Telecommunications and Radio Engineering, 70(10):843-856 (2011)

MATHEMATICAL METHODS IN ELECTROMAGNETIC THEORY

PROJECTION APPROXIMATIONS TO THE MATRIX SCATTERING OPERATORS AND RELATIVE CONVERGENCE PHENOMENON

I.V. Petrusenko & Yu.K. Sirenko[♦]

A. Usikov Institute of Radio Physics and Electronics,
National Academy of Sciences of Ukraine
12, Academician Proskura St., Kharkiv 61085, Ukraine

[♦]Address all correspondence to Yu.K. Sirenko E-mail: yks@ire.kharkov.ua

A matrix-operator model of the mode-matching technique is examined for a scalar problem of mode diffraction on abrupt discontinuity in the waveguide. The convergence of approximations, which are found via a truncation procedure, to the matrix operators of mode reflection and transmission is studied analytically. Various estimates for these projection approximations are obtained. On the basis of these results, the unconditional and relative convergence phenomena are analyzed. The impact of the "Mittra rule" upon the rate of convergence is validated analytically. The uniform-convergence condition for scattering operator approximations is stated. Properties of the condition number for the truncated matrix-operator equation are discussed. The results thus obtained may well be used to substantiate a highly practically class of mathematical models related to the mode-matching technique.

KEY WORDS*: mode-matching technique, truncation procedure, Cayley transform*

1. INTRODUCTION

The findings of recent papers [1-3] indicate that the mode-matching technique for scalar problem on mode-diffraction by abrupt discontinuity in the waveguide yields not only the infinite systems of linear algebraic equations, as it has been earlier believed, but also the Fresnel formulae for matrix operators of mode reflection and transmission. Furthermore, the popular matrix models are found to be set up as a sequel of the specifically formulated mode-diffraction problem; they represent a reduced version of the operator-based Fresnel formulae and result in a fatal loss of knowledge about the properties of the sought-for solution and of the given matrix operator.

In other words, it was found that the mode-matching technique for the scalar mode-diffraction problems is of matrix-operator nature and an adequate mathematical apparatus for this technique is based on the theory of operators in the Hilbert space and in the Pontrjagin space [4].

The above reasoning and the derived operator-based form of the power conservation law (see [1,4]) have allowed us to justify the correctness of mathematical models related to the mode-matching technique for the mentioned class of mode-diffraction problems [3], namely, to solve the problem that had remained topical far as long as the period of applying this technique permitted.

In the present work we proceed with further evolving the theory of mode-matching technique just from this new standpoint. Herein we go into another open question known as the problem of validity of the truncation procedure for matrix models of the mode-matching technique. The pivotal feature about this problem is the relative convergence phenomenon [5], which is not adequately studied.

The aim of the present work is to construct projection approximations for operator-based Fresnel formulae and to make an analytical study into the qualitative characteristics of their convergence. We derive different estimates for these approximations, and by their convergence we imply the projection convergence (or P-convergence), which is of particular interest in practical computations (see the reference to P-convergence theory in [6,7]).

The compositional principle of the present paper is based upon the following sequence of the results obtained in the course of the analytical study: (a) the basic theorem about the unconditional strong P-convergence of approximations; (b) the theorem about the function of the "Mittra rule"; (c) the condition for the uniform P-convergence of approximations; (d) the estimates of the condition number for the truncated matrix equation.

All the matrix operators under consideration belong to the Banach algebra $\mathcal{B}(H)$ of bounded operators defined on the whole Hilbert space H. Throughout the text $H \equiv \ell_2$ except for the proof of Lemma 1, in which the weighted spaces are utilized.

2. PROBLEM FORMULATION AND NOTATIONS

The abrupt discontinuity under study is made up of the plane junction of two regular waveguides of the rectangular cross-section $\Omega_p \times d$, $p = 1,2$, which have an equal height d (hereinafter without any loss of generality we set $\Omega_2 \subseteq \Omega_1$). The reference plane is brought into coincidence with an aperture of discontinuity Ω_0. The waveguide walls are assumed to be perfect electric conductors; the waveguides are filled with a homogeneous lossless medium and terminated in matching loads. Both of the waveguides are the simple partial regions in which independent field sources of the frequency ω are positioned. The time-factor $\exp(i\omega t)$ is omitted throughout.

