Control and Cybernetics

vol. **25** (1996) No. 5

Sensitivity computation and automatic differentiation

by

Ph. Guillaume and M. Masmoudi
INSA-UPS, CNRS-UMR 9974 MIP,
Avenue de Rangueil, 31077 Toulouse, France

Abstract: In shape optimization problems, each computation of the cost function by the finite element method leads to an expensive analysis. It is possible to make use of this analysis for getting more information using higher order derivatives. The additional cost of this computation is low with respect to the cost of the analysis. Moreover, automatic differentiation tools make it easy to implement, and provide exact derivatives of the discrete problem.

1. Introduction

What is shape optimal design? Consider for instance an elastic rod, with cross-section Ω (Fig. 1.).

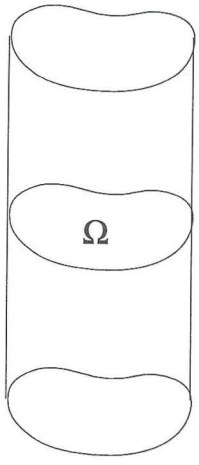


Figure 1. Elastic rod

The torsional rigidity of the rod is given by

$$j(\Omega) \quad = \quad J(\Omega, y_\Omega)$$

$$J(\Omega, y) = \int_\Omega |\nabla y_\Omega|^2 \, dx$$

where y_Ω is the solution to the partial differential equation

$$\begin{aligned} -\Delta y_\Omega &= 2 \quad \text{in} \quad \Omega \\ y_\Omega &= 0 \quad \text{on} \quad \Gamma \end{aligned}$$

We denote by Γ the boundary of the domain Ω which is supposed to be simply connected. We assume also that the measure of Ω is a fixed number. A classical shape optimization problem is to find Ω which maximizes the torsional rigidity of the rod (the solution to this problem is well-known: it is a circle).

The basic sheme of a shape optimization problem is the following:

$$\Omega \longmapsto y_\Omega \longmapsto j(y_\Omega)$$

and the real unknown of the problem is the domain Ω itself. We do not discuss here the existence of an optimal shape. This is a difficult problem which is yet not completely solved. The aim of this paper is to describe some basic tools for shape optimization, and essentially a new method of *higher order derivatives*. Indeed, in most problems the maps $\Omega \longmapsto y_\Omega$ and $y_\Omega \longmapsto j(y_\Omega)$ (or at least the map $\Omega \longmapsto y_\Omega$) are smooth. But one can ask the following questions:
- are higher order derivatives expensive to calculate?
- are they complicated to use?
- are they imprecise?
- are they useless?

At first sight, the answer seems to be positive, but classical results of Strassen (1990) and Morgenstern (1985) tell us that the higher order derivatives are not expensive and can be computed automatically. The purpose of this lecture is to give an answer to the third question by showing that the higher order derivatives of a function can be computed with the same precision as the function itself. We show also that the derivatives so computed are equal to the derivatives of the discrete problem. We call the discrete problem the finite dimensional problem processed by the computer. This result allows the use of automatic differentiation, which works only on discrete problems. Furthermore, the numerical results which are proposed at the end of the lecture give an answer to the last question

We give in Section 2. some basic results about differentiation with respect to the shape. We describe in Section 3. how to compute the derivatives and we propose in Section 4. a method which gives intrinsic expressions of these derivatives. The *mth* order derivatives are given in Section 5. The discretization error is studied in Section 6. and we illustrate the higher order derivatives method in Section 10.

2. Derivation with respect to shape

We consider now a general cost function $\mathbf{j}(\omega) = \mathbf{J}(\omega, y_\omega)$ where y_ω is the solution to a partial differential equation defined on a variable bounded domain ω of $\mathbb{R}^N$.

We suppose here that y_ω is the solution in a Hilbert space $\mathcal{V}(\omega)$ to the variational equation

$$\mathbf{a}(\omega, y_\omega, v) + \mathbf{l}(\omega, v) = 0 \quad \forall v \in \mathcal{V}(\omega). \tag{1}$$

Here $\mathbf{a}(\omega, ., .)$ is a continuous bilinear form satisfying the $\mathcal{V}(\omega)$-*ellipticity* property (Ciarlet, 1978), and $\mathbf{l}(\omega, .)$ is a continuous linear form on $\mathcal{V}(\omega)$.

EXAMPLE 2.1 *In many cases the forms* $\mathbf{a}(\omega, ., .)$ *and* $\mathbf{l}(\omega, .)$ *can be written as follows: for* $y, v \in H_0^1(\omega)$,

$$\begin{aligned} \mathbf{a}(\omega, y, v) &= \int_\omega \sum_{i,j} a_{i,j}\, \partial_i y\, \partial_j v\, dx \\ \mathbf{l}(F, v) &= -\int_\omega f\, v\, dx \end{aligned}$$

where $f \in L^2(\mathbb{R}^N)$ *and the functions* $a_{i,j} \in L^\infty(\mathbb{R}^N)$ *satisfy the following ellipticity condition: there exists a constant* $\alpha > 0$ *such that for any* x *and* y *in* $\mathbb{R}^N$

$$\sum_{i,j} a_{i,j}(x)\, y_i\, y_j \geq \alpha\, ||y||^2_{R^N}$$

which implies that there is a unique solution y_ω *to equation (1).*

The problem is studied on a fixed bounded domain Ω (Guillaume, Masmoudi, 1993; Masmoudi, 1987; Murat et Simon, 1976), and instead of looking for an optimal domain we are looking for a perturbation F such that the domain $F(\Omega)$ is optimal. Let $B(0, R)$ be an open ball of $\mathbb{R}^N$ containing $\overline{\Omega}$ and consider the Sobolev space of transformations $W^{1,\infty}(B(0, R); \mathbb{R}^N)$ equipped with the norm

$$||F|| = ||F||_{L^\infty(B(0,R);R^N)} + ||DF||_{L^\infty(B(0,R);\mathcal{L}(R^N))}$$

This choice allows us to start with any kind of domain Ω, even with an irregular boundary.

The set $U \subset W^{1,\infty}(B(0, R); \mathbb{R}^N)$ of all the maps F which are homeomorphisms from $B(0, R)$ to $F(B(0, R))$ with Lipschitz inverse, is open in $W^{1,\infty}(B(0, R); \mathbb{R}^N)$ (Murat et Simon; 1976, Guillaume, Masmoudi, 1993). We define a (respectively l, J, j) on $U \times \mathcal{V}(\Omega) \times \mathcal{V}(\Omega)$ (respectively $U \times \mathcal{V}(\Omega)$, $U \times \mathcal{V}(\Omega)$, U) by

$$\begin{aligned} a(F, y, v) &= \mathbf{a}(F(\Omega), y \circ F^{-1}, v \circ F^{-1}) \\ l(F, v) &= \mathbf{l}(F(\Omega), v \circ F^{-1}) \\ J(F, y) &= \mathbf{J}(F(\Omega), y \circ F^{-1}) \\ j(F) &= \mathbf{j}(F(\Omega)). \end{aligned}$$

This change of variable allows the use of classical differentiation tools in the normed space $W^{1,\infty}(B(0,R);\mathbb{R}^N)$, and we can easily derive higher order variations of j.

It will be assumed throughout the paper that

(A) *For any $y \in \mathcal{V}(\Omega)$ and any $F \in U$, $y \circ F^{-1}$ belongs to $\mathcal{V}(F(\Omega))$ and the linear mapping $y \mapsto y \circ F^{-1}$ is an homeomorphism from $\mathcal{V}(\Omega)$ onto $\mathcal{V}(F(\Omega))$.*

For instance this is the case when $\mathcal{V}(F(\Omega)) = H^1(F(\Omega))/\mathbb{R}$ (Neumann problem) or $\mathcal{V}(F(\Omega)) = H_0^1(F(\Omega))$ (Dirichlet homogeneous problem). That is also often the case when $H_0^1(F(\Omega)) \subset \mathcal{V}(F(\Omega)) \subset H^1(F(\Omega))$ (mixed problem). If we use a *conforming* finite element method (Ciarlet, 1978) then assumption (A) is also fulfilled for the discrete problem.

Hence $a(F,.,.)$ is a continuous and bilinear form satisfying the $\mathcal{V}(\Omega)$-*ellipticity* property and $l(F,.)$ a continuous linear form on $\mathcal{V}(\Omega)$. For $F \in U$ and $\omega = F(\Omega)$ equation (1) reads now ("direct state")

$$a(F, y_F, v) + l(F, v) = 0 \qquad \forall v \in \mathcal{V}(F(\Omega)) \tag{2}$$

where

$$y_F = y_{F(\Omega)} \circ F \tag{3}$$

The cost function becomes $j(F) = J(F,\, y_F)$.

THEOREM 2.1 *Let $y_F \in \mathcal{V}(\Omega)$ be the solution to equation (2). If the mappings $F \mapsto a(F,.,.)$ and $F \mapsto l(F,.)$ are of class $\mathcal{C}^m$ $(m \geq 0)$ on U then the mapping $F \mapsto y_F$ is of class $\mathcal{C}^m$. Moreover if J is of class $\mathcal{C}^m$ then j is also of class $\mathcal{C}^m$.*

Proof: Define $A_F \in \mathcal{L}(\mathcal{V}(F(\Omega)))$ and $L_F \in \mathcal{V}(F(\Omega))$ by

$$\begin{aligned} a(F,\, y,\, v) &= ((A_F y,\, v)) \quad \forall v \in \mathcal{V}(F(\Omega)) \\ l(F,\, v) &= ((L_F,\, v)) \quad \forall v \in \mathcal{V}(F(\Omega)) \end{aligned}$$

where $((.,.))$ denotes the scalar product on $\mathcal{V}(F(\Omega))$. Equation (2) reads

$$y_F = (A_F)^{-1} L_F$$

Hence the result is a simple consequence of the composed maps theorem.

In example 2.1, if the functions $a_{i,j} \in \mathcal{C}^m(\mathbb{R}^N)$ and $f \in H^m(\mathbb{R}^N)$ then the first condition of the theorem is fulfilled (cf. following lemma). For instance in the case of the Laplace equation $-\Delta\, y_{F(\Omega)} = f$, if the function f is of class $\mathcal{C}^\infty$, then the mapping $F \mapsto y_F$ is also of class $\mathcal{C}^\infty$.

LEMMA 2.1 *Murat and Simon (1976). For $m \geq 0$ the map $\Psi : F \longmapsto g \circ F$ is of class $\mathcal{C}^m$ from U into*
(i) $C^0(B(0,R))$ if $g \in \mathcal{C}^m(B(0,R))$
(ii) $\mathrm{L}^2(B(0,R))$ if $g \in \mathrm{H}^m(B(0,R))$
and, for $V^1, \ldots, V^m \in W^{1,\infty}(B(0,R);\mathbb{R}^N)$, its derivative is:

$$\Big(\mathrm{D}^m\Psi(F)(V^1, \ldots, V^m)\Big)(x) = \mathrm{D}^m g(F(x))(V^1(x), \ldots, V^m(x)) \qquad a.e.$$

(iii) The map $F \longmapsto \mathrm{D}F^{-1}$ is of class $\mathcal{C}^\infty$ from U into $\mathrm{L}^\infty(B(0,R))^{N^2}$.

3. Computation of the derivatives

There are two basic methods for computing the derivatives of the cost function j: the *direct* method and the *Lagrangian* method. For the first order derivative the second method is much more efficient than the first one. Both of them can be generalized at higher order derivatives. However, the complexity of the second method grows much faster than for the first one, so the best solution is to use the Lagrangian method once and then only the direct method.

