Hybrid methods for inverse scattering problems

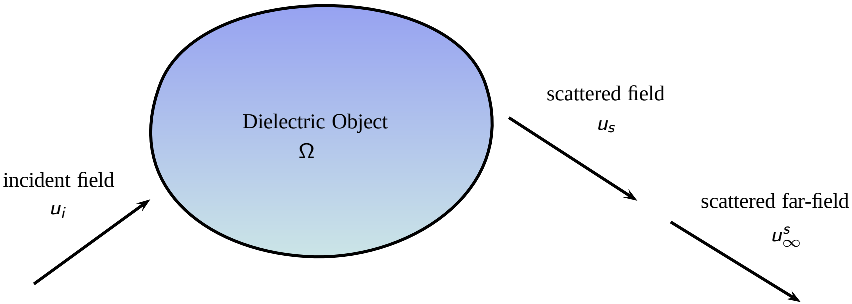
Coupling linear sampling, gradient, topological gradient and quasi-newton methods

Dimitri Nicolas
joint work w/ G. Allaire and H. Haddar

MMSN 2012
CMAP, Ecole Polytechnique

January, 9th 2012

Goal : Determining the shape of an object from measurements of the electromagnetic far-field it scattered



Direct scattering problem : find u_s knowing u_i and Ω
Inverse scattering problem : find Ω knowing u_s (u^s_∞ from far-field point of view)

Summary

Two main approaches

Two traditionnaly approaches to solve those kind of problems :

1. Sampling methods : find the solution of a linear Fredholm equation of the first kind
2. Non-linear optimization schemes : reconstruction is performed iteratively from an initial guess

Sampling methods

Sampling methods such as the linear sampling method (LSM) :

- No need to solve the direct nor adjoint scattering problem
- Quick solve
- Very little a priori information is needed (nor the number of scatterers, nor if the object is penetrable or not, nor which kind of boundary conditions is satisfied by the total field on the boundary)
- Lack of precision
- Only provides a reconstruction of the support of the scatterer (no way to find the point value of the index of refraction in the case of inhomogenious scattering)

Iterative methods

Iterative methods such as the gradient method (with level-set framework) :

- Can reach a good accuracy
- Slow solve
- Local convergence, so need a "not too bad" first approximation
- Need a priori information on the scatterer
- Need to solve forward and adjoint problems

Topological gradient method

Topological gradient method :

- Can create and destroy inclusions : topology changes naturaly handled during the process of optimization
- Can be used iteratively or not
- Need a priori information on the scatterer as the gradient method.
- Need to solve forward and adjoint problems

Second order shape derivative method

Second derivative of the shape functionnal (iterative method too) :

- Can reach a better accuracy than first order or accelerate the convergence rate
- Slower solving than iterative method : need to solve more forward and adjoint problems
- Local convergence too
- Need a priori information on the scatterer too
- How to approximate the best way the second order shape derivative.

Idea : Coupling between the linear sampling method, the gradient method with the level-set formalism and the topological gradient method and/or second order schemes.

Idea : Coupling between the linear sampling method, the gradient method with the level-set formalism and the topological gradient method and/or second order schemes.

- The linear sampling method used to find quickly an initial guess whose accuracy depends on a cutoff parameter.

Idea : Coupling between the linear sampling method, the gradient method with the level-set formalism and the topological gradient method and/or second order schemes.

- The linear sampling method used to find quickly an initial guess whose accuracy depends on a cutoff parameter.
- The level-set method used to go closer to the true shape (if close enough already...)

Idea : Coupling between the linear sampling method, the gradient method with the level-set formalism and the topological gradient method and/or second order schemes.

- The linear sampling method used to find quickly an initial guess whose accuracy depends on a cutoff parameter.
- The level-set method used to go closer to the true shape (if close enough already...)
- Whenever the use is needed (by a posteriori information first), use of the topological gradient method to get other connected component and to accelerate the convergence.