The diffracting waveguide modes $LM_{m0}, m=\overline{1,\infty}$ or $LE_{m1}, m=\overline{0,\infty}$ are taken as real-valued cross-eigenfunctions that are collected in the column-vector $\varphi_p = \left\{\varphi_m^{(p)}(x)\right\}_{m=(0)1}^{\infty}, x\in\Omega_p$, and as propagation constants $\left\{\gamma_m^{(p)}\right\}_{m=(0)1}^{\infty}$, $p=1,2$, such that $\operatorname{Im}\gamma_m^{(p)}>0$ for propagating waves and $\operatorname{Re}\gamma_m^{(p)}>0$ for evanescent modes. The nuclear self-adjoint integral operators discussed below are likewise characterized by the column-vector of eigenfunctions $\varphi_0 = \left\{\varphi_m^{(0)}(x)\right\}_{m=1}^{\infty}, x\in\Omega_0$ (see formula (22)). Henceforward, we shall proceed from the fact all the above eigenfunctions constitute the maximal orthonormal systems. Specifically, in terms of the identity operators we have

$$\begin{aligned}&\varphi_q^{\dagger}(x)\varphi_q(x') = \delta(x-x'),\ q=0,1,2;\\&\left(\varphi_p,\varphi_p^T\right)_{\Omega_p} = \mathbf{I};\qquad \left(\varphi_0,\varphi_0^{\dagger}\right)=\mathbf{I}.\end{aligned}\tag{1}$$

The scalar (bilinear) product of functions is hereafter symbolized by parentheses and, when integrating over the aperture of discontinuity, the subscript Ω_0 is omitted; the dagger " † " denotes the Hermitian conjugation and the superscript " T " is for transposition.

Now introduce the Hilbert spaces of sequences of complex quantities

$$h_{\pm1} \equiv \left\{\tilde{\mathbf{b}} = \left\{\tilde{b}_n\right\} : \sum_{n=1}^{\infty} n^{\pm1}\left|\tilde{b}_n\right|^2 < \infty\right\},\quad \ell_2 \equiv \left\{\mathbf{b} = \left\{b_n\right\} : \sum_{n=1}^{\infty}\left|b_n\right|^2 < \infty\right\},$$

for which the inclusions of $h_{+1}\subset \ell_2 \subset h_{-1}$ take place. The afore-mentioned wave propagation constants originate the diagonal "matrix operators of similarity" $\mathbf{I}_{\gamma p}^{\pm1/2} = \left\{\left(\gamma_m^{(p)}\right)^{\pm1/2}\delta_{mn}\right\}$, $p=1,2$, which, as it is easy to check, are bounded on the pairs of spaces $h_{+1}\rightleftarrows \ell_2$, $\ell_2 \rightleftarrows h_{-1}$ (here the symbol δ_{mn} is the Kronecker delta, whereas the cut-off points for which $\exists m: \gamma_m^{(p)}=0$ are excluded as nonphysical ones). In subsequent reasoning we rely upon the following

Lemma 1. The matrix operator $\mathbf{V}_{pq} = \left(\varphi_p,\varphi_q^{\dagger}\right)$, $p,q=0,1,2$ is bounded in each of the spaces $h_{\pm1}$ and ℓ_2.

Proof. Indeed, the matrix operators

$$\begin{aligned}&\mathbf{W}_p \equiv \mathbf{V}_{pq}\mathbf{V}_{pq}^{\dagger} = \left(\varphi_p,\varphi_p^{\dagger}\right);\\&\mathbf{W}_q \equiv \mathbf{V}_{pq}^{\dagger}\mathbf{V}_{pq} = \left(\varphi_q,\varphi_q^{\dagger}\right)\end{aligned}\tag{2}$$

are defined in each of these Hilbert spaces; they are continuous, self-adjoint and idempotent, i.e., $\mathbf{W}^2 = \mathbf{W}$ by virtue of properties (1). Consequently, $\mathbf{W}$ is an ortho-projector and, hence, $\|\mathbf{W}\| = 1$. From relations (2) then follows $\|\mathbf{V}_{pq}\| = 1$.

Corollary. The matrix operators of the form $\mathbf{I}_{\gamma p}^{\pm 1/2}\left(\varphi_p, \varphi_k^{\dagger}\right)\mathbf{I}_{\gamma q}^{\mp 1/2}$, $p, q = 1, 2;\ k = 0, p, q$, are bounded in ℓ_2.

