ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Seminar for Applied Mathematics SAM

Comparison of Approximate Shape Gradients

R. Hiptmair and A. Paganini and S. Sargheini

Research Report No. 2013-30
September 2013

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Funding: This research was supported by ETH Zurich through the CHIRP project CH1-02 11-1

Comparison of Approximate Shape Gradients

R. Hiptmair, A. Paganini and S. Sargheini

September 23, 2013

Abstract

Shape gradients of PDE constrained shape functionals can be stated in two equivalent ways, both of which rely on the solution of two boundary value problems (BVPs). Usually, these two BVPs can only be solved approximatively, for instance, by finite element methods. However, when used with finite element solutions, the equivalence of the two formulations breaks down. By means of a comprehensive convergence analysis, we establish that one expression for the shape gradient offers better accuracy in a finite element setting. The results are confirmed by several numerical experiments.

1 Introduction

Shape calculus studies the "differentiation of shape functionals with respect to the variation of a domain they depend upon". Over the last two decades this notion has been made rigorous, notably by the introduction of the velocity method by Zolesio [8, 22] and the domain perturbation method by Simon [21] and Eppler [9, 10]. Shape calculus has also become important as a key tool in the field of optimization, where it supplies the so-called shape gradient, that is, the first order shape derivative, for use in the framework of descent methods. Yet, methods for shape optimization are outside the scope of this article and we refer the reader to the monographs [1,7,12,15,16,20,22].

Shape optimization entails the approximate numerical computation of shape gradients. This step will be the focus of this article. Of course, a vast diversity of shape functionals is conceivable, leading to vastly different types of shape gradients. Thus, we have to adopt a "case study approach" and restrict our study to a special, albeit important, class of shape functionals.

The shape functionals under scrutiny are least squares output functionals for solutions of scalar second-order elliptic boundary value problems. They belong to the category of PDE constrained shape functionals and have widely been considered in articles on shape optimization [2, 13].

In [1], for instance, formulas have been derived for the associated shape gradients. They are based on solutions u and p of two boundary value problems , called state and adjoint problem. Starting point for our investigations was the insight that the formulas can be stated in two equivalent ways, (i) as expressions involving traces of u and p on the boundary of the domain, and (ii) by means of volume integrals on the domain, see [4, Sect. 6].

The situation resembles that faced for quite a few common output functionals depending on solutions of second-order elliptic PDEs. Examples are total heat flux in heat conduction, lift functionals for potential flow [13], far field functionals [18, 19], and electromagnetic force functionals [17]. All these functionals can be stated as integrals over boundaries or over parts of the domain, and the same value is obtained when inserting exact solutions of state and adjoint BVPs. Both kinds of formulas can also be used in the context of finite element approximation, but when applied to discrete solutions, they fail to give the same answer. More strikingly, the volume integrals often display much faster convergence and provide superior accuracy compared to their boundary based counterparts. An explanation is that the expressions featuring volume integrals enjoy continuity in energy norm, whereas integrals of traces are not well-defined on the natural variational spaces. This makes a crucial difference, because we can benefit from superconvergence, when evaluating continuous functionals for Galerkin solutions [3, Sect. 2].

This made us suspect that similar effects could be observed for the different expressions for shape gradients and their use with finite element solutions. The analysis and numerical experiments of this article largely confirm our expectation that volume based expressions for the shape gradient often offer better accuracy than the use of formulas involving traces on boundaries. This is the message of both the a priori convergence estimates developed in Section 3, see Theorems 1 and 2, and of the numerical tests reported in Section 4.

What compounds the difficulties of gauging the quality of formulas for shape gradients is the fact that they must be viewed as linear functionals on spaces of infinitesimal deformations. Of course, one can switch back to functions via the Riesz representation theorem, but the choice of the underlying inner product is somewhat arbitrary and might bias the outcome. Thus, we decide to study the errors of shape gradients directly in the relevant dual norms.

2 Shape Gradients

Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be an open bounded domain with piecewise smooth boundary $\partial\Omega$, and let $\mathcal{J}(\Omega) \in \mathbb{R}$ be a real-valued quantity of interest associated to it. One is often interested in its shape sensitivity, which quantifies the impact of relatively small perturbations of $\partial\Omega$ on the value $\mathcal{J}(\Omega)$.

For this purpose, we model perturbations of the domain Ω through maps of the form

$$T_{\mathcal{V}} \coloneqq \mathcal{I} + \mathcal{V}, \tag{1}$$

where $\mathcal{I}$ is the identity operator and $\mathcal{V}$ is a vector field in $C^1(\mathbb{R}^d;\mathbb{R}^d)$. It can easily be proven that the map (1) is a diffeomorphism for $\|\mathcal{V}\|_{C^1} < 1$ [1, Lemma 6.13]. Therefore, it is natural to consider $\mathcal{J}(\Omega)$ as the realization of a shape functional, a real map

$$\mathcal{J} : \mathcal{A} \to \mathbb{R}$$

defined on the family of admissible domains

$$\mathcal{A} \coloneqq \left\{ T_{\mathcal{V}}(\Omega)\,; \mathcal{V} \in C^1(\mathbb{R}^d;\mathbb{R}^d)\,, \|\mathcal{V}\|_{C^1} < 1 \right\}.$$

The sensitivity of $\mathcal{J}(\Omega)$ with respect to the perturbation direction $\mathcal{V}$ can be expressed through the *Eulerian derivative* of the shape functional $\mathcal{J}$ in the direction $\mathcal{V}$, that is,

$$d\mathcal{J}(\Omega;\mathcal{V}) \coloneqq \lim_{s \searrow 0} \frac{\mathcal{J}\left(T_{s\cdot\mathcal{V}}(\Omega)\right) - \mathcal{J}(\Omega)}{s}\,. \tag{2}$$

It goes without saying that it is desirable that (2) exists for all possible perturbation directions $\mathcal{V}$. It is therefore natural to define a shape functional $\mathcal{J}$ to be *shape differentiable* at Ω if the mapping

$$d\mathcal{J}(\Omega;\cdot) : C^1(\mathbb{R}^d;\mathbb{R}^d) \to \mathbb{R}, \qquad \mathcal{V} \mapsto d\mathcal{J}(\Omega;\mathcal{V})\,. \tag{3}$$

defined by (2) is linear and continuous on $C^1(\mathbb{R}^d;\mathbb{R}^d)$. In literature, the mapping $d\mathcal{J}(\Omega;\mathcal{V})$ is called *shape gradient* of $\mathcal{J}$ at Ω, as it is the Gâteaux derivative in $0 \in C^1(\mathbb{R}^d;\mathbb{R}^d)$ of the map

$$\mathcal{V} \mapsto \mathcal{J}\left(T_{\mathcal{V}}(\Omega)\right)\,,$$

see [8, Chapter 9, Definition 2.2]. Note that Formula (2) is well-defined for any vector field in the Banach space $C^1(\mathbb{R}^d;\mathbb{R}^d)$, and the shape gradient is an element of its dual space.

