Pontryagin Approximations for Optimal Design

JESPER CARLSSON

Licentiate Thesis in Numerical Analysis
Stockholm, Sweden 2006

Nada är en gemensam institution mellan
Kungliga Tekniska högskolan och Stockholms universitet.

Pontryagin Approximations for Optimal Design

JESPER CARLSSON

Avhandling som med tillstånd av Kungl Tekniska Högskolan
framlägges till offentlig granskning för avläggande av teknologie licentiatexamen
fredagen den 8 september 2006 kl 10.00
i sal D41, Huvudbyggnaden, Lindstedtsvägen 3,
Kungliga Tekniska högskolan, Stockholm.

ISBN 91-7178-417-9
TRITA-CSC-A 2006:11
ISSN-1653-5723
ISRN-KTH/CSC/A--06/11--SE

Akademisk avhandling som med tillstånd av Kungl Tekniska högskolan framlägges till offentlig granskning för avläggande av teknologie licentiatexamen fredagen den 8 september kl 10.00 i sal D41, Huvudbyggnaden, Lindstedtsvägen 3, Stockholm.

ISBN 91-7178-417-9
TRITA-CSC-A 2006:11
ISSN-1653-5723
ISRN-KTH/CSC/A--06/11--SE

Abstract

This thesis concerns the approximation of optimally controlled partial differential equations for applications in optimal design and reconstruction. Such optimal control problems are often ill-posed and need to be regularized to obtain good approximations. We here use the theory of the corresponding Hamilton-Jacobi-Bellman equations to construct regularizations and derive error estimates for optimal design problems. The constructed Pontryagin method is a simple and general method where the first, analytical, step is to regularize the Hamiltonian. Next its stationary Hamiltonian system, a nonlinear partial differential equation, is computed efficiently with the Newton method using a sparse Jacobian. An error estimate for the difference between exact and approximate objective functions is derived, depending only on the difference of the Hamiltonian and its finite dimensional regularization along the solution path and its L^2 projection, i.e. not on the difference of the exact and approximate solutions to the Hamiltonian systems. In the thesis we present solutions to applications such as optimal design and reconstruction of conducting materials and elastic structures.

Preface

This thesis consists of an introduction and two papers.

Paper 1: Carlsson J, Szepessy A and Sandberg M. Symplectic Pontryagin Approximations for Optimal Design, *preprint*, 2006.

The author of this thesis contributed to the ideas presented, performed the numerical computations and wrote sections 3, 4, and parts of section 2, of the manuscript.

Paper 2: Carlsson J. Pontryagin Approximation for Optimal Design of Elastic Structures, *preprint*, 2006.

Acknowledgments

I wish to thank my supervisor, professor Anders Szepessy, for his support and guidance throughout this work. I am also grateful to Mattias Sandberg and Erik von Schwerin for the fruitful discussions and cooperation we have had. Finally, I also wish to thank all my colleagues at NADA.

Financial support from VR is gratefully acknowledged.

Contents

Chapter 1

The Optimal Design Problem

Optimal design has with the increase of computational capacity and commercial software for solving partial differential equations become an important industrial field, with applications in virtually all fields of science. Two important applications are optimal design of material structures, and inverse optimal reconstruction of physical properties from experimental data, see e.g. [3] and [4], respectively.

Mathematically, optimal design can be described as the particular inverse problem of controlling one or more a partial differential equations to meet some design criteria in an optimal way. For example, consider the general problem to find a bounded open set $D \subset \Omega \subset \mathbb{R}^d$ such that

$$\inf_{D \in \mathcal{D}_{ad}} \left\{ \int_D F(u) \, \mathrm{d}x \;\middle|\; G(u) = 0 \text{ in } D \right\}, \tag{1.1}$$

where the design criteria is described by the functional $F : \mathbb{R}^n \to \mathbb{R}$, the state variable $u : \Omega \to \mathbb{R}^n$ satisfies the partial differential equation $G(u) = 0$ in D, and $\mathcal{D}_{ad}$ describes a set of admissible designs. Typically, the partial differential operator G here describes a physical state, while the design criteria consists of some energy to minimize or, for a reconstruction problem, an error functional relating the solution u to measurements. Note, that we here only deal with time independent problems governed by elliptic or at least coercive partial differential equations.

The above problem (1.1) is usually referred to as an optimal shape problem [15] and is in general ill-posed in the sense that small perturbations of data lead to large changes in the solution [9, 17]. Also, for a too large set of admissible designs $\mathcal{D}_{ad}$, the infimum in (1.1) may not even be attained.

An alternative way to write the optimal shape problem (1.1) is as a parameter design problem

$$\inf_{\chi \in \chi_{ad}} \left\{ \int_\Omega \chi F(u) \, \mathrm{d}x \;\middle|\; G\chi(u) = 0 \text{ in } \Omega \right\}, \tag{1.2}$$

where the domain Ω is fixed and the infimum is taken is over all characteristic functions $\chi : \Omega \to \{0, 1\}$ in the admissible set χ_{ad}. The partial differential operator G_χ here indicates that the parameter χ is included as a coefficient in the partial differential equation.

