Journal of Engineering and Natural Sciences
Mühendislik ve Fen Bilimleri Dergisi

Sigma
2004/2

Invited Review Paper

OPTIMAL DESIGN OF NONLINEAR MAGNETIC SYSTEMS USING FINITE ELEMENTS

Levent OVACIK*

Istanbul Technical University, Department of Electrical Engineering, Maslak-ISTANBUL

Geliş Tarihi: 15.06.2004

SONLU ÖĞELER İLE DOĞRUSAL OLMAYAN MANYETİK DİZGELERİN ENİYİ TASARIMI

ÖZET

Manyertik bir cihaz içerisindeki belirlenmiş noktalarda ve doğrultuda arzu edilen bir manyetik akı yoğunluğu dağılımını yaklaşık olarak elde etmek için, manyetik bir cihazın eniyi tasarımını elde edebilen bir ters sonlu öğeler yöntermi geliştirilmiştir. Ağırlaştırılmış Lagrange çarpanları yöntermi kullanılarak en küçük kareler fonksiyonundan oluşan bir amaç fonksiyonu ile bir dizi kısıt denklemlerinden oluşan bir eniyilik problemi kısıtsız bir eniyilik problemine dönüştürülmüştür. Lagrange-Newton yöntemine dayalı ikinci mertebeden bir yaklaşım ile dönüştütülmüş problemin çüzülmesiyle, manyetik cihazın tasarımı adım adım iyileştirilmiştir. Manyetik malzemede doymanın gözönünde bulundurulması durumunda, ikinci mertebeden duyarlılık analizi için gerekli türevlerin hesaplanması oldukça zor bir işlemdir. Bu makalede, geliştirilen yeni bir yöntem ile eniyileştirme işlemi ile doğrusal olmayan statik manyetik alan probleminin çözümüne ilşkin doğrusal olmayan sonlu öğeler denklemleri uygun bir biçimde birleştirilerek, karmaşık duyarlılık analizi verimli bir şekilde yapılabilmiştir. Cihazın geometrisi için parametrik bir modelinin oluşturulması ve lineer olmayan manyetik malzemelerin matemtiksel modellerinin kurulmasına ilişkin yöntemler araştırılarak, geliştirilen bir bilgisayar programının içerisinde kullanılmasıyla, karalı bir çözüm elde edilebilmiştir. Elektrik makinalarının tasarımına ilişkin çeşitli problemler üzerinde yapılan denemlerde, bu makalede geliştirilen yöntemlerin sonuçları sunulmuştır.
Anahtar Sözcükler: Sonlu elemanlar (öğeler) yöntemi, optimizasyon (eniyilik), manyetik sistemler (dizgeler)

ABSTRACT

An inverse finite element method was developed to find optimal geometric parameters of a magnetic device to approximate a desired magnetic flux density distribution at certain test points and directions selected in the device. The augmented Lagrange multipliers method was utilized to transform the constrained problem consisting of a least-square objective function and a set of constraint equations to the unconstrained problem. A second-order approach based on the Lagrange-Newton method was used to minimize the unconstrained problem to improve the design iteratively. Numerical calculation of derivatives in the second-order design sensitivity analysis becomes a difficult task if saturation in material properties is accounted. A novel approach is developed to minimize the computational effort by directly combining the optimization process with the nonlinear finite element equations. The best capabilities to parametrize the device geometry and to model the nonlinear material characteristics were incorporated into the optimization program for rapid sensitivity analysis. Demonstration of various test cases arising from optimally designing electrical machinery verified the validity of the overall theory and developments.
Keywords: Finite element method, numerical optimization, magnetic systems

* e-mail:ovacik@elk.itu.edu.tr, Tel: (212) 285 6765

1. INTRODUCTION

The theory of the finite element method (FEM) for calculating flux distribution in electromagnetic devices has been well established. Finding a vector potential solution for a given geometry, material properties, excitation sources, and so on is called a "forward" problem. Due to the competitive world market, designers attempt to reduce cost, weight, and/or to improve efficiency and reliability of electromagnetic devices. Therefore, the designer is more interested in finding the geometry of a certain portion of a device, such as iron parts, sizes and locations/positions of excitation coils to satisfy a given field, torque or force pattern, rather than finding how magnetic flux is distributed inside the device. Since this requires back-calculation of device-descriptive design parameters, these types of design optimization problems are classified as "inverse" problems.

Although the theory of mathematical programming has been widely developed since the 1950's, the initial research on optimal design problems emerged from the area of structural mechanics in the early 1970's. During the last 20 years, the optimal design techniques combining optimization methods with solution of governing partial differential equations (PDE) have been extensively studied. The results can be found in the textbooks, for example, Gallagher and Zienkiewicz [1], Haug and Arora [2], Kirch [3], Pironneau [4], Vanderplaats [5], and Haslinger and Neittaanmaki [6].

In electromagnetics, there have recently been diverse applications of design optimization problems appearing in the literature [7]-[46]. Some applications are concerned with optimizing device output forces (Gitosusastro et al. [9], and Saldanha et al. [10], [11]) while others are concerned with minimizing eddy current losses in conductive parts of devices (Kasper [12]), or minimizing production cost (Arpino et al. [13], Appelbaum et al. [14], [15]). Many of the applications deal with optimizing field patterns where an objective function in the form of the squared sum of the differences between calculated and desired values of field quantities at selected points is minimized. The results have appeared in the recent papers (Simkin and Trowbridge [16], Weeber [17], Weeber and Hoole [18], [19], Koh et al. [20], Park et al. [21], [22], Subramaniam et al. [23], and Vasconcelos et al. [24]). In this type of problems, the objective function is expected to achieve a minimum at zero in the absence of constraint equations. However, in the presence of constraint equations, the optimum is not at zero, but the best possible solution approaching the desired criterion is found. In some problems, a unique solution may not exist or there may be no solution.

There have been different field analysis methods used in electromagnetic optimization. Initially, integral methods were performed for field calculations while gradients with respect to design parameters were determined by the finite difference method (FDM) (Gottvald [44], [45], [46], Simkin and Trowbridge [16], and Park et al. [21]). Perhaps, one of the earliest attempts using the finite element method (FEM) in an optimization process were presented by Salon and Istfan [7], and Istfan and Salon [8]. In their work, they cast a sensitivity analysis based on direct differentiation of the finite element matrices with respect to nodal displacements describing a device's geometry. Weeber [17], Weeber and Hoole. [18], [19], Koh et al. [20], Park et al. [21], and Hoole [25] subsequently used this direct differentiation scheme for sensitivity analysis. The boundary element method (BEM), developed for analysis of open boundary problems, has recently been used in sensitivity analysis (Koh et al. [20], Park et al. [21], Enokizono and Tsuchida [26], and Enokizono et al. [27])

Since the objective function is a highly nonlinear function of design parameters, a wide variety of nonlinear optimization strategies have been investigated, and tested on different problems for their speed, convergence and efficiency (Preis et al. [28], and Gottvald [44]). These optimization methods are classified into two main categories: 1) deterministic methods, and 2) stochastic methods. The deterministic methods include the first-order methods such as the steepest-descent, conjugate gradient and quasi-Newton methods (see, e.g., [5], [47] and [48]).

These methods are based on finding the optimum in the decreasing gradient direction, requiring only the first-order derivatives. The first-order derivatives are calculated by means of sensitivity analysis incorporating various field solution techniques, such as the finite difference, and the finite element methods. Stochastic methods such as simulated annealing (SA) [16], [49], and the methods based on evolution strategies [45], and genetic algorithms (GA) [49] are zeroth-order (derivative-free) methods, requiring evaluation of the objective function at so many points that, depending on the problem complexity, they may become computationally exhaustive. However, a major advantage is that these two methods have been shown to be globally convergent (Gottvald [46]).

Independently carrying out the field analysis and optimization steps has some advantages and disadvantages. One advantage is that commercially available efficient optimization packages save time for program development, especially for program modification when the optimization problem is changed, (e.g., when the objective function is modified or new constraint equations are added to the problem). Another advantage is that the program might have more flexibility in choosing another optimization algorithm when one cannot perform as well as the other. On the other hand, the applications in which more sophisticated derivative calculations are involved result in excessive number of calls of field analysis programs, increasing the computational cost.

A great deal of design optimization tools developed so far conventionally use formal optimization techniques using the first-order or gradient information to find an optimization direction to progress the design towards the optimum solution. The first-order derivatives of the objective function as well as the design constraints are formulated in a way that the field potential is considered as a function of design variables. In most problems, this relationship is implicit and the necessary derivatives are obtained by means of the design sensitivity analysis. Following the forward solution to the field potentials outside the optimization process, the derivatives of the field potentials with respect to the design parameters are numerically calculated by perturbing the field equations. While the first-order methods are computationally affordable and can be efficiently solved for large-sized optimization problems, extension to the second-order optimization methods become cumbersome, requiring intensive computational effort with increasing number of design parameters and constraint equations. Especially, if nonlinearity of material properties is considered for more realistic design problems, this numerical procedure becomes even more complicated when forming the large matrices from the perturbed field equations.

2. BASIC CONCEPT OF OPTIMAL DESIGN

The design optimization process requires basically two main modules: a module for field solution which utilizes an analytical (seldom) or numerical approximation (often) based on differential or integral approaches, such as finite difference methods, finite element method, boundary element method or hybrid methods (BEM-FEM); a module which employs optimization strategies ranging from crude “trial and error” strategies to robust mathematical programming techniques.

Traditional computer-aided design (CAD) systems integrate these two modules to search for an optimal solution based on simple, trial-and-error principles (see, for example, Hoole [50], Lowther and Silvester [51], and Binns et al. [52]). To find an optimal design, the optimization parameters are modified for each possible trial design state, and then the field analysis is performed for the new design. Then, the objective function is evaluated and the constraint equations are checked if they are satisfied. This process is continued until a desired performance is achieved.

This process is expected to grow as n^q (where n is the number of design variables and q is the number of states for each design variable). Consider a simple problem to be optimized with respect to three independent design variables, and each variable has three possible design

states. Also, suppose that each call of field analysis program requires one-tenth of a CPU second on a computer. Thus, the analysis program is called $3^3 = 9$ times to evaluate the objective function for all possible design states requiring a total of 0.9 CPU second on the computer. This would probably be considered as an economical and efficient solution. However, most practical design problems typically require as many as 10 variables and 10 design states for each variable. In this case, a total of 10^{10} analysis calls are made to evaluate the objective for all possible design states. Suppose a relatively more accurate field analysis is required for these problems, and each analysis call requires one second. It will be expected 320 years of computation on the same computer! Despite the fact that these possible trial design states are eliminated by an experienced design engineer or an expert system using some knowledge-based heuristics, the process is computationally exhaustive and its use is still impractical for complex engineering problems.

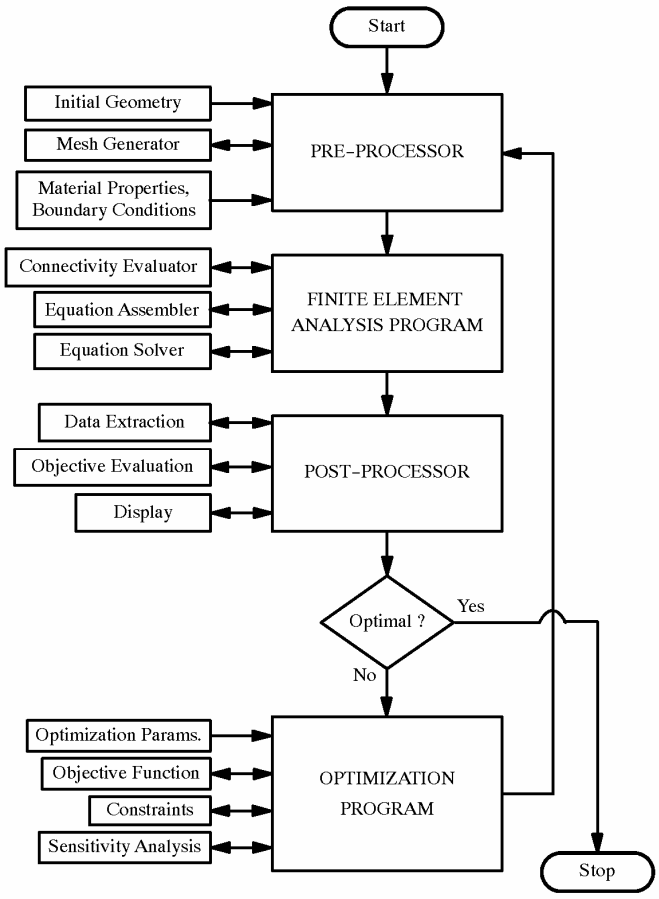

Figure 1. Automated Design Process

3. AUTOMATED OPTIMAL DESIGN

In recent years, traditional design procedures have been automated using pre- and post-processing modules cooperating with commercially available field analysis and numerical optimization

software. The principle of an automated optimal design (AOD) system is based on a "tight" integration of the modules into an iterative loop in which the unknown design parameters are progressively updated to advance the design towards the optimum solution. The block diagram in Figure 1 shows the relations among the basic modules during the iterative loop is in progress. Starting with a given initial geometry, the pre-processor module generates the initial mesh data used in the finite element analysis module. After the field potential is accurately solved, the objective function is evaluated using the field or other quantities computed within the post-processing module. If the desired objective is not satisfied, the field solution is passed to the optimization module to perform a line search procedure to obtain an optimum point along the gradient direction computed by the design sensitivity analysis. Once the new values of the design variables are computed, the mesh coordinates are updated, and the new quantities are computed to check if the desired performance is achieved. If the new result is not satisfactory, then iterations are carried out until a satisfactory result is attained.

While traditional optimal design procedures require substantial amount of computational time and human-computer interaction, AOD systems using numerical optimization techniques offer a logical strategy to approach the optimal solution in a systematic way. Although the discipline of nonlinear programming is well established, applications to real engineering systems are quite new and engineering ingenuity is required for adaptation of these techniques to a variety of design problems. In the remainder of this paper, methodologies for developing a computer tool for design optimization of magnetic systems will be discussed.

4. FINITE ELEMENT MODEL OF MAGNETOSTATIC FIELD

4.1. Field Equation

The general form of the governing partial differential equation of a magnetostatic problem is derived from Maxwell's equations. Neglecting the high frequency effects, consider

$$\nabla \times (\nu \, \nabla \times \vec{\mathbf{A}}) = \vec{\mathbf{J}} \,, \tag{1}$$

where displacement currents are neglected (i.e., no energy is stored in electric field). This equation represents the most general case of magnetostatic phenomenon which takes place in mediums with nonlinear magnetic reluctivity characteristics. Since this study focuses strictly on two-dimensional cases, the current density vector, $\vec{\mathbf{J}}$, possesses only the longitudinal (z-directed) component. Thus, Equation (1) is reduced to the scalar nonlinear Poisson's equation

$$\nabla \cdot (\nu \, \nabla A) = -J \,, \tag{2}$$

where $A(x, y)$ and $J(x, y)$ denote the z-directed components of $\vec{\mathbf{A}}$ and $\vec{\mathbf{J}}$ respectively.

4.2. Local Element Matrices

For numerical solution of Equation (2), FEM can be furnished based on variational principles by which the correct potential minimizes the energy functional [53]:

$$\Im(A) = \iint_{S(\Omega)} \left(\tfrac{1}{2} \int_0^{B^2} \nu(B^2) \; \mathrm{d}(B^2) - J \cdot A \right) \mathrm{dS} \,, \tag{3}$$

where Ω is the problem domain in which potential field takes place.

