DESIGN OF ELECTROMAGNETIC DEVICES USING SENSITIVITIES COMPUTED WITH THE ADJOINT VARIABLE METHOD

D N Dyck, D A Lowther   E M Freeman

McGill University, Canada   Imperial College, England

The inverse problem in electromagnetics often deals with quantities of the form:

$$F = \int_V f(\boldsymbol{E}, \boldsymbol{B})\, dV \tag{1}$$

where f is a function of the electric and magnetic fields at a point. The fields are themselves implicit functions of the system parameters: the conductivity, permeability, source current density, etc. In general, the fields and the system parameters will take on different values at each point in space.

Efficient optimization of a quantity such as (1) requires derivative information. This information is the sensitivity of F to changes in the material parameters and sources. A powerful method of computing these sensitivities, which is widely used in structural engineering (Haug et al (1)), and recently in the magnetostatic case (Park et al (2)), is the adjoint variable method. This method allows the sensitivities with respect to all system parameters to be computed at the cost of a solution to an adjoint problem.

This paper derives expressions for the sensitivities in terms of the fields of the original and the adjoint problems. Expressions for the sources to the adjoint problem are also derived. These expressions are then applied to an example in which the objective is to maximize the levitating force on a conducting ring.

THE ADJOINT VARIABLE METHOD

The Method of Weighted Residuals

This derivation of the sensitivity expressions is closely tied to the method of weighted residuals, which is a method for solving the system equations (see, for example, Silvester and Ferrari (3)). In the weighted residual approach, a field quantity ϕ is a solution if it satisfies the following equation:

$$a(u;\phi,\phi') = l(u;\phi') \tag{2}$$

for all functions ϕ', defined over the problem domain, which satisfy the boundary conditions. In this equation, u is the vector of parameters which specifies the physical structure of the system, e.g. one parameter in time harmonic systems is the conductivity. The left hand side of (2) is called the bilinear term, and the right hand side is the forcing term. Equation (2) is derived by taking the system equation, multiplying both sides by a weighting function ϕ', and integrating over the problem domain. Thus both $a(u;\phi,\phi')$ and $l(u;\phi')$ are integrals which depend on the values of ϕ and ϕ' everywhere in the problem domain.

The field solution is obtained using (2) by expressing the solution as a weighted sum of basis functions. The coefficients of the basis functions can be found using Galerkin's method, which generates a system of linear equations by substituting each basis function for ϕ' in (2).

The bilinear term possesses certain properties which are relied upon to derive the sensitivity expressions. As the name suggests, the bilinear term is linear with respect to ϕ and ϕ'. It is also symmetric with respect to ϕ and ϕ': $a(u;\phi,\phi') = a(u,\phi',\phi)$. Finally, the bilinear term must be continuous with respect to the parameters u.

Derivation of the Sensitivity Expressions

This derivation closely follows the derivation given in Haug et al (1). Consider a quantity of interest which depends on the solution ϕ:

$$F = \int_V f(\phi)\, dV \tag{3}$$

The perturbation in this quantity due field perturbations (which in turn are a result of perturbations in the system parameters) is:

$$\delta F = \int_V \frac{\partial f}{\partial \phi}\, \delta\phi\, dV \tag{4}$$

For reasons which will become clearer later on, it is convenient to introduce an adjoint problem. The adjoint problem will have the same bilinear term, but the forcing term will be defined as:

$$l_A(u;\phi_A') = \int_V \frac{\partial f}{\partial \phi}\, \phi_A'\, dV \tag{5}$$

so that:

$$\delta F = l_A(u;\delta\phi) \tag{6}$$

The solution to the adjoint problem is ϕ_A, which is defined as the function which satisfies:

$$a(u;\phi_A,\phi'_A) = l_A(u;\phi'_A) \tag{7}$$

for all functions ϕ'_A, defined on the problem domain, and with homogeneous boundary conditions. Now $\delta\phi$ is a function defined on the problem domain, and it satisfies homogeneous boundary conditions since the boundary conditions are not affected by parameter perturbations. Therefore $\delta\phi$ can be substituted for ϕ'_A in (7) to yield:

$$a(u;\phi_A,\delta\phi) = l_A(u;\delta\phi) \tag{8}$$

substituting this into (6) yields:

$$\delta F = a(u;\phi_A,\delta\phi) \tag{9}$$

Now the symmetry property of $a()$ can be applied to (9) resulting in:

$$\delta F = a(u;\delta\phi,\phi_A) \tag{10}$$

The property of linearity of $a()$ allows this to be re-written as:

$$\delta F = a(u;\phi+\delta\phi,\phi_A) - a(u;\phi,\phi_A) \tag{11}$$

And since $a()$ is continuous in u, this also holds if the parameters u are perturbed slightly (this amounts to neglecting the second order term $a(\delta u;\delta\phi,\phi_A)$):

$$\delta F = a(u+\delta u;\phi+\delta\phi,\phi_A) - a(u+\delta u;\phi,\phi_A) \tag{12}$$

But $\phi+\delta\phi$ is a solution to the problem with perturbed parameters $u+\delta u$, and so it satisfies:

$$a(u+\delta u;\phi+\delta\phi,\phi') = l(u+\delta u;\phi') \tag{13}$$

for all functions ϕ', under the same conditions as before. Since ϕ_A qualifies as one of these functions, (13) can be substituted into (12) with $\phi'=\phi_A$ to yield:

$$\delta F = l(u+\delta u;\phi_A) - a(u+\delta u;\phi,\phi_A) \tag{14}$$

If both terms are linear in u, then since ϕ satisfies (2) with $\phi'=\phi_A$, (14) reduces to:

$$\delta F = l(\delta u;\phi_A) - a(\delta u;\phi,\phi_A) \tag{15}$$

This is the general form of the sensitivity equation for any variational formulation. Note that (15) does not depend on the perturbed solution. Therefore once the solution to the adjoint problem is found, (15) allows the perturbation to F to be computed for all small parameter perturbations.

LINEAR TIME-HARMONIC SYSTEMS

Weighted Residual Formulation

For the linear time harmonic case, Maxwell's equations are:

$$\begin{aligned} \nabla\times\hat{\boldsymbol{E}} &= -j\omega\hat{\boldsymbol{B}} \\ \nabla\times(\nu\hat{\boldsymbol{B}}) &= (\sigma+j\omega\epsilon)\hat{\boldsymbol{E}} + \hat{\boldsymbol{J}}_S \end{aligned} \tag{16}$$

The system parameters in this case are: the source current density $\hat{\boldsymbol{J}}_S$, the conductivity σ, the permittivity ϵ, and the reluctivity ν. The bilinear and forcing terms corresponding to the weighted residual formulation of these equations are:

$$a(\hat{\boldsymbol{J}}_S,\sigma,\epsilon,\nu;\hat{\boldsymbol{E}},\hat{\boldsymbol{E}}') = -\int_V\left[(\sigma+j\omega\epsilon)\hat{\boldsymbol{E}}\cdot\hat{\boldsymbol{E}}' + (j\omega)^{-1}\nu(\nabla\times\hat{\boldsymbol{E}})\cdot(\nabla\times\hat{\boldsymbol{E}}')\right]dV \tag{17a}$$

$$l(\hat{\boldsymbol{J}}_S,\sigma,\epsilon,\nu;\hat{\boldsymbol{E}}') = \int_V\hat{\boldsymbol{J}}_S\cdot\hat{\boldsymbol{E}}'\,dV \tag{17b}$$

