THE MERITS OF ANALYTICAL REGULARIZATION AND ANALYTICAL PRECONDITIONING

Alexander I. Nosich

Institute of Radio-Physics and Electronics NASU
Ul. Proskury 12, Kharkov 61085, Ukraine. E-mail: alex@emt.kharkov.ua

1. Introduction

Usually antennas and many other electromagnetic devices have to be analyzed in infinite spatial domains. This leads to finding solutions to exterior problems for the electromagnetic fields and waves; additionally, in many applications, simulating the time-harmonic-field performance ($\sim e^{-ikct}$, k and c being the free-space wavenumber and light velocity, respectively) is crucial. Thus, one comes to wave scattering and radiation analysis based on boundary-value problems for the harmonic-Maxwell and Helmholtz equations. Uniqueness of their solution is normally guaranteed by a sufficient set of boundary, edge, and radiation conditions. Although one may try to find it by using finite-difference discretizations of the partial differential equations, associated problems of domain truncation, "good" exterior meshing, and solving enormous matrices are hardly compatible with high accuracy. To avoid these, boundary-element and Green's function methods can be used, applied to integral-equation (IE) formulations. Of the IE advantages, two main points are to be emphasized: the radiation condition is automatically taken into account, and only the finite domains or their boundaries need to be discretized. However, this frequently generates ill-conditioned dense matrices, and so something should be done to adapt these formulations to meeting the combined challenge of the speed and accuracy of computations.

Many textbooks and journal papers on computational electromagnetics deal with IEs for determining the surface or polarization currents of two-dimensional and three-dimensional metallic or dielectric scatterers, respectively, given the incident field. Such IEs are obtained from the boundary conditions, and many types of them are first-kind and normally have logarithmic-type or higher singular kernels. We shall assume that when passing from boundary-value problem to an IE, uniqueness of solution is preserved, although some types of IEs loose it on discrete sets of frequencies. IEs are further discretized for a numerical solution with Method-of-Moments (MoM) by using subdomain (collocations) or entire-domain basis functions. Although this commonly brings meaningful and useful results, unfortunately, there are not any theorems proving general MoM convergence, or even the existence of an exact solution, for such IEs [1]. A rule-of-thumb of taking at least 10 mesh points per wavelength is only a rule-of-thumb, and by no means does it guarantee any number of correct digits.

A simple example of this sort is the logarithmic-singular IE in the two-dimensional (2D) E-wave scattering from a PEC (perfectly electrically conducting) flat or curved strip. It is well known that here the condition number of conventional MoM matrix grows with the number of equations, thus making the matrix impossible to solve for accuracy better than several digits and a strip greater than 10-20 wavelengths [3]. Nearly the same can be said of the MoM discretizations of second kind IEs having strongly singular kernels. A good demonstration of what may sometimes happen to such algorithms was given in [2]. By the simple example of 2D plane-wave scattering from a tubular circular cylinder, it was shown that the MoM and FDTD solutions could be 1000% or more in error in a vicinity of sharp resonance. The "pain points" of the conventional MoM approach have been excellently reviewed in [1]; since then, essentially nothing has changed. The final statement of [1] is worth reciting: "It is misleading to refer to the result as *solution* when in fact it is *numerical approximation* with no firm mathematical estimate of nearness to solution." Although it is possible to eliminate ill conditioning by using specialized discretization schemes, based on the Sobolev-space inner products, this appears to have a limited range of application.

Meanwhile, there exists a general approach to obtaining second kind IEs of the Fredholm type, with a smoother kernel, from first-kind equations. Discretization of these new equations, either by collocation or by a Galerkin-type projection on a set of basis functions, generates matrix equations, whose condition numbers remain small when the number of mesh points or "impedance-matrix" size is progressively increased. The approach mentioned is collectively called the Method of Analytical Regularization (MAR) [4,5]. The term has apparently been introduced by Muskhelishvili [6]; sometimes *semi-inversion* is used as a synonym. It is based on the identification and analytical inversion of the whole singular part of the original IE or its most singular component. It must be admitted that the whole idea of MAR can be traced back to the pioneering work of the founders of singular IEs theory, Hilbert, Poincare and Noether, well before the first appearance of a computer.

