Physics

Electricity & Magnetism fields

Okayama University *Year* 2004

The Newton-Raphson method accelerated by using a line search - comparison between energy functional and residual minimization

Koji Fujiwara[*] Yoshifumi Okamoto[†]
Akihisa Kameari[‡] Akira Ahagon[**]

[*]Okayama University
[†]Okayama University
[‡]Science Solutions International Laboratory, Incorporated
[**]Science Solutions International Laboratory Incorporated

The Newton–Raphson Method Accelerated by Using a Line Search—Comparison Between Energy Functional and Residual Minimization

Koji Fujiwara[1], Yoshifumi Okamoto[1], Akihisa Kameari[2], *Member, IEEE*, and Akira Ahagon[3]

[1]Department of Electrical and Electronic Engineering, Okayama University, Okayama 700-8530, Japan
[2]Research Group of Science Solutions International Laboratory, Inc., Tokyo 153-0065, Japan

A line search was combined with the Newton–Raphson method to accelerate the convergence of the iterative calculation in nonlinear magnetic field analysis. As a method for determining a step size for update, the minimization of an energy functional and a square of 2-norm of residual obtained from the finite-element discretization was investigated. It was demonstrated that the energy functional minimization is superior to the residual minimization from the viewpoint of computational cost. The line search is effective even in the magnetic vector potential formulation, which is said to be stable usually.

***Index Terms*—Energy functional, line search, minimization, Newton–Raphson method, nonlinear magnetic field analysis, residual.**

I. Introduction

In order to take account of the nonlinearity of magnetic properties of magnetic materials, such as electrical steel sheets and permanent magnets, the iterative calculation using the Newton–Raphson method is adopted in general. However, in the case of strong nonlinearity, a large number of iterations may be required, and the iterative procedure will be failed to converge in the worst. The convergence characteristic is also dependent on a kind of unknown variable selected for discretization. In [1] and [2], it was reported that the nonlinear iteration was failed to converge, when the TEAM Workshop Problem 13 (a nonlinear magnetostatic model) was analyzed by using the magnetic scalar potential formulation. Then, simple under-relaxation was introduced to get the stable convergence by utilizing a square of 2-norm of residual. An under-relaxed step size for update is determined at every iteration so that the residual obtained from the finite element discretization can decrease monotonously with iteration. It was fairly effective to the analysis of Problem 13 using the magnetic scalar potential formulation. However, it is not so attractive to the other problems having relatively good convergence characteristic because it allows only the under-relaxation.

In this paper, a method combining a line search [3], [4] with the Newton–Raphson method was investigated to improve the convergence characteristic by allowing over-relaxation as well as under-relaxation even in case of well-posed problems. In the line search procedure, a step size for update was determined by minimizing an objective function such as an energy functional or a square of 2-norm of residual. A typical electromagnet model was analyzed to examine behavior of the objective function and the step size during the nonlinear iteration. Computational cost was also compared.

Digital Object Identifier 10.1109/TMAG.2005.846048

II. Nonlinear Magnetic Field Analysis Based on the Newton–Raphson Method

In this section, first, the ordinary Newton–Raphson method is described briefly to understand the following explanation clearly. Second, the Newton–Raphson method with a line search is discussed. In the finite element discretization, the magnetic vector potential formulation is used. In order to determine a step size for update in the line search procedure, the energy functional or the square of 2-norm of residual is adopted.

A. Ordinary Newton–Raphson Method

Fig. 1 shows the flowchart of nonlinear magnetic field analysis based on the Newton–Raphson method. $\boldsymbol{A}$ is the magnetic vector potential. ν is the reluctivity. $\boldsymbol{B}$ is the flux density. $\boldsymbol{G}$ is the residual. $\partial\boldsymbol{G}/\partial\boldsymbol{A}$ is the coefficient matrix. $\delta\boldsymbol{A}$ is the increment of $\boldsymbol{A}$. α is the step size for update. $\delta\boldsymbol{B}(=\mathrm{rot}(\alpha\delta\boldsymbol{A}))$ is the increment of $\boldsymbol{B}$. Step 5 is required only in case of the line search to determine an effective α at every iteration. α in the ordinary Newton–Raphson method is equal to unity. Explanations of the other steps are skipped because they are nothing special. In this paper, the ordinary Newton–Raphson method is abbreviated to "normal NR."

