On second order methods in shape optimization.

L. Afraites, M. Dambrine, K. Eppler, H. Harbrecht, D. Kateb

1 Introduction

Let us consider the following prototype problem. We are interested in solving

$$E(\Omega) = \int_{\Omega} f(u_\Omega, \nabla u_\Omega) \rightarrow \min,$$

where the integrand satisfies $f \in C^3(\mathbb{R} \times \mathbb{R}^d, \mathbb{R})$. The state $u$ solves the boundary value problem

$$-\text{div} (A \nabla u_\Omega) = k \text{ in } \Omega \text{ and } u_\Omega = 0 \text{ on } \partial \Omega$$

where $A \in C^2(\mathbb{R}^d, M_n(\mathbb{R}))$ is such that $L = -\text{div} (A \nabla .)$ is uniformly (in space) elliptic and continuous and $k \in C^{0,\alpha}(\mathbb{R}^d, \mathbb{R})$ is known. We will study the problem for shapes in the class $\Omega \in C^{2,\alpha}$ for $0 < \alpha \leq 1$. Then, by means of a standard calculus, the existence of second order shape derivatives is provided in terms of regular boundary integrals, see e.g., Sokolowski and Zolesio [12].

By investigating second order methods, one aims to verify either second order stability of critical shapes, the sufficient second order condition (SSOC), or one is interested in constructing a Newton method for numerical solution. As an elementary prerequisite, limitation to (scalar) boundary variational approaches, like radial parametrization for starshaped domains or parametrization by the normal field of a reference shape, is useful by at least two reasons. On the one hand, it provides a one-to-one correspondence between the class of shapes and elements from certain function spaces, at least locally around a reference shape. Consequently, useful function space topologies are directly applicable for the study of sufficient second order conditions and their consequences for e.g., the ill-posedness of shape problems. On the other hand, it ensures the invertibility of the shape Hessian in a generic way and simplifies technicalities for realizing the updates.

In Section 2, we explain how to ensure sufficient conditions in case of well-posed problems. It appeared for example in the context of magnetic shaping of liquid metal. Let us further mention that the two-norm discrepancy (the hessian is coercive in the norm which is strictly weaker than the norm of differentiability) is strongly related to the fact that the shape Hessian naturally extends to a continuous bilinear form on that weaker space. We call this space the energy space of the shape Hessian.

---

* LTM, UMR 5129, Université Joseph Fourier et CEA Grenoble, F-38054 Grenoble, France
† LMA, UMR 5142, Université de Pau et des Pays de l’Adour, CNRS, F-64013 Pau, France
‡ TU Dresden, NuMath, Zellescher Weg 12–14, D-01062 Dresden, Germany
§ Universität Bonn, INS, Wegelerstr. 6, D-53115 Bonn, Germany
¶ LMAC, EA 2222, Université de Technologie de Compiègne, F-60200 Compiègne, France
The second case presented in Section 3 comes from inverse problems. The question is to recover an unknown inclusion inside a material from boundary measurement. There is a huge numerical literature devoted to the numerical study of this question in the field of inverse problems; the numerical experiments insist on the ill-posedness of this problem. The main feature consists in the following fact: even if the second derivative of the criterion to be optimized is strictly positive, its Riesz operator can be compact. Such problems are highly ill-posed and some regularization strategies are necessary to get a convenient shape.

2 Stability results.

Sufficient conditions. We explain first the so-called two-norm discrepancy problem.

In [5], J. Descloux studies the electromagnetic shaping of liquid metals by shape optimization methods and addresses the question of stability for the liquid bubble. He is led to the following Taylor-Young formula

\[ E(h) - E(0) = DE(0)[h] + \frac{1}{2}D^2E(0)[h, h] + o(\|h\|^2_{C^{2,\alpha}}). \]

Here \( h \) stands for the normal deformation while \( 0 \) stands for the critical shape. The Banach space \( C^{2,\alpha} \) appears as the space of differentiability. At a critical shape, the Euler equation holds and the term \( DE(0)[h] \) disappears. The variation of the criterion \( E(h) - E(0) \) is positive as soon as the term \( D^2E(0)[h, h] \) is greater than \( o(\|h\|^2_{C^{2,\alpha}}) \). J. Descloux remarks that the former situation can not appear: he proves the existence of two constants \( C_1, C_2 > 0 \) such that

\[ C_1\|h\|_{H^{1/2}}^2 \leq D^2E(0)[h, h] \leq C_2\|h\|_{H^{1/2}}^2. \] (3)