2. MAR as conversion to a Fredholm second kind IE

To present a formal scheme of MAR, assume that the boundary condition and some suitable integral representation of the unknown field function (e.g., potential theory) generate a first kind singular IE. In operator notation, this can be written as

$$GX = Y , \tag{1}$$

where X and Y stand for the unknown and given function, respectively. A direct analytical inversion of G is normally not possible, while a numerical inversion, as has been mentioned, has no guaranteed convergence. Split operator G into two parts: $G=G_1+G_2$. Assume now that the first of these has a known inverse, $G = G_1^{-1}$. Then, by acting with this operator on the original equation, one obtains a second kind IE:

$$X + AX = B , \tag{2}$$

where $A=G_1^{-1}G_2$ and $B = G_1^{-1}Y$. However, this scheme is mathematically justified only if the resulting IE (2) is of the Fredholm type. This means that the operator A must be compact, e.g., has a bounded L_2-norm, $\| A \|_{L_2} < \infty$, and the right-hand side B must belong to the same space L_2. This inherently implies that the inverted operator G_1 is more singular than G_2. Then, all the power of the Fredholm theorems generalized for operators [7,8] can be exploited, proving both the existence of an exact solution, $X = (I + A)^{-1} B$ (I is the identity operator), and the point-wise convergence of discretization schemes, without resorting to residual-error estimations like for first-kind IEs [1]. Indeed, suppose that we have discretized the second-kind equation (2) by projecting it on some set of local or global expansion functions. Besides of obvious observation that such a set should possess the completeness in L_2, note that matrix counterpart of (2) can be equivalent to IE only if this set is infinite, and in this case the latter is the Fredholm second kind in the discrete space l_2 for the unknown expansion coefficients. Considering its "truncated" counterpart, with the matrix A^N filled in with zeros off the $N \times N$ square, it is easy to show the following estimation for the relative error, by the norm in l_2:

$$e(N) = \| X - X^N \| (\| X \|)^{-1} \leq \| (I + A)^{-1} \| \| A - A^N \| \tag{3}$$

It is clear that (3) destined to go to zero with $N \to \infty$, as the first factor in the right-hand part above is a bounded constant, while the second is decreasing. Theoretically, in finite-digit arithmetic, this decrement is limited by the machine precision and, in the practical computations, by the accuracy of intermediate operations (e.g., numerical integration for filling in A). The rate of decay of $e(N)$ determines the *cost of the algorithm,* and this can be different for different ways of selecting the invertible singular part, G_1. Here comes a key question: how to select the operator G_1? It is apparently possible to point out at least three or four basic principles for extracting an invertible singular part of the original operator. These are extracting the *static part*, the *high-frequency part* (in fact, this is about half-plane scattering, which can be solved by the Wiener-Hopf method), and the frequency-dependent *canonical-shape* part, which corresponds to either a circle, in 2D, or to a sphere, in 3D, solvable by separation of variables. Note that a half plane is also a sort of canonical-shape scatterer, but it has an infinite surface in terms of any length parameter, including the wavelength. The fourth principle is to invert the *small-contrast* part of the problem, as used in the scattering by dielectric objects.

3. Preconditioning as conversion to a Fredholm second-kind matrix equation avoiding intermediate IE

Although inversion of the static or high-frequency part seems to be based on quite specialized functional techniques, it is useful to point out one general feature in all the above cases. If it is possible to find a *set of orthogonal eigenfunctions* of the separated singular operator G_1, then the Galerkin-projection technique applied to original singular IE of the first kind, i.e. (1), with these functions as a basis, immediately results in a regularized discretization scheme (i.e., yields a Fredholm second-kind infinite matrix-operator equation). This is especially evident with MAR based on canonical-shape inversion, as in this case, the orthogonal eigenfunctions are just trigonometric (entire-order azimuth exponents) or spherical polynomials (products of the former with the Legendre functions). Another simple example is the 2D scattering by flat PEC strip – here the weighted Chebyshev polynomials invert logarithmic or hyper-singular static part of the full-wave IE operator, as discussed in [5]. Such a judicious projection, in fact, combines both regularization (semi-inversion) and discretization in one single procedure, and avoids intermediate step of a Fredholm second kind IE. One may easily see that it bridges the gap between MAR and conventional MoM solutions. Indeed, the intuitive idea that a good choice of expansion functions in MoM can facilitate convergence obtains the form of a clear mathematical rule: to have the convergence guaranteed, take the expansion functions as orthogonal eigenfunctions of G_1. The procedure of finding such functions is called *diagonalization* of a singular integral operator. From the viewpoint of numerical analysis, this procedure plays the role of a perfect pre-conditioning of the original IE, the direct discretizations of which are ill conditioned. Pre-conditioning eigenfunctions can be global in the case of simpler shapes such as strip. However, one can view complicated boundaries as collections of *M* simpler elements, introduce locally-global basis functions diagonalizing integral operators at these elements, and then obtain a $M \times M$ block-type infinite-matrix equation of the Fredholm second kind, in the space of sequences l_2^M. To distinguish it from classical MAR of the IE theory, one can call the outlined scheme a Method of Analytical Preconditioning (MAP).