Observe, that (1.1) and (1.2) can also represent designing a boundary between two separate regions, by adding a design criteria and a partial differential equation for the domain $\Omega \setminus D$ in (1.1), or by replacing χ with $\alpha\chi + \beta(1 - \chi)$ in (1.2).

Example 1.1 (Optimal design in conductivity)**.** *Consider the problem of minimizing the power loss in an electric conductor, by placing a given amount C of conducting material in*

a given domain $\Omega \subset \mathbb{R}^d$, for a given surface current $q : \partial\Omega \to \mathbb{R}$. In the shape optimization setting this can be formulated as finding the conducting domain $D \subset \Omega$ such that

$$\inf_D \left\{ \int_D |\nabla\varphi|^2 \ \mathrm{d}x \ \middle| \ -\mathrm{div}(\nabla\varphi)\big|_D = 0, \ \frac{\partial\varphi}{\partial n}\big|_{\partial D\setminus\Gamma} = 0, \ \frac{\partial\varphi}{\partial n}\big|_\Gamma = q, \ \int_D \ \mathrm{d}x = C \right\}, \quad (1.3)$$

where $\partial/\partial n$ denotes the normal derivative on the boundary, $\Gamma \equiv \{x \in \partial\Omega : q(x) \neq 0\}$ and $\varphi \in V \equiv \{v \in H^1(\Omega) : \int_\Omega v \ \mathrm{d}x = 0\}$ is the electric potential. Note that the boundary ∂D must contain Γ.

A corresponding parameter design problem can be formulated as to find the characteristic conductivity function $\sigma : \Omega \to \{0, 1\}$ such that

$$\inf_\sigma \left\{ \int_\Omega \sigma|\nabla\varphi|^2 \ \mathrm{d}x \ \middle| \ -\mathrm{div}(\sigma\nabla\varphi)\big|_\Omega = 0, \ \sigma\frac{\partial\varphi}{\partial n}\big|_{\partial\Omega} = q, \ \int_\Omega \sigma \ \mathrm{d}x = C \right\}. \quad (1.4)$$

This parameter design problem is studied in detail in [7].

1.1 Existence of Solutions

Without any restrictions on the class of admissible designs, optimal design problems often do not admit any solutions. A simple example is the problem to find the set $D \subset \Omega \in \mathbb{R}^2$ that minimizes $1/l(D)$, where $l(D)$ is the length of the boundary ∂D. This unconstrained minimization problem clearly has no minimizer although the minimum tends to zero, and to attain a minimizer we must add extra constraints on for example the shape of the domain D, or the boundary ∂D.

To understand why the set of admissible designs is so important we review some conditions on the existence of minimizers, see [15]: To assure existence of a solution D with a corresponding state variable u to the minimization problem (1.1), a necessary condition is that there exists a minimizing sequence D_m to (1.1) such that $\bar{D}_m \to \bar{D}$, in the Hausdorff sense. This does not imply that the corresponding characteristic functions $\chi_{D_m} : \Omega \to L^\infty(\Omega)$ converges pointwise or even weakly * to a characteristic function χ_D (see Definition 1.1 for weak * convergence). However, there always exists a minimizing sequence such that the characteristic functions χ_{D_m} converges in the weak * sense to a limit *not* belonging to the class of characteristic functions. For the problem (1.1) this means that even if the state variables u_m, corresponding to the minimizing sequence of shapes D_m, satisfies the constraint $G(u_m) = 0$, the limit u may not be a solution to the original partial differential constraint $G(u) = 0$.

Definition 1.1. *By weak * convergence of $\chi_m : \Omega \to L^\infty(\Omega)$ to $\chi : \Omega \to L^\infty(\Omega)$ we mean that*

$$\lim_{m\to\infty} \int_\Omega \chi_m(x)\phi(x) \ \mathrm{d}x = \int_\Omega \chi(x)\phi(x) \ \mathrm{d}x,$$

*for all test functions $\phi : \Omega \to L^1(\Omega)$. The notation 'weak *' is here used since $L^1(\Omega)$ is not the dual space of $L^\infty(\Omega)$.*

To find a minimizing sequence of characteristic functions that converges to a characteristic function, we can either alter the original optimal design problem by adding penalty terms in the design criterion, or change the set of admissible designs, for example by adding conditions on the smoothness of the boundary, e.g. only allowing Lipschitz boundaries. One problem is that this restriction usually gives a minimum different from the infimum of the original problem, i.e. the problem has been altered significantly. Another approach is to

extend the admissible set χ_{ad} in (1.2) to include not only characteristic functions, e.g. by introducing composites of laminated materials as in the homogenization method [1]. Such composites describes periodic material micro-structures and can for certain laminations give a minimum that coincides with the true infimum. It is worth to mention that even if a solution exists, optimal design problems may be ill-posed in the sense that small perturbations of data lead to large changes in the solution.

In Chapter 2, a different approach more connected with optimal control and calculus of variations, is used for finding a regularization. For some problems we can derive sufficient conditions for a minimizer [6, 7].