Since the norm \( \| \cdot \|_{H^{1/2}} \) is strictly weaker than the norm \( \| \cdot \|_{C^{2,\alpha}} \), there is no way to control the remainder \( o(\|h\|^2_{C^{2,\alpha}}) \) with \( D^2E(0)[h, h] \). Such a fact is also remarked by K. Eppler in [6] for the more elementary case of shape functionals, not depending on a state. The required analysis has to be more precise and to use the Taylor formula with integral rest. This strategy imposes:

- the construction of a path \( t \mapsto \Omega(t) \) inside the admissible shapes on which the formula

\[ e(1) = e(0) + \int_0^1 (1 - t)e''(t)dt \]

can be written for the function \( e(t) = E(\Omega(t)) \),

- the propagation of the positiveness \( e''(0) > 0 \), obtained by assumption at the initial point of the path – the critical geometry – to the whole path. Knowing that \( e''(t) > 0 \) ensures that \( e(1) > e(0) \).

Let \( L = -\text{div}(A(x)\nabla u) \) be a strictly and uniformly elliptic operator with \( A \in C^2(\mathbb{R}^d, M_d) \). Let \( j \) and \( f \) be functions in \( C^{0,\alpha}(\mathbb{R}^d) \) and \( C^3(\mathbb{R}^{d+1}) \) respectively. For each domain \( \Omega \), with a \( C^{2,\alpha} \) boundary, we define the state function \( u_\Omega \) as the solution of the boundary values problem (2) while the criterion \( E \) under consideration takes the form (1). The main point of [3] is the following continuity result.
Theorem 2.1 There exist a real \( \eta_0 \) and a modulus of continuity \( \omega : (0, \eta_0) \rightarrow \mathbb{R}^+ \), which depends only on \( \Omega_0, L, f \) and \( j \) such that for all \( \eta \in (0, \eta_0) \) and all \( \Theta \in C^{2,\alpha}(\mathbb{R}^d, \mathbb{R}^d) \) with
\[
\|\Theta - Id_{\mathbb{R}^d}\|_{2,\alpha} \leq \eta
\]
there exists a vector field \( X_{\Theta} \) whose flow defines a path \( (\Omega(t))_{t \in [0,1]} \) connecting \( \Omega_0 \) and \( \Theta(\Omega_0) \) with the following estimate: for all \( t \in [0,1] \),
\[
|e''(t) - e''(0)| \leq \omega(\eta)\|X_{\Theta} \cdot n_{\partial \Omega_0}\|^2_{H^{1/2}(\partial \Omega_0)}, \tag{4}
\]
As an application, we obtain the stability result.

Theorem 2.2 If \( \Omega_0 \) is a critical shape of a criterion \( E \) of the previous type and the Hessian \( D^2 E(\Omega_0) \) is coercive in the weak norm \( \| \cdot \|_{H^{1/2}(\partial \Omega_0)} \), then \( \Omega_0 \) is a stable equilibrium shape: there is an open ball around \( \Omega_0 \) in the \( C^{2,\alpha} \) topology on which \( E \) has a local strict minimum at \( \Omega_0 \).

Other functionals such as the volume \( V \) of \( \Omega \) or its perimeter \( P \) can be considered. In that case, the estimates corresponding to (4) are obtained in other norms. One has
\[
|V''(t) - V''(0)| \leq \omega(\eta)\|X_{\Theta} \cdot n_{\partial \Omega_0}\|^2_{L^2(\partial \Omega_0)},
\]
\[
|P''(t) - P''(0)| \leq \omega(\eta)\|X_{\Theta} \cdot n_{\partial \Omega_0}\|^2_{H^{1}(\partial \Omega_0)}.
\]

Extension of these results to the case of Neumann boundary conditions is studied in [4] with the same conclusions. Let us mention again the refined remainder estimate stated in Theorem 2.1 is the key point.

