# Regularization of the Moment Matrix Solution by a Nonquadratic Conjugate Gradient Method

Luc Knockaert, *Member, IEEE,* and Daniël De Zutter, *Senior Member, IEEE*

***Abstract*—Inspired by Tikhonov regularization, a nonlinear conjugate gradient method is proposed with the purpose of simultaneously regularizing and solving the moment matrix equation. The procedure is based on a nonquadratic conjugate gradient algorithm with exact line search, restart, and rescale. Applied to the problem of TM scattering by perfectly conducting rectangular cylinders, the method is shown to exhibit a fast convergence rate.**

***Index Terms*—Conjugate gradient method, method of moments.**

## I. Introduction

It is widely known that the solution of the electric field integral equation (EFIE) [1] and, more generally speaking, the solution of Fredholm integral equations of the first kind [2] are ill-posed problems [3]. The main reason for this is that the kernel or Green's function of the integral equation represents a compact operator [4] for which the eigenvalues always cluster in the vicinity of the origin.

As a consequence, when simple subsectional basis and testing functions such as pulses are utilized or when the geometry of the scatterer is nonsmooth, the resulting moment matrix is often very ill conditioned. This implies that the convergence rate of the conjugate gradient (CG) method as applied to the normal equations [5], [6] will, in general, be too low [7] and, therefore, CG in general performs not as well as direct methods such as LU decomposition.

As stated by Nashed [3], the philosophy of resolution of ill-posed problems involves one or more of the following intuitive ideas.

- Change the spaces and/or topologies.
- Modify the equation or the problem itself.
- Change the notion of what is meant by a solution.

An instance of the first idea, change the topology, i.e., the norm, has been proposed in [8]. The second idea corresponds with the use of various matrix preconditioners [1] and the third idea corresponds with the theory of Tikhonov regularization [9].

Our approach is a mixture of the first and third ideas. It is based on Tikhonov regularization, but unlike Tikhonov regularization, where a nearby linear system is solved, it solves the exact problem. Unlike the Euclidean norm minimization which always leads to normal equations [6], we minimize a more general nonquadratic and bounded functional. The key point is that it is possible to apply the CG method successfully to a large class of nonquadratic minimization problems [10]–[12]. The main difficulties in nonquadratic CG reside with exact line searches [11] and restart [13] procedures. It is shown that the line-search, restart, and rescale procedures can be implemented in an efficient and simple way, with the same computational complexity, i.e., two matrix–vector products per step, as normal equations CG. This results in two algorithms, the first based on the Polak–Ribière accelerator [11] with restart–rescale cycling and the second on a accelerator resulting from total rescaling. Finally, the algorithms are tested on the problem of TM scattering by rectangular perfectly conducting (PEC) cylinders.

Manuscript received December 10, 1997; revised December 16, 1999.

The authors are with the Department of Information Technology INTEC, B-9000 Gent, Belgium (e-mail: knokaert@intec.rug.ac.be).

Publisher Item Identifier S 0018-926X(00)04371-4.

## II. Regularization

Consider the moment matrix equation

$$\mathtt{Z}x - y = 0. \tag{1}$$

Since $\mathtt{Z}$ is in general ill conditioned but nonsingular, a Tikhonov regularization scheme [9] can be proposed with the purpose of minimizing the functional

$$J_\tau = \|\mathtt{Z}x - y\|^2 + \|\mathtt{L}(x - x_t)\|^2 \tag{2}$$

where $\|\cdot\|$ is the Euclidian norm. The reason for the second term in $J_\tau$ is that in some cases we do not want to obtain a solution which is too far away from a certain target solution $x_t,$ filtered by an appropriate matrix $\mathtt{L}$. The minimum of (2) is obtained when

$$\mathtt{Z}^H(\mathtt{Z}x - y) + \mathtt{L}^H\mathtt{L}(x - x_t) = 0 \tag{3}$$

where $\mathtt{Z}^H$ is the Hermitian transpose of $\mathtt{Z}$.