3.1. Notations

We use the notation $\mathrm{D}^m g(F).(V_1, \ldots, V_m)$ for the m-th order derivative of a function g defined on U, evaluated in the directions $V_1, \ldots, V_m \in W^{1,\infty}(B(0,R);\mathbb{R}^N)$. This notation could lead to some confusion when taking the derivative of the map $F \longmapsto y_F$. Thus we denote by $\mathrm{D}_F^m y_F.(V_1, \ldots V_m)$ the m-th order derivative of the map $F \longmapsto y_F$ from U into $\mathcal{V}(\Omega)$ in the directions $V_1, \ldots V_m \in W^{1,\infty}(B(0,R);\mathbb{R}^N)$. For $m = 1$ or 2 we follow the usual notation, i.e. for $V, W \in W^{1,\infty}(B(0,R);\mathbb{R}^N)$:

$$\dot{y}_F.V := \mathrm{D}_F\, y_F.V, \quad \ddot{y}_F.(V,W) := \mathrm{D}_F^2\, y_F.(V,W).$$

We refer to these derivatives as the *total* derivatives (also called *material* derivatives). They are not to be confounded with the ordinary m-th order differential of a function y defined on an open set of $\mathbb{R}^N$: the m-th order derivative of the map $x \longmapsto y(x)$ in the directions $h_1, \ldots, h_m \in \mathbb{R}^N$ is denoted by $\mathrm{D}^m y(x).(h_1, \ldots, h_m)$ or simply by $\nabla y(x).h_1$ if $m = 1$ and $\mathrm{D}^2 y(x) h_1.h_2$ if $m = 2$ ($x.y$ denotes here the usual scalar product in $\mathbb{R}^N$). For $V_1, \ldots V_m \in W^{1,\infty}(B(0,R);\mathbb{R}^N)$ we denote by $\mathrm{D}^m y.(V_1, \ldots, V_m)$ the function defined by $(\mathrm{D}^m y.(V_1, \ldots, V_m))(x) = \mathrm{D}^m y(x).(V_1(x), \ldots, V_m(x))$.

3.2. The direct method

The direct method is very simple: use the chain rule. When the problem is differentiable, deriving the expression $j(F) = J(F, y_F)$ yields

$$Dj(F).V = D_1 J(F, y_F).V + D_2 J(F, y_F).(\dot{y}_F.V)$$

where the derivative $\dot{y}_F.V \in \mathcal{V}(\Omega)$ of y_F with respect to F in the direction V is the solution to the equation (obtained by deriving equation (2) with respect to F)

$$a(F, \dot{y}_F.V, v) + D_1 a(F, y_F, v).V + D_1 l(F, v).V = 0 \qquad \forall\, v \in \mathcal{V}(\Omega)$$

(we denote by D_i the partial derivative with respect to the ith argument).

When solving the discrete problem the transformations F of the domain are chosen in a finite subspace of $W^{1,\infty}(B(0,R);\mathbb{R}^N)$ (for instance, when using a P_1 finite element method, F is a continuous function which is linear on each

triangle). Let M be the dimension of this subspace. The derivative $\dot{y}_F.V$ has to be computed in M independent directions $V_1, \ldots, V_M$. Hence we need to solve $M+1$ systems for computing $j(F)$ and $Dj(F)$. For higher order derivative we need to solve $(M+k)!/(M!k!)$ systems for computing $j(F), \ldots, D^{(k)}j(F)$.

3.3. First order Lagrangian method

The Lagrangian method has been introduced by Céa (1986). It allows the elimination of the derivative of y_F with respect to F and leads to fast computation of the derivative when solving the discrete problem.

The Lagrangian is defined on $U \times \mathcal{V}(\Omega) \times \mathcal{V}(\Omega)$ by

$$\mathcal{L}(F, y, v) = J(F, y) + a(F, y, v) + l(F, v)$$

THEOREM 3.1 *Assume that the conditions of Theorem 2.1 hold with $m = 1$. Let $v_F \in \mathcal{V}(\Omega)$ be the solution to the equation (called "adjoint state")*

$$a(F, q, v_F) + \mathrm{D}_2 J(F, y_F).q = 0 \qquad \forall q \in \mathcal{V}(\Omega)$$

Then for all V in $W^{1,\infty}(B(0,R); \mathbb{R}^N)$ we have

$$\mathrm{D}j(F).V = \mathrm{D}_1 \mathcal{L}(F, y_F, v_F).V$$

Proof: Let v be an arbitrary element in $\mathcal{V}(\Omega)$; for all F in U we have

$$\mathcal{L}(F, y_F, v) = J(F, y_F) = j(F)$$

Hence if we differentiate this expression with respect to F and choose then $v = v_F$ we get the above-mentioned expression.

When solving the discrete problem we need now solving only two systems (one for y_F and one for v_F) instead of $M+1$ in the direct method. This reduces greatly the computer time spent for the computation of $\mathrm{D}j(F)$. For higher order derivatives the best solution is to use the Lagrangian method once and then to apply the direct method. We need hence to solve $2(M+k-1)!/(M!(k-1)!)$ systems for computing $j(F), \ldots, D^{(k)}j(F)$ (instead of $(M+k)!/(M!k!)$ when using only the direct method).

3.3.1. Example of the Laplace equation

Let f be a function in $H^1(\mathbb{R}^N)$ and let $y_{F(\Omega)} \in H_0^1(F(\Omega))$ be the solution to the equation

$$-\Delta y_{F(\Omega)} = f \qquad \text{in } F(\Omega)$$

Consider the cost function (energy):

$$\mathbf{j}(F(\Omega)) = \mathbf{J}(F(\Omega), y_{F(\Omega)}) = \int_{F(\Omega)} f\, y_{F(\Omega)}\, dx$$

The transported solution $y_F = y_{F(\Omega)} \circ F$ is the solution in $H_0^1(\Omega)$ to the equation

$$\begin{aligned} a(F, y_F, v) + l(F, v) &= 0 \qquad \forall v \in H_0^1(\Omega) \\ a(F, y, v) &= \int_\Omega \mathrm{D}F^{-T}\nabla y.\mathrm{D}F^{-T}\nabla v \, JF \, dx \\ l(f, v) &= -\int_\Omega f \circ F \; v \, JF \, dx \end{aligned}$$

(JF is the Jacobian of F and $\mathrm{D}F^{-T}$ is the transposed inverse of $\mathrm{D}F$) and the cost function is

$$\begin{aligned} j(F) &= J(F, y_F) \\ J(F, y) &= \int_\Omega f \circ F \; y \; JF \, dx \end{aligned}$$

According to Lemma 2.1 the function j is of class $\mathcal{C}^1$.

Using the direct method yields for all $V \in W^{1,\infty}(B(0,R); \mathbb{R}^N)$:

$$\mathrm{D}j(I).V = \int_\Omega y_I \, \nabla f.V + f \, \dot{y}_I.V + f \, y_I \, \nabla.V \, dx \tag{4}$$

(we denote by I the identity map on $B(0,R)$ and $\nabla.V = \sum_i \partial_i V^i$, $V = (V^1, \ldots, V^N)$).

There is another expression of the first order derivative involving the *local* derivative the definition of which is here recalled (Murat and Simon, 1976):

DEFINITION 3.1 *If for all $\omega \subset\subset \Omega$ the map $F \mapsto \left(y_{F(\Omega)}\right)_{|\omega}$ defined from a neighborhood of the identity I into $\mathrm{L}^2(\omega)$ is differentiable at $F = I$ then the map $F \mapsto y_{F(\Omega)}$ is said to be locally differentiable at I and the local derivative $y'_\Omega.V \in \mathrm{L}^2_{loc}(\Omega)$ is defined onto the whole domain Ω by:*

$$y'_\Omega.V = \lim_{t \to 0} \frac{\left(y_{(I+tV)(\Omega)}\right)_{|\omega} - \left(y_\Omega\right)_{|\omega}}{t} \quad \textit{for all } \omega \subset\subset \Omega$$

Note the difference with $\dot{y}_I.V = \lim_{t\to 0} \frac{(y_{(I+tV)(\Omega)}) \circ (I+tV) - y_I}{t}$. When $\dot{y}_I.V$ is well defined, the relationship between these two functions is:

$$\dot{y}_I.V = y'_\Omega.V + \nabla y_I.V \tag{5}$$

Thus (4) can also be written:

$$\mathrm{D}j(I).V = \int_\Omega \mathrm{div}(y_I \, f \, V) + f \, y'_\Omega.V \, dx$$

and using $y_I \in \mathrm{H}_0^1(\Omega)$ yields

$$\mathrm{D}j(I).V = \int_\Omega f \, y'_\Omega.V \, dx \tag{6}$$

The Lagrangian and its first order derivative are given for all $F \in U$ and $V \in W^{1,\infty}(B(0,R);\mathbb{R}^N)$ by the following expressions (with $A : B = \sum_{i,j} A^i_j B^i_j$):

$$\begin{aligned}
\mathcal{L}(F,y,v) &= \int_\Omega ((y-v)f\circ F + \mathrm{D}F^{-T}\nabla y.\mathrm{D}F^{-T}\nabla v)JF\,dx \\
\mathrm{D}_1\mathcal{L}(F,y,v).V &= \int_\Omega \Big(((y-v)f\circ F + \mathrm{D}F^{-T}\nabla y.\mathrm{D}F^{-T}\nabla v)\,\mathrm{D}F^{-T} : \mathrm{D}V \\
&\quad +(y-v)\nabla f\circ F.V - \mathrm{D}F^{-T}\mathrm{D}V^T\mathrm{D}F^{-T}\nabla y.\mathrm{D}F^{-T}\nabla v \\
&\quad - \mathrm{D}F^{-T}\nabla y.\mathrm{D}F^{-T}\mathrm{D}V^T\mathrm{D}F^{-T}\nabla v\Big)\,JF\,dx
\end{aligned}$$

In this particular case we find here that $v_I = -y_I$ (this is due to the fact that $J(F,y) = -l(F,y)$), and thus for all $V \in W^{1,\infty}(B(0,R);\mathbb{R}^N)$ the derivative of j at $F = I$ is:

$$\mathrm{D}j(I).V = \int_\Omega 2\,y_I\,\nabla f.V + (2\,f\,y_I - |\nabla y_I|^2)\,\nabla.V + 2\,\mathrm{D}V\,\nabla y_I.\nabla y_I\,dx \quad (7)$$

4. Intrinsic expression of the derivatives

The derivative of the cost function j in the direction $V \in W^{1,\infty}(B(0,R);\mathbb{R}^N)$ depends only on the values of V on the boundary Γ. Indeed, because of its very definition, if $F(\Omega) = G(\Omega)$ then $j(F) = j(G)$. Hence it is a consequence of the following Lemma from Murat and Simon (1976):

LEMMA 4.1 *Let g be a differentiable function defined on U. If $g(F) = g(G)$ when $F(\Omega) = G(\Omega)$ then for all $V \in W^{1,\infty}(B(0,R);\mathbb{R}^N)$ vanishing on Γ and all $F \in U$ one has*

$$\mathrm{D}g(F).V = 0.$$

(In fact when Γ and V are sufficiently smooth $\mathrm{D}g(I).V$ depends only on $V.n$). We describe in this section how to obtain intrinsic expressions of the first order derivative. What is meant by intrinsic is a little loose. On the one hand, as far as the first order derivative depends only on the restriction of V on Γ, one can consider as intrinsic an expression involving only the values of V on Γ. On the other hand we have also (recall $\dot{y}_I.V = y'_\Omega.V + \nabla y_I.V$), Simon (1980):

$$V = 0 \text{ on } \Gamma \quad \Longrightarrow \quad y'_\Omega.V = 0 \quad (8)$$

(whereas $\dot{y}_I.V \neq 0$ if $\nabla y_I.V \neq 0$) so one can consider intrinsic as well an expression involving $y'_\Omega.V$ rather than $\dot{y}_I.V$.

4.1. Extension of the local derivative

DEFINITION 4.1 *If the map $F \longmapsto y_F$ is differentiable from U into $\mathcal{V}(\Omega)$ then the local derivative $y'_F.V \in \mathrm{L}^2(\Omega)$ is defined for all $V \in W^{1,\infty}(B(0,R);\mathbb{R}^N)$ by:*

$$\begin{aligned}
\dot{y}_F.V &= y'_F.V + \nabla(y_{F(\Omega)})\circ F.V \\
&= y'_F.V + \mathrm{D}F^{-T}\nabla y_F.V \quad (9)
\end{aligned}$$

(recall that $\nabla(y \circ F^{-1}) \circ F = \mathrm{D}F^{-T}\nabla y$ *when* $F \in U$ *and* $y \in \mathrm{H}^1(F(\Omega))$*, Necas, 1967).*

REMARK 4.1 *Using definition 3.1 one finds the following relation:*

$$y'_F.V = \left(y'_{F(\Omega)}.(V \circ F^{-1})\right) \circ F$$

and particularly for $F = I$ *one has* $y'_I.V = y'_\Omega.V$.