The shape Hessian as a pseudo-differential operator. The following considerations are based on the one-to-one identification between domains \( \Omega \in U_\delta(\Omega_0) \) and elements \( h \in U_\delta(0) \) of a certain Banach space \( X \). In particular, the choice \( X = C^{2,\alpha}(\partial \Omega_0) \) is convenient, for the parametrization via a (smooth) reference manifold. Then, we have for a twice continuously differentiable shape functional \( J : X \rightarrow \mathbb{R} \) that \( \nabla J(h) \in X^* \), \( D^2 J(h) \in L(X, X^*) \). Equivalently, the shape Hessian defines a continuous bilinear form on \( X \times X \)
\[
|D^2 J(h)[dh_1, dh_2]| \leq C\|dh_1\|_{X} \|dh_2\|_{X}, \quad \text{uniformly for all } h \in U_\delta(0).
\]
However, by investigating the structure of shape Hessians in more detail, they naturally extend to continuous bilinear forms on weaker spaces. The notation \( H^s \), \( L_2(= H^0) \) is used as abbreviation of \( H^s(\partial \Omega_0) \).

Theorem 2.3 For the objective function \( E \) of (1), the shape Hessian \( D^2 E \) satisfies the estimate
\[
|D^2 E(\Omega)[dh_1; dh_2]| \leq c_0 \|dh_1\|_{H^{1/2}} \cdot \|dh_2\|_{H^{1/2}}, \tag{5}
\]
and no similar estimate with respect to a weaker space is possible in general.

Of course, this corresponds with the observation (3). The analogue estimates for elementary domain and boundary integrals like volume and perimeter are
\[
|D^2 V(\Omega)[dh_1; dh_2]| \leq c_0 \|dh_1\|_{L_2} \cdot \|dh_2\|_{L_2},
\]
\[
|D^2 P(\Omega)[dh_1; dh_2]| \leq c_0 \|dh_1\|_{H^{1}} \cdot \|dh_2\|_{H^{1}}.
\]
In conclusion, if a shape functional $J$ is twice continuously differentiable at $h \in X$ with a boundary (parameter) integral representation for the shape Hessian $D^2 J(h)$, this bilinear form extends continuously to $H^s \times H^s$ for a certain $s \geq 0$,

$$|D^2 J(h)[dh_1; dh_2]| \leq C_0 \|dh_1\|_{H^s} \cdot \|dh_2\|_{H^s}, \quad \text{uniformly for all } h \in U_\delta(0).$$  \hfill (6)

The value $s$ is characteristic for each type of functionals. Moreover, the following important consequences are valid, cf. [11].

**Theorem 2.4** The shape Hessian $D^2 J(h)$ defines a continuous linear mapping from $H^s$ to $H^{-s} := (H^s)^*$, i.e., it is a pseudo-differential operator of order $2s$. Furthermore, it holds $DJ(h) \in H^{-s}$ for all $h \in U_\delta(0)$, and the gradient is locally Lipschitz as a mapping in the $(H^{-s}, H^s)$-duality, that is

$$\|DJ(h + dh) - DJ(h)\|_{H^{-s}} \leq c_0 \|dh\|_{H^s}$$  \hfill (7)

for all $h, h + dh \in B_X^U(0)$.

Of course, coercivity of the related bilinear form in the energy norm implies the nature of the related Riesz operator which is a *regular* pseudo-differential operator of order $2s$. Vice versa, if the compactness of the Riesz operator is shown, the related bilinear form cannot be strictly coercive. The latter is used in the subsequent investigations in section 3. We refer to [10] for further examples of typical structures of shape Hessians.

**Convergence for (semi)discretized problems.** It has been shown in [11] that strict coercivity in the energy norm is also the key assumption for proving the convergence of optimal solutions of (semi)discretized problems to the solution of the original problem (1), (2). For convenience, we skip the technicalities. Furthermore, we refer to [9] for further details about application of Newton’s method to a particular shape problem with additional functional constraints.

### 3 An unstable case.

This situation was first investigated by K. Eppler and H. Harbrecht on a model situation in Electrical Impedance Tomography in [7]. A strategy is developed in the paper to prove compactness of the Riesz operator associated to the shape Hessian based on a factorization in terms of layer potentials. This idea has then been successfully used in [1, 2]. We present here the general problem of E.I.T. studied in [2].