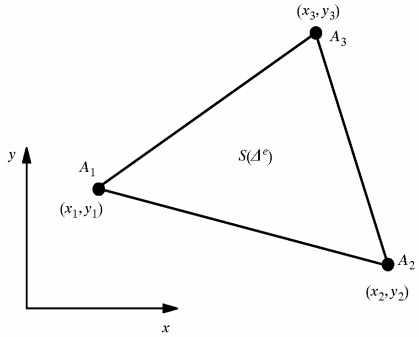

Figure 2. A triangular element with three vertex coordinates

First, the problem region is discretized into two-dimensional triangular elements (see Figure 2) and the potential function inside each element is approximated by a linear polynomial

$$A(x,y) = \alpha_1 + \alpha_2 x + \alpha_3 y \,, \tag{4}$$

where the α 's are the coefficients of this polynomial and defined as

$$\begin{aligned} \alpha_1 &= \frac{1}{2\Delta^e}\left(a_1 A_1 + a_2 A_2 + a_3 A_3\right) \\ \alpha_2 &= \frac{1}{2\Delta^e}\left(b_1 A_1 + b_2 A_2 + b_3 A_3\right) \\ \alpha_3 &= \frac{1}{2\Delta^e}\left(c_1 A_1 + c_2 A_2 + c_3 A_3\right), \end{aligned} \tag{5}$$

where A_1 , A_2 and A_3 are the vector potentials on the vertices of the triangular element, and Δ^e is the area of this element and is defined by

$$2\Delta^e = a_1 + a_2 + a_3 = b_1 c_2 - b_2 c_1 \,. \tag{6}$$

The geometrical coefficients, *a*'s *b*'s and *c*'s, are defined as

$$\begin{aligned} a_1 &= x_2 y_3 - x_3 y_2 & b_1 &= y_2 - y_3 & c_1 &= x_3 - x_2 \\ a_2 &= x_3 y_1 - x_1 y_3 & b_2 &= y_3 - y_1 & c_2 &= x_1 - x_3 \\ a_3 &= x_1 y_2 - x_2 y_1 & b_3 &= y_1 - y_2 & c_3 &= x_2 - x_1 \,, \end{aligned} \tag{7}$$

where *x*'s and *y*'s are the coordinates of the element vertices (see Figure 2). Thus, the components of the element flux density are calculated by

$$B^2 = \alpha_2^2 + \alpha_3^2, \quad \text{where} \quad B_x = \frac{\partial A}{\partial y} = \alpha_3; \quad B_y = -\frac{\partial A}{\partial x} = -\alpha_2 \tag{8}$$

which implies that the flux density is constant throughout the element. The energy functional given in Equation (3) can be written for a triangular element

$$\Im^e(A) = \iint_{S(\Delta^e)} \left(\tfrac{1}{2} \int_0^{B^2} \nu(B^2)\, \mathrm{d}(B^2) - J^e A \right) \mathrm{d}S \,. \tag{9}$$

To apply the minimization principle, the partial derivatives of Equation (9) with respect to the potentials at the vertices of this triangular element are set to zero

$$\frac{\partial \Im^e(A)}{\partial A_k} = 0 \quad \text{for} \quad (k = 1,2,3) \,. \tag{10}$$

Then, applying the differentiation into Equation (9) directly yields the following integral equation

$$\frac{1}{2}\iint_{S(\Delta^e)} \nu(B^2)\frac{\partial B^2}{\partial A_k}\mathrm{d}x\mathrm{d}y = \iint_{S(\Delta^e)} J^e \frac{\partial A}{\partial A_k}\mathrm{d}x\mathrm{d}y \,. \tag{11}$$

Obtaining the closed form evaluation of integrals in Equation (11) for each nodes (i.e.,) results in a set of three nonlinear equations

$$\mathbf{K}^e_{3\times 3}(\mathbf{A}^e_{3\times 1})\mathbf{A}^e_{3\times 1} = \mathbf{F}^e_{3\times 1} \,, \tag{12}$$

where $\mathbf{K}^e$ is the element stiffness matrix, $\mathbf{A}^e$ is the vector potentials at the element vertices, and $\mathbf{F}^e$ is the source vector representing the excitation sources. The entries of the element stiffness matrix in Equation (12) can be written more explicitly as

$$\mathbf{K}^e_{3\times 3} = \nu(B^2)\mathbf{P}^e_{3\times 3} \,, \tag{13}$$

where $\mathbf{P}^e$ is the element geometric coefficient matrix whose entries are calculated by

$$P^e_{ij} = \frac{1}{4\Delta^e}\left(b_i b_j + c_i c_j\right) \tag{14}$$

and the entries of the source vector are calculated by

$$F^e_i = \frac{\Delta^e}{3}J^e \,, \tag{15}$$

where J^e denotes the element current density and is assumed constant throughout this element.

4.3. Global Field Equations

The total energy functional is calculated by summing the individual contributions of *NE* elements inside the problem region. Thus, the total energy inside the system is expressed by

$$\Im(A) = \sum_{e=1}^{NE}\int_{\Omega_e}\left(\tfrac{1}{2}\int_0^{B^2}\nu(B^2)\ \mathrm{d}(B^2) - J^e A\right)\mathrm{d}\Omega_e, \tag{16}$$

where Ω_e denotes the domain of the element-e and *B* denotes the magnitude of the magnetic flux density inside this element. The functional in Equation (16) is a function of all vector potentials at the nodes of the triangular elements defined. Finally, minimization of the energy functional with respect to the nodal vector potentials (i.e., $\partial\Im(A_1, A_2\cdots A_N)/\partial A_i$ for $i = 1,2,\cdots,N$) results in the well-known nonlinear global field equations of the finite element method

$$\mathbf{K}_{N\times N}(\mathbf{A}_{N\times 1})\mathbf{A}_{N\times 1} = \mathbf{F}_{N\times 1}, \tag{17}$$

where $\mathbf{K}$ is the global stiffness matrix, $\mathbf{A}$ includes the vector potentials of *N* nodes, and $\mathbf{F}$ is the source term due to applied excitation currents to create the magnetic field in the system. The global stiffness matrix in Equation (17) is a function of the magnetic vector potentials because of the nonlinear reluctivity characteristics of saturable iron parts.

5. THE OPTIMAL DESIGN PROBLEM

This section is concerned with presenting the basic mathematical steps in formulating the optimization algorithm proposed in [5]. First, the optimization problem will be described in a standard form. It will be shown how the objective function and the constraint equations are normalized to improve the condition of the optimization problem. Then, the constrained problem is transformed to an unconstrained problem by forming the augmented Lagrange function as described in [5]. The optimality conditions are imposed to the augmented Lagrange function to find a minimizer for the unconstrained optimization problem. The Lagrange-Newton equations are obtained by linearizing the nonlinear equations from the necessary conditions.

The iterative optimization procedure developed in this section alters the geometry of certain desired parts of the magnetic system until the calculated flux densities are approximately matched at selected locations. The desired performance is obtained by minimizing the objective function expressed in the least-squares sense

$$\Phi(\mathbf{A},\mathbf{d}) = \sum_{\ell=1}^{M} \left[B_{c,\ell}(\mathbf{A},\mathbf{d}) - B_{s,\ell} \right]^2 , \tag{18}$$

where $\boldsymbol{B}_{s,\ell}$ and $\boldsymbol{B}_{c,\ell}$ denote the specified and calculated magnetic flux densities at M test points, $\mathbf{d}$ and $\mathbf{A}$ denote the nodal displacements, the unknown deflections of selected nodes on the optimized geometry from the initial device geometry, respectively.

During the iterative process, it is required that the field equations be satisfied as the geometry is altered. By adding the field equations given in Equation (17) to the optimization problem as nonlinear equality constraints, the objective function becomes a function of both the geometric parameters and the magnetic vector potentials. On the other hand, the geometric design parameters are subjected to some constraints in the problem region (e.g., the excitation coils cannot be larger than some dimensions, and the excitation currents are limited by some magnitude due to thermal constraints). With these assumptions, the optimization problem is expressed as minimization of an objective function subject to a set of nonlinear constraint equations:

$$\text{minimize}: \quad \Phi(\mathbf{A},\mathbf{d}) = \sum_{\ell=1}^{M} \left[B_{c,\ell}(\mathbf{A},\mathbf{d}) - B_{s,\ell} \right]^2 \tag{19}$$

$$\text{subject to}: \quad \sum_{n=1}^{N} K_{kn}(\mathbf{A},\mathbf{d}) A_n - F_k(\mathbf{d}) = 0 \quad \text{for} \quad k = 1,\cdots,N \tag{20}$$

$$d_k^L \le d_k \le d_k^U \quad \text{for} \quad k = 1,\cdots,DF , \tag{21}$$

where d_k^L and d_k^U are lower and upper bounds of the nodal displacement vector, d_k, assigned to some selected nodes on the device geometry to be optimized.

5.1. Normalization Of Objective Function And Constraints

The objective and the constraint functions in Equations (19) and (21) are of different dimensions. These functions directly depend on magnitude of excitation sources (such as applied excitation current density in the field windings) and the scale in which the physical device dimensions is defined (in this study all dimensions are defined in meter). The dimensions of the objective function in Equation (19), the residual of global finite element equations in Equation (20), and the side constraints in Equation (21) are the square of Tesla (Tesla is the dimension of magnetic flux density), Ampere, and meter, respectively. For example, if the applied field current is doubled (ignoring the effects of saturation in iron materials), the objective function quadruples and the residual of the finite element field equations doubles, while the inequality constraints in Equation (21) will remain the same because they directly depend on the scale of the device dimensions defined. This unbalance among the different functions causes the following undesired effects:

- numerical difficulties in solving the constrained optimization problem when a constraint function or the objective function dominates the optimization process;
- dependence of the objective function on physical device parameters prevents the user from correctly interpreting optimization performance computed for different excitation currents.

Objective Function. A proper normalization of the objective function and the constraint equations is therefore necessary to improve the conditioning of the unconstrained minimization process in augmented Lagrange multiplier method.

One way for normalizing the least-squares error or the objective function in Equation (19) is to divide each individual contribution by the corresponding $B_{s,\ell}$. This normalization

method is not always numerically suitable because it is limited to a specific case that all values must be different from zero. Alternatively, a better normalization is to multiply Equation (19) by a normalization factor

$$\Phi(\mathbf{A},\mathbf{d}) = k_{\Phi}\sum_{\ell=1}^{M}\left[B_{c,\ell}(\mathbf{A},\mathbf{d}) - B_{s,\ell}\right]^2 \tag{22}$$

where the normalization factor for the objective is calculated by $k_{\Phi} = \sum_{1}^{M} B_{s,\ell}^2$. The objective function becomes dimensionless since it is normalized relative to the squared summation of specified flux densities. The form of normalized square root of the error function ($\sqrt{\Phi}$) indicates the relative average error per test point and it will be used in all results demonstrated in this paper.

Field Equations. The global finite element field equations in Equation (20) are normalized by multiplying the overall equations by a constant

$$\mathrm{k}_f\left[\sum_{n=1}^{N} K_{kn}(\mathbf{A},\mathbf{d})A_n - F_k(\mathbf{d}) = 0\right] \quad \text{for} \quad k = 1,\cdots,N \tag{23}$$

where k_f is calculated by $k_f = 1/\max[F_i]$ (where $1 \le i \le N$) from the assembled global source vector $\mathbf{F}_{N\times 1}$. Therefore, the dependency of the residual of the global finite element equations on the applied currents is eliminated.

Costraints. To express the optimization problem in the form of a standard constrained optimization problem, the linear, double-sided inequality constraints in Equation (21) are converted into a set of quadratic but single-sided inequalities as

$$d_k^L \le d_k \le d_k^U \quad \rightarrow \quad g_k(d_k) \equiv \frac{\left(d_k^L - d_k\right)\left(d_k^U - d_k\right)}{\left(d_k^U - d_k^L\right)^2} \tag{24}$$

This inequality constraint function on the right side of Equation (24) has two valuable properties: the first, it takes a negative value as long as the nodal displacement stays between d_k^L and d_k^U; the second, it is dimensionless since it is divided by the term $(d_k^U - d_k^L)^2$.

5.2. Augmented Lagrange Function

Considering the description of the augmented Lagrange function and the objective function and the set of constraint equations, the augmented Lagrange function is formed by adding the equality and the inequality constraints to the objective function as

$$\begin{aligned} L_A(\mathbf{A},\mathbf{d},\boldsymbol{\lambda},\boldsymbol{\mu},\mathbf{r}_p) = {} & k_{\Phi}\sum_{\ell=1}^{M}\left[B_{c,\ell}(\mathbf{A},\mathbf{d}) - B_{s,\ell}\right]^2 + k_f\sum_{k=1}^{N}\lambda_k\left[\sum_{n=1}^{N} K_{kn}(\mathbf{A},\mathbf{d})A_n - F_k(\mathbf{d})\right] \\ & + \left[\sum_{k=1}^{DF}\mu_k\theta_k + r_p\theta_k^2\right], \end{aligned} \tag{25}$$

where DF denotes the total number of nodal displacements assigned to the geometry to be optimized, λ_k and μ_k denote the Lagrange multipliers corresponding to the equality and inequality constraints, r_p is the penalty multiplier for the inequality constraints used in the augmented Lagrange multiplier method, and θ_k is the augmented inequality function expressed as

$$\theta_k = \max\left[g_k(d_k) \equiv \frac{\left(d_k^L - d_k\right)\left(d_k^U - d_k\right)}{\left(d_k^U - d_k^L\right)^2}, \quad -\frac{\mu_k}{2r_p}\right]. \tag{26}$$

Note that the equality constraints from the finite element field equations are not augmented to avoid complex second-order derivative calculations of the squared equality constraints. Thus, the Lagrange multipliers associated with the equality constraint equations are calculated in the iterative optimization process.

5.3. First-Order Necessary Conditions

The augmented Lagrange function given in Equation (25) is a nonlinear function of the magnetic vector potentials and the unknown geometric design parameters describing the optimized device geometry. This standard optimization problem with nonlinear constraints is widely studied in the field of mathematical programming. The solution is obtained by setting the optimality conditions on the Lagrangian. Then, the system of resulting nonlinear equations is linearized, and the unknown variables describing the geometry and the magnetic field are obtained by the Newton-Raphson method.