It is obvious that Tikhonov regularization means that we completely abandon the original moment matrix formulation for a new one. This makes sense in problems such as deconvolution in the presence of noise, where the matrix $\mathtt{L}$ is strongly related with the noise covariance matrix [14], but it seems less appropriate in electromagnetic problems, where the moment matrix is mostly noise free with a high information content. Also, the introduction of the parameters $\mathtt{L}$ and $x_t$ supposes that we have some *a priori* information on how to select these items, which in practice we do not have. Nevertheless, we can learn a lot from the regularized equation. The solution to (3) obeys the easily proved norm inequality

$$\|\mathtt{Z}x - y\| \le \|\mathtt{Z}^{-H}\mathtt{L}^H\|_2 \|\mathtt{L}(x - x_t)\| \tag{4}$$

where $\|\mathtt{Z}\|_2$ is the matrix $p$-norm with $p = 2$ [7], i.e., in our case, the largest singular value of $\mathtt{Z}$.

Inequality (4) can be written as

$$J_\mu = \frac{||\mathrm{Z}x-y||^2}{||\mathrm{L}(x-x_t)||^2} \leq ||\mathrm{Z}^{-H}\mathrm{L}^H||_2^2. \tag{5}$$

The functional $J_\mu$ in (5) represents a quotient of two positive definite inhomogeneous quadratic forms and, hence, a logical extension of Tikhonov's method is to consider the minimization of $J_\mu$ instead of the minimization of $J_\tau$. It is important to note that the minimum of $J_\mu$ is zero, obtained when $x = x_e$, the exact solution of $\mathrm{Z}x = y$, whereas the minimum of $J_\tau$ does not, in general, correspond with the exact solution.

Unfortunately, $J_\mu$ cannot be considered as a "good" functional. The reason for this is that $J_\mu = \infty$ when $x = x_t$ and, hence, if the exact solution happens to be close to $x_t$, the functional $J_\mu$ will not be anywhere close to zero, as continuity would require.

In other words, what we really need is a functional represented by a bounded quotient of two positive definite inhomogeneous quadratic forms such that the minum occurs at the exact solution. The existence of such a functional follows easily from a generalization of the triangle inequality which is valid for all norms:

$$||x-y||^p \leq 2^{p-1}(||x||^p + ||y||^p) \qquad p > 0. \tag{6}$$

Taking $p = 2$, splitting $\mathrm{Z}$ and $y$ according to

$$\mathrm{Z} = \mathrm{Z}_1 + \mathrm{Z}_2$$
$$y = y_1 + y_2 \tag{7}$$

and applying (6) yields

$$J_{12} = \frac{||\mathrm{Z}x-y||^2}{||\mathrm{Z}_1x-y_1||^2 + ||\mathrm{Z}_2x-y_2||^2} \leq 2. \tag{8}$$

It is seen that the functional $J_{12}$ is appropriate, provided that the solution sets of $\mathrm{Z}_1x - y_1 = 0$ and $\mathrm{Z}_2x - y_2 = 0$ do not intersect. This is certainly the case if we take $\mathrm{Z}_1 = \mathrm{Z}$ and $y_2 = y$. Therefore, we take as our objective functional

$$\Phi = \frac{||\mathrm{Z}x-y||^2}{||\mathrm{Z}x||^2 + ||y||^2} \leq 2. \tag{9}$$

Note that $\Phi = 0$, the minimum, is attained when $x = x_e$ and $\Phi = 2$, the maximum, is attained when $x = -x_e$. It is easily seen that $\Phi$ can be written as

$$\Phi = 1 - 2\frac{\Re[v^H x]}{||\mathrm{Z}x||^2 + ||y||^2} \tag{10}$$

where $\Re$ stands for the real part of a complex number and

$$v = \mathrm{Z}^H y. \tag{11}$$

From (10) we see that the hyperplane $\Re[v^H x] = 0$ divides $x$-space in two half-spaces where $\Phi \leq 1$ and $\Phi \geq 1$, respectively. Hence, in order to have a successful search for a minimum, we must require

$$\Re[v^H x] > 0. \tag{12}$$

Requirement (12) enables us to confine $\Phi$ to the interval $[0, 1]$. This also indicates that $x = v = \mathrm{Z}^H y$ is a logical initial value for $x$. In terms of conditioning, we can interpret the initial value $x_{ini} = \mathrm{Z}^H y$ as the exact solution if $\mathrm{Z}$ were perfectly conditioned, i.e., proportional to a unitary matrix. In the above context it is important to note that the equality