We can now generalize property (8) in the following way:

LEMMA 4.2 *If the map* $F \longmapsto y_F$ *is differentiable from* U *into* $\mathcal{V}(\Omega)$ *then for all* $F \in U$ *and for all* $V \in W^{1,\infty}(B(0,R);\mathbb{R}^N)$ *vanishing on* Γ *one has:*

$$\dot{y}_F.V = \nabla(y_{F(\Omega)}) \circ F.V \tag{10}$$

Consequently the local derivative y'_F *has the same property as* y'_Ω*:*

$$V = 0 \text{ on } \Gamma \quad \Longrightarrow \quad y'_F.V = 0. \tag{11}$$

Proof: This property comes from the fact that when $F(\Omega) = G(\Omega)$ then $y_{F(\Omega)} = y_{G(\Omega)}$ (recall $y_{F(\Omega)} \circ F = y_F$). Let $F \in U$ be a fixed element, E_0 the set of all $V \in W^{1,\infty}(B(0,R);\mathbb{R}^N)$ vanishing on Γ, $\mathcal{A}$ the affine space $\mathcal{A} = F + \mathrm{E}_0$ and U' a neighborhood of F in $\mathcal{A}$ such that for all $G \in U'$ one has $G(\Omega) = F(\Omega)$. For $G \in U'$ one has $y_{G(\Omega)} = y_{F(\Omega)}$ thus $y_{G(\Omega)} \circ G = y_{F(\Omega)} \circ G$ which reads:

$$y_G = y_{F(\Omega)} \circ G.$$

Let $V \in \mathrm{E}_0$. From the differentiability hypothesis the left side has for derivative $\dot{y}_F.V$ at the point $G = F$. This is when deriving from U' into $\mathcal{V}(\Omega) \supset \mathrm{L}^2(\Omega)$, thus the same expression holds when deriving from U' into $\mathrm{L}^2(\Omega)$. Because $y_{F(\Omega)} \in \mathrm{H}^1(\Omega)$ we know from Lemma 2.1 that right side is differentiable with respect to G from U' into $\mathrm{L}^2(\Omega)$ and has $\nabla(y_{F(\Omega)}) \circ F.V$ for derivative at $G = F$. Hence equaling the two sides completes the proof. ■

REMARK 4.2 *Equality (10) tells that for* $F = I$ *and* $V \in W^{1,\infty}(B(0,R);\mathbb{R}^N)$ *vanishing on* Γ *on has:*

$$\nabla y_I.V = \dot{y}.V.$$

One rediscovers here in a surprising way a regularity result: as soon as the derivative exists the map $\dot{y}_I.V$ *belongs to* $\mathrm{H}^1(\Omega)$*, thus we have* $\nabla y_I.V \in \mathrm{H}^1(\Omega)$*. This means that* $(y_I)_{|\omega} \in \mathrm{H}^2(\omega)$ *for all* $\omega \subset\subset \Omega$*. However, observe that* $y'_\Omega.V$ *belongs only to* $\mathrm{L}^2(\Omega)$*: there is a loss of regularity when dealing with the intrinsic local derivative.*

4.2. From the domain to the boundary

We suppose here that the maps j and $F \longmapsto y_F$ are differentiable at $F = I$. We know that if $V = 0$ on Γ then $\mathrm{D}j(I).V = 0$ (lemma 4.1). This property will allow the use of the following theorem, leading to elegant expressions of the derivatives (recall that $x.y$ denotes the usual scalar product in $\mathbb{R}^N$, and $X : Y$ the natural contraction $X : Y = \sum_{i,j=1}^{N} X_{ij} Y_{ij}$. This notation stands also for vector-valued functions).

THEOREM 4.1 *Let $l \in W^{1,\infty}(B(0,R);\mathbb{R}^N)'$. Suppose that:*
(i) if $V = 0$ on Γ, then $l(V) = 0$,
(ii) the boundary Γ is piecewise of class $\mathcal{C}^1$,
(iii) the linear form l can be written

$$l(V) = \int_\Omega a\, \dot{y}_I.V + A : \mathrm{D}V + L.V \; dx \tag{12}$$

where the functions A, L and a belong respectively to $\mathrm{W}^{1,1}(\Omega)^{N\times N}$, $\mathrm{L}^1(\Omega)^N$ and $\mathrm{L}^2(\Omega)$. Then

$$l(V) = \int_\Omega a\, y'_\Omega.V \; dx + \int_\Gamma A : (V\, n^T)\; ds.$$

Thus one just has to substitute $y'_\Omega.V$ for $\dot{y}_I.V$, eliminate the term with $L.V$, and substitute $V\, n^T$ on the boundary for $\mathrm{D}V$ on the domain.

Proof: The proof of the theorem uses the following elementary lemma:

LEMMA 4.3 *Let $l_0 \in W^{1,\infty}(B(0,R);\mathbb{R}^N)'$. Suppose that:*
(i) if $V = 0$ on Γ, then $l_0(V) = 0$,
(ii) there is $L_0 \in \mathrm{L}^1(\Omega)^N$ such that $l_0(V) = \int_\Omega L_0.V \; dx$.
Then $l_0 = 0$.

Proof: Considering l_0 as a distribution, the support of l_0 is a subset of Γ so L_0 must vanish on Ω and $l_0 = 0$.
Proof of the theorem: Using (5) equation (12) reads:

$$l(V) = \int_\Omega a\,(y'_\Omega.V + \nabla y_I.V) + A : \mathrm{D}V + L.V \; dx$$

and using Green's formula yields:

$$l(V) \;=\; \int_\Omega \sum_{i,j} (-\partial_j A_{ij} + L_i + a\,\partial_i y_I) V^i + a\, y'_\Omega.V \, dx + \int_\Gamma A : (V\, n^T)\; ds$$

Let

$$l_0(V) = l(V) - \int_\Omega a\, y'_\Omega.V \, dx - \int_\Gamma A : (V\, n^T)\; ds$$

Due to (8) the linear form l_0 satisfies condition (i) of the precedent lemma. It satisfies also condition (ii). Thus $l_0 = 0$, which is the required result. ■

4.2.1. Application to the Laplace equation

Assume that the boundary is piecewise of class $\mathcal{C}^1$ and that $f \in \mathrm{H}^1(\mathbb{R}^N)$ in the example of the Laplace equation. Applying Theorem 4.1 to equation (4) we get directly the expression (6) involving the partial derivative $y'_\Omega.V$ instead of the total derivative $\dot{y}_I.V$.

More interesting, if the boundary is of class $\mathcal{C}^1$ then the solution y_I is in $\mathrm{H}^1_0(\Omega) \cap \mathrm{H}^2(\Omega)$. Applying Theorem 4.1 to equation (7) we obtain the boundary expression

$$\mathrm{D}j(I).V = \int_\Gamma -|\nabla\, y_I|^2\, V.n \; + 2(V.\nabla y_I)(n.\nabla y_I)\; ds. \tag{13}$$

Using the fact that $\nabla y_I(s)$ and the normal $n(s)$ are dependent this integral reduces to:

$$\mathrm{D}j(I).V = \int_\Gamma |\nabla\, y_I|^2\, V.n\; ds. \tag{14}$$

which is *not* a direct consequence of Green's formula and can be useful in a descent algorithm.

4.2.2. Application to electromagnetism

When dealing with antenna shape optimization, one often uses integral equations for solving the problem. Hence it is necessary to have a boundary expression of the derivative. In Millo (1991, Th. 3.5 p. 49), one finds the following expression for the derivative of a cost function j:

$$\begin{aligned}\mathrm{D}j(I).V = \mathrm{Re}\Big(& -k^2 \int_\Omega \Big((e.p) + \mathrm{curl}e.\mathrm{curl}p + \nabla.e\; \overline{\nabla.p}\Big)\; \nabla.V\; dx \\ & - \int_\Omega \mathrm{D}e\, \mathrm{D}V\, .\, (\overline{\mathrm{D}p - {}^t\mathrm{D}p}) + (\mathrm{D}e - {}^t\mathrm{D}e)\, .\, \overline{\mathrm{D}p}\, \mathrm{D}V\; dx \\ & - \int_\Omega \overline{\nabla.p}\; {}^t\mathrm{D}e\, .\, \mathrm{D}V + \nabla.e\; {}^t\overline{\mathrm{D}p}\, .\, \mathrm{D}V\; dx\Big).\end{aligned}$$

It was not easy to derive a boundary expression of this integral, and needed a few lemmas and complicated Green formulas. If we apply our method, we obtain directly

$$\begin{aligned}\mathrm{D}j(I).V = \mathrm{Re}\Big(& \int_{\partial\Omega} [(\mathrm{curl}e.\mathrm{curl}p + \nabla.e\; \overline{\nabla.p}) - k^2\, (e.p)]\, V.n\; ds \\ & - \int_{\partial\Omega} [({}^t\overline{\mathrm{D}p - \mathrm{D}p})\mathrm{D}e + ({}^t\mathrm{D}e - \mathrm{D}e)\overline{\mathrm{D}p}]\, V.n\; ds \\ & - \int_{\partial\Omega} (\overline{\nabla.p}\; \mathrm{D}e + \nabla.e\; \overline{\mathrm{D}p})\, V.n\; ds\Big),\end{aligned}$$

which is the same expression as the one found by Millo (Millo, 1991, Prop. 3.9 p. 55).

5. The mth order derivatives

Our purpose is now to generalize the Lagrangian method, and to express the mth order derivative of j in the form

$$D^m j(F).V = D_1\mathcal{L}^m(F,\, y_{F,V},\, p_{F,V}).V_m \tag{15}$$

where $V = (V_1, ..., V_m) \in (W^{1,\infty}(B(0,R);\mathbb{R}^N))^m$, and $y_{F,V}, p_{F,V} \in (\mathcal{V}(\Omega))^{2^{m-1}}$.

Notations.

We denote by $\mathcal{V}$ the space $\mathcal{V}(\Omega)$; for all nonnegative integer m, we set $s = 2^m$; we denote by $\mathcal{V}_m$ the Cartesian product $\mathcal{V}^s$, which is equipped with the norm

$$||(y_1, ..., y_s)||_{\mathcal{V}_m} = ||y_1||_{\mathcal{V}} + ... + ||y_s||_{\mathcal{V}}\,.$$

- Let $y = (y_1, ..., y_s)$ be an element of $\mathcal{V}_m$; the element $(0, ..., 0, y_i, 0, ..., 0) \in \mathcal{V}_m$ (y_i is the ith component) will be written simply y_i , and $(y_1, y_2, ..., y_i, 0, ..., 0) \in \mathcal{V}^i \times \{0_{\mathcal{V}}\}^{s-i} \subset \mathcal{V}_m$ will be written y^i , with the convention $y^0 = 0$.
- For $y \in \mathcal{V}$, we denote by y_i the element $(0, ..., 0, y, 0, ..., 0) \in \mathcal{V}_m$ (y is the ith component).

Note that in both cases, $y_i \in \{0_{\mathcal{V}}\}^{i-1} \times \mathcal{V} \times \{0_{\mathcal{V}}\}^{s-i} \subset \mathcal{V}_m$.

Induction formulas.