Let $\Omega$ be a bounded open set with smooth boundary in $\mathbb{R}^2$ or $\mathbb{R}^3$. Fix $d_0 > 0$ and consider inclusions $\omega$ such that $\omega \subset \Omega_{d_0} = \{x \in \omega, \ d(x, \partial \Omega) > d_0\}$. We also assume that the boundary $\partial \omega$ is of class $C^{4,\alpha}$. Consider two non-negative real numbers $\sigma_1$, $\sigma_2$ and the $L^\infty$ function $\sigma(x) = \sigma_1 1_\omega + \sigma_2 1_{\Omega \setminus \omega}$. Consider the boundary value problem

$$-\text{div} (\sigma(x) \nabla u) = 0 \quad \text{in } \Omega,$nabla u) = 0 \quad \text{in } \Omega,$$

$$u = f \quad \text{on } \partial \Omega. \quad \hfill (9)$$

Define the Dirichlet-to-Neumann map as

$$\Lambda_{\sigma} : f \mapsto \sigma (\partial_n u)|_{\partial \Omega},$$

4
where $u$ solves (8),(9) and $n$ is the outer unit normal to $\partial \Omega$. The inverse problem arises when one has access to the normal derivative of the potential $u$ that solves (8),(9). Assume that one knows

$$
\partial_n u = g \text{ on } \partial \Omega,
$$

(10)

then the problem (8)-(10) is overdetermined. The electrical impedance tomography problem we consider is to recover the shape of $\omega$ from the knowledge of the single Cauchy pair $(f,g)$.

In order to recover the shape of the inclusion $\omega$, a usual strategy is to minimize a cost function. Many choices are possible, it turns out that a Kohn and Vogelius type objective leads to a minimization problem with nicer properties than the least squares fitting approaches. We now define this criterion that involves two state functions $u_D$ and $u_N$: the state $u_D$ solves Dirichlet boundary value problem (8),(9) while $u_N$ solves Neumann boundary value problem (8),(10). Kohn-Vogelius objective $J_{KV}$ is defined as:

$$
J_{KV}(\omega) = \int_{\Omega} \sigma |\nabla (u_D - u_N)|^2
$$

(11)

Similar expressions hold for $u_N$. To compute the derivatives of the objective, we recall the derivatives of the state functions. Let $\Omega$ be a open subset of $\mathbb{R}^d$ ($d = 2$ or 3) with a $C^2$ boundary and $\omega$ a element of $\Omega_{\delta_0}$ with a $C^{4,\alpha}$ boundary. The state $u_D$ is shape differentiable and its shape derivative $u'_D \in H^1(\Omega \setminus \omega) \cup H^1(\omega)$ satisfies

$$
\begin{cases}
\Delta u'_D = 0 \text{ in } \Omega \setminus \omega \text{ and in } \omega, \\
[u'_D] = h_n \frac{[\sigma]}{[\sigma_1]} \partial_n u_D \text{ on } \partial \omega, \\
[\sigma \partial_n u'_D] = [\sigma] \text{div}_\tau (h_n \nabla_\tau u_D) \text{ on } \partial \omega, \\
u'_D = 0 \text{ on } \partial \Omega.
\end{cases}
$$

(12)

Then, one checks that the derivative of Kohn-Vogelius objective in the direction of a deformation field $h$ is given by:

$$
DJ_{KV}(\omega; h) = \int_{\partial \omega} \frac{\sigma_1}{\sigma_2} (|\partial_n u'_D|^2 - |\partial_n u'_N|^2) + |\nabla_\tau u_D|^2 - |\nabla_\tau u_N|^2 \big) h_n.
$$

(13)

To that end, we consider an admissible inclusion $\omega^*$ and solve both (8),(9) and (8),(10) to obtain the corresponding measurements $f^*$ and $g^*$. The domain $\omega^*$ realizes the absolute minimum of the criterion $J_{KV}$ since by construction $u_D = u_N$ in $\Omega$ and hence $J_{KV}(\omega^*) = 0$. We check that the Euler equation

$$
DJ_{KV}(\omega^*)(h) = 0,
$$

holds. Set $v = u_D - u_N$. We have proved in [2] that

$$
D^2 J_{KV}(\omega^*)(h,h) = \int_{\Omega} \sigma |\nabla v'|^2.
$$

(14)