$$\Phi(x) = \xi \leq 1 \tag{13}$$

where $\Phi$ is considered as a function of $x$, can be rewritten as

$$||(1-\xi)\mathrm{Z}x - y|| = ||y||\sqrt{1-(1-\xi)^2}. \tag{14}$$

Equation (14) follows from

$$\begin{aligned} ||(1-\xi)\mathrm{Z}x-y||^2 &= (1-\xi)^2||\mathrm{Z}x||^2 + ||y||^2 + (1-\xi) \\ &\quad \cdot (||\mathrm{Z}x-y||^2 - ||\mathrm{Z}x||^2 - ||y||^2) \\ &= (1-\xi)^2||\mathrm{Z}x||^2 + ||y||^2 + (1-\xi) \\ &\quad \cdot (\xi-1)(||\mathrm{Z}x||^2 + ||y||^2) \\ &= ||y||^2[1-(1-\xi)^2]. \end{aligned}$$

This means that for $\xi \to 0$, we observe a slightly biased behavior: the residual $||\mathrm{Z}x - y||$ does not approach zero directly, but rather indirectly with $x$ scaled to $(1-\xi)x$.

In preparation of the next sections it is useful to obtain an expression for the gradient of $\Phi(x)$. We have

$$g = \frac{2}{||\mathrm{Z}x||^2 + ||y||^2}[(1-\Phi)\mathrm{Z}^H\mathrm{Z}x - v]. \tag{15}$$

If $\mathrm{Z}$ is nonsingular and $y \neq 0$ then it is straightforward to show that $g = 0$ if and only if $\Phi = 0, 2$. This is important because it proves that there are no local extrema. Moreover, the gradient is bounded, as it is not too hard to show that

$$||g|| \leq 2\frac{||\mathrm{Z}^H||_2}{||y||}\left\{(1-\Phi)\sqrt{\Phi} + \Phi\right\} \leq 2\frac{||\mathrm{Z}^H||_2}{||y||} \tag{16}$$

if we remain in the appropriate half-space $\Re[v^H x] > 0$. Even more important, from (16) we deduce that

$$||g|| \leq 2.5\frac{||\mathrm{Z}^H||_2}{||y||}\sqrt{\Phi}. \tag{17}$$

This shows that the gradient is closely bounded by $\sqrt{\Phi}$.

## III. Algorithm 1: Polak–Ribiere

The nonquadratic conjugate gradient algorithm as advocated by Polak–Ribière [15] and modified by Gilbert–Nocedal [11] consists in the following two initialization and four cycle steps.

- $$x_1 = x_{ini} \tag{18}$$
- $$p_1 = -g_1 \tag{19}$$
- $$\alpha_r = \arg\min_{\alpha \in R} \Phi(x_r + \alpha p_r) \tag{20}$$
- $$x_{r+1} = x_r + \alpha_r p_r \tag{21}$$

- $$\beta_r = \max\left\{\Re\left[g_{r+1}^H(g_{r+1}-g_r)\right]/\|g_r\|^2,\, 0\right\} \tag{22}$$
- $$p_{r+1} = -g_{r+1} + \beta_r p_r. \tag{23}$$

It is seen that the line search (20) is the most important step in the nonquadratic conjugate gradient method.

We now show how to find an explicit formula for the real parameter $\alpha$ in our case. Minimization of $\Phi$ along the line $x_r + \alpha p_r$ is equivalent to the maximization problem

$$\alpha = \arg\max \frac{A+\alpha B}{\alpha^2 + 2\alpha C + D} \tag{24}$$

where

$$\begin{aligned} E &= \|\mathtt{Z}p_r\|^2 \\ A &= \Re\left[v^H x_r\right]/E \\ B &= \Re\left[v^H p_r\right]/E \\ C &= \Re\left[(\mathtt{Z}p_r)^H \mathtt{Z}x_r\right]/E \\ D &= \left[\|\mathtt{Z}x_r\|^2 + \|y\|^2\right]/E. \end{aligned} \tag{25}$$

After some calculus, the solution to (24) is obtained as

$$\alpha_r = -\frac{A}{B} + \frac{1}{B}\sqrt{A^2 - 2ABC + B^2 D}. \tag{26}$$

Note that in the limit for $B \to 0$, the above formula remains valid since $\lim_{B\to 0}\alpha_r = -C$, corresponding with the minimum of the denominator of (24). Note also that, since $A + \alpha_r B > 0$, the new value of the unknown vector, i.e., $x_{r+1}$, is always in the appropriate half-space.