In order to simplify the writings of subsequent mathematical manipulations we introduce the following special notations:

$$\begin{aligned} {}^{(\pm)}\mathbf{F}_{pq} &= \mathbf{I}_{\gamma p}^{\pm 1/2}\left(\varphi_p, \varphi_q^{T}\right)\mathbf{I}_{\gamma q}^{\mp 1/2}; \quad {}^{(\pm)}\mathbf{F}_{p} \equiv {}^{(\pm)}\mathbf{F}_{pp}; \\ {}^{(\pm)}\mathbf{F}_{0p} &= \mathbf{I}_{\gamma p}^{\pm 1/2}\left(\varphi_0, \varphi_p^{T}\right)\mathbf{I}_{\gamma p}^{\mp 1/2}; \quad {}^{(\pm)}\mathbf{F}_{p0} = \mathbf{I}_{\gamma p}^{\pm 1/2}\left(\varphi_p, \varphi_0^{\dagger}\right)\mathbf{I}_{\gamma p}^{\mp 1/2}, \end{aligned} \quad p, q = 1, 2,$$

for frequently used matrix operators $\ell_2 \to \ell_2$.

3. APPROXIMATIONS TO FRESNEL FORMULAE

The first operator-based Fresnel formula in Weyl's notation [3]

$$\mathbf{R} = \frac{\mathbf{D} - \mathbf{I}}{\mathbf{D} + \mathbf{I}}; \quad \mathbf{R}, \mathbf{D} : \ell_2 \to \ell_2, \tag{3}$$

which is also known as the Cayley transformation (e.g. see [8,9]), is characterized by the following properties:

(a) $\mathbf{R}^T = \mathbf{R} \Leftrightarrow \mathbf{D}^T = \mathbf{D}$;

(b) $\|\mathbf{R}\| < 1 \Leftrightarrow \operatorname{Re}\mathbf{D} > 0$;

(c) $\mathbf{R} = \mathbf{I} - 2\mathbf{A}$, where $\mathbf{A} \equiv \left(\mathbf{D} + \mathbf{I}\right)^{-1} : \mathbf{A}^T = \mathbf{A}; \quad \operatorname{Re}\mathbf{A} > \mathbf{A}\mathbf{A}^{\dagger}; \quad \|\mathbf{A}\| < 1$.

In order to find a mode reflection operator for the first partial region ${}^{(1)}\mathbf{R}$ the following relations

$$\mathbf{R} = (\pm)^{(1)}\mathbf{R}; \quad \mathbf{D} = \mathbf{D}_0\mathbf{D}_0^T; \quad \mathbf{D}_0 = {}^{(\pm)}\mathbf{F}_{12} \quad \begin{pmatrix} LM \\ LE \end{pmatrix} \tag{4}$$

need to be substituted into (3) (see [3]).

We obtain the reflection operator for the second region by substituting $(\pm) \to (\mp)$, $1 \rightleftarrows 2$ in relations (4). The Fresnel formula for the transmission operator may then be written as

$$\mathbf{T} = \left(\mathbf{I} - \mathbf{R}\right)\mathbf{D}_0 = 2\mathbf{A}\mathbf{D}_0 . \tag{3a}$$

Now introduce the orthoprojectors $\mathbf{P}_K \equiv \left\{ P_{mn}^{(K)} = \sum_{p=(0)1}^{K} \delta_{mp}\delta_{pn} \right\}$ and $\mathbf{Q}_K \equiv \mathbf{I} - \mathbf{P}_K$, where $K = M, N$ stands for the number of waveguide modes taken into account in two partial regions. In what follows, when constructing finite-dimensional approximations, we assume the field in the region related to the sought-for reflection operator to be reduced to sum of M modes, whereas N modes are allowed for in the adjacent region.

We derive the matrix (i.e., the truncated $M \times M$) Fresnel formula by two steps. In the initial stage we introduce an approximation for the given operator in the form of $\widehat{\mathbf{D}} \equiv \mathbf{D}_0 \mathbf{P}_N \mathbf{D}_0^T : \ell_2 \to \ell_2$. By implementing the conventional projection procedure [6,7] we make use of the second step to construct the final finite-dimensional $M \times M$ approximation

$$\tilde{\mathbf{R}} = \frac{\tilde{\mathbf{D}} - \mathbf{P}_M}{\tilde{\mathbf{D}} + \mathbf{P}_M}; \quad \tilde{\mathbf{D}} \equiv \mathbf{P}_M \widehat{\mathbf{D}} \mathbf{P}_M , \tag{5}$$

which is also characterized by three basic properties identical to those described above:

(a′) $\tilde{\mathbf{R}}^T = \tilde{\mathbf{R}} \Leftrightarrow \tilde{\mathbf{D}}^T = \tilde{\mathbf{D}}$;

(b′) $\left\|\tilde{\mathbf{R}}\right\| < 1 \Leftrightarrow \operatorname{Re}\tilde{\mathbf{D}} > 0$;

