutions w.r.t. the lower bound on the minimal structural dimension has received significant attention in the literature (see, e.g., the bibliographical notes (16) in [BeS03] on the "stress singularity phenomenon"). Despite the abundant amount of literature on the topology optimization of linearly elastic continuous systems, similar studies have not been conducted in this case.

Recently, topology optimization techniques have been applied to optimization problems in flow mechanics [BoP03], where traditionally shape optimization methods were prevealing (see the pioneering works of Pironneau [Pir73, Pir74] on the optimality conditions for shape optimization in fluid mechanics; see also the bibliographical notes (2) in [BeS03] for classical references). The benefits of using topology optimization (or control in coefficients) over shape optimization include easier implementation and sensitivity analysis, and better integration with existing FEM codes. Borvall and Petersson [BoP03] considered the optimal design of flow domains for minimizing the total power of the Stokes flows. The set of admissible designs is a set of porous materials with a spatially varying Darcy permeability tensor, under a constraint on the total volume of fluid in the control region. The appearence of internal walls in the domain (regions with pure solid material, not permitting flow; these can be interpreted as "holes in the flow") is not permitted. Thus, the *topology*, i.e., connectivity of the flow region does not change, and, carrying over the terminology from optimization in solid mechanics, we will refer to this case as that of a "sizing" optimization.

In the present paper we study the “real” topology optimization case of the Stokes flow, i.e., pure solid and pure flow regions are allowed. We show that the relaxed problem of distributing porous material, as well as the pure solid–void (zero–one) problem, possesses optimal solutions. Furthermore, we show that the sizing optimal solutions have limits as the permeability of the porous material is allowed to vanish (i.e., converge to the permeability of solid material).

The outline of the present paper is as follows. In the next section, we describe the necessary notation and state precisely the weak formulation of the governing equations, its interpetation, and the objective functional. Section 3.3 is dedicated to the proof of the existence of the optimal solutions to the relaxed problem, while in Section 3.4 we introduce a well-posed formulation of the zero-one optimal problem and establish the well-posedness of the latter. Two numerical approaches for the solution of the zero-one control problem are the topics of Sections 3.5 and 3.6. In Section 3.7, we show that for functionals other than the total power of the flow, the control problem might be ill-posed, even if rather strong continuity requirements are imposed on the objective functional. We end the paper with a brief discussion of further research topics.

3.2 Prerequisutes

3.2.1 Notation

We follow standard engineering practice and will denote vector quantities, such as vectors and vector-valued functions, using the **bold** font. However, for functional spaces of both scalar- and vector-valued functions we will use regular font.

Let Ω be a connected bounded domain of $\mathbb{R}^d$, $d \in \{2,3\}$ with a Lipschitz continuous boundary Γ. In this domain we would like to control the Darcy-Stokes, or Brinkman, equations [NiB99] with the prescribed flow velocities $\mathbf{g}$ on the boundary, and forces $\mathbf{f}$ acting in the domain by adjusting the inverse permeability α of the medium occupying Ω, which depends on the control function ρ:

$$\begin{cases} \left.\begin{aligned} -\nu\Delta\mathbf{u} + \alpha(\rho)\mathbf{u} + \nabla p = \mathbf{f}, \\ \operatorname{div}\mathbf{u} = 0 \end{aligned}\right\}, & \text{in } \Omega, \\ \mathbf{u} = \mathbf{g}, & \text{on } \Gamma. \end{cases} \tag{3.1}$$

In the system (3.1), $\mathbf{u}$ is a flow velocity, p is a pressure, and ν is a kinematic viscosity. Of course, the function $\mathbf{g}$ must satisfy the compatibility condition

$$\int_\Gamma \mathbf{g}\cdot\mathbf{n} = 0,$$

where $\mathbf{n}$ denotes the outward unit normal. If $\alpha(\rho(x)) = +\infty$ for some $x \in \Omega$, we simply require $\mathbf{u}(x) = 0$ in the first equation of (3.1).

Our control set $\mathscr{H}$ is defined as follows:

$$\mathscr{H} = \{\, \rho \in L^\infty(\Omega) \mid 0 \le \rho \le 1, \text{a.e. in } \Omega, \int_\Omega \rho \le \gamma|\Omega| \,\},$$

where $0 < \gamma < 1$ is the maximal volume fraction that can be occupied by the fluid. Every element $\rho \in \mathscr{H}$ describes the scaled Darcy permeability tensor of the medium at a given point $x \in \Omega$ in the following (informal) way: $\rho(x) = 0$ corresponds to zero permeability at x (i.e., solid, which does not permit any flow at a given point), while $\rho(x) = 1$ corresponds to infinite permeability (i.e., 100% flow region; no structural material is present). Formally, we relate the permeability α^{-1} to ρ using a convex, decreasing, and nonnegative function $\alpha : [0,1] \to \mathbb{R}_+ \cup \{+\infty\}$, defined as

$$\alpha(\rho) = \rho^{-1} - 1.$$

Modelling the Stokes flow, we are interested only in the two extreme values of permeability, $\alpha^{-1} = 0$, or $\alpha^{-1} = +\infty$. For this purpose, we will introduce the following subset of $\mathscr{H}$:

$$\widetilde{\mathscr{H}} = \{\, \rho \in \mathscr{H} \mid \rho \in \{0,1\}, \text{a.e. in } \Omega \,\}.$$

However, both from the analytical and computational points of view, it is impossible to state the control problem in the set $\widetilde{\mathscr{H}}$, because it is nonconvex, and not weakly* closed. Therefore, we first study the properties of the relaxed control problem posed over the set $\mathscr{H}$.

In the rest of the paper, we will use the symbol χ_A for $A \subset \Omega$ to denote the characteristic function of A: $\chi_A(\mathbf{x}) = 1$ for $\mathbf{x} \in A$; $\chi_A(\mathbf{x}) = 0$ otherwise. Also, for $\mathbf{u} \in H^1(\Omega)$, we define a set $\Omega_{\mathrm{nz}}(\mathbf{u}) := \{\, \mathbf{x} \in \Omega \mid \mathbf{u}(x) \neq \mathbf{0} \,\}$. Finally, let $\Omega(\mathbf{u}) \subset \Omega$ be such that (i) $\mathbf{x} \in \Omega(\mathbf{u}) \implies \forall \mathbf{y} \in \Omega \setminus \Omega(\mathbf{u}) : \mathbf{u}(\mathbf{x}) \cdot \mathbf{u}(\mathbf{x}) \geq \mathbf{u}(\mathbf{y}) \cdot \mathbf{u}(\mathbf{y})$, and (ii) $|\Omega(\mathbf{u})| = \gamma|\Omega|$.

3.2.2 Weak formulation

To state the problem in a more analytically suitable way, and to incorporate the special case $\alpha = +\infty$ into the first equation of the system (3.1), we introduce a weak formulation of the equations. Let us consider the set of admissible flow velocities and test functions

$$\begin{aligned} \mathscr{U} &= \{\, \mathbf{v} \in H^1(\Omega) \mid \mathbf{v} = \mathbf{g} \text{ on } \Gamma \,\}, \\ \mathscr{V} &= \{\, \mathbf{v} \in H^1(\Omega) \mid \mathbf{v} = \mathbf{0} \text{ on } \Gamma \,\}, \end{aligned}$$

and pressures

$$L^2_0(\Omega) = \{\, q \in L^2(\Omega) \mid \int_\Omega q = 0 \,\}.$$

Then, the weak formulation of (3.1) reads as follows: for $\mathbf{f} \in L^2(\Omega)$, compatible $\mathbf{g} \in H^{1/2}(\Gamma)$, and $\rho \in \mathscr{H}$ find $(\mathbf{u}, p) \in \mathscr{U} \times L^2_0(\Omega)$ such that

$$\begin{aligned} \nu \int_\Omega \nabla\mathbf{u} \cdot \nabla\mathbf{v} + \int_\Omega \alpha(\rho)\mathbf{u} \cdot \mathbf{v} - \int_\Omega p \operatorname{div} \mathbf{v} = \int_\Omega \mathbf{f} \cdot \mathbf{v}, \qquad & \forall \mathbf{v} \in \mathscr{V}, \\ \int_\Omega q \operatorname{div} \mathbf{u} = 0, \qquad & \forall q \in L^2_0(\Omega) \end{aligned} \tag{3.2}$$

(In the system above we use the usual convention $\infty \cdot 0 = 0$.)

Allowing designs with zero permeability significantly increases the complexity of the control problem. From the purely technical side, the inverse permeability α may be infinite on sets of positive measure, and thus does not belong to any of the common functional spaces. Even worse, internal walls that do not permit flows with the given boundary conditions might appear as limits of admissible designs, making the design space not closed. The latter difficulty is demonstrated in the following example.

Example 3.1 (Diminishing permeability). Let $\mathbf{g}$ be some compatible non-zero boundary condition, $\mathbf{f}$ be arbitrary in $L^2(\Omega)$. Let $\rho_k \equiv 1/k$ in Ω, $\rho \equiv 0$ in Ω, so that $\rho_k \to \rho$ in $L^\infty(\Omega)$ as $k \to \infty$. It is not difficult to check (this follows from the standard theory for the Stokes equations as well as from the results in [BoP03]) that for each $k = 1, 2, \ldots$, there is a solution $(\mathbf{u}_k, p_k)$ to (3.2). However, since $\alpha(\rho) \equiv +\infty$ in Ω, from the first equation in (3.2) it follows that $\mathbf{u} \equiv \mathbf{0}$ in Ω, which is clearly not compatible with the boundary conditions. In other words, there is no solution $(\mathbf{u}, p)$ to (3.2) corresponding to the limiting control ρ, which means that the set of admissible controls is not closed even in the strong topology of $L^\infty(\Omega)$!

This is in vast contrast with the sizing case, which can be modelled by requiring $\underline{\rho} \le \rho \le \overline{\rho}$, a.e. in Ω, for some constants $0 < \underline{\rho} \le \overline{\rho} \le 1$. Under these conditions, Borrvall and Petersson [BoP03] show that the set of admissible controls is closed in the weak* toplogy of $L^\infty(\Omega)$. (In fact, the case $\rho = 1$ or $\alpha = 0$ is not allowed in the cited work; however, the arguments used there work for this case as well because, owing to Fredrichs' inequality, the semi-norm $|\cdot|_1$ is equivalent to the norm of $H^1(\Omega)$ in the problem we consider; see also Theorem 3.4).

Example 3.1 demonstrates that the lower semicontinuity of the objective functional alone is not sufficient for the topology optimization of the Darcy-Stokes flow to possess optimal solutions; e.g., take the problem of minimizing the "volume of the flow" $\int_\Omega \rho$ to recover a situation similar to that of Example 3.1. Hovewer, if the objective functional also enjoys an inf-compactness property w.r.t. the set of admissible controls, every minimizing sequence converges, thus making the control problem well-posed. In what follows we establish that the power functional, introduced below, for the Darcy-Stokes flow is both lower semi-continuous and inf-compact, thus extending the results of [BoP03] from sizing to topology optimization.

Let $\mathcal{J}^{\mathcal{S}} : \mathcal{U} \to \mathbb{R}$ denote the potential power of the Stokes flow:

$$\mathcal{J}^{\mathcal{S}}(\mathbf{u}) = \frac{1}{2}\int_\Omega \nabla\mathbf{u} \cdot \nabla\mathbf{u} - \int_\Omega \mathbf{f} \cdot \mathbf{u}.$$

Let us further define the additional power dissipation $\mathcal{J}^{\mathcal{D}} : \mathcal{H} \times \mathcal{U} \to \mathbb{R} \cup \{+\infty\}$, due to the presence of the porous medium:

$$\mathcal{J}^{\mathcal{D}}(\rho, \mathbf{u}) = \frac{1}{2}\int_\Omega \alpha(\rho)\mathbf{u} \cdot \mathbf{u}.$$

Finally, let $\mathcal{J}(\rho, \mathbf{u}) = \mathcal{J}^{\mathcal{S}}(\mathbf{u}) + \mathcal{J}^{\mathcal{D}}(\rho, \mathbf{u})$ denote the total power of the Darcy-Stokes flow.

Assuming $\alpha(\rho) < +\infty$, one can derive the variational formulation of the system (3.1) (cf. [BoP03]):

$$\begin{aligned} \phi(\rho) = \min_{\mathbf{u}\in\mathcal{U}} \mathcal{J}(\rho, \mathbf{u}), \\ \text{s.t. } \operatorname{div}\mathbf{u} = 0, \ \text{weakly in } \Omega, \end{aligned} \tag{3.3}$$

the system (3.2) being the first order necessary optimality conditions for (3.3). In particular, the pressure $p \in L^2_0(\Omega)$ is defined as a Lagrange multiplier for the constraint $\operatorname{div}\mathbf{u} = 0$. In what follows, we will denote the feasible set of the problem (3.3) by $\mathscr{U}_{\text{div}}$.

Now, assume that for a given $\rho \in \mathscr{H}$ there is a solution $\mathbf{u} \in H^1(\Omega)$ to the variational problem (3.3). Define the new domain $\Omega_{\text{nz}}(\mathbf{u}) = \{\mathbf{x} \in \Omega \mid \mathbf{u}(\mathbf{x}) \neq \mathbf{0}\}$. Clearly, $\alpha < +\infty$, a.e. in $\Omega_{\mathbf{u}}$, and $\mathbf{u}$ solves the variational problem (3.3) in the domain $\Omega_{\text{nz}}(\mathbf{u})$ with the boundary conditions $\mathbf{u} = \mathbf{g}$ on Γ, $\mathbf{u} = \mathbf{0}$ on $\partial\Omega_{\text{nz}}(\mathbf{u}) \setminus \Gamma$. Therefore, there must be an associated pressure $p : \Omega_{\text{nz}}(\mathbf{u}) \to \mathbb{R}$ such that the pair $(\mathbf{u}, p)$ solves the weak formulation of the Darcy-Stokes equation in the domain $\Omega_{\text{nz}}(\mathbf{u})$ with the already described boundary condition. (In particular, if $\alpha = 0$ a.e. in $\Omega_{\text{nz}}(\mathbf{u})$, then $(\mathbf{u}, p)$ is a weak solution to the Stokes equation in the domain $\Omega_{\text{nz}}(\mathbf{u})$.) With this interpretation, we will use the variational formulation (3.3) of the problem instead of (3.2) in the development that follows.

3.2.3 Objective functional

The objective functional in our problem will be to minimize the total potential power of the flow, which in the case of $\mathbf{f} = \mathbf{0}$ amounts to minimizing the power dissipated by the flow. (The same problem can be interpreted as a minimization of the average pressure drop, provided $\mathbf{f} = \mathbf{0}$ and $\mathbf{g} = g\mathbf{n}$ [BoP03].)

Therefore, the optimization problem we consider can be written as follows:

$$\min_{\rho \in \mathscr{H}} \phi(\rho), \tag{3.4}$$

where $\phi : \mathscr{H} \to \mathbb{R} \cup \{+\infty\}$ is defined in (3.3).

As has been announced above, with this functional the control problem (3.4) possesses optimal solutions *despite* the fact that the set of admissible controls is not closed (see Corollary 3.5). Furthermore, in contrast to the situation in the case of linear elasticity, the "discrete" problem of minimizing the total power of the Stokes flow with controls in $\widetilde{\mathscr{H}}$ possesses optimal solutions. However, special approximation techniques are necessary to find them (see Sections 3.5 and 3.6).

3.3 Existence of optimal solutions

In this section we prove that the problem (3.4) admits optimal solutions; see Theorem 3.4 and its Corollary. However, we need a few auxiliary results first.

Proposition 3.2. *The function $h : [0,1] \times \mathbb{R}^d \to \mathbb{R}_+ \cup \{+\infty\}$ defined as $h(x, \mathbf{y}) = x^{-1}\mathbf{y} \cdot \mathbf{y}$ with the conventions $0^{-1} = +\infty$ and $+\infty \cdot 0 = 0$ is convex and lower semicontinuous.*

Proof. The proof is elementary and can be found in [Roc70, p. 83]. □

Lemma 3.3. *Let $\{(\rho_k, \mathbf{u}_k)\} \subset \mathscr{H} \times \mathscr{U}_{\text{div}}$ be such that:*
- $\liminf_{k\to+\infty} \mathscr{J}^{\mathscr{D}}(\rho_k, \mathbf{u}_k) = C$*, for some constant* $C < +\infty$*;*
- $\text{w}^*\text{-}\lim_{k\to+\infty} \rho_k = \rho$ *in* $L^\infty(\Omega)$*;*
- $\text{w-}\lim_{k\to\infty} \mathbf{u}_k = \mathbf{u}$ *in* $H^1(\Omega)$.

Then, the pair $(\rho,\mathbf{u}) \in \mathscr{H} \times \mathscr{U}_{\mathrm{div}}$, *and* $\mathscr{J}^{\mathscr{D}}(\rho,\mathbf{u}) \le C$.

Proof. The first claim is obvious.

Since $\int_\Omega \alpha(\rho_k)\mathbf{u}_k \cdot \mathbf{u}_k = \int_\Omega h(\rho_k,\mathbf{u}_k) - \int_\Omega \mathbf{u}_k \cdot \mathbf{u}_k$, where h is defined in Proposition 3.2, and the last integral converges to $\int_\Omega \mathbf{u}\cdot\mathbf{u}$, it remains to estimate $\liminf_{k\to+\infty}\int_\Omega h(\rho_k,\mathbf{u}_k)$. The weak lower semicontinuity of $(\rho,\mathbf{u}) \mapsto \int_\Omega h(\rho,\mathbf{u})$ follows from the (pointwise) convexity and lower semicontinuity of h (Proposition 3.2), Fatou's Lemma and Corollary 2.2 in [EkT99]. $\square$

Now we are ready to establish the existence result.

Theorem 3.4 (Existence of optimal solutions). *The optimization problem*

$$\min_{(\rho,\mathbf{u})\in\mathscr{H}\times\mathscr{U}_{\mathrm{div}}} \mathscr{J}(\rho,\mathbf{u}). \tag{3.5}$$

possesses at least one optimal solution $(\rho^*,\mathbf{u}^*)$.

Proof. Let $\mathbf{u}_0$ be the solution to the Stokes problem in Ω (i.e., the solution to (3.3) corresponding to $\rho \equiv 1$ in Ω); set $\rho_0 \equiv \gamma/|\Omega|$. Then $(\rho_0,\mathbf{u}_0) \in \mathscr{H}\times\mathscr{U}_{\mathrm{div}}$, and $\mathscr{J}(\rho_0,\mathbf{u}_0) < +\infty$. Furthermore, for all $(\rho,\mathbf{u}) \in \mathscr{H}\times\mathscr{U}_{\mathrm{div}}$ it holds that $\mathscr{J}(\rho,\mathbf{u}) \ge \mathscr{J}(1,\mathbf{u}_0) > -\infty$, i.e., the problem (3.5) is feasible and $\mathscr{J}$ is proper w.r.t. its feasible set.