By (10), the new value of $\Phi$ is given by

$$\Phi_{r+1} = 1 - \frac{2A + 2\alpha_r B}{\alpha_r^2 + 2\alpha_r C + D}. \tag{27}$$

The new gradient $g_{r+1}$ is then given by

$$g_{r+1} = \nu_{r+1}\left\{(1-\Phi_{r+1})\mathtt{Z}^H[\mathtt{Z}x_r + \alpha_r \mathtt{Z}p_r] - v\right\} \tag{28}$$

where

$$\nu_{r+1} = \frac{1-\Phi_{r+1}}{AE + \alpha_r BE}. \tag{29}$$

At first sight it would seem that we have to perform three matrix–vector multiplications per cycle step, but this is incorrect since at stage $r+1$ we can utilize the fact that

$$\mathtt{Z}x_{r+1} = \mathtt{Z}x_r + \alpha_r \mathtt{Z}p_r. \tag{30}$$

Hence, the computational load is approximately two matrix–vector multiplications per cycle step, exactly as in the normal equations version of the conjugate gradient method [5].

After some cycle steps the conjugate gradient algorithm is likely to stall and, therefore, the algorithm is restarted after a number of steps with the current value of $x$. At restart time it is also judicious to rescale $x$, i.e., to find the complex parameter $\gamma$ such that $\Phi(\gamma x)$ is minimal and then to replace $x$ with $\gamma x$. It is easy to prove that the explicit expression for $\gamma$ is

$$\gamma = e^{-i\theta}\frac{\|y\|}{\|\mathtt{Z}x\|} \tag{31}$$

where

$$\theta = \arg\left[v^H x\right]. \tag{32}$$

The value of $\Phi$ after rescaling is

$$\Phi = 1 - \frac{|v^H x|}{\|\mathtt{Z}x\|\|y\|}. \tag{33}$$

Since $\Phi$ is always $\geq 0$, (33) actually constitutes another alternative proof of the Cauchy–Schwartz inequality. It is important to note that after rescaling we have $\|\mathtt{Z}x\| = \|y\|$, which means that $x$ belongs to the surface of the ball $\|x\|_Z = \|y\|$ with respect to the norm $\|x\|_Z = \|\mathtt{Z}x\|$.

## IV. Algorithm 2: Total Rescaling

The weak point of Algorithm 1 is the possible suboptimal value of the accelerator $\beta_r$ in (23). To alleviate this, we combine scaling, complex line search and accelerator all in one, i.e., we modify the equations (18)–(23) to yield

- $$x_1 = x_{ini} \tag{34}$$
- $$p_0 = 0 \tag{35}$$
- $$(\gamma_r, \alpha_r, \beta_{r-1}) = \arg\min_{(\gamma,\alpha,\beta)\in C^3} \Phi(\gamma x_r + \alpha(-g_r + \beta p_{r-1})) \tag{36}$$
- $$p_r = -g_r + \beta_{r-1}p_{r-1} \tag{37}$$
- $$x_{r+1} = \gamma_r x_r + \alpha_r p_r. \tag{38}$$

In this manner, we do not need a formula for the accelerator $\beta_r$, since we can obtain the optimal complex accelerator $\beta_{r-1}$ at the previous stage by solving the minimization problem (36). This seemingly difficult problem has a surprisingly simple solution. Defining the $N \times 3$ matrix

$$\mathtt{V} = [\mathtt{Z}x_r,\, \mathtt{Z}g_r,\, \mathtt{Z}p_{r-1}] \tag{39}$$

and the $3 \times 1$ vector

$$u = \begin{bmatrix} \gamma_r \\ -\alpha_r \\ \alpha_r\beta_{r-1} \end{bmatrix} \tag{40}$$

it is seen that

$$\mathtt{Z}x_{r+1} = \mathtt{V}u. \tag{41}$$

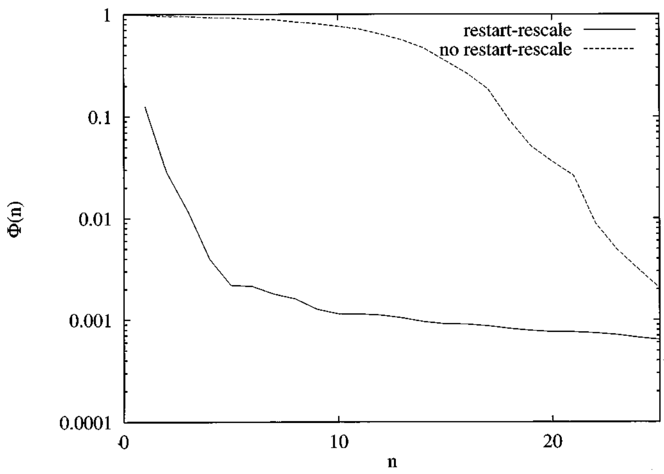


Fig. 1. Convergence rate of algorithm 1 with and without restart–rescale.

