Some applications of the adjoint variable method in electromagnetic optimization and inverse problems

Ivan Cimrák*
NfaM^2 Research Group, Dep. Mathematical Analysis, Ghent University
Galglaan 2, 9000 Ghent, Belgium, Tel:+32 9 2644920, Fax:+32 9 2644987

Main area of the submission: Optimization
Type of the submission: Oral presentation

Abstract

In this paper we present two applications of the adjoint variable method (AVM). First we consider a design optimization problem in magnetic shielding. The objective is to reduce the magnetic stray field of an axisymmetric induction heating device for the heat treatment of aluminum discs. We involve two types of shielding, the passive and the active shielding. In the former, one needs to optimize the geometry of the passive shield. In the latter, the position of all coils and the real and imaginary components of the currents (when working in the frequency domain) must be determined.

Second application involves determination of the dissipation parameter in micromagnetic model of ferromagnetism. The micromagnetic model governed by the Landau-Lifshitz equation includes the dissipation parameter α that in some cases can be a space dependent function. The actual distribution of α however can be unknown and must be determined by measurements of the magnetization in the workpiece.

Using AVM method, one obtains the derivative of cost functional in terms of an adjoint variable. The main advantage is that the number of direct problem simulations needed to evaluate the derivative is independent of the number of parameters.

1 Introduction

From the point of view of accuracy and time-efficiency in finding the optimum solution in design space, the design sensitivity analysis (DSA) appears to be very competitive compared with other optimization methods. We can distinguish between two types according to the technique used to compute the derivative of an objective function [1]: the discrete DSA, where the gradient information is obtained from direct differentiation of the discretized algebraic system matrix, and the continuum DSA, where an analytically derived sensitivity formula is used for the gradient information. This formula uses the adjoint variable. We exploit both methods.

Direct differentiation method. Derivatives of the state variables with respect to the design variables are predicted by introducing small variation of the design variable, re-evaluating the cost and approximating the derivative by dividing the subtraction of both cost values by the perturbation step. Then design variables are properly updated and process simulation is performed again. The procedure is repeated until an optimal design is achieved. The method is ideal for dealing with a general class of optimal design problems.

Adjoint variable method. The method differs from the direct differentiation method in that it calculates the design sensitivity by introducing adjoint variables, to avoid calculating the derivatives of the state variables with respect to the design variables. The method becomes computationally more efficient than the direct differentiation method as the number of design variables is increased.

*The author is supported by the FWO Fund for Scientific Research - Flanders (Belgium).

2 Simple example of AVM

In order to understand the meaning of the adjoint system, we briefly describe the key ingredients of the derivation and of the use of such an adjoint system. We demonstrate the use of AVM on a simple example:

Example 1. *For an insulated metal rod with conductivity coefficient k consider the following Neumann problem of the heat conduction*

$$\left.\begin{array}{ll} A_t - kA_{xx} = f, & for\ (x,t) \in \Omega_T := (0,1) \times (0,10) \\ A_x(0,t) = A_x(1,t) = 0, & for\ t \in I := (0,10) \\ A(x,0) = A_0, & for\ x \in \Omega := (0,1), \end{array}\right\} \tag{1}$$

where k, f are constant functions. For a given function A_f determine two scalar values k, f in order to minimize the cost functional $\frac{1}{2}\int_{\Omega_T}(A - A_f)^2 dxdt$.

In this example the design space is a vector $P = (k, f)$ and the cost functional takes the form $F(A, P) = \frac{1}{2}\int_{\Omega_T}(A - A_f)^2 dxdt$. We denote by δ a formal derivative of a function with respect to P in direction μ, namely $\delta u = \frac{\partial u}{\partial P}\mu$. Differentiation of F with respect to P gives $\delta F = \int_{\Omega_T} \delta A(A - A_f)dxdt$.

We derive the so called *sensitivity equation* by formal differentiation of (1) with respect to P in direction μ

$$\left.\begin{array}{l} \delta A_t - k\delta A_{xx} - \delta k A_{xx} = \delta f, \qquad \text{for } (x,t) \in \Omega_T \\ \delta A_x(0,t) = \delta A_x(1,t) = 0, \text{ for } t \in I \quad \text{and} \quad \delta A(x,0) = 0, \text{ for } x \in \Omega. \end{array}\right\} \tag{2}$$

The adjoint system looks like

$$\left.\begin{array}{l} -\varphi_t - k\varphi_{xx} = A - A_f, \qquad \text{for } (x,t) \in \Omega_T \\ \varphi_x(0,t) = \varphi_x(1,t) = 0, \text{ for } t \in I \quad \text{and} \quad \varphi(x,10) = 0, \text{ for } x \in \Omega. \end{array}\right\} \tag{3}$$