B. Newton–Raphson Method With a Line Search

In the iterative procedure for the Newton–Raphson method with a line search, only step 5 mentioned above is added to the ordinary Newton–Raphson procedure. At step 5, a step size for update is determined at every iteration so that an objective function can decrease monotonously or can be minimized in the direction of the increment of solution. Hereinafter, two methods are investigated, in which a square of 2-norm of residual or an energy functional is used as the objective function.

1) Square of 2-Norm of Residual:

a) Step Size With Monotonous Decrease: When the nonlinear iteration can reach to the converged state, the residual for each unknown approaches to zero ideally. Therefore, as one of the robust methods, a step size for update is determined at every

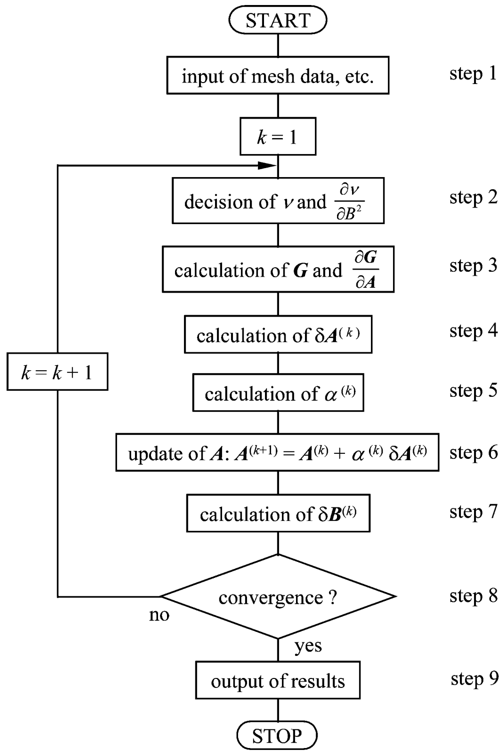


Fig. 1. Flowchart of nonlinear magnetic field analysis based on the Newton–Raphson method.

```
    α(k)=1
10  A(k+1) = A(k) + α(k)δA(k)
    if (||G(k+1)||² > ||G(k)||²) then
      α(k) = α(k) × 0.5
      go to 10
    end if
```

Fig. 2. Method for determining a step size in the simple NR.

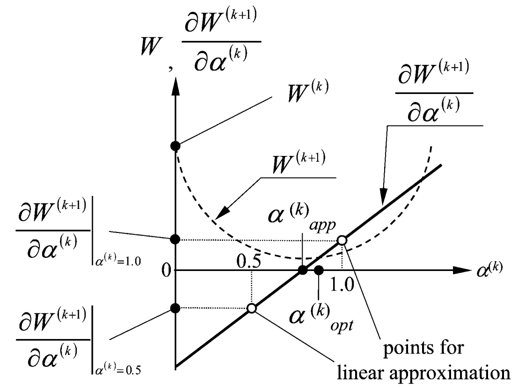


Fig. 3. Method for determining a step size by using linearization.

iteration so that a square of 2-norm of residual $(\|\boldsymbol{G}\|_2)^2$ can decrease monotonously with iteration [3]. Fig. 2 shows the calculation procedure. The initial value of step size at the kth nonlinear iteration $\alpha^{(k)}$ is set to be unity. If $(\|\boldsymbol{G}^{(k+1)}\|_2)^2$ at the $(k+1)$th iteration is larger than $(\|\boldsymbol{G}^{(k)}\|_2)^2, \alpha^{(k)}$ is reduced to half of the previous value. As the finally obtained $\alpha^{(k)}$ is less than unity, the update is under-relaxed at every iteration. Therefore, when the ordinary Newton–Raphson method can perform the monotonous decrease of $(\|\boldsymbol{G}\|_2)^2$, the introduction of the above-mentioned method has no meaning. However, the method is fairly effective to ill-posed problems because of the guarantee of the monotonous decrease of $(\|\boldsymbol{G}\|_2)^2$. The method is abbreviated to "simple NR."