To minimize the unconstrained optimization problem in Equation (25), unknown variables are comprised in a vector

$$\mathbf{X} = \left\{\mathbf{A}, \mathbf{d}, \boldsymbol{\lambda}\right\}^{\mathrm{T}} \tag{27}$$

Let the optimal solution be $\mathbf{X}^* = \left\{\mathbf{A}^*, \mathbf{d}^*, \boldsymbol{\lambda}^*\right\}^{\mathrm{T}}$. Then, the stationary points, $\mathbf{X}^*$, are obtained by applying the first order necessary conditions. These conditions are thus met by taking the first partial derivatives with respect to the unknown variables and setting them equal to zero (i.e., $\nabla L(\mathbf{X}^*) = 0$):

$$f_{1,i} = \frac{\partial L_A}{\partial \lambda_i} = k_f \left[\sum_{n=1}^{N} K_{kn}(\mathbf{A}, \mathbf{d}) A_n - F_k(\mathbf{d}) \right] = 0 \tag{28}$$

$$f_{2,i} = \frac{\partial L_A}{\partial A_i} = 2k_\Phi \sum_{\ell=1}^{M} \left[B_{c,\ell}(\mathbf{A}, \mathbf{d}) - B_{s,\ell} \right] \frac{\partial B_{c,\ell}}{\partial A_i} + k_f \sum_{k=1}^{N} \lambda_k \left[K_{ki} + \sum_{n=1}^{N} \frac{\partial K_{km}}{\partial A_i} A_m \right] = 0 \tag{29}$$

$$f_{3,i} = \frac{\partial L_A}{\partial d_i} = 2k_\Phi \sum_{\ell=1}^{M} \left[B_{c,\ell}(\mathbf{A}, \mathbf{d}) - B_{s,\ell} \right] \frac{\partial B_{c,\ell}}{\partial d_i} + k_f \sum_{k=1}^{N} \lambda_k \left[\sum_{n=1}^{N} \frac{\partial K_{km}}{\partial d_i} A_m - \frac{\partial F_k}{\partial d_i} \right]$$
$$+ \delta_i \left[\mu_i + 2r_p \left(\frac{\left(d_i^L - d_i\right)\left(d_i^U - d_i\right)}{\left(d_i^U - d_i^L\right)^2} \right) \right] \left(\frac{2d_i - \left(d_i^U + d_i^L\right)}{\left(d_i^U - d_i^L\right)^2} \right) = 0 \tag{30}$$

The $2N + DF$ nonlinear expressions obtained in Equations (28)-(30) are explicitly expressed in terms of the finite element matrices, the magnetic field quantities, and their derivatives as:

$$\begin{aligned} &\frac{\partial L_A}{\partial \lambda_i} = 0 \quad \text{for} \quad i = 1, \cdots, N \\ &\frac{\partial L_A}{\partial A_i} = 0 \quad \text{for} \quad i = 1, \cdots, N \\ &\frac{\partial L_A}{\partial d_i} = 0 \quad \text{for} \quad i = 1, \cdots, DF \end{aligned} \tag{31}$$

where an indicator flag is defined as

$$\delta_k = \begin{cases} 1 & \text{if} \quad g_k(d_k) > -\dfrac{\mu_k}{2r_p} \\ 0 & \text{otherwise} \end{cases}. \tag{32}$$

The value of this flag takes a value of either 1 or 0 depending on the state of the augmented inequality constraint given in Equation (26). If the inequality constraint becomes active, its derivatives with respect to the design parameters are nonzero [17].

The resulting nonlinear equation system can be uniquely solved for independent variables for an initial starting point $\mathbf{X}^0 = \{\mathbf{A}^0, \mathbf{d}^0, \boldsymbol{\lambda}^0\}^T$. It should be noted that the initial estimate for the optimization variables may be very critical depending on the problem. In the case of multiple solutions, convergence to the global optimum solution is not guaranteed unless the initial guess is close enough to the optimal point. Otherwise, the solution may be trapped at a local optimum solution if multiple solutions exist.

5.4. Lagrange-Newton Equations

The multi-dimensional system given in Equations (28)-(30) is nonlinear and it is necessary to linearize it for the iterative solution. Given the initial solution $\mathbf{X}^k$, the solution for the next iteration is determined from the multi-dimensional Taylor's expansion by keeping only the linear terms

$$\nabla L_A(\mathbf{X}) = \nabla L_A(\mathbf{X}^k) + \nabla^2 L_A(\mathbf{X}^k)\Delta\mathbf{X}^k = \mathbf{0}, \tag{33}$$

where the residual term is defined as $\Delta\mathbf{X}^k = \mathbf{X}^{k+1} - \mathbf{X}^k$, and obtained from the solution of the linear system

$$\nabla^2 L_A(\mathbf{X}^k)\Delta\mathbf{X}^k = -\nabla L_A(\mathbf{X}^k). \tag{34}$$

The right hand side matrix, $\nabla^2 L_A(\mathbf{X}^k)$, is called the Hessian of the Lagrangian and the right hand side is obtained by evaluating the gradients of the augmented Lagrange function as defined in Equations (28)-(30). If the maximum entry of the residual term is greater than a tolerance, the solution for the the *k*-th iteration is updated by $\mathbf{X}^{k+1} = \mathbf{X}^k + \Delta\mathbf{X}^k$.

If the set of equations is arranged and the necessary differentiations are obtained, the residual vector $\Delta\mathbf{X}^k$ for the optimization variables is calculated by solving the non-symmetric sparse linear system

$$\begin{bmatrix} \mathbf{S}_{N\times N} & \mathbf{0}_{N\times N} & \mathbf{D}_{N\times DF} \\ \mathbf{E}_{N\times N} & \mathbf{F}_{N\times N} & \mathbf{G}_{N\times DF} \\ \mathbf{H}_{DF\times N} & \mathbf{D}^T_{DF\times N} & \mathbf{J}_{DF\times DF} \end{bmatrix} \begin{Bmatrix} \Delta\mathbf{A}_{N\times 1} \\ \Delta\boldsymbol{\lambda}_{N\times 1} \\ \Delta\mathbf{d}_{DF\times 1} \end{Bmatrix} = \begin{Bmatrix} \mathbf{f}_{1N\times 1} \\ \mathbf{f}_{2N\times 1} \\ \mathbf{f}_{3N\times 1} \end{Bmatrix}, \tag{35}$$

Where, the block matrix form in Equation (35) is the Hessian matrix which includes the information for the size and the direction of the decreasing gradient vector in the optimization procedure. The block matrices are sparse and stored by an uncompressed pointer storage scheme in one-dimensional arrays. Detailed formulations for block matrices can be found in Ovacik [54].

6. PARAMETRIZATION OF DEVICE GEOMETRY

The task of optimization is to iteratively calculate the new values of the selected optimization parameters to modify the actual shape of the device, and therefore to minimize the objective function in order to achieve a desired performance. Once the values of the design parameters are calculated, the geometry is modified by small increments from its previous shape to a new shape, and new iterations are performed until these incremental changes do not significantly effect the variation of the objective function. Perhaps, one of the most difficult problems in shape optimization problems is to link the geometric optimization parameters to the physical device geometry described by the certain physical dimensions and the contours of the device (the

material interfaces, such as iron-to-air or copper-to-air). This section will explain the parametrization method used in the development of the optimization algorithm.

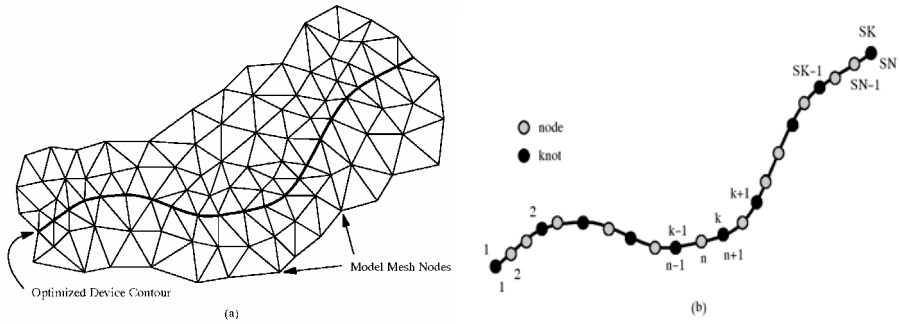

Figure 3. Modeling the optimized contour: a) triangular mesh and device contour; b) numbering nodes on the device contour and the user defined knots

As the new values of the geometric parameters are calculated, the device geometry changes and thus the new finite element mesh has to be generated for the next iteration. For this purpose, an automatic mesh generation algorithm is called whenever the device geometry is modified (Subramaniam et al. [23]). This however has certain drawbacks, especially when the gradient error due to the discretization is significantly large in regions containing coarse elements (Weeber and Hoole [18]). In this case, the convergence behavior of the optimization algorithm is significantly influenced by the discretization error due to discretization of the problem domain. This may jeopardize obtaining smoothly converging results since spikes on the objective function gradient changes the decreasing direction of optimization. The previous work showed that this problem can be greatly circumvented by maintaining the same mesh topology during the iterative modification of the geometry (Pironneau [4], Haslinger and Neittaanmaki [6], Weeber and Hoole [18], and Weeber [17]). The nodal coordinates of the finite element mesh are mapped to the geometric parameters. Thus, mesh nodes are smoothly moved from one position to another ensuring that the discretization error smoothly changes during this procedure.

During the iterative modification of the device geometry, four types of mesh nodes are considered in terms of restrictions on their moving abilities:

- *Principal nodes:* nodes on the device contour which describes the optimized geometry of the device. Principal nodes are allowed to move only in the direction of assigned displacements;
- *Associated nodes:* mesh nodes which are selected by the user and are critical for the geometry modification; they move along with the deflected surface in any direction (unless the restrictions are specified by the user) so as to prevent any possible excessive element deformation or overlapping in the finite element region;
- *Constrained nodes:* nodes which are constrained either horizontally or vertically to avoid any possible violation of device's physical geometry.
- *Fixed nodes:* nodes which are either far from the deflected surface (not quite influenced by the surface deflection) or on the fixed device boundaries which are not allowed to move in any direction to avoid any unwanted alteration of the problem geometry.

6.1. Optimized Device Contour

To explain how the optimization geometry is modeled, consider the optimized geometry and the finite element mesh nodes shown in Figure 3(a). There are SN nodes on the optimized portion of

the device geometry, and some of them are directly linked to the shape optimization parameters (the nodal displacements as defined earlier). The displacements directly control the geometry of the device while altering the geometry (see Figure 3(b)). The displacements are applied to some of the user-selected SK knots (principal nodes) on the surface. Since the optimized surface may be complicated in most problems, these knots are chosen from these nodes, defining the optimized geometry. The remaining $SN - SK$ nodes on the geometry surface are simply the associated nodes whose displacement weights are determined from the weights of the knots by cubic-spline interpolation. Once the incremental displacements applied to the surface nodes the surface is deflected. The total deflection in the *p*-direction is expressed by

$$\Delta D_p(s) = \sum_{i=1}^{SK} Q_{i(p)} \Delta d_{i(p)} \tag{36}$$

where s is the parametric distance on the surface, $Q_{p,i}(s)$ is the basis of the displacement weight associated to the *i*-th knot on the surface, and $\Delta D_p(s)$ is the total deflection of the surface after applying the incremental nodal displacements Δd_p to the knots. Figure 4 shows the basis functions associated with 5 knots selected out of 41 surface nodes on an optimized surface. The weights take unit values at the associated knot and vanish on the other knots. For the surface nodes, the displacement weights associated to each displacement is calculated from the basis functions as

$$\beta_{i(p)} = Q(s_i) \tag{37}$$

Here, using cubic spline interpolation is found to be a good choice to smooth the "jagged" contours on the optimized geometry. As the new nodal displacements are updated the knots are moved. Therefore, the independent nodal displacements on the optimized surface are linked to the internal mesh nodes in such a way that any displacements of the surface nodes also effect the nodes of the sub-region.

6.2. Mesh Coordinates

The finite element matrices and the magnetic flux densities inside the elements are functions of some geometric coefficients expressed in terms of the vertex coordinates of elements. If the optimization surface is deflected, the elements inside a certain region are deformed. Therefore, the element matrices and the magnetic flux densities change with the deflection parameters. This section provides an estimated analytical expression between the mesh coordinates and the nodal displacement parameters applied to the optimized surface. The relationship between the displacements and the coordinates of the finite element mesh is nonlinear since the nodes nearby the deflected region move more than the ones far from the surface. The expression suggested is used for closed form derivative calculations and is valid for small displacements.

The described mesh deformation methodology uses a fixed mesh topology during the iterative modification of the device geometry. As the new values of nodal displacements are calculated, the optimization surface is deflected, and the internal nodes in the user-defined sub-region move along with the surface nodes. To avoid any overlapping elements, values between 0 and 1 are assigned to the displacement weights, in such a way that the weights smoothly decrease as the distance of the nodes to the deflected surface increase (while the weights of nodes outside the sub-region remain zero). The mesh coordinates of the k-th iteration are calculated from the coordinates of the initial mesh plus the sum of the products of nodal weights and corresponding displacements in that direction:

$$x_p^k = x_p^{k-1} + \sum_{i=1}^{DF_p} \beta_{i(p)}^k \Delta d_{i(p)}^k, \tag{38}$$

where DF_p is the number of nodal displacements in the p-direction, $\beta^k_{i(p)}$ is the displacement weight of the *i*-th node. In every iteration the displacement weight, and $\Delta d_{i(p)}$, the computed increment for the displacement associated with each nodal displacement, is first determined, and then the relation between the new mesh nodes and the old mesh nodes is expressed in terms of the new coordinates of the new mesh. Therefore, in calculation of the derivative terms in the Hessian from Equation (35), the partial derivative of a geometry-dependent function $f(\mathbf{d}_p, \mathbf{x}_p)$ with respect to the displacement $d_{i(p)}$ is obtained by the chain-rule differentiation [55]-[57]:

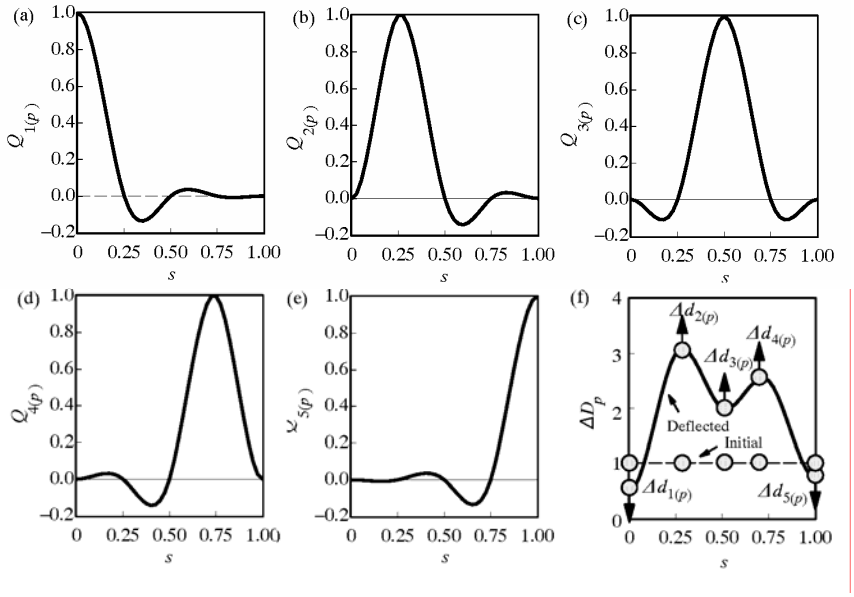

Figure 4. Modeling a five-knot device contour using basis functions: (a)-(e) the cubic-spline basis functions of each knot; (f) the initial contour (dashed line) is deflected after applying incremental displacements to each knot

$$\frac{df}{dd_{i(p)}} = \frac{\partial f}{\partial d_{i(p)}} + \sum_{j=1}^{N_L} \frac{\partial f}{\partial \bar{x}_{j(p)}} \frac{\partial \bar{x}_{j(p)}}{\partial d_{i(p)}}, \tag{39}$$

where the vector $\bar{x}_p$ contains N_L element nodes from x_p (i.e., $\bar{x}_p \in R^{N_L} \subset x_p$), and f is any function which may be the flux density or the local finite element matrix entry of an element. The term $\partial \bar{x}_{j(p)} / \partial d_{i(p)}$ can be obtained by directly differentiating Equation (38).