(c′) $\tilde{\mathbf{R}} = \mathbf{P}_M - 2\mathbf{A}$, where $\tilde{\mathbf{A}} \equiv \left(\tilde{\mathbf{D}} + \mathbf{P}_M\right)^{-1} : \tilde{\mathbf{A}}^T = \tilde{\mathbf{A}}; \quad \operatorname{Re}\tilde{\mathbf{A}} > \tilde{\mathbf{A}}\tilde{\mathbf{A}}^{\dagger}; \quad \left\|\tilde{\mathbf{A}}\right\| < 1.$

Again, an approximation for the transmission operator takes the form of $\tilde{\mathbf{T}} = 2\tilde{\mathbf{A}}\mathbf{D}_0\mathbf{P}_N$.

Note that the properties of the approximations listed above are not dependent upon the relationship between the sizes of truncation M, N and the number of propagating modes in both of the waveguides. This fact is an immediate corollary of the continuity condition for the power flow through the aperture of discontinuity. This condition is met both for the exact and approximate solution of the problem.

Making use of the properties (c) and (c′) we then derive the following estimate for the projection approximations already constructed:

$$\left\{ \begin{matrix} \left\|\left(\mathbf{P}_M \mathbf{R} - \tilde{\mathbf{R}}\right)\mathbf{b}^T\right\| \\ \left\|\left(\mathbf{P}_M \mathbf{T} - \tilde{\mathbf{T}}\right)\mathbf{b}^T\right\| \end{matrix} \right\} < \left\|\left(\mathbf{P}_M \mathbf{D} - \tilde{\mathbf{D}}\right)\mathbf{d}^T\right\|, \quad \mathbf{d} = 2\left\{ \begin{matrix} \mathbf{b}\mathbf{A} \\ \mathbf{b}\mathbf{D}_0^T\mathbf{A} \end{matrix} \right\}, \quad \forall \mathbf{b} \in \ell_2 . \tag{6}$$

So, a strong P-convergence of the approximations in question to the scattering operators is completely determined by the strong P-convergence of matrix $\tilde{\mathbf{D}}$ to the given operator $\mathbf{D}$.

Hence, the problem is to examine the conditions of convergence of the difference of two known operators $\mathbf{P}_M\mathbf{D}-\tilde{\mathbf{D}}\equiv\Lambda_{MN}$ to the null operator.

Lemma 2. Operator Λ_{MN} shows a strong convergence to zero.

Proof. Using definitions (4) and (5) we write this operator in the form of

$$\Lambda_{MN}=\mathbf{P}_M\mathbf{D}_0\mathbf{Q}_N\mathbf{D}_0^T+\mathbf{P}_M\mathbf{D}_0\mathbf{P}_N\mathbf{D}_0^T\mathbf{Q}_M\,. \quad (7)$$

Our assertion follows immediately from representation (7) as an corollary of the strong (but non-uniform) convergence of orthoprojector $\mathbf{P}_K$ to the unit operator in ℓ_2:

$$\lim_{K\to\infty}\left\|\mathbf{Q}_K\mathbf{b}^T\right\|=0,\quad \forall\mathbf{b}\in\ell_{\mathbf{2}};\quad K=M,N. \quad (8)$$

Lemma 2 and the derived estimate (6) lead to the following principal result:

Theorem 1. Projection approximations $\tilde{\mathbf{R}}$ and $\tilde{\mathbf{T}}$ always exhibit a strong P-convergence to the corresponding scattering operators.

In other words, for the mathematical model of the mode-matching technique given as the operator-based Fresnel formulae (3) and (3a) the relative (strong) P-convergence is absent.

The representations of operator Λ_{MN} other than formula (7) can be used to make an in-depth study into the convergence properties of projection approximations at hand for a variety of essential applications. In the text below, we shall present formulae for the mode reflection operator only, because these relations for the transmission operator are similar-in-kind.

4. EQUIVALENT REPRESENTATION FOR THE OPERATOR Λ_{MN}

In carrying out a subsequent analysis the key role is played by distributions:

$$^{(\mp)}G^{(p)}\left(x,x'\right)=\varphi_p^T\left(x\right)\mathbf{I}_{\gamma p}^{\mp1}\varphi_p\left(x'\right),\quad x,x'\in\Omega_0;\quad p=1,2\,, \quad (9)$$

which have the meaning of traces of the Green's function for the p-th partial region (the upper sign "minus") and of its second derivative (the lower sign "plus") on the aperture of discontinuity. Distributions (9) induce the integral operator of the Hilbert-Schmidt type $^{(-)}G^{(p)}$, the hypersingular integral operator $^{(+)}G^{(p)}$ and the operators of difference $^{(\mp)}B^{(qp)}={}^{(\mp)}G^{(q)}-{}^{(\mp)}G^{(p)}$, $p,q=1,2;\ p\neq q$.