From now on, we choose in $W^{1,\infty}(B(0,R);\mathbb{R}^N)$ m directions of derivation $V_1, ..., V_m$. We define l^1 on $U \times \mathcal{V} \times \mathcal{V}$ by $l^1(F, y, p) = l(F, p)$. The function J, the bilinear form a and the Lagrangian $\mathcal{L}$ will be denoted J^1, a^1, $\mathcal{L}^1$; let $Y = (y, p)$ and $P = (q, r)$ be functions belonging to $\mathcal{V}_m = \mathcal{V}_{m-1} \times \mathcal{V}_{m-1}$ ($m \geq 1$); whenever possible, we define by induction (in m)

$$J^{m+1}(F, Y) = D_1\mathcal{L}^m(F, y, p).V_m \tag{16}$$

$$a^{m+1}(F, Y, P) = a^m(F, y, r) + a^m(F, q, p) \tag{17}$$

$$l^{m+1}(F, Y, P) = l^m(F, y, r) + D_2J^m(F, y).q + D_2l^m(F, y, p).q \tag{18}$$

$$\mathcal{L}^{m+1}(F, Y, P) = J^{m+1}(F, Y) + a^{m+1}(F, Y, P) + l^{m+1}(F, Y, P) \tag{19}$$

REMARK 5.1 *If J, a and l are of class $\mathcal{C}^d$, then for $0 \leq m \leq d$, the mappings $\mathcal{L}^{m+1}$, J^{m+1} and l^{m+1} are at least of class $\mathcal{C}^{d-m}$, $a^{m+1}(F, \bullet, \bullet)$ is a continuous bilinear form and $l^{m+1}(F, Y, \bullet)$ a continuous linear form on $\mathcal{V}_m$.*

It will be seen that the element $Y_{F,V} = (y_{F,V}, p_{F,V}) \in \mathcal{V}_m$ involved in (15) is the unique solution to the equation

$$a^{m+1}(F, Y_{F,V}, P) + l^{m+1}(F, Y_{F,V}, P) = 0 \quad \forall P \in \mathcal{V}_m\,.$$

THEOREM 5.1 *Let a, l and J be of class $\mathcal{C}^d$. Then for $0 \leq m \leq d$, the equation*

$$a^{m+1}(F, Y, P) + l^{m+1}(F, Y, P) = 0 \quad \forall P \in \mathcal{V}_m \tag{20}$$

has a unique solution $Y_{F,V} \in \mathcal{V}_m$, and the map $F \mapsto Y_{F,V}$ is of class $\mathcal{C}^{d-m}$ on U. Moreover, solving equation (20) is equivalent to solving the following triangular system of variational equations:

$$\begin{aligned}
&a(F, y_1, p) + l^{m+1}(F, y^0, p_s) = 0 && \forall p \in \mathcal{V}\\
&a(F, p, y_2) + l^{m+1}(F, y^1, p_{s-1}) = 0 && \forall p \in \mathcal{V}\\
&a(F, y_{2i-1}, p) + l^{m+1}(F, y^{2i-2}, p_{s-2i+2}) = 0 && \forall p \in \mathcal{V}\\
&a(F, p, y_{2i}) + l^{m+1}(F, y^{2i-1}, p_{s-2i+1}) = 0 && \forall p \in \mathcal{V}\\
&a(F, y_{s-1}, p) + l^{m+1}(F, y^{s-2}, p_2) = 0 && \forall p \in \mathcal{V}\\
&a(F, p, y_s) + l^{m+1}(F, y^{s-1}, p_1) = 0 && \forall p \in \mathcal{V}
\end{aligned}$$

Proof: The equivalence of (20) with this system of equations can be proved by induction in m. Then, according to remark 5.1, l^{m+1} is of class $\mathcal{C}^{d-m}$; thus we just have to apply theorem 2.1 to each equation of the system. ■

REMARK 5.2 *This theorem and formulas (17) and (18) allow us to set (20) under the following equivalent form:*
$Y_{F,V} = (y_{F,V}, p_{F,V}) \in \mathcal{V}_m$ is the unique solution to (20) if and only if $y_{F,V} \in \mathcal{V}_{m-1}$ is the unique solution to the equation

$$a^m(F, y_{F,V}, r) + l^m(F, y_{F,V}, r) = 0 \qquad \forall r \in \mathcal{V}_{m-1}$$

and $p_{F,V} \in \mathcal{V}_{m-1}$ the unique solution to the equation

$$a^m(F, q, p_{F,V}) + D_2 J^m(F, y_{F,V}).q + D_2 l^m(F, y_{F,V}, p_{F,V}).q = 0$$

$$\forall q \in \mathcal{V}_{m-1}.$$

Note that these equations generalize the direct and adjoint states described in section 1.

THEOREM 5.2 *We assume that a, J and l are of class $\mathcal{C}^m$.*
Let $Y_{F,V} = (y_{F,V}, p_{F,V}) \in \mathcal{V}_m$ be the unique solution to the equation

$$a^{m+1}(F, Y_{F,V}, P) + l^{m+1}(F, Y_{F,V}, P) = 0 \quad \forall P \in \mathcal{V}_m;$$

then we have

$$D^m j(F).V = D_1 \mathcal{L}^m(F, y_{F,V}, p_{F,V}).V_m \overset{def}{=} J^{m+1}(F, Y_{F,V}).$$

This can be proved in the same way as Theorem 3.1.

Note that these results are still valid if we replace $W^{1,\infty}(B(0,R); \mathbb{R}^N)$ by a subspace $E \subset W^{1,\infty}(B(0,R); \mathbb{R}^N)$ and U by $U \cap E$; this is especially the case when one uses the finite element method.

6. Discretization of the problem

Let us write Taylor's expansion of the cost function j:

$$j(F+V) = j(F) + \sum_{i=1}^{m} \frac{1}{i!} D^i j(F).V^{(i)} + o(||V||^m).$$

We denote by $V^{(i)}$ the element $(V, ..., V) \in (W^{1,\infty}(B(0,R); \mathbb{R}^N))^i$.

According to Remark 5.2 and Theorem 5.2, each term $D^i j(F).V^{(i)}$ can be written in the form $J^{i+1}(F, Y_{F,V})$ where $Y_{F,V}$ is the solution to (20).

Therefore, the mth order Taylor's expansion of $j(F+V)$ depends on $Y_{F,V}$ in the following way:

$$j(F) + \sum_{i=1}^{m} \frac{1}{i!} D^i j(F).V^{(i)} = T_m(Y_{F,V}),$$

where we have defined for $y \in \mathcal{V}_m$

$$T_m(y) = j(F) + \sum_{i=1}^{m} \frac{1}{i!} J^{i+1}(F, y).$$

Note that $J^{i+1}(F, y)$ is only applied to the 2^i first components of y.

Now the problem arises that we cannot calculate $T_m(Y_{F,V})$; the only thing that we can do is to compute the value of $T_m(Z_{F,V})$ (which we call the finite element approximation of Taylor's expansion of $j(F+V)$), where $Z_{F,V}$ is the unique solution in $(\mathcal{V}_h)_m$ to the equation

$$a^{m+1}(F, Z_{F,V}, P) + l^{m+1}(F, Z_{F,V}, P) = 0 \quad \forall P \in (\mathcal{V}_h)_m. \tag{21}$$

Here $\mathcal{V}_h = \mathcal{V}_h(\Omega)$ is the finite element space (Ciarlet, 1978), and we assume that $\mathcal{V}_h$ is a subspace of $\mathcal{V}$, so that $T_m(Z_{F,V})$ is well defined. Thus we approximate the cost function by $T_m(Z_{F,V})$, which can be seen as the "discretization of the derivatives of the continuous problem".

One could naturally ask the following question: do we get the same result if we first approximate the problem by the finite element method, and then differentiate the approximate cost function?
As we see it now, the answer is yes.

Recall that $j(F) = J(F, y_F)$, where $y_F \in \mathcal{V}$ is the solution to the equation

$$a(F, y_F, p) + l(F, p) = 0 \quad \forall p \in \mathcal{V}.$$

It is what we call the "continuous problem".

We introduce now the discrete cost function $j_h(F) = J(F, z_F)$, where $z_F \in \mathcal{V}_h$ is the solution to the equation

$$a(F, z_F, p) + l(F, p) = 0 \quad \forall p \in \mathcal{V}_h.$$

This is the "discrete problem".

From now on we suppose that $F \in U$ is a fixed element. If the assumption (A) is still valid for the space $\mathcal{V}_h(F(\Omega))$ (this means that the linear map $y \mapsto y \circ F^{-1}$ is a homeomorphism from $\mathcal{V}_h(\Omega)$ onto $\mathcal{V}_h(F(\Omega))$), we can substitute $\mathcal{V}_h(\Omega)$ for $\mathcal{V}(\Omega)$ and j_h for j in sections 3. and 5., so that the result still holds, i.e. we have

$$D^m j_h(F).V = J^{m+1}(F, Z_{F,V})$$

where $Z_{F,V} \in (\mathcal{V}_h)_m$ is the solution to (21). This equality means that the result does not depend on the order in which discretization and differentiation are made.

7. Error estimate on the approximate solution

We assume now that Ω is an open set of $\mathbb{R}^2$ or $\mathbb{R}^3$, and it is a finite union of simplicial or curved elements. Equation (21) is solved by using Lagrange's finite elements.

Notations.

- The norms of the spaces $H^p(\Omega)$ and $(H^p(\Omega))^r$ are denoted by $|| \; ||_p$.
- $\Pi_h \; : \; \mathcal{V} \cap \mathcal{C}^0(\Omega) \longrightarrow \mathcal{V}_h$ denotes the Lagrange's interpolation operator (Ciarlet, 1978).
- We denote by $y = (y_1, \ldots, y_s)$ the solution $Y_{F,V}$ to (20) and by $z = (z_1, \ldots, z_s)$ the solution $Z_{F,V}$ to (21).
- For $p \in \mathcal{V}$, $q \in \mathcal{V}_m$ and $1 \leq i \leq s$, we define g_i by

$$g_i(q, p) = l^{m+1}(F, q, p_{s-i+1}).$$

- For the sake of simplicity, we denote $a(F, p, r)$ by $a(p, r)$ and we suppose that the bilinear form a is symmetric.

Thus, the system of Theorem 5.1 becomes:

$$\left\{ \begin{array}{l} a(y_1, p) + g_1(y^0, p) = 0 \quad \forall p \in \mathcal{V} \\ a(y_2, p) + g_2(y^1, p) = 0 \quad \forall p \in \mathcal{V} \\ \qquad \vdots \\ a(y_i, p) + g_i(y^{i-1}, p) = 0 \quad \forall p \in \mathcal{V} \\ \qquad \vdots \\ a(y_s, p) + g_s(y^{s-1}, p) = 0 \quad \forall p \in \mathcal{V} \end{array} \right. \tag{22}$$

and z is the solution to the system obtained by substituting $\mathcal{V}_h$ for $\mathcal{V}$.

Let $k \geq 1$ be the degree of the finite element space. Recall that when the family of finite elements is *regular* (Ciarlet, 1978), the following property holds:

(P) There exists a constant c such that for any function $u \in \mathcal{V} \cap H^{k+1}(\Omega)$:

$$||u - \Pi_h u||_1 \leq c\, h^k ||u||_{k+1}.$$

THEOREM 7.1 *Assume that the family of finite elements is regular (Ciarlet, 1978), that the maps a, J and l are of class C^m and that their m-th order derivatives satisfy a Lipschitz condition. If y belongs to $(H^{k+1}(\Omega))^s$, then*

$$||y - z||_1 = O(h^k) .$$

Proof: It is sufficient to prove that for $1 \leq j \leq s$, we have

$$||y^j - z^j||_1 = O(h^k) .$$

We prove it by induction in j.

For $j = 1$, it is a classical result (Ciarlet, 1978).

Let the result be true for $1 \leq j - 1 \leq s - 1$; it follows from Strang lemma (Ciarlet, 1978) that there exists a constant C such that:

$$||y_j - z_j||_1 \leq C\, h^k\, ||y_j||_{k+1} + ||g_j(z^{j-1}, \bullet) - g_j(y^{j-1}, \bullet)||_{\mathcal{V}'}.$$

Moreover, taking into account the definition of g and the Lipschitz condition in the theorem, there exists $\eta > 0$ and $L > 0$ such that

$$||u - y^{j-1}||_1 \leq \eta \Longrightarrow ||g_j(u, \bullet) - g_j(y^{j-1}, \bullet)||_{\mathcal{V}'} \leq L\, ||u - y^{j-1}||_1 .$$

Now it follows from the induction hypothesis that the inequality

$$||y^{j-1} - z^{j-1}||_1 = \sum_{i=1}^{j-1} ||y_i - z_i||_1 \leq \eta$$

holds for h small enough; therefore, for such a small h we have

$$||y_j - z_j||_1 \leq C\, h^k\, ||y_j||_{k+1} + L\, ||y^{j-1} - z^{j-1}||_1,$$

and the result is obtained by using the induction hypothesis once more. ■

8. Error estimate on the approximate derivative

Let us define the error on the approximate derivative by

$$|J^{m+1}(F, y) - J^{m+1}(F, z)|.$$

A first estimate of this error is given by the following theorem:

THEOREM 8.1 *Assume that the family of finite elements is regular, that the maps a, J et l are $m + 1$ times differentiable and that their $(m + 1)$th order derivatives are locally bounded. If y belongs to $(H^{k+1}(\Omega))^s$, then*

$$J^{m+1}(F, y) - J^{m+1}(F, z) = O(h^k) .$$

The proof is easily obtained by using the mean value theorem and Theorem 7.1.