Now it is easy to show that the minimization of

$$\Phi = \frac{||\mathbb{V}u - y||^2}{||\mathbb{V}u||^2 + ||y||^2} \tag{42}$$

yields

$$u = \frac{1}{1-\Phi}(\mathbb{V}^H\mathbb{V})^{-1}\mathbb{V}^H y = \frac{1}{1-\Phi}\,\tilde{u} \tag{43}$$

where $\tilde{u}$ is the least squares solution pertaining to the minimization of $||\mathbb{V}\tilde{u} - y||$. After some calculations we obtain the minimum value of $\Phi$

$$1 - \Phi = ||\mathbb{V}\tilde{u}||/||y|| \tag{44}$$

implying that $||\mathbb{V}u|| = ||\mathbb{Z}x_{r+1}|| = ||y||$. Note that the complex accelerator for the total rescaling algorithm is $\beta_{r-1} = -u_3/u_2$.

## V. Numerical Simulations

First, we consider the TM scattering by a plane wave of a square PEC cylinder at diagonal incidence. The circumference of the square is $20\lambda$ and, hence, with the conventional pulse basis and point matching method [16] with pulse width $\lambda/10$ we obtain a symmetric moment matrix of dimension $N = 200$. The condition number is $\kappa = 150.8$. The convergence rate of algorithm 1 without and with restart–rescale after every fifth step is shown in Fig. 1. It is seen that restart-rescale is imperative in order to obtain a decent convergence rate. Next we compare algorithm 2 (total rescaling) with algorithm 1 (restart–rescale) and normal CG [6]. The convergence rate is shown in Fig. 2. It is seen that algorithm 2 produces the best results.

In Fig. 3 we plot the condition number $\kappa$ as a function of the dimension $N$, for a square cylinder of increasing circumference. It is seen that $\kappa$ is a wildly oscillating function of $N$. Additionally, we see a strong maximum of $\kappa = 7863$ at $N = 412$.

The convergence rate for worst case dimension $N = 412$ for the three algorithms, under the same conditions as the first example, is shown in Fig. 4. Again, algorithm 2 seems to outperform the other algorithms.

As a last example we again take $N = 412$, but with an excitation due to a line source situated on the diagonal at a distance

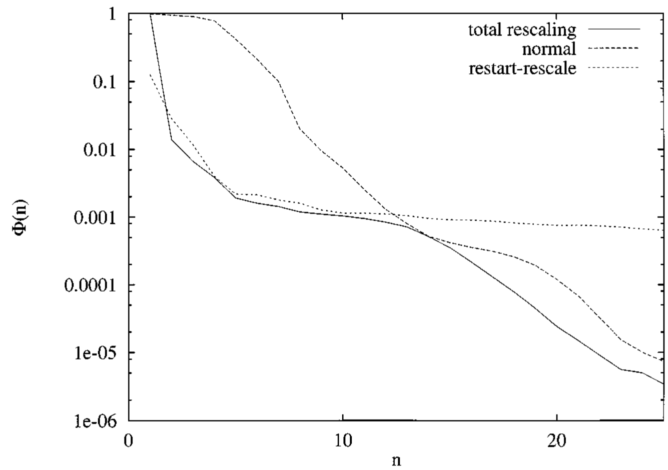


Fig. 2. Comparison of convergence rates of algorithm 2, normal CG and algorithm 1, for $N = 200$ and plane wave excitation.