Remark 1. *In literature, perturbations as in* (1) *are known as* perturbations of the identity*. From a differential geometry point of view, this approach is less general than the so called* velocity method*, which is, for instance, introduced in [8, Chapter 4]. However, both methods lead to the same formula for the shape gradient, which merely takes into account first order perturbations of the shape functional* $\mathcal{J}$ *[8, Chapter 9, Thm 3.2].*

An interesting property of shape gradients is expressed in the *Hadamard structure theorem* [8, Chapter 9, Thm 3.6]. If $\partial\Omega$ is smooth, $d\mathcal{J}(\Omega,\cdot)$ admits a representative $\mathfrak{g}(\Omega)$ in the space of distributions $\mathcal{D}^k(\partial\Omega)$

$$d\mathcal{J}(\Omega,\mathcal{V}) = \langle \mathfrak{g}(\Omega), \gamma_{\partial\Omega}\mathcal{V}\cdot\mathbf{n}\rangle_{\mathcal{D}^k(\partial\Omega)}\,, \tag{4}$$

where $\gamma_{\partial\Omega}\mathcal{V}\cdot\mathbf{n}$ is the normal component of $\mathcal{V}$ on the boundary $\partial\Omega$. This implies that only normal displacements of the boundary have an impact on the value of $\mathcal{J}(\Omega)$. However, we should take into account that this is no longer true if the boundary $\partial\Omega$ is only piecewise smooth.

We are particularly interested in PDE constrained shape functionals of the form

$$\mathcal{J}(\Omega) = \int_\Omega j(u)\,d\mathbf{x}\,, \tag{5}$$

where $j:\mathbb{R}\to\mathbb{R}$ is a Lipschitz continuous-differentiable function and u is the solution of the *state problem*, a scalar elliptic equation with Neumann or Dirichlet boundary conditions

$$\begin{cases} \mathcal{L}(u) &= f & \text{in } \Omega\,,\\ u &= g \text{ or } \frac{\partial u}{\partial\mathbf{n}} = g & \text{on } \partial\Omega\,. \end{cases} \tag{6}$$

The functions f and g are assumed to belong to $L^2(\mathbb{R}^d)$ ($H^1(\mathbb{R}^d)$ for aNeumann BVP) and $H^2(\mathbb{R}^d)$, respectively, and they are identified with their restriction on Ω and $\partial\Omega$.

Explicit formulas for $d\mathcal{J}(\Omega)$ can easily be derived both for unconstrained and PDE constrained shape functionals, cf. [8, Chapter 9, Section 4.3, and Chapter 10, Section 2.5]. In the case of PDE constrained shape functionals, the formula involves the integration of u, the solution of (6), and p, the solution of an adjoint problem

$$\begin{cases} \mathcal{L}(p) &= j'(u) & \text{in } \Omega\,,\\ p &= 0 \text{ or } \frac{\partial p}{\partial\mathbf{n}} = 0 & \text{on } \partial\Omega\,. \end{cases} \tag{7}$$

As different $\mathcal{L}$ lead to different formulas for the Eulerian derivative, from now on we only consider the model elliptic operator

$$\mathcal{L}(u) = -\Delta u + u\,, \tag{8}$$

which should be regarded as a representative for the class of scalar elliptic differential operators of order two.

As mentioned in the introduction, $d\mathcal{J}(\Omega;\mathcal{V})$ can be formulated as an integral in volume, as well as an integral on the boundary. For example, the formula for the PDE constrained shape functional (5) with elliptic operator (8) and Dirichlet boundary conditions $u = g$ on $\partial\Omega$ reads (see the Appendix for the derivation)

$$\begin{aligned} d\mathcal{J}(\Omega,\mathcal{V}) = \int_\Omega \Big(&\nabla u \cdot (\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla p - f\mathcal{V}\cdot\nabla p \\ &+ \operatorname{div}\mathcal{V}(j(u) - \nabla u\cdot\nabla p - up) \\ &+ (j'(u) - p)(\nabla g\cdot\mathcal{V}) - \nabla p\cdot\nabla(\nabla g\cdot\mathcal{V})\Big)\,d\mathbf{x}, \end{aligned} \tag{9}$$

and can be recast as

$$d\mathcal{J}(\Omega,\mathcal{V}) = \int_{\partial\Omega} (\mathcal{V}\cdot\mathbf{n})\left(j(u) + \frac{\partial p}{\partial\mathbf{n}}\frac{\partial(u-g)}{\partial\mathbf{n}}\right) dS\,. \tag{10}$$

The volume integral (9) and the boundary integral (10) are equivalent representations of the shape gradient $d\mathcal{J}(\Omega;\mathcal{V})$. However, the bulk of literature mainly considers (10) and does not pay attention to (9), probably because the former better matches the theoretical result (4). Only recently it has been realized that the volume representation (9) might be better suited for computations, see [4] and [8, Chapter 10, Remark 2.3].

Remark 2. *In the case of Neumann boundary conditions* $\frac{\partial p}{\partial\mathbf{n}} = 0$ *on smooth domains, the counterparts of Formulas* (9) *and* (10) *read*

$$\begin{aligned} d\mathcal{J}(\Omega,\mathcal{V}) = \int_\Omega \Big(&(\nabla f\cdot\mathcal{V})p + \nabla u\cdot(\mathbf{D}\mathcal{V} + \mathbf{D}\mathcal{V}^T)\nabla p \\ &+ \operatorname{div}\mathcal{V}(fp + j(u) - \nabla u\cdot\nabla p - up)\Big)\,d\mathbf{x} \\ &+ \int_{\partial\Omega} (\nabla g\cdot\mathcal{V})p + gp\operatorname{div}_\Gamma\mathcal{V}\,dS\,, \end{aligned} \tag{11}$$

where $\operatorname{div}_\Gamma$ *denotes the tangential divergence on* $\partial\Omega$*, and*

$$d\mathcal{J}(\Omega,\mathcal{V}) = \int_{\partial\Omega} \mathcal{V}\cdot\mathbf{n}\left(j(u) - \nabla u\cdot\nabla p - up + fp + \frac{\partial gp}{\partial\mathbf{n}} + \mathrm{K}gp\right) dS\,, \tag{12}$$

where K *is the mean curvature of* $\partial\Omega$*.*

3 Approximation of Shape Gradients

In this section we investigate the approximation of the shape gradient $d\mathcal{J}$. For the sake of readability, we perform the analysis for the elliptic operator (8) in 2D with Dirichlet boundary conditions only. The results can easily be extended to general elliptic operators in divergence form, in general dimensions (with a correction on the convergence rates), and both with Dirichlet and Neumann boundary conditions.