Under additional weak assumptions, it is possible to improve the error estimate. For that purpose, we recall the lemma introduced in Masmoudi (1987).

LEMMA 8.1 *Assume that the family of finite elements is regular. Let b be a continuous linear form on $\mathcal{V}$, and assume that the solution $w(b) \in \mathcal{V}$ to the equation*

$$a(p, w(b)) + b(p) = 0 \quad \forall p \in \mathcal{V}$$

belongs to $H^{k+1}(\Omega)$. If y_1 belongs to $H^{k+1}(\Omega)$, then

$$b(y_1 - z_1) = O(h^{2k}).$$

Proof: We recall from (22) that

$$a(y_1, p) + l(F, p) = 0 \qquad \forall p \in \mathcal{V}$$

$$a(z_1, p) + l(F, p) = 0 \qquad \forall p \in \mathcal{V}_h,$$

which gives

$$a(y_1 - z_1, p) = 0 \qquad \forall p \in \mathcal{V}_h.$$

By using the definition of $w(b)$ and the previous equality, we have

$$b(y_1 - z_1) = a(z_1 - y_1, w(b)) = a(z_1 - y_1, w(b) - \Pi_h(w(b))).$$

Hence, the continuity of the form a, property (P) and theorem 7.1 allow us to conclude. ■

The proof of the previous lemma is based on the fact that for all $p \in \mathcal{V}_h$, we have $a(y_1 - z_1, p) = 0$. But this is not the case when the linear form is not computed exactly, as in

$$a(y_2, p) + g_2(y^1, p) = 0 \quad \forall p \in \mathcal{V}$$

$$a(z_2, p) + g_2(z^1, p) = 0 \quad \forall p \in \mathcal{V}_h.$$

However, if g_2 is regular, one can get a similar result for $b(y_2 - z_2)$. Thus, we obtain the following lemma:

LEMMA 8.2 *Assume that the family of finite elements is regular and that the maps a, J et l are $m+2$ times differentiable. Let b be a continuous linear form on $\mathcal{V}$. Assume that the solution $w(b) \in \mathcal{V}$ to the equation*

$$a(p, w(b)) + b(p) = 0 \quad \forall p \in \mathcal{V}$$

belongs to $H^{k+1}(\Omega)$, and that for any $v \in H^{k+1}(\Omega)$, the solutions $w_{i,j}(v) \in \mathcal{V}$ to the equation

$$\sum_{i=1}^{j-1} a(w_{i,j}(v), p_i) + D_1 g_j(y^{j-1}, v).p = 0 \quad \forall p \in \mathcal{V}_m$$

belong also to $H^{k+1}(\Omega)$ for $1 \le i \le j-1 \le s-1$. If y belongs to $H^{k+1}(\Omega)$, then

$$b(y_i - z_i) = O(h^{2k}) \quad \forall i \quad 1 \le i \le s.$$

We can now give a better estimate of the error with the following theorem:

THEOREM 8.2 *If the conditions of lemma 8.2 hold for all the linear maps* $p_i \mapsto D_2 J^{m+1}(F, y).p_i$ $(1 \leq i \leq s)$, *then*

$$J^{m+1}(F, y) - J^{m+1}(F, z) = O(h^{2k}).$$

Proof: Let us write the second order Taylor's expansion of $J^{m+1}(F, \bullet)$:

$$J^{m+1}(F, z) - J^{m+1}(F, y) = D_2 J^{m+1}(F, y).(z - y) +$$

$$+ D^2_{22} J^{m+1}(F, y).(z - y)^{(2)} + o(||y - z||^2);$$

The bilinear map $D^2_{22} J^{m+1}(F, y)$ is continuous and y belongs to $H^{k+1}(\Omega)$, so we just have to use Theorem 7.1 and to apply Lemma 8.2 to $D_2 J^{m+1}(F, y).(z - y)$ to get the conclusion. ■

REMARK 8.1 *It can easily be proved that, in example (E), if* $f \in H^m(\mathbb{R}^N)$ *and if the functions* $a_{i,j} \in W^{m+1,\infty}(\mathbb{R}^N)$, *then the assumptions of Lemma 8.2 hold when the boundary of* Ω *is regular enough, Grisvard (1985).*

9. Direct computing of the mth order derivative

Instead of using the Lagrangian method, it is possible to get the mth order derivative of j by using the chain rule in the expression $j(F) = J(F, y_F)$; this time we need the derivatives of y_F with respect to F.
We shall denote by $D^m y_F$ the mth order derivative of the map $F \mapsto y_F$ (note that this is not the ordinary derivative of the function y_F).

9.1. Computing $D^m y_F.(V_1, ..., V_m)$

In this section, let $l^0(F, p) = l(F, p)$ and define by induction (in m)

$$l^{m+1}(F, p) = D_1 a(F, D^m y_F.(V_1, ..., V_m), p).V_{m+1} + D_1 l^m(F, p).V_{m+1}\,.$$

It is easy to check that if a, J and l are of class $\mathcal{C}^d$, then for $0 \leq m \leq d$, then $D^m y_F.(V_1, ..., V_m) \in \mathcal{V}$ is the solution to the equation

$$a(F, q, p) + l^m(F, p) = 0 \qquad \forall p \in \mathcal{V}\,.$$

9.2. Discretization of the direct method

Given any function φ defined on U, let

$$G(J, F, \varphi(F), D\varphi(F), ..., D^m\varphi(F), V)$$

be the formal expression obtained when differentiating at order m the mapping $F \mapsto J(F, \varphi(F))$ in the direction $V = (V_1, ..., V_m)$; for instance we have

$$D^m j(F).V = G(J, F, y_F, Dy_F, ..., D^m y_F, V) .$$

Recall that $z_F \in \mathcal{V}_h$ is the solution to the discrete problem:

$$a(F, z_F, p) + l(F, p) = 0 \qquad \forall p \in \mathcal{V}_h .$$

As we did it previously for the continuous problem, let $l_h^0(F, p) = l(F, p)$ and define by induction (in m)

$$l_h^{m+1}(F, p) = D_1 a(F, D^m z_F.(V_1, ..., V_m), p).V_{m+1} + D_1 l_h^m(F, p).V_{m+1} .$$

The function $D^m z_F.(V_1, ..., V_m) \in \mathcal{V}_h$ is then the solution to the equation

$$a(F, q, p) + l_h^m(F, p) = 0 \qquad \forall p \in \mathcal{V}_h .$$

If one uses this direct method, the discrete approximation of $D^m j(F).V$ should naturally be

$$G(J, F, z_F, Dz_F, ..., D^m z_F, V) ,$$

but is it the same as the one made with the Lagrangian method? Fortunately it is:

THEOREM 9.1 *Assume that a, J et l are of class C^m. Then we have*

$$G(J, F, z_F, Dz_F, ..., D^m z_F, V) = J^{m+1}(F, z) .$$

As a consequence, the approximations made by using either of the two methods are the same; especially, if the conditions of Theorem 8.1 hold, we have

$$D^m j(F).V - G(J, F, z_F, Dz_F, ..., D^m z_F, V) = O(h^k) ,$$

and if the conditions of Theorem 8.2 hold, we have

$$D^m j(F).V - G(J, F, z_F, Dz_F, ..., D^m z_F, V) = O(h^{2k}) .$$

Proof: Recall that $j_h(F) = J(F, z_F)$; according to the definition of G,

$$D^m j_h(F).V = G(J, F, z_F, Dz_F, ..., D^m z_F, V) .$$

The result is then a simple consequence of the equality $D^m j_h(F).V = J^{m+1}(F, z)$, proved in section 5. ∎

Thus one can differentiate after making discretization or discretize after differentiating with either of the two methods.

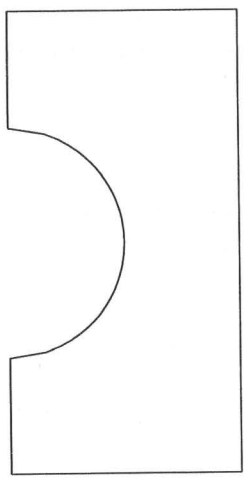
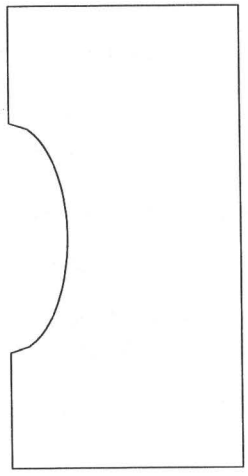


Figure 2. Initial design



modified design $(I+V)(\Omega)$

10. Examples

10.1. Numerical results for the torsional rigidity example

With intent to illustrate numerically the use of higher order derivatives, let us take the popular example of the torsional rigidity of an elastic bar, whose cross section is an open set $(I+V)(\Omega)$. Ω is the initial domain (Fig. 2), and $(I+V)(\Omega) := \{x + V(x);\, x \in \Omega\}$ is a perturbation of this domain.

The cost function (torsional rigidity) for a cross section $(I+V)(\Omega)$ is

$$j(I+V) = 2 \int_{(I+V)(\Omega)} y_{(I+V)(\Omega)}\, dx\,,$$

where $y_{(I+V)(\Omega)} \in H_0^1\Big((I+V)(\Omega)\Big)$ is the solution to the Laplacian equation

$$-\Delta\, y_{(I+V)(\Omega)} = 2\,.$$

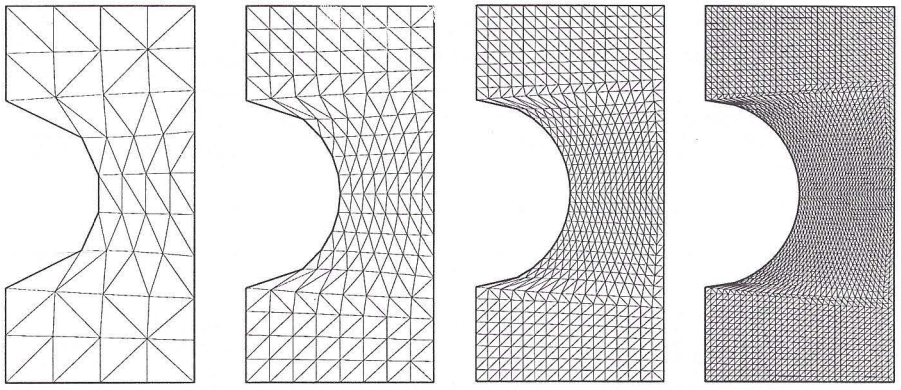
This cost function j is of class $\mathcal{C}^\infty$.

10.1.1. First design perturbation

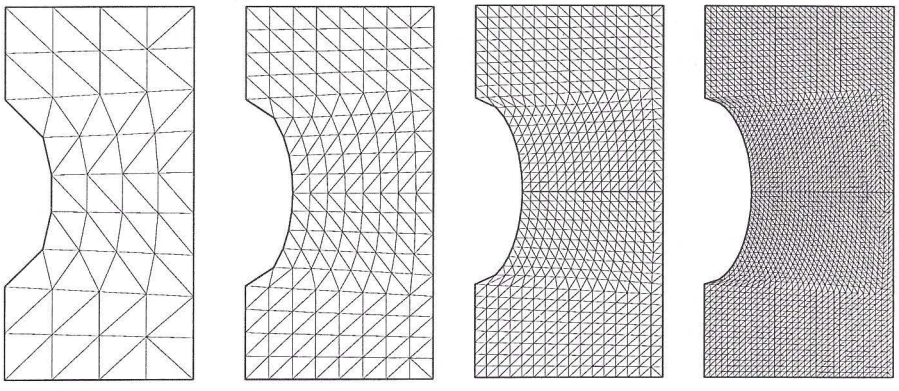
Here the modified design is given in Fig. 2. We are using finite elements of degree 1, using smaller and smaller elements: at each step, the size of the elements becomes half of the previous ones, as shown on Fig. 3.