4. Analytical solutions

As has been emphasized, MAR/MAP solutions, based on the Fredholm second-kind matrix equations, have a guaranteed point-wise convergence, and thus a controlled accuracy of numerical results. Provided that all intermediate computations necessary for filling in the matrix and the right-hand part have been done with superior accuracy, the parameter controlling the final accuracy is just the size of the matrix, i.e., the order of its truncation. Depending on the nature of the inverted part, the number of equations needed for a practical two-to-three-digit accuracy is usually slightly greater than, respectively, the electrical dimension of the scatterer, or its inverse value, or the normalized deviation of the surface from the canonical shape, in terms of both distance and curvature. In fact, the norm of the compact operator *A* is always proportional to one of the above-mentioned values, denoted as, say, κ, so that $\| A(\kappa) \|_{L_2} < const\, \kappa$. This enables one to exploit an important feature of the Fredholm second-kind equations. Provided that $\| A(\kappa) \|_{L_2} < 1$, which can always be satisfied for a small enough κ, an iterative solution to (2) is given by the Neumann-operator series

$$X = \sum_{s=0}^{\infty} [-A(\kappa)]^s B \, , \tag{4}$$

which converges by norm to the exact solution. Hence, one can avoid inverting (2), at least in a certain domain of parameters. If, for example, the IEs based on the static and high-frequency parts inversion have overlapping domains of the Neumann series convergence then the need of solving a matrix is completely eliminated – one has only to perform matrix-vector multiplication. Besides numerical efficiency, this has another attractive consequence. On expanding (4) in terms of the power series of κ, one obtains, analytically, rigorous asymptotic formulas for the low-frequency or high-frequency scattering, or the scattering from a nearly-canonical or small-contrast object. Such asymptotics have been published for PEC flat and circular strips, finite pipes, disks, a spherical cap, and PEC and imperfect strip gratings. What is worth noting is that this can be done for various excitations specified by *B:* plane or cylindrical waves, a complex source-point beam, a surface wave in the layered-media scattering, etc.

5. Eigenvalue problems

These problems are closely tied to the wave-scattering problems. They can be classified as either natural-frequency or natural-wave problems, although other eigen-parameters can be also considered. The natural-wave problems appear only in the analysis of infinite cylindrical geometries, assuming a traveling-wave field solution, i.e. $\sim e^{ihz-ikct}$. Correspondingly, the (complex-valued) eigenvalue to be determined is *k* or the modal wavenumber, *h* (the propagation constant). What is important, in either case, is that a MAR solution leads to a homogeneous equation analogous to the scattering problem:

$$X + A(h,k)X = 0 \tag{4}$$

This is a Fredholm operator equation*,* with compact operator *A* normally being a continuous function of the geometrical parameters and a meromorphic function of the material parameters, *k*, and *h*. Hence, due to the Steinberg theorems [8], it is guaranteed that the eigenvalues form a discrete set on a complex *k* plane (in the three-dimensional case), or on a logarithmic Riemann surface of $Ln\,k$ or $Ln\,h(k)$ - in the two-dimensional case and for the natural waves, respectively. There are no finite accumulation points; eigenvalues can appear or disappear only at those values of the other parameters where continuity or analyticity of A is lost. Moreover, after discretization, the determinant of the infinite matrix $Det[I + A(k,h)]$ exists as a function of parameters, and its zeroes are the needed k or h eigenvalues. The latter are piece-wise continuous or piece-wise analytic functions of geometrical and material parameters: these properties can be lost only at the points where two or more eigenvalues coalesce. From a practical viewpoint, it is important that eigenvalues can be determined numerically: the convergence of discretization schemes is guaranteed, the number of equations needed being dependent on the desired accuracy and the nature of the inverted part. No spurious eigenvalues appear, unlike many approximate numerical methods. Note that nothing of the above can be established for an infinite-matrix equation of the first kind, which is common in conventional MoM analyses.