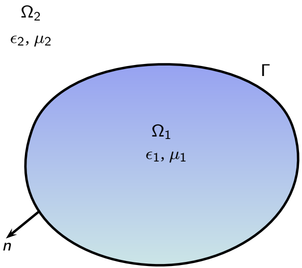
Idea : Coupling between the linear sampling method, the gradient method with the level-set formalism and the topological gradient method and/or second order schemes.

- The linear sampling method used to find quickly an initial guess whose accuracy depends on a cutoff parameter.
- The level-set method used to go closer to the true shape (if close enough already...)
- Whenever the use is needed (by a posteriori information first), use of the topological gradient method to get other connected component and to accelerate the convergence.
- Use of second order approximation to get a better accuracy.

Helmholtz in the dielectric case

We look first for the scattered plane wave $u \in H^2_{loc}(\mathbb{R}^2)$ such as :

$$\begin{cases} \nabla.(\frac{\nabla u}{\mu}) + k^2 \epsilon u = 0 & in \quad \mathbb{R}^2 \\ u = u^i + u^s & in \quad \mathbb{R}^2 \\ \lim_{R \to \infty} \int_{S_R} |\partial_r u^s - iku^s|^2 ds = 0 & \end{cases}$$



The minimization problem

We can write u_s as :

$$u_s = \frac{e^{ik|x|}}{|x|^{\frac{1}{2}}}(u^s_\infty(\hat{x}) + O(\frac{1}{|x|}))$$

with $\hat{x} = \frac{x}{|x|} \in S^1$ and u^s_∞ the far-field up to a constant.

For an incident plane wave $u^i(x) = e^{ikx.d}$, we measure in $d \in S^1$ directions

- $u^{s,mes}_\infty(\hat{x}, d)$ the far field computed for the true shape
- $u^s_\infty(\Gamma)(\hat{x}, d)$ the far field computed for the current iteration shape

The problem :

Find Γ^{min} which minimize the following functionnal :

$$\mathcal{J}(\Gamma) := \frac{1}{2}\|u_\infty(\Gamma) - u^{mes}_\infty\|^2_{L^2(S^1\times S^1)}$$

The Linear Sampling method

The far field pattern u^s_∞ defines the far-field operator $F : L^2(S^1) \to L^2(S^1)$ by

$$(Fg)(\hat{x}) := \int_{S^1} u^s_\infty(\hat{x}, d) g(d) ds(d)$$

The Linear Sampling method is to find $g = g(., z) \in L^2(S^1)$ of :

$$Fg_z(\hat{x}) = \phi_\infty(\hat{x}, z)$$

where $z \in \mathbb{R}^2$ and $\phi_\infty(., z)$ is the far-field pattern of the fundamental solution $\phi(., z)$ of the Helmholtz equation.

[Colton D., Kress R., Inverse acoustic and electromagnetic scattering theory, Second edition, Applied Mathematical Science, 93. Springer-Verlag, Berlin, 1998]

The Linear Sampling method

The far field pattern u^s_∞ defines the far-field operator $F : L^2(S^1) \to L^2(S^1)$ by

$$(Fg)(\hat{x}) := \int_{S^1} u^s_\infty(\hat{x}, d) g(d) ds(d)$$

The Linear Sampling method is to find $g = g(., z) \in L^2(S^1)$ of :

$$Fg_z(\hat{x}) = \phi_\infty(\hat{x}, z)$$

where $z \in \mathbb{R}^2$ and $\phi_\infty(., z)$ is the far-field pattern of the fundamental solution $\phi(., z)$ of the Helmholtz equation.

- The more g_z **is high**, the more there is a chance $z \in \Omega$.

[Colton D., Kress R., Inverse acoustic and electromagnetic scattering theory, Second edition, Applied Mathematical Science, 93. Springer-Verlag, Berlin, 1998]

The Linear Sampling method

The far field pattern u^s_∞ defines the far-field operator $F : L^2(S^1) \to L^2(S^1)$ by

$$(Fg)(\hat{x}) := \int_{S^1} u^s_\infty(\hat{x}, d) g(d) ds(d)$$

The Linear Sampling method is to find $g = g(., z) \in L^2(S^1)$ of :

$$Fg_z(\hat{x}) = \phi_\infty(\hat{x}, z)$$

where $z \in \mathbb{R}^2$ and $\phi_\infty(., z)$ is the far-field pattern of the fundamental solution $\phi(., z)$ of the Helmholtz equation.