The set $\mathscr{H}$ is weakly* compact in $L^\infty(\Omega)$, and the set $\mathscr{U}_{\mathrm{div}}$ is weakly closed in $H^1(\Omega)$.

Owing to the weak lower semicontinuity of $\mathscr{J}^{\mathscr{S}}$ in $H^1(\Omega)$ (cf. [Dac89, Theorem 2.3]), and lower semicontinuity of $\mathscr{J}^{\mathscr{D}}$ in the weak$^*\times$weak topology of $L^\infty(\Omega)\times H^1(\Omega)$ (cf. Lemma 3.3), it remains to show that every minimizing sequence $\{(\rho_k,\mathbf{u}_k)\}$ of (3.5) has bounded second components.

The valid inequality $+\infty > \limsup_{k\to+\infty}\mathscr{J}(\rho_k,\mathbf{u}_k) \ge \limsup_{k\to+\infty}\mathscr{J}(1,\mathbf{u}_k) = \limsup_{k\to+\infty}\mathscr{J}^{\mathscr{S}}(\mathbf{u}_k)$ implies that $\{|\mathbf{u}_k|_1\}$ is bounded. Since Ω is bounded, and $\mathbf{u}_k|_\Gamma = \mathbf{g}$, Fredrichs' inequality implies that $\{\|\mathbf{u}_k\|\}$ is bounded. $\square$

Corollary 3.5. *The optimization problem* (3.4) *possesses at least one optimal solution.*

Proof. Let $(\rho^*,\mathbf{u}^*)$ be optimal solution to (3.5); then, ρ^* is optimal in (3.4). $\square$

3.4 Existence of black-white solutions

From the engineering point of view, it is important to find optimal solutions to the problem (3.4) that also lie in $\widetilde{\mathscr{H}}$. Such optimal solutions are traditionally called *zero-one*, or *black-white*, solutions in the topology optimization literature. Zero-one optimal solutions are easy to interpret and to manufacture (e.g., one does not need to include microstructures into the final design in linear elasticity, or materials with varying porosity in Darcy-Stokes flow mechanics).

Let $\mathbf{u}^*$ be a flow that is optimal in the problem (3.4). We can always obtain an optimal control ρ^* for this flow as a solution to the following opimization problem:

$$\min_{\rho\in\mathscr{H}} \mathscr{J}(\rho,\mathbf{u}^*). \tag{3.6}$$

For the problem (3.6) to admit optimal solutions at the extreme points of the control set $\mathscr{H}$, i.e., in $\widetilde{\mathscr{H}}$, it is necessary for the inverse permeability α to depend on ρ in a *concave* way. At the same time, the lower semicontinuity of the objective functional $\mathscr{J}$ depends on the fact that α (in fact, h, cf. Lemma 3.3) depends on its arguments in a *convex* manner. Clearly, there is no function mapping $[0,1]$ onto $[0,+\infty]$ satisfying both requirements. Therefore, we need to specify the requirement that there must be at least one solution to (3.4) in $\widetilde{\mathscr{H}}$ as an additional constraint. As will be shown in Theorem 3.6, this can be achieved by adding a requirement of zero energy dissipation due to the flow through the porous material, i.e., $\mathscr{J}^{\mathscr{D}}(\rho,\mathbf{v}) = 0$.

On the other hand, in the case of the sizing optimization problems considered in [BoP03], the design space $\mathscr{H}$ describes inverse permeabilities α which belong to the bounded subset $\{0 < \underline{\alpha} \leq \alpha \leq \overline{\alpha} < +\infty\}$ of $L^\infty(\Omega)$. Therefore, one has a freedom to choose an affine mapping (that is, both convex and concave) $\alpha^{(\ell)}(\rho) = \overline{\alpha} + (\underline{\alpha} - \overline{\alpha})\rho$ to describe the dependence of the inverse permeability on the design; with such a choice, there is always an optimal solution $\rho^* \in \widetilde{\mathscr{H}}$ to the sizing optimization problem (cf. Corollary 3.1 in [BoP03]). Hovever, the zero-one optimal solutions obtained in [BoP03] are not black-white in the traditional interpretation (i.e., black denotes solid material, and white is its opposite: void in linear elasticity, or flow region in flow mechanics), but rather "dark-grey – light-grey"! Namely, they are composed of two porous materials with high and low permeabilities, respectively. A priori, it is not clear how close they are to the real black-white solutions (if any of the latter exist).

Therefore, our further goals are as follows. In this section, we show how to set up, in an analytically suitable manner, an optimization problem for minimizing the potential power of the Stokes flow that possesses black-white solutions. This problem is not suitable for numerical computations though, because the zero-one solution requirement is posed as a complementarity condition between the inverse permeability and the velocity of the flow. (Complementarity conditions are known to generate highly non-convex feasible sets, which often violate standard constraint qualifications [LPR96] and are therefore extremely hard to solve to global or even local optimality.) As a remedy, in the two subsequent sections we propose two computational approaches to the zero-one problem: one is based on a penalty function, with the viscosity of the flow playing the role of a penalty parameter; the other one is based on the aforementioned "dark-grey – light-grey" approximations.

Theorem 3.6 (Existence of 0–1 solutions). *The optimization problem*

$$\begin{cases} \min\limits_{(\rho,\mathbf{u})\in\mathscr{H}\times\mathscr{U}_{\text{div}}} \mathscr{J}^{\mathscr{S}}(\mathbf{u}), \\ s.t. \quad \mathscr{J}^{\mathscr{D}}(\rho,\mathbf{u}) = 0, \end{cases} \tag{3.7}$$

possesses at least one optimal solution $(\widetilde{\rho},\mathbf{u}^*) \in \widetilde{\mathscr{H}} \times \mathscr{U}_{\text{div}}$.

Proof. The constraint of the problem (3.7) can be equivalently written as $\mathscr{J}^{\mathscr{D}}(\rho,\mathbf{u}) \leq 0$, which, together with Lemma 3.3, implies the closedness of the feasible set of the problem (3.7) in the weak* $\times$ weak topology of $L^\infty(\Omega) \times H^1(\Omega)$. Therefore, following the proof of Theorem 3.4, we can establish existence of the optimal solution $(\rho^*,\mathbf{u}^*) \in \mathscr{H} \times \mathscr{U}_{\text{div}}$, provided there is at least one solution that is feasible in (3.7).

To construct a feasible solution, we choose a closed set $\Omega_0 \subset \Omega$, such that $|\Omega_0| = (1-\gamma)|\Omega|$ and $\Omega \setminus \Omega_0$ is connected and has a Lipschitz continuous boundary. Let $\mathbf{u}^{\mathscr{S}}$ be the Stokes flow in $\Omega \setminus \Omega_0$ with boundary conditions $\mathbf{u}^{\mathscr{S}} = \mathbf{g}$ on Γ and $\mathbf{u}^{\mathscr{S}} = \mathbf{0}$ on $\partial(\Omega \setminus \Omega_0) \setminus \Gamma$; set $\rho^{\mathscr{S}} = \chi_{\Omega\setminus\Omega_0}$. When, $\mathscr{J}^{\mathscr{S}}(\mathbf{u}^{\mathscr{S}}) < +\infty$ and $\mathscr{J}^{\mathscr{D}}(\rho^{\mathscr{S}}, \mathbf{u}^{\mathscr{S}}) = 0$.

Now, let $\widetilde{\rho} = \chi_{\Omega_{\mathrm{nz}}(\mathbf{u}^*)}$, where $\Omega_{\mathrm{nz}}(\mathbf{u}^*) = \{\, \|\mathbf{u}^*\| \neq 0 \,\}$. Then, $\int_\Omega \widetilde{\rho} \le \int_\Omega \rho^*$ and $\mathscr{J}^{\mathscr{D}}(\widetilde{\rho}, \mathbf{u}^*) = 0$, yielding an optimal solution $(\widetilde{\rho}, \mathbf{u}^*) \in \widetilde{\mathscr{H}} \times \mathscr{U}_{\mathrm{div}}$. □

We would like to stress the fact that, owing to Theorem 3.6, for every optimal solution to (3.7), there is an optimal solution to the following zero-one problem

$$\begin{cases} \displaystyle\min_{(\rho,\mathbf{u})\in L^\infty(\Omega)\times\mathscr{U}_{\mathrm{div}}} \mathscr{J}^{\mathscr{S}}(\mathbf{u}), \\ \text{s.t.} \begin{cases} \rho(x) = 0 \implies \mathbf{u}(x) = \mathbf{0}, \text{a.e. in } \Omega, \\ \mathbf{u}(x) \neq \mathbf{0} \implies \rho(x) = 1, \text{a.e. in } \Omega, \\ \displaystyle\int_\Omega \rho \le \gamma|\Omega|, \end{cases} \end{cases} \tag{3.8}$$

having the same objective value. Therefore, every optimal solution to (3.8) is also optimal in (3.7). In this sense, the problems (3.8) and (3.7) are *equivalent*, i.e., neither one is a relaxation nor a restriction of the other. Such an equivalence is a very important and unique fact about the topology optimization of Stokes flows. We recall that the zero-one problem "as is" in linear elasticity is ill-posed, and either relaxation or restriction is *necessary* to guarantee the existence of optimal solutions (cf. the biliographical notes (8) in [BeS03] for an extensive account of relaxation and restriction methods in topology optimization in solid mechanics).

3.5 Black-white solutions via increasing the viscosity

There is a school of thought arguing that under some circumstances the viscosity ν and permeability α^{-1} in the system (3.1) alone do not adequately describe the Stokes flow in porous media. An additional parameter μ is introduced into the first PDE as follows [NiB99]:

$$-\nu\boldsymbol{\Delta}\mathbf{u} + \mu\alpha(\rho)\mathbf{u} + \boldsymbol{\nabla} p = \mathbf{f}.$$

Now, the parameter μ is the viscosity of the flow, while ν is an "effective viscosity". Repeating the arguments of Section 3.1, we then arrive at the following formulation of the optimization problem (3.4):

$$\min_{(\rho,\mathbf{v})\in\mathscr{H}\times\mathscr{U}_{\mathrm{div}}} \mathscr{J}^{\mathscr{S}}(\mathbf{v}) + \mu\,\mathscr{J}^{\mathscr{D}}(\rho,\mathbf{v}). \tag{3.9}$$

Clearly, this is nothing but the exterior penalty reformulation of the problem (3.7), with the viscosity μ playing the role of a penalty parameter. The arguments of Theorem 3.4 are applicable to the problem (3.9) as well, so that there exists a family of optimal solutions $\{\rho_\mu^*, \mathbf{u}_\mu^*\}$, $\mu > 0$ to (3.9). From the standard theory for nonlinear programs (cf. Theorem 9.2.2, [BSS93]), it follows that every weak* $\times$ weak limit point of this sequence as $\mu \to +\infty$ (and there is at least one) is an optimal solution to (3.7).

We note that the problem (3.9) does not contain any complicating state constraints, and thus is much easier to solve than (3.7). While the penalty method might converge quite slowly, and the approximating designs might contain quite a large amount of porous material with intermediate values of permeability, we think it is instructive to mention this approach, owing to its clear mathematical and physical interpretations (compare with, e.g., the most popular "SIMP" approach [BeS99] in the topology optimization of elastic materials, or the more material science-compatible "RAMP" method [StS01]; see also the discussion in [BeS03, p. 64]).

3.6 Black-white solutions as limits of "dark-grey – light-grey" solutions

In this section we will approximate the zero-one problem (3.7) using the aforementioned two-value "dark-grey – light-grey" optimal controls obtained in [BoP03]. To perform such an approximation, we introduce two sequences, $\{\underline{\alpha}_k\} \downarrow 0$ and $\{\overline{\alpha}_k\} \uparrow +\infty$, of extreme inverse permeabilities. Further, we let $\underline{\rho}_k = (\overline{\alpha}_k + 1)^{-1}$, $\overline{\rho}_k = (\underline{\alpha}_k + 1)^{-1}$, and define an affine function $\alpha^{(\ell,k)} : [\underline{\rho}_k, \overline{\rho}_k] \to \mathbb{R}_+$ so that $\alpha^{(\ell,k)}(\underline{\rho}_k) = \overline{\alpha}_k$, $\alpha^{(\ell,k)}(\overline{\rho}_k) = \underline{\alpha}_k$. To simplify the discussion somewhat, we assume that the sequence $\{(\underline{\alpha}_k, \overline{\alpha}_k)\}$ is chosen so that the inequality $\overline{\rho}_k \gamma + \underline{\rho}_k(1-\gamma) \le \gamma$ is satisfied. Then, we can also define the approximating control sets $\mathscr{H}_k = \{\, \rho \in \mathscr{H} \mid \underline{\rho}_k \le \rho \le \overline{\rho}_k, \text{a.e. in } \Omega \,\}$, and $\widetilde{\mathscr{H}_k} = \{\, \rho \in \mathscr{H} \mid \rho \in \{\underline{\rho}_k, \overline{\rho}_k\}, \text{a.e. in } \Omega \,\}$. Finally, we define $\mathscr{J}_k^{\mathscr{D}}(\rho, \mathbf{v}) = 1/2 \int_\Omega \alpha^{(\ell,k)}(\rho)\mathbf{v}\cdot\mathbf{v}$, and $\mathscr{J}_k(\rho,\mathbf{v}) = \mathscr{J}^{\mathscr{S}}(\mathbf{v}) + \mathscr{J}_k^{\mathscr{D}}(\rho,\mathbf{v})$.

The main result of this section is Theorem 3.9, establishing the convergence (under some arguably mild conditions) of the "dark-grey – light-grey" approximations towards the black–white limits. We begin with some auxiliary results.

The following lemma allows us to define a "limiting" design $\widetilde{\rho} \in \widetilde{\mathscr{H}}$, corresponding to the limiting flow $\mathbf{u}$, even though the sequence of "dark-grey – light-grey" controls $\{\rho_k\}$ might have no limit points in $\widetilde{\mathscr{H}}$ in the usual weak* sense.

Lemma 3.7. *Let $\{\mathbf{u}_k\} \subset H^1(\Omega)$ weakly converge to $\mathbf{u} \in H^1(\Omega)$. Define $\rho_k = \overline{\rho}_k \chi_{\Omega(\mathbf{u}_k)} + \underline{\rho}_k \chi_{\Omega\setminus\Omega(\mathbf{u}_k)}$, and assume that $\rho_k \in \widetilde{\mathscr{H}_k}$ (i.e., $\int_\Omega \rho_k \le \gamma|\Omega|$), and that*

$$\liminf_{k\to+\infty} \mathscr{J}^{\mathscr{D}}(\rho_k, \mathbf{u}_k) = \liminf_{k\to+\infty} \frac{1}{2}\left[\underline{\alpha}_k \int_{\Omega(\mathbf{u}_k)} \mathbf{u}_k\cdot\mathbf{u}_k + \overline{\alpha}_k \int_{\Omega\setminus\Omega(\mathbf{u}_k)} \mathbf{u}_k\cdot\mathbf{u}_k\right] \le C,$$

for some constant $C < +\infty$. Then, there is $\widetilde{\rho} \in \widetilde{\mathscr{H}}$ such that

$$\mathscr{J}^{\mathscr{D}}(\widetilde{\rho}, \mathbf{u}) = 0. \tag{3.10}$$

In particular, $|\Omega_{nz}(\mathbf{u})| \le \gamma|\Omega|$.

Proof. The existence of limit points follows from the inclusion $\widetilde{\mathscr{H}_k} \subset \mathscr{H}$, $k = 1, 2, \ldots$, and the weak*-compactness of the latter. Therefore, we will assume that the original sequence $\{\rho_k\}$ weakly* converges to $\rho \in \mathscr{H}$.

The control function ρ_k is a solution to the following optimization problem with a linear objective functional and weak*-compact feasible set:

$$\max_{\rho\in\mathscr{H}_k}\int_\Omega \rho\mathbf{u}_k\cdot\mathbf{u}_k, \tag{3.11}$$

Since $\{\mathbf{u}_k\cdot\mathbf{u}_k\}$ converges strongly in $L^1(\Omega)$, from Proposition 4.4 in [BoS00] it follows that ρ must solve the following optimization problem:

$$\max_{\rho\in\mathscr{H}}\int_\Omega \rho\mathbf{u}\cdot\mathbf{u}, \tag{3.12}$$

Further, since the objective functional of (3.12) is linear (in ρ), the problem possesses a zero-one optimal solution $\widetilde{\rho}\in\widetilde{\mathscr{H}}$; we can always take $\widetilde{\rho}=\chi_{\Omega(\mathbf{u})}$.

Clearly,

$$2C\ge\liminf_{k\to+\infty}\int_\Omega\alpha(\rho_k)\mathbf{u}_k\cdot\mathbf{u}_k=\liminf_{k\to+\infty}\overline{\alpha}_k\int_{\Omega\setminus\Omega(\mathbf{u}_k)}\mathbf{u}_k\cdot\mathbf{u}_k,$$

which implies that

$$0=\liminf_{k\to+\infty}\int_{\Omega\setminus\Omega(\mathbf{u}_k)}\mathbf{u}_k\cdot\mathbf{u}_k=\lim_{k\to+\infty}\int_\Omega\rho_k\mathbf{u}_k\cdot\mathbf{u}_k=\int_\Omega\widetilde{\rho}\mathbf{u}\cdot\mathbf{u}=\int_{\Omega\setminus\Omega(\mathbf{u})}\mathbf{u}\cdot\mathbf{u},$$

where we used the convergence of optimal values for the problems (3.11) to the one of the problem (3.12) as k goes to $+\infty$ (again, by Proposition 4.4 in [BoS00]). We conclude that $\mathbf{u}\equiv\mathbf{0}$ on $\Omega\setminus\Omega(\mathbf{u})$, which implies (3.10). □

Corollary 3.8. *In addition to the assumptions of Lemma 3.7, assume that* $|\Omega_{nz}(\mathbf{u})|=\gamma|\Omega|$. *Then, the sequence* $\{\rho_k\}$ *converges to* $\widetilde{\rho}\in\widetilde{\mathscr{H}}$ *strongly in* $L^1(\Omega)$.

Proof. The additional assumption implies that the problem (3.12) possesses the only optimal solution $\widetilde{\rho}=\chi_{\Omega(\mathbf{u})}=\chi_{\Omega_{\mathrm{nz}}(\mathbf{u})}$. This implies the weak* convergence of the sequence $\{\rho_k\}$ towards $\widetilde{\rho}$ in $L^\infty(\Omega)$. Strong convergence in $L^1(\Omega)$ then follows from Corollary 3.2 in [Pet99]. □

Now, the main result of this section can be established.