To better stress the dependence of $d\mathcal{J}$ on the solution of the state and adjoint problem u and p, as well as to distinguish between Formula (9) and (10), we introduce the notation

$$\begin{aligned} d\mathcal{J}(\Omega,u,p;\mathcal{V})^{\mathrm{Vol}} := \int_\Omega \Big(&\nabla u\cdot(\mathbf{D}\mathcal{V}+\mathbf{D}\mathcal{V}^T)\nabla p - f\mathcal{V}\cdot\nabla p \\ &+ \operatorname{div}\mathcal{V}(j(u)-\nabla u\cdot\nabla p - up) \\ &+ (j'(u)-p)(\nabla g\cdot\mathcal{V}) - \nabla p\cdot\nabla(\nabla g\cdot\mathcal{V})\Big)\, d\mathbf{x}\,, \end{aligned} \tag{13}$$

$$d\mathcal{J}(\Omega,u,p;\mathcal{V})^{\mathrm{Bdry}} := \int_{\partial\Omega} \mathcal{V}\cdot\mathbf{n}\left(j(u)+\frac{\partial p}{\partial\mathbf{n}}\frac{\partial(u-g)}{\partial\mathbf{n}}\right) dS\,. \tag{14}$$

Note that, provided u and p are exact solutions of (6) and (7),

$$d\mathcal{J}(\Omega,\mathcal{V}) = d\mathcal{J}(\Omega,u,p;\mathcal{V})^{\mathrm{Vol}} = d\mathcal{J}(\Omega,u,p;\mathcal{V})^{\mathrm{Bdry}}\,. \tag{15}$$

The operator $d\mathcal{J}(\Omega,\cdot)$ can be approximated by replacing the functions u and p with Ritz–Galerkin Lagrangian finite elements solutions of (6) and (7) respectively. We consider approximations based on discretization with finite elements, as this approach is very popular in shape optimization due to its flexibility for engineering applications. Yet, approximations based on boundary element methods are also possible and they are thoroughly investigated [11, 14, 23].

Equality (15) certainly breaks down when the functions u and p are approximated with finite elements. Thus, a natural question is which formula of (13) and (14) is better suited for an approximation of $d\mathcal{J}(\Omega,\cdot)$ in the sense of linear operator. The following theorem shows that Formula (13) achieves the superconvergence offered by Galerkin approximation. Throughout we tacitly assume that shape-regular and quasi-uniform families of meshes are employed [6, Def. (4.4.13)].

Theorem 1. *Let u_h and p_h be Ritz–Galerkin linear Lagrangian finite elements approximations of the solutions u and p of* (6) *and* (7)*. Furthermore,* *assume that the boundary value problem* (6) *is at least* H^2*-regular [5, Chapter II, Definition 7.1], that the source function* f *is in* $H^1(\Omega)$ *and that the boundary data* g *is a restriction of a function in* $H^2(\mathbb{R}^N)$*. Then*[1]

$$|d\mathcal{J}(\Omega, \mathcal{V}) - d\mathcal{J}(\Omega, u_h, p_h; \mathcal{V})^{\mathrm{Vol}}| \leq C\|\mathcal{V}\|_{W^{1,\infty}}\mathcal{O}(h^2),$$

where h *is the meshwidth.*

Proof. From the equality $d\mathcal{J}(\Omega, \mathcal{V}) = d\mathcal{J}(\Omega, u, p; \mathcal{V})^{\mathrm{Vol}}$ and the linearity of $d\mathcal{J}(\Omega, \mathcal{V})$ in $\mathcal{V}$, we immediately get

$$\begin{aligned}
&|d\mathcal{J}(\Omega, \mathcal{V}) - d\mathcal{J}(\Omega, u_h, p_h; \mathcal{V})^{\mathrm{Vol}}| \\
&\quad \leq \|\mathcal{V}\|_{W^{1,\infty}} \Bigg(|\int_\Omega (-\nabla g \cdot \mathbf{1})(p - p_h)\, d\mathbf{x}| \\
&\qquad + |\int_\Omega j(u) - j(u_h) + (j'(u) - j'(u_h))\, \nabla g \cdot \mathbf{1}\, d\mathbf{x}| \\
&\qquad + |\int_\Omega \nabla u \cdot \nabla p + up - \nabla u_h \cdot \nabla p_h - u_h p_h\, d\mathbf{x}| \\
&\qquad + |\int_\Omega \nabla(p - p_h) \cdot (\nabla(\nabla g \cdot \mathbf{1}) + (\nabla g \cdot \mathbf{1} - f)\mathbf{1})\, d\mathbf{x}| \\
&\qquad + 2|\int_\Omega \nabla u \cdot \mathbb{1} \nabla p - \nabla u_h \cdot \mathbb{1} \nabla p_h\, d\mathbf{x}| \Bigg),
\end{aligned} \tag{16}$$

where $\mathbb{1}$ is a matrix and $\mathbf{1}$ is a vector with entries equal 1. The proof boils down to bounding each integral in the previous inequality and applying standard FEM convergence and interpolation estimates.

With the Cauchy–Schwarz inequality we get

$$|\int_\Omega (-\nabla g \cdot \mathbf{1})(p - p_h)\, d\mathbf{x}| \leq \|g\|_{H^1(\Omega)} \|p - p_h\|_{L^2(\Omega)}\,.$$

Thanks to the Lipschitz continuity of the function j and, again, by the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned}
&|\int_\Omega j(u) - j(u_h) + (j'(u) - j'(u_h))\, \nabla g \cdot \mathbf{1}\, d\mathbf{x}| \\
&\qquad \leq C\|j\|_{W^{1,\infty}_{\mathbb{R}}} \|g\|_{H^1(\Omega)} \|u - u_h\|_{L^2(\Omega)} + \mathcal{O}(\|u - u_h\|^2_{L^2(\Omega)})\,.
\end{aligned}$$

[1] We write C for generic constants, whose value may differ between different occurrences. They may depend only on Ω, shape-regularity and quasi-uniformity.

The third integral in (16) can conveniently be split into

$$
\begin{aligned}
&\int_\Omega \nabla u \cdot \nabla p + up - \nabla u_h \cdot \nabla p_h - u_h p_h \, d\mathbf{x} \\
&= \int_\Omega \big(\nabla p \cdot \nabla(u - u_h) + \nabla u_h \cdot \nabla(p - p_h) + p(u - u_h) + u_h(p - p_h)\big) \, d\mathbf{x}\,, \\
&= \int_\Omega \nabla p \cdot \nabla(u - u_h) + p(u - u_h) \, d\mathbf{x} + \int_\Omega \nabla u \cdot \nabla(p - p_h) + u(p - p_h) \, d\mathbf{x} \\
&\quad + \int_\Omega \nabla(p - p_h) \cdot \nabla(u_h - u) + (p - p_h)(u_h - u) \, d\mathbf{x}\,. \qquad (17)
\end{aligned}
$$