The nodal displacement weights are unknown for each moving node in the finite element region. They should be assigned before each Newton-Raphson iteration is performed. A quick way to approximately determine these weights is to define a strategy such that the displacement weights assigned on the optimization surface and the restrictions are applied on the nodes in the region and the displacement weights on the associated nodes are determined from basis functions applied to the nodes on the optimized geometry. The following describes the outline of this algorithm.

Step 1. Read the coordinates of the user defined frame, the global node numbers of the surface nodes which describe the deflected surface, direction of the displacements, indices of the selected knots from the surface nodes, the knot indices assigned to the displacements and the global node numbers describing the fixed geometries in the user defined geometries.

Step 2. Initialize the displacement weights of the nodes.

Step 3. Assign unit value of displacement weights to the knots which are associated with the displacements, and interpolate the displacement weights for the other surface nodes which are not selected as knots.

Step 4. Iteratively calculate the displacement weights of the other nodes which are neither a surface node nor a node describing the fixed geometries. Iterate for the nodal displacement weights which do not belong to the displaced surface as well as the nodes which belong to the fixed geometries.

In Step 4, the displacement weights of the associated nodes are determined from the displacement weights of its neighboring nodes as illustrated in Figure 5. These values are calculated by weighted averages considering the geometric distances of the neighboring nodes

$$\beta_{o(p)} \sum_{i=1}^{NG} w_{i(p)} \beta_{i(p)} , \tag{40}$$

where the term $w_{i(p)}$ denotes the geometric weight and is calculated by

$$w_{i(p)} = \frac{\ell_{i(p)}}{\sum_{i=1}^{NG} \ell_{i(p)}} . \tag{41}$$

The iterative procedure in Step 4 is continued until a minor change occurs between the displacement weights obtained in two successive iterations.

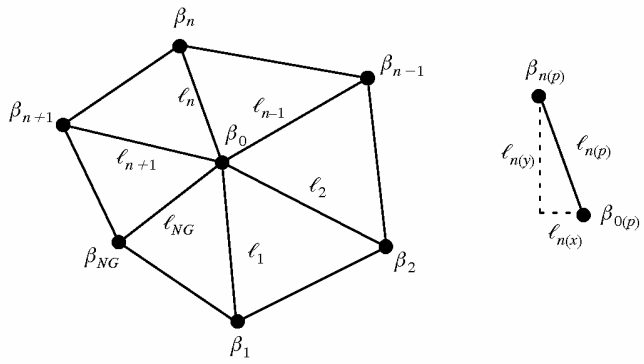

Figure 5. Iterative scheme to calculate displacement weights of free-moving nodes from those of their neighbors

6.3. Updating Mesh Coordinates

The robustness and efficiency of numerical shape optimization algorithms strongly depend on the geometric mapping of electromagnetic problem (Weeber [17] and Weeber and Hoole [18]). Once

the incremental displacements are computed within the Newton-Raphson procedure, the next step is to proceed towards the optimal design by applying these changes to the current device geometry. In this step, care has to be taken that the new coordinates of the nodes on the optimized device contour and the associated nodes in the finite element mesh are properly deflected without violating the constraints for the device's physical dimensions as well as without forming any overlapping or excessively distorted elements. Then, this deflected mesh structure is used in subsequent Newton-Raphson iterations; thereby maintaining a constant mesh topology throughout the optimization process to avoid discontinuities in derivatives of both objective function and constraints caused by discretization error. Although limited but somewhat quite attractive for simple geometries, analytical mapping expressed in terms of nodal displacement weights was used earlier by Marrocco and Pironneau [59], and Istfan and Salon [8]. However, satisfying all constraints of complex geometries is avoidably difficult for general applications.

A numerical mapping technique handling more complicated geometries is used in the DOPT2D program. The mesh structure is deflected based on structural laws of elasticity using an elastic body analogy to the structural subdomain including the optimized device contour and its neighborhood where the shape of elements may be critically distorted due to changes in geometric design. Since the deflected geometry in the *x*-*y* plane does not vary along the *z*-direction (the axial direction of the optimized device), the *elasticity* problem can be treated as two-dimensional. Therefore, the structural deflection is described by displacement vector $\overline{\mathbf{U}}(x,y)$ with its components $u(x,y)$ and $v(x,y)$ in the planar coordinate directions *x* and *y*, respectively. The analysis of the plane stress problem is carried out using the finite element method. The total incremental deflection of the optimized contour is applied as boundary conditions to obtain displacements of associated nodes in the structural sub-problem region.

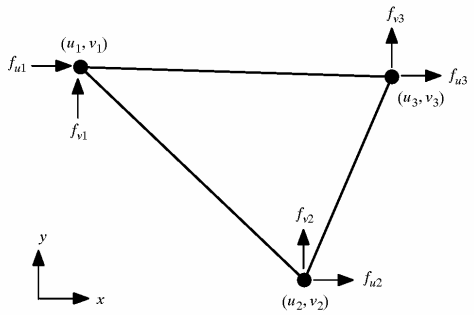

Figure 6. Deformation of an elastic element subject to applied point forces

Because of their nodal compatibility, the first-order triangular elements identical to those for the electromagnetic model are also used in the finite element model of the structural sub-problem. Consider the triangular element shown in Figure 6. The relation between the applied point forces and the resulting displacements at the element vertices is expressed by

$$\mathbf{S}^e_{6\times6}\mathbf{U}^e_{6\times1} = \mathbf{F}^e_{6\times1} \tag{42}$$

where $\mathbf{S}^e$ is the element stiffness matrix, $\mathbf{U}^e$ is the displacement vector and $\mathbf{F}^e$ is the nodal point source vector including the components in the directions of and

$$\mathbf{U}^e_{6\times1} = \left\{ u_1 \quad v_1 \quad u_2 \quad v_2 \quad u_3 \quad v_3 \right\}^{\mathrm{T}} ; \; F^e_{6\times1} = \left\{ f_{x1} \quad f_{y1} \quad f_{x2} \quad f_{y2} \quad f_{x3} \quad f_{y3} \right\}^{\mathrm{T}} \tag{43}$$

The derivation of the element matrices is not pursued in this paper. For the theory and a detailed formulation, the reader should refer to [53] and [61].

The mesh data including the vertex coordinates of triangular elements and connectivity indices are extracted from the global mesh data used for the electromagnetic model. To solve the structural sub-problem having N_s nodes, all the element matrices are calculated and then assembled to form the global system equations

$$\mathbf{S}_{N_S \times N_S} \mathbf{U}_{N_S \times 1} = \mathbf{F}_{N_S \times 1} . \tag{44}$$

Then, the boundary conditions are applied as either the specified point forces or the specified displacements. Using the latter is more advantageous because the device descriptive parameters regarding the deflected and the constrained geometry are defined as displacements. In this manner, the $\mathbf{U}$ vector of displacements is partitioned into the vector of specified displacements $\mathbf{U}_s$ and the vector of unknown displacements $\mathbf{U}_u$ so that the global equations are written as

$$\begin{bmatrix} \mathbf{S}_{uu} & \mathbf{S}_{us} \\ \mathbf{S}_{su} & \mathbf{S}_{ss} \end{bmatrix} \begin{Bmatrix} \mathbf{U}_u \\ \mathbf{U}_s \end{Bmatrix} = \begin{Bmatrix} \mathbf{0} \\ \mathbf{0} \end{Bmatrix}, \tag{45}$$

where the subscripts *u* and *s* represent the nodes of unknown and specified displacements. Thus, the state of the structural sub-problem is determined from Equation (45)

$$\mathbf{S}_{uu} \mathbf{U}_u = -\mathbf{S}_{us} \mathbf{U}_s . \tag{46}$$

The structural finite element analysis adopted herein is used for geometric mapping purposes and no emphasis has been placed on accurate solution of the elasticity problem. It should be pointed out that the element stiffness matrix in Equation (45) is derived based on variational principles with infinitesimal small displacements, assuming constant strain-displacement relation inside the elastic element. In actuality, however, the elasticity problem is *materially* nonlinear, i.e., element strain is not constant for large displacements. With this assumption, if large displacements are applied, serious numerical errors may occur in solution: unrealistically high strain energy is stored in the elements nearby the boundaries of displaced surface, causing uneven displacement distribution over the structural domain. This eventually will result in very distorted elements after some number of successive geometry modifications.

To avoid distorted elements, the elasticity problem is *linearized* by applying a fraction of the total deflection (say one part in 1000). Then, the algebraic system in Equation (46) is assembled and solved using the boundary condition. Typically, the applied boundary conditions $\mathbf{U}_s$ include:

- a fraction of the total deflection of primary nodes determined from Equation (36);
- the displacements of the nodes constrained by physical device geometry (displacement components in constrained directions are forced to be zero);
- displacements of the boundary nodes (fixed nodes) of the structural subproblem domain (displacement components in both directions are forced to be zero).

Once Equation (46) is solved for the unknown displacements $\mathbf{U}_u$ of the associated nodes, the obtained result is divided by this fraction and then applied to the mesh coordinates.

The DOPT2D program described in this paper utilizes the subroutines of the MODEL library (Akin [60]). The size of structural subproblem region is determined by the user based on his/her previous experience. It should be kept in mind however that this size should be chosen as small as possible to minimize the additional computational cost.

7. MODELING NONLINEAR MATERIAL CHARACTERISTICS

Since most devices are designed to operate in saturation, accurately modeling reluctivity characteristic of nonlinear ferromagnetic materials plays an important role in synthesis of magnetic problems. In most cases, even under normal operating conditions, magnetic devices are designed to operate in saturation region. Therefore, considering a linear magnetization curve in such devices is unrealistic. Mathematical modeling of such characteristics should be taken care of by an appropriate method.

Magnetic saturation may be understood graphically by examining a curve of magnetic flux density *B* versus magnetic field intensity, *H* as shown in Figure 7. In general, magnetic characteristics of materials are represented by magnetization curves relating the magnetic field intensity H, to that of magnetic flux density. These curves of different magnetic materials are experimentally determined and tabulated as a set of *B-H* curves in the manufacturer's catalogue.

Forming the finite element stiffness matrix (Equation (13)) requires material reluctivity values. The reluctivity characteristics used in the formulations are expressed in terms of the square of the magnitude of the flux density in each element as $\nu(B^2)$. During Newton-Raphson iterations, the reluctivity and its derivatives are repeatedly evaluated while assembling the local element matrices for computing complex calculations of the Jacobian and the Hessian matrices. Efficiently computing the reluctivity characteristic and its derivatives is a key element to reduce the computational cost.

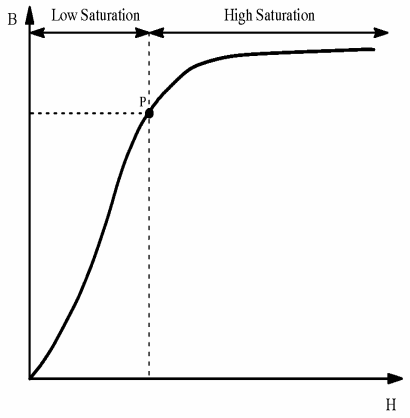

Figure 7. A Typical B-H characteristic curve of a nonlinear magnetic material with low- and high-saturation regions

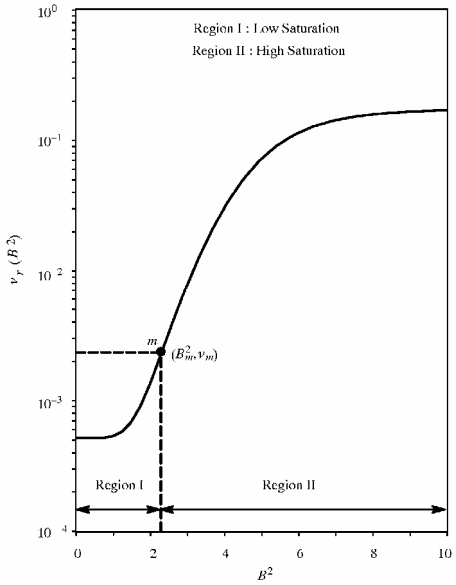

Figure 8. Variation of the nonlinear relative reluctivity in low-saturation (Region I) and high-saturation (Region II) regions

To guarantee a quadratic, smooth convergence to a unique solution of the optimization problem, reluctivity characteristic curves must be at least twice differentiable. There are several approximation methods offered in the literature. Using piece-wise cubic polynomials (cubic spline method) used by Silvester et al. [62] has become popular in solution of the forward

problem. Since the characteristic curve is represented by cubic polynomials within given segments, the derivatives may be oscillatory when large segments are selected. When a large number of smaller segments are used to represent the same curve, however, more effort is needed to find the associated segment for a given magnetic flux density. Series expansion methods used by El-Sherbiny [63] suggest a relatively more accurate, but more expensive-to-evaluate, analytical approach using summation of exponential functions. Using a single exponential function suggested in Hoole and Hoole [64] is relatively low in cost to evaluate but it can represent the curve in a limited range of saturation.

A typical reluctivity characteristic of a ferromagnetic material is shown in Figure 8. The relative reluctivity is expressed as a function of B^2 because it provides an easy differentiation to avoid relatively expensive square root operation. The relative reluctivity curve varies very slowly in the linear region, and then linearly increases in the moderate saturation level (Region I). In the high-saturation level (Region II), however, the linearly increasing curve slightly bends and keeps slowly increases to the limit value of $\nu_o = 1/\mu_o$.

There are several good reasons for using squared values of these independent variables B or H, rather than their magnitudes. First, these variables are usually derived from potentials in vector component form, so that finding the magnitude involves first finding the squares of the components and then extracting the square root of their sum. The relatively expensive square root operation is avoided in this way. The second, perhaps more important, is the stability problem: working with the square tends to emphasize the behavior of the curve at high flux densities or fields where higher precision is usually required.

The suggested model in this paper can only accurately represent the reluctivity characteristic up to a certain saturation level of in the low-saturation region. The fitting values begin to deviate from the actual values beyond a certain level of magnetization B_m. At this point, the reluctivity function

$$\nu_r(B^2) = 1 - (c_1 + c_2 B^2)\exp(-c_3 B^2) \tag{47}$$

is used for representation at high-saturation levels in Region II. Here, the coefficients c_1, c_2, and c_3 are determined by imposing the continuity conditions of the reluctivity function itself and its first and second derivatives at the intersection of the two models

$$\begin{aligned}
\nu_m &\equiv \nu_r(B^2)\big|_{B_m} = 1 - (c_1 - c_2 B_m^2)\exp(-c3B_m^2) \\
\nu'_m &\equiv \frac{\partial \nu_r(B^2)\big|}{\partial B^2}_{B_m} = \left[-c_2 + c_3(c_1 + c_2 B_m^2)\right]\exp(-c3B_m^2) \\
\nu''_m &\equiv \frac{\partial^2 \nu_r(B^2)\big|}{\partial (B^2)^2}_{B_m} = \left[2c_2c_3 - c_3^2(c_1 + c_2 B_m^2)\right]\exp(-c3B_m^2)
\end{aligned} \tag{48}$$

where the unknown coefficients c_1, c_2, and c_3 are determined by simultaneously solving Equation (48):

$$\begin{aligned}
c_1 &= 1 - \nu_m \exp(c_3 B_m^2) - c_2 B_m \\
c_2 &= \frac{c_3 \nu'_m + \nu_m}{c_3}\exp(c_3 B_m^2) - c_2 B_m \\
c_2 &= \frac{\nu'_m}{1-\nu_m}\sqrt{\left(\frac{\nu'_m}{1-\nu_m}\right)^2 + \frac{\nu''_m}{1-\nu_m}}\,.
\end{aligned} \tag{49}$$

Therefore, a more realistic model representing the nonlinear reluctivity characteristics of iron parts in both low- and high-saturation cases may be utilized in the solution of the field potentials.