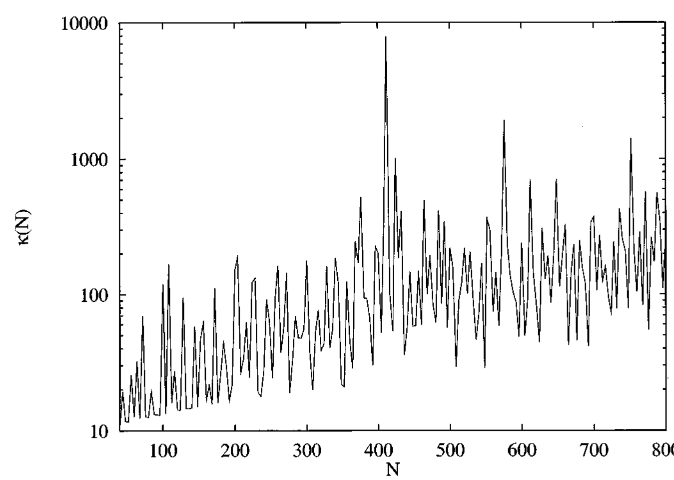


Fig. 3. The condition number as a function of dimension $N$.

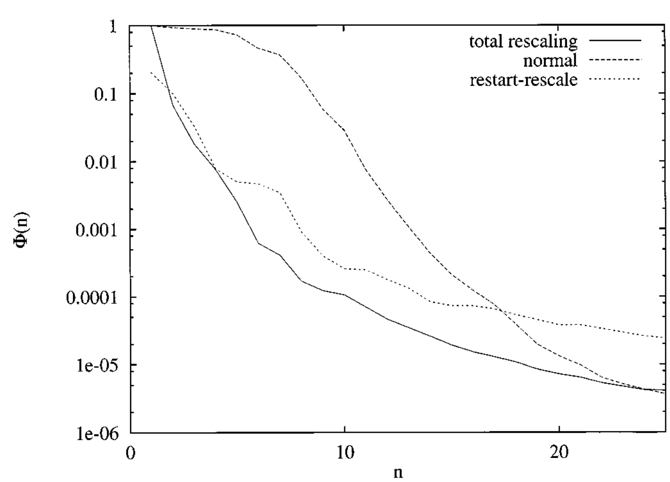


Fig. 4. Comparison of convergence rates of algorithm 2, normal CG and algorithm 1, for $N = 412$ and plane wave excitation.

$4\lambda$ from the tip of the square cylinder. The convergence rate for the three algorithms is shown in Fig. 5.

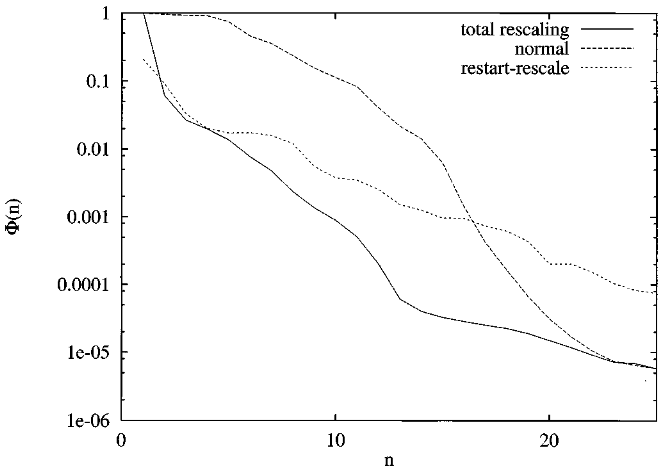


Fig. 5. Comparison of convergence rates of algorithm 2, normal CG and algorithm 1, for $N = 412$ and line-source excitation.

It is seen that the overall performance of algorithm 2 (total rescaling) is better than normal CG, and that algorithm 1 (Polak–Ribière with restart–rescale) is not as performant as normal CG. Hence, we conclude from our numerical simulations that the total rescaling algorithm is the better one in the sense that its convergence rate has a very early plunge region, leading to an acceptably small value of the objective function $\Phi$ at an early stage of the algorithm.

## VI. Conclusion

The main advantage of the nonquadratic CG method is that we minimize a bounded functional. Hence, in contradistinction with the normal CG method, which minimizes an unbounded functional, we know at every step of the algorithm exactly how close we are to the minimum.

By performing exact line-searches and total rescaling procedures, the algorithm keeps the functional in a tight grip on its strictly decreasing path to the minimum. The relationship with Tikhonov regularization ensures the numerical stability of the method.

## References

[1] F. X. Canning and J. F. Scholl, "Diagonal preconditioners for the EFIE using a wavelet basis," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 1239–1246, Sept. 1996.

[2] D. Colton and R. Kress, *Integral Equation Methods in Scattering Theory*. New York: Wiley, 1983.