- The more g_z **is high**, the more there is a chance $z \in \Omega$.
- Need of a **"cutoff"** value.

[Colton D., Kress R., Inverse acoustic and electromagnetic scattering theory, Second edition, Applied Mathematical Science, 93. Springer-Verlag, Berlin, 1998]

The Gradient method

The general form of a shape derivative in the direction θ is

$$\mathcal{J}'(\Omega)(\theta) = \int_{\partial\Omega} \nu\theta.n ds$$

where the function ν defines the descent direction θ with :

$$\theta = -\nu n$$

Update of the shape with :

$$\Omega_{n+1} = (Id + t\theta)(\Omega_n)$$

where $t > 0$ is a small descent step. Formally :

$$\mathcal{J}(\Omega_{n+1}) = \mathcal{J}(\Omega_n) - t \int_{\partial\Omega} \nu^2 ds + O(t^2)$$

The Gradient method : Level-Set framework

Shape Ω is **characterized by the zero level-set** of a function ϕ.

The level-set method will move the shape Ω (hence ϕ) by solving the following **Hamilton-Jacobi** advection equation

$$\frac{\partial \phi}{\partial t} + V \mid \nabla \phi \mid = 0$$

The Gradient method : Level-Set framework

Shape Ω is **characterized by the zero level-set** of a function ϕ.

The level-set method will move the shape Ω (hence ϕ) by solving the following **Hamilton-Jacobi** advection equation

$$\frac{\partial \phi}{\partial t} + V \mid \nabla \phi \mid = 0$$

1. Find the descent direction V by computing the **shape derivative** of the functionnal we want to minimize : advection velocity $V = -\nu$

The Gradient method : Level-Set framework

Shape Ω is **characterized by the zero level-set** of a function ϕ.

The level-set method will move the shape Ω (hence ϕ) by solving the following **Hamilton-Jacobi** advection equation

$$\frac{\partial \phi}{\partial t} + V \mid \nabla \phi \mid = 0$$

1. Find the descent direction V by computing the **shape derivative** of the functionnal we want to minimize : advection velocity $V = -\nu$
2. Solve the H-J equation to get the new moved shape characterized by $\phi = 0$.

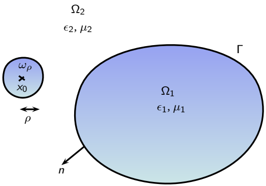
The Topological Gradient method

Look at each point x_0 of the mesh **if it is fruitful to create a small infinitesimal inclusion**.

Two problems : the non-perturbed problem and the one perturbed by the add of a ball of size ρ at the point x_0.

Definition

- χ the characteristic function of Ω_1
- χ_{ω_ρ} the characteristic function of ω_ρ
- $\chi_\rho = \chi + \chi_{\omega_\rho}$



The Topological Gradient method

Definition

The topological gradient is defined by the Taylor expansion :

$$\mathcal{J}(\chi_\rho) = \mathcal{J}(\chi) + \rho^d DJ(x_0) + o(\rho^d)$$

with the cost function $\mathcal{J}$ defined before.

We then get a map $x_0 \to DJ(x_0)$. The more $DJ(x_0)$ is negative, the more we should create a small inclusion at the point x_0.

Need of a "cutoff" value also

[Sokolowski, A. Zochowski, On the topological derivative in shape optimization, SIAM J. Control Optim., 37, pp.1251-1272 (1999)]
[Céa J., Garreau S., Guillaume P., Masmoudi M., The shape and topological optimizations connection, Compute. Methods Appl. Mech. Engrg 188, 713-726 (2000)]

Computation of DJ

We get the topological gradient $DJ(x_0)$ (For each point x_0 of the underlying grid) for an inclusion of medium n°1 (μ_1, ϵ_1) into a matrix of medium n°2 (μ_2, ϵ_2) :

Topological Gradient expression

$$DJ(x_0) = \Re\{\frac{-2\pi(\mu_2 - \mu_1)}{\mu_2(\mu_1 + \mu_2)}\nabla\overline{u}_\chi(x_0).\nabla p_\chi(x_0) - \pi k^2(\epsilon_2 - \epsilon_1)\overline{u}_\chi(x_0)p_\chi(x_0)\}$$

with u_χ and p_χ solutions of the forward (2) and adjoint (3) problems respectively.