8. OUTLINE OF THE DOPT2D PROGRAM

The flow diagram in Figure 9 shows the basic steps of the optimal design process controlled by the main program (DOPT2D). The main program cooperates with three modules: the pre-processor module readily available in the MICROFLUX program; the electromagnetic finite element analysis module (FORWARD); the structural finite element analysis module in the MODEL program library. The MICROFULX program is used for generating the data including: mesh, material, excitation current and material properties for the electromagnetic finite element model of the magnetostatic device to be optimized. The data generated are written in a file later read by the DOPT2D program. The FORWARD program is used to obtain a finite element solution for the magnetic vector potentials in the nonlinear magnetostatic problem: it is called prior to the optimization process for providing an initial solution to the vector potentials for the initial geometry. The MODEL program is used to solve structural subproblem described in Section 6.3 to compute the new coordinates of the moving nodes whenever a new device shape is calculated. All routines described herein are implemented using the standard FORTRAN-77 language.

The outline of the overall optimization process is as follows. The necessary data files are read from the DOPT2D program. These include the mesh coordinates, specified flux densities, their specified directions and the material properties (either linear or nonlinear). The preliminary computations are then carried out. In this step, element connectivity indices including the indices of neighboring elements and nodes are determined and stored in arrays. The parameters for nonlinear saturation characteristics of nonlinear materials in low- and high-saturation regions are computed. These parameters are later used in direct differentiation of the element matrices in forming the Lagrange-Newton equations.

In the nonlinear optimization process, the augmented Lagrange function is linearized and to obtain the system of equations to calculate the update of unknown variables including the design parameters, vector potentials and the Lagrange multipliers for the equality constraints (the field equations). Once the updates for the unknown displacements are computed, they are passed to the MODEL program to obtain new coordinates of the mesh nodes of the finite element model. The nonlinearity of the field equations are determined calculating the maximum residual of the global field equations $R_{\max}$. For severely saturated magnetostatic field, the computed updates for the vector potentials are not projected correctly using the updates calculated by the Newton-Raphson process. If the nonlinearity is severe (i.e., $R_{\max} > \varepsilon_R$), the geometry is updated and then the global finite element equations are solved using the FORWARD problem. These computed vector potentials are used to update the vector potentials. Once the vector potentials are updated for the new geometry, the flux densities at the test points are evaluated and the least-squares error is computed for the next iteration.

9. RESULTS AND DISCUSSUIONS

This section is concerned with the optimal design of synchronous machinery by using the DOPT2D program developed and implemented in this paper. The design objective used for the problems arises from the requirement of the air-gap flux density to vary sinusoidally along outer periphery of the airgap region. The objective function is minimized, in the sense of least-squares, with respect to the most sensitive geometric parameters subject to geometric constrains specified. The case studies undertaken for unsaturated and saturated salient-pole, and saturated round-rotor

synchronous generators will be presented to demonstrate that the DOPT2D program can be used as an optimal design tool for minimizing space harmonics in the air-gap flux waveforms.

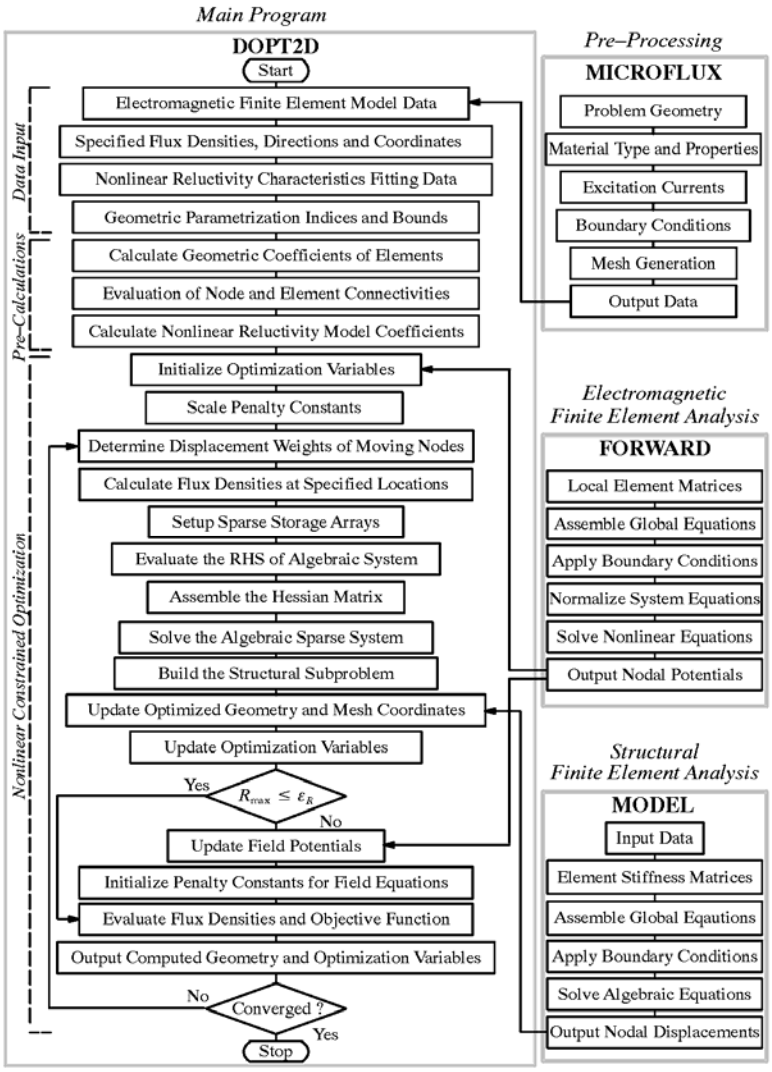

Figure 9. Flow diagram showing the basic steps of optimal design procedure

9.1. Salient-Pole Synchronous Generator

The full pole-pitch of a salient-pole machine in Figure 10 (a), including the stator, rotor, field-winding and air-gap regions, is used to demonstrate optimization of the pole face of salient-pole synchronous generators. The magnetic-circuit model used in this study is adopted from the previous work presented by Weeber [17], who originally described this model for testing constrained optimization algorithms. The geometric dimensions and the design parameters

defined in one-half of the pole pitch of a salient-pole synchronous generator are shown in Figure 10(b). The design of the pole geometry in this figure is described by six geometric dimensions: τ_p is one-half of the pole pitch; w_p is one-half of the pole width; δ_0 is the air gap at the middle of the pole; δ_t is the air-gap size at the pole tip; h_t is the height of the pole tip; δ_w is the gap between the pole and the field winding.

Design Objective and Problem Definition. The design objective of the optimal design problem is to synthesize the pole geometry to achieve a sinusoidally-distributed flux density along the air-gap line A-B in Figure 10 (b). In constructing the data for the optimization problem, before computing the geometric design parameters, similar procedures explained in [54] are carried out to construct the optimization problem. First, the problem domain is discretized using the triangular mesh generator of the MICROFLUX finite element analysis program. The first-order, triangular mesh including 622 elements and 340 nodes shown in Figure 11 is used for the finite element model. Then, six nodal displacements assigned to some of the 24 mesh nodes on pole-face contour for parametrization of the optimization geometry. The directions of the displacement parameters and the boundary conditions for the magnetic vector potential are shown in Figure 12. Finally, the y-component of magnetic field density, B_y, is specified at 31 points located in regular intervals on the line A-B, and calculated by using the expression $B_y = B_0 \sin(\pi x / 2\tau_p)$.

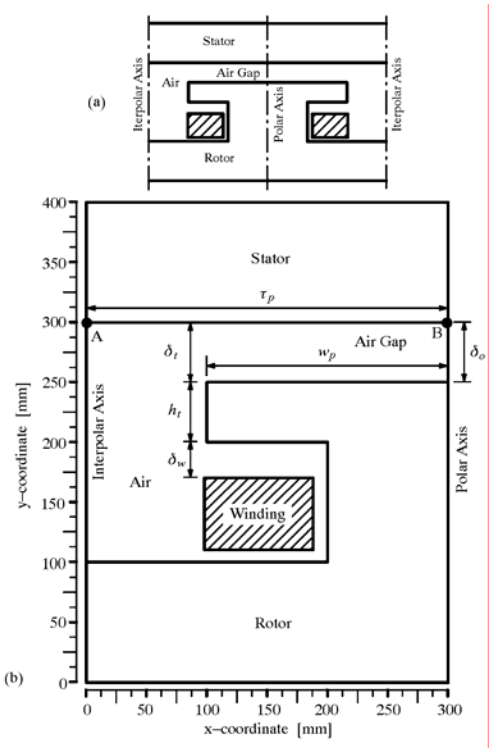

Figure 10. Magnetic-circuit model of a salient-pole synchronous generator: (a) the full pole-pitch geometry including the stator, rotor, air, pole and field-winding regions; (b) the solution domain and geometric dimensions of one-half of the pole pitch

Selection of Case Studies. The case studies considered herein are selected primarily based on the foregoing discussions made on the effect of saturation in [54]. In retrospect, the increasing flux leakage due to the increasing level of saturation in the nonlinear characteristics is *critical* to the performance of the optimization algorithm. These cases must be investigated to determine the limits of the optimization algorithm.

Interpretation of the saturated magnetic-circuit model in the sense of leakage flux is worthwhile considering when identifying the most critical conditions. For the problem in Figure 10, the air-gap flux density B_0 mainly depends on the air-gap size δ_0 and the Ampere-turns NI in the field windings. If, for example, saturation is not included in the magnetic-circuit model, the air-gap flux density is directly proportional to the Ampere-turns, and inversely proportional to the airgap size δ_0 as expressed by $B_0 \approx \mu_0 NI / \delta_0$ (the mmf drops in the iron regions are neglected by assuming that iron is much more permeable than air). It is therefore obvious that a constant ratio of is maintained by the ratio NI / δ_0 in the form of any combination of two variables NI and δ_0.

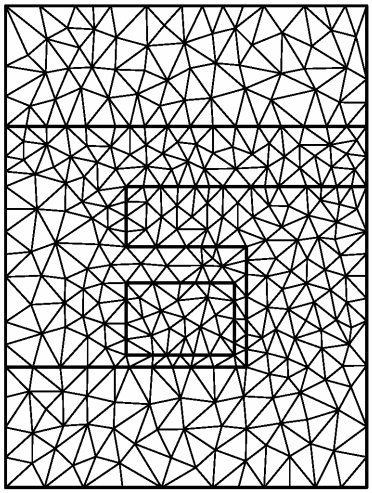

Figure 11. Initial triangulation of the problem geometry in Figure 10. The finite element mesh consists of 622 first-order triangular elements and 340 nodes

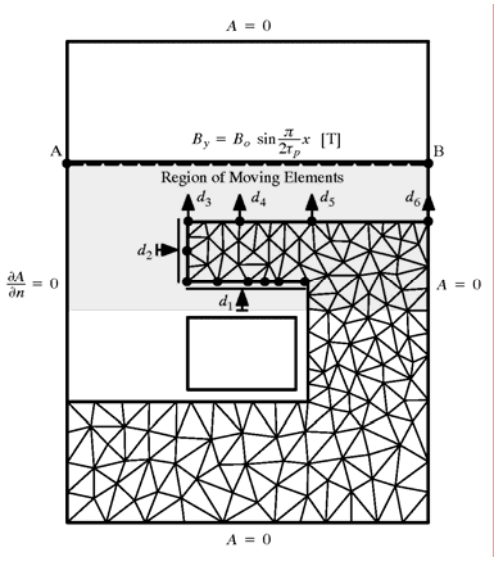

Figure 12. Definition of the field and optimization parameters applied to the mesh model in Figure 11. Dirichlet and Neumann boundary conditions, and the user defined rectangular region (the light-grey region) of moving mesh including 158 nodes

This relationship, on the other hand, does not apply to the saturated problems because increasing saturation level increases the magnitude of the leakage flux through the air, yielding a nonlinear relationship between the ratios B_0 / δ_0 and NI / δ_0. For heavily-saturated cases, varying the ratio NI / δ_0 has a very little effect on B_0. Therefore, increasing flux leakage decreases the sensitivity of the objective function to the critical sizes of iron parts, eventually leading to a very *ill-conditioned* optimization problem for which obtaining a stable numerical

solution is very difficult, sometimes even impossible without employing effective regularization techniques. Based on these observations, the following cases are worth considering:

- an unsaturated generator (Case A) to yield $B_0 = 1.0$ T in the air gap;
- a moderately-saturated generator (Case B) to yield $B_0 = 0.75$ T in the air gap;
- an heavily-saturated generator (Case C) to yield $B_0 = 1.0$ T in the air gap, which is a typical operating condition expected for a high-performance generator.

Using the geometry and field parameters described in Figure 12, the DOPT2D program is used to optimize the pole shape of the synchronous generator for three different saturation levels which have been found to have significant influence on the performance of the optimization procedure. The results obtained for these cases will be presented and discussed within the text as follows.

Unsaturated Generator (Case A). Since air-gap magnetic flux density depends on the the excitation current, a proper current density value should be determined. The value of 4.0 A/mm^2 was found to be an appropriate excitation current value which may yield a starting point which is not significantly far from the optimal solution. This current value approximately determined on the trial-and-error basis by evaluating the objective function for different current density values. This current density value will produce an air-gap flux density whose peak is sufficiently close to the specified magnitude $B_0 = 1.0$ T.

Table 1. Initial and final values of the optimization variables for the unsaturated, salient-pole generator (Case A) to obtain sinusoidal flux-density distribution with magnitude of 1.0 T.

Iter.	Φ	d^L[mm]	d [mm]	d^U[mm]	μ	g	r_p
0	0.3921818	-0.0100	0.0000E+0	0.0100	0.3845E-2	-0.2500	0.5391E-4
		-0.0150	0.0000E+0	0.0100	0.3691E-2	-0.2400	
		-0.0200	0.0000E+0	0.0300	0.3691E-2	-0.2400	
		-0.0200	0.0000E+0	0.0300	0.3691E-2	-0.2400	
		-0.0200	0.0000E+0	0.0300	0.3691E-2	-0.2400	
		-0.0200	0.0000E+0	0.0300	0.3691E-2	-0.2400	
20	0.0363778	-0.0100	-0.1662E-2	0.0100	0.000E+0	-0.2431	0.1552E+2
		-0.0150	-0.1500E-1	0.0100	-0.155E-2	0.0000	
		-0.0200	0.7726E-3	0.0300	0.000E+0	-0.2428	
		-0.0200	0.1155E-1	0.0300	0.000E+0	-0.2328	
		-0.0200	0.1985E-1	0.0300	0.000E+0	-0.1618	
		-0.0200	0.2337E-1	0.0300	0.000E+0	-0.1150	

Computations were carried out to 20 iterations for a relatively simple case without allowing any saturation in the stator and rotor iron. The optimization variables for the initial and the optimal geometry obtained after the 20th iteration are tabulated in Table 1. Note that all displacements computed remained in the feasible domain, and the inequality constraint, g_2, (for the second displacement parameter determining the pole width, w_p) became active at the optimal solution. This substantiates the robustness of the augmented Lagrange multiplier method, forcing the geometric variables to satisfy the inequality constraints during the optimization process.