Adjoint System Forcing Term

To clarify the procedure, the forcing term will first be derived for objective functions of the form:

$$F = \int_V f(\hat{\boldsymbol{E}})\,dV \tag{18}$$

The expression for the adjoint forcing term follows by a direct substitution of the system variables into (5):

$$l_A(\hat{\boldsymbol{J}}_S,\sigma,\epsilon,\nu;\hat{\boldsymbol{E}}'_A) = \int_V(\nabla_{\hat{E}}f)\cdot\hat{\boldsymbol{E}}'_A\,dV \tag{19}$$

where $\nabla_{\hat{E}}$ is the gradient with respect to the components of the electric field $\hat{\boldsymbol{E}}$. Note that this forcing term is equivalent to driving the adjoint system with a source current density of:

$$\hat{\boldsymbol{J}}_{SA} = \nabla_{\hat{E}}f \tag{20}$$

Sensitivities in Linear Time Harmonic Systems

To simplify the final expressions, the bilinear term in (17a) can be re-written as:

$$a(\hat{\boldsymbol{J}}_S,\sigma,\epsilon,\nu;\hat{\boldsymbol{E}},\hat{\boldsymbol{E}}') = -\int_V\left[(\sigma+j\omega\epsilon)\hat{\boldsymbol{E}}\cdot\hat{\boldsymbol{E}}' + j\omega\nu\hat{\boldsymbol{B}}\cdot\hat{\boldsymbol{B}}'\right]dV \tag{21}$$

The sensitivities are derived by simply substituting the expressions for the forcing term in (17b) and the bilinear term in (21) into (15), which yields:

$$\delta F = \int_V\left[\delta\hat{\boldsymbol{J}}_S\cdot\hat{\boldsymbol{E}}_A + (\delta\sigma+j\omega\delta\epsilon)\hat{\boldsymbol{E}}\cdot\hat{\boldsymbol{E}}_A + j\omega\delta\nu\hat{\boldsymbol{B}}\cdot\hat{\boldsymbol{B}}_A\right]dV \tag{22}$$

From this equation, the expressions for the sensitivity of F with respect to each parameter can be identified:

$$S_\sigma = \hat{\boldsymbol{E}}\cdot\hat{\boldsymbol{E}}_{\mathrm{A}} \qquad S_\epsilon = j\omega\hat{\boldsymbol{E}}\cdot\hat{\boldsymbol{E}}_{\mathrm{A}}$$
$$S_{J_S} = \hat{\boldsymbol{E}}_{\mathrm{A}} \qquad S_\nu = j\omega\hat{\boldsymbol{B}}\cdot\hat{\boldsymbol{B}}_{\mathrm{A}} \tag{23}$$

More General Objective Functions

The procedure outlined above, to derive sensitivity expressions, can be extended to apply to objective functions of the form:

$$F = \int_V f(\hat{\boldsymbol{E}},\hat{\boldsymbol{B}})dV \tag{24}$$

The perturbation experienced by this objective function, due to field perturbations, is:

$$\delta F = \int_V [\,(\nabla_{\hat{E}} f)\cdot\delta\hat{\boldsymbol{E}} + (\nabla_{\hat{B}} f)\cdot\delta\hat{\boldsymbol{B}}\,]dV \tag{25}$$

Using algebra and a vector identity, and assuming that either f or $\hat{\boldsymbol{E}}$ tangential vanishes on the boundary, this can be manipulated into the following form:

$$\delta F = \int_V (\,\nabla_{\hat{E}} f - (j\omega)^{-1}\nabla_{xyz}\times\nabla_{\hat{B}} f)\cdot\delta\hat{\boldsymbol{E}}\,dV \tag{26}$$

where the curl and gradient operators are differentiating with respect to different variables. The forcing term is again derived by substituting the integrand in the above expression into (5):

$$l(\hat{\boldsymbol{J}}_S,\sigma,\epsilon,\nu;\hat{\boldsymbol{E}}_A') = \int_V (\,\nabla_{\hat{E}} f - (j\omega)^{-1}\nabla_{xyz}\times\nabla_{\hat{B}} f)\cdot\hat{\boldsymbol{E}}_{\mathrm{A}}'\,dV \tag{27}$$

This forcing term is equivalent to a source current density of:

$$\hat{\boldsymbol{J}}_{SA} = \nabla_{\hat{E}} f - (j\omega)^{-1}\nabla_{xyz}\times\nabla_{\hat{B}} f \tag{28}$$

APPLICATION TO DESIGN

The theoretical results derived in the previous sections have been applied to the design of electromagnetic devices. The following sub-section explains how the optimization is set up for a simple objective function. The approach used for this example will clarify the method of application of the technique to other cases of interest.