Remark 1.1. *For the particular example of minimizing energy in Example 1.1, there exists a unique minimizer without any restriction on the shape [15]. On the other hand, changing the 'inf' for a 'sup' needs regularization to admit a solution. This particular maximization problem has is addressed in [7], and can be regularized by convexification or homogenization [1, 11, 12, 13, 14].*

1.2 Solution Methods

Roughly, the computational methods solving for optimal design problems can be divided into two classes: Methods with optimality conditions derived from (1.1), and methods based on approximation of the characteristic function χ in (1.2).

In the first class we find the classical method of shape derivatives, which derives the optimal variation of the boundary. Topological derivatives, or the bubble method, is a similar method that derives optimality conditions for the creation of holes in the domain, i.e not only moves the boundary but also changes the topology. The shape optimization methods commonly use a finite element or finite difference discretization of the domain D to solve the partial differential equation $G(u) = 0$ and update both D and the discretization from the optimality conditions. Alternatively, a fixed mesh and a mapping onto the domain D can be used. Another method that uses the shape derivative, the topological derivative, or a combination of both is the level-set method. A level-set function is then used to indicate the boundary, and boundary movement and creation of holes is done by solving a transport equation for the level-set function on the whole domain Ω.

The second class of computational methods is based on the formulation (1.2) and relaxes the class of admissible designs to allow a global minimum, either by smooth approximation of χ, or as in the homogenization method, by a special class of admissible controls χ_{ad} based on periodic micro-structures. Since these methods uses a discretization of the whole region Ω it is here often necessary to use a weak material to mimic void, i.e. $\chi > 0$. Also, to produce sharp boundaries between, in this case, the weak and the solid phase, some penalization procedure is often added. This may seem counter productive, but the hope is to first reach a global minimum to the relaxed problem, followed by a penalization which removes existence of a global minimum but forces the solution to a nearby local minimum.

In this presentation, we only deal with the continuous problem, and do not discuss any of the many optimization methods dealing with the discretized versions of (1.1) and (1.2). An introduction to discrete methods concerning optimal design of material structures can be found in [3].

Shape and Topological Derivative

Consider the problem (1.1) and define the objective functional

$$J(D) \equiv \int_D F(u) \, \mathrm{d}x,$$

where $u : D \to V$ is the solution, belonging to some Hilbert space V, to the partial differential equation $G(u) = 0$ in (1.1)

For a small perturbation $\theta : \mathbb{R}^d \to \mathbb{R}^d$ of the domain $D \subset \mathbb{R}^d$ into $D+\theta = \{x+\theta(x), x \in D\}$ the shape derivative in the direction θ can be defined as

$$\delta J(D;\theta) = \int_{\partial D} L\big(u(s), \lambda(s)\big) \; \theta(s) \cdot n \;\; \mathrm{d}s, \tag{1.5}$$

where n denotes the outward boundary normal. The functional L is here a certain problem dependent functional which is described for an example below, see Example 1.2. The variable $\lambda : D \to V$ is here the solution to a corresponding adjoint problem. One way to define the adjoint problem is from the Lagrangian

$$\mathcal{L}(D, u, \lambda) \equiv J(D, u) + \langle \lambda, G(u) \rangle,$$

where $\langle v, w \rangle$ is the duality pairing on V, which reduces to the L^2 inner product if $v, w \in L^2(D)$. The Gâteau derivative with respect to λ gives the original constraint $G(u) = 0$, in the distribution sense, while the Gâteau derivative with respect to u gives the dual problem for λ. The shape derivative (1.5) gives the sensitivity of the value function J with respect to change in the domain, and indicates how to move the boundary $\partial\Omega$, or the individual mesh points in the discretization of D.

Example 1.2. *Consider a simplified version of the conductivity optimization problem* (1.3)*, given in Example 1.1, where the objective functional now is*

$$J(D) = \int_D |\nabla\varphi|^2 \;\; \mathrm{d}x + \eta \int_D \;\; \mathrm{d}x,$$

and the state variable φ solves

$$-\mathrm{div}(\nabla\varphi)\big|_D = 0, \quad \frac{\partial\varphi}{\partial n}\big|_{\partial D\setminus\Gamma} = 0, \quad \frac{\partial\varphi}{\partial n}\big|_{\Gamma} = q.$$

The shape derivative is then given by

$$\delta J(D;\theta) = \int_{\partial D} (\nabla u \cdot \nabla\lambda) \; \theta \cdot n \;\; \mathrm{d}s + \eta \int_{\partial D} \theta \cdot n \;\; \mathrm{d}s$$

where the dual solution is given by $\lambda = \varphi$, see [15].

Unfortunately, the shape derivative does not deal with changes in the topology, e.g. nucleation of holes in the domain. A method which does consider topological changes is the method of topological derivatives, see e.g. [8]. The topological derivative is an extension of the shape derivative, and derives an expression for the change in the value function with respect to the creation of a small hole inside the domain.