Exploiting Galerkin orthogonality of $u - u_h$ and $p - p_h$ to the finite dimensional trial space V_h in the first two integrals in (17), and applying the Cauchy–Scharz inequality on the third one, we obtain the following bound

$$
\begin{aligned}
&\int_\Omega \nabla u \cdot \nabla p + up - \nabla u_h \cdot \nabla p_h - u_h p_h \, d\mathbf{x} \\
&\leq \inf_{w_h \in V_h} \|p - w_h\|_{H^1(\Omega)} \|u - u_h\|_{H^1(\Omega)} + \inf_{w_h \in V_h} \|u - w_h\|_{H^1(\Omega)} \|p - p_h\|_{H^1(\Omega)} \\
&\quad + \|u - u_h\|_{H^1(\Omega)} \|p - p_h\|_{H^1(\Omega)}\,.
\end{aligned}
$$

The fourth and the fifth integral in (16) can be bounded with standard duality techniques. We give the details for the fourth integral in (16) only, since the procedure for the remaining one is similar. We introduce the function w as solution to the adjoint PDE

$$
\begin{cases}
-\Delta w + w = -\operatorname{div}\left(\nabla(\nabla g \cdot \mathbf{1}) + (\nabla g \cdot \mathbf{1} - f)\mathbf{1}\right) & \text{in } \Omega\,, \\
\frac{\partial w}{\partial \mathbf{n}} = \left(\nabla(\nabla g \cdot \mathbf{1}) + (\nabla g \cdot \mathbf{1} - f)\mathbf{1}\right) \cdot \mathbf{n} & \text{on } \partial\Omega\,.
\end{cases} \qquad (18)
$$

Then, by exploiting the Galerkin orthogonality of $p - p_h$ to the finite dimensional trial space V_h

$$
\begin{aligned}
&\left|\int_\Omega \nabla(p - p_h) \cdot \left(\nabla(\nabla g \cdot \mathbf{1}) + (\nabla g \cdot \mathbf{1} - f)\mathbf{1}\right) \, d\mathbf{x}\right| \\
&\qquad = \left|\int_\Omega \nabla w \cdot \nabla(p - p_h) + w(p - p_h) \, d\mathbf{x}\right|\,, \\
&\qquad = \inf_{w_h \in V_h} \left|\int_\Omega \nabla(w - w_h) \cdot \nabla(p - p_h) + (w - w_h)(p - p_h) \, d\mathbf{x}\right|\,, \\
&\qquad \leq \|p - p_h\|_{H^1(\Omega)} \inf_{w_h \in V_h} \|w - w_h\|_{H^1(\Omega)}\,.
\end{aligned}
$$

Owing to the H^2-regularity of (6), we conclude the proof with standard FEM convergence estimates. □

Remark 3. *The quadratic rate of convergence in Theorem 1 depends on the regularity of the functions u and p. If the assumption on the H^2-regularity of* (6) *is not fulfilled, the rate of convergence deteriorates to fractional values, but the formula is still well-defined, as long as $H^1(\Omega)$-weak solutions exist. On the other hand, if the functions u and p satisfy stronger regularity conditions, the convergence may be improved by increasing the polynomial order of the basis functions of FEM.*

For Formula (14), the following holds:

Theorem 2. *Let u_h and p_h be Ritz–Galerkin linear Lagrangian finite elements approximations of the solutions u and p of* (6) *and* (7)*. In addition to the hypothesis of Theorem 1 let us assume that*

$$\|u\|_{W^2_p(\Omega)} \leq C\|f\|_{L^p(\Omega)}$$

for some $p > d$, where d is the dimension of Ω. Then

$$|d\mathcal{J}(\Omega, \mathcal{V}) - d\mathcal{J}(\Omega, u_h, p_h; \mathcal{V})^{\mathrm{Bdry}}| \leq C\|\mathcal{V}\cdot\mathbf{n}\|_{L^\infty(\Omega)}\mathcal{O}(h),$$

where h is the meshwidth.

Proof. The result follows straightforwardly from the $W^{1,\infty}(\Omega)$ approximation properties of the finite element method

$$\|u - u_h\|_{W^{1,\infty}(\Omega)} \leq Ch\,,$$
$$\|p - p_h\|_{W^{1,\infty}(\Omega)} \leq Ch\,,$$

cf. [6, Corollary 8.1.12]. □

Remark 4. *For $d\mathcal{J}(\Omega, u, p; \mathcal{V})^{\mathrm{Bdry}}$ to be well-defined, the functions u and p must be smoother than merely belonging to $H^1(\Omega)$.*

4 Numerical Experiments

We numerically study the approximation of the shape gradient for the quadratic shape functional

$$\mathcal{J}(\Omega) = \int_\Omega u^2\,d\mathbf{x}$$

under the PDE constraint

$$\begin{cases} -\Delta u + u &= f \quad \text{in } \Omega\,, \\ u &= g \quad \text{on } \partial\Omega\,. \end{cases} \tag{19}$$

It is challenging to numerically investigate convergence rates in the $C^1(\mathbb{R}^d;\mathbb{R}^d)$ dual norm. Therefore, we only consider an operator norm over a finite dimensional space of vector fields in $\mathcal{P}_{3,3}(\mathbb{R}^2)$, whose components are multivariate product polynomials of degree three. Additionally, the $C^1(\mathbb{R}^d;\mathbb{R}^d)$–norm is replaced with the $H^1(\Omega)$–norm, which is more tractable computationally. The convergence studies are performed monitoring the approximate dual norms

$$\mathrm{err}^{\mathrm{Vol}} := \left(\max_{\mathcal{V}\in\mathcal{P}_{3,3}} \frac{1}{\|\mathcal{V}\|^2_{H^1(\Omega)}} |d\mathcal{J}(\Omega,\mathcal{V}) - d\mathcal{J}(\Omega,u_h,p_h;\mathcal{V})^{\mathrm{Vol}}|^2 \right)^{1/2}$$

and

$$\mathrm{err}^{\mathrm{Bdry}} := \left(\max_{\mathcal{V}\in\mathcal{P}_{3,3}} \frac{1}{\|\mathcal{V}\|^2_{H^1(\Omega)}} |d\mathcal{J}(\Omega,\mathcal{V}) - d\mathcal{J}(\Omega,u_h,p_h;\mathcal{V})^{\mathrm{Bdry}}|^2 \right)^{1/2}$$

on different meshes generated through uniform refinement[2].

Although analytical values are in some cases computable, the reference values $d\mathcal{J}(\Omega,\mathcal{V})$ are approximated by evaluating $d\mathcal{J}(\Omega,u_h,p_h;\mathcal{V})^{\mathrm{Vol}}$ on a mesh with an extra level of refinement. This gives us much flexibility in the selection of test cases (the same code can be used for different geometries Ω, source functions f and g, and vector fields $\mathcal{V}$). Agreement with the theoretical predictions of Theorem 1 and a numerical study in the third numerical experiment confirm the viability of this approach.

In the implementation, we opt for linear Lagrangian finite elements on quasi–uniform triangular meshes with nodal basis functions [3]. Integrals in the domain are computed by 7 point quadrature rule in each triangle, while line integrals with 6 point Gauss quadrature on each segment. The boundary of the computational domains is approximated by a polygon, which is generally believed not to affect the convergence of linear finite elements [6, Section 10.2].