Parameters

- Implementation in scilab and Fortran.
- Simulations over a uniform 60x40 grid
- 10 incident plane waves and 10 measures over 360°
- 1% noise
- Wavelenght $\lambda = 1$

A way to handle topology changes

Big square is 0.5 wide and little square 0.3 wide. Both separated by 0.3

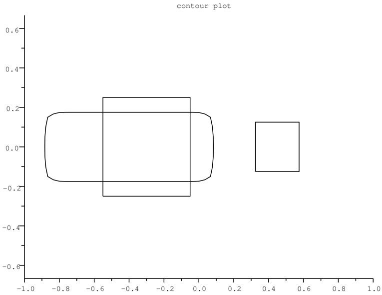


FIGURE: Initialisation of a square with a rectangle (little square missed on purpose)

A way to handle topology changes

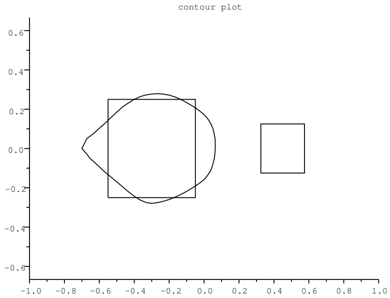


FIGURE: 4th iteration with the Level-Set method

A way to handle topology changes

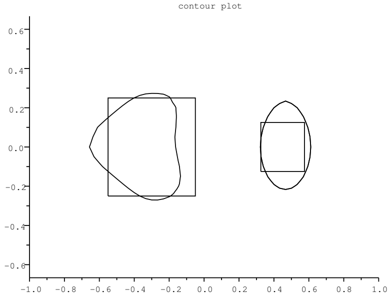


FIGURE: 5th iteration with the Topological Gradient method

A way to handle topology changes

Topological gradient succeeded in finding the little square with

- very few measurements
- a wavelenght 3 times larger than the size of the little square