These integral operators, in their turn, originate the matrix operators:

$$\begin{aligned}&\mathbf{D}\equiv\mathbf{D}_0\mathbf{D}_0^T=\mathbf{I}_{\gamma p}^{\pm1/2}\left(\varphi_p,{}^{(\mp)}G^{(q)}\varphi_p^T\right)\mathbf{I}_{\gamma p}^{\pm1/2};\\&{}^{(\pm)}\mathbf{F}_p{}^{(\mp)}\mathbf{F}_p=\mathbf{I}_{\gamma p}^{\pm1/2}\left(\varphi_p,{}^{(\mp)}G^{(p)}\varphi_p^T\right)\mathbf{I}_{\gamma p}^{\pm1/2};\\&\mathbf{B}=\mathbf{I}_{\gamma p}^{\pm1/2}\left(\varphi_p,{}^{(\mp)}B^{(qp)}\varphi_p^T\right)\mathbf{I}_{\gamma p}^{\pm1/2}\end{aligned}\qquad\begin{Bmatrix}p\\q\end{Bmatrix}=\begin{Bmatrix}1\\2\end{Bmatrix},\begin{Bmatrix}2\\1\end{Bmatrix}. \tag{10}$$

Hereinafter, when the first partial region is concerned, the choice of the upper sign corresponds to LM -modes, whereas the lower sign is for LE -modes, and vice versa for the second region.

If the finite number of modes is taken into account in regular waveguides, then the functions (9) take on the such form of expression:

$${}^{(\mp)}G_K^{(p)}\left(x,x'\right)=\left[\varphi_p^T\left(x\right)\mathbf{P}_K\right]\mathbf{I}_{\gamma p}^{\mp1}\left[\mathbf{P}_K\varphi_p\left(x'\right)\right],\ K=M,N,$$

and the corresponding matrix operators are given as

$$\begin{aligned}&\widehat{\mathbf{D}}\equiv\mathbf{D}_0\mathbf{P}_N\mathbf{D}_0^T=\mathbf{I}_{\gamma p}^{\pm1/2}\left(\varphi_p,{}^{(\mp)}G_N^{(q)}\varphi_p^T\right)\mathbf{I}_{\gamma p}^{\pm1/2};\\&{}^{(\pm)}\mathbf{F}_p\mathbf{P}_K{}^{(\mp)}\mathbf{F}_p=\mathbf{I}_{\gamma p}^{\pm1/2}\left(\varphi_p,{}^{(\mp)}G_K^{(p)}\varphi_p^T\right)\mathbf{I}_{\gamma p}^{\pm1/2};\\&\mathbf{B}_{NK}=\mathbf{I}_{\gamma p}^{\pm1/2}\left(\varphi_p,{}^{(\mp)}B_{NK}^{(qp)}\varphi_p^T\right)\mathbf{I}_{\gamma p}^{\pm1/2},\end{aligned}\qquad\begin{Bmatrix}p\\q\end{Bmatrix}=\begin{Bmatrix}1\\2\end{Bmatrix},\begin{Bmatrix}2\\1\end{Bmatrix}, \tag{11}$$

where, by definition, ${}^{(\mp)}B_{NK}^{(qp)}={}^{(\mp)}G_N^{(q)}-{}^{(\mp)}G_K^{(p)}$.

Using operators (10) and (11) we find the desired equivalent representation of the matrix operator (7) written as

$$\Lambda_{MN}=\mathbf{P}_M\left(\mathbf{B}-\mathbf{B}_{NK}\right)\mathbf{P}_M+\mathbf{P}_M\mathbf{B}\mathbf{Q}_M+\mathbf{\Xi}_{MK}, \tag{12}$$

where $K=M$ or N and the following notation is introduced:

$$\mathbf{\Xi}_{MK}=\mathbf{P}_M{}^{(\pm)}\mathbf{F}_p\left(\mathbf{Q}_K{}^{(\mp)}\mathbf{F}_p+\mathbf{P}_K{}^{(\mp)}\mathbf{F}_p\mathbf{Q}_M\right).$$

As regards the found representation (12), we shall take interest in the convergence of difference of two known operators $\mathbf{B}-\mathbf{B}_{NK}\equiv\Delta\mathbf{B}_{NK}$ to the null operator.