The initial, and the computed pole geometry obtained after the 20th iteration are shown in Figure 13. Figure 13(a) shows the initial and the deflected mesh for the optimal geometry, respectively. The flux lines computed for the initial and the optimal geometries are shown in

Figure 13 (b). Substantial improvements in the waveform of magnetic flux distribution along the air-gap line A-B for the optimized pole geometry are obtained. It is clearly seen in Figure 14 that the obtained flux density for the new design perfectly matches to the ideal sine-wave distribution at 31 points when compared to that for the initial geometry.

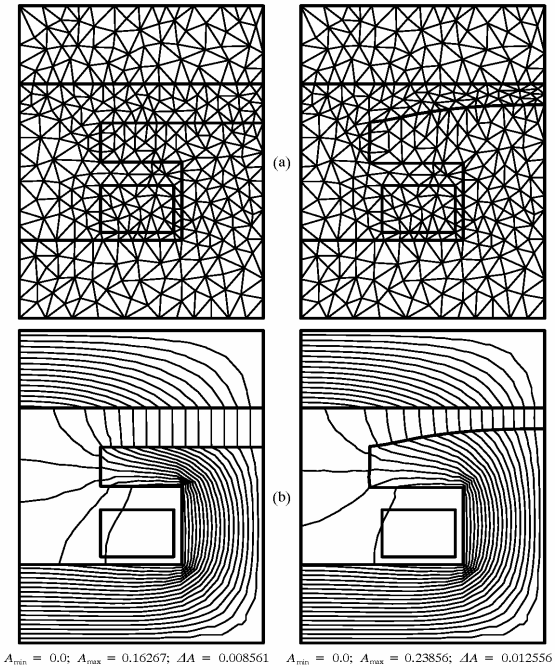

Figure 13. Computational results for the parametrization with 6 variables (in Figure 12) applied to the unsaturated, salient-pole generator to obtain a sinusoidal flux-density distribution with the magnitude of 1.0 T: (a) deflected mesh geometry; (b) flux lines for the initial and the optimized pole shapes

A more quantitative evaluation of these computational results is made by comparing the FFT analyses of the air-gap flux waveforms given in Figure 15. The harmonic ratios computed from Figure 15 are tabulated in Table 2 for numerical comparisons. The improvements in minimizing the harmonics are substantial. The major reduction is obtained for the 3rd harmonic: the 3rd harmonic ratio (where, harmonic ratio is defined as the ratio of the corresponding harmonic to the fundamental component) is reduced from 12% to 1.37% after the optimization. Relatively low but noticeable reductions are also obtained for the 5th and 7th harmonic ratios: they are reduced from approximately 4.64% to 1.74%, and from 4.78% to 1.52%, respectively. A slight increase in the 9th harmonic is obtained: from 1.31% to 1.59%.

Table 2. Comparison of the harmonic ratios of the initial and the optimal design (using the FFT analysis for the unsaturated, salient-pole generator in Case A)

Harmonic Ratio [%]	3rd	5th	7th	9th	11th
Initial	12.00	4.64	4.78	1.31	0.53
Optimal	1.37	1.74	1.52	1.59	0.04

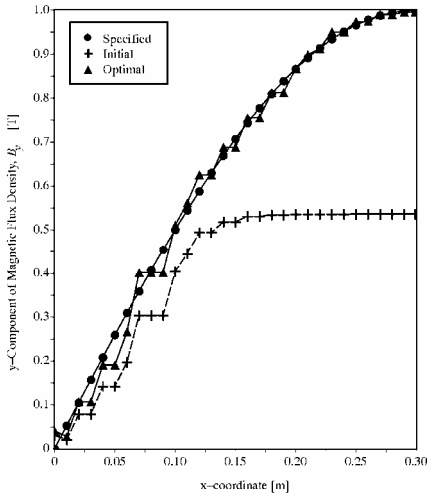

Figure 14. Variation of the y-component of the magnetic flux density along the air-gap line A-B for results in Figure 13(b)

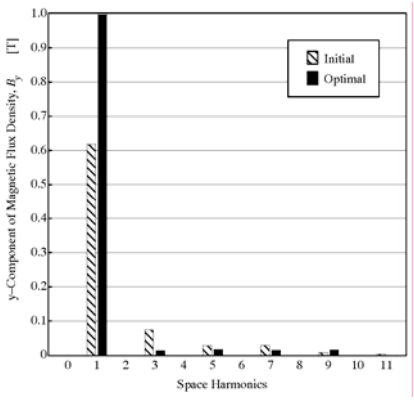

Figure 15. FFT analysis showing the space-harmonics of the air-gap magnetic flux density distribution in Figure 14

Moderately-Saturated Generator (Case B). The solution of the optimization problem related to the unsaturated problem in Case A was stably converged to the optimal solution yielding efficient and accurate results. This case involves the solution of the same problem including the effects of saturation in the finite element model. To obtain a moderate saturation level, the applied winding current density is proportionally decreased to 3.0 A/mm^2 for the magnitude of the desired sinusoidal flux density $B_0 = 0.75$ T. Then, the computations were carried out to 15 iterations. The initial and the final values of the optimization variables for this case are tabulated in Table 3. Note that the normalized least squares error, Φ, is reduced from 0.4000451 to 0.0425484, showing that convergence is achieved properly in even saturated is allowed in the stator and rotor regions.

Table 3. Initial and final values of the optimization variables for the moderately-saturated, salient-pole generator (Case B) to obtain sinusoidal flux-density distribution with magnitude of 0.75 T

Iter.	Φ	d^L[mm]	d [mm]	d^U[mm]	μ	g	r_p
0	0.4000451	-0.0100	0.0000E+0	0.0100	0.4000E-2	-0.2500	0.5609E-3
		-0.0150	0.0000E+0	0.0100	0.3845E-2	-0.2400	
		-0.0200	0.0000E+0	0.0300	0.3845E-2	-0.2400	
		-0.0200	0.0000E+0	0.0300	0.3845E-2	-0.2400	
		-0.0200	0.0000E+0	0.0300	0.3845E-2	-0.2400	
		-0.0200	0.0000E+0	0.0300	0.3845E-2	-0.2400	
15	0.0425484	-0.0100	-0.1361E-2	0.0100	0.000E+0	-0.2454	0.3281E+1
		-0.0150	-0.1500E-1	0.0100	-0.194E-2	0.0000	
		-0.0200	0.5573E-2	0.0300	0.000E+0	-0.2499	
		-0.0200	0.1689E-1	0.0300	0.000E+0	-0.1935	
		-0.0200	0.2393E-1	0.0300	0.000E+0	-0.1067	
		-0.0200	0.2684E-1	0.0300	0.000E+0	-0.0592	

Examining the mesh and the field potential lines for the initial design and the optimal design shown in Figure 16 indicates that the obtained geometry and the flux distribution did not significantly differ from the unsaturated problem. For the optimal pole face geometry obtained herein, the flux distribution for the moderately saturated case is almost perfectly approximated to a sine-distribution as shown in Figure 17. The FFT analysis results for this case are shown in Figure 18. The computed harmonic ratios are given in Table 4. The improvements for the harmonic ratios obtained for the moderate level of saturations are approximately in the same order but slightly lower than those obtained for the unsaturated case.

Table 4. Comparison of the harmonic ratios of the initial and the optimal design (using the FFT analysis for the moderately-saturated, salient-pole generator in Case B).

Harmonic Ratio [%]	3rd	5th	7th	9th	11th
Initial	11.61	4.53	4.72	1.33	0.55
Optimal	1.83	2.22	2.00	1.93	0.16

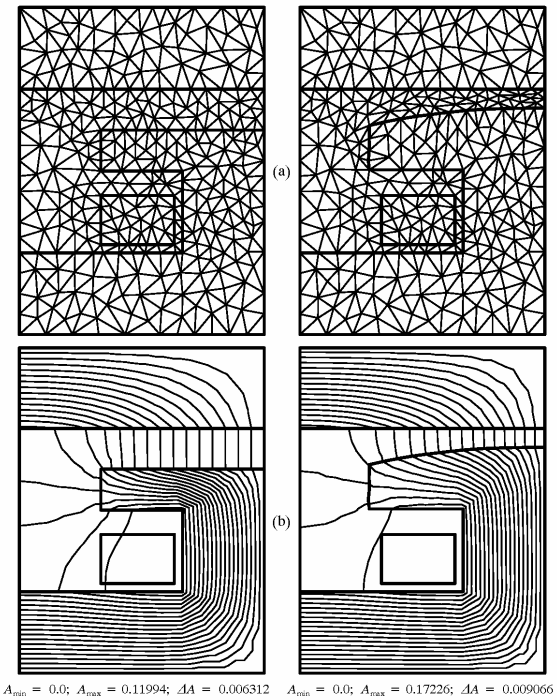

Figure 16. Computational results for the parametrization with 6 variables (in Figure 12) applied to the moderately-saturated, salient-pole generator to obtain a sinusoidal flux-density distribution with the magnitude of 0.75 T: (a) deflected mesh geometry; (b) flux lines for the initial and the optimized pole shapes

Heavily-Saturated Generator (Case C). Recognizing the fact that moderate saturation in iron regions affect the optimization results slightly, the optimization algorithm needs to be tested for

the heavy saturation. In this case, the current density value for is increased to 10.0 A/mm^2 to properly compensate the flux leakage through the air region, and the magnitude of the desired sinusoidal flux distribution is increased to $B_0 = 1.0$ T. The computations were carried out to 20 iterations. The optimization variables for the initial and the final geometry are tabulated in Table 5.

The mesh and the flux lines for the initial and the optimal geometry are shown in Figure 19(a) and (b), respectively. Figure 20 shows the comparison of flux distribution in the air-gap flux distribution before and after the optimization. Note in this example that the initial flux distribution is not flat as obtained in the previous case for $B_0 = 0.75$ T (see Figure 17). The FFT analysis of the flux distribution in Figure 21 shows that a very little improvement is achieved in the third and the higher harmonics (compare Figure 18 for the moderately saturated 0.75 T case and Figure 15 for the unsaturated case of 1.0 T). The computed harmonic ratios for this case are given in Table 6. It appears that the 3rd, 5th and 9th harmonics in the optimized air-gap flux waveform are increased relative to the initial design. Very little improvement is obtained in the 7th and 11th harmonic ratios. It is interesting to see that even if the normalized least-squares error was reduced more than 50% compared to the initial design, the overall improvement in the flux densities is very little.

Table 5. Initial and final values of the optimization variables for the heavily-saturated, salient-pole generator (Case B) to obtain sinusoidal flux-density distribution with magnitude of 1.0 T

Iter.	Φ	d^L[mm]	d [mm]	d^U[mm]	μ	g	r_p
0	0.1328825	-0.0100	0.0000E+0	0.0100	0.2207E-2	-0.2500	0.3095E-4
		-0.0150	0.0000E+0	0.0100	0.2119E-2	-0.2400	
		-0.0200	0.0000E+0	0.0300	0.2119E-2	-0.2400	
		-0.0200	0.0000E+0	0.0300	0.2119E-2	-0.2400	
		-0.0200	0.0000E+0	0.0300	0.2119E-2	-0.2400	
		-0.0200	0.0000E+0	0.0300	0.2119E-2	-0.2400	
20	0.0535845	-0.0100	-0.3648E-3	0.0100	0.000E+0	-0.2497	0.8911E+1
		-0.0150	-0.1500E-1	0.0100	-0.191E-2	0.0000	
		-0.0200	0.8937E-2	0.0300	0.000E+0	-0.2438	
		-0.0200	0.1970E-1	0.0300	0.000E+0	-0.1636	
		-0.0200	0.2617E-1	0.0300	0.000E+0	-0.0707	
		-0.0200	0.2846E-1	0.0300	0.000E+0	-0.0298	

Table 6. Comparison of the harmonic ratios of the initial and the optimal design (using the FFT analysis for the heavily-saturated, salient-pole generator in Case C)

Harmonic Ratio [%]	3rd	5th	7th	9th	11th
Initial	1.57	1.62	3.72	2.08	0.88
Optimal	2.53	2.74	2.54	2.29	0.28

Comparison and Evaluation of Results. The numerical results of the case studies are concerned with the optimal pole shape design of synchronous generator for three different levels of saturation in iron materials resulted in qualitatively similar but quantitatively varying geometric dimensions. The summary of the geometric dimensions of the initial and optimized pole piece

design are listed for all three cases in Table 7. The following conclusions may be drawn from these results with respect to the increase in saturation:

- The air gap δ_0 for the optimized pole shape decreased as the saturation level is increased because of the decreasing permeance as the iron saturates.
- The height of the pole tip h_p increased with the increasing saturation in iron, leading to larger cross-sectional area of the pole piece as the saturation level is increased.
- In all optimization, the ratio of w_p / τ_p remained the same because one half of the pole width τ_p could not exceed the geometric constraints imposed in the optimization problem. It is expected however that relaxation of this constraint to a broader range may result in a further improvement in the design.

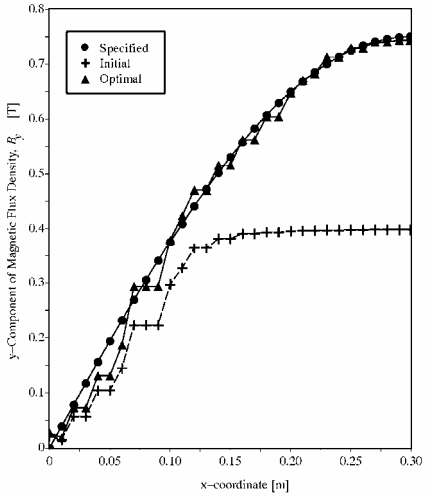

Figure 17. Variation of the y-component of the magnetic flux density along the air-gap line A-B for the results in Figure 16(b)

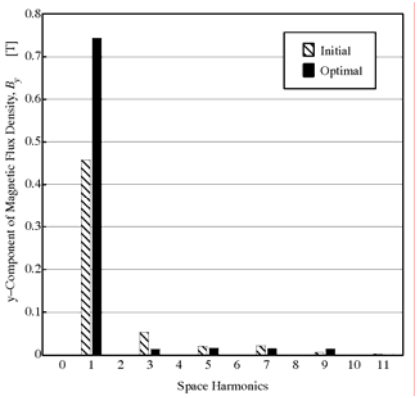

Figure 18. FFT analysis showing the space-harmonics of the air-gap magnetic flux density distribution in Figure 17

Furthermore, the prescribed design criterion was satisfied differently for each case. Of the cases investigated, the best performance in terms of the design criterion and the stability of the converged iterative solution was obtained for the unsaturated generator (Case A). The optimization for the moderate saturation level in Case B resulted in a reasonably good design and performance; the obtained results are quite close to those of the unsaturated generator. While the the maximum value of the harmonic ratios is approximately 12% for both cases. This maximum value after the optimization the maximum is as much as 1.74% for the unsaturated generator and 2.22% for the moderately-saturated generator.

On the contrary to the results obtained for Cases A and B, optimization with heavy saturation in iron (Case C) did not reduce all the harmonic ratios even though the winding current was increased by a factor of 2.5 (relative to that in the unsaturated case) to compensate the leakage flux.