Level-Set Methods

The level-set method, conveniently connects the two problems (1.1) and (1.2) by parameterizing the boundary between the phases using a level-set function $\psi : \Omega \times [0,T] \to \mathbb{R}$, given by

$$\begin{cases} \psi(x,\cdot) > 0, & x \in \Omega - D, \\ \psi(x,\cdot) = 0, & x \in \partial D, \\ \psi(x,\cdot) < 0, & x \in D, \end{cases}$$

where the normal n of ∂D is given by $\nabla\psi/|\nabla\psi|$ and the curvature by $\operatorname{div}(\nabla\psi/|\nabla\psi|)$. The time is here an artificial variable used to evolve the shape towards its optimum, by the dynamics of the Hamilton-Jacobi equation

$$\partial_t \psi + V|\nabla\psi| = 0 \text{ in } \Omega \tag{1.6}$$

where $V : \Omega \times [0,T] \to \mathbb{R}$ denotes the normal velocity of ∂D. Here, the normal velocity can be chosen according to the shape or topological derivatives, see [2, 5], and the time T corresponds to the length of the gradient step. In practice, the T is chosen such that the normal and curvature of the level-set function does not become too distorted. From the solution $\psi(\cdot, T)$, a reinitialization where the partial differential equation in (1.2) is solved, gives new initial data $\psi(\cdot, 0)$ for solving (1.6) again. The level-set method requires using a weak phase to mimic void when solving the partial differential equation in (1.2), and extra computational work is introduced from introducing the additional function ψ. Also, a fixed discretization of the whole domain Ω is used for both (1.6) and the partial differential equation constraint in (1.2).

Homogenization

The previous methods all tried to find an optimal domain $D \in \Omega$, which may not exist for certain problems, unless some restriction is put on the shape of the boundary ∂D. The homogenization method, on the other hand, looks for optimal designs in the class of periodic micro-structures. Such structures do not in general form sharp boundaries, but instead share the property that there exists a minimum which coincides in average with the infimum of the original problem, as mentioned in Section 1.1.

To exemplify, we state the problem briefly mentioned in Remark 1.1: Find the conductivity function $\sigma : \Omega \to \{\sigma_-, 1\}$ that maximizes the power loss in an electric conductor, i.e.

$$\sup_{\sigma:\Omega\to\{\sigma_-,1\}} \left\{ \int_\Omega \sigma|\nabla\varphi|^2 \ \mathrm{d}x \ \middle| \ -\operatorname{div}(\sigma\nabla\varphi)\big|_\Omega = 0, \ \sigma\frac{\partial\varphi}{\partial n}\big|_{\partial\Omega} = q, \ \int_\Omega \sigma \ \mathrm{d}x = C \right\}. \tag{1.7}$$

Note that we have here filled the void with a weak phase $\sigma_- > 0$. This maximization problem lacks maximizers, but can be relaxed to allow the existence a maximizer by simply using $\sigma : \Omega \to [\sigma_-, 1]$ instead of $\sigma : \Omega \to \{\sigma_-, 1\}$. A more clever approach is to use the homogenization method for laminated materials. We then look at the problem

$$\max_{\theta,\phi} \left\{ \int_\Omega \sigma^*|\nabla\varphi|^2 \ \mathrm{d}x \ \middle| \ -\operatorname{div}\big(\sigma^*\nabla\varphi\big)\big|_\Omega = 0, \ \sigma^*\frac{\partial\varphi}{\partial n}\big|_{\partial\Omega} = q, \ \int_\Omega \theta \ \mathrm{d}x = C \right\}. \tag{1.8}$$

with $\theta : \Omega \to [0,1]$, $\phi : \Omega \to [0,\pi]$ and the rank-1 laminate tensor

$$\sigma^*(\theta,\phi) = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} \lambda_\theta^+ & 0 \\ 0 & \lambda_\theta^- \end{pmatrix} \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix},$$

with

$$\lambda_\theta^- = \left(\frac{\theta}{\sigma_-} + \frac{1-\theta}{\sigma_+} \right)^{-1}, \quad \lambda_\theta^+ = \theta\sigma_- + (1-\theta)\sigma_+.$$

The tensor σ^* is obtained from rotation and mixing of the two tensor valued controls $\sigma_- I$ and $\sigma_+ I$ in proportions θ and $1-\theta$ and direction ϕ, see Figure 1.1. We have thus enlarged the set of admissible controls by introducing two new parameters θ, ϕ describing a laminated material. The effective conductivities in the principal directions of the material is λ_θ^+ and λ_θ^-, while $(\lambda_\theta^+)^{-1}$ and $(\lambda_\theta^-)^{-1}$ correspond to the total resistances for resistors connected in parallel and in series, respectively. The homogenization method has the advantage that a maximizer (θ, ϕ) is found and that that the value of (1.8) coincides with (1.7). This particular problem uses a rank-1 laminate, but higher rank laminates, sufficient to find minimizers (or maximizers) for many important optimal design problems, can be found [1].