We give in Table 1 the results of the computing of $j(I+V)$ obtained when using the Taylor's expansion of j at the point I, which have to be compared with the ones obtained when computing directly $j(I+V)$ on the modified domain.

Observe that the nodal table is made of the components of the map I in an appropriate basis. When V is a perturbation of the identity, the modified design



triangulation 1.1 triangulation 2.1 triangulation 3.1 triangulation 4.1



triangulation 1.2 triangulation 2.2 triangulation 3.2 triangulation 4.2

Figure 3. Triangulations of initial design and modified designs

Order of exp.	triangulation 1	triangulation 2	triangulation 3	triang. 4
0	.165754759809873	.17144321842656	.17416741916434	.1751508015
1	.169	.178	.181	.182
3	.1754	.1840	.1870	.1879
5	.174584	.183198	.18630	.18719
10	.1745975	.183210	.1862835	.1871720
20	.1745992758	.18320866	.186283765	.1871722234
30	.174599276026168	.18320864866	.18628376224	
40	.174599276026178	.1832086486832	.1862837622883	
50	.174599276026178	.18320864868347	.18628376228882	
j(I+V)	.174599276026178	.18320864868347	.18628376228882	.1871722233

Table 1. Behavior of Taylor's expansion of the cost function

Order of exp.	triangulation 1	triangulation 2	triangulation 3	triang. 4
1	.073309714253548	.086073617452509	.093782161947154	.0968730219
3	.006718268316071	.013093233335998	.014979462707718	.0149311039
5	.001722788508621	.002539750030623	.004808110366787	.0053616606
10	.000028892586130	.000129843146119	.000248661883022	.0002795846
20	.000000002534305	.000000923092796	.000001335313315	.0000013535
30	.000000000000226	.000000002219957	.000000006721870	.0000000087
40	.000000000000001	.000000000010874	.000000000134758	
50	.000000000000001	.000000000000040	.000000000000581	

Table 2. Behavior of the solution

nodal table is made of the components of the map $I + V$. Thus the derivatives of the map $V \longmapsto j(I + V)$ are exactly the derivatives of j with respect to the nodal table.

We give in Table 2 the relative error for the norm L^∞ between the solution y_{I+V} computed on the modified domain and the approximation of the latter by Taylor's expansion of $y_(I + V)$ at the point I.

The convergence of the series seems to depend only slightly on the size of the elements. On the other hand, when the number of nodes increases, it becomes more and more advantageous to solve some linear systems where the LU decomposition has already been done, than to compute the LU decomposition of the new stiffness matrix at the point $I + V$; this means that the use of higher order derivatives is particularly valuable when solving large scale problems.

This is shown by Table 3; the CPU time spent to compute $j(I)$ (which is also the one spent to compute $j(I + V)$ directly on the new domain) appears in the column $j(I)$, and the additional CPU time spent to compute $j(I + V)$ when using Taylor's expansion of j at the order k appears in the columns Tj(k). We have done those computations on a processor MISP 6000.

Triangulation	j(X)	Tj(1)	Tj(3)	Tj(5)	Tj(10)	Tj(20)	Tj(30)
1	.5	.5	2.5	3.5	6.5	11.5	21.5
2	5	5	9	12	22	60	109
3	46	9	33	50	127	311	527
4	543	68	159	248	756	1908	

Table 3. CPU time (s)

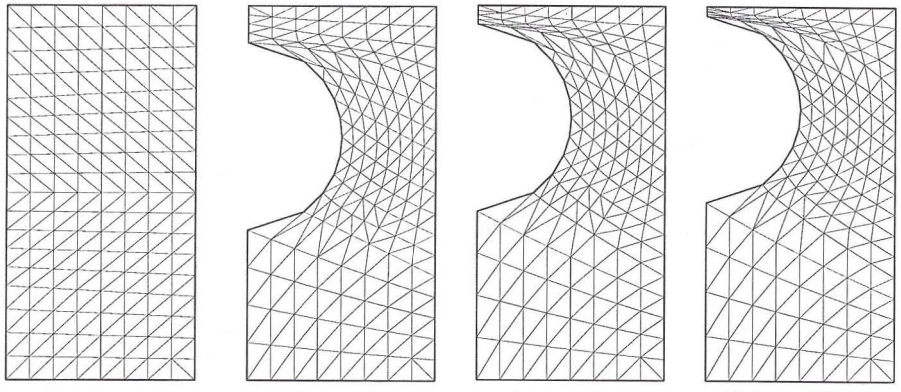
One can see that a Taylor's expansion at the order 3 or 4, which gives a sufficient precision for the engineer, leads to a shorter computation of the result on and after the third triangulation. In the numerical algorithm, we have taken into account the fact that the stiffness matrix is a band matrix, which decreases the ratio

cost of the LU decomposition / cost of solving of the linear system.

When solving a three dimensional problem, this ratio is larger (the band of the matrix is larger), as well as the size of the problem himself. It follows that the use of the higher order derivatives will be more efficient in dimension 3 than in dimension 2.

10.1.2. Other design perturbations

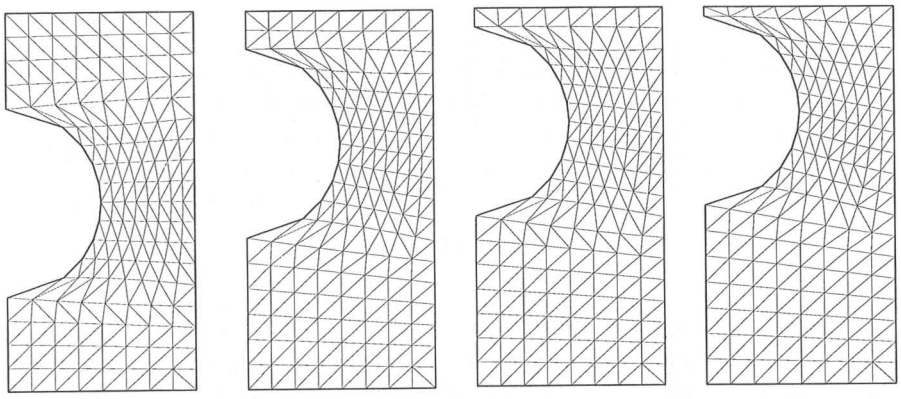
We are now interested in the circle of convergence of Taylor's series, as well as in the domain of validity of the method. Recall that analyticity results have been proved in Destuynder (1976) for a similar problem. The chosen perturbations can damage the triangulation, as shown in Fig. 4.



triangulation 2.1 triangulation 2.3 triangulation 2.4 triangulation 2.5

Figure 4. Damaged triangulations

A correct triangulation of the new domains would be given by Fig. 5.



triangulation 2.1 triangulation 2.3.1 triangulation 2.4.1 triangulation 2.5.1

Figure 5. Correct triangulations

Order of expansion	triangulation 2.3	triangulation 2.4	triangulation 2.5
1	0.17	0.17	0.17
3	0.1951	0.2134	0.225
5	0.19471	0.2122	0.2228
10	0.1945934	0.21151	0.22119
20	0.1945949508	0.21153	0.22082
30	0.1945949526	0.21154	0.2196
50	0.19459494967	0.21153	0.187
100	0.194594949635326	0.21132	-178.7
j(I+V)	0.194594949635325	0.21158	0.22146
jct(I+V)	0.194553646999355	0.21287	0.22313

Table 4. Behavior of Taylor's series of the cost function with a bad triangulation

We give in Table 4 the results of the computing of $j(I+V)$ obtained when using Taylor's expansion of j at the point I for these different perturbations, which have to be compared on the one hand with the direct computation of $j(I+V)$ on those perturbations with bad triangulation (triangulations 2.3, 2.4, 2.5), on the other hand with the direct computation of $j(I+V)$ on the correct triangulation of those perturbations (triangulations 2.3.1, 2.4.1, 2.5.1), denoted by $jct(I+V)$. One can see that the Taylor's series converges for the perturbation 2.3, and is obviously divergent for the large perturbation 2.5; however, one can see that even in this extreme case, one can get a good approximation of $j(I+V)$ by choosing a correct order of Taylor's expansion (here between 5 and 10), and better, a good approximation of $jct(I+V)$.

We give in Table 5 the relative errors e1 and e2 for the norm L^∞:
- e1 is the error between Taylor expansion of $y_(I+V)$ at the point I and the

Order of expansion	0	1	3	5	10	20	30	50	70
e1	0.408	0.102	0.054	0.016	0.016	0.034	0.127	3.990	126.4
e2	0.425	0.105	0.069	0.057	0.061	0.061	0.146	3.955	124.7

Table 5. Behavior of Taylor's series of the solution with a bad triangulation

solution y_{I+V} computed on the triangulation 2.5;
- e2 is the error between Taylor expansion of $y_(I+V)$ at the point I and the solution computed on the correct triangulation 2.5.1, the relative error between the solutions computed on the two triangulations being 0.060. One can see here that even in the case of an important perturbation of the triangulation, the use of the higher order derivatives leads to quite a good approximation (relative error of six per cent), which is often sufficient in practice.

10.2. Numerical tests for Maxwell's equations

10.3. Introduction

The sources of some Spatial Antennas are a network of waveguides. The topology of such a network is obtained using a recent patent (Masmoudi, Brunet, Dusseux and Saury, 1993). The aim of this section is to study an element of this network, i.e. the junction of two rectangular waveguides. The shape of this junction is of most importance in a telecommunication satellite, where room, weight and performance are crucial.

We show that the use of higher order derivatives of the discrete problem (with respect here to the frequency and the shape of the junction) (Guillaume, Masmoudi, 1994) leads to a very performant method for the numerical simulation of the waveguide.

10.4. The 3D problem

Let us roughly describe the physical problem. Two rectangular waveguides G and G' meet together as shown on Fig. 6.

We denote by $\mathcal{G}$ the inside of the complete waveguide. The inner boundary Γ is supposed to be a perfect conductor.

We state the problem as follows: an incident wave is given, which propagates in the waveguide G toward the junction; a part of it is reflected, the other part being transmitted in the waveguides G'. Our goal is to obtain a Taylor expansion of the reflected wave with respect to the frequency and the shape of the junction. This expansion will be used to perform shape optimization, in order to minimize the modulus of the reflected wave on a large frequency scale.

Following usual assumptions are made:
• The electromagnetic field is time-harmonic, i.e. the time dependence occurs through a factor $\exp(i\omega t)$ with $\omega = 2\pi f$ (f is the frequency); more precisely,

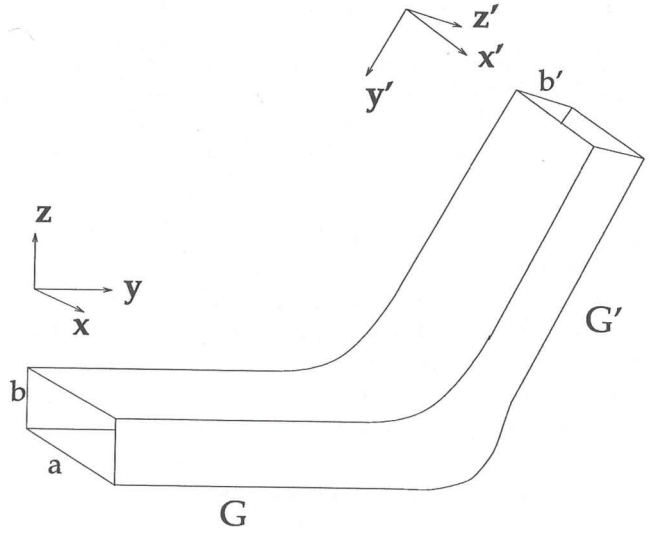


Figure 6. The 3D waveguide

the physical electromagnetic field is the real part of a complex field $E\exp(i\omega t)$, where E is time-independent.

• The incident wave is not modified by the scattered electromagnetic field. This defines in each waveguide G and G' an electric incident field E^i, which vanishes in G', and is expressed in G by

$$E^i = \sin\frac{\pi x}{a}(0,\ 0,\ \mathrm{e}^{-iky})$$

where k is the wave number, defined as (c is the light speed in vacuum):

$$k = \sqrt{\beta^2 - \frac{\pi^2}{a^2}}, \quad \beta = \frac{2\pi f}{c} \tag{23}$$

It is the fundamental mode of the rectangular waveguide (Dautrey and Lions, 1988; Vassallo 1985) and the only one which does not vanish at infinity if we assume that

$$\frac{\pi}{a} < \beta < \min(\frac{\pi}{b},\ \frac{\pi}{b'}). \tag{24}$$

This hypothesis is not really necessary, and is just made to avoid heavy formulation.