The **first numerical experiment** is constructed starting from the solution

$$u(x,y) = \cos(x)\cos(y)$$

and setting f and g accordingly. The computational domain is a circle with radius $\sqrt{\pi}$ (see Figure 1 , left). The predicted quadratic and linear convergence with respect to the meshwidth h for, respectively, Formulas (13) and (14) are evident in Figure 2 (left).

[2]In experiments 1 and 4 we consider domains with curved boundaries. In this case the new mesh is always adjusted to fit the boundary.

[3]The experiments are perfomed in MATLAB and are based on the library LehrFEM developed at the ETHZ.

The **second experiment** is performed on a triangle with corners located at $(-\pi, -\pi)$, $(\pi, -\pi)$, and $(0, \pi)$ (see Figure 1, right). The source functions and the boundary data are chosen as follows:

$$f(x, y) = x^2 - y^2\,, \quad g(x, y) = x + y\,.$$

Again, the rates of convergence predicted in Theorems 1 and 2 are confirmed by the experiment, see Figure 2 (right).

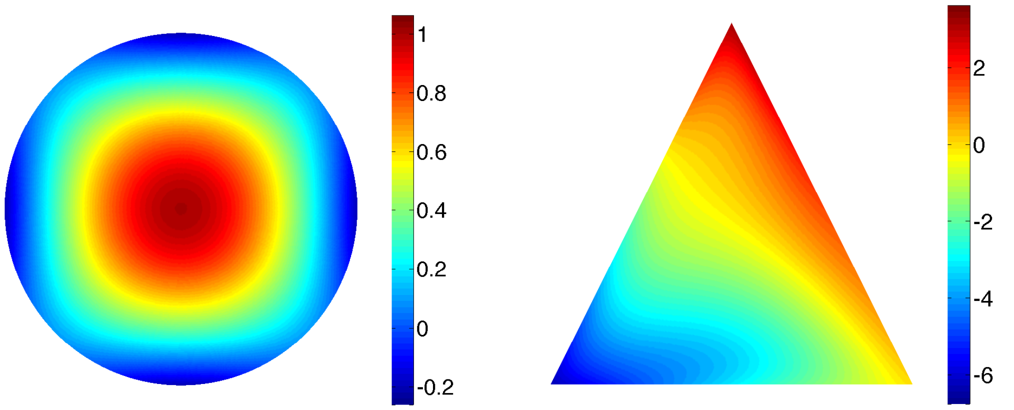

Figure 1: Plot of the state problem solution u in the computational domain Ω for the first (left) and the second (right) numerical experiment.

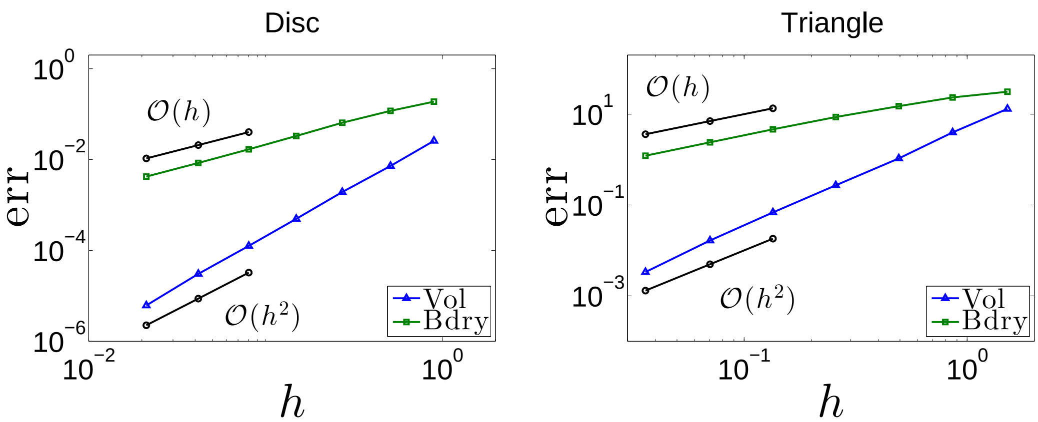

Figure 2: Convergence study for the first (left) and the second (right) numerical experiment. Obviously, Formula (13) is better suited for a finite element approximation of the Eulerian derivative $d\mathcal{J}(\Omega, \mathcal{V})$ than Formula (14).

The **third numerical experiment** is conducted on a domain which does not guarantee H^2-regularity of the state problem (6), see Figure 3 (left). The source and the boundary functions are, respectively, $f(\mathbf{x}) = 1$ and $g(\mathbf{x}) = 0$. As expected, the convergence rates of the formulation in volume deteriorates by a factor of 0.5, because the reentrant corner has an interior angle of amplitude $2\pi \cdot 60/61$ which affect the regularity of the functions u and p. Note also that, due to the poor regularity, the formulation of $d\mathcal{J}(\Omega; \mathcal{V})$ as a boundary integral is barely defined. Moreover, to show that this observation is not due to a poor reference solution, we compute the maximal absolute error err^{Vol} of the reference solution used in the convergence study with respect to a more accurate approximation of $d\mathcal{J}(\Omega, \mathcal{V})$ obtained performing an additional refinement of the mesh. Since this error is $4.8 \cdot 10^{-5}$, the reference solution used in the experiment is accurate enough.

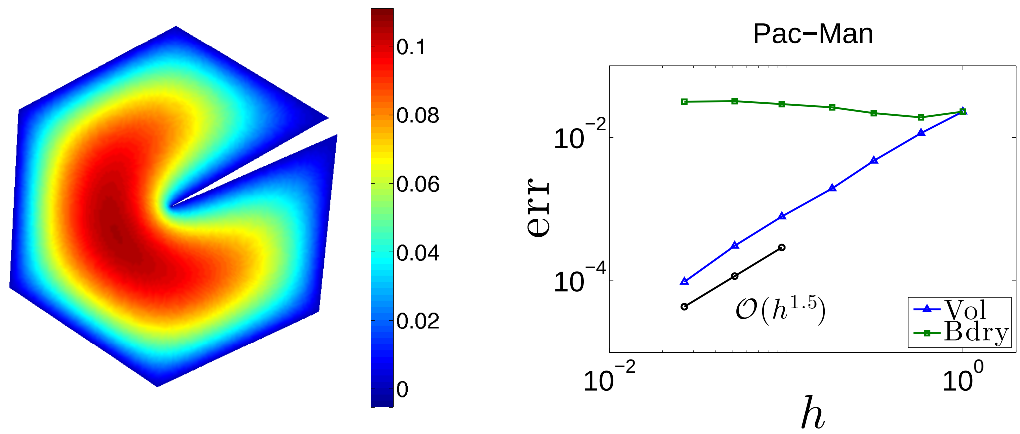

Figure 3: Plot of the state problem solution u in the computational domain Ω (left) for the third numerical experiment, and corresponding convergence study (right). Due to the poor regularity of the functions u and p, the convergence rate of $d\mathcal{J}(\Omega, u_h, p_h; \mathcal{V})^{\text{Vol}}$ deteriorates while $d\mathcal{J}(\Omega, u_h, p_h; \mathcal{V})^{\text{Bdry}}$ does not seem to converge.