A way to improve the speed of convergence

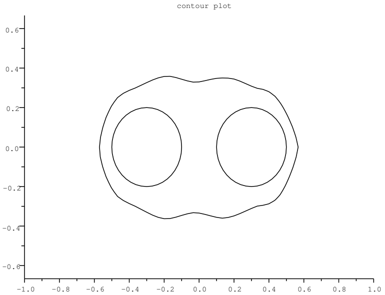


FIGURE: Initialisation with the Linear Sampling method

A way to improve the speed of convergence

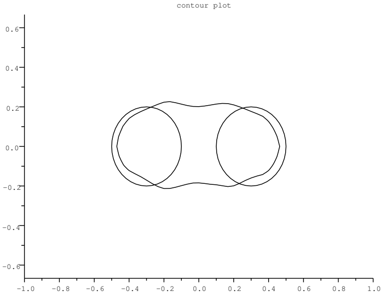


FIGURE: 2nd iteration with the Level-Set method

A way to improve the speed of convergence

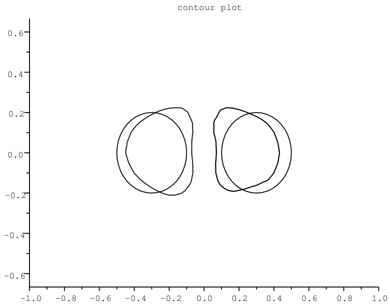


FIGURE: 3rd iteration with the Topological Gradient method

A way to improve the speed of convergence

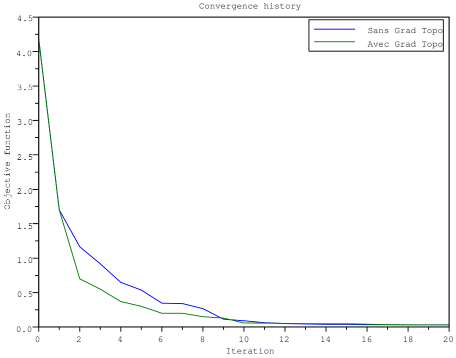


FIGURE: Convergence history with and without the use of the Topological Gradient

The second order derivative

Let $\Omega_{n+1} = (I + \tilde{\theta})(\Omega_n)$. The Taylor series of $\mathcal{J}'(\Omega_{n+1})(\theta)$ yields for each θ

$$\mathcal{J}'(\Omega_{n+1})(\theta) = \mathcal{J}'(\Omega_n)(\theta) + (\mathcal{J}')'(\Omega_n)(\theta, \tilde{\theta}) + \ldots$$

So we got to find $\tilde{\theta}$ by solving

$$-\mathcal{J}'(\Omega_n)(\theta) = (\mathcal{J}')'(\Omega_n)(\theta, \tilde{\theta}) \qquad \forall\theta$$

But in this above formula, $(\mathcal{J}')'$ is complicated. We got, for θ fixed, to solve a PDE for each $\tilde{\theta}$ we set. So we use a BFGS approximation in finite dimension to simplify and see what happens.

BFGS approximation

A way to find this $\tilde{\theta}$ is to use a Quasi-Newton method with an approximation of the inverse of the so-called Hessian matrix $(\mathcal{J}')'$.

We use the following BFGS approximation (in finite dimension) :

BFGS : approximation of the inverse of the Hessian

$$H_{k+1}^{-1} = (I - \frac{d_k \delta_k^T}{d_k^T \delta_k})^T H_k^{-1} (I - \frac{d_k \delta_k^T}{d_k^T \delta_k}) + \frac{\delta_k \delta_k^T}{d_k^T \delta_k}$$

with $d_k = \nabla \mathcal{J}(x_{k+1}) - \nabla \mathcal{J}(x_k)$ and $\delta_k = x_{k+1} - x_k$. We parametrize $\Gamma_k = \partial \Omega_k$ so that $\forall k$, the number of points of Γ_k be the same. Hence we can approximate $\mathcal{J}'(\Gamma_k)$ by $\nabla \mathcal{J}(x_k)$, $\forall k$ such that $x_k \in \Gamma_k$.
Then,