Theorem 3.9 (Convergence of "dark-grey – light-grey" approximations). *Consider the sequence of sizing optimization problems:*

$$\min_{(\rho,\mathbf{v})\in\mathscr{H}_k\times\mathscr{U}_{\mathrm{div}}}\mathscr{J}_k(\rho,\mathbf{v}),\quad k=1,\dots, \tag{3.13}$$

Let $\{(\rho_k^*,\mathbf{u}_k^*)\}$ *be a sequence of "dark-grey – light-grey" optimal solutions to* (3.13) *(i.e.,* $(\rho_k^*,\mathbf{u}_k^*)\in\widetilde{\mathscr{H}_k}\times\mathscr{U}_{\mathrm{div}}$, $k=1,2,\dots$*), which exists by Corollary 3.1 in [BoP03]. Then, an arbitrary weak limit point* $\mathbf{u}$ *of the sequence* $\{\mathbf{u}_k^*\}\subset H^1(\Omega)$ *(and there is at least one) defines a control* $\rho=\chi_{\Omega(\mathbf{u})}\in\widetilde{\mathscr{H}}$ *such that* $(\rho,\mathbf{u})$ *is an optimal solution to the problem* (3.7).

If, in addition, $|\Omega_{nz}(\mathbf{u})|=\gamma|\Omega|$, *then* $\{\rho_k\}$ *strongly converges to* ρ *in* $L^1(\Omega)$.

Proof. Let $\mathbf{u}^{\mathscr{S}}$ be the Stokes flow constructed in the proof of Theorem 3.6; set $\rho_k = \overline{\rho}_k \chi_{\Omega_{\mathrm{nz}}(\mathbf{u}^{\mathscr{S}})} + \underline{\rho}_k \chi_{\Omega \setminus \Omega_{\mathrm{nz}}(\mathbf{u}^{\mathscr{S}})}$. Then, $(\rho_k, \mathbf{u}^{\mathscr{S}})$ is feasible in (3.13), $k = 1, 2, \dots$. Therefore, the following inequalities hold:

$$\begin{aligned}\limsup_{k\to+\infty} \mathscr{J}_k(\rho_k^*, \mathbf{u}_k^*) &\le \limsup_{k\to+\infty} \mathscr{J}_k(\rho_k, \mathbf{u}^{\mathscr{S}}) \\ &\le \mathscr{J}^{\mathscr{S}}(\mathbf{u}^{\mathscr{S}}) + \lim_{k\to+\infty} 1/2\underline{\alpha_k}\|\mathbf{u}^{\mathscr{S}}\|_{L^2(\Omega)} \\ &= \mathscr{J}^{\mathscr{S}}(\mathbf{u}^{\mathscr{S}}) < +\infty.\end{aligned}$$

This directly implies the boundedness of the sequence $\{\mathbf{u}_k^*\}$; we therefore assume that the original sequence weakly converges to $\mathbf{u}$. Furthermore, owing to Lemma 3.7, the pair $(\rho, \mathbf{u})$, with $\rho = \chi_{\Omega(\mathbf{u})}$, is feasible in (3.7).

Let $(\rho^*, \mathbf{u}^*) \in \widetilde{\mathscr{H}} \times \mathscr{U}_{\mathrm{div}}$ be an arbitrary zero-one optimal solution to (3.7). By the weak lower semicontinuity of $\mathscr{J}^{\mathscr{S}}$, we have:

$$\mathscr{J}^{\mathscr{S}}(\rho^*, \mathbf{u}^*) \le \mathscr{J}^{\mathscr{S}}(\rho, \mathbf{u}) \le \liminf_{k\to+\infty} \mathscr{J}^{\mathscr{S}}(\mathbf{u}_k^*) \le \liminf_{k\to+\infty} \mathscr{J}_k(\rho_k^*, \mathbf{u}_k^*).$$

On the other hand, letting $\widetilde{\rho}_k = \overline{\rho}_k \chi_{\Omega(\mathbf{u}^*)} + \underline{\rho}_k \chi_{\Omega\setminus\Omega(\mathbf{u}^*)}$, we obtain the reverse inequality:

$$\mathscr{J}^{\mathscr{S}}(\rho^*, \mathbf{u}^*) = \mathscr{J}(\rho^*, \mathbf{u}^*) = \lim_{k\to+\infty} \mathscr{J}_k(\widetilde{\rho}_k, \mathbf{u}^*) \ge \limsup_{k\to+\infty} \mathscr{J}_k(\rho_k^*, \mathbf{u}_k^*),$$

owing to the feasibility of $(\widetilde{\rho}_k, \mathbf{u}^*)$ in (3.13), $k = 1, 2, \dots$. This establishes the optimality of $(\rho, \mathbf{u})$ in (3.7).

The last claim is a simple application of Corollary 3.8. □

Now we are ready to discuss the additional assumption of Theorem 3.9 (the assumption of Corollary 3.8), which guarantees the strong convergence of the optimal approximating controls. This condition necessarily holds if the flow volume constraint $\int_\Omega \rho \le \gamma|\Omega|$ is active (binding) at *every* control that is optimal in (3.7). While we do not know if this condition holds in every instance of the problem (3.7), it can always be satisfied by decreasing the flow volume factor γ, if the convergence towards the flow $\mathbf{u}$ with $|\Omega_{\mathrm{nz}}(\mathbf{u})| < \gamma|\Omega|$ is observed, and resolving the problem.

There is an obstacle, however, which might prevent this from working in practice: each of the approximating problems (3.13) is nonconvex, and, therefore, we cannot expect them to be solved to global optimality by numerical algorithms. (Many structural optimization problems are rather difficult to approximate due to the inherent nonconvexity of the approximating problems; see [StS01].) Despite this fact, in realistic instances of (3.7) we expect the flow volume constraint to be binding.

3.7 Bilevel programming in flow mechanics: a possible generalization?

Assume that we are interested in the optimal control of the Darcy-Stokes equations with respect to an alternative objective functional $\mathscr{F} : \mathscr{H} \times H^1(\Omega) \to \mathbb{R} \cup \{\infty\}$, where $\mathscr{H}$ denotes the abstract control set. Formally, we would like to solve the following *bilevel* (cf. [LPR96, page 10]) programming problem:

$$\begin{cases} \min\limits_{(\rho,\mathbf{u})\in\mathscr{H}\times H^1(\Omega)} \mathscr{F}(\rho,\mathbf{u}), \\ \text{s.t.} \quad \mathbf{u} \in \operatorname*{argmin}\limits_{\mathbf{v}\in\mathscr{U}_{\text{div}}} \mathscr{J}(\rho,\mathbf{v}), \end{cases} \tag{3.14}$$

Similarly, if we are interested only in pure Stokes flows, the optimization problem can be posed as follows:

$$\begin{cases} \min\limits_{(\rho,\mathbf{u})\in\mathscr{H}\times H^1(\Omega)} \mathscr{F}(\rho,\mathbf{u}), \\ \text{s.t.} \begin{cases} \mathbf{u} \in \operatorname*{argmin}\limits_{\mathbf{v}\in\mathscr{U}_{\text{div}}} \mathscr{J}(\rho,\mathbf{v}), \\ \mathscr{J}^{\mathscr{D}}(\rho,\mathbf{u}) = 0. \end{cases} \end{cases} \tag{3.15}$$

Of course, the minimization of the power function is the simplest problem one can consider in flow topology optimization, owing to the fact that we can join the lower-level and upper-level optimization problems into one: then, the bilevel program (3.14) reduces to (3.5). This fact allows us to minimize the objective functional $\mathscr{F} \equiv \mathscr{J}$ *simultaneously* w.r.t. $(\rho,\mathbf{u})$, resulting in a problem with an inf-compact, l.s.c. functional (w.r.t. the weak* $\times$ weak topology of $L^\infty(\Omega)\times H^1(\Omega)$) that, thus, possesses optimal solutions. In the bilevel case, the mapping $\rho \to \operatorname{argmin}_{\mathbf{v}\in\mathscr{U}_{\text{div}}} \mathscr{J}(\rho,\mathbf{v})$ is not closed in the weakly* $\times$ weakly topology of $L^\infty(\Omega)\times H^1(\Omega)$. The next example shows that this mapping is not closed even in the strong topology of $L^1(\Omega)\times H^1(\Omega)$, which in particular prevents us from using the weak* topology of $BV(\Omega)$ (or even $SBV(\Omega)$, cf. [AFP00]) for the design space of the problems (3.14) and (3.15).

Example 3.10 (Disappearing wall in the driven cavity flow problem). Let $\Omega = (0,1)\times(-1,1)\subset\mathbb{R}^2$, $\Omega_+ = (0,1)\times(0,1)$, $\Omega_- = \Omega\setminus\Omega_+$, $\mathbf{f}\equiv 0$ in Ω, $\mathbf{g}\equiv(1,0)$ on the "upper" boundary (the line connecting the points $(0,1)$ and $(1,1)$), and $\mathbf{g}\equiv\mathbf{0}$ otherwise. Define $\mathbf{u}_+$ to be the solution to the "lid-driven cavity flow" problem (see, e.g., [Jia98, page 146]) in Ω_+, $\mathbf{u}_+ = \mathbf{0}$ in Ω_-.

Consider a sequence $\{\rho_k\}\subset L^\infty(\Omega)\cap BV(\Omega)$, with $\rho_k \equiv 1-\chi_{(1,0)\times(-1/k,0)}$ in Ω, $k=1,2,\dots$. The solution to the Darcy-Stokes problem (3.2) in this case is $\mathbf{u}_k = \mathbf{u}_+$; thus $\{(\rho_k,\mathbf{u}_k)\}\to(1,\mathbf{u}_+)$ strongly in $L^1(\Omega)\times H^1(\Omega)$. At the same time, the flow corresponding to $\rho\equiv 1$ in Ω is the solution to the driven cavity flow problem in Ω, which is not equal to $\mathbf{u}_+$. Thus, the mapping $\rho\to\operatorname{argmin}_{\mathbf{v}\in\mathscr{U}_{\text{div}}}\mathscr{J}(\rho,\mathbf{v})$ is not closed even in the strong topology of $L^1(\Omega)\times H^1(\Omega)$, even though $\limsup_{k\to+\infty}\mathscr{J}(\rho_k,\mathbf{u}_k) < +\infty$.

Now, define $\mathscr{F}(\rho,\mathbf{v}) = \|1-\rho\|_{BV(\Omega)} + \|\mathbf{v}-\mathbf{u}_+\|_{H^1(\Omega)}$, $\mathscr{H} = \{\rho\in BV(\Omega) \mid 0\le\rho\le 1, \text{a.e. in } \Omega\}$. Then, the sequence $\{(\rho_k,\mathbf{u}_k)\}$ is a minimizing sequence for both problem (3.14) and (3.15), which does not converge to a feasible point of either of the problems. Therefore, the classic "flow tracking problem" posed as a bilevel topology optimization problem of Darcy-Stokes flow has no solutions.

If we restrict the set of admissible controls so that $\rho\ge\underline{\rho}>0$ in Ω, the problem (3.14) becomes well-posed for every enough continuous objective functional; however, making such a restriction we arrive at a less interesting for us sizing case. Therefore, the problem of choosing practically interesting and well-posed formulations of the topology optimization of Stokes flows with objective functionals other than the total power $\mathscr{J}$ remains open.

3.8 Conclusions and further research

We have shown that the topology optimization problem of the Darcy-Stokes equations w.r.t. total power minimization admits optimal solutions, even if the limiting zero and infinite permeabilities are included in the design domain. We have further established that the problem of finding a zero-one optimal control, or optimal pure Stokes flow, can be set up in a well-posed way; no additional restriction techniques are necessary in contrast with the case of linear elastisity (cf. [BeS03]). Two techniques were proposed for solving the zero-one optimal control problem. We have also shown that the topology optimization problem w.r.t. alternative functionals might be ill-posed, and might lack optimal solutions.

It would be particularly interesting to study the zero-one topology optimization problem of Navier-Stokes or Euler flows. For the Navier-Stokes flows, which are of much engineering interest, one can take the same design parametrization as for the Stokes flows. The problematic part, as it is typical in topology optimization, is to establish the inf-compactness property of the chosen objective functional on the set of admissible designs. The theory for the sizing case is straightforward, and only the numerical part needs to be investigated. For the Euler flows, even the design parametrization is unclear, partly due to the fact that flows of inviscid fluids through porous media are not so well investigated in the literature.

As for the Stokes flow, the further study of bilevel optimization problems might be interesting, as well as the consideration of alternative flow boundary conditions (cf. [Jia98, Section 8.2.2]).

Acknowledgement: The author is grateful to Michael Patriksson for his careful reading of the manuscript and suggestions that improved the presentation. This research is supported by the Swedish Research Council (grant 621-2002-5780).

References

[AFP00] Luigi Ambrosio, Nicola Fusco, and Diego Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 2000.

[BeS99] Martin P. Bendsøe and Ole Sigmund. Material interpolation schemes in topology optiization. *Archives of Applied Mechanics*, 69(9–10):635–654, 1999.

[BeS03] Martin P. Bendsøe and Ole Sigmund. *Topology Optimization: Theory, Methods, and Applications*. Springer-Verlag, Berlin, 2003.

[BoP03] Thomas Borrvall and Joakim Petersson. Topology optimization of fluids in Stokes flow. *Internat. J. Numer. Methods Fluids*, 41(1):77–107, 2003.

[BoS00] J. Frédéric Bonnans and Alexander Shapiro. *Perturbation Analysis of Optimization Problems*. Springer-Verlag, New York, 2000.

[BSS93] Mokhtar S. Bazaraa, Hanif D. Sherali, and C. M Shetty. *Nonlinear Programming*. John Wiley & Sons, Inc, New York, 1993.

[Dac89] Bernard Dacorogna. *Direct methods in the calculus of variations*, volume 78 of *Applied Mathematical Sciences*. Springer-Verlag, Berlin, 1989.

[EkT99] Ivar Ekeland and Roger Témam. *Convex analysis and variational problems*, volume 28 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, english edition, 1999. Translated from the French.

[Jia98] Bo-nan Jiang. *The least-squares finite element method*. Scientific Computation. Springer-Verlag, Berlin, 1998. Theory and applications in computational fluid dynamics and electromagnetics.

[LPR96] Zhi-Quan Luo, Jong-Shi Pang, and Daniel Ralph. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, Cambridge, 1996.

[NiB99] Donald A. Nield and Adrian Bejan. *Convection in porous media*. Springer-Verlag, New York, second edition, 1999.

[Pet99] Joakim Petersson. A finite element analysis of optimal variable thickness sheets. *SIAM J. Numer. Anal.*, 36(6):1759–1778, 1999.

[Pir73] O. Pironneau. On optimum profiles in Stokes flow. *J. Fluid Mech.*, 59:117–128, 1973.

[Pir74] O. Pironneau. On optimum design in fluid mechanics. *J. Fluid Mech.*, 64:97–110, 1974.

[Roc70] R. Tyrrell Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, N.J., 1970.

[StS01] M. Stolpe and K. Svanberg. On the trajectories of penalization methods for topology optimization. *Struct. and Multidisc. Optim.*, 21(2):128–139, 2001.

Paper 4

TOPOLOGY OPTIMIZATION OF SLIGHTLY COMPRESSIBLE FLUIDS

Anton Evgrafov*

Abstract

We consider the problem of optimal design of flow domains for Navier–Stokes flows in order to minimize a given performance functional. We attack the problem using topology optimization techniques, or control in coefficients, which are widely known in structural optimization of solid structures for their flexibility, generality, and yet ease of use and integration with existing FEM software. Topology optimization rapidly finds its way into other areas of optimal design, yet until recently it has not been applied to problems in fluid mechanics. The success of topology optimization methods for the minimal drag design of domains for Stokes fluids (see the study of Borrvall and Petersson [Internat. J. Numer. Methods Fluids, vol. 41, no. 1, pp. 77–107, 2003]) has lead to attempts to use the same optimization model for designing domains for incompressible Navier–Stokes flows. We show that the optimal control problem obtained as a result of such a straightforward generalization is ill-posed, at least if attacked by the direct method of calculus of variations.

We illustrate the two key difficulties with simple numerical examples and propose changes in the optimization model that allow us to overcome these difficulties. Namely, to deal with impenetrable inner walls that may appear in the flow domain we slightly relax the incompressibility constraint as typically done in penalty methods for solving the incompressible Navier–Stokes equations. In addition, to prevent discontinuous changes in the flow due to very small impenetrable parts of the domain that may disappear, we consider so-called filtered designs, that has become a “classic” tool in the topology optimization toolbox. Technically, however, our use of filters differs significantly from their use in the structural optimization problems in solid mechanics, owing to the very unlike design parametrizations in the two models.

We rigorously establish the well-posedness of the proposed model and then discuss related computational issues.

Keywords. Topology optimization, Fluid mechanics, Navier–Stokes flow, Domain identification, Fictitious domain.

AMS subject classification. 76D55, 76N25, 62K05, 49J20, 49J45.

*Department of Mathematics, Chalmers University of Technology, SE-412 96 Göteborg, Sweden, e-mail: toxa@math.chalmers.se

4.1 Introduction

THE optimal control of fluid flows has long been receiving considerable attention by engineers and mathematicians, owing to its importance in many applications involving fluid related technology; see, e.g., the recent monographs [Gun03, MoP01], and articles [Fei03, Ton03b, Ton03a, vBS02, ChG02, GuM02, Kim01, GKM00, OkK00, DzZ99, GuK98, DRSS96, Sue96, BoB95, NRS95, StS94, BeD92, BFCS92], including the pioneering works of Pironneau [Pir73, Pir74] on the optimality conditions for shape optimization in fluid mechanics. According to a well-established classification in structural optimization [BeS03, page 1], the absolute majority of works dealing with optimal design of flow domains fall into the category of shape optimization. (See the bibliographical notes (2) in [BeS03] for classic references in shape optimization.) In the framework of *shape optimization*, the optimization problem formulation can be stated as follows: choose a flow domain out of some family so as to maximize an associated performance functional. The family of domains considered may be as rich as that of all open subsets of a given set satisfying some regularity criterion (see, e.g. [Fei03]), or as poor as the ones obtained from a given domain by locally perturbing some part of the boundary in a Lipschitz manner (cf. [Ton03b, GKM00, GuK98]). Unfortunately, it is typically only the problems in the latter group that can be attacked numerically. On the other hand, *topology optimization* (or, control in coefficients) techniques are known for their flexibility in describing the domains of arbitrary complexity (e.g., the number of connected components need not to be bounded), and at the same time require relatively moderate efforts from the computational part. In particular, one may completely avoid remeshing the domain as the optimization algorithm advances, which eases the integration with existing FEM codes, and simplifies and speeds up sensitivity analysis.