b) Optimal Step Size Obtained From Minimization: $(\|\boldsymbol{G}\|_2)^2$ seems to have a minimum. Therefore, the optimal step size, $\alpha_{\rm opt}^{(k)}$, can be calculated by the minimization of $(\|\boldsymbol{G}\|_2)^2$ using the following equation:

$$\frac{\partial\left(\left\|\boldsymbol{G}^{(k+1)}\right\|_2\right)^2}{\partial\alpha^{(k)}} = \frac{\partial\left(\left\|\boldsymbol{G}^{(k+1)}\right\|_2\right)^2}{\partial\boldsymbol{A}^{(k+1)}}\frac{\partial\boldsymbol{A}^{(k+1)}}{\partial\alpha^{(k)}} = 2\boldsymbol{G}^{(k+1)^T}\frac{\partial\boldsymbol{G}^{(k+1)}}{\partial\boldsymbol{A}^{(k+1)}}\delta\boldsymbol{A}^{(k)} = 2\sum_{i=1}^{\nu}G_i^{(k+1)}\left(\sum_{j=1}^{\nu}\frac{\partial G_i^{(k+1)}}{\partial A_j^{(k+1)}}\delta A_j^{(k)}\right) = 0 \quad (1)$$

where nu is the number of unknowns. The superscript T means the transpose. The subscripts i and j correspond to the unknown numbers. As the method has no restrictions on a search range of $\alpha^{(k)}$, over-relaxation, as well as under-relaxation, can be performed.

If $(\|\boldsymbol{G}^{(k+1)}\|_2)^2$ has complete quadratic change with $\alpha^{(k)}, \partial(\|\boldsymbol{G}^{(k+1)}\|_2)^2/\partial\alpha^{(k)}$ become a linear function and can be represented by using two values at arbitrary two $\alpha^{(k)}$s. Therefore, $\alpha_{\rm opt}^{(k)}$ can be calculated easily. Even in the case when the change of $(\|\boldsymbol{G}^{(k+1)}\|_2)^2$ is not completely quadratic, an approximated value $\alpha_{\rm app}^{(k)}$ can be obtained by the linearization of $\partial(\|\boldsymbol{G}^{(k+1)}\|_2)^2/\partial\alpha^{(k)}$ at typical two $\alpha^{(k)}$s of, for example, 0.5 and 1.0, as shown in Fig. 3. $W^{(k+1)}$ means the objective function and corresponds to $(\|\boldsymbol{G}^{(k+1)}\|_2)^2$. The method is abbreviated to "residual NR."

2) Energy Functional: It seems that the energy functional χ can be an alternative to $(\|\boldsymbol{G}\|_2)^2$ because the solution is searched by minimizing χ in the finite element method [4]. $\alpha_{\rm opt}^{(k)}$ can be obtained from the following equation. The way to calculate the approximated value $\alpha_{\rm app}^{(k)}$ of $\alpha_{\rm opt}^{(k)}$ is the same as before. The method is abbreviated to "functional NR"

$$\frac{\partial\chi^{(k+1)}}{\partial\alpha^{(k)}} = \left(\frac{\partial\chi^{(k+1)}}{\partial\boldsymbol{A}^{(k+1)}}\right)\frac{\partial\boldsymbol{A}^{(k+1)}}{\partial\alpha^{(k)}} = \boldsymbol{G}^{(k+1)^T}\delta\boldsymbol{A}^{(k)} = \sum_{i=1}^{\nu}G_i^{(k+1)}\delta A_i^{(k)} = 0. \quad (2)$$

III. Test Example

An electromagnet with a small gap shown in Fig. 4 was selected as a test example. One second of the whole model is illustrated because of the symmetry. The material of yoke is S45C steel used for structural purposes. Its magnetic property is shown in Fig. 5 and Table I. The property of $0 \leq B \leq 2.048$ T (B: flux density) obtained from the measurement was interpolated by using piecewise cubic polynomials. In the region of

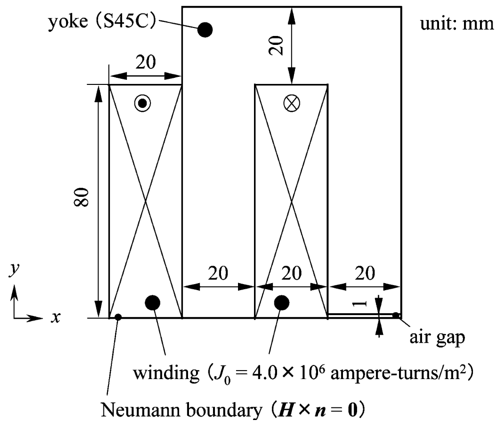


Fig. 4. Analyzed model.