$$\tilde{\theta}_k = -H_k^{-1} \nabla \mathcal{J}(x_k)$$

Better accuracy and speed of convergence

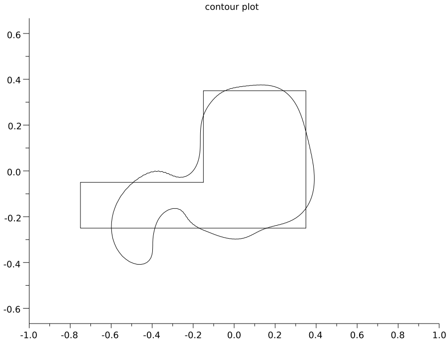


FIGURE: 15th iteration with a usual first order method

Better accuracy and speed of convergence

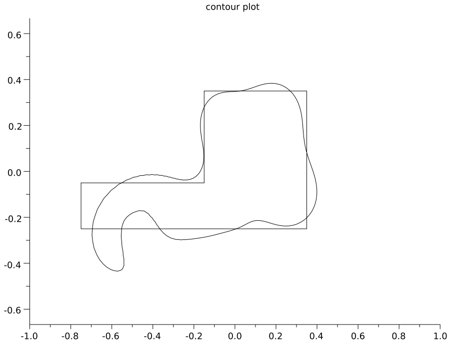


FIGURE: 15th iteration with BFGS method

Better accuracy and speed of convergence

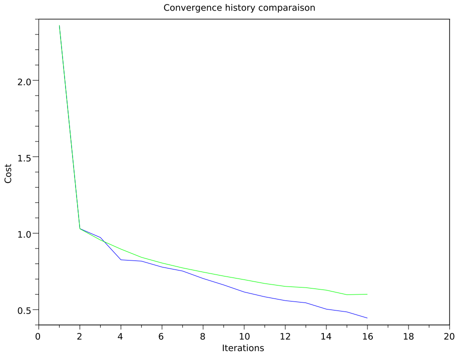


FIGURE: Blue : Second order derivative, green : normal method

Conclusions

- The Linear Sampling method allow to get a first approximation. The Level-Set method, which converges only locally, can get us closer to the true shape.
- The Topological Gradient allow us to get other connected components we eventually missed from the beginning.
- The topological Gradient can improve the speed of convergence of the algorithm if we use it at the right moment.
- Second order schemes is only slightly better than first order method on first tests.

$\Rightarrow$ Still the question to know when to use the topological gradient effciently.

$\Rightarrow$ Better implementation of the topological gradient method for faster computation.

$\Rightarrow$ Find the true second order shape derivative and use it in an iterative method (approximations will be needed).

Some references

- H. Ammari, "An introduction to mathematics of emerging biomedical imaging", Collection : Mathématiques et Applications, Vol. 62, Springer (2008).
- M. Bonnet, "Inverse acoustic scattering by small-obstacle expansion of misfit function". Inverse Problems, 24 :035022 (2008)
- F. Cakoni, D. Colton, "Qualitative Methods in Inverse Scattering Theory". 2006. Springer, Berlin.
- M. Bonnet, B. B. Guzina, "Elastic-wave identification of penetrable obstacles using shape-material sensitivity framework",J. Comput. Phys. 228 (2009) 294–311.
- G. Allaire, F. Jouve, A-M. Toader, "Structural optimization using sensisivity analysis and a level-set method", J. Comp. Phys. 194 (2004) 363-393.
- J. Sokolowski, A. Zochowski, On the topological derivative in shape optimization, SIAM J. Control Optim., 37, pp.1251-1272 (1999)
- S. Garreau, P. Guillaume, M. Masmoudi, The topological asymptotic for PDE systems : the elasticity case. SIAM J. Control Optim. 39, no. 6, pp.1756-1778 (2001)

Questions ?

Thank you !

A Lagrangian method

Use of the following lagrangian to determine the adjoint

$$\mathcal{L}(s,t) = \frac{1}{2} \parallel s_\infty(s) - s_\infty^{mes} \parallel^2_{L^2(S^1 \times S^1)} + \Re\{\int_{B_R} [-\frac{1}{\mu}\nabla s \nabla \bar{t} + k^2 \epsilon s \bar{t}] - \int_{S_R} [T_R(s) - g]\bar{t}\}$$

that leads to the forward and adjoint problem (2) and (3). By taking $v = u_{\chi_\rho} - u_\chi$ and $q = p_{\chi_\rho} - p_\chi$, where u_{χ_ρ} and p_{χ_ρ} are the respective solution of (4) and (5), v is then solution of the following problem :

$$\begin{cases} \nabla.(\frac{\nabla v}{\mu_{\chi_\rho}}) + k^2 \epsilon_{\chi_\rho} v = \nabla.(\chi_{\omega_\rho}[\frac{1}{\mu}]\nabla u_\chi) + k^2[\epsilon]\chi_{\omega_\rho} u_\chi \quad in \quad B_R \\ v = u^s_{\chi_\rho} - u^s_\chi \quad in \quad B_R \\ \frac{1}{\mu}\frac{\partial v}{\partial n} + T_R(v) = 0 \quad on \quad S_R \end{cases} \tag{1}$$

q is solution of the adjoint problem of this above problem (not written here)