In the **fourth numerical experiment**, we investigate the Neumann problem and the accuracy of Formulas (11) and (12), for which we expect results similar to the Dirichlet case. As in the first numerical experiment, we consider the solution

$$u(x, y) = \cos(x)\cos(y)$$

and we choose f and g accordingly. The computational domain is again a disc with radius $\sqrt{\pi}$. Surprisingly, we observe that Formula (12) performs as well as Formula (11), showing quadratic convergence in the meshwidth h too, see Figure 4.

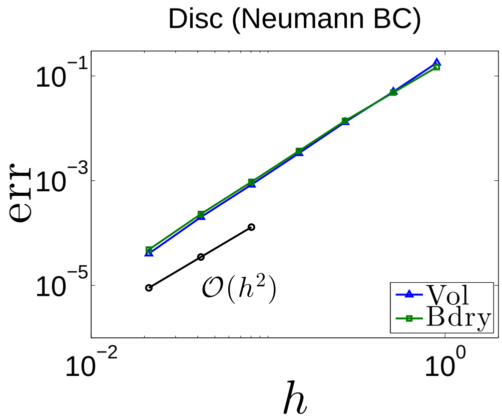

Figure 4: Convergence study for a non-homogeneous Neumann boundary value problem. The quadratic convergence of $d\mathcal{J}(\Omega, u_h, p_h; \mathcal{V})^{\mathrm{Bdry}}$ is unexpected and still defies a theoretical explanation.

A closer look at Formula (12) reveals a cancellation of the normal derivatives of u and p, so that the formula is equivalent to

$$d\mathcal{J}(\Omega, \mathcal{V}) = \int_{\partial\Omega} \mathcal{V} \cdot \mathbf{n} \left(j(u) - \nabla_\Gamma u \nabla_\Gamma p - up + fp + \mathrm{K}gp \right) \, dS \,, \tag{20}$$

where ∇_Γ stands for the tangential derivative. To further investigate we split Formula (20) as follows

$$\begin{aligned} d\mathcal{J}(\Omega, \mathcal{V}) = &\int_{\partial\Omega} \mathcal{V} \cdot \mathbf{n} \left(j(u) - up + fp + \mathrm{K}gp \right) \, dS \\ &- \int_{\partial\Omega} \mathcal{V} \cdot \mathbf{n} \left(\nabla_\Gamma u \nabla_\Gamma p \right) \, dS \,. \end{aligned} \tag{21}$$

An approximation of the first integral in (21) by FEM converges quadratically in h. This can be shown similarly as in the proof of Theorem 1, due to the continuity of the Dirichlet trace operator with respect to $H^1(\Omega)$. On the other hand, the good approximation of the tangential derivative of u and p still defies a theoretical explanation.

Finally, all experiments are repeated considering the operator norm on the subspace of multivariate polynomials of degree two instead of three. The results qualitatively agree with the observations made above, see Figure 5. Thus, the arbitrary choice of restricting the operator norm on the finite dimensional subspace of multivariate polynomial vector fields of degree three is not so severe.

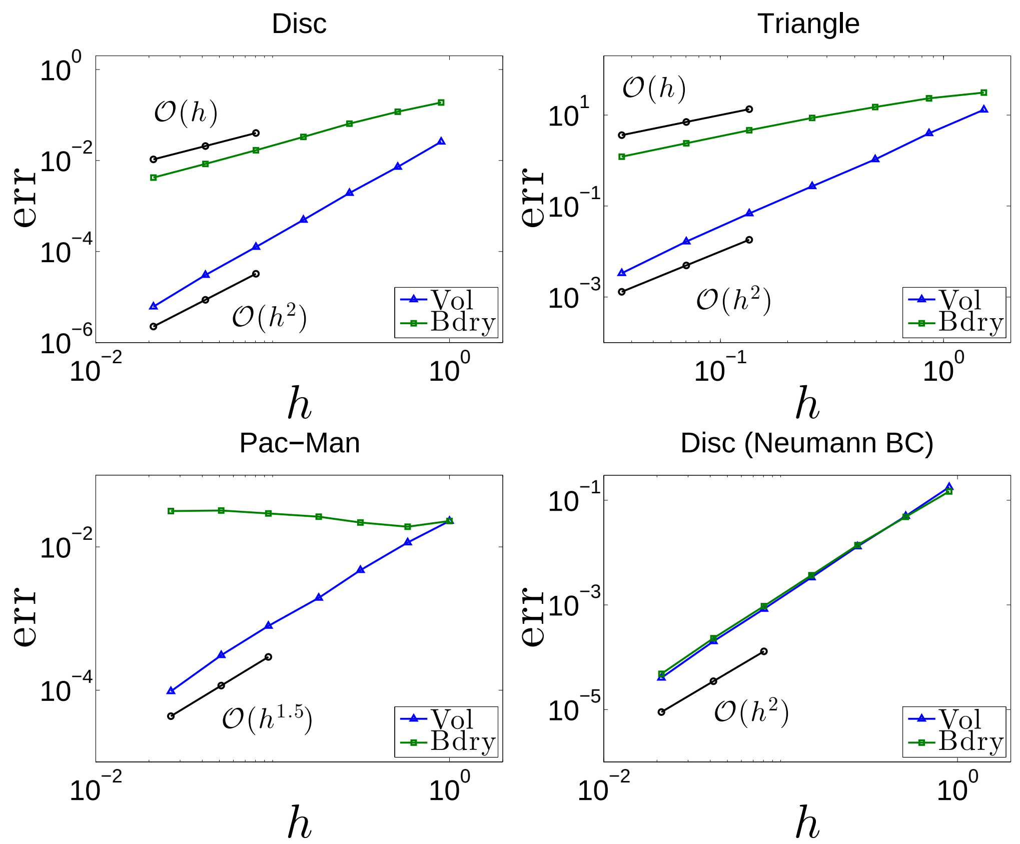

Figure 5: Convergence study for the first (left,up), the second (right, up), the third (left, down), and the fourth (right, down) numerical experiment, when considering the operator norm on the subspace of multivariate polynomials of degree two. The results agree with those obtained with cubic polynomials.

5 Conclusion

The shape gradient of shape differentiable PDE constrained shape functionals is an element of the dual space of $C^1(\mathbb{R}^d;\mathbb{R}^d)$, and it can be expressed either as an integration in volume or as an integration on the boundary. Theorems in Section 3 and numerical experiments in Section 4 confirm that it is advisable to evaluate the shape gradient with an integration in volume, when the underlying approximation method is FEM.