Notice the pseudosource $A - A_F$ appearing in the adjoint system. This term comes from the expression for δF. Next we integrate over time-space domain the multiplication of the principal equation, first from (2) with φ, and second from (3) with δA. Resulting two equations we subtract to obtain

$$\int_{\Omega_T} \delta k A_{xx}\varphi dxdt + \int_{\Omega_T} \delta F\varphi = \int_{\Omega_T}(A - A_f)\delta A dxdt = \frac{\partial F}{\partial P}\mu. \tag{4}$$

In such a way we obtained an explicit expression for the derivative $\frac{\partial F}{\partial P}\mu$ for an arbitrary μ involving one solution of the adjoint system denoted by φ. So if P has 1000 components, only one solution of an adjoint system is necessary in order to obtain $\frac{\partial F}{\partial P}\mu$. In the case of direct differentiation, it is necessary to evaluate 1000 times the direct problem. Note, that the direct problem has the same structure as the adjoint problem, so for the solution of an adjoint system one can use the direct solver with different pseudo-source and no re-implementation is necessary.

3 Magnetic shielding optimization problem

We applied the above described AVM in the design optimization problem of the magnetic shielding. The setting is described in Figure 1. We consider the axisymmetric case with the axis of rotation being the left part pf the boundary Ω. Similar setting was explored in [6].

The governing equations in time harmonic domain read as

$$\nabla\left[(\mu(\mathbf{P}))^{-1}\nabla A(\mathbf{P})\right] + \partial/\partial r\left((r\mu(\mathbf{P}))^{-1}A(\mathbf{P})\right) - j\omega\sigma(\mathbf{P})A(\mathbf{P}) = -J_e(\mathbf{P}),$$

equipped with the following boundary conditions $A = 0$ on Γ_1 and $\nabla A \cdot \mathbf{n} = 0$ on Γ_2. Further developments will be done using weak formulation of the above problem.

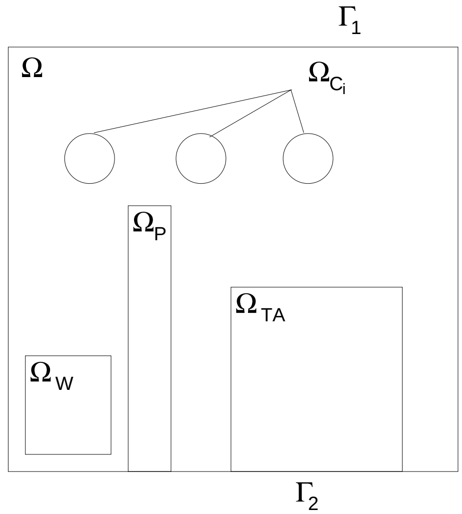


Figure 1: Domain Ω. Γ_2 is the bottom horizontal boundary of Ω and Γ_1 is the rest of the boundary of Ω; Ω_W is the workpiece generating the stray field, Ω_P is the passive shield; Ω_{TA} is the target area that has to be shielded; Ω_{C_i} are the compensation coils of the active shield.

The objective value is a function of the scalar potential A and of ∇A. First of all, it takes into account the magnetic field in the target area (F_B): the main objective is to minimize the magnetic induction $B = \nabla \times A$ in the reference region Ω_{TA} which in the axisymmetric case can be taken as $B = \nabla A$. Other objective values in the multi-objective optimization are the dissipation in the passive shield (F_P), the dissipation in the active shield (F_A), the change of the heating of the workpiece by the adding of shields (F_W) and the volume of the passive shield as an investment cost (F_V). The complete cost functional takes the form

$$F(\mathbf{P}) = w_1 F_B + w_2 F_P + w_3 F_A + w_4 F_W + w_5 F_V,$$

where $w_i, i = 1, \dots, 5$ are appropriate weights and the corresponding subfunctionals are given by

$$F_B = \frac{1}{2}\|\nabla A\|^2_{\Omega_{TA}}, \quad F_P = \frac{1}{2}\sigma\omega^2\|A\|^2_{\Omega_P}, \quad F_A = \frac{\pi r \rho}{S_A}\sum_{i=1}^{n}(I^2_{\mathrm{r},i} + I^2_{\mathrm{i},i}),$$
$$F_W = P_{W,0} - \frac{1}{2}\sigma\omega^2\|A\|^2_{\Omega_W}, \quad F_V = 2\pi r_p h_p t_p.$$

Herein, S_A is the cross section of an active shield coil and ρ is the resistivity of the coil material. The constant $P_{W,0}$ is the power dissipated in the workpiece without shields present. Further, r_p is the position and h_p the height of the passive shield with thickness t_p in steel, and $I_{\mathrm{r},i}, I_{\mathrm{i},i}, i = 1, \dots, n$ are the (complex) currents of the n active shield coils in the axisymmetric problem.