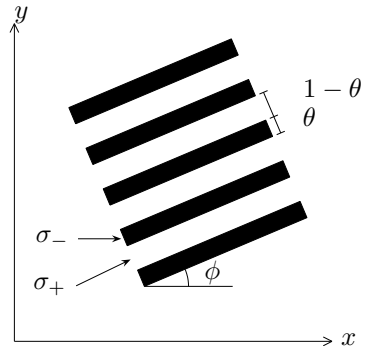

Figure 1.1: The rank-1 laminate used in (1.8)

Chapter 2

Optimal Control and the Pontryagin Method

In the previous chapter we saw that optimal design problems often need to be regularized to obtain good approximations, and that regularization may also be necessary to assure the mere existence of a solution. In this chapter we present a method for optimal design using a regularization derived from the Hamilton-Jacobi-Bellman equations for the corresponding optimal control problem. We first describe the method for control of a system of ordinary differential equations, and then apply the methodology to control partial differential equations.

2.1 Dynamic Programming

Consider an optimal control problem for a controlled ordinary differential equation

$$\inf_{\alpha\in\mathcal{A}}\left\{g\big(X(T)\big)+\int_0^T h\big(X(s),\alpha(s)\big)\,\mathrm{d}s \;\middle|\; X'(t)=f\big(X(t),\alpha(t)\big),\ X(0)=X_0\right\}, \tag{2.1}$$

with given data $g:\mathbb{R}^n\to\mathbb{R}$, $h:\mathbb{R}^n\times B\to\mathbb{R}$, $f:\mathbb{R}^n\times B\to\mathbb{R}^n$, $X_0\in\mathbb{R}^n$, the state variable $X:[0,T]\to\mathbb{R}^n$ and a set of controls $\mathcal{A}=\{\alpha:[0,T]\to B\subset\mathbb{R}^m\}$. Optimal control problems like (2.1) can be solved by dynamic programming or by the Lagrange principle. From the dynamic programming approach, a value function $u:\mathbb{R}^n\times[0,T]\to\mathbb{R}$, defined by

$$u(x,t)\equiv\inf_{X(t)=x,\alpha\in\mathcal{A}}\left\{g\big(X(T)\big)+\int_t^T h\big(X(s),\alpha(s)\big)\,\mathrm{d}s\right\}, \tag{2.2}$$

is the unique viscosity solution (see Definition 2.1 and [10]) of the nonlinear Hamilton-Jacobi-Bellman partial differential equation

$$\begin{aligned}\partial_t u(x,t)+H\big(\partial_x u(x,t),x\big)&=0, \qquad (x,t)\in\mathbb{R}^n\times(0,T),\\ u(x,T)&=g(x), \qquad x\in\mathbb{R}^n,\end{aligned} \tag{2.3}$$

where the Hamiltonian function $H:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ is defined by

$$H(\lambda,x)\equiv\min_{\alpha\in B}\big\{\lambda\cdot f(x,\alpha)+h(x,\alpha)\big\}. \tag{2.4}$$

The value function (2.2) indicates the least cost from starting at a point (x,t) and following an optimal path $X(s)$ and control $\alpha(s)$ for the remaining time $s\in[t,T]$, and the infimum

of (2.1) is given by the solution to (2.3) in the point $(X_0, 0)$. Although we can here find a global minimum, the Hamilton-Jacobi equation can in practice not be solved numerically for high dimensional problems where $n \gg 1$.

Definition 2.1. *(Viscosity solution) A bounded uniformly continuous function u is a viscosity solution to* (2.3)*, if $u(\cdot, T) = g(\cdot)$, and for each $v \in C^\infty(\mathbb{R}^n \times (0,T))$*

- *$\partial_t v(x,t) + H\big(\partial_x v(x,t), x\big) \geq 0$ when $u - v$ has a local maximum in (x,t), and*
- *$\partial_t v(x,t) + H\big(\partial_x v(x,t), x\big) \leq 0$ when $u - v$ has a local minimum in (x,t).*

The viscosity solution u is also unique, see [10].

2.2 The Pontryagin Principle

To derive information on the optimal path $X(t)$ and the corresponding optimal control $\alpha(t)$, we consider the Pontryagin (Minimum) Principle, which states the following necessary condition for an optimal control to (2.1): Assuming that f, g, h are differentiable, then given an optimal path $X(t)$ with an optimal control $\alpha(t)$, there exists a path $\lambda(t)$ such that

$$\begin{aligned} \partial_t X(t) &= f\big(X(t), \alpha(t)\big), \\ X(0) &= X_0, \\ -\partial_t \lambda_i(t) &= \partial_{x_i} f\big(X(t), \alpha(t)\big) \cdot \lambda(t) + \partial_{x_i} h\big(X(t), \alpha(t)\big), \\ \lambda(T) &= \partial_x g\big(X(T)\big), \end{aligned} \tag{2.5}$$

and

$$\lambda(t) \cdot f\big(X(t), \alpha(t)\big) + h\big(X(t), \alpha(t)\big) \leq \lambda(t) \cdot f\big(X(t), a\big) + h\big(X(t), a\big), \quad a \in B,$$

or equivalently

$$\alpha(t) \in \operatorname{argmin}_{a \in B} \big\{\lambda(t) \cdot f\big(X(t), a\big) + h\big(X(t), a\big)\big\}. \tag{2.6}$$