This observation might be of relevance for shape optimization, because, in the words of M. Berggren, "the sensitivity information - directional derivatives of objective functions and constraints - needs to be very accurately computed in order for the optimization algorithms to fully converge" [4]. However, shape optimization techniques usually rely on function representative of the shape gradient on the boundary. If volume based formulas are used, it takes an extension of boundary deformations into the interior of the domain, in order to obtain those. It remains to be seen whether the superiority of volume formulas persists under these conditions.

Appendix

We give a detailed derivation of Formula (9). Let u be the $H^1(\Omega)$–weak solution of the following state problem

$$\left\{ \begin{array}{rcll} -\Delta u + u &=& f & \text{in } \Omega\,, \\ u &=& g & \text{on } \partial\Omega\,. \end{array} \right. \tag{22}$$

It is assumed that the Dirichlet problem (22) is H^2-regular, so that its solution u is at least in $H^2(\Omega)$. We consider the shape functional

$$\mathcal{J}(\Omega) = \int_\Omega j(u)\, d\mathbf{x}\,,$$

and we introduce the Lagrangian

$$\mathscr{L}(\Omega, v, q, \lambda) := \int_\Omega j(v) + (\Delta v - v + f) q \, d\mathbf{x} + \int_{\partial\Omega} \lambda (g - v)\, dS\,, \tag{23}$$

where the functions v, q and λ are in $H^2(\mathbb{R}^d)$. Performing integration by parts, the Lagrangian can be rewritten as

$$\begin{aligned} \mathscr{L}(\Omega, v, q, \lambda) &= \int_\Omega j(v) - \nabla v \cdot \nabla q - v\, q + f\, q \, d\mathbf{x} + \int_{\partial\Omega} \frac{\partial v}{\partial \mathbf{n}} q + \lambda(g - v)\, dS\,, \\ &= \int_\Omega j(v) + (\Delta q - q) v + f\, q\, d\mathbf{x} + \int_{\partial\Omega} \frac{\partial v}{\partial \mathbf{n}} q - \frac{\partial q}{\partial \mathbf{n}} v + \lambda(g - v)\, dS\,, \end{aligned}$$

The saddle point of $\mathscr{L}(\Omega, \cdot, \cdot, \cdot)$ is characterized by

$$\langle \frac{\partial \mathscr{L}(\Omega, v, q, \lambda)}{\partial v}, \phi \rangle = \langle \frac{\partial \mathscr{L}(\Omega, v, q, \lambda)}{\partial q}, \phi \rangle = \langle \frac{\partial \mathscr{L}(\Omega, v, q, \lambda)}{\partial \lambda}, \phi \rangle = 0$$

for all $\phi \in H^2(\mathbb{R}^d)$, which, by density, leads to

$$\left\{ \begin{array}{rcll} -\Delta v + v &=& f & \text{in } \Omega\,, \\ v &=& g & \text{on } \partial\Omega\,, \end{array} \right. \tag{24a}$$

$$\left\{ \begin{array}{rcll} -\Delta q + q &=& j'(v) & \text{in } \Omega\,, \\ q &=& 0 & \text{on } \partial\Omega\,, \end{array} \right. \tag{24b}$$

$$\lambda = -\frac{\partial q}{\partial \mathbf{n}} \quad \text{on } \partial\Omega\,, \tag{24c}$$

weakly in $H^1(\mathbb{R}^d)$. Thus, for Ω fixed,

$$\mathcal{J}(\Omega) = \min_{v\in H^2(\mathbb{R}^d)} \max_{q,\lambda\in H^2(\mathbb{R}^d)} \mathscr{L}(\Omega, v, q, \lambda)\,, \tag{25}$$

because

$$\mathcal{J}(\Omega) = \mathscr{L}(\Omega, u, q, \lambda) \quad \text{for all } q, \lambda \text{ in } H^2(\mathbb{R}^d)\,.$$

The material derivative of a generic function f with respect to the deformation $T_{\mathcal{V}}$ is defined as

$$\dot{f} \coloneqq \lim_{s\searrow 0} \frac{f\circ T_{s\cdot\mathcal{V}} - f}{s}\,.$$

Note that, if f is independent of Ω, $\dot{f} \in H^1(\mathbb{R}^d)$ for $f \in H^2(\mathbb{R}^d)$.

To compute the Eulerian derivative of $\mathcal{J}(\Omega)$, the Correa-Seeger theorem can be applied on the righthand side of (25) [8, Chapter 10, Section 6.3], so that a formula for $d\mathcal{J}(\Omega)$ can be obtained by evaluating the Eulerian derivative of the Lagrangian (23) in its saddle point. For $T_{\mathcal{V}}(\mathbf{x}) \coloneqq \mathbf{x} + \mathcal{V}(\mathbf{x})$, the Eulerian derivative of (23) reads

$$\begin{aligned}
&\lim_{s\searrow 0} \frac{\mathscr{L}(T_{s\cdot\mathcal{V}}(\Omega), v, q, \lambda) - \mathscr{L}(\Omega, v, q, \lambda)}{s} = \\
&= \int_\Omega \big(j'(v)\dot{v} - \nabla\dot{v}\cdot\nabla q - \nabla v\cdot\nabla\dot{q} + \nabla v\cdot(D\mathcal{V} + D\mathcal{V}^T)\nabla q \\
&\qquad - \dot{v}\,q - v\,\dot{q} + \dot{f}q + f\dot{q} + \operatorname{div}(\mathcal{V})\,(j(v) - \nabla v\cdot\nabla q - v\,q + fq)\big)\,d\mathbf{x} \\
&\quad + \int_{\partial\Omega} \dot{\frac{\partial v}{\partial\mathbf{n}}} q + \frac{\partial v}{\partial\mathbf{n}}\dot{q} + \lambda(\dot{g} - \dot{v}) + \dot{\lambda}(g - v) + \operatorname{div}_\Gamma(\mathcal{V})\left(\frac{\partial v}{\partial\mathbf{n}} q + \lambda(g - v)\right) dS \\
&= \int_\Omega j'(v)\dot{v} + \Delta q\,\dot{v} - q\,\dot{v}\,d\mathbf{x} + \int_\Omega \Delta v\,\dot{q} - v\,\dot{q} + f\,\dot{q}\,d\mathbf{x} \\
&\quad + \int_{\partial\Omega} \dot{\frac{\partial v}{\partial\mathbf{n}}} q + \dot{\lambda}(g - v) + \operatorname{div}_\Gamma(\mathcal{V})\left(\frac{\partial v}{\partial\mathbf{n}} q + \lambda(g - v)\right) dS \\
&\quad + \int_\Omega \nabla v\cdot(D\mathcal{V} + D\mathcal{V}^T)\nabla q + \dot{f}q + \operatorname{div}(\mathcal{V})\,(j(v) - \nabla v\cdot\nabla q - v\,q + fq)\,d\mathbf{x} \\
&\quad + \int_{\partial\Omega} \lambda(\dot{g} - \dot{v}) - \frac{\partial q}{\partial\mathbf{n}}\dot{v}\,dS\,.
\end{aligned}$$