Following the idea presented in Example 1 one can construct the following adjoint problem formulated in weak sense: Find $\theta \in H_{0,\Gamma_1}(\Omega)$ such that for all $\varphi \in H_{0,\Gamma_1}(\Omega)$ the following holds

$$(\mu^{-1}\nabla\varphi, \nabla\xi) + (j\omega\sigma\varphi, \xi) + \big((\mu r)^{-1}\varphi, \partial\xi/\partial r\big) = a(\varphi). \tag{5}$$

Denoting by $(u, v)^* = 1/2[(u^*, v) + (u, v^*)]$ the complex conjugate product, we define

$$a(\varphi) = (\nabla\varphi, \nabla A)^*_{\Omega_{TA}} + \sigma\omega^2(\varphi, A)^*_{\Omega_P} - \sigma\omega^2(\varphi, A)^*_{\Omega_W}$$

The explicit expression using the solution of the adjoint problem reads as

$$\begin{aligned}\delta F \;=\; & (\delta J_e, \xi) - (\delta(\mu^{-1})\nabla A, \nabla\xi) - (j\omega\delta\sigma A, \xi) - \big(r^{-1}\delta(\mu^{-1})A, \partial\xi/\partial r\big) \\ & + \frac{1}{2}\delta\sigma\omega^2\|A\|^2_{\Omega_P} + \frac{1}{2}\sigma\omega^2\int_{\delta\Omega_P} A \cdot A^* \mathrm{d}(\delta\Omega) + \frac{1}{2}\delta\sigma\omega^2\|A\|^2_{\Omega_W} + 2\pi r_p t_p \delta h_p \\ & + \frac{2\pi r\rho}{S_A}\sum_{i=1}^{n}\big((\delta I_{\mathrm{r},i})I_{\mathrm{r},i} + (\delta I_{\mathrm{i},i})I_{\mathrm{i},i}\big) + \delta r\frac{\pi\rho}{S_A}\sum_{i=1}^{n}(I^2_{\mathrm{r},i} + I^2_{\mathrm{i},i}).\end{aligned}$$

	Optimized parameters		Conv. DDM		Grad+AVM	
	Pas.	Act.	CPU	$F_{\min}$	CPU	$F_{\min}$
1	h_{p}	-	6' 28"	2.9562	9' 24"	3.1376
2	h_{p}	$I_{\mathrm{r}1}, I_{\mathrm{i}1}, r_1$	8' 13"	2.9100	8' 15"	2.9681
3	-	$I_{\mathrm{r}i}, I_{\mathrm{i}i}, r_i, i = 1, 2, 3$	1h 16' 54"	3.0523	16' 13"	2.9832

Table 1: Comparison of optimization techniques concerning calculation time and optimal cost value $F_{\min}$ for several shielding problems

For more details on derivation of this formula we refer to [5].

Numerical study

We performed a numerical study aiming at the comparison of direct differentiation method (DDM) and adjoint variable method. We carried out three cases. We optimize:

The height of the passive shield. In optimization 1, the passive shield has conductivity $5.9 \times 106 S/m$ and permeability of $\mu_r = 372$. Table 1 shows that the optimum at 100.1 mm with its cost 2.9562 is found by the conventional gradient DDM only. The adjoint method finds an approximation 71.7 of the optimum and a slightly higher objective value.

For such a high value of μ the AVM is very sensible on numerical errors because of the term $\delta\mu$. The computed gradients are less accurate for AVM then for DDM. This explains, why AVM does not find the optimum. However, the costs differ only slightly, so that the solution obtained by the adjoint method is still an acceptable solution.

Both the passive and the active shield with one coil. In optimization 2, the complex current $I_{\mathrm{r}1} + jI_{\mathrm{i}1}$, the horizontal position r_1 and the height of the passive shield is optimized (4 parameters in total). The following starting value was chosen: $[h_p, I_{\mathrm{r}1}, I_{\mathrm{i}1}, r_1] = [0.060, 200, -121, 0.300]$. The gradient method using the adjoint system ends up with a higher cost, because the gradient of the adjoint method crosses zero and since the gradient is less accurate than in DDM case, the AV method does not move anymore from the approximated solution. The calculation time of both gradient algorithms has the same magnitude.