Moreover, if $h_n \neq 0$, then $D^2 J_{KV}(\omega^*)(h,h) > 0$. Nevertheless, (14) does not mean that the minimization problem is well posed. In fact, the following theorem explains the instability of standard minimization algorithms.
Theorem 3.1 Assume that $\omega^*$ is a critical shape of $J_{KV}$ for which the additional condition $u_N = u_D$ holds, then the Riesz operator corresponding to $D^2J_{KV}(\omega^*)$ defined from $H^{1/2}(\partial \omega^*)$ with values in $H^{-1/2}(\partial \omega^*)$ is compact. Moreover, the minimization is severely ill-posed in the following sense: if $\lambda_n$ denotes the $n^{th}$ eigenvalue of $D^2J_{KV}(\omega^*)$, then $\lambda_n = o(n^{-s})$ for all $s > 0$.

Let us briefly explain how to extend the idea from [8]. We start from

$$D^2J_{KV}(\omega^*)(h, h) = -2 \int_{\partial \omega^*} \left[ \sigma \left( u'_D \partial_n u' - \partial_n u'_N v' \right) \right]$$

$$= 2 [\sigma] \left( u'_D - u'_N, \text{div}_\tau (h_n \nabla_\tau u_D) \right) - \frac{\sigma_1}{\sigma_2} \left( \partial_n u_D h_n, \partial_n (u'_D - u'_N)^+ \right)$$

obtained after two integrations by parts. Here $\langle \cdot, \cdot \rangle$ denotes the duality between $H^{1/2}(\partial \omega^*) \times H^{-1/2}(\partial \omega^*)$. With clear notations, the Hessian can then be written as:

$$D^2J_{KV}(\omega^*)(h, h) = 2 [\sigma] \left( \langle M_1(h), T_1(h) \rangle \right) = 2 [\sigma] \left( \frac{\sigma_1}{\sigma_2} \langle T_2(h), M_2(h) \rangle \right) .$$

Let us study now the compactness of $D^2J_{KV}$. Operators $T_1$ and $T_2$ are continuous as multipliers by a smooth function but $M_1$ and $M_2$ are compact. For example, let us prove that $M_1$ is a compact operator. We introduce the space

$$H^{1/2}(\partial \Omega) = \left\{ \phi \in H^{1/2}(\partial \Omega) : \int_{\partial \Omega} \phi = 0 \right\},$$

and remark that $M_1$ is the composition $R_1 \circ R_2$ of the operators:

$$R_1 : H^{1/2}(\partial \omega^*) \rightarrow H^{1/2}(\partial \Omega) \quad \text{and} \quad R_2 : H^{1/2}(\partial \Omega) \rightarrow H^{1/2}(\partial \omega^*)$$

$$h \rightarrow R_1(h) = -u'_N \quad \phi \rightarrow R_2(\phi) = \psi$$

where $\psi$ is the trace on $\partial \omega^*$ of $\Psi$ solution of

$$\begin{cases}
-\Delta u = 0 & \text{in } \Omega \setminus \overline{\omega^*} \text{ and in } \omega^*, \\
[u] = 0 & \text{on } \partial \omega^*, \\
[\sigma \partial_n u] = 0 & \text{on } \partial \omega^*, \\
u = \phi & \text{on } \partial \Omega.
\end{cases}$$

We check that $R_1$ is continuous. Let us show that $R_2$ is compact. To express $u|_{\partial \omega^*} = \psi$, we introduce the single and double layer potentials associated with the fundamental solution $\Gamma$ of Laplace equation and the surface $\partial \omega$ and $\partial \Omega$

$$S_{\partial \Omega \partial \omega} u(x) = \int_{\partial \Omega} \Gamma(x, y) u(y), \quad K_{\partial \Omega \partial \omega} u(x) = \int_{\partial \Omega} \partial_n \Gamma(x, y) u(y),$$

$$S_{\partial \omega \partial \Omega} u(x) = \int_{\partial \omega} \Gamma(x, y) u(y), \quad K_{\partial \omega \partial \Omega} u(x) = \int_{\partial \omega} \partial_n \Gamma(x, y) u(y).$$