[3] M. Z. Nashed, "Operator-theoretic and computational approaches to ill-posed problems with applications to antenna theory," *IEEE Trans. Antennas Propagat.*, vol. AP-29, pp. 220–231, Mar. 1981.

[4] J. R. Ringrose, *Compact Non-Self-Adjoint Operators*. London, U.K.: Van Nostrand, 1971.

[5] A. F. Peterson, C. F. Smith, and R. Mittra, "Eigenvalues of the moment-method matrix and their effect on the convergence of the conjugate gradient algorithm," *IEEE Trans. Antennas Propagat.*, vol. 36, pp. 1177–1179, Aug. 1988.

[6] L. Knockaert, "A note on the relationship between the conjugate gradient method and polynomials orthogonal over the spectrum of a linear operator," *IEEE Trans. Antennas Propagat.*, vol. AP-35, pp. 1089–1091, Sept. 1987.

[7] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: Johns Hopkins Univ., 1996.

[8] G. C. Hsiao and R. E. Kleinman, "Mathematical foundations for error estimation in numerical solutions of integral equations in electromagnetics," *IEEE Trans. Antennas Propagat.*, vol. 45, pp. 316–328, Mar. 1997.

[9] R. J. Santos, "Equivalence of regularization and truncated iteration for general ill-posed problems," *Linear Algebra Applicat.*, vol. 236, pp. 25–33, 1996.

[10] R. Fletcher and C. M. Reeves, "Function minimization by conjugate gradients," *Comput. J.*, vol. 7, pp. 149–154, 1964.

[11] J. C. Gilbert and J. Nocedal, "Global convergence properties of conjugate gradient methods for optimization," *SIAM J. Optimization*, vol. 2, no. 1, pp. 21–42, Feb. 1992.

[12] A. Edelman and S. T. Smith, "On conjugate gradient-like methods for eigen-like problems," *BIT*, vol. 36, no. 3, pp. 494–508, Oct. 1996.

[13] M. J. D. Powell, "Restart procedures for the conjugate gradient method," *Math. Programming*, vol. 12, pp. 241–254, 1977.

[14] T. K. Sarkar, D. D. Weiner, and V. K. Jain, "Some mathematical considerations in dealing with the inverse problem," *IEEE Trans. Antennas Propagat.*, vol. AP-29, pp. 373–379, Mar. 1981.

[15] E. Polak and G. Ribière, "Note sur la convergence de méthodes de directions conjuguées," *Revue Française d'Informatique et de Recherche Opérationelle*, vol. 16, pp. 35–43, 1969.

[16] R. L. Wagner and W. C. Chew, "A study of wavelets for the solution of electromagnetic integral equations," *IEEE Trans. Antennas Propagat.*, vol. 43, pp. 802–810, Aug. 1995.

**Luc Knockaert** (M'81) received the Physical Engineer, Telecommunications Engineer, and Ph.D. degrees in electrical engineering from the University of Gent, Belgium, in 1974, 1977, and 1987, respectively.

From 1979 to 1984 and from 1988 to 1995, he was working in North–South cooperation and development projects at the Universities of Zaire and Burundi. He is presently a Senior Researcher at INTEC-IMEC. He has contributed to about 25 international journal papers as author or coauthor . His current interests are in the application of statistical and linear algebra methods in signal identification, Shannon theory, matrix compression, reduced-order modeling, and the use of integral transforms in electromagnetic problem solving.

Dr. Knockaert is a member of ACM.

**Daniël De Zutter** (M'92–SM'96) was born in Eeklo, Belgium, November 8, 1953. He received the Electrical Engineering and Ph.D. degrees from the University of Gent, Belgium, in 1976 and 1981, respectively.

From September 1976 to September 1984, he was a Research and Teaching Assistant in the Laboratory of Electromagnetism and Acoustics (now the Department of Information Technology) at the same university. In the spring of 1984 he completed a thesis leading to a degree equivalent to the French Aggrégation or the German Habilitation. He is now a Full Professor at the Department of Information Technology, University of Gent. He has contributed to about 60 international journal papers and 70 papers in conference proceedings as author or coauthor. Most of his earlier scientific work dealt with the electrodynamics of moving media, with emphasis on the Doppler effect and on Lorentz forces. His research now focuses on all aspects of circuit and electromagnetic modeling of high-speed and high-frequency interconnections, electromagnetic compatibility and electromagnetic interference topics, and indoor propagation.

Dr. De Zutter received the 1995 Microwave Prize Award.