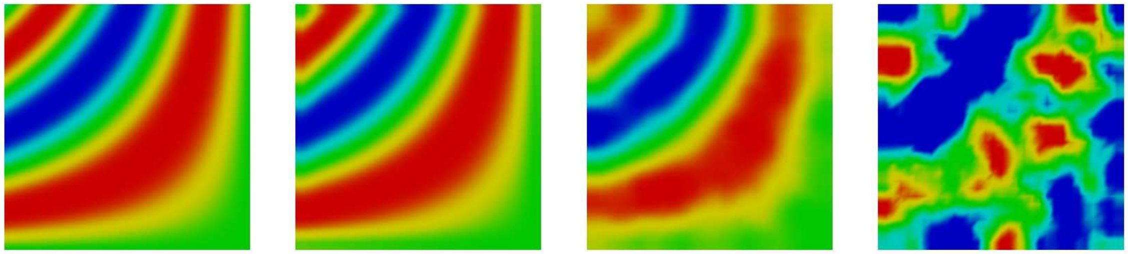
An active shield consisting of 3 coils. Optimization 3 results in 9 optimization parameters (three times real part of the current, imaginary part of the current and horizontal position). The starting positions were $0.3, 0.4$ and $0.6m$, the starting values for the currents were $I_{\mathrm{r}} = [300, -100, 50]$ and $I_{\mathrm{i}} = -I_{\mathrm{r}}/1.65$. For 9 variables, the approach using the adjoint variable is much faster than the conventional DDM algorithm. Moreover, it finds a lower cost value although the gradients don't deviate much from the conventional ones.

The detailed results can be found in Table 1. One can clearly see that as soon as the number of optimized parameters growth, AVM method becomes much more effective in computational costs and even gives better results then conventional DDM.

4 Determination of micromagnetic parameters

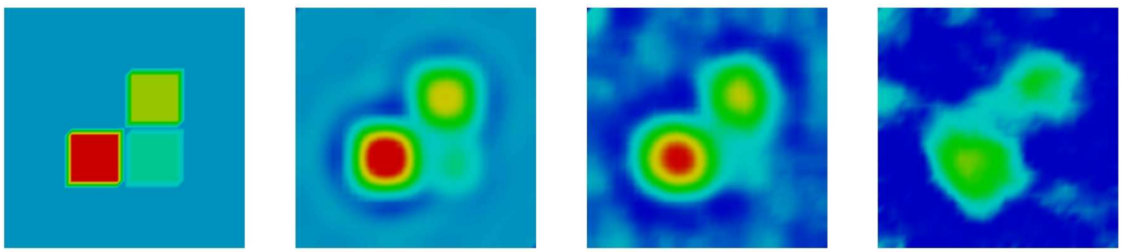
In the micromagnetics, AVM was already successfully used in the design optimization of the ferromagnetic core in MRAM memories [3]. We applied AVM in the case of determination of a dissipation parameter in micromagnetics. The micromagnetic model described by the Landau-Lifshitz equation models the electromagnetic behaviour on very small time and space scales. The governing equation reads as

$$\partial \mathbf{m}/\partial t = -\gamma \left(\mathbf{m} \times \mathbf{H}_{\mathrm{eff}} + \alpha \mathbf{m} \times (\mathbf{m} \times \mathbf{H}_{\mathrm{eff}}) \right), \quad \text{in } \Omega \qquad (6)$$



(a) exact (b) 0% noise (c) 1% noise (d) 10%noise

Figure 2: Results for smooth exact solution.



(a) exact (b) 0% noise (c) 1% noise (d) 10%noise

Figure 3: Results for discontinuous piecewise constant exact solution.

where the gyromagnetic factor γ is a space and time constant, $\alpha(\mathbf{x})$ is a space variable and time constant function, and $\mathbf{H}_{\text{eff}}$ is effective field including anisotropy, demagnetizing, exchange and applied field effects. System is completed with the initial and boundary conditions $\mathbf{m}(\mathbf{x}, 0) = \mathbf{m}_0(\mathbf{x})$ and $\partial \mathbf{m}/\partial \mathbf{n} = 0$ on $\partial\Omega$.

The forward problem has been extensively studied in [4] from the numerical point of view and in [7] from computational point of view.