Force on a Jumping Ring

This example "designs" the support structure for a jumping ring. Fig. 1 shows the basic geometry of the system. The system has translational symmetry, so the ring is actually an infinitely long rectangle. In the

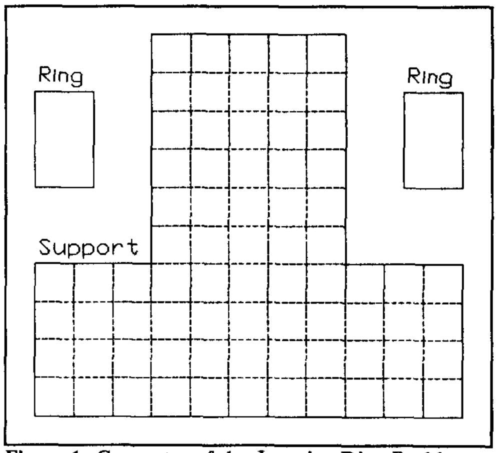

Figure 1: Geometry of the Jumping Ring Problem

subdivided region, each square is allowed to contain source current (either positive or negative) or a highly permeable material (i.e. laminated iron), or some mixture of the two. Note that air is included among the possibilities (zero current).

The objective here is to maximize the time average of the upward force on the ring. The force on the ring is computed by integrating $\boldsymbol{J}\times\boldsymbol{B}$ over the cross section of the ring. The upward force is of the form (25), although, because of translational symmetry, the volume integral reduces to a surface integral over the cross section of the ring:

$$F_y = \int_{ring} \sigma\hat{\boldsymbol{E}}\times\hat{\boldsymbol{B}}^*\cdot\vec{y}\,dS \tag{29}$$

where $\vec{y}$ is the unit vector in the upward direction. The quantity of interest is the time average of the upward force, which is simply:

$$F = \frac{1}{2}\mathrm{Re}(F_y) \tag{30}$$

In the conductor, the electric field is entirely $\vec{z}$ directed (because of translational symmetry), and therefore:

$$F_y = \int_{ring} \sigma\hat{E}_z\hat{B}_x^*\,dS \tag{31}$$

To compute the sensitivity, the source must be set according to (29). However, here it becomes apparent that f is not exactly in the required form, since it is a function, not of $\hat{\boldsymbol{H}}$, but of the complex conjugate of $\hat{\boldsymbol{H}}$. To evaluate this sensitivity would normally require solving an additional adjoint system with a complex conjugated source, and using the complex conjugate of the sensitivity expressions. However, in this case, only the real part of F_y is used, so a simpler approach is possible. Consider the perturbation in F_y resulting from perturbations in the fields:

$$\delta F_y = \int_V \left[\left(\nabla_{\hat{E}} f_y \right) \cdot \delta \hat{E} + \left(\nabla_{\hat{B}^*} f_y \right) \cdot \delta \hat{B}^* \right] dV \quad (32)$$

where f_y is the integrand of F_y. Now consider:

$$\delta \tilde{F}_y = \int_V \left[\left(\nabla_{\hat{E}} f_y \right) \cdot \delta \hat{E} + \left(\nabla_{\hat{B}} f_y^* \right) \cdot \delta \hat{B} \right] dV \quad (33)$$