Also, assuming that the Hamiltonian H defined in (2.4) is differentiable, the Pontryagin Principle (2.5) and (2.6), equals the Lagrange principle, i.e. an optimal path $X(t)$ satisfies the Hamiltonian boundary value system

$$\begin{aligned} \partial_t X(t) &= \partial_\lambda H\big(\lambda(t), X(t)\big), \quad X(0) = X_0, \\ -\partial_t \lambda(t) &= \partial_x H\big(\lambda(t), X(t)\big), \quad \lambda(T) = \partial_x g(X(T)), \end{aligned} \tag{2.7}$$

which in fact is the method of characteristics for the Hamilton-Jacobi equation (2.3) provided $\lambda(t) \equiv \partial_x u(X(t), t)$ exists. The Lagrange principle has the advantage that high dimensional problems, $n \gg 1$ can be solved computationally and the drawback is that in practice only local minima can be found computationally. When using (2.7) to solve the minimization problem (2.1) it is assumed that the Hamiltonian is explicitly known and differentiable. In general, Hamiltonians are only Lipschitz continuous for smooth f, g and h.

Many optimal control problems lead to non-smooth optimal controls, which occur by two reasons: the Hamiltonian is in general only Lipschitz continuous, even though f, g, h are smooth, and backward optimal paths $X(t)$ may collide. To be able to use the computational advantage of solving the Hamiltonian boundary value system (2.7) a regularized problem with a $\mathcal{C}^2(\mathbb{R}^n \times \mathbb{R}^n)$ λ-concave approximation H_δ of the Hamiltonian H, is introduced in [16]. This approximation not only gives meaning to (2.7), but is well defined in the sense

that the corresponding approximated value function u_δ is close to the original value function u, see [16]. In [16], error analysis yields the estimate

$$\|u_\delta - u\|_{L^\infty(\mathbb{R}^n \times \mathbb{R}_+)} = \mathcal{O}(\delta), \tag{2.8}$$

for the real and approximate value functions u and u_δ, and with a regularization parameter δ, such that $\|H_\delta - H\|_{L^\infty(\mathbb{R}^n \times \mathbb{R}^n)} = \mathcal{O}(\delta)$. This error estimate is not explicitly dependent on the dimension n, which makes the regularization suitable for optimal control of discretized partial differential equations. Observe that $\|u_\delta - u\|_{L^\infty(\mathbb{R}^n \times \mathbb{R}_+)} \to 0$ does not necessarily imply convergence of the optimal paths $X(t)$ or the controls $\alpha(t)$.

2.3 Pontryagin Approximations for Optimal Design

In [7], the above analysis for optimal control of ordinary differential equations is extended to control of a time dependent partial differential equation

$$\begin{aligned} \partial_t \varphi(x,t) &= f\big(\varphi(x,t), \alpha(x,t)\big), \quad (x,t) \in \Omega \times (0,T) \\ \varphi(x,0) &= \varphi_0, \qquad\qquad\qquad\quad x \in \Omega \end{aligned}$$

where f is a partial differential operator, $\Omega \subset \mathbb{R}^n$, and $\varphi(\cdot,t)$ belongs to some Hilbert space V on Ω. The minimization problem corresponding to (2.1) then becomes

$$\begin{aligned} \inf_{\alpha:\Omega\times[0,T]\to B} \Big\{ g\big(\varphi(\cdot,T)\big) + \int_0^T h\big(\varphi(\cdot,t), \alpha(\cdot,t)\big)\ \mathrm{d}t \ \Big| \\ \partial_t \varphi = f\big(\varphi(\cdot,t), \alpha(\cdot,t)\big),\ \varphi(\cdot,0) = \varphi_0 \Big\}, \end{aligned} \tag{2.9}$$

The Hamiltonian $H : V \times V \to \mathbb{R}$ is defined as

$$H(\lambda, \varphi) \equiv \min_{\alpha:\Omega\to B} \{\langle \lambda, f(\varphi,\alpha)\rangle + h(\varphi,\alpha)\}, \tag{2.10}$$

and the value function $u : V \times [0,T] \to \mathbb{R}$,

$$\begin{aligned} u(\phi,\tau) \equiv \inf_{\alpha:\Omega\times[0,T]\to B} \Big\{ g\big(\varphi(\cdot,T)\big) + \int_\tau^T h\big(\varphi(\cdot,t), \alpha(\cdot,t)\big)\ \mathrm{d}t \ \Big| \\ \partial_t \varphi = f\big(\varphi(\cdot,t), \alpha(\cdot,t)\big),\ \varphi(\cdot,\tau) = \phi \in V \Big\} \end{aligned}$$

solves the Hamilton-Jacobi-Bellman equation

$$\partial_t u(\phi,t) + H\big(\partial_\phi u(\phi,t), \phi\big) = 0, \quad u(\cdot,T) = g. \tag{2.11}$$

Here, ∂ now denotes Gâteaux derivatives (except for ∂_t), and $\langle v, w\rangle$ is the duality pairing on V, which reduces to the $L^2(\Omega)$ inner product if $v, w \in L^2(\Omega)$. The Lagrange principle gives the Hamiltonian system

$$\begin{aligned} \partial_t \varphi &= \partial_\lambda H(\lambda,\varphi), \qquad \varphi(\cdot,0) = \phi \\ \partial_t \lambda &= -\partial_\varphi H(\lambda,\varphi), \quad \lambda(\cdot,T) = \partial_\varphi g\big(\varphi(\cdot,T)\big). \end{aligned} \tag{2.12}$$

In [6, 7], the time-independent version of Equation (2.12) is solved for φ, λ defined on a finite element subspace $\bar{V} \subset V$ and using a $\mathcal{C}^2$ regularized approximate Hamiltonian $\bar{H}_\delta$.