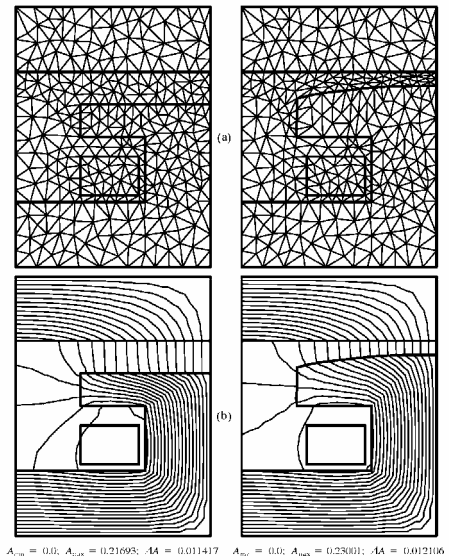

Figure 19. Computational results for the parametrization with 6 variables (in Figure 1.15) applied to the heavily-saturated, salient-pole generator to obtain a sinusoidal flux-density distribution with the magnitude of 1.0 T: (a) deflected mesh geometry; (b) flux lines for the initial and the optimized pole shapes

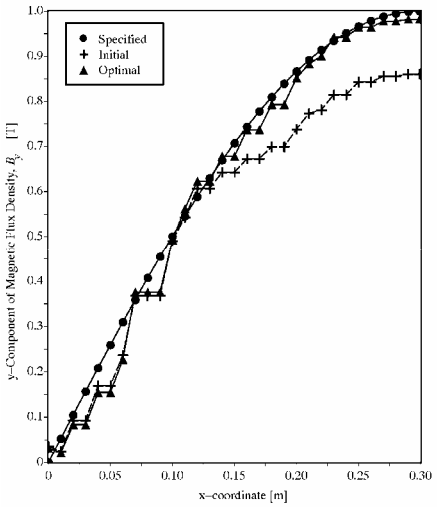

Figure 20. Variation of the y-component of the magnetic flux density along the air-gap line A-B for the results in Figure 19

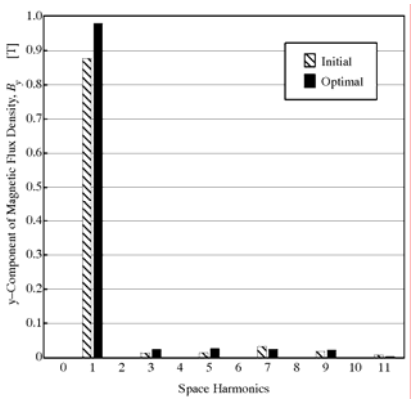

Figure 21. FFT analysis showing the space-harmonics of the air-gap magnetic flux density distribution in Figure 20

Table 7. Summary of the optimal dimensions of the salient-pole, synchronous generator for three different test cases

Case		δ_0 [mm]	δ_t [mm]	δ_w [mm]	h_t [mm]	w_p / τ_p [%]
A	Initial	50.00	50.00	30.00	50.00	66.67
	Optimal	26.63	49.23	23.34	50.77	71.67
B	Initial	50.00	50.00	30.00	50.00	66.67
	Optimal	23.16	44.44	23.64	55.57	71.67
C	Initial	50.00	50.00	30.00	50.00	66.67
	Optimal	21.54	41.06	24.64	58.94	71.67

9.2. ROUND-ROTOR GENERATOR

The study presented in this section is concerned with the design optimization of a round-rotor generator. The magnetic-circuit model including stator, rotor, air and rotor slot regions described in the full pole-pitch of the round-rotor generator is shown in Figure 22. The results presented herein are used to demonstrate the basic principles of the methodologies developed in this study; the geometric details, such as slot geometry and rotor curvature are not included in the model. Handling more complicated geometries, however, is possible by rigorously formulating the parametrization and the finite element models for the magnetic field and the structural sub-problem.

Considering the symmetry conditions about the polar and interpolar axes, the computations are carried out in the half pole-pitch geometry. Figure 23 shows the geometry and design dimensions on one-half of the pole pitch. The seven design dimension describing the design are given: τ_p is one-half of the pole pitch; δ_i (i=1,..,5) is the depth of the corresponding slot number i; t_1 is the distance from the left side of the first slot to the polar axis; t_2 is the distance between the first and the second slots.

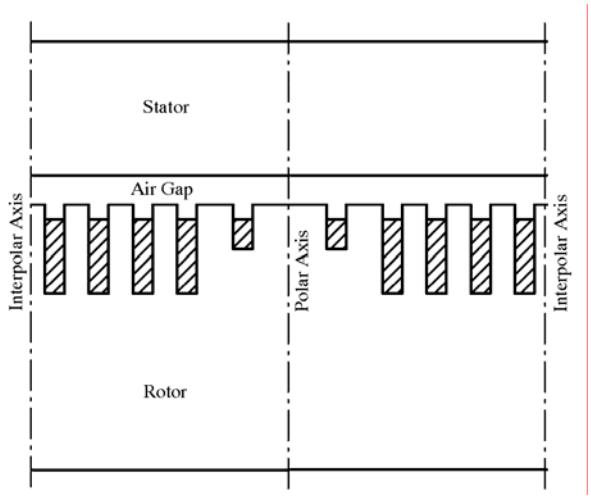

Figure 22. Full pole-pitch geometry of a round-rotor synchronous generator

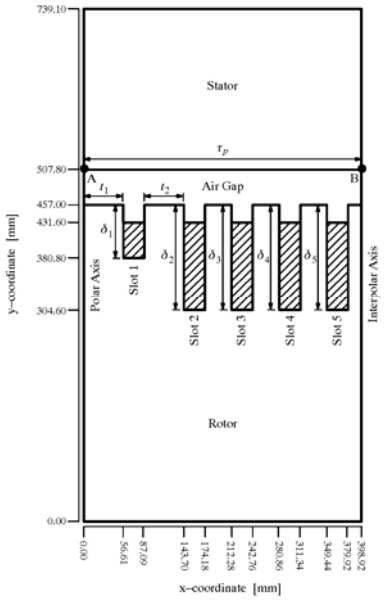

Figure 23. Geometry and dimensions of the part of a round-rotor synchronous machine

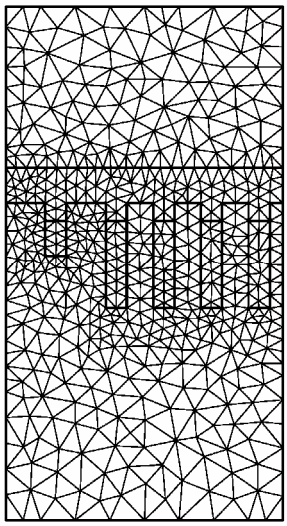

Figure 24. Initial triangulation of the problem domain in Figure 23. The finite element mesh includes 1124 first-order elements and 605 nodes

Design Objective and Problem Definition. The task of optimally designing the round-rotor generator is to determine the geometric dimensions to minimize the space harmonics in the y-component of the flux density along the air-gap line A-B shown in Figure 23. The first step for constructing the optimal design problem is to discretize the problem domain in Figure 23. The triangular mesh of 1124 first-order elements and 605 nodes is shown in Figure 24. The geometry of the rotor slots and location of the first slot on the rotor surface is parametrically defined using six displacements properly assigned to some of the 107 nodes on the rotor surface. Figure 25 shows the assignment of these displacements, the boundary conditions and the boundaries of the structural subproblem including 424 nodes. When choosing the number and the directions of the displacements, maximum sensitivity and minimum computational cost are expected to have by selecting the optimization parameters as follows:

- the displacement d_1 is to vary the slot location by moving in the horizontal direction (left or right) so that the flux waveform is modified;
- the displacements d_2 through d_6 are to vary the depth of each slot to control the Ampere-turns determining the magnitude of the flux waveform on line A-B.

Therefore, minimum number of degrees of freedom is provided for altering both the magnitude and the shape of the air-gap flux waveform.

The sine-wave flux distribution is specified at 51 points (approximately one point in each element attaching the line A-B shown by a dashed line in Figure 25) selected in regular

intervals on the air-gap line A-B. The expression $B_y = B_0 \sin(\pi x / 2\tau_p)$ is used to evaluate the flux densities at 51 points for $B_0 = 1.0$ T.

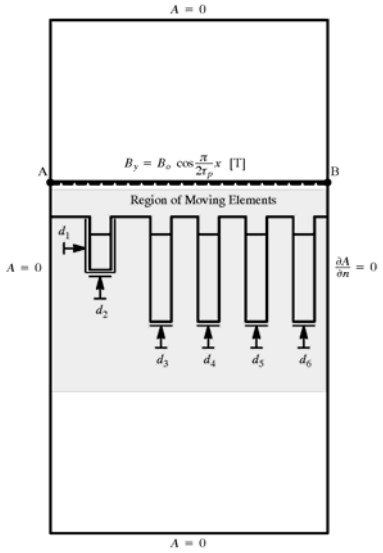

Figure 25. Definition of the field and optimization parameters used for the mesh shown in Figure 23. Dirichlet and Neumann boundary conditions and the user defined rectangular region of moving mesh including 424 nodes

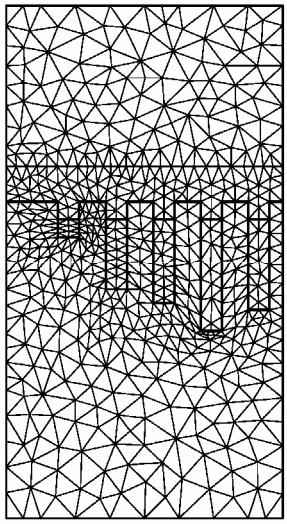

Figure 26. Deflected mesh geometry

Table 8. Initial and final values of the optimization variables for the saturated, round-rotor generator to obtain sinusoidal flux-density distribution with magnitude of 1.0 T

Iter.	Φ	d^L [mm]	d [mm]	d^U [mm]	μ	g	r_p
0	0.0525691	-0.0250	0.0000E+0	0.0250	0.6909E-3	-0.2500	0.1036E-2
		-0.0250	0.0000E+0	0.0250	0.6909E-3	-0.2500	
		-0.0400	0.0000E+0	0.0400	0.6909E-3	-0.2500	
		-0.0400	0.0000E+0	0.0400	0.6909E-3	-0.2500	
		-0.0400	0.0000E+0	0.0400	0.6909E-3	-0.2500	
		-0.0400	0.0000E+0	0.0400	0.6909E-3	-0.2500	
9	0.0220618	-0.0250	0.1790E-1	0.0250	0.000E+0	-0.1219	0.6156E-1
		-0.0250	0.2474E-1	0.0250	0.000E+0	-0.0052	
		-0.0400	0.2509E-1	0.0400	0.000E+0	-0.1517	
		-0.0400	0.4723E-2	0.0400	0.000E+0	-0.2465	
		-0.0400	-0.3536E-1	0.0400	0.000E+0	-0.0546	
		-0.0400	-0.4641E-2	0.0400	0.000E+0	-0.2466	

Numerical Results. A converged solution for the six displacements to minimize the objective function for the saturated round-rotor generator was obtained after 9 iterations. The initial and the final values of the optimization variables are listed in Table 8. The finite element mesh is properly deflected during the iterative modification of the generator geometry. The deflected mesh for the final geometry is shown in Figure 26. The flux lines for the initial and the optimal geometries are shown in Figure 27(a) and (b), respectively.

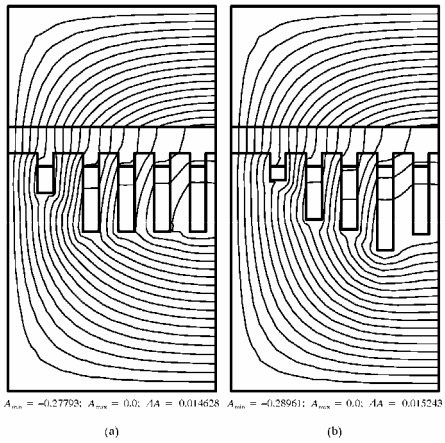

Figure 27. Flux lines for: (a) the initial geometry; (b) the optimized geometry

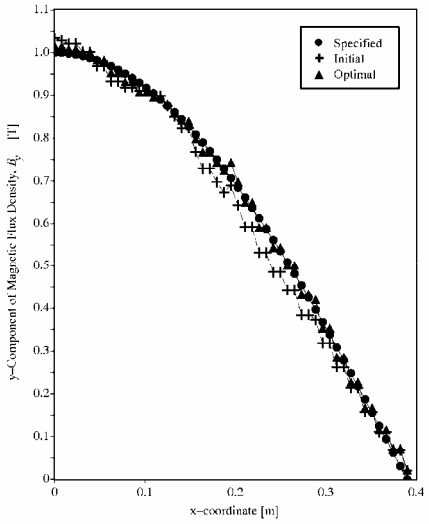

Figure 28. Variation of the y-component of the magnetic flux density along the air-gap line A-B for the results in Figure 27

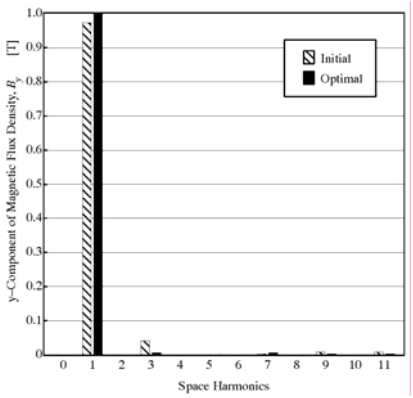

Figure 29. FFT analysis showing the space-harmonics of the air-gap magnetic flux density distribution in Figure 28

Table 9. Comparison of the harmonic ratios of the initial and the optimal design (using the FFT analysis for the saturated, round-rotor generator in Figure 29)

Harmonic Ratio [%]	3rd	5th	7th	9th	11th
Initial	4.27	0.10	0.33	0.88	0.90
Optimal	0.64	0.15	0.71	0.28	0.34

The optimal design dimensions computed by using the six displacements are given in Table 10. Examining the optimal geometry in Figure 27(b) and the dimensions in Table 10 indicate that the design objective is best satisfied by automatically changing the slot sizes, and the moving the first slot in the horizontal direction to improve the top portion of the initial waveform to match the sine-distribution. More geometric parameters can be introduced into the problem at the expense of the computational cost and modeling the very complex structured sub-problem. The difficulty encountered in the previously presented results for the heavily-saturated salient-pole generator was not experienced in this problem because the air-gap flux is mainly determined by the sized of the slots instead of the air-gap size.

Table 10. Geometric dimensions of the round-rotor generator for the initial and the optimized design

Design Parameters	t_1[mm]	t_2[mm]	δ_1[mm]	δ_2[mm]	δ_3[mm]	δ_4[mm]	δ_5[mm]
Initial	56.61	56.61	76.20	152.40	152.40	152.40	152.40
Optimal	74.51	38.71	51.46	127.31	147.68	187.76	157.04

10. CONCLUSIONS

This paper has summarized the effort for analytical and numerical developments in formulating a second-order, constrained optimization algorithm using the augmented Lagrange multipliers method (ALMM). The basic principles of the method are described in detail. The optimization algorithm implemented in the design optimization program, DOPT2D, is specifically developed for optimal shape design of two-dimensional magnetostatic devices accounting for the saturation characteristics of iron materials. This optimization tool minimizes the least-squares objective function subject to global field equations (equality constraints), and the geometric constraints specified by design standards or physical device dimensions (inequality constraints) to compute optimal values of geometric parameters of a device. The DOPT2D program developed in this study is capable of determining optimal geometry of selected material interfaces, such as iron-to-air material interface or geometric dimensions and location of excitation windings, to approximate a desired field distribution prescribed at certain coordinates. The augmented Lagrange multipliers method (ALMM) is partially applied to the inequality constraints to ensure that the computed geometric design parameters stay in the feasible domain when the optimization procedure is completed.