It is easy to check that the real parts of (32) and (33) are the same, i.e.: $\mathrm{Re}(\delta F_y) = \mathrm{Re}(\delta \tilde{F}_y)$. Therefore, (33) can be used instead of (25) to derive an expression for the source current density in the adjoint system. This turns out to be:

$$\hat{J}_{SA} = \nabla_{\hat{E}} f_y - (j\omega)^{-1} \nabla_{xyz} \times \nabla_{\hat{B}} f_y^* \quad (34)$$

which reduces to:

$$\hat{J}_{SAz} = \sigma \hat{B}_x^* + \frac{1}{j\omega} \frac{\partial}{\partial y} \left(\sigma E_z^* \right) \quad (35)$$

Once the adjoint system is solved with this source, the sensitivity of F_y to perturbations in any of the system parameters are given by the expressions in (23). The sensitivity of the objective function is found from these expressions by applying the chain rule to (30), and turns out to be simply one half the real part of these expressions:

$$\begin{aligned} S_\mu &= \tfrac{1}{2} \mathrm{Re}\left(j\omega \hat{B} \cdot \hat{B}_A \right) = -\tfrac{\omega}{2} \mathrm{Im}\left(\hat{B} \cdot \hat{B}_A \right) \\ S_{J_s} &= \tfrac{1}{2} \mathrm{Re}\left(\hat{E}_A \right) \end{aligned} \quad (36)$$

Results

Using this derivative information, the optimization converged to a partial design after 21 iterations, as shown in Fig. 2. The optimization was constrained to limit the peak current density, and the reluctivity was also constrained to be between its value in air and its value in iron.

Note that the final design is not composed of solid material everywhere. However, the basic structure is clearly visible: an iron shell completely filled in with current carrying windings (the left and right currents are of opposite sign).

CONCLUSION

This paper derives expressions to compute the sensitivity of electromagnetic objective functions to perturbations in material parameters and sources.

The sensitivity expressions are applied to optimize the support structure of a jumping ring.

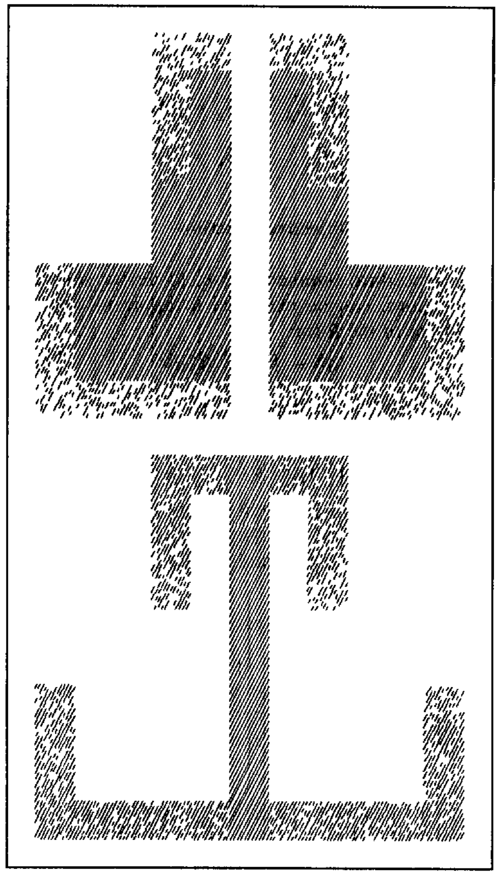

Figure 2: Design of Jumping Ring Support Structure
Top: Current Density Distribution,
Bottom: Distribution of Permeable Material.

[1] Haug, E.J., Choi, K.K., Komkov, V., 1986, "Design Sensitivity Analysis of Structural Systems", Academic Press, Orlando, USA

[2] Park, I.H., Coulomb, J.L., Hahn, S.Y., 1992, Journal de Physique III, 2, 2045-2053

[3] Silvester, P.P, Ferrari, R.L., 1983, "Finite Elements for Electrical Engineers," Second Edition, Cambridge University Press, Cambridge, UK