The global electric field is a solution to the equations (which derive directly from Maxwell's equations):

$$\text{(P)} \qquad \begin{cases} \operatorname{curl}\operatorname{curl}\mathrm{E} - \beta^2\mathrm{E} = 0 & \text{in } \mathcal{G} \\ E \wedge n = 0 & \text{on } \Gamma \\ \text{(RC) conditions} \end{cases}$$

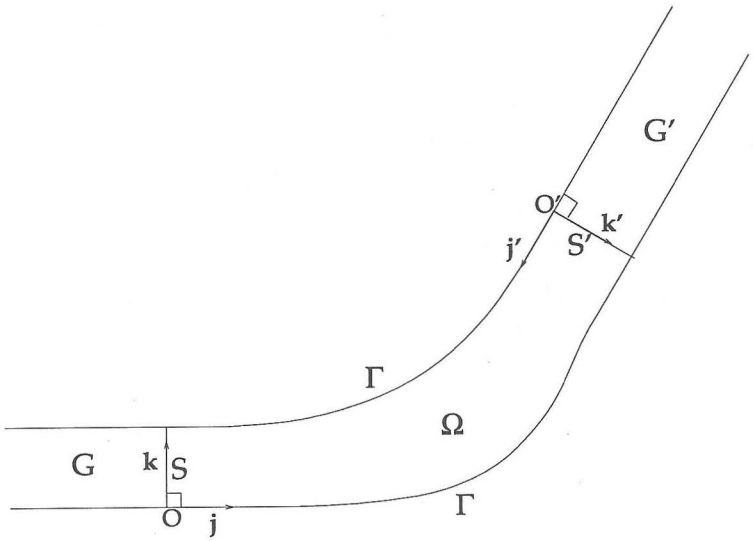


Figure 7. The domain Ω

We denote by n the outward unit normal to Γ, and the condition $E \wedge n = 0$ reflects the fact that the boundary is a perfect conductor. For $e_y = (0, 1, 0)$ in G and $e_{y'} = (0, 1, 0)$ in G', the radiation conditions (RC) (Sommerfeld conditions) are

$$\text{(RC)} \qquad \lim_{y \to -\infty} \operatorname{curl} (\mathrm{E} - \mathrm{E}^{\mathrm{i}}) \wedge \mathrm{e_y} - \mathrm{ik}\,(\mathrm{E} - \mathrm{E}^{\mathrm{i}}) = 0 \quad \text{in G}$$
$$\lim_{y' \to -\infty} \operatorname{curl} \mathrm{E} \wedge \mathrm{e_{y'}} - \mathrm{ik}\,\mathrm{E} = 0 \quad \text{in G}'$$

These equations express that the wave $E - E^i$ is outgoing, and behaves at infinity like the fundamental mode.

The 2D problem

The 2D waveguide (Fig. 7) is the median plane of the 3D waveguide (i.e. the intersection of the 3D waveguide with the plane $x = a/2$). We still denote by $\mathcal{G}$ the inside of the 2D waveguide, and by Γ its boundary. We also denote by Ω a bounded part of the junction delimited by two cross-sections S and S'.

For all scalar field g and all vector-valued field $u = (u_y, u_z)$, let

$$\operatorname{Curl} \mathrm{g} = (\partial_{\mathrm{z}} \mathrm{g},\ -\partial_{\mathrm{y}} \mathrm{g})$$
$$\operatorname{curl} \mathrm{u} = \partial_{\mathrm{y}} \mathrm{u_z} - \partial_{\mathrm{z}} \mathrm{u_y}.$$

and denote by u_t the tangential component of u (i.e. $u = u_n\ n + u_t\ t$ where (n, t) is a direct oriented orthonormal basis).

It can be proved (Guillaume, Masmoudi, 1995) that when the width a of the waveguide is constant, then the 3D problem reduces to a 2D problem. Moreover, the solution of the 2D problem is well approximated by the solution of the following problem:

$$\text{(P)} \qquad \begin{cases} \text{Curl curl u} - \text{k}^2\text{u} = 0 & \text{in } \Omega \\ u_t = 0 & \text{on } \Gamma \\ \text{curl u} + \text{ik u}_\text{t} = \text{curl u}^\text{i} + \text{ik u}_\text{t}^\text{i} & \text{on } S \cup S \end{cases}$$

The incident wave u^i (which is the restriction of E^i to the median plane) vanishes in G', and is expressed in G by

$$u^i = (0,\ \text{e}^{-iky}).$$

We attempt to calculate the reflexion coefficient S_{11} and the transmission coefficient S_{12}. These are given by

$$\begin{aligned} S_{11} &= \frac{1}{b}\int_S (u_t - u_t^i)\ ds \\ S_{12} &= \frac{1}{b'}\int_{S'} u_t\ ds \end{aligned}$$

where u is the solution to problem (P).

10.5. Multi-frequency analysis

We use an H(curl) conforming finite element method (Nédélec, 1980). Hence we have to solve a linear system of equations

$$A(k)X(k) = B(k) \tag{25}$$

Thanks to the form of the approximate problem, the matrix $A(k)$ and the vector $B(k)$ are polynomial in k, that is

$$A(k) = C - k^2 S + ikF$$

(C like curl, S like scalar and F like frontier) and

$$B(k) = kL$$

where C, S, F, L do not depend on k. Observe that higher than second order derivatives of A and B vanish.

Let k_g be a given frequency. The solution to eq. (25) with $k = k_g$ is computed with a direct method, using a LDL^T Cholesky-Crout factorization of the matrix $A(k)$.

In order to compute $X(k)$ for other values of k *without computing a new factorization of the matrix* $A(k)$, we use the Taylor expansion

$$X(k) = X(k_g) + \sum_{j\geq 1} \frac{X^{(j)}(k_g)}{j!}(k - k_g)^j. \tag{26}$$

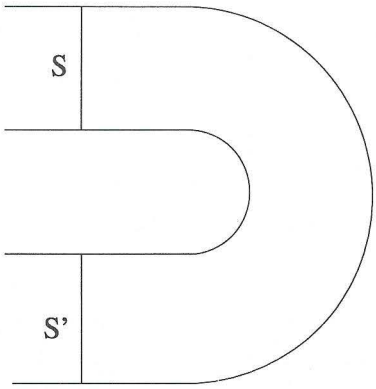


Figure 8. Circular guide

For this purpose, we need to compute the derivatives of $X(k)$ with respect to k. This is done by solving the following systems, which are obtained by taking the successive derivatives of the system (25):

$$A(k)X'(k) = -A'(k)X(k) + L \tag{27}$$

$$A(k)X^{(n)}(k) = -A'(k)X^{(n-1)}(k) + SX^{(n-2)}(k) \quad \forall n \geq 2 \tag{28}$$

Note that the second member of (28) is calculated from the non zero elements of the matrices $A'(k) = -2kS + iF$ and $A''(k) = -2S$ (these are at most five on each line).

The approximation of the coefficients $S_{11}(k)$ and $S_{12}(k)$ becomes

$$S_{11}(k) = \frac{1}{2ikb} B.\overline{X}(k) - 1 = \frac{1}{2ikb} AX(k).\overline{X}(k) - 1$$

$$S_{12}(k) = \frac{1}{2ikb'} B'.\overline{X}(k)$$

where B' is defined in a similar way as B, and $X(k)$ is approximated through (26).

10.6. First example

Our first example is a circular waveguide (see Fig. 8), for which physical measures were supplied.

Computation has been performed on a HP 9000/867S. Time spent is given in Table 6. It can be seen that the multi-frequency analysis (i.e. the computation of the Taylor expansion, here up to the order 50) is about so expensive as a single ordinary analysis.

By classical methods, the knowledge of the reflexion coefficient needs about hundred analyses.

Number of degrees of freedom	CPU time (s) for Cholesky	CPU time (s) for Taylor(50)
799	1	18
3264	20	166
10853	571	678

Table 6. Computation time

	Frequency (GHz)	$\|S_{11}\|$ (dB)	$\|S_{12}\|$ (dB)	phase of S_{11} (deg)	phase of S_{12} (deg)
Taylor(30)	10	-30.918855	-0.248347	-136.899087	18.637913
Taylor(50)	10	-37.476134	-0.001559	-71.579726	18.761669
direct computation	10	-37.496066	0.000773	-71.204879	18.763322
Taylor(30)	15	-24.902031	0.029737	-100.536434	-11.838846
Taylor(50)	15	-33.595172	-0.001504	-99.458462	-9.562543
direct computation	15	-33.657307	-0.001960	-99.606644	-9.553712

Table 7. Convergence of the Taylor expansion on the interval 10-15 GHz

10.6.1. Convergence of the Taylor expansion

Results of convergence of the Taylor expansion are given in Tables 7 and 8, with k_g corresponding to the frequency 12.3 GHz. Comparison is made with the solution computed by solving directly the system (25) for different values of k (10, 11, 14 and 15 Ghz). Of course, the rate of convergence is better on 11-14 Ghz than on 10-15 Ghz.

10.6.2. Comparison with experimental measures

Comparison with the experimental measures (non smooth curves) is given in Figures 9 and 10. Take into account that the reflected wave is very weak, thus computation as well as experimental measure are quite inaccurate (Figs. 9(a) and 10(a)). In contrast, the computation of the transmitted wave is so precise that the result of the computation of the phase (Fig. 10(b)) is exactly the same as the one of the measures (the slight difference in the modulus (Fig. 9(b)) comes from non perfect conduction of the real waveguide).

	Frequency (GHz)	$\|S_{11}\|$ (dB)	$\|S_{12}\|$ (dB)	phase of S_{11} (deg)	phase of S_{12} (deg)
Taylor(30)	11	-37.960780	-0.000695	22.602659	-67.361454
Taylor(50)	11	-37.960780	-0.000695	22.602660	-67.361454
direct computation	11	-37.960780	-0.000695	22.602660	-67.361454
Taylor(30)	14	-32.765963	-0.002320	-26.661061	63.463939
Taylor(50)	14	-32.765955	-0.002320	-26.660971	63.463938
direct computation	14	-32.765955	-0.002320	-26.660971	63.463938