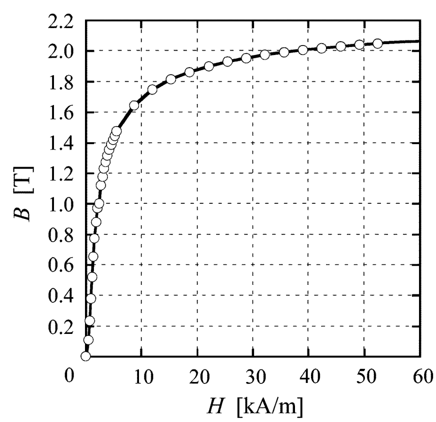


Fig. 5. Magnetic property of S45C steel.

2.048 T $\leq B \leq B_s$ (B_s: saturation flux density), a quadratic polynomial was assigned. In the region of $B_s \leq B$, the magnetic field strength H changes monotonously with B of which the slope is $1/\mu_0$ (μ_0: permeability of vacuum). A dc current of 4.0×10^6 ampere-turns/m^2 was applied to the winding and then the 2-D magnetostatic analysis was carried out. The initial value of relative permeability of yoke was set to be 1000. As boundary conditions, the Neumann condition is assigned on the boundary of symmetry. The other boundaries have the fixed condition of $A = 0$. Fig. 6 shows the mesh. The numbers of elements, nodes, and unknowns of the first order triangular mesh are 6948, 3506, and 3496, respectively. Linear equations were solved by both the Gaussian elimination method and ICCG method. Fig. 7 shows the flux distribution and distribution of flux density to understand saturation level.

IV. Results and Discussion

Fig. 8 shows the changes of $(\|\boldsymbol{G}^{(k+1)}\|_2)^2$ and $\partial(\|\boldsymbol{G}^{(k+1)}\|_2)^2/\partial\alpha^{(k)}$ with $\alpha^{(k)}$ in the ordinary Newton–Raphson procedure. Results at the first iteration shown in Fig. 8 (a) are identical to those obtained from the linear analysis because the initial value of $\boldsymbol{A}$ is set to be zero. $\partial(\|\boldsymbol{G}^{(k+1)}\|_2)^2/\partial\alpha^{(k)}$ is linearized at two $\alpha^{(k)}$s of 0.5 and 1.0. A step size obtained from the linearization $\alpha_{\mathrm{app}}^{(k)}$ is different from $\alpha_{\mathrm{opt}}^{(k)}$. However, as $(\|\boldsymbol{G}^{(k+1)}\|_2)^2$s at $\alpha^{(k)} = \alpha_{\mathrm{opt}}^{(k)}$ and $\alpha_{\mathrm{app}}^{(k)}$ are very close to each other, it is not problematic from the practical viewpoint.

TABLE I
Details of Magnetic Property of S45C Steel.
(a) Measured Data. (b) Extrapolated Data

(a)

no.	B [T]	H [A/m]	no.	B [T]	H [A/m]
1	0	0	18	1.417	4880
2	0.109	520	19	1.446	5203
3	0.235	731	20	1.474	5539
4	0.378	944	21	1.642	8661
5	0.520	1151	22	1.746	12000
6	0.653	1378	23	1.814	15301
7	0.773	1617	24	1.862	18706
8	0.880	1860	25	1.900	22094
9	0.972	2137	26	1.929	25461
10	1.000	2418	27	1.953	28818
11	1.118	2700	28	1.974	32212
12	1.177	3000	29	1.990	35673
13	1.230	3300	30	2.004	39001
14	1.274	3614	31	2.017	42400
15	1.316	3923	32	2.028	45793
16	1.354	4240	33	2.039	49171
17	1.388	4564	34	2.048	52489

(b)