The aim is to determine $\alpha(\mathbf{x})$ from the measurements of $\mathbf{m}$ available over time and space domain. First we define a cost functional $F(\alpha) = \int_{\Omega_T} \|\mathbf{m} - \mathbf{m}_{\text{mes}}\|^2 - \|\alpha_n - \alpha_{n-1}\|^2$. First term in the integral minimizes the difference between the measured data and the computed solution, the second term serves as a regularization. So we end up with a minimization problem described by the cost functional F subject to PDE (6) describing the relation between state variable $\mathbf{m}$ and the design variable α. The subscript n in α_n refers to the nth iteration of the minimization process.

In this case, we approximate α_n by the Lagrange first order finite elements. This however means, that in the case of a regular triangulation of the unit square in 2D $\Omega = (0,1) \times (0,1)$, where one side of the square is divided into N segments, we have $(N+1)^2$ degrees of freedom, so our design space has dimension $(N+1)^2$. The direct differentiation method can not be used in this case, since the computation of one gradient would require $(N+1)^2$ evaluations of the direct problem. In this particular example, one can see the strength of the adjoint variable method.

We do not provide the actual form of the adjoint problem, for the details we refer to [2]. We provide the numerical simulations.

Computations

We set up two scenarios. First we take the exact solution to be a smooth function and then we test our algorithm on an exact solution with discontinuities.

Smooth exact solution Take the following exact solution

$$\alpha_{\text{exact}} = 0.02 + 0.01 \sin(\pi^2 xy), \text{for } (x,y) \in \Omega = (0,1) \times (0,1)$$

In Figure 2a the exact solution is depicted. The reconstructed solution can be seen in Figure 2b. We added noise in order to test the reliability of the algorithm and in Figures 2c and 2d we see that with 10% noise the results become unreliable.

Piecewise constant exact solution Take the following exact solution

$$\alpha_{\text{exact}} = \begin{cases} 0.01 & (x,y) \in \Omega_1 = (0.25, 0.5) \times (0.25, 5) \\ 0.02 & (x,y) \in \Omega_2 = (0.5, 0.75) \times (0.25, 5) \\ 0.03 & (x,y) \in \Omega_3 = (0.25, 0.5) \times (0.5, 0.75) \\ 0 & (x,y) \in \Omega \setminus (\Omega_1 \cup \Omega_2 \cup \Omega_3) \end{cases}$$

In Figure 3a the exact solution is depicted. The reconstructed solution can be seen in Figure 3b. Comparing the results with noise one can see, that the acceptable noise level is less than 1%. This is due to the nature of regularization. The L^2 regularization used above is more suitable for smooth functions. In order to obtain better results for discontinuous solution, one need to use total variation regularization.

5 Conclusions

From our first example, clearly the adjoint variable method is effective for the case when the design space is higher dimensional. We have seen that optimization using the adjoint variables is slower than the conventional gradient method in case of less than three parameters to optimize, comparable in case of three or four parameters and faster in case of more than four parameters. This is due to the fact that one needs two more additional evaluations of adjoint problem for real and imaginary part of the unknown.

In our second example we verified that ADV is effective also in the case of inverse problems. Our design space included 441 degrees of freedom which correspond to the FEM approximation of the unknown function $\alpha(\mathbf{x})$ on a regular mesh consisting of 2×20^2 triangles. Still, AVM method was able to capture $\alpha(\mathbf{x})$ acceptably well.

References

[1] T. Burczynski, J. H. Kane, and C. Balakrishna. Comparison of shape design sensitivity analysis formulations via material derivative-adjoint variable and implicit differentiation techniques for 3-d and 2-d curved boundary element. *Comput. Meth. Appl. Mech. Eng.*, 142:89–109, 1997.

[2] I. Cimrák and V. Melicher. Determination of precession and dissipation parameters in the micromagnetics. Preprint.

[3] I. Cimrák and V. Melicher. Sensitivity analysis framework for micromagnetism with application to optimal shape design of MRAM memories. *Inverse Problems*, 23:563–588, 2007.

[4] A. Prohl. *Computational Micromagnetism.* Advances in Numerical Mathematics. B. G. Teubner, Stuttgart, 2001.

[5] P. Sergeant, I. Cimrák, V. Melicher, L. Dupré, and R. Van Keer. Adjoint variable method for the study of combined active and passive magnetic shielding. submitted.

[6] P. Sergeant, L. Dupré, M. De Wulf, and J. Melkebeek. Optimizing Active and Passive Magnetic Shields in Induction Heating by a Genetic Algorithm. *IEEE Trans.Magn.*, 39(6):3486–3496, 2003.

[7] D. Suess, J. Fidler, and T. Schrefl. Micromagnetic simulation of magnetic materials. In K.H.J. Buschow, editor, *Handbook of Magnetic Materials*, volume 16, pages 41–125. Elsevier, 2006.