While the field of topology optimization is very well established for optimal design of solids and structures, surprisingly little work has been done for optimal design of fluid domains. Borrvall and Petersson [BoP03] considered the optimal design of flow domains for minimizing the total power of the incompressible Stokes flows, using inhomogeneous porous materials with a spatially varying Darcy permeability tensor, under a constraint on the total volume of fluid in the control region. Later, this approach has been generalized to include both limiting cases of the porous materials, i.e., pure solid and pure flow regions have been allowed to appear in the design domain as a result of the optimization procedure [Evg03]. (We also cite the work of Klarbring et al. [KPTK03], which however studies the problem of optimal design of flow networks, where design and state variables reside in finite-dimensional spaces; in some sense this is an analogue of truss design problems if one can carry over the terminology and ideas from the area of optimal design of structures and solids.)

However, applications of the Stokes flows are rather limited, while the Navier–Stokes equations are now regarded as the universal basis of fluid mechanics [Dar02]. Therefore, it has been suggested that the optimization model proposed by Borrvall and Petersson [BoP03] (with straightforward modifications), in particular the same design parametrization should be used for the topology optimization of the incompressible Navier–Stokes equations [GH03]. Essentially, in this model we control the Brinkman-type equations including the nonlinear convection term [All90a] (which will be referred to as nonlinear Brinkman equations in the sequel) by varying a coefficient before the zeroth order velocity term. Setting the control coefficient to zero is supposed to recover the Navier–Stokes equations; at the same time, infinite values of the coefficient are supposed to model the impenetrable inner walls in the domain. In Section 4.3 we illustrate the difficulties inherent in this approach, namely that the design-to-flow mapping is not closed, leading to ill-posed control problems.

It turns out that if we employ the idea of *filter* [Sig97, SiP98] (which has become quite a standard technique in topology optimization, see [Bou01, BrT01] for the rigorous mathematical treatment) *in addition* to relaxing the incompressibility constraint (which is unique to the topology optimization of fluids) we can establish the continuity of the resulting design-to-flow mapping, and therefore the existence of optimal designs for a great variety of design functionals; this is discussed in Section 4.4. Not going into details yet, we comment that our use of filters significantly differs from the traditional one in the topology optimization. Namely, not only do we use filters to forbid small features from appearing in our designs and thus to transform weak(-er) design convergence into a strong(-er) one (cf. Proposition 4.5), but also to verify certain growth conditions near impenetrable walls [see inequality (4.4) and Proposition 4.27], which later guarantees the embedding of certain weighted Sobolev spaces into classic ones (see inequality (4.14) in the proof of Proposition 4.12), and finally allows us to prove the continuity of design-to-flow mappings in Section 4.5. The existence of optimal designs, formally established in Section 4.6, is an easy corollary of the continuity of the design-to-flow mappings.

Some computational techniques are introduced in Section 4.7. Namely, in Subsection 4.7.1 we discuss a standard topic of approximating the topology optimization problems with so-called sizing optimization problems (also known as “ε-perturbation”), which in our case reduces to approximation of the impenetrable walls with materials of very low permeability. In Subsection 4.7.2 we touch upon techniques aimed at reducing the amount of porous material in the optimal design. We conclude the paper by discussing possible extensions of the presented results, open questions, and further research topics in Section 4.8. Proofs of some results are found in Appendix 4.A.

4.2 Prerequisites

4.2.1 Notation

We follow standard engineering practice and will denote vector quantities, such as vectors and vector-valued functions, using the **bold** font. However, for functional spaces of both scalar- and vector-valued functions we will use regular font.

Let Ω be a connected bounded domain of $\mathbb{R}^d$, $d \in \{2,3\}$ with a Lipschitz continuous boundary Γ. In this domain we would like to control the nonlinear Brinkman equations [All90a] with the prescribed flow velocities $\mathbf{g}$ on the boundary, and forces $\mathbf{f}$ acting in the domain by adjusting the inverse permeability α of the medium occupying Ω, which depends on the control function ρ:

$$\begin{cases} \left.\begin{aligned} -\nu\boldsymbol{\Delta}\mathbf{u} + \mathbf{u}\cdot\boldsymbol{\nabla}\mathbf{u} + \alpha(\rho)\mathbf{u} + \boldsymbol{\nabla} p = \mathbf{f}, \\ \operatorname{div}\mathbf{u} = 0 \end{aligned}\right\}, & \text{in } \Omega, \\ \mathbf{u} = \mathbf{g}, & \text{on } \Gamma. \end{cases} \tag{4.1}$$

In the system (4.1), $\mathbf{u}$ is the flow velocity, p is the pressure, and ν is the kinematic viscosity. Of course, owing to the incompressibility of $\mathbf{u}$, the function $\mathbf{g}$ must satisfy the compatibility condition

$$\int_\Gamma \mathbf{g}\cdot\mathbf{n} = 0, \tag{4.2}$$

where $\mathbf{n}$ denotes the outward unit normal. If $\alpha(\rho(\mathbf{x})) = +\infty$ for some $\mathbf{x}\in\Omega$, we simply require $\mathbf{u}(\mathbf{x}) = 0$ in the first equation of (4.1).

Our control set $\mathscr{H}$ is defined as follows:

$$\mathscr{H} = \{\, \rho \in L^\infty(\Omega) \mid 0 \le \rho \le 1, \text{a.e. in } \Omega, \int_\Omega \rho \le \gamma|\Omega| \,\},$$

where $0 < \gamma < 1$ is the maximal volume fraction that can be occupied by the fluid. Every element $\rho \in \mathscr{H}$ describes the scaled Darcy permeability tensor of the medium at a given point $\mathbf{x}\in\Omega$ in the following (informal) way: $\rho(\mathbf{x}) = 0$ corresponds to zero permeability at $\mathbf{x}$ (i.e., solid, which does not permit any flow at a given point), while $\rho(\mathbf{x}) = 1$ corresponds to infinite permeability (i.e., 100% flow region; no structural material is present). Formally, we relate the permeability α^{-1} to ρ using a convex, decreasing, and nonnegative function (cf. [BoP03, Evg03]) $\alpha : [0,1] \to \mathbb{R}_+ \cup \{+\infty\}$, defined as

$$\alpha(\rho) = \rho^{-1} - 1.$$

In the rest of the paper we will use the symbol χ_A for $A \subset \Omega$ to denote the characteristic function of A: $\chi_A(\mathbf{x}) = 1$ for $\mathbf{x}\in A$; $\chi_A(\mathbf{x}) = 0$ otherwise.

4.2.2 Variational formulation

To state the problem in a more analytically suitable way and to incorporate the special case $\alpha = +\infty$ into the first equation of the system (4.1), we introduce a weak formulation of the equations. Let us consider the sets of admissible flow velocities:

$$\begin{aligned}\mathscr{U} &= \{\, \mathbf{v} \in H^1(\Omega) \mid \mathbf{v} = \mathbf{g} \text{ on } \Gamma \,\},\\ \mathscr{U}_{\text{div}} &= \{\, \mathbf{v} \in \mathscr{U} \mid \operatorname{div}\mathbf{v} = 0, \text{weakly in } \Omega \,\}.\end{aligned}$$

Let $\mathscr{J}^{\mathscr{S}} : \mathscr{U} \to \mathbb{R}$ denote the potential power of the viscous flow:

$$\mathscr{J}^{\mathscr{S}}(\mathbf{u}) = \frac{\nu}{2}\int_\Omega \nabla\mathbf{u}\cdot\nabla\mathbf{u} - \int_\Omega \mathbf{f}\cdot\mathbf{u}.$$

Let us further define the additional power dissipation $\mathscr{J}^{\mathscr{D}} : \mathscr{H}\times\mathscr{U} \to \mathbb{R}\cup\{+\infty\}$, due to the presence of the porous medium (we use the standard convention $0\cdot +\infty = 0$):

$$\mathscr{J}^{\mathscr{D}}(\rho,\mathbf{u}) = \frac{1}{2}\int_\Omega \alpha(\rho)\mathbf{u}\cdot\mathbf{u}.$$

Finally, let $\mathscr{J}(\rho,\mathbf{u}) = \mathscr{J}^{\mathscr{S}}(\mathbf{u}) + \mathscr{J}^{\mathscr{D}}(\rho,\mathbf{u})$ denote the total power of the Brinkman flow. Then, the requirement "$\alpha(\rho) = +\infty \implies \mathbf{u} = 0$" is automatically satisfied if $\mathscr{J}^{\mathscr{D}}(\rho,\mathbf{u}) < +\infty$.

We will use epi-convergence of optimization problems as a main theoretical tool in the subsequent analysis, thus it is natural to study the following variational formulation (cf., e.g., [Evg03]) for Darcy-Stokes flows [i.e., obtained by neglecting the convection term $\mathbf{u}\cdot\nabla\mathbf{u}$ in (4.1)]: for $\mathbf{f}\in L^2(\Omega)$, compatible $\mathbf{g}\in H^{1/2}(\Gamma)$, and $\rho\in\mathscr{H}$, find $\mathbf{u}\in\mathscr{U}_{\mathrm{div}}$ such that

$$\mathbf{u}\in\operatorname*{argmin}_{\mathbf{v}\in\mathscr{U}_{\mathrm{div}}}\mathscr{J}(\rho,\mathbf{v}).$$

Naturally, taking convection into account, this can be generalized to the following fixed point-type formulation of (4.1) (see Subsection 4.5.2 for the rigorous discussion of its well-posedness): for $\mathbf{f}\in L^2(\Omega)$, compatible $\mathbf{g}\in H^{1/2}(\Gamma)$, and $\rho\in\mathscr{H}$ find $\mathbf{u}\in\mathscr{U}_{\mathrm{div}}$ such that

$$\mathbf{u}\in\operatorname*{argmin}_{\mathbf{v}\in\mathscr{U}_{\mathrm{div}}}\left\{\mathscr{J}(\rho,\mathbf{v})+\int_\Omega(\mathbf{u}\cdot\nabla\mathbf{u})\cdot\mathbf{v}\right\}. \tag{4.3}$$

4.3 Problems with the existing approach

When we allow impenetrable walls to appear and to disappear in the design domain, we create two particular types of difficulties, each related to a corresponding change in topology (see Subsection 4.3.1 and 4.3.2). We note that in the “sizing” case, which can be modeled by introducing an additional design constraint $\rho\geq\varepsilon$, a.e. in Ω (for some small $\varepsilon>0$) these difficulties do not appear. (In fact, it is an easy exercise to verify that under such circumstances the design-to-flow mapping is closed w.r.t. strong convergence of designs, e.g., in $L^1(\Omega)$, and $H^1(\Omega)$-weak convergence of flows.) Such a distinct behavior of the sizing and topology optimization problems may indicate that the former is not a useful approximation of the latter in this case.

4.3.1 Disappearing walls

For the sake of simplicity, in this subsection we assume that the objective functional in our control problem (which is not formally stated yet) is the power $\mathscr{J}$ of the incompressible Navier–Stokes flow. This functional is interesting from at least two points of view. Firstly, in many cases the resulting control problem is equivalent to the minimization of the drag force or pressure drop, which is very important in engineering applications [BoP03]. Secondly, while it is intuitively clear that impenetrable inner walls of vanishing thickness change the flow in a discontinuous way, for the Stokes flows the total potential power is lower semi-continuous w.r.t. such changes, which allows us to apply the Weierstrass theorem and ensure the existence of optimal designs (cf. [Evg03, Theorem 3.3]). In this subsection we consider two examples illustrating the discontinuity of the flow as well as non-lower semicontinuity of the power functional in the case of the incompressible Navier–Stokes equations; this means that the corresponding control problem of minimizing the potential power is ill-posed, at least from the point of view of the direct method of calculus of variations.

Example 4.1 (Infinitely thin wall). We consider a variant of the backstep flow with $\nu=1.0\cdot10^{-3}$ (which corresponds to the Reynolds number $\mathrm{Re}=1000$), as shown in Figure 4.1. We specify $\mathbf{u}$ on the inflow boundary to be $(0.25-(y-0.5)^2,0.0)^t$, on the outflow boundary we require $u_y=0$ as well as $p=0$; on the rest of the boundary the no-slip condition $\mathbf{u}=\mathbf{0}$ is assumed. We consider a sequence of the domains containing a

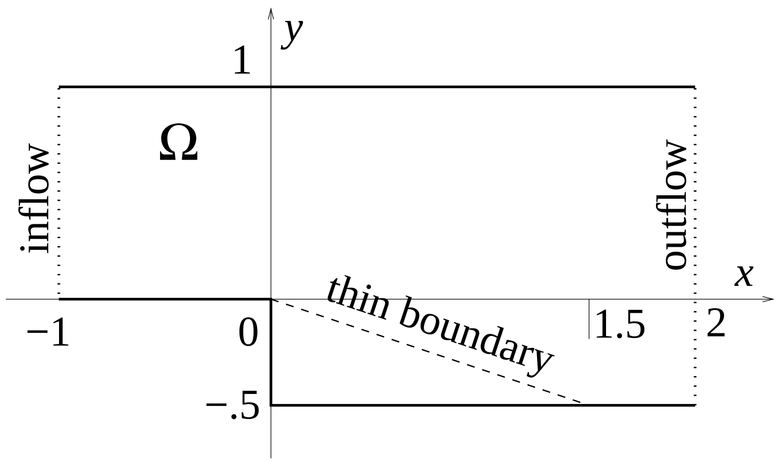


Fig. 4.1: Flow domain for the backstep flow.

thin but impenetrable wall of vanishing thickness (as shown in Figure 4.1 by dashed line). The limiting domain is the usual backstep shown with the solid line. Direct numerical computation in Femlab (see Figure 4.2 showing the flows) shows that for the domains with thin wall we have $\mathscr{J} \approx 0.8018$, while for the limiting domain $\mathscr{J} \approx 0.8263$. This demonstrates the non-lower semicontinuity of the total power functional in the case of incompressible Navier–Stokes equations.

We note that while the “jump” of the power functional may seem negligible in this example, other examples may be constructed where this jump is much bigger.

It may be argued that in the example above the thin wall may be substituted by the complete filling of the resulting isolated subdomain with impenetrable material, and the following example is more peculiar and demonstrates that we can control the behavior of the Navier–Stokes flow with an infinitesimal amount of material. It is interesting to note that the example is based on the construction of Allaire [All90a], which in some sense is “opposite” to our design parametrization. Namely, we try to control the Navier–Stokes equations by adjusting the coefficients in the nonlinear Brinkman equations, while the sequence of perforated domains considered in Example 4.2 has been used to obtain the nonlinear Brinkman equations starting from the Navier–Stokes equations in a periodically perforated domain as a result of the homogenization process.

Example 4.2 (Perforated domains with tiny holes). We assume that the boundary Γ is *smooth* and impenetrable (i.e., the homogeneous boundary conditions $\mathbf{g} = \mathbf{0}$ hold), and that the viscosity ν is large enough relatively to the force $\mathbf{f}$ to guarantee the existence of a unique solution to the Navier–Stokes system in Ω. Let Ω^ε denote a perforated domain, obtained from Ω by taking out spheres of radius $r_d(\varepsilon)$ with centers $\varepsilon\mathbb{Z}^d$, where $\lim_{\varepsilon\to+0} r_d(\varepsilon)/\varepsilon = 0$; see Figure 4.3. Let $(\widetilde{\mathbf{u}}^\varepsilon, \widetilde{p}^\varepsilon)$ denote a solution to the Navier–Stokes problem inside Ω^ε with homogeneous boundary conditions $\widetilde{\mathbf{u}}^\varepsilon = \mathbf{0}$ on $\partial\Omega^\varepsilon$. We extend $\widetilde{\mathbf{u}}^\varepsilon$ onto the whole Ω by setting it to zero inside each sphere; we further denote by $\mathbf{u}^\varepsilon$ this extended solution. For every small $\varepsilon > 0$ it holds that $\mathbf{u}^\varepsilon$ solves the problem (4.3) for

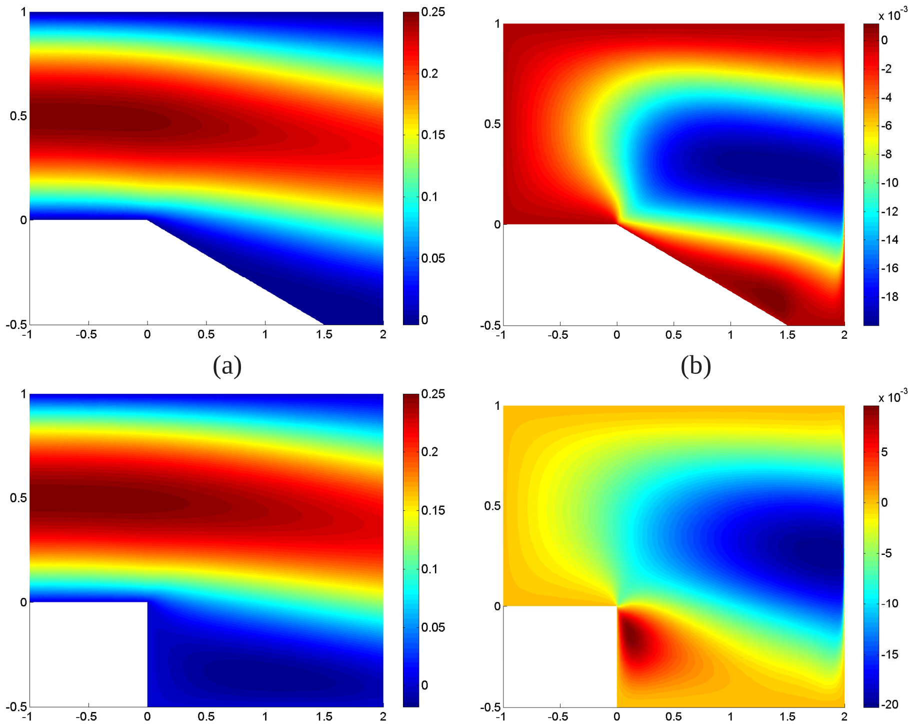


Fig. 4.2: Backstep flow: Example 4.1. (a), (b): x- and y-components, respectively, of the flow velocity when the impenetrable wall has arbitrary but positive thickness (only the part of the domain with nontrivial flow is shown); (c), (d): x- and y-components, respectively, of the flow velocity as the impenetrable wall disappears. *Note the different color scales.*

$\rho^\varepsilon = \chi_{\Omega^\varepsilon}$. Allaire [All90a] has shown that depending on the limit $C = \lim_{\varepsilon\to+0} r_d(\varepsilon)/\varepsilon^3$ for $d = 3$, or $C = \lim_{\varepsilon\to+0} -\varepsilon^2 \log(r_d(\varepsilon))$ for $d = 2$, there are three limiting cases:

$C = 0$: $\{\mathbf{u}^\varepsilon\}$ converges strongly in $H^1(\Omega)$ towards the solution to the Navier–Stokes problem in the unperforated domain Ω, i.e., the solution to the problem (4.3) corresponding to $\rho = 1$ (see [All90b, Theorem 3.4.4]);

$C = +\infty$: $\{\mathbf{u}^\varepsilon\}$ converges towards 0 strongly in $H^1(\Omega)$ (in fact, there is more information about $\{\mathbf{u}^\varepsilon\}$ available, see [All90b, Theorem 3.4.4]);

$0 < C < +\infty$: $\{\mathbf{u}^\varepsilon\}$ converges weakly in $H^1(\Omega)$ towards the solution to the nonlinear Brinkman problem in the unperforated domain Ω, i.e., the solution of the problem (4.3) corresponding to $\rho = \sigma$, for a computable constant $\sigma(d,\nu,C) > 0$ (see [All90a, Main Theorem]).