Lemma 3. Operator $\Delta\mathbf{B}_{NK}$ shows a strong convergence to zero.

Proof. With the properties of the strong convergence of orthoprojector (8) being allowed for, this assertion is immediately evident from the definition in (10) and (11) (or from the representation (12)).

Further, we take a closer look of the convergence properties $\Delta\mathbf{B}_{NK}\to0$, $N,K\to\infty$, at the specific examples.

5. FUNCTION OF THE "MITTRA RULE"

We first consider the canonical problem on a step in a rectangular waveguide $\left(\Omega_2 \subset \Omega_1\right)$. For straight guides we assume $\Omega_1 \equiv \left(0,a\right)$ and $\Omega_2 \equiv \left(0,b\right)$, where $b \le a$, and for continuously curved waveguides we put $\Omega_1 \equiv \left(r, r+a\right)$ and $\Omega_2 \equiv \left(r, r+b\right)$, where r is the curvature radius of their common wall.

By extracting the main (or static) part ${}^{(\mp)}g^{(p)}\left(x,x'\right)$ in distribution (9), for the induced integral operators we obtain the representation of ${}^{(\mp)}G^{(p)} = {}^{(\mp)}g^{(p)} + {}^{(\mp)}\theta^{(p)}$, $p = 1,2$, where remainders of ${}^{(-)}\theta^{(p)}$ and ${}^{(+)}\theta^{(p)}$ are a nuclear operator and a Hilbert-Schmidt operator, respectively. In case of the finite number of waveguide modes in partial regions we have degenerate-kernel integral operator ${}^{(\mp)}G_K^{(p)} = {}^{(\mp)}g_K^{(p)} + {}^{(\mp)}\theta_K^{(p)}$, $K = M, N$.

By picking out the value of $K = M$ we obtain the following representation:

$$\mathbf{\Lambda}_{MN} = \mathbf{P}_M \left(\mathbf{B} - \mathbf{B}_{NM}\right)\mathbf{P}_M + \mathbf{P}_M \mathbf{B}\mathbf{Q}_M + \mathbf{\Xi}_M . \tag{13}$$

Now consider the difference of matrix operators in the first summand of this equality:

$$\Delta\mathbf{B}_{NM} = \mathbf{I}_{\gamma p}^{\pm 1/2}\left(\varphi_p , {}^{(\mp)}\Delta B_{NM}^{(qp)}\, \varphi_p^T\right)\mathbf{I}_{\gamma p}^{\pm 1/2}, \tag{14}$$

where

$${}^{(\mp)}\Delta B_{NM}^{(qp)} = {}^{(\mp)}\Delta g_{NM}^{(qp)} + {}^{(\mp)}\Delta\theta_{NM}^{(qp)}; \tag{15}$$

$$\begin{matrix} {}^{(\mp)}\Delta g_{NM}^{(qp)} = {}^{(\mp)}\Delta g_N^{(q)} - {}^{(\mp)}\Delta g_M^{(p)}; \\ {}^{(\mp)}\Delta\theta_{NM}^{(qp)} = {}^{(\mp)}\Delta\theta_N^{(q)} - {}^{(\mp)}\Delta\theta_M^{(p)} \end{matrix} \quad \begin{Bmatrix} p \\ q \end{Bmatrix} = \begin{Bmatrix} 1 \\ 2 \end{Bmatrix}, \begin{Bmatrix} 2 \\ 1 \end{Bmatrix};$$

$${}^{(\mp)}\Delta g_K^{(p)} = {}^{(\mp)}g^{(p)} - {}^{(\mp)}g_K^{(p)}; \ {}^{(\mp)}\Delta\theta_K^{(p)} = {}^{(\mp)}\theta^{(p)} - {}^{(\mp)}\theta_K^{(p)}; \ K = M, N$$

Let us introduce six bounded matrix operators using the formula

$$\begin{Bmatrix} {}^{(\mp)}\mathbf{C}^{(p)} \\ {}^{(\mp)}\mathbf{\Delta C}_{NM}^{(qp)} \\ {}^{(\mp)}\mathbf{\Delta \Theta}_{NM}^{(qp)} \end{Bmatrix} = \mathbf{I}_{\gamma p}^{\pm 1/2} \left\{ \varphi_p \begin{Bmatrix} {}^{(\mp)}g^{(p)} \\ {}^{(\mp)}\Delta g_{NM}^{(qp)} \\ {}^{(\mp)}\Delta\theta_{NM}^{(qp)} \end{Bmatrix} \varphi_p^T \right\} \mathbf{I}_{\gamma p}^{\pm 1/2} . \tag{16}$$

Then the relation (14) can be rewritten in the form of

$$\Delta\mathrm{B}_{NM} = {}^{(\mp)}\mathbf{\Delta C}_{NM}^{(qp)} + {}^{(\mp)}\mathbf{\Delta\Theta}_{NM}^{(qp)} .$$

Here the first term is the main part of the operator under study, whereas the remainder is the nuclear operator such that

$$\lim_{M,N\to\infty}\left\| {}^{(\mp)}\mathbf{\Delta\Theta}_{NM}^{(qp)} \right\| = 0, \quad \forall N / M \ .$$

Consequently, the strong convergence $\Delta \mathrm{B}_{NK} \to 0$ will be determined by the estimate:

$$\left\| \Delta\mathbf{B}_{NM}\mathbf{b}^T \right\| \le \left\| {}^{(\mp)}\mathbf{\Delta C}_{NM}^{(qp)}\mathbf{b}^T \right\| + \left\| {}^{(\mp)}\mathbf{\Delta\Theta}_{NM}^{(qp)} \right\| \left\| \mathbf{b} \right\|, \quad \forall \mathbf{b} \in \ell_{\mathbf{2}} \ . \tag{17}$$

Next, if the first region is to be considered, then we assume $N = tM$ (for the second region we put $M = tN$), where $t > 0$, and sum the kernel of the integral operator ${}^{(\mp)}\Delta g_{NM}^{(qp)}$ (i.e., the kernel of the main part of operator ${}^{(\mp)}\Delta B_{NM}^{(qp)}$) over the remaining index $M\left(\text{or } N\right) = \overline{1,\infty}$. For the straight rectangular waveguides, in which it is convenient to introduce dimensionless variables $\alpha = \pi x / b,\ \alpha' = \pi x' / b$ and the geometrical parameter $\tau = b / a \le 1$, this summation procedure results in the formula

$$\begin{aligned}&\sum_{M=1}^{\infty} {}^{(\mp)}\Delta g_{tM,M}^{(21)}\left(\alpha,\alpha'\right) = -\sum_{N=1}^{\infty} {}^{(\mp)}\Delta g_{N,tN}^{(12)}\left(\alpha,\alpha'\right) = \left(\mp\right)\left(\tau^{-1} - t^{-1}\right)\begin{Bmatrix} \delta\left(\alpha - \alpha'\right) \\ \delta''_{xx}\left(\alpha - \alpha'\right) \end{Bmatrix} + \\ &+ {}^{(\mp)}g^{(1)}\left(\alpha,\alpha'\right) - \frac{1}{2}\left[{}^{(\mp)}g^{(2)}\left(\alpha,\alpha'\right) + {}^{(\mp)}g\left(t\alpha,t\alpha'\right) \right],\end{aligned} \tag{18}$$

where the notation is as follows:

$${}^{(\mp)}g\left(t\alpha,t\alpha'\right) = \begin{cases} -\dfrac{2}{\pi}\displaystyle\int \Phi\left(t\alpha,t\alpha'\right)d\alpha; \\ \dfrac{2\pi}{b^2}\dfrac{\partial}{\partial\alpha}\Phi\left(t\alpha,t\alpha'\right); \end{cases} \Rightarrow {}^{(\mp)}g\left(t\alpha,t\alpha'\right)\Big|_{t=\tau} = \frac{1}{\tau}\,{}^{(\mp)}g^{(1)}\left(\alpha,\alpha'\right);$$

$$\Phi\left(t\alpha,t\alpha'\right) = \frac{1}{2}\left[\operatorname{co}\tan\left(t\frac{\alpha - \alpha'}{2} \right) \mp \operatorname{co}\tan\left(t\frac{\alpha + \alpha'}{2} \right) \right] \quad \begin{pmatrix} LM \\ LE \end{pmatrix}.$$

The result thus obtained and definition (16) may well be employed to construct an operator series that will be convergent or divergent depending upon the ratio M / N. It is exactly on condition $t = \tau$ that we are able to obtain the convergent operator series:

$$2\sum_{M=1}^{\infty} {}^{(\mp)}\mathbf{\Delta C}_{\tau M,M}^{(21)} = -2\sum_{N=1}^{\infty} {}^{(\mp)}\mathbf{\Delta C}_{N,\tau N}^{(12)} = \frac{2\tau - 1}{\tau}\,{}^{(\mp)}\mathbf{C}^{(1)} - {}^{(\mp)}\mathbf{C}^{(2)} . \tag{19}$$

If, otherwise, $t \neq \tau$, it is easy to notice that δ-singularities in the sum of series (18) generate ℓ_2 -unbounded matrix operator of the form:

$$const \cdot \left(\tau^{-1} - t^{-1}\right) \mathbf{I}_{\gamma p}^{\pm 1/2} \left(\varphi_p , \begin{Bmatrix} \varphi_p^T \\ \varphi_p''^T \end{Bmatrix} \right) \mathbf{I}_{\gamma p}^{\pm 1/2}, \quad p = 1, 2 .$$

Hence, in this case, the formula (18) cannot be used to construct the convergent operator series in $\mathcal{B}\left(\ell_2\right)$.

Thus, the rate of the first summand decrease in estimate (17) given as $\left\| {}^{(\mp)}\Delta\mathbf{C}_{NM}^{(qp)}\mathbf{b}^T \right\| \to 0,\ M, N \to \infty$, is dependent upon M / N and can be qualitatively evaluated as being sufficient for the series (19) to converge $\left(t = \tau\right)$ or not $\left(\forall t \neq \tau\right)$.

Note that for other regular waveguides this argumentation will remain valid if the formula (18) holds true in terms of the result from asymptotic summation involving the correction of geometrical meaning of the parameter included in the formulae above. For instance, this will be true in the case of continuously curved rectangular waveguides for which (as it was ascertained in [10]) $\tau = \ln\left(1 + b / r\right) / \ln\left(1 + a / r\right)$.

So, the “Mittra rule” for scattering operators may be laid down as follows:

Theorem 2. With the relation $N / M = \tau$, where τ is the certain geometrical parameter of the problem ($\tau = b / a$ for straight rectangular waveguides, etc.) the rate of approximation convergence will be higher than with any other proportions of mode numbers being taken into account.

Based upon formulae (6) and (13) we can make the inference that in this case, at $N / M = \tau,$ the previously established strong P -convergence of projection approximations is characterized by the following estimate:

$$\left\| \left(\mathbf{P}_M \mathbf{R} - \tilde{\mathbf{R}} \right) \mathbf{b}^T \right\| < 2 \left\| \Delta\mathbf{B}_{\tau M, M} \mathbf{P}_M \mathbf{d}^T \right\| + \left\| \left(\mathbf{P}_M \mathbf{B} \mathbf{Q}_M + \mathbf{\Xi}_M \right) \mathbf{d}^T \right\| , \qquad (20)$$

where $\mathbf{d} = 2\mathbf{b}\mathbf{A},\ \forall \mathbf{b} \in \ell_2$.

6. RELATIVE UNIFORM CONVERGENCE

Assume that the geometry of abrupt discontinuity be such that $\Omega_1 = \Omega_2$. The canonical discontinuity of this type is the junction of the straight and uniformly curved waveguides of identical rectangular cross-section (i.e., the break of guide curvature).

In terms of the properties of the integral operators under study this signifies that the kernel

$${}^{(\pm)}B^{(qp)}\left(x, x'\right) = {}^{(\pm)}G^{(q)}\left(x, x'\right) - {}^{(\pm)}G^{(p)}\left(x, x'\right) . \qquad (21)$$

is continuous with respect to both variables $x, x' \in \Omega_0$ and, as a consequence, the induced integral operator ${}^{(\pm)}B^{(qp)}$ is a nuclear one.

Let the column-vector ${}^{(\pm)}\boldsymbol{\varphi}_0 \equiv \left\{ {}^{(\pm)}\varphi_m^{(0)}(x) \right\}_{m=1}^{\infty}$ represents a maximal orthonormal system of eigenfunctions of the positive operator $\left\{ \left[{}^{(\pm)}B^{(qp)} \right]^{\dagger} \left[{}^{(\pm)}B^{(qp)} \right] \right\}^{1/2}$, whereas $\left\{ {}^{(\pm)}\alpha_m \right\}_{m=1}^{\infty} \in \ell_1$ are the corresponding nonnegative eigenvalues. An expansion similar to formula (9) (see, e.g., [9]) is then valid for the kernel (21):

$$ {}^{(\pm)}B^{(qp)}(x, x') = {}^{(\pm)}\boldsymbol{\varphi}_0^{\dagger}(x)\, \mathbf{I}_{\alpha}\, {}^{(\pm)}\boldsymbol{\varphi}_0(x'), \quad x, x' \in \Omega_0 . \tag{22} $$

Substituting expression (22