Additionally, if a corresponding parameter κ is small, then the matrix A is quasi-diagonal, and the eigenvalues of k or h can be obtained in the form of an asymptotic series. Such an analytical study has been done for a PEC axially slotted cylinder and a spherical cap, assuming a narrow slot or a small circular aperture, respectively. These asymptotics serve as a perfect starting guess when searching for eigenvalues numerically, with a Newton or another iterative algorithm – this has been done, e.g., for the modes on planar and circular-cylindrical strip and slot lines, and for arbitrary-cross-section dielectric waveguides.

6. Conclusions

Summarizing, by using the MAR/MAP it is possible to overcome many difficulties encountered in conventional MoM treatments. Theoretical merits of the MAR are numerous: exact solution existence is established, convergence is guaranteed, and rigorous asymptotic formulas can be derived. Computationally, the MAR results in a small matrix size for a practical accuracy, and sometimes no numerical integrations are needed for filling in the matrix. Thus, the *cost of MAR algorithms is low* in terms of both CPU time and memory. A frequent feature is that both power conservation and reciprocity are satisfied at the machine-precision level, independently of the number of equations, whatever it is. As $(I + A)^{-1}$ is bounded, the condition number is small and stable, not growing with mesh refinement or increasing with the number of basis functions. The latter fact means that conjugate-gradient numerical algorithms are very promising even in spite of a possible squaring of the condition number, as has already been emphasized in [9]. Using fast iterative methods, applied to the MAR matrix equations with static-part inversion, it is probably possible to perform an accurate full-wave desktop analysis of the Arecibo reflector. However, the same thing can be done more economically by using a high-frequency inversion.

From a practical viewpoint, it is also important that the accuracy of the MAR/MAP is uniform, including resonances, both in near-field and far-field predictions. Here, one must be reminded that near sharp resonances, conventional Moment-Method and FDTD solutions suffer heavy inaccuracy [2], which cannot be removed, in principle. All this makes MAR/MAP algorithms perfect candidates for CAD software in the numerical optimization of multielement 2D and 3D scatterers in the so-called resonant range, where interaction between separate elements plays an important role, and in the quasioptical range, where both ray-like and mode-like phenomena coexist. In fact, quite complicated 2D models of reflector and lens antennas, radomes, and open resonators have been already accurately studied, showing a variety of features not predicted by ray-tracing techniques and inaccessible by rough numerical approximations.

References

1. D.G. Dudley, Error minimization and convergence in numerical methods, *Electromagnetics*, vol. 5, no 2-3, pp. 89-97, 1985.
2. G.L. Hower, R.G. Olsen, J.D. Earls, J.B. Schneider, Inaccuracies in numerical calculation of scattering near natural frequencies of penetrable objects, *IEEE Trans. Antennas Propagat*., vol. AP-41, no 7, pp. 982-986, 1993.
3. C. Klien, R. Mittra, Stability of matrix equations arising in electromagnetics, *IEEE Trans. Antennas Propagat.*, vol. 21, no 6, pp. 902-905, 1973.
4. A.I. Nosich, The MAR in wave-scattering and eigenvalue problems: foundations and review of solutions, *IEEE Antennas Propagat. Magazine*, vol. 41, no 3, pp. 25-49, 1999.
5. G. Fikioris, A note on the MAR, *IEEE Antennas Propagat. Magazine*, vol. 43, no 2, pp. 34-40, 2001.
6. N.I. Muskhelishvili, *Singular Integral Equations*, Groningen, Noordhoff, 1953.
7. D. Colton, R. Kress, *Integral Equation Methods in Scattering Theory*, New York, Wiley, 1983, ch. 4.
8. S. Steinberg, Meromorphic families of compact operators, *Arch. Rat. Mechanics Analysis*, vol. 31, no. 5, pp. 372-379, 1968.
9. V. Rokhlin, Rapid solution of IEs of scattering theory in 2D, *J. Computat. Physics*, vol. 86, no 2, pp. 414-439, 1990.