We note that in all three cases the sequence of designs $\{\rho^\varepsilon\}$ strongly converges to zero in $L^1(\Omega)$, while only in the case $C = 0$ the corresponding sequence of flows converges to the "correct" flow. As for the other two cases, we can either completely stop ($C = +\infty$) or

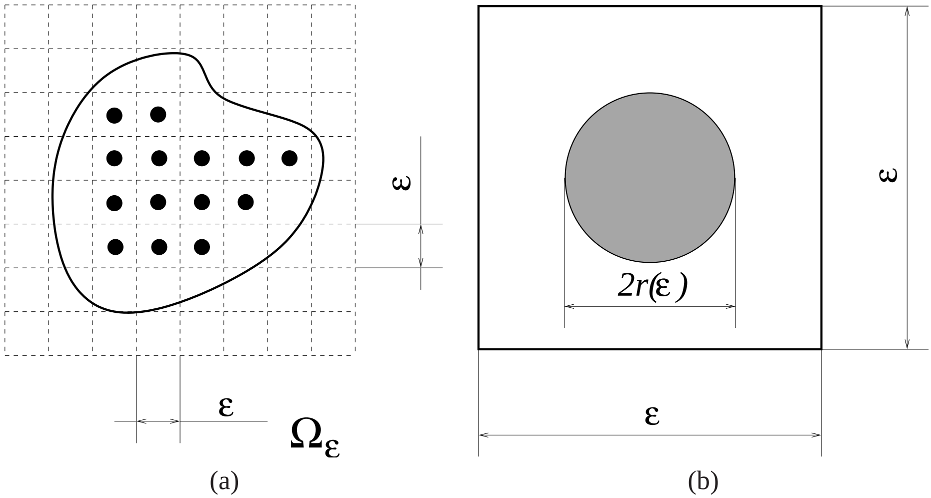


Fig. 4.3: The perforated domain (a) and a periodic cell (b).

just slow ($0 < C < +\infty$) the flow using only infinitesimal amounts of structural material (recall that $r_d(\varepsilon)/\varepsilon \to +0$). Moreover, the sequence of perimeters of ρ^ε converges to zero, and therefore the perimeter constraint cannot enforce the convergence of flows in this case (contrary to the situation in linear elasticity, cf. [BeS03, p. 31]). In the same spirit, the regularized intermediate density control method considered by Borrvall and Petersson [BoP01] classifies the designs ρ^ε as regular for all enough small $\varepsilon > 0$ (since they are indeed close to a regular design $\rho \equiv 0$ in the strong topology of $L^p(\Omega)$, $1 \leq p < \infty$); thus the latter method also fails to recognize the pathological cases illustrated in the present example.

4.3.2 Appearing walls

Walls that appear in the domain as a result of the optimization process may break the connectivity of the flow domain (or some parts of it), so that the incompressible Navier–Stokes system may not admit any solutions in the limiting domain (resp., some parts of it). While obtaining such results may seem to be a failure of the optimization procedure, completely stopping the flow might be interesting (or even optimal) with respect to some engineering design functionals.

The following example is purely artificial and its only purpose is to demonstrate the possible non-closedness of the design-to-flow mapping when new walls appear in the domain. It essentially repeats [Evg03, Example 2.1], but we include it here for convenience of the reader.

Example 4.3 (Domain with diminishing permeability). Let $\Omega = (0,1)^2$, $\mathbf{g} \equiv (1,0)^t$, and $\mathbf{f} \equiv \mathbf{0}$. Let further $\rho_k \equiv 1/k$ in Ω, $k = 1,2,\dots$, $\rho \equiv 0$ in Ω, so that $\rho_k \to \rho$, strongly in $L^\infty(\Omega)$ as $k \to \infty$. Then, $\mathbf{u} \equiv (1,0)^t$ is a solution of the problem (4.3) for all $k = 1,2,\dots$; clearly, $(\rho_k, \mathbf{u}) \to (\rho, \mathbf{u})$, strongly in $L^\infty(\Omega) \times H^1(\Omega)$. At the same time, it is not difficult to verify that the problem (4.3) has no solutions for the limiting design ρ, which means that the design-to-flow mapping is not closed even in the strong topology of $L^\infty(\Omega)\times H^1(\Omega)$!

The problem related to the appearence of walls completely stopping the flow in some domains has been solved for Stokes flows by (implicitly) introducing an additional constraint $\mathscr{J}(\rho,\mathbf{u})\le C$, for a suitable constant C. Owing to the coercivity of $\mathscr{J}$ on $H_0^1(\Omega)$, this keeps the flows in some bounded set. However, in view of the non-lower semicontinuity of the power functional for the Navier–Stokes flows (see Example 4.1), this set is not necessarily closed, making the problems with appearing walls much more severe in the present case.

We consider the next example in some detail, even though it is quite similar to the previous one, because we will return to it later in Subsection 4.4.2.

Example 4.4 (Channel with a porous wall). We consider a channel flow at Reynolds number $\mathrm{Re}=1000$ ($\nu=1.0\cdot10^{-3}$) through a wall made of porous material with vanishing permeability appearing in the middle of the channel (see Figure 4.4). We specify $\mathbf{u}$

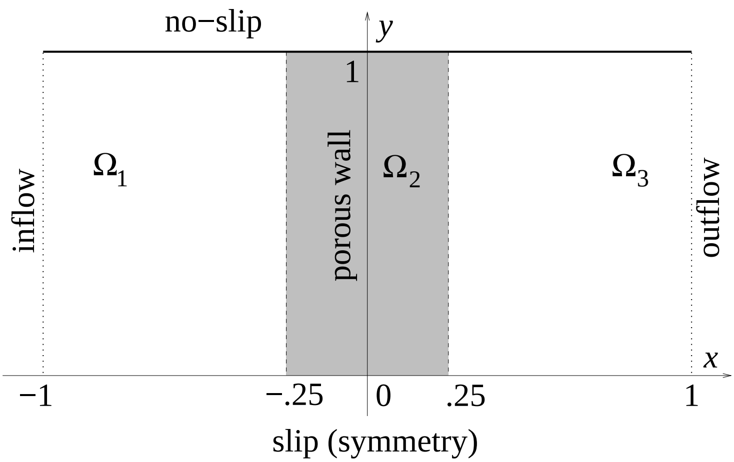


Fig. 4.4: Flow domain of Example 4.4.

on the inflow boundary to be $(1-y^2,0.0)^t$, on the outflow boundary we require $u_y=0$ as well as $p=0$; on the rest of the boundary the no-slip condition $\mathbf{u}=\mathbf{0}$ is assumed except that on the “lower” edge we have slip (i.e., only $u_y=0$) due to the symmetry.

We choose ρ so that $\alpha(\rho)=0$ on $\Omega_1\cup\Omega_3$ and $\alpha(\rho)=\alpha$ on Ω_2, where α assumes values 1.0, $1.0\cdot10^2$, $1.0\cdot10^4$, $+\infty$. The corresponding flows (calculated in Femlab) are shown in Figure 4.5; the incompressible Navier–Stokes problem in the last (limiting as $\alpha\to+\infty$) domain admits no solutions.

To summarize, even though the sequence of designs $\rho_\alpha\to\chi_{\Omega_1\cup\Omega_3}$, strongly in $L^\infty(\Omega)$, the corresponding sequence of flows does not converge to the flow corresponding to the limiting design, simply because the latter does not exist.

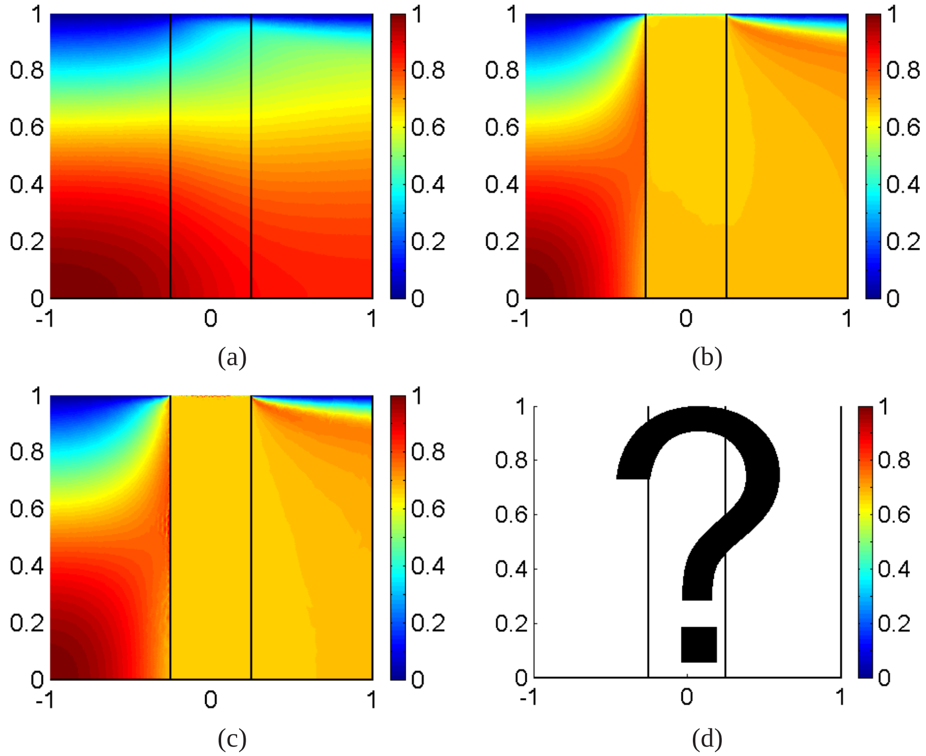


Fig. 4.5: Incompressible flow through the porous wall: (a) $\alpha = 1.0$, (b) $\alpha = 1.0 \cdot 10^2$, (c) $\alpha = 1.0 \cdot 10^4$, (d) $\alpha = +\infty$.

4.4 Proposed solutions to the difficulties outlined

Difficulties inherent in the straightforward generalization of the methodology proposed by Borrvall and Petersson [BoP03] for Stokes flows to incompressible Navier–Stokes flows have been outlined in Section 4.3. One possible solution, which allows us to avoid these difficulties, is simply to forbid topological changes and to perform sizing optimization, interpreting optimal designs as distributions of porous materials with spatially varying permeability (cf. [All90a, All90b]; see also [Hor97]). As it has already been mentioned the resulting designs may or may not accurately describe the domains obtained by substituting the materials with high permeability by void, and those with low permeability by impenetrable walls. Furthermore, if we decide to keep the porous material, it is questionable whether such designs can be easily manufactured and thus it is unclear whether they are "better" from the engineering point of view. Thus we do not employ this approach but instead try to slightly modify the design parametrization as well as the underlying state equations with the ultimate goal to rigorously obtain a closed design-to-flow mapping while maintaining a clear engineering/physical meaning of our optimization model.

4.4.1 Filters in the topology optimization

In both examples in Subsection 4.3.1 we constructed the sequences of designs having very small details, which disappear in the limit. Using the notion of a filter [Sig97, SiP98] we can control the minimal scale of our designs; we will employ this technique, which has become quite standard in topology optimization of linearly elastic materials [BeS03].

Following Bourdin [Bou01], and Bruns and Tortorelli [BrT01], we define a *filter* $F : \mathbb{R}^d \to \mathbb{R}$ *of characteristic radius* $R > 0$ to be a function verifying the following properties:

$$F \in C^{0,1}(\mathbb{R}^d), \qquad \operatorname{supp} F \Subset B_R, \qquad \operatorname{supp} F \text{ is convex},$$
$$F \geq 0 \text{ in } B_R, \qquad \int_{B_R} F = 1,$$

where B_R denotes the open ball of radius R centered in origo. We denote the convolution product by a $*$ sign, i.e.

$$(F * \rho)(\mathbf{x}) = \int_{\mathbb{R}^d} F(\mathbf{x}-\mathbf{y})\rho(\mathbf{y})d\mathbf{y}.$$

Owing to the Lipschitz continuity of F, $F * \rho$ is a continuous function (cf. [Bre83, Proposition IV.19]).

In order to compute the convolution between the filter and a given design ρ the latter must be defined not only on Ω, but also on the whole space $\mathbb{R}^d$. Therefore, in the sequel we consider the following redefined design domain:

$$\mathscr{H} = \{\, \rho \in L^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d) \mid 0 \leq \rho \leq 1, \text{a.e. in } \mathbb{R}^d, \int_{\mathbb{R}^d} \rho \leq V \,\},$$

for a given $V > 0$.

One of the consequences of the fact that F is Lipschitz continuous in $\mathbb{R}^d$ and not just in B_R is that the following important *growth condition* is verified (see Proposition 4.27):

$$(F * \chi_{\mathbb{R}^d \setminus \operatorname{supp} F})(\mathbf{x}) \leq C|\mathbf{x}|^2, \tag{4.4}$$

as $|\mathbf{x}| \to 0$, for some appropriate constant $C > 0$, which implies that $\alpha((F*\rho)(\cdot))$ grows at least as fast as $\operatorname{dist}^{-2}(\cdot, \{\, F*\rho = 0\,\})$ arbitrarily near to impenetrable walls. It is this condition that allows us to prove an approximation result, Proposition 4.12, which is in turn the key ingredient in the proof of our closedness theorems.

For notational convenience we set $\mathscr{J}^F(\rho, \mathbf{u}) = \mathscr{J}(F*\rho, \mathbf{u})$. As a consequence of the introduction of the filter, we can demonstrate the following simple claim, which translated to normal language says that impenetrable walls cannot disappear in the limit. In the following Proposition, $\operatorname{Lim\,sup}$ is understood in the sense of Painlevé-Kuratowski, see [AuF90, Definition 1.4.6], or [BoS00, Definition 2.52].

Proposition 4.5. *Consider an arbitrary sequence of designs* $\{\rho_k\} \subset \mathscr{H}$, *such that* $\rho_k \rightharpoonup \rho$, *weakly in* $L^1_{\text{loc}}(\mathbb{R}^d)$, *for some* $\rho \in \mathscr{H}$. *Define a sequence* $\{\Omega_0^k\}$ *of subsets of* Ω *as*

$$\Omega_0^k = \{\, \mathbf{x} \in \Omega \mid (F*\rho_k)(\mathbf{x}) = 0 \,\},$$
$$\Omega_0^\infty = \{\, \mathbf{x} \in \Omega \mid (F*\rho)(\mathbf{x}) = 0 \,\}.$$

Then, $\operatorname{Lim\,sup}_{k\to\infty} \Omega_0^k \subset \Omega_0^\infty \cup \Gamma$.

Proof. Let $I \subset \mathbb{N}$ be an infinite subsequence of indices, such that for some $\mathbf{x}_k \in \Omega_0^k$, $k \in I$, there exists $\mathbf{x} \in \mathbb{R}^d$ such that $\mathbf{x} = \lim_{k\in I} \mathbf{x}_k$. We know that $\rho_k \equiv 0$ a.e. on $\mathbf{x}_k + \operatorname{supp} F$, $k \in I$. Then, $\rho \equiv 0$ a.e. on $\mathbf{x} + \operatorname{supp} F$, i.e., $(F * \rho)(\mathbf{x}) = 0$. Clearly, $\mathbf{x} \in \operatorname{cl}\Omega$, which finishes the proof. □

Remark 4.6. The convergence of flow domains $\Omega \setminus \Omega_0^k$ induced by the weak convergence of designs (which implies strong convergence of filtered designs) can be compared to the convergence of domains in some topology defined for set convergence, e.g., the complementary Hausdorff topology. It is known, in general, that the latter topology is weaker (see, e.g., [SoZ92, Section 2.6.2]). However, such a comparison is not quite fair in the present situation, where the domains we deal with can be rather irregular (e.g., lie on two sides of their boundaries), and, more importantly, the domains in the sequence may have different connectivity compared to the "limiting" domain.

Later we will see that we need even stronger convergence of $\Omega_0^k \to \Omega_0^\infty$ to obtain closedness of the design-to-flow mappings.

The use of filtered designs $F * \rho$ in place of ρ in problem (4.3) allows us to overcome the difficulties caused by disappearing walls. While we delay the formal statement of this fact until Section 4.5, at this point we can consider an example that illustrates the effect of using filters.

Example 4.7 (Example 4.2 revisited). Consider an arbitrary filter F and a sequence of designs $\{\rho_\varepsilon\}$ defined in Example 4.2. Let for every $\varepsilon > 0$ extend the definition of ρ_ε (that has been defined only on Ω) by setting $\rho_\varepsilon(\mathbf{x}) = 1$ for all $\mathbf{x} \in (\Omega + \operatorname{supp} F) \setminus \Omega$, and $\rho_\varepsilon(\mathbf{x}) = 0$ for all $\mathbf{x} \in \mathbb{R}^d \setminus (\Omega + \operatorname{supp} F)$. Then, $F * \rho_\varepsilon \to 1$ as $\varepsilon \to +0$, uniformly in $\operatorname{cl}\Omega$, and the corresponding sequence of flows converges to a pure Navier–Stokes flow in the domain Ω (case $C = 0$ in Example 4.2).

4.4.2 Slightly compressible fluids

While it seems difficult to imagine a reasonable cure for Example 4.3, because the limiting flow must be zero on Ω with nonzero trace on Γ, we can at least try to get a closed design-to-flow mapping if impenetrable walls do not appear too close to the boundary with non-homogeneous Dirichlet conditions on velocity, as in Example 4.4. The difficulty in the latter example is that in our model the porous wall does not stop, or slow, the incompressible fluid while we use material with positive permeability. At the same time, the limiting domain does not permit any incompressible flow through it, because it is not connected.