・ $2.048 \leqq B \leqq B_s$	・ $B_s \leqq B$
$H = aB^2 + bB + c$	$H = (B - M_s)/\mu_0$
$a = 2.8087 \times 10^5$	※ $M_s = 2.1580$
$b = -7.5129 \times 10^5$	
$c = 4.1309 \times 10^5$	
※ $B_s = 2.7541$,	
$H_s = 4.7434 \times 10^5$	

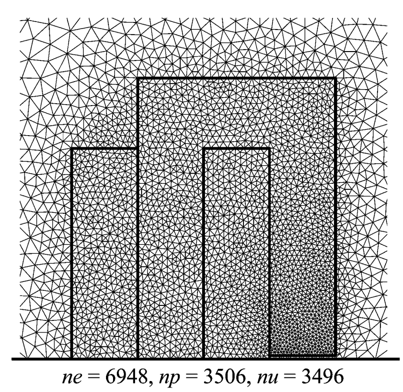
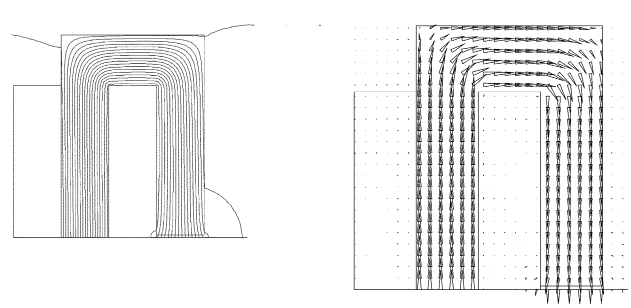


Fig. 6. Mesh.



(a) (b)

Fig. 7. Analyzed results. (a) Flux distribution. (b) Distribution of flux density vectors.

At the first iteration, the update can be under-relaxed because of $\alpha_{\mathrm{app}}^{(k)} < 1$. At the fifth iteration, the update can be

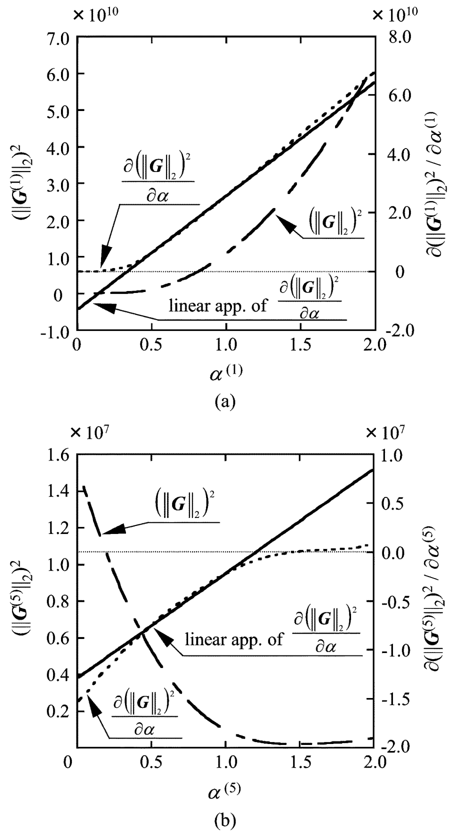


Fig. 8. Changes of a square of 2-norm of residual and its derivative with step size. (a) First iteration ($\alpha_{\rm opt}^{(1)} = 0.10, \alpha_{\rm app}^{(1)} = 0.34$). (b) Fifth iteration ($\alpha_{\rm opt}^{(5)} = 1.47, \alpha_{\rm app}^{(5)} = 1.21$).

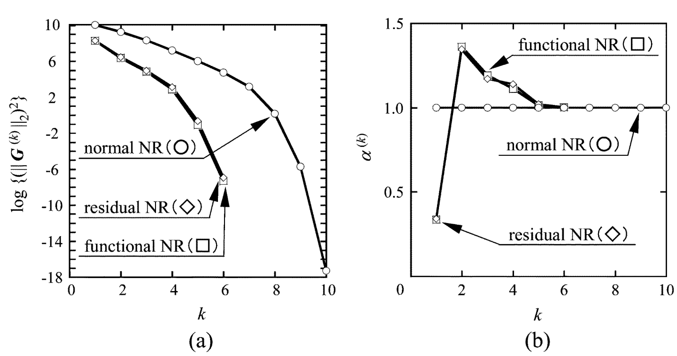


Fig. 9. Changes of square of 2-norm of residual and step size with iteration. (a) Square of 2-norm of residual. (b) Step size.

over-relaxed because of $\alpha_{\rm app}^{(k)} > 1$. The changes of $\chi^{(k+1)}$ and $\partial\chi^{(k+1)}/\partial\alpha^{(k)}$ with $\alpha^{(k)}$ have a similar tendency.