So, in the saddle point defined by (24), we have

$$\lim_{s\searrow 0}\frac{\mathscr{L}(T_{s\cdot\mathcal{V}}(\Omega),v,q,\lambda)-\mathscr{L}(\Omega,v,q,\lambda)}{s}=$$

$$=\int_{\Omega}\nabla v\cdot(D\mathcal{V}+D\mathcal{V}^T)\nabla q+\dot{f}q+\operatorname{div}(\mathcal{V})\left(j(v)-\nabla v\cdot\nabla q-v\,q+fq\right)\,d\mathbf{x}$$

$$+\int_{\partial\Omega}-\frac{\partial q}{\partial\mathbf{n}}\dot{g}\,dS$$

$$=\int_{\Omega}\big(\nabla v\cdot(D\mathcal{V}+D\mathcal{V}^T)\nabla q+\dot{f}q+(j'(v)-q)\dot{g}-\nabla q\cdot\nabla\dot{g}$$

$$+\operatorname{div}(\mathcal{V})\left(j(v)-\nabla v\cdot\nabla q-v\,q+fq\right)\big)\,d\mathbf{x}\,,$$

which corrisponds to Formula (9). Formula (10) can be retrieved from Formula (9) with integration by parts on $\partial\Omega$ [22, Section 3.8] and Gauss's theorem.

References

[1] Grégoire Allaire. *Conception optimale de structures*, volume 58 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin, 2007. With the collaboration of Marc Schoenauer (INRIA) in the writing of Chapter 8.

[2] Grégoire Allaire, Frédéric de Gournay, François Jouve, and Anca-Maria Toader. Structural optimization using topological and shape sensitivity via a level set method. *Control Cybernet.*, 34(1):59–80, 2005.

[3] R. Becker and R. Rannacher. An optimal control approach to *a posteriori* error estimation in finite element methods. *Acta Numerica*, 10:1–102, 2001.

[4] Martin Berggren. A unified discrete-continuous sensitivity analysis method for shape optimization. In *Applied and numerical partial differential equations*, volume 15 of *Comput. Methods Appl. Sci.*, pages 25–39. Springer, New York, 2010.

[5] Dietrich Braess. *Finite elements*. Cambridge University Press, Cambridge, third edition, 2007. Theory, fast solvers, and applications in elasticity theory, Translated from the German by Larry L. Schumaker.

[6] Susanne C. Brenner and L. Ridgway Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, third edition, 2008.

[7] Dorin Bucur and Giuseppe Buttazzo. *Variational methods in shape optimization problems*. Progress in Nonlinear Differential Equations and their Applications, 65. Birkhäuser Boston Inc., Boston, MA, 2005.

[8] M. C. Delfour and J.-P. Zolésio. *Shapes and geometries*, volume 22 of *Advances in Design and Control*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 2011. Metrics, analysis, differential calculus, and optimization.

[9] K. Eppler. Second derivatives and sufficient optimality conditions for shape functionals. *Control Cybernet.*, 29(2):485–511, 2000.

[10] Karsten Eppler. Boundary integral representations of second derivatives in shape optimization. *Discuss. Math. Differ. Incl. Control Optim.*, 20(1):63–78, 2000. German-Polish Conference on Optimization—Methods and Applications (Żagań, 1999).

[11] Karsten Eppler and Helmut Harbrecht. Coupling of FEM and BEM in shape optimization. *Numer. Math.*, 104(1):47–68, 2006.

[12] Max D. Gunzburger. *Perspectives in flow control and optimization*, volume 5 of *Advances in Design and Control*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2003.

[13] Helmut Harbrecht. On output functionals of boundary value problems on stochastic domains. *Math. Methods Appl. Sci.*, 33(1):91–102, 2010.

[14] Helmut Harbrecht and Johannes Tausch. On the numerical solution of a shape optimization problem for the heat equation. *SIAM J. Sci. Comput.*, 35(1):A104–A121, 2013.

[15] J. Haslinger and R. A. E. Mäkinen. *Introduction to shape optimization*, volume 7 of *Advances in Design and Control*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2003. Theory, approximation, and computation.

[16] Emmanuel Laporte and Patrick Le Tallec. *Numerical methods in sensitivity analysis and shape optimization*. Modeling and Simulation in Science, Engineering and Technology. Birkhäuser Boston Inc., Boston, MA, 2003. With 1 CD-ROM (Unix, Macintosh).

[17] S. McFee, J.P. Webb, and D.A. Lowther. A tunable volume integration formulation for force calculation in finite-element based computational magnetostatics. *IEEE Trans. Magnetics*, 24(1):439–442, 1988.

[18] P Monk. *Finite Element Methods for Maxwell's Equations.* Clarendon Press, Oxford, UK, 2003.

[19] Peter Monk and Endre Süli. The adaptive computation of far-field patterns by a posteriori error estimation of linear functionals. *SIAM J. Numer. Anal.*, 36(1):251–274, 1999.

[20] O. Pironneau. *Opitmal shape design for elliptic systems.* Springer Series in Computational Physics. Springer, New York, 1984.

[21] J. Simon. Differentiation with respect to the domain in boundary value problems. *Numer. Funct. Anal. Optim.*, 2(7-8):649–687 (1981), 1980.

[22] Jan Sokołowski and Jean-Paul Zolésio. *Introduction to shape optimization*, volume 16 of *Springer Series in Computational Mathematics.* Springer-Verlag, Berlin, 1992. Shape sensitivity analysis.

[23] Rajitha Udawalpola, Eddie Wadbro, and Martin Berggren. Optimization of a variable mouth acoustic horn. *Internat. J. Numer. Methods Engrg.*, 85(5):591–606, 2011.

Recent Research Reports

Nr.	Authors/Title
2013-20	R. Kruse Consistency and Stability of a Milstein-Galerkin Finite Element Scheme for Semilinear SPDE
2013-21	J. Sukys Adaptive load balancing for massively parallel multi-level Monte Carlo solvers
2013-22	R. Andreev and A. Lang Kolmogorov-Chentsov theorem and differentiability of random fields on manifolds
2013-23	P. Grohs and M. Sprecher Projection-based Quasiinterpolation in Manifolds
2013-24	P. Grohs and S. Keiper and G. Kutyniok and M. Schaefer α-Molecules: Curvelets, Shearlets, Ridgelets, and Beyond
2013-25	A. Cohen and A. Chkifa and Ch. Schwab Breaking the curse of dimensionality in sparse polynomial approximation of parametric PDEs
2013-26	A. Lang Isotropic Gaussian random fields on the sphere
2013-27	Ch. Schwab and C. Schillings Sparse Quadrature Approach to Bayesian Inverse Problems
2013-28	V. Kazeev and I. Oseledets The tensor structure of a class of adaptive algebraic wavelet transforms
2013-29	J. Dick and F.Y. Kuo and Q.T. Le Gia and D. Nuyens and Ch. Schwab Higher order QMC Galerkin discretization for parametric operator equations