We can solve this problem by relaxing the incompressibility requirement $\operatorname{div}\mathbf{u} = 0$ in the system (4.1) [of course, we do not need to require the compatibility condition (4.2) in this case]. For example, we may assume that the fluid is *slightly compressible*, i.e., choose a small $\delta > 0$ and let $\operatorname{div}\mathbf{u} + \delta p = 0$. In fact, it is known that for a fixed domain admitting an incompressible flow, the difference between the regular incompressible and slightly compressible flows is of order δ, i.e., we change model only slightly if δ is small enough. The slightly compressible Navier–Stokes equations are often used as approximations of incompressible ones in so-called *penalty algorithms* [Gun89, Chapter 5]. On the other hand, with the gained maturity of mixed finite element methods, the incompressible system can be equally well solved to approximate the behavior of slightly compressible fluids [Tem01].

Whether one considers slightly compressible Navier–Stokes fluids to be the most suitable mathematical model of the underlying physical flow (see Remark 4.9) or just an accurate approximation of the incompressible Navier–Stokes equations, we make an assumption of slight compressibility because it allows us to achieve the ultimate goal of this paper: to obtain a closed design-to-flow mapping. Again, delaying the precise formulations until Section 4.5, we revisit Example 4.4 to illustrate our point.

Example 4.8 (Example 4.4 revisited). We choose $\delta = 1.0 \cdot 10^{-3}$ and resolve the flow problem of Example 4.4 for $\alpha \in \{1.0, 1.0 \cdot 10^2, 1.0 \cdot 10^4, +\infty\}$. The corresponding flows (calculated in Femlab) are shown in Figure 4.6; in contrast with the incompressible Navier–Stokes case we can see the convergence of flows as domains converge (i.e., as α increases) to a limiting flow, which exists in the compressible case. Note that for small values of α and δ the incompressible and the slightly compressible flows look similar.

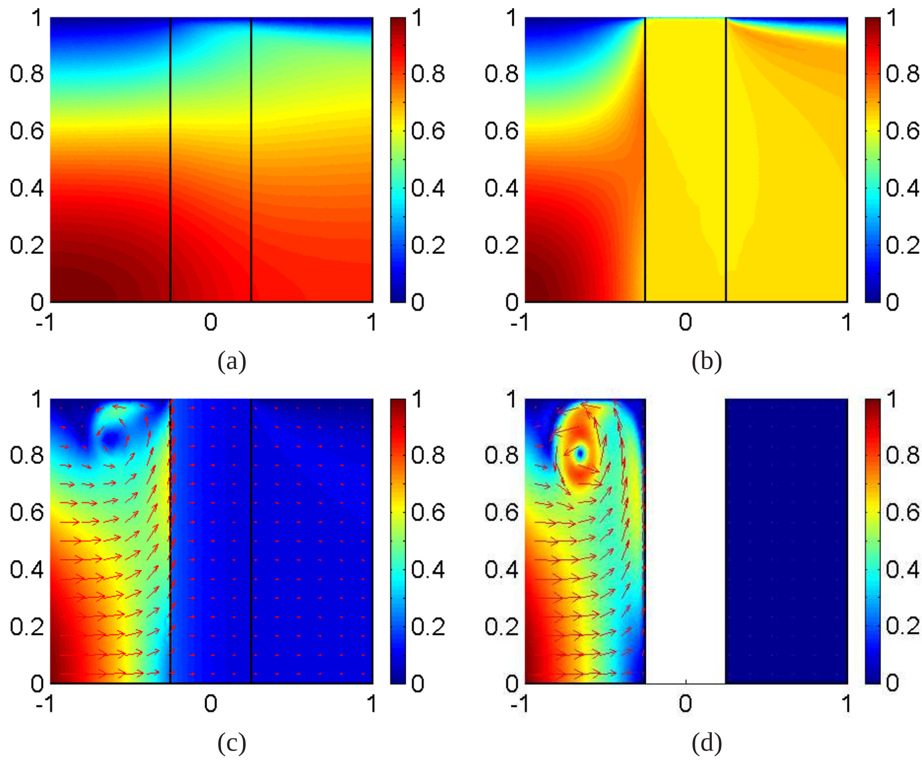


Fig. 4.6: Compressible flow through the porous wall: (a) $\alpha = 1.0$, (b) $\alpha = 1.0 \cdot 10^2$, (c) $\alpha = 1.0 \cdot 10^4$, (d) $\alpha = +\infty$. Compare with Figure 4.5.

Remark 4.9. It is known that the pseudo-constitutive relation $\operatorname{div}\mathbf{u} + \delta p = 0$ lacks an adequate physical interpretation for many important physical flows (e.g., see [HeV95]).

In particular, there is no physical pressure field compatible with the flow shown in Figure 4.6 (d). On the other hand, the pseudo-constitutive relation resulting from the penalty method can still be used as a mathematical method of generating flows approximating those of incompressible viscous fluids. Moreover, the idea of relaxing the incompressibility contraint may also be useful for topology optimization in fluid *dynamics*, where the corresponding relation $\operatorname{div}\mathbf{u} + \delta dp/dt = 0$ is known to be physical.

4.5 Continuity of the design-to-flow mapping

4.5.1 Stokes flows

We start by showing the closedness of the design-to-flow mapping for slightly compressible Stokes flows with homogeneous boundary conditions, and then show the necessary modifications for the inhomogeneous boundary conditions. For the compressible Stokes system the variational formulation is as follows. Given $\rho \in \mathscr{H}$, find the solution to the following minimization problem:

$$\min_{\mathbf{v}\in\mathscr{U}} \left\{ \mathscr{J}^F(\rho,\mathbf{v}) + (2\delta)^{-1}\int_\Omega (\operatorname{div}\mathbf{v})^2 \right\}. \tag{4.5}$$

We note that in the case of homogeneous boundary conditions we have $\mathscr{U} = H_0^1(\Omega)$.

Remark 4.10. Since the condition $\operatorname{div}\mathbf{u} = 0$ is violated, we should replace the term $\int_\Omega |\nabla\mathbf{u}|^2$ in the definition of $\mathscr{J}^\mathscr{S}$ with $\int_\Omega |E(\mathbf{u})|^2$, where $E(\mathbf{u}) = (\nabla\mathbf{u} + \nabla\mathbf{u}^t)/2$ is the linearized rate of strain tensor (cf. [Gun89, Section 4.3]). However, both quadratic forms give rise to equivalent norms on $H_0^1(\Omega)$ (cf. [CaK84, Bre83]) and thus do not affect our theoretical developments in any way. Therefore, we choose to keep the definition of $\mathscr{J}^\mathscr{S}$ for notational simplicity.

In fact, one can go one step further and replace the term $\int_\Omega |\nabla\mathbf{u}|^2$ with $\int_\Omega \mathscr{P}(|E(\mathbf{u})|)$, where $\mathscr{P}$ is a positive convex function verifying certain growth assumptions, thus including non-Newtonian flows into the discussion [FuS00, Chapters 3 and 4]. For some functionals this will not affect the discussion, while for others (e.g., Prandtl-Eyring fluids) we must reconsider the very basic problem statements [such as (4.5)]. Therefore, in this paper we consider Newtonian fluids only (that is, the case $\mathscr{P}(x) = x^2$) and discuss possible extensions in Section 4.8.

Proposition 4.11. *For every design $\rho \in \mathscr{H}$ the optimization problem* (4.5) *has a unique solution $\mathbf{v} \in H^1(\Omega)$ whenever its objective functional is proper w.r.t. $\mathscr{U}$, in particular if $\mathscr{U} = H_0^1(\Omega)$.*

Proof. See Appendix 4.A. □

The proof of the main theorem of this section, Theorem 4.13, which establishes the continuity of the design-to-flow mapping in the case of Stokes flow with homogeneous boundary conditions, heavily depends on the following approximation result. Its proof can be found in the Appendix 4.A.

Proposition 4.12. *Let* $\mathbf{u} \in H_0^1(\Omega)$, $\rho \in \mathscr{H}$, *and* $\mathscr{J}^F(\rho,\mathbf{u}) \le M < +\infty$. *Define also* $\Omega_0 = \{\,\mathbf{x} \in \Omega \mid (F*\rho)(\mathbf{x}) = 0\,\}$. *Then, there exists a sequence* $\{\mathbf{u}_k\} \subset H_0^1(\Omega)$ *such that:*

(*i*) $\operatorname{supp}\mathbf{u}_k \Subset (\Omega \setminus \Omega_0)$;
(*ii*) $\lim_{k\to+\infty}\mathbf{u}_k = \mathbf{u}$, *strongly in* $H_0^1(\Omega)$;
(*iii*) $\limsup_{k\to+\infty}\mathscr{J}^F(\rho,\mathbf{u}_k) \le M$.

Theorem 4.13. *Consider a sequence of designs* $\{\rho_k\} \subset \mathscr{H}$ *and the corresponding sequence of flows* $\{\mathbf{u}_k\} \subset H_0^1(\Omega)$, $k = 1,2,\dots$ (*i.e.*, $\mathbf{u}_k$ *solves the problem* (4.5) *for* ρ_k). *Assume that* $\rho_k \to \rho_0$, *strongly in* $L^1(\Omega + B_R)$, *and* $\mathbf{u}_k \rightharpoonup \mathbf{u}_0$, *weakly in* $H_0^1(\Omega)$. *Then,* $\mathbf{u}_0$ *is the flow corresponding to the limiting design* ρ_0.

Proof. Throughout the proof we denote the optimal value of the optimization problem (4.5) for a given design ρ as $\operatorname{val}(\rho)$. Owing to the weak lower-semicontinuity of $\mathscr{J}$ (cf. [Evg03, Lemma 3.2]) and the weak lower-semicontinuity of $\int_\Omega(\operatorname{div}\mathbf{u})^2$ (cf. [EkT99, Corollary 2.2]) we have that

$$\operatorname{val}(\rho_0) \le \mathscr{J}^F(\rho_0,\mathbf{u}_0) + (2\delta)^{-1}\int_\Omega(\operatorname{div}\mathbf{u}_0)^2 \le \liminf_{k\to\infty}\operatorname{val}(\rho_k). \tag{4.6}$$

If we can also show that $\operatorname{val}(\rho_0) \ge \limsup_{k\to\infty}\operatorname{val}(\rho_k)$, then since $\operatorname{val}(\rho_0) < +\infty$ (owing to Proposition 4.11) we must have equality throughout in (4.6), which means that $\mathbf{u}_0$ solves (4.5) for ρ_0.

Without any loss of generality, we assume that $\operatorname{val}(\rho_0) = \lim_{k\to\infty}\operatorname{val}(\rho_k)$. Let $\widetilde{\mathbf{u}}_0$ be the optimal solution of (4.5), and consider a sequence $\{\mathbf{u}_0^n\} \subset H_0^1(\Omega)$ constructed in Proposition 4.12 for ρ_0 and $\widetilde{\mathbf{u}}_0$. Due to the properties of $\{\mathbf{u}_0^n\}$, for every $\varepsilon > 0$ there exists $N(\varepsilon) \in \mathbb{N}$ such that for all $n > N(\varepsilon)$ it holds that

$$\operatorname{val}(\rho_0) + \varepsilon > \mathscr{J}^F(\rho_0,\mathbf{u}_0^n) + (2\delta)^{-1}\int_\Omega(\operatorname{div}\mathbf{u}_0^n)^2. \tag{4.7}$$

Moreover, strong convergence of ρ_k together with Lipschitz continuity of F imply uniform convergence of $F*\rho_k$ towards $F*\rho_0$ on $\operatorname{cl}\Omega$ (cf. [Bre83, Théorème IV.15]). Since $\mathbf{u}_0^n \Subset \Omega \setminus \{\,\mathbf{x} \in \Omega \mid (F*\rho_0)(\mathbf{x}) = 0\,\}$, it holds that $\alpha(F*\rho_k)$ uniformly converges towards $\alpha(F*\rho_0)$ on $\operatorname{supp}\mathbf{u}_0^n$, and thus there is $K(n,\varepsilon) \in \mathbb{N}$ such that for all $k > K(n,\varepsilon)$ we have

$$\mathscr{J}^F(\rho_0,\mathbf{u}_0^n) + (2\delta)^{-1}\int_\Omega(\operatorname{div}\mathbf{u}_0^n)^2 + \varepsilon > \mathscr{J}^F(\rho_k,\mathbf{u}_0^n) + (2\delta)^{-1}\int_\Omega(\operatorname{div}\mathbf{u}_0^n)^2 \ge \operatorname{val}(\rho_k), \tag{4.8}$$

where the last inequality is due to the feasibility of $\mathbf{u}_0^n$ in (4.5) for the design ρ_k. Combining (4.7) and (4.8), and letting k grow to infinity in the latter we get

$$\operatorname{val}(\rho_0) + 2\varepsilon > \lim_{k\to\infty}\operatorname{val}(\rho_k).$$

Finally, letting ε go to zero, we finish the proof. □

Remark 4.14. Theorem 4.13 shows the epi-convergence of the objective functionals of the ρ-parametric optimization problem (4.5) as the parameters strongly converge in $L^1(\Omega + B_R)$ (cf. [BoS00, p. 41]).

Remark 4.15. We use strong convergence on the space of designs in order to guarantee the Lipschitz continuity (cf. [AuF90, Definition 1.4.5]) of the family of walls $\{\mathbf{x} \in \Omega \mid (F * \rho_k)(\mathbf{x}) = 0\}$, parametrized by $k \in \mathbb{N}$, which is a stronger property than upper-semicontinuity (cf. Proposition 4.5). We need Lipschitz continuity to justify (4.8).

In the case of non-homogeneous boundary conditions, the proof is essentially the same provided we can keep the walls away from the regions of the boundary where injection/suction of the fluid is performed; see Subsection 4.4.2 and Example 4.3 for motivations.

Theorem 4.16. *Consider a sequence of designs* $\{\rho_k\} \subset \mathscr{H}$ *and the corresponding sequence of flows* $\{\mathbf{u}_k\} \subset \mathscr{U}$, $k = 1, 2, \dots$ (*i.e.,* $\mathbf{u}_k$ *solves the problem* (4.5) *for* ρ_k). *Assume that* $\rho_k \to \rho_0$, *strongly in* $L^1(\Omega + B_R)$, *and* $\mathbf{u}_k \rightharpoonup \mathbf{u}_0$, *weakly in* $H^1(\Omega)$. *Further assume that for some positive constants* ε, τ *it holds that*

$$\inf\{\, (F * \rho_k)(\mathbf{x}) \mid k \in \mathbb{N}, \mathbf{x} \in \Omega \cap (\operatorname{supp} \mathbf{g} + B_\varepsilon)\,\} \geq \tau. \tag{4.9}$$

Then, $\mathbf{u}_0$ *is the flow, corresponding to the limiting design* ρ_0 (*i.e.,* $\mathbf{u}_0$ *solves the problem* (4.5) *for* ρ_0).

Proof. Let $\mathbf{w} \in \mathscr{U}$ be a function with $\operatorname{supp} \mathbf{w} \Subset \Omega \cap (\operatorname{supp} \mathbf{g} + B_\varepsilon)$. Then, owing to the additional condition (4.9), the objective functional of (4.5) is finite when evaluated at $\mathbf{w}$, for every ρ_k, $k \in \mathbb{N}$, as well as for ρ_0. Therefore, for every ρ_k, $k \in \mathbb{N}$, (resp., for the limiting design ρ_0) the optimization problem (4.5) admits a unique optimal solution, which can be written as $\mathbf{u}_k = \mathbf{w} + \mathbf{v}_k$, $\mathbf{v}_k \in H^1_0(\Omega)$ (resp., $\widetilde{\mathbf{u}}_0 = \mathbf{w} + \widetilde{\mathbf{v}}_0$, $\widetilde{\mathbf{v}}_0 \in H^1_0(\Omega)$).

The epi-convergence of the mappings $H^1_0(\Omega) \ni \mathbf{v} \to \mathscr{J}^F(\rho, \mathbf{w} + \mathbf{v}) + (2\delta)^{-1} \int_\Omega (\operatorname{div}(\mathbf{w} + \mathbf{v}))^2$ as the parameters ρ strongly converge in $L^1(\Omega + B_R)$ keeping (4.9) true, can be shown exactly as in the proof of Theorem 4.13. The latter implies the claim. □

Remark 4.17. We note that the condition (4.9) is automatically verified for Stokes problems with homogeneous boundary conditions, because the infimum is taken over the empty set in this case ($\operatorname{supp} \mathbf{g} = \emptyset$).

4.5.2 Navier–Stokes flows

In the case of the Navier–Stokes equations things get much more complicated, because we do not seek a minimizer of some functional anymore, and we cannot apply epi-convergence results directly. Nevertheless, we can utilize them to show the closedness of the design-to-flow mappings even in the Navier–Stokes case.

We introduce a general fixed-point framework related to the optimization problem (4.5), and then show (at least for the case of homogeneous boundary conditions) that the slightly compressible Navier–Stokes equations can be considered in this framework.

Let $A(\mathbf{u}, \mathbf{v}) : \mathscr{U} \times \mathscr{U} \to \mathbb{R}$ be a weakly continuous functional, and consider the problem of finding a fixed point of the point-to-set mapping $T_\rho : \mathscr{U} \rightrightarrows \mathscr{U}$ defined for $\rho \in \mathscr{H}$ as

$$T_\rho(\mathbf{u}) = \operatorname*{argmin}_{\mathbf{v} \in \mathscr{U}} \left\{ \mathscr{J}^F(\rho, \mathbf{v}) + (2\delta)^{-1} \int_\Omega (\operatorname{div} \mathbf{v})^2 + A(\mathbf{u}, \mathbf{v}) \right\}. \tag{4.10}$$

Theorem 4.18. *Consider a sequence of designs* $\{\rho_k\} \subset \mathscr{H}$ *and the corresponding sequence of fixed points* $\{\mathbf{u}_k\} \subset \mathscr{U}$, $k = 1, 2, \ldots$ *(i.e.,* $\mathbf{u}_k \in T_{\rho_k}(\mathbf{u}_k)$ *for* T_{ρ_k} *defined by* (4.10)*). Assume that* $\rho_k \to \rho_0$*, strongly in* $L^1(\Omega + B_R)$, $\mathbf{u}_k \rightharpoonup \mathbf{u}_0$*, weakly in* $H^1(\Omega)$*, and* $T(\mathbf{u}_0) \neq \emptyset$*. Further assume that for some positive constants* ε, τ *the condition* (4.9) *is satisfied. Then,* $\mathbf{u}_0 \in T_{\rho_0}(\mathbf{u}_0)$.