As an example of a time-independent optimal control problem for partial differential equations we review problem (1.4) in Example 1.1, which using Gauss theorem and a prescribed multiplier $\eta \in \mathbb{R}$ corresponding to the volume constraint C, can be written as

$$\inf_{\sigma:\Omega\to\{0,1\}} \left\{ \int_{\partial\Omega} q\varphi \ \ \mathrm{d}s + \eta \int_\Omega \sigma \ \ \mathrm{d}x \ \middle| \ -\mathrm{div}(\sigma\nabla\varphi)\big|_\Omega = 0, \ \sigma\frac{\partial\varphi}{\partial n}\big|_{\partial\Omega} = q \right\}. \tag{2.13}$$

I this case, the Hamiltonian becomes

$$\begin{aligned} H(\lambda,\varphi) &= \min_{\sigma:\Omega\to\{0,1\}} \left\{ \int_\Omega \sigma(\underbrace{\eta - \nabla\varphi\cdot\nabla\lambda}_{v}) \ \ \mathrm{d}x + \int_{\partial\Omega} q(\varphi+\lambda) \ \ \mathrm{d}s \right\} \\ &= \int_\Omega \underbrace{\min_{\sigma\in\{0,1\}} \{\sigma v\}}_{\mathfrak{h}(v)} \ \ \mathrm{d}x + \int_{\partial\Omega} q(\varphi+\lambda) \ \ \mathrm{d}s. \end{aligned}$$

By replacing $\mathfrak{h}$ with a smooth function $\mathfrak{h}_\delta$ (see Figure 2.1) the time-independent version of the Hamiltonian system (2.12) can by symmetry $\varphi = \lambda$ be reduced to the non-linear partial differential equation

$$\begin{cases} -\mathrm{div}\Big(\mathfrak{h}_\delta'(\eta - |\nabla\varphi|^2)\nabla\varphi\Big) = 0, & \text{in } \ \Omega \\ \qquad\qquad \mathfrak{h}_\delta'(\eta - |\nabla\varphi|^2)\dfrac{\partial\varphi}{\partial n} = q, & \text{on } \ \partial\Omega \end{cases}$$

The regularization is here similar to adding a standard Tichonov penalty, c.f. [9], on the L^2-norm of σ in problem (2.13), which combined with allowing intermediate conductivities $\sigma : \Omega \to [0,1]$ gives the problem

$$\inf_{\sigma:\Omega\to[0,1]} \left\{ \int_{\partial\Omega} q\varphi \ \ \mathrm{d}s + \eta \int_\Omega \sigma \ \ \mathrm{d}x + \delta\int_\Omega \sigma^2 \ \ \mathrm{d}x \ \middle| \ -\mathrm{div}(\sigma\nabla\varphi)\big|_\Omega = 0, \ \sigma\frac{\partial\varphi}{\partial n}\big|_{\partial\Omega} = q \right\},$$

with a regularization parameter $\delta > 0$. The Hamiltonian then becomes

$$\begin{aligned} H(\lambda,\varphi) &= \min_{\sigma:\Omega\to[0,1]} \left\{ \int_\Omega \sigma(\underbrace{\eta - \nabla\varphi\cdot\nabla\lambda}_{v} + \delta\sigma) \ \ \mathrm{d}x + \int_{\partial\Omega} q(\varphi+\lambda) \ \ \mathrm{d}s \right\} \\ &= \int_\Omega \sigma^*(v) \ v \ \ \mathrm{d}x + \int_{\partial\Omega} q(\varphi+\lambda) \ \ \mathrm{d}s, \end{aligned}$$

with the control

$$\sigma^*(v) = \begin{cases} 1, & v < -2\delta, \\ \dfrac{-v}{2\delta}, & -2\delta \le v \le 0, \\ 0, & 0 < v, \end{cases}$$

see, Figure 2.1.