Note that all these operators have a smooth kernel since the boundaries $\partial \omega$ and $\partial \Omega$ are assumed to have no common point; they are all compact. We also denote

$$S_{\Omega} u(x) = \int_{\partial \Omega} \Gamma(x, y) u(y), \quad K_{\Omega} u(x) = \int_{\partial \Omega} \partial_n \Gamma(x, y) u(y),$$

$$S_{\omega} u(x) = \int_{\partial \omega} \Gamma(x, y) u(y), \quad K_{\omega} u(x) = \int_{\partial \omega} \partial_n \Gamma(x, y) u(y).$$

Note that all these operators have a smooth kernel since the boundaries $\partial \omega$ and $\partial \Omega$ are assumed to have no common point; they are all compact.
Let us recall the regularity properties of these operators: for $s \geq 1/2$ and less than the regularity of the boundaries $\partial \omega$ and $\partial \Omega$ (here assumed to be $C^{4,\alpha}$ respectively $C^{2}$), the single layer operators $S_{\Omega}$ and $S_{\omega}$ are invertible from $H^{s-1}$ into $H^{s}$ while the double layer operators $K_{\Omega}$ and $K_{\omega}$ are compacts from $H^{s}$ into $H^{s}$.

We use the integral representation formula of $u$ in terms of the layer potentials:

$$
\begin{bmatrix}
\frac{1}{2}I + \mu K_{\omega}^* & \frac{\sigma_{1}}{\sigma_{2} + \sigma_{1}} S_{\partial \Omega \partial \omega}^* \\
\mu K_{\partial \omega \partial \Omega} & \frac{\sigma_{1}}{\sigma_{2} + \sigma_{1}} S_{\Omega}
\end{bmatrix}
\begin{bmatrix}
(u)|_{\partial \omega}^* \\
(\partial_{n} u)|_{\partial \Omega}
\end{bmatrix} = \frac{\sigma_{1}}{\sigma_{1} + \sigma_{2}}
\begin{bmatrix}
K_{\partial \Omega \partial \omega}^* \phi \\
(- \frac{1}{2} + K_{\Omega}) \phi
\end{bmatrix}.
$$

We remark that the matrix operator arising in this equation has a continuous inverse by the Fredholm alternative. Let us express $u|_{\partial \omega}^* = \psi$:

$$
\begin{bmatrix}
\frac{1}{2}I + \mu K_{\omega}^* - \mu S_{\partial \Omega \partial \omega}^* S_{\partial \omega \partial \Omega}^{-1} K_{\partial \omega \partial \Omega} \\
K_{\partial \Omega \partial \omega}^* - S_{\partial \Omega \partial \omega}^* S_{\Omega}^{-1} \left( \frac{1}{2}I - K_{\Omega} \right)
\end{bmatrix} \psi = \frac{\sigma_{1}}{\sigma_{1} + \sigma_{2}}
\begin{bmatrix}
K_{\partial \Omega \partial \omega}^* \phi \\
\left( \frac{1}{2}I - K_{\Omega} \right) \phi
\end{bmatrix}.
$$

Since the operators $K_{\partial \Omega \partial \omega}^*$ and $S_{\partial \Omega \partial \omega}^*$ are compact, $R_{2}$ then $M_{1}$ are compact.

The natural question is then how is this optimization problem ill-posed. This question is directly in relation with the rate at which the singular values of the Hessian operator are decreasing. Equation (3) shows that this rate is the one of the operators $K_{\partial \Omega \partial \omega}^*$ and $S_{\partial \Omega \partial \omega}^*$. Now, since for every $u \in H^{1/2}(\partial \Omega)$, the functions $K_{\partial \Omega \partial \omega}^* u$ and $S_{\partial \Omega \partial \omega}^* u$ are harmonic outside of $\partial \Omega$ and therefore in $\Omega$, their restrictions on $\partial \omega^*$ are $C^\infty$ then belongs to each $H^s(\partial \omega^*)$ for $s > 1/2$.

In figure 1, we depict a typical situation from numerical experiments in [7]. In particular, the plot exhibits clearly the exponential decay of the eigenvalues. The $\ell^2$-condition number of the discrete Hessian is about $10^9$.

![Figure 1: Logarithmic moduli of the coefficients of the discrete Hessian $d^2 J(\Omega^*)[dr_1, dr_2]$ and its eigenvalue distribution.](image)
References