Proof. It is enough to show that the objective functionals of the parametric optimization problems (4.10) epi-converge as $(\rho_k, \mathbf{u}_k)$ converge towards $(\rho_0, \mathbf{u}_0)$. This follows from Theorem 4.16, the continuity of A, and [RoW98, Exercise 7.8.(a)]. □

Remark 4.19. In fact, weak continuity of $A(\mathbf{u}, \mathbf{v})$ is an unnecessarily strong requirement. We can repeat the arguments of Theorem 4.13 with straightforward modifications and prove Theorem 4.18 under the following weaker assumptions on A:

(*i*) $A(\mathbf{u}, \mathbf{u}) \leq \liminf_{k\to\infty} A(\mathbf{u}_k, \mathbf{u}_k)$ whenever $\mathbf{u}_k \rightharpoonup \mathbf{u}$, weakly in $\mathscr{U}$; and

(*ii*) $A(\mathbf{u}, \mathbf{v}) \geq \limsup_{k\to\infty} A(\mathbf{u}_k, \mathbf{v}_k)$ whenever $\mathbf{u}_k \rightharpoonup \mathbf{u}$, weakly in $\mathscr{U}$, and $\mathbf{v}_k \to \mathbf{v}$, strongly in $\mathscr{U}$.

As an example application of Theorem 4.18, we consider a particular penalty formulation of the incompressible Navier–Stokes equations with homogeneous boundary conditions studied in [CaK84]. A more general treatment is of course possible, including inhomogeneous boundary conditions and variants of slightly compressible Navier–Stokes equations; the main difference is in the number of technical details to be covered.

To put the penalty formulation considered in [CaK84] (of course, without the control term α) into the framework of (4.10) we define

$$A(\mathbf{u}, \mathbf{v}) = \int_\Omega (\mathbf{u} \cdot \nabla \mathbf{u}) \cdot \mathbf{v} + 2^{-1} \int_\Omega (\mathbf{u} \cdot \mathbf{v}) \operatorname{div} \mathbf{u}. \tag{4.11}$$

We note that the last integral adds an additional stability to the penalty algorithm [CaK84] and identically equals zero in the incompressible case; we can thus expect that the effects of its presence can be almost neglected in the slightly compressible case. Owing to [CaK84, Lemma 2.7], the functional A defined in (4.11) is weakly continuous on $H_0^1(\Omega) \times H_0^1(\Omega)$, and in order to apply Theorem 4.18 it remains to establish an analogue of Proposition 4.11.

Proposition 4.20. *With* $\mathscr{U} = H_0^1(\Omega)$ *and* A *defined by* (4.11)*, the fixed-point problem* (4.10) *admits solutions for every* $\rho \in \mathscr{H}$.

Proof. The functional $A(\mathbf{u}, \cdot)$ is linear and continuous on $H_0^1(\Omega)$. Applying Lemma 4.29 to the "force" $\langle \mathbf{f}, \cdot \rangle + A(\mathbf{u}, \cdot) \in H^{-1}(\Omega)$, we conclude that for every $\rho \in \mathscr{H}$ the operator $T_\rho(\mathbf{u})$ is single-valued and completely continuous.

Now, assume that $\mathbf{w} = \sigma T_\rho(\mathbf{w})$ for some $\mathbf{w} \in H_0^1(\Omega)$ and $0 < \sigma \leq 1$. Then, using the fact that $A(a\mathbf{w}, b\mathbf{w}) = a^2 b A(\mathbf{w}, \mathbf{w}) = 0$ for all $a, b \in \mathbb{R}$, where the last equality is by [CaK84, Lemma 2.4], and evaluating the objective function of (4.10) at $\sigma^{-1}\mathbf{w}$ (the optimal solution) and $\mathbf{0} \in H_0^1(\Omega)$ we get the inequality

$$\frac{\nu}{2\sigma^2} \int_\Omega |\nabla \mathbf{w}|^2 + \frac{1}{2\sigma^2 \delta} \int_\Omega (\operatorname{div} \mathbf{w})^2 + \frac{1}{2\sigma^2} \int_\Omega \alpha(F * \rho)|\mathbf{w}|^2 - \frac{1}{\sigma} \int_\Omega \mathbf{f} \cdot \mathbf{w} \leq 0,$$

which implies that $\|\mathbf{w}\|_{H_0^1(\Omega)} \le C\|\mathbf{f}\|_{L^2(\Omega)}$ holds, for some constant C independent of σ. An application of the Leray-Schauder Theorem (cf. [GrD03, § 6, Theorem 5.4]) concludes the proof. □

Remark 4.21. While the mapping $(\rho, \mathbf{u}) \to T_\rho(\mathbf{u})$ is in many cases single-valued for every pair $(\rho, \mathbf{u})$, there might be more than one solution to the fixed point problem (4.10) with this operator. In other words, we do not assume that the compressible Navier-Stokes system admits a unique solution.

Remark 4.22. We can use another popular weak formulation of slightly compressible Navier–Stokes equations (e.g., see [LCW95]), identifying

$$A(\mathbf{u},\mathbf{v}) = \frac{1}{2}\left(\int_\Omega (\mathbf{u}\cdot\nabla\mathbf{u})\cdot\mathbf{v} - \int_\Omega (\mathbf{u}\cdot\nabla\mathbf{v})\cdot\mathbf{u} \right).$$

Our results hold even in this case without any changes.

Remark 4.23. Of course, the fixed-point framework (4.10) is not bounded to Navier–Stokes equations. For example, putting, for some $\mathbf{u}_0 \in \mathbb{R}^d$,

$$A(\mathbf{u},\mathbf{v}) = \int_\Omega (\mathbf{u}_0\cdot\nabla\mathbf{u})\cdot\mathbf{v},$$

we can show continuity results for Oseen flows [Gal94]. This type of flow is probably not very interesting in bounded domains Ω, but illustrates the possible uses of the fixed-point formulation (4.10). Finally, we note that setting $A \equiv 0$ we recover the original Stokes problem.

4.6 Existence of optimal solutions

4.6.1 Ensuring strong convergence of designs and condition (4.9)

The results established in Section 4.5 all require strong convergence of designs in $L^1(\Omega + B_R)$. In order to guarantee convergence we need to embed our controls $\mathscr{H}$ into some space that is more regular than $L^\infty(\mathbb{R}^d)\cap L^1(\mathbb{R}^d)$. The most appropriate choice, in our opinion, is the space $SBV(\mathbb{R}^d)$ (cf. [AFP00]), which is typically used for perimeter constrained topology optimization (see [BeS03, p. 31] and references therein; see also [FGR99, Pet99, HBJ96]). Other choices are possible, including $W^{1,1}(\mathbb{R}^d)\cap L^\infty(\mathbb{R}^d)$ (that is, imposing "slope constraints" on the design space; see [PeS98], but also [Bor01, BoC03]). Bounds on the perimeter, or slope, may be introduced into the problem directly as constraints, or added as penalties to the objective function.

Regardless of the particular method used, we get the required property: $\rho_k \rightharpoonup \rho$, weakly in $\mathscr{H}$, implies $\rho_k \to \rho$, strongly in $L^1(\Omega + B_R)$, allowing us to establish the closedness of the design-to-flow mappings.

As for the condition (4.9), it can be easily verified if we require in addition that every design $\rho \in \mathscr{H}$, satisfying the bounds $0 \le \rho \le 1$, a.e. on $\mathbb{R}^d$, also satisfies $\rho \ge \tau$, a.e. on $\operatorname{supp}\mathbf{g} + B_{R+\varepsilon}$, for some positive constants ε, τ.

4.6.2 An abstract flow topology optimization problem

Now we are ready to formally discuss the well-posedness of an abstract flow topology optimization problem:

$$\begin{aligned}
&\min_{(\rho,\mathbf{u})} \mathscr{F}(\rho,\mathbf{u}),\\
&\text{s.t.} \begin{cases} (\rho,\mathbf{u}) \in \mathscr{Z}, \\ \quad\quad \mathbf{u} \in T_\rho(\mathbf{u}). \end{cases}
\end{aligned} \tag{4.12}$$

The previous results imply the following theorem.

Theorem 4.24. *Let $\mathscr{Z}$ be a nonempty weakly compact subset of $\mathscr{H} \times \mathscr{U} \subset SBV(\mathbb{R}^d) \times H^1(\Omega)$, and let for all $\rho \in \mathscr{H}$ the assumption* (4.9) *be verified* (*see the discussion in Subsection 4.6.1*). *We also assume that A* [*which defines the mapping T_ρ via* (4.10)] *enjoys the conditions of Remark 4.19, and that for every $\rho \in \mathscr{H}$ the fixed-point problem* (4.10) *admits solutions. Finally, let $\mathscr{F} : SBV(\mathbb{R}^d) \times H^1(\Omega) \to \mathbb{R}$ be weakly l.s.c. Then, there exists at least one optimal solution to the abstract flow topology optimization problem* (4.12).

Proof. Essentially this is a Weierstrass Theorem adapted to our specific notation, because the hypotheses and Theorem 4.18 imply that the feasible set of (4.12) is nonempty and weakly compact. □

Remark 4.25. If the assumptions of Theorem 4.24 about the flow model are satisfied, we may set

$$\mathscr{Z} = \{\, (\rho,\mathbf{u}) \in \mathscr{Z}_0 \times \mathscr{U} \mid \mathscr{G}(\rho,\mathbf{u}) \leq C \,\},$$

where $\mathscr{Z}_0$ is a nonempty weakly compact subset of $\mathscr{H} \subset SBV(\mathbb{R}^d)$ verifying condition (4.9), $\mathscr{G}(\rho,\mathbf{u})$ is an arbitrary weakly l.s.c. functional, which is in addition coercive in $\mathbf{u}$, uniformly w.r.t. ρ, and $C \in \mathbb{R}$ is an arbitrary constant but such that $\mathscr{Z} \neq \emptyset$.

In particular, we may set $\mathscr{G} = \mathscr{J}$, or $\mathscr{G} = \mathscr{J}^F$ (cf. [Evg03, Lemma 3.2]).

At last, we note that assumptions of Theorem 4.24 about the solvability of the fixed-point problem for every feasible design ρ are verified in many practical situations. For example, we have shown that they are satisfied for Stokes equations (see Proposition 4.11 and Remark 4.23) and for Navier–Stokes equations with homogeneous boundary conditions (see Proposition 4.20).

4.7 Computational issues

In this section we briefly discuss two topics that are standard in topology optimization with specialization to flow topology optimization problems. Throughout the section we will use problem (4.12) as a model example, and we assume that the assumptions of Theorem 4.24 are verified without further notice.

4.7.1 Approximation with sizing optimization problems

Clearly, no finite element software can be applied to the problem (4.10) if $\alpha(F*\rho)$ is allowed to become arbitrarily large; from the practical point of view the theory of Section 4.5 implying the existence of optimal solutions to (4.12) is pointless, unless we can describe a computational procedure capable of finding approximations of these optimal solutions. In fact, once we have proved Theorem 4.18 the latter goal can be easily accomplished. For arbitrary $\varepsilon > 0$, consider the set $\mathscr{Z}_\varepsilon = \{\,(\rho,\mathbf{u}) \in \mathscr{Z} \mid \rho \geq \varepsilon, \text{a.e.}\}$, i.e., only designs with porosity uniformly bounded away from zero are allowed, implying in particular the uniform bound $\alpha(F*\rho) \leq \varepsilon^{-1} - 1$, for every $(\rho,\mathbf{u}) \in \mathscr{Z}_\varepsilon$.

Then, the following easy statement holds.

Proposition 4.26. *Assume that the sequence $\{\mathscr{Z}_\varepsilon\}$ is lower-semicontinuous in Painlevé-Kuratowski sense (topology in $\mathscr{H} \times \mathscr{U}$ being the strong one), namely*

$$\operatorname*{Lim\,inf}_{\varepsilon\to+0} \mathscr{Z}_\varepsilon = \mathscr{Z}, \tag{4.13}$$

(*in particular, $\mathscr{Z}_\varepsilon \neq \emptyset$ for all small $\varepsilon > 0$*). *Let further, for every small $\varepsilon > 0$, $(\rho_\varepsilon,\mathbf{u}_\varepsilon)$ denote a globally optimal solution of an approximating problem, obtained from* (4.12) *substituting $\mathscr{Z}_\varepsilon$ in place of $\mathscr{Z}$. Then, an arbitrary limit point of $\{(\rho_\varepsilon,\mathbf{u}_\varepsilon)\}$ (and there is at least one) is a globally optimal solution of the limiting problem* (4.12).

Proof. All claims easily follow from the uniform inclusion $\mathscr{Z}_\varepsilon \subset \mathscr{Z}$, Theorems 4.24 and 4.18, and finally [BoS00, Proposition 4.4]. □

The assumption (4.13) is probably easier to check in every particular case rather than to develop a general sufficient condition implying it; we only mention that for typical constraints in topology optimization, such as constraints on volume and on the perimeter, it is easily verified.

In general, there is a substantial amount of literature on the topic of approximation of topology optimization problems using sizing ones. (See the bibliographical notes (16) in [BeS03] for a survey of the situation in the topology optimization of linearly elastic materials; also see [Evg03, Section 6] for results on incompressible stokesian flows, and [KPTK03, Appendix A.2] for a similar problem arising in the design of flow networks.) Cases of interest in such literature are when some of the underlying assumptions of Proposition 4.26 are violated, such as the compactness of $\mathscr{Z}$ or $\mathscr{Z}_\varepsilon$, or the assumption (4.13); in some particular situations it is nevertheless possible to prove statements similar to Proposition 4.26. We do not try to generalize our result in this direction, because computationally the problem (4.12) is already extremely demanding for realistic flows, and complicated constraints violating (4.13) are hardly necessary in practical situations.

4.7.2 Control of intermediate densities

Starting with the problem of distributing the solid material inside a control volume Ω so as to minimize some objective functional dependent on the flow, we expect an optimal design of the type $\rho = \chi_A$, where $A \subset \Omega$ is a flow region ("black–white" designs). Usually, this is a very naïve expectation [BeS03, Section 1.3.1]; however, there are some exceptions, such as the minimum-power design of domains for Stokes flows [BoP03, Evg03], or the design of one-dimensional wave-guides for stopping wave propagation [Bel03].

However, if we use a filter, it is simply impossible to obtain optimal distributions of material assuming *only* values zero or one (not counting the trivial designs $F * \rho \equiv 0$ and $F * \rho \equiv 1$), because $F * \chi_A$ is a continuous function, and the "edges" ∂A will be "smoothed out" by the filter. One possible way to reduce the amount of porous material in the final optimal design $F * \rho$ is to use a filter of a smaller radius. This may or may not work as expected — since the control problem (4.12) is non-convex, the optimal designs may change significantly as we vary the radius only slightly.

Another possibility is to add a penalty term $\mu \mathscr{J}^{\mathscr{D}}(F * \rho, \mathbf{u})$, for some positive μ, requiring that the power dissipation due to the flow through the porous part of the domain should be relatively small [Evg03, Section 5]. We must warn that increasing penalty μ might lead to unexpected results, because as we have already mentioned, the presence of the filter *requires* the presence of porous regions in the domain (except for trivial cases), thus the sequence of designs may converge to either one of those trivial designs, or $\mu \mathscr{J}^{\mathscr{D}}(F * \rho, \mathbf{u})$ may grow indefinitely. Therefore, suitable values of μ should be obtained in each case experimentally.

At last, various restriction or regularization techniques that are designed to control the amount of "microstructural material" in topology optimization of linearly elastic structures may be used for similar purposes in our case. We already mentioned the regularized intermediate density control method [BoP01]; other possible choices may be found in [BeS03, bibliographical notes (8)].

4.8 Conclusions and further research

We have considered the topology optimization of fluid domains in a rather abstract setting, and established the closedness of design-to-flow mappings for a general family of slightly compressible fluids, whose behavior is characterized by the fixed-point formulation (4.10). We used the notion of epi-convergence of optimization problems as a main analytical tool (cf. [BoS00, Proposition 4.6]) that allows us to treat very ill-behaving functionals, which arise due to the fact that we allow completely impenetrable walls to appear in the design domain.

It is of course of great engineering interest to perform numerical experiments with topology optimization of slightly compressible fluids for various objective functionals, theoretical foundations for which are established in this paper. Provided a stable solver of the underlying flow problem is available, it should not be a difficult task to combine it with the optimization code; in the end, the ease of integration with FEM software is one of the main reasons why topology optimization techniques are widely accepted and still gain popularity in many fields of physics and engineering [BeS03]. In fact, one such successful attempt of integrating topology optimization with Femlab is done for incompressible Navier–Stokes fluids [GH03]. Unfortunately, at the time of writing this code was not available to the author. We hope to be able to perform numerical computations in the near future.

The motivation for relaxing the incompressibility requirement is found in Section 4.3.2; however, if one is not convinced, and for the sake of completeness it would be interesting to prove the main approximation result, Proposition 4.12, for divergence-free functions, from which the rest of the theory should follow for incompressible fluids as well.

The method we used is of course not bound to Newtonian fluids. It seems that our results should hold for many common non-Newtonian fluids, including power-law, Bingham, and Powell-Eyring models (cf. [FuS00, Chapter 3]), without any major modifications (cf. Remark 4.10). Additional work is obviously needed for fluids of Prandtl-Eyring type [FuS00, Chapter 4]; we however feel that the special treatment this (mathematically) exotic type of fluids deserves lies well outside the scope of this paper.

At last, but not the least, we feel it is important to establish the existence of solutions, or construct a disproving counter-example, for the "original" problem of power minimization for incompressible Navier–Stokes fluids without the use of filtered designs. While we have shown that this problem looks ill-posed and is probably unsuitable for practical numerical computations, knowing whether optimal solutions exist would greatly contribute to the deeper understanding of Navier–Stokes flows and affect the further development in the area of topology optimization of fluids.

Acknowledgement: This research is supported by the Swedish Research Council (grant 621-2002-5780).

References

[AFP00] Luigi Ambrosio, Nicola Fusco, and Diego Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 2000.

[All90a] Grégoire Allaire. Homogenization of the Navier-Stokes equations in open sets perforated with tiny holes. I. Abstract framework, a volume distribution of holes. *Arch. Rational Mech. Anal.*, 113(3):209–259, 1990.

[All90b] Grégoire Allaire. Homogenization of the Navier-Stokes equations in open sets perforated with tiny holes. II. Noncritical sizes of the holes for a volume distribution and a surface distribution of holes. *Arch. Rational Mech. Anal.*, 113(3):261–298, 1990.

[AuF90] Jean-Pierre Aubin and Hélène Frankowska. *Set-Valued Analysis*. Birkhäuser Boston Inc., Boston, MA, 1990.