Fig. 9(a) shows the convergence characteristic of $(\|\boldsymbol{G}^{(k)}\|_2)^2$. When the change of the flux density is less than 10^{-3} T for each element, the nonlinear iteration is terminated. The functional and residual NRs require the six iterations, whereas the normal NR requires the ten iterations. Therefore, the former is superior to the latter. It is noted that the simple NR is not performed in this calculation because the normal NR could give the monotonous decrease of $(\|\boldsymbol{G}^{(k)}\|_2)^2$.

Fig. 9(b) shows the behavior of $\alpha^{(k)}$ during the nonlinear iteration. The update at the first iteration is under-relaxed in both the functional and residual NRs because the flux density is over-estimated in the linear analysis. From the second iteration, the updates are over-relaxed and $\alpha^{(k)}$ finally approaches to unity. It is understood from Fig. 9 that the residual NR is practically equivalent to the functional NR.

TABLE II
COMPUTATIONAL COST

method	number of iterations (normlized)	time per iteration [s] (normalized)		total time [s] (normalized)	
		Gauss	ICCG	Gauss	ICCG
normal NR	10 (1)	0.40 (1)	0.056 (1)	3.95 (1)	0.56 (1)
residual NR	6 (0.6)	0.43 (1.08)	0.070 (1.25)	2.59 (0.66)	0.42 (0.75)
functional NR	6 (0.6)	0.41 (1.05)	0.068 (1.21)	2.45 (0.62)	0.41 (0.73)

computer used: Pentium 4 / 3.2 GHz with 2 GB RAM

Table II shows the number of nonlinear iterations and CPU time required to get the convergence. The value in parentheses means that it is normalized by the value obtained from the normal NR. The functional NR requires the shorter CPU time for calculating a step size than the residual NR. This is also obvious from the comparison of (1) and (2). Therefore, the functional NR is superior to the residual NR from the viewpoint of computational cost. In case of the Gaussian elimination method, about 40% of the CPU time is reduced. In case of the ICCG method, about 30% is reduced. As the ICCG method can solve linear equations for considerably shorter time than the Gauss elimination method, the percentage of the time for calculating a step size become larger. When larger-scale problems are solved, the reduction of time for the ICCG method approaches to the same value as the Gaussian elimination method.

V. Conclusion

In order to improve the convergence characteristic of the nonlinear magnetic field analysis, a line search was combined with the Newton–Raphson method, in which an energy functional or a square of 2-norm of residual was selected as an objective function and minimized. It can be concluded that the line search is fairly effective, and the minimization of the energy functional is superior to the residual minimization from the viewpoint of the computational cost.

References

[1] T. Nakata, N. Takahashi, K. Fujiwara, N. Okamoto, and K. Muramatsu, "Improvements of convergence characteristics of Newton-Raphson method for nonlinear magnetic field analysis," *IEEE Trans. Magn.*, vol. 28, no. 2, pp. 1048–1051, Mar. 1992.

[2] K. Fujiwara, T. Nakata, N. Okamoto, and K. Muramatsu, "Method for determining relaxation factor for modified Newton-Raphson method," *IEEE Trans. Magn.*, vol. 29, no. 2, pp. 1962–1965, Mar. 1993.

[3] O. C. Ziewnkiewicz and R. L. Taylor, *The Finite Element Method, Fourth Edition, Volume 2: Solid and Fluid Mechanics Dynamics and Non-Linearity*. New York: McGraw-Hill, 1989.

[4] M. Schinnerl, J. Schöberl, M. Kaltenbacher, and R. Lerch, "Multigrid methods for the three-dimensional simulation of nonlinear magnetomechanical systems," *IEEE Trans. Magn.*, vol. 38, no. 3, pp. 1497–1511, May 2002.

Manuscript received June 8, 2004.