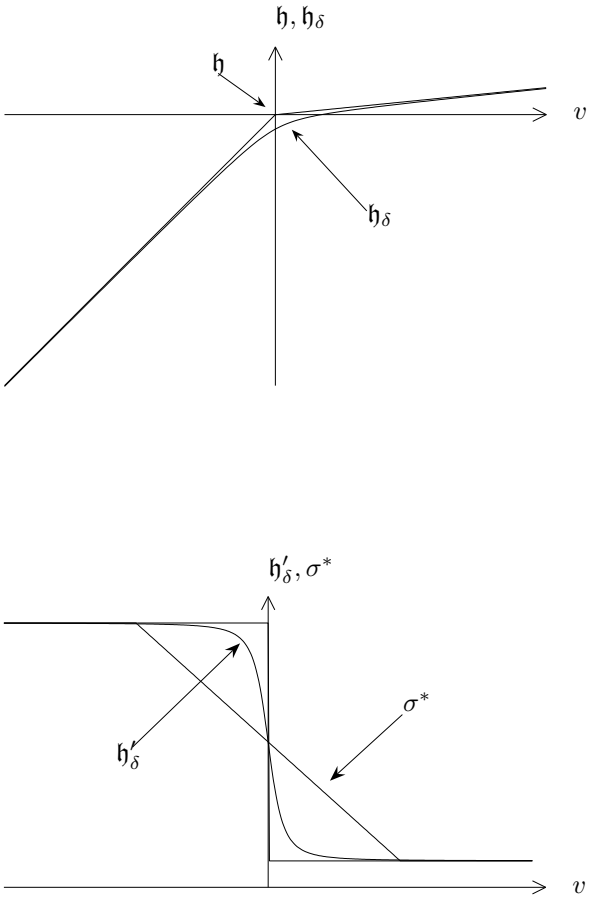

Figure 2.1: Top: The function $\mathfrak{h}$ and its regularization $\mathfrak{h}_\delta$ with respect to v. Bottom: The approximation $\mathfrak{h}'_\delta$ compared to a control σ^* obtained from adding a Tichonov type penalty $\delta \int_\Omega \sigma^2 \, \mathrm{d}x$ to (2.13) with $\sigma : \Omega \to [\sigma_-, \sigma_+]$.

Chapter 3

Summary of Papers

3.1 Paper 1: Pontryagin Approximations for Optimal Design

In this paper we use the Pontryagin method presented in Chapter 2 to solve three different typical optimal design problems; one scalar concave maximization problem in conductivity, one scalar non-concave maximization problem in elasticity, and one inverse reconstruction problem in impedance tomography. An error estimate for the difference in the true and approximated value functions, using only the difference of the true and approximated Hamiltonians along the same paths, is also derived. This estimate gives an error estimate which in practice can be bounded in terms of the regularization parameter and the finite element mesh size, such that the value functions converge even though the optimal paths do not.

3.2 Paper 2: Pontryagin Approximation for Optimal Design of Elastic Structures

In this paper we test the derived Pontryagin method on problems in optimal design of elastic structures, such as distributing a limited amount of material to minimize its compliance, or to detect interior material distributions from surface measurements.

Bibliography

[1] Grégoire Allaire. *Shape optimization by the homogenization method*, volume 146 of *Applied Mathematical Sciences*. Springer-Verlag, Nw York, 2002.

[2] Grégoire Allaire, François Jouve, and Anca-Maria Toader. Structural optimization using sensitivity analysis and a level-set method. *J. Comput. Phys.*, 194(1):363–393, 2004.

[3] M. P. Bendsøe and O. Sigmund. *Topology optimization*. Springer-Verlag, Berlin, 2003. Theory, methods and applications.

[4] Liliana Borcea. Electrical impedance tomography. *Inverse Problems*, 18(6):R99–R136, 2002.

[5] Martin Burger and Stanley J. Osher. A survey on level set methods for inverse problems and optimal design. *European J. Appl. Math.*, 16(2):263–301, 2005.

[6] Jesper Carlsson. Pontryagin approximations for optimal design of elastic structures. *preprint*, 2006.

[7] Jesper Carlsson, Mattias Sandberg, and Anders Szepessy. Symplectic Pontryagin approximations for optimal design. *preprint*, 2006.

[8] Jean Céa, Stéphane Garreau, Philippe Guillaume, and Mohamed Masmoudi. The shape and topological optimizations connection. *Comput. Methods Appl. Mech. Engrg.*, 188(4):713–726, 2000. IV WCCM, Part II (Buenos Aires, 1998).

[9] Heinz W. Engl, Martin Hanke, and Andreas Neubauer. *Regularization of inverse problems*, volume 375 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1996.

[10] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

[11] Jonathan Goodman, Robert V. Kohn, and Luis Reyna. Numerical study of a relaxed variational problem from optimal design. *Comput. Methods Appl. Mech. Engrg.*, 57(1):107–127, 1986.

[12] Robert V. Kohn and Gilbert Strang. Optimal design and relaxation of variational problems. I. *Comm. Pure Appl. Math.*, 39(1):113–137, 1986.

[13] Robert V. Kohn and Gilbert Strang. Optimal design and relaxation of variational problems. II. *Comm. Pure Appl. Math.*, 39(2):139–182, 1986.

[14] Robert V. Kohn and Gilbert Strang. Optimal design and relaxation of variational problems. III. *Comm. Pure Appl. Math.*, 39(3):353–377, 1986.

[15] Olivier Pironneau. *Optimal shape design for elliptic systems*. Springer Series in Computational Physics. Springer-Verlag, New York, 1984.

[16] M. Sandberg and A. Szepessy. Convergence rates of symplectic Pontryagin approximations in optimal control theory. *M2AN*, 40(1), 2006.

[17] Curtis R. Vogel. *Computational methods for inverse problems*, volume 23 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. With a foreword by H. T. Banks.