Table 8. Convergence of the Taylor expansion on the interval 11-14 GHz

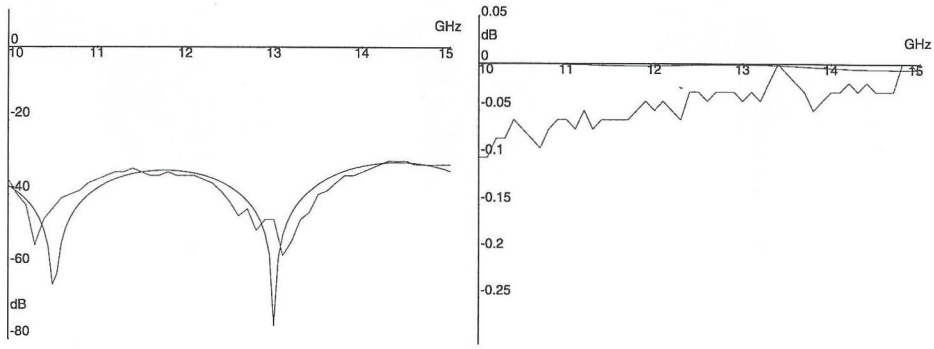


Figure 9. (a) modulus of S11 (b) modulus of S12

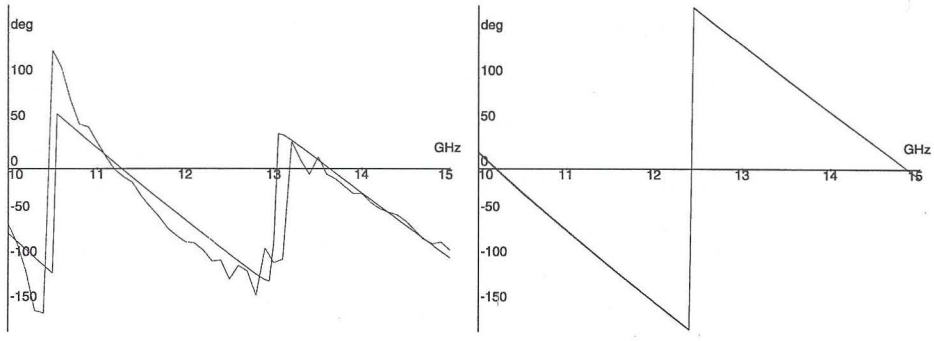


Figure 10. (a) phase of S11 (b) phase of S12

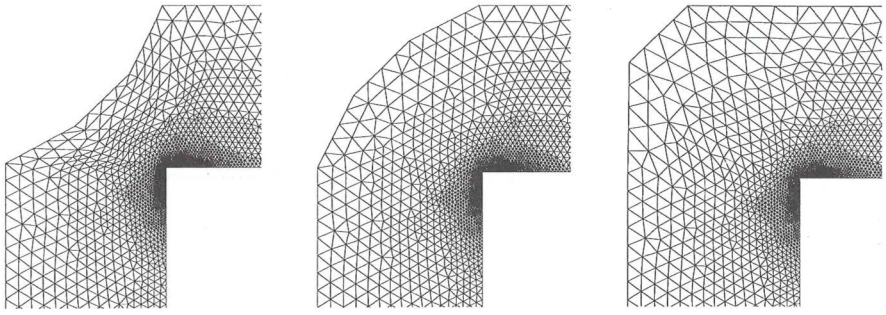


Figure 11. (a) shape at $t = -2$ (b) shape at $t = 0$ (c) shape at $t = 2$

	Frequency (GHz)	$\|S_{11}\|$ (dB)	$\|S_{12}\|$ (dB)	phase of S_{11} (deg)	phase of S_{12} (deg)
Taylor(20,0)	10	-28.36168	0.06291	-5.23238	-48.04583
Taylor(20,3)	10	-11.11887	-0.29565	-135.08447	-44.57097
Taylor(20,5)	10	-11.07522	-0.34964	-135.87184	-44.63012
Taylor(20,10)	10	-11.09239	-0.35196	-135.92671	-44.59784
Taylor(20,20)	10	-11.09178	-0.35187	-135.92933	-44.59680
direct computation	10	-11.09184	-0.35185	-135.92960	-44.59692
Taylor(20,0)	15	-36.93388	-0.00209	13.10028	-117.14689
Taylor(20,3)	15	-6.01872	-1.31288	164.36942	-104.22770
Taylor(20,5)	15	-5.08636	-2.07613	172.37182	-99.06069
Taylor(20,10)	15	-4.83339	-1.75674	169.51543	-98.01031
Taylor(20,20)	15	-4.84270	-1.72714	169.23757	-98.05877
direct computation	15	-4.84289	-1.72714	169.23763	-98.05965

Table 9. Convergence of the Taylor expansion at $t = -2$ and $f = 15GHz$

10.7. Second example

Our second example is shown in Fig. 11. As in previous section, the coefficients S_{11} and S_{12} are computed by using a Taylor expansion, but now with respect to the frequency *and* the shape: the position of the middle facet depends upon a parameter t. The shapes corresponding to the values $t = -2$, $t = 0$ and $t = 2$ are shown respectively in Fig. 11(a), Fig. 11(b) and Fig. 11(c).

The difference here is that the dependence of the matrix $A(k,t)$ is no more polynomial in t. Thus the successive derivatives of $t \longmapsto A(k,t)$ are computed by using automatic differentiation on the elementary matrices Morgenstern (1985), Griewank (1989). A polynomial in two variables $P(k,t)$ is obtained. The convergence of the Taylor expansion is given in Table 9. Taylor(p,q) indicates a derivation at order p with respect to the frequency, and at order q with respect to the shape.

The graph of the map $(f,t) \longmapsto |P(k(f),t)|$ is shown in Fig. 12 and Fig. 13. It is worth to note that a simple view on Fig. 13 gives the value of t for the solution to the non differentiable problem:

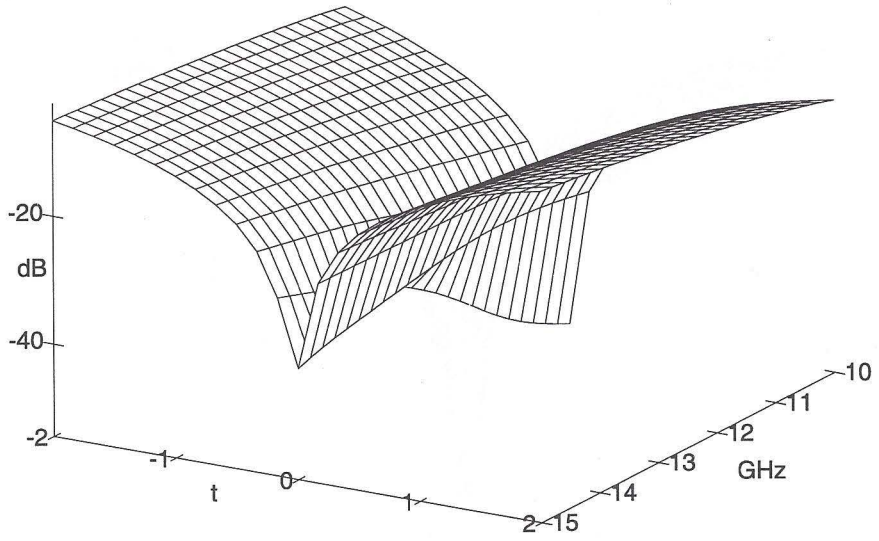


Figure 12. Graph of $(f,t) \longmapsto |S_{11}(f,t)|$

minimize $j(t)$ with

$$j(t) = \sup_{10\,GHz \le f \le 15\,GHz} |S_{11}(f,t)|$$

The curve $t \longmapsto j(t)$ is the superior envelope of all the curves, and the optimal t is near zero (which corresponds to the initial design!).

11. Conclusion

Those numerical results show clearly the efficiency of the higher order derivatives method and automatic differentiation.

It is necessary, in order to introduce shape optimization methods into industry, to lower their cost of implementation. There is, at least, one mean to achieve this objective: the use of automatic differentiation to compute first and higher order derivatives.

The communication between CAD (Computer Aided Design) and computing environments needs a lot of human time; the *high order derivatives* method reduces the number of analyses and the number of conversions of CAD models to finite element models.

The designer can obtain in real time the solution of the modified domain (polynomial evaluation), and uses at each moment his own know-how for updating parameters.

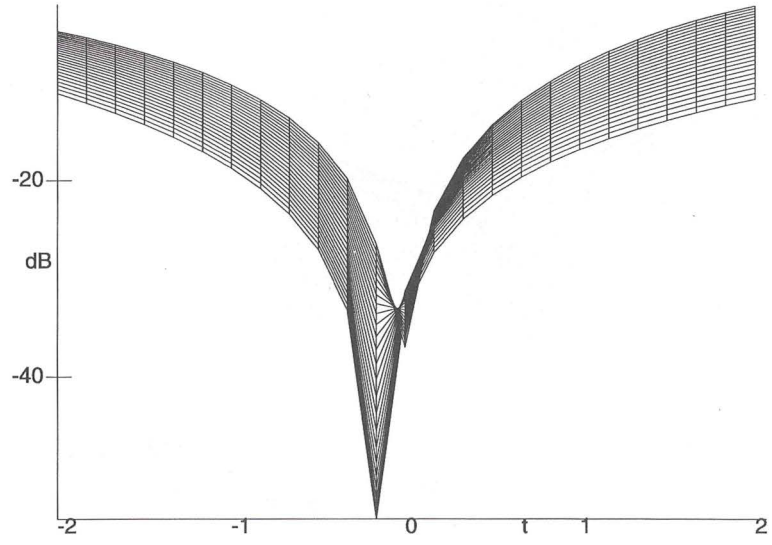


Figure 13. Projection of the graph on the plane $f = 10\,GHz$

In conclusion, using the suggested methods, the designer can obtain a satisfactory design in a short time.

References

Brezis, H. (1987) *Analyse Fonctionnelle.* Masson.

Cea, J. (1986) Conception optimale ou identification de formes, calcul rapide de la dérivée directionnelle de la fonction coût. *M.A.A.N*, **20**, 3.

Ciarlet, G. (1978) *The Finite Element Method for Elliptic Problems.* North-Holland.

Dautrey, R. and Lions, J.L. (1988) *Analyse Mathématique et Calcul Numérique pour les Sciences et les Techniques.* Masson, Paris.

Destuynder, P. (1976) Thèse de troisième cycle. Université PARIS VI.

Fujii, N. (1986) Necessary conditions for a domain optimization problem in elliptic boundary value problems, *SIAM. J. Control and Optimisation*, **24**, 346-360.

Grisvard, P. (1985) *Elliptic Problems in Nonsmooth Domains.* Pitman.

Griewank, A. (1989) On automatic differentiation. In: M. Iri and K. Tanabe (editors), *Mathematical Programming: Recent Developments and Applications*, 83-108, Kluwer Academic Publishers, Dordrecht.

Guillaume, P. (1994) *Dérivées d'ordre supérieur en conception optimale de forme.* Thèse de doctorat, Université Paul Sabatier, Toulouse, France.

Guillaume, Ph., Masmoudi, M. (1992) Dérivées d'ordre supérieur en optimisation de domaines. *C.R. Acad. Sci. Paris*, **315**, Série I, 859-862.

GUILLAUME, PH., MASMOUDI, M. (1993) Calcul numérique des dérivées d'ordre supérieur en conception optimale de forme. *C.R. Acad. Sci. Paris*, **316**, série 1, 1091-1096.

GUILLAUME, PH., MASMOUDI, M. (1994) Computation of high order derivatives in optimal shape design. *Numerische Mathematik*, **67**, 231-250.

GUILLAUME, PH., MASMOUDI, M. (1995) Solution to the time-harmonic Maxwell's equations in a waveguide, use of higher order derivatives for solving the discrete problem, to appear in *SIAM Journal on Numerical Analysis.*

KENNAN,?., SMITH, T. (1983) *Primer of Modern Analysis.* Springer Verlag.

MASANAO, T., FUJII, N. (1992) Second order necessary conditions for domain optimization problems in elastic structures. *Journal of Optimization Theory and Applications*, **72**, 2, February.

MASMOUDI, M. (1987) *Outils pour la conception optimale de formes.* Thèse d'état, Nice.

MASMOUDI, M., BRUNET, P., DUSSEUX, T. and SAURY, M. (1993) *Décomposition d'un répartiteur orthogonal en coupleurs élémentaires.* Patent with Alcatel Espace No 19711.

MILLO, F. (1991) *Conception optimale de structures rayonnantes.* Thèse, Nice-Sophia Antipolis.

MORGENSTERN, J. (1985) How to compute fast a function and all its derivatives, a variation on the theorem of Baur-Strassen. *SIGACT News.*

MURAT, F. and SIMON, J. (1976) *Sur le contrôle par un domaine géométrique.* Thèse d'état, Paris VI.

NECAS, J. (1967) *Les Méthodes Directes en Théorie des Équations Elliptiques.* Masson.

NÉDÉLEC, J.C. (1980) Mixed finite elements in $\mathbb{R}^3$. *Numerische Mathematik*, **35**, 315-341.

RAVIART, P.A. and THOMAS, J.M. (1983) *Introduction à l'Analyse des Équations aux Dérivées Partielles.* Masson.

ROUSSELET, B. (1987) Shape design sensitivity from partial differential equation to implementation. *Eng. Opt.*, **11**, 151-171.

SIMON, J. (1980) Differentiation with respect to the domain in boundary value problems. *Numerical Functional Analysis and Optimization.* **2**, 7-8, 649-687.

SIMON, J. (1989) Second variation for domain optimization problems. *International Series of Numerical Mathematics*, **91**, Birkhauser.

STRASSEN, V. (1990) Algebraic complexity theory. In: J. van Leeuwen (editor), *Handbook of Theoritical Computer Science, volume A: Algorithms and complexity,* Elsevier, Amsterdam.

VASSALLO, CH. (1985) *Théorie des guides d'ondes électromagnétiques*, Collection Technique et Scientifique des Telécommunications, Eyrolles, France.