[BeD92] François Beux and Alain Dervieux. Exact-gradient shape optimization of a 2-D Euler flow. *Finite Elem. Anal. Des.*, 12(3-4):281–302, 1992.

[Bel03] José Carlos Bellido. Existence of classical solutions for a one-dimensional optimal design problem in wave propagation, 2003. Preprint, Mathematical Institute, University of Oxford, Oxford, UK. Submitted for publication.

[BeS03] Martin P. Bendsøe and Ole Sigmund. *Topology Optimization: Theory, Methods, and Applications*. Springer-Verlag, Berlin, 2003.

[BFCS92] Juan A. Bello, Enrique Fernández Cara, and Jacques Simon. Optimal shape design for Navier-Stokes flow. In *System modelling and optimization (Zurich, 1991)*, volume 180 of *Lecture Notes in Control and Inform. Sci.*, pages 481–489. Springer, Berlin, 1992.

[BoB95] Jeff Borggaard and John Burns. A sensitivity equation approach to shape optimization in fluid flows. In *Flow control (Minneapolis, MN, 1992)*, volume 68 of *IMA Vol. Math. Appl.*, pages 49–78. Springer, New York, 1995.

[BoC03] Blaise Bourdin and Antonin Chambolle. Design-dependent loads in topology optimization. *ESAIM Control Optim. Calc. Var.*, 9:19–48, 2003.

[BoP01] Thomas Borrvall and Joakim Petersson. Topology optimization using regularized intermediate density control. *Comput. Methods Appl. Mech. Engrg.*, 190(37–38):4911–4928, 2001.

[BoP03] Thomas Borrvall and Joakim Petersson. Topology optimization of fluids in Stokes flow. *Internat. J. Numer. Methods Fluids*, 41(1):77–107, 2003.

[Bor01] T. Borrvall. Topology optimization of elastic continua using restriction. *Arch. Comput. Methods Engrg.*, 8(4):351–385, 2001.

[BoS00] J. Frédéric Bonnans and Alexander Shapiro. *Perturbation Analysis of Optimization Problems*. Springer-Verlag, New York, 2000.

[Bou01] Blaise Bourdin. Filters in topology optimization. *Internat. J. Numer. Methods Engrg.*, 50(9):2143–2158, 2001.

[Bre83] Haïm Brezis. *Analyse fonctionnelle*. Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree]. Masson, Paris, 1983. Théorie et applications. [Theory and applications].

[BrT01] Tyler E. Bruns and Daniel A. Tortorelli. Topology optimization of non-linear elastic structures and compliant mechanisms. *Comput. Methods Appl. Mech. Engrg.*, 190(26–27):3443–3459, 2001.

[CaK84] G. F. Carey and R. Krishnan. Penalty finite element method for the Navier-Stokes equations. *Comput. Methods Appl. Mech. Engrg.*, 42(2):183–224, 1984.

[ChG02] A. Chakib and T. Ghemires. Résolution d'un problème d'écoulement par la méthode d'optimisation de forme. *Math-Rech. Appl.*, 4:47–64, 2002.

[Dar02] Olivier Darrigol. Between hydrodynamics and elasticity theory: the first five births of the Navier-Stokes equation. *Arch. Hist. Exact Sci.*, 56(2):95–150, 2002.

[DRSS96] Quang Vinh Dinh, Gilbert Rogé, Cyrille Sevin, and Bruno Stoufflet. Shape optimization in computational fluid dynamics. *Rev. Européenne Élém. Finis*, 5(5-6):569–594, 1996.

[DzZ99] Raja Dziri and Jean-Paul Zolésio. Dynamical shape control in non-cylindrical Navier-Stokes equations. *J. Convex Anal.*, 6(2):293–318, 1999.

[EkT99] Ivar Ekeland and Roger Témam. *Convex analysis and variational problems*, volume 28 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, english edition, 1999. Translated from the French.

[Evg03] Anton Evgrafov. On the limits of porous materials in the topology optimization of Stokes flows. Blå serien, Department of Mathematics, Chalmers University of Technology, Gothenburg, Sweden, 2003. Submitted for publication.

[Fei03] Eduard Feireisl. Shape optimization in viscous compressible fluids. *Appl. Math. Optim.*, 47(1):59–78, 2003.

[FGR99] P. Fernandes, J. M. Guedes, and H. Rodrigues. Topology optimization of three-dimensional linear elastic structures with a constraint on "perimeter". *Comput. & Structures*, 73(6):583–594, 1999.

[FuS00] Martin Fuchs and Gregory Seregin. *Variational methods for problems from plasticity theory and for generalized Newtonian fluids*, volume 1749 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2000.

[Gal94] Giovanni P. Galdi. *An introduction to the mathematical theory of the Navier-Stokes equations. Vol. I*, volume 38 of *Springer Tracts in Natural Philosophy*. Springer-Verlag, New York, 1994. Linearized steady problems.

[GH03] Allan Gersborg-Hansen. Topology optimization of incompressible Newtonian flows at moderate Reinolds numbers. Master's thesis, Department of Mechanical Engineering, Technical University of Denmark, December 2003.

[GKM00] Max D. Gunzburger, Hongchul Kim, and Sandro Manservisi. On a shape control problem for the stationary Navier-Stokes equations. *M2AN Math. Model. Numer. Anal.*, 34(6):1233–1258, 2000.

[GrD03] Andrzej Granas and James Dugundji. *Fixed point theory*. Springer Monographs in Mathematics. Springer-Verlag, New York, 2003.

[GuK98] Max D. Gunzburger and Hongchul Kim. Existence of an optimal solution of a shape control problem for the stationary Navier-Stokes equations. *SIAM J. Control Optim.*, 36(3):895–909, 1998.

[GuM02] M. Gunzburger and S. Manservisi. Flow matching by shape design for the Navier-Stokes system. In *Optimal control of complex structures (Oberwolfach, 2000)*, volume 139 of *Internat. Ser. Numer. Math.*, pages 279–289. Birkhäuser, Basel, 2002.

[Gun89] Max D. Gunzburger. *Finite element methods for viscous incompressible flows*. Computer Science and Scientific Computing. Academic Press Inc., Boston, MA, 1989.

[Gun03] Max D. Gunzburger. *Perspectives in flow control and optimization*. Advances in Design and Control. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2003.

[HBJ96] R. B. Haber, M. P. Bendsøe, and C. S. Jog. Perimeter constrained topology optimization of continuum structures. In *IUTAM Symposium on Optimization of Mechanical Systems (Stuttgart, 1995)*, volume 43 of *Solid Mech. Appl.*, pages 113–120. Kluwer Acad. Publ., Dordrecht, 1996.

[HeV95] J. C. Heinrich and C. A. Vionnet. The penalty method for the Navier-Stokes equations. *Arch. Comput. Methods Engrg.*, 2(2):51–65, 1995.

[Hor97] Ulrich Hornung, editor. *Homogenization and porous media*, volume 6 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, 1997.

[Kim01] Hongchul Kim. Penalized approach and analysis of an optimal shape control problem for the stationary Navier-Stokes equations. *J. Korean Math. Soc.*, 38(1):1–23, 2001.

[KPTK03] Anders Klarbring, Joakim Petersson, Bo Torstenfelt, and Matts Karlsson. Topology optimization of flow networks. *Comput. Methods Appl. Mech. Engrg.*, 192(35-36):3909–3932, 2003.

[Kuf80] Alois Kufner. *Weighted Sobolev spaces*, volume 31 of *Teubner-Texte zur Mathematik [Teubner Texts in Mathematics]*. BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1980. With German, French and Russian summaries.

[LCW95] San Yih Lin, Yan Shin Chin, and Tsuen Muh Wu. A modified penalty method for Stokes equations and its applications to Navier-Stokes equations. *SIAM J. Sci. Comput.*, 16(1):1–19, 1995.

[MoP01] B. Mohammadi and O. Pironneau. *Applied shape optimization for fluids*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, 2001.

[NRS95] P. Neittaanmäki, V. Rivkind, and G. Serëgin. About optimal shape design in fluid dynamics. *Optimal Control Appl. Methods*, 16(2):143–148, 1995.

[OkK00] H. Okumura and M. Kawahara. Shape optimization of body located in incompressible Navier-Stokes flow based on optimal control theory. *CMES Comput. Model. Eng. Sci.*, 1(2):71–77, 2000.

[PeS98] Joakim Petersson and Ole Sigmund. Slope constrained topology optimization. *Internat. J. Numer. Methods Engrg.*, 41(8):1417–1434, 1998.

[Pet99] J. Petersson. Some convergence results in perimeter-controlled topology optimization. *Comput. Methods Appl. Mech. Engrg.*, 171(1-2):123–140, 1999.

[Pir73] O. Pironneau. On optimum profiles in Stokes flow. *J. Fluid Mech.*, 59:117–128, 1973.

[Pir74] O. Pironneau. On optimum design in fluid mechanics. *J. Fluid Mech.*, 64:97–110, 1974.

[RoW98] R. Tyrrell Rockafellar and Roger J.-B. Wets. *Variational Analysis*. Springer-Verlag, Berlin, 1998.

[Sig97] Ole Sigmund. On the design of compliant mechanisms using topology optimization. *Mech. Struct. Mach.*, 25(4):493–524, 1997.

[SiP98] Ole Sigmund and Joakim Petersson. Numerical instabilities in topology optimization: A survey on procedures dealing with checkerboards, mesh-dependencies and local minima. *Struct. Multidisc. Optim.*, 16(1):68–75, 1998.

[SoZ92] Jan Sokołowski and Jean-Paul Zolésio. *Introduction to shape optimization*, volume 16 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1992.

[StS94] Srdjan Stojanovic and Thomas P. Svobodny. A variational approach to shape optimization for the Navier-Stokes equations. In *Optimal control of differential equations (Athens, OH, 1993)*, volume 160 of *Lecture Notes in Pure and Appl. Math.*, pages 281–291. Dekker, New York, 1994.

[Sue96] A. P. Suetov. A shape optimization algorithm for the minimum drag problem in Stokes flow. In *Modelling and optimization of distributed parameter systems (Warsaw, 1995)*, pages 375–383. Chapman & Hall, New York, 1996.

[Tem01] Roger Temam. *Navier-Stokes equations*. AMS Chelsea Publishing, Providence, RI, 2001. Reprint of the 1984 edition.

[Ton03a] B. A. Ton. Optimal shape control problem for the Navier-Stokes equations. II. *J. Optim. Theory Appl.*, 118(1):135–156, 2003.

[Ton03b] Bui An Ton. Optimal shape control problem for the Navier-Stokes equations. *SIAM J. Control Optim.*, 41(6):1733–1747, 2003.

[Tri78] Hans Triebel. *Interpolation theory, function spaces, differential operators*, volume 18 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, 1978.

[vBS02] E. H. van Brummelen and A. Segal. Adjoint shape optimization for steady free-surface flows. *Internat. J. Numer. Methods Fluids*, 40(3-4):605–614, 2002. ICFD Conference on Numerical Methods for Fluid Dynamics, Part II (Oxford, 2001).

4.A Appendix

Proposition 4.27. *Under the assumptions on the filter F made in this paper the condition* (4.4) *is satisfied.*

Proof. Let $|\mathbf{x}| = h$ and define $S_h = \{\mathbf{y} \in \operatorname{supp} F \mid \operatorname{dist}(\mathbf{y}, \partial \operatorname{supp} F) \leq h\}$. $|S_h| \leq h\mathscr{H}_{d-1}(\partial \operatorname{supp} F)$, where $\mathscr{H}_{d-1}(\partial \operatorname{supp} F)$ is the Hausdorff measure of $\partial \operatorname{supp} F$ (i.e., perimeter of $\partial \operatorname{supp} F$). Moreover, $\sup_{\mathbf{y}\in S_h} F(\mathbf{y}) \leq Lh$, where L is the Lipschitz constant for F. Thus,

$$(F * \chi_{\mathbb{R}^d \setminus \operatorname{supp} F})(\mathbf{x}) \leq \int_{S_h} F \leq h^2 L \mathscr{H}_{d-1}(\partial \operatorname{supp} F). \qquad \square$$

Proof (of Proposition 4.11). The function $\mathscr{U} \ni \mathbf{v} \to \mathscr{J}^F(\rho,\mathbf{v}) + (2\delta)^{-1}\int_\Omega(\operatorname{div}\mathbf{v})^2$ is strongly convex and l.s.c. (in particular, owing to Poincaré inequality [Bre83, Corollaire IX.19]). Of course, if it is also proper w.r.t. $\mathscr{U}$ we get both existence and uniqueness of solutions. In particular, the last property holds in the case $\mathbf{0} \in \mathscr{U}$. □

We will make use of the following statement, which can be found in the proof of [Kuf80, Theorem 9.7]. We remark that Ω needs not to be regular for this to hold (cf. [Tri78, Section 3.2.3]).

Lemma 4.28. *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain. Then, for every $h > 0$ there is a cut-off function $F_h \in C_0^\infty(\Omega)$ such that:*
 (i) $0 \le F_h \le 1$,
 (ii) $\forall \mathbf{x} \in \Omega : |\boldsymbol{\nabla} F_h(x)| \le C_1 h^{-1}$ *for a suitable constant $C_1 > 0$, and*
 (iii) $F_h \equiv 1$ *on $\Omega \setminus \Omega_h$, where $\Omega_h = \{\, \mathbf{x} \in \Omega \mid \operatorname{dist}(\mathbf{x},\partial\Omega) \le h \,\}$.*
We can always (and in fact will) assume that $F_h \in C_0^\infty(\mathbb{R}^d)$, extending F_h by zero on $\mathbb{R}^d \setminus \Omega$.

The proof of the Proposition 4.12 essentially mimics the proof of [Kuf80, Theorem 9.7]; however, we adapt it to our notation. The most important difference is the fact that the specific growth condition holds not on the whole boundary of our domain but rather only on part of it, therefore we cannot apply the cited theorem directly.

Proof (of Proposition 4.12). We apply Lemma 4.28 to a set $(\Omega + B_R) \setminus \Omega_0$ to obtain a family of "cut-off" functions $\{F_h\} \subset C_0^\infty((\Omega + B_R) \setminus \Omega_0)$, $h > 0$, and set $\mathbf{u}_h = F_h\mathbf{u}$ on Ω.

Defining u_h in such a way implies that $\{\mathbf{u}_h\} \subset H_0^1(\Omega)$ and clearly gives us (i) and the uniform estimation $\int_\Omega \alpha(F*\rho)|\mathbf{u}_h|^2 \le \int_\Omega \alpha(F*\rho)|\mathbf{u}|^2$. Thus it suffices to verify (ii) to obtain (iii) as well.

Define $\Omega_h = \{\, \mathbf{x} \in \Omega \setminus \Omega_0 \mid \operatorname{dist}(\mathbf{x},\Omega_0) \le h \,\}$. Since $\mathbf{u} - \mathbf{u}_h = (1-F_h)\mathbf{u}$, $\boldsymbol{\nabla}\mathbf{u} - \boldsymbol{\nabla}\mathbf{u}_h = (1-F_h)\boldsymbol{\nabla}\mathbf{u} - \boldsymbol{\nabla}(1-F_h)\cdot\mathbf{u}$, and $\operatorname{supp}(1-F_h) \subset \Omega_h$, it is necessary to estimate the differences only on Ω_h.

$$\lim_{h\to+0} \|\mathbf{u} - \mathbf{u}_h\|^2_{L^2(\Omega)} \le \lim_{h\to+0} \int_{\Omega_h} (1-F_h)^2|\mathbf{u}|^2 = 0,$$

because $0 \le F_h \le 1$, $\mathbf{u} \in H^1(\Omega)$, and $|\Omega_h| \to 0$ as $h \to +0$.

Similarly,

$$\lim_{h\to+0} \|\boldsymbol{\nabla}\mathbf{u} - \boldsymbol{\nabla}\mathbf{u}_h\|^2_{L^2(\Omega)} \le \lim_{h\to+0} \int_{\Omega_h} (1-F_h)^2|\boldsymbol{\nabla}\mathbf{u}|^2 + \lim_{h\to+0} \int_{\Omega_h} |\boldsymbol{\nabla}F_h|^2|\mathbf{u}|^2,$$

and the first limit is zero, as before, since $0 \le F_h \le 1$, $\mathbf{u} \in H^1(\Omega)$, and $|\Omega_h| \to 0$ as $h \to +0$. We estimate the last integral as

$$\int_{\Omega_h} |\boldsymbol{\nabla}F_h|^2|\mathbf{u}|^2 \le C_1^2 \int_{\Omega_h} h^{-2}|\mathbf{u}|^2 \le C_1^2 C^{-1} \int_{\Omega_h} \alpha(F*\rho)|\mathbf{u}|^2, \tag{4.14}$$

where the constant C_1 is given by Lemma 4.28, and the last inequality holds owing to the filter growth condition (4.4). Owing to the bound $\mathscr{J}^F(\rho,\mathbf{u}) < +\infty$, the last integral converges to zero as h does, and thus the proof is complete. □

The following fact is very well known for elliptic forms; we only show that possible infinite values if α do not change it.

Lemma 4.29. *For every $\rho \in \mathscr{H}$,*

$$H^{-1}(\Omega) \ni \mathbf{f} \rightrightarrows \operatorname*{argmin}_{\mathbf{v}\in H_0^1(\Omega)} \left\{ \frac{\nu}{2}\int_\Omega \nabla\mathbf{v}\cdot\nabla\mathbf{v} + \frac{1}{2\delta}\int_\Omega (\operatorname{div}\mathbf{v})^2 + \frac{1}{2}\int_\Omega \alpha(F*\rho)\mathbf{v}\cdot\mathbf{v} - \langle \mathbf{f},\mathbf{v}\rangle \right\},$$

where $\langle\cdot,\cdot\rangle$ denotes the canonical pairing between the $H^{-1}(\Omega)$ and $H_0^1(\Omega)$ is single-valued, linear, and completely continuous.

Proof. Both existence and uniqueness follow from Proposition 4.11. It is an easy exercise to verify the linearity of $\mathbf{f} \to \mathbf{u}(\mathbf{f})$. Furthermore, comparing objective functionals at $\mathbf{u}(\mathbf{f})$ and at $\mathbf{0}$, and using Poincaré inequality [Bre83, Corollaire IX.19] we get the inequality $\|\mathbf{u}(\mathbf{f})\|^2_{H_0^1(\Omega)} \leq C\langle \mathbf{f},\mathbf{u}(\mathbf{f})\rangle$, for some C independent of $\mathbf{f}$. Of course, it implies complete continuity at zero, which owing to the linearity is equivalent to complete continuity at every point. □