Forward and adjoint problems

$$\begin{cases} \nabla.(\frac{\nabla u_\chi}{\mu_\chi}) + k^2 \epsilon_\chi u_\chi = 0 & in \quad B_R \\ u_\chi = u^i + u^s_\chi & in \quad B_R \\ \frac{1}{\mu}\frac{\partial u_\chi}{\partial n} + T_R(u_\chi) = g & on \quad S_R \end{cases} \tag{2}$$

$$\begin{cases} \nabla.(\frac{\nabla p_\chi}{\mu_\chi}) + k^2 \epsilon_\chi p_\chi = 0 & in \quad B_R \\ p_\chi = p^i + p^s_\chi & in \quad B_R \\ \frac{1}{\mu}\frac{\partial p_\chi}{\partial n} + \overline{T}_R(\overline{p_\chi}) = S^\chi_\infty & on \quad S_R \end{cases} \tag{3}$$

Perturbed forward and adjoint problems

$$\begin{cases} \nabla.(\frac{\nabla u_{\chi_\rho}}{\mu_{\chi_\rho}}) + k^2 \epsilon_{\chi_\rho} u_{\chi_\rho} = 0 & in \quad B_R \\ u_{\chi_\rho} = u^i + u^s_{\chi_\rho} & in \quad B_R \\ \frac{1}{\mu} \frac{\partial u_{\chi_\rho}}{\partial n} + T_R(u_{\chi_\rho}) = g & on \quad S_R \end{cases} \tag{4}$$

$$\begin{cases} \nabla.(\frac{\nabla p_{\chi_\rho}}{\mu_{\chi_\rho}}) + k^2 \epsilon_{\chi_\rho} p_{\chi_\rho} = 0 & in \quad B_R \\ p_{\chi_\rho} = p^i + p^s_{\chi_\rho} & in \quad B_R \\ \frac{1}{\mu} \frac{\partial p_{\chi_\rho}}{\partial n} + \overline{T}_R(\overline{p_{\chi_\rho}}) = S^{\chi}_{\infty} & on \quad S_R \end{cases} \tag{5}$$

with

$$S^{\chi}_{\infty} = \mu \overline{T}_R(\overline{\mathcal{H}}(u_\infty(\chi) - u^{mes}_\infty)) + \frac{\partial \overline{\mathcal{H}}}{\partial n}(u_\infty(\chi) - u^{mes}_\infty)$$

Difference between $(\mathcal{J}')'$ and $\mathcal{J}''$

If $\mathcal{J}(x)$ and x lives in a linear space then we have $(\mathcal{J}')' = \mathcal{J}''$. But we have $\mathcal{J}(\Omega)$ and Ω is not a linear space. We use the parameter θ to define the variations of Ω. And even if θ lives in a linear space, we have the following :

$$(\Omega + \theta) + \tilde{\theta} \neq \Omega + (\theta + \tilde{\theta})$$

Indeed by definition $\Omega + \theta = (I + \theta)(\Omega)$, so

$$(\Omega+\theta)+\tilde{\theta} = (I+\tilde{\theta})\circ(I+\theta)(\Omega) = (I+\theta+\tilde{\theta}\circ(I+\theta))(\Omega) = \Omega+(\theta+\tilde{\theta}\circ(I+\theta)(\Omega))$$

We got the following relation between $(\mathcal{J}')'$ and $\mathcal{J}''$:

$$\mathcal{J}''(\Omega,\theta,\tilde{\theta}) = (\mathcal{J}')'(\Omega,\theta,\tilde{\theta}) - \mathcal{J}'(\Omega,\theta.\nabla\theta)$$

[J. Simon. Second variation for domain optimization problems. In Control and estimation of distributed parameter systems,Birkhäuser (1989), 361-378]

Plot of exterior normals

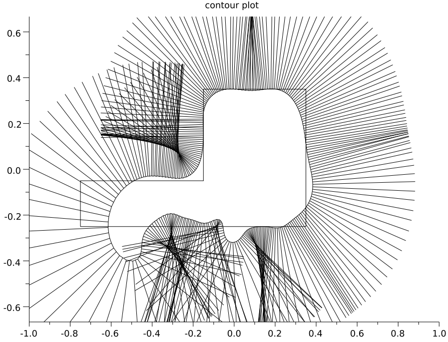


FIGURE: 15th iteration with 2nd order method and plot of exterior normals