Demonstrated results were devoted to optimal design of synchronous machinery. Both qualitative and quantitative comparisons of the case studies involving optimal pole design in a salient-pole synchronous generator for three different saturation levels yielded different levels of optimal design objective. The conclusions drawn from these results are:

- The performance of the iterative solution of the nonlinear constrained optimization problem also showed different performances for all the cases.

- Computations carried out for saturated and unsaturated-salient pole generator case showed that the leakage flux play an essential role in convergence of iterative procedures. The results obtained for unsaturated and moderately-saturated salient-pole generator cases showed almost identical pole shapes. Space harmonics of the air-gap flux distribution was significantly eliminated yielding almost perfectly sinusoidal air-gap flux distribution in the *y*-direction.
- The results for the heavily-saturated salient-pole generator optimization, on the other hand, did not yield results as good as those for unsaturated and moderately-saturated generator cases.
- Application to a saturated round-rotor synchronous generator demonstrated the following capabilities:
- The computational experiments performed on a saturated round-rotor synchronous generator case did not have the same difficulties previously experienced in the heavily-saturated, salient-pole generator case since the iron boundaries were not modified.
- Considering overall experiments the following conclusions can be drawn for the DOPT2D program:
- The overall results showed that the DOPT2D program is capable of modifying iron-to-air and copper-to-air or copper-to-iron boundaries.
- When modifying heavily-saturated iron-to-air boundaries, care has to be taken. To achieve a reliably converging results, the residual error of the nonlinear field equations is checked after a significant geometric modification. Field analysis should be performed externally to update the field potential if it is necessary.
- Computational requirements can grow with the number of nodes in the mesh model, the number of nodes in the substructural problem and the number of design parameters.

REFERENCES

[1] Gallagher, R., and Zienkiewicz, O. (Editors.), *Optimum Structural Design*, John Wiley & Sons, New York, NY, 1973.

[2] Huag, E. J., and Arora, J. S., *Applied Optimal Design: Mechanical and Structural Systems,* John Wiley & Sons, New York, NY, 1979.

[3] Kirch, U., *Optimum Structural Design*, McGraw-Hill, New York, NY, 1981.

[4] Pironneau, O., *Optimal Shape Design for Elliptic Systems,* Springer-Verlag, New York, NY, 1984.

[5] Vanderplaats, G. N., *Numerical Optimization Techniques for Engineerinig Design,* McGraw-Hill, New York, NY, 1984.

[6] Haslinger, J., and Neittaanmaki, P., *Finite Element Approximation for Optimal Shape Design : Theory and Applications,* John Wiley & Sons, Ltd., New York, NY, 1988.

[7] Salon, S. J. and Istfan, B., "Inverse Nonlinear Finite Element Problems," IEEE Trans. Mag., Vol. 22, No. 5, September 1986, pp. 817-818.

[8] Istfan, B. and Salon, S. J., "Inverse Nonlinear Finite Element Problems with Local and Global Constraints," IEEE Trans. Mag., Vol. 24, No. 6, November 1988, pp. 2568-2572.

[9] Gitosusastro, S., Coulomb, J. L., and Sabonnadiere, J. C., "Performance Derivative Calculations and Optimization Process," IEEE Trans. Mag., Vol. 25, No. 4, July 1989, pp. 2834-2839.

[10] Saldanha, R. R., Coulomb, J. -L., Foggia, A., and Sabonnadiere, J. -C., "A Dual Method for Constrained Optimization Design in Magnetostatic Problems," IEEE Trans. Mag., Vol. 27, No. 5, March 1991, pp. 4136-4141.

[11] Saldanha, R. R., Coulomb, J. -L., and Sabonnadiere, J. -C., "An Ellipsoid Algorithm for The Optimum Design of Magnetostatic Problems," IEEE Trans. Mag., Vol. 28, No. 2, March 1992, pp. 1573-1576.

[12] Kasper, M., "Shape Optimization by Evolution Strategy," IEEE Trans. Mag., Vol. 28, No. 2, March 1992, pp. 1556-1560.

[13] Arpino, F., Di Barba, P., Navarra, P., and Savini A., "3D Global Optimum of an Electromechanical Device," IEEE Trans. Mag., Vol. 30, No. 5, September 1994, pp. 3427-3430.

[14] Appelbaum, J., Fuchs, E. F., and White, J. C., "Optimization of Three-Phase Induction Motor Design, Part I: Formulation of the Optimization Technique," IEEE Trans. Energy Conv., Vol. 2, No. 3, September 1987, pp. 407-414.

[15] Appelbaum, J., Khan, I. A., and Fuchs, E. F., "Optimization of Three-Phase Induction Motor Design, Part II: The Efficiency and Cost of an Optimal Design," IEEE Trans. Energy Conv., Vol. 2, No. 3, September 1987, pp. 415-422.

[16] Simkin, J., and Trowbridge, C. W., "Optimizing Electromagnetic Devices Combining Direct Search Methods with Simulated Annealing," IEEE Trans. Mag., Vol. 28, No. 2, March 1992, pp. 1545-1548.

[17] Weeber, K., *Parametrization, Efficiency and Formulations for the Shape Optimization of Electromagnetic Devices with Finite Elements*, PhD Thesis, Institut National Polytechnique de Grenoble, France, June 1992.

[18] Weeber, K., Hoole, S. R. H., "Geometric Parametrization and Constrained Optimization Techniques in the Design of Salient Pole Synchronous Machines," IEEE Trans. Mag., Vol. 28, No. 4, July 1992, pp. 1948-1960.

[19] Weeber, K., and Hoole, S. R. H., "The Subregion Method in Magnetic Field Analysis and Design Optmization," Trans. Mag., Vol. 28, No. 2, March 1992, pp. 1561-1564.

[20] Koh, C-S, Chung T-K, and Jung H-K, "A Sensitivity Analysis Using Boundary Element Method for Shape Optimization of Electromagnetic Devices," IEEE Trans. Mag., Vol. 28, No. 2, March 1992, pp. 1577-1580.

[21] Park, I., Lee, B., and Hahn, S., "Sensitivity Analysis Based on Analytic Approach for Shape Optimization of Electromagnetic Devices: Interface Problem of Iron and Air," IEEE Trans. Mag., Vol. 27, No. 5, September 1991, pp. 4142-4145.

[22] Park, I., Lee, B., and Hahn, S., "Design Sensitivity Analysis for Nonlinear Magnetostatic Problems Using Finite Element Method," IEEE Trans. Mag., Vol. 28, No. 2, September 1992, pp. 1533-1536.

[23] Subramaniam, S., Arkadan, A. A., and Hoole, S. R. H., "Optimization of a Magnetic Pole Face Using Linear Constraints to Avoid Jagged Contours," IEEE Trans. Mag., Vol. 30, No. 5, September 1994, pp. 3455-3458.

[24] Vasconcelos, J. A., Krahenbuhl, L., Nicolas, L., and Nicolas, A., "Design Optimization in Electrostatic Field Analysis Using the BEM and the Augmented Lagrangian Method," IEEE Trans. Mag., Vol. 30, No. 5, September 1994, pp. 3443-3446.

[25] Hoole, S.R.H., "Optimal Design, Inverse Problems and Parallel Computers," IEEE Trans. Mag., Vol. 27, No. 5, September 1991, pp. 4146-4149.

[26] Enokizono, M., and Tsuchida, Y., "Optimal Design Method by Boundary Element with Fuzzy Interface," IEEE Trans. Mag., Vol. 30, No. 5, September 1994, pp. 3447-3450.

[27] Enokizono, M., Todaka, T., and Nakamura, M., "Optimal Design Method by Boundary Element Analysis for Plural Permanent Magnets," IEEE Trans. Mag., Vol. 28, No. 2, March 1992, pp. 1597-1600.

[28] Preis, K., Biro, O., Friedrich, M., Gottvald, A., and Magele, C., "Comparison of Different Optimization Strategies in the Design of Electromagnetic Devices," IEEE Trans. Mag., Vol. 27, No. 5, September 1991, pp. 4154-4157.

[29] Albanese, R., Ambrosino, G., Martone, R., and Pironti, A., "PF Coil Voltage Optimization for Start-up Scenerios in Air Core Tokamaks," IEEE Trans. Mag., Vol. 30, No. 5, September 1994, pp. 3423-3426.

[30] Bellina, F., Campostrini, P., Chitarin, G., Stella, A., and Trevisan, F., "Automated Optimal Design Techniques for Inverse Electromagnetic Problems," IEEE Trans. Mag., Vol. 28, No. 2, March 1992, pp. 1549-1552.

[31] Drago, G., Manella, A., Nervi, M., Repetto, M., and Secondo, G., "A Combined Strategy for Optimization in Non-Linear Magnetostatic Problems Using Simulated Annealing And Search Techniques," IEEE Trans. Mag., Vol. 28, No. 2, March 1992, pp. 1541-1544.

[32] Hoole, S. R. H., and Subramaniam, S., "Higher Finite Element Derivatives for the Quick Synthesis of Electromagnetic Devices," IEEE Trans. Mag., Vol. 28, No. 2, March 1992, pp. 1565-1568.

[33] Hoole, S. R. H., and Subramaniam, S., "Inverse Problems with Boundary Elements: Synthesizing a Capacitor," IEEE Trans. Mag., Vol. 28, No. 2, March 1992, pp. 1529-1532.

[34] Hameyer, K., and Hanitsch, R., "Numerical Optimization of the Electromagnetic Field by Stochastic Search and MEC-Model," IEEE Trans. Mag., Vol. 30, No. 5, September 1994, pp. 3431-3434.

[35] Ishiyama, A., Shimizu, K., and Sakahara, A., "An Optimal Design Method for Multi-section Superconducting Magnets Using Modified Simulated Annealing," IEEE Trans. Mag., Vol. 30, No. 5, September 1994, pp. 3435-3438.

[36] Kent, R. D., "Degausing with BEM and MFS," IEEE Trans. Mag., Vol. 30, No. 5, September 1994, pp. 3451-3454.

[37] Kitamura, M., Yamamoto, H., Tomeoku, H., and Maki, N., "Optimal Design and Field Measurements of a Dipole Magnet Model for Compact electron Storage Ring," IEEE Trans. Mag., Vol. 28, No. 2, March 1992, pp. 1569-1572.

[38] Marchesi, M. L., Molinary, G., and Repetto, M., "A Parallel Simulated Annealing Algorithm for the Design of Magnetic Structures," IEEE Trans. Mag., Vol. 30, No. 5, September 1994, pp. 3439-3442.

[39] Russenschuck, S., "Application of Lagrange Multiplier Estimation to the Design Optimization of Permanent Magnet Synchronous Machines," IEEE Trans. Mag., Vol. 28, No. 2, March 1992, pp. 1525-1528.

[40] Russenschuck, S., and Tortschanoff, T., "Mathematical Optimization of Superconducting Accelerator Magnets," IEEE Trans. Mag., Vol. 30, No. 5, September 1994, pp. 3419-3422.

[41] Reeve, P. A., "Field Optimization Using an Analytic Model Feedback Technique," IEEE Trans. Mag., Vol. 28, No. 2, March 1992, pp. 1553-1555.

[42] Saldanha, R. R., Pelissier, S., Kadded, K., Yonnet, Y .P., and Coulomb, J. -L., "Nonlinear Optimization Methods Applied to Actuators Design," IEEE Trans. Mag., Vol. 28, No. 2, March 1992, pp. 1581-1584.

[43] Watterson, P. A., Zhu, J. G., and Ramsden, V.S., "Optimization of Permanent Magnet Motors Using Field Calculations of Increasing Precision," Trans. Mag., Vol. 28, No. 2, March 1992, pp. 1589-1592.

[44] Gottvald, A., "Comperative Analysis of Optimization Methods for Magnetostatics," IEEE Trans. Mag., Vol. 24, No. 1, January 1988, pp. 411-414.

[45] Gottvald, A., "Optimal Magnet Design for NMR," IEEE Trans. Mag., Vol. 26, No. 2, March 1990, pp. 399-402.

[46] Gottvald, A., "Global Optimization Methods for Computational Electromagnetics," IEEE Trans. Mag., Vol. 28, No. 1, March 1992, pp. 1537-1540.

[47] Ecker, J. G., and Kupferschmid, M., *Introduction to Operations Research*, Krieger Publishing Co., Malabar, FL, 1988.

[48] Luenberger, D. G., *Linear and Nonlinear Programming*, Addison-Wesley, Reading, MA, 1989.

[49] Scales, J. A., Smith, M. L., and Fischer, T. L., "Global Optimization Methods for Multimodal Inverse Problems," J. Comp. Phys., Vol. 103, 1992, pp. 258-268.

[50] Hoole, S.R.H., *Computer-Aided Analysis and Design of Electromagnetic Devices*, Elsevier Science Publishing Co., Inc., New York, NY, 1989.

[51] Lowther, D. A., and Silvester, P. P., *Computer-Aided Design in Magnetics*, New York, NY, 1985.

[52] Binns, K. J., Lawrenson, P. J., and Trowbridge, C. W., *The Analytical and Numerical Solution of Electric and Magnetic Fields*, John Wiley and Sons, Ltd., New York, NY, 1992.

[53] Bathe, K. J., *Finite Element Procedures in Engineering Analysis*, Prentice Hall, New York, NY, 1982.

[54] Ovacik, L, *Extensions to the finite element Method for the Analysis of Inverse Problems in electromagnetic Devices*, PhD Thesis, Rensselaer Polytechnic Institute, USA, December 1998.

[55] Haug, E. J., Choi, K. K., and Komkov, V., *Design Sensitivity Analysis of Structural Systems*, Academic Press, Orlando, FL, 1986.

[56] Coulomb, J. L., "A Methodology for the Determination of Global Electromechanical Quantities from a Finite Element Analysis and Its Application to the Evaluation of Magnetic Forces, Torques and Stiffness," IEEE Trans. Mag., Vol. 19, No. 6, November 1983, pp. 2514-2519.

[57] Biedinger, J. M., and Lemoine, D., "Shape Sensitivity Analysis of Magnetic Forces," IEEE Trans. Mag., Vol. 33, No. 3, May 1997, pp. 2309-2316.

[58] Zienkiewicz, O. C., and Taylor, R. L., *The Finite Element Method*, 4th Edit., MacGraw-Hill, New York, NY, 1988.

[59] Marrocco, A., and Pironneau, O., "Optimum Design with Lagrangian Finite Elements: Design of an Electromagnet," Comp. Meth. Appl. Mech. Eng., Vol. 15, 1978, pp. 277-308.

[60] Akin, J. E., *Finite Element for Analysis and Design*, Academinc Press, Inc., San Diego, CA, 1994.

[61] Zienkiewicz, O. C., and Taylor, R. L., *The Finite Element Method*, 4th Edit., MacGraw-Hill, New York, NY, 1988.

[62] Silvester, P. P., Cabayan, H. S. and Browne, B. T., "Efficient Techniques for Finite Element Analysis of Electric Machines," IEEE Trans. Power Appar. Syst., Vol. 92, No. 4, July/August 1973, pp. 1274-1281.

[63] El-Sherbiny, M. K., "Representation of the Magnetization Characteristic by a Sum of Exponentials," IEEE Trans. Mag., Vol. 9, No. 1, March 1973, pp. 60-61.

[64] Hoole, S. R. H. and Hoole, N. R. G., "Reluctivity Characteristics in Nonlinear Finite Element Analysis of Magnetostatic Fields," IEEE Trans. Mag., Vol. 22, No. 5, September 1986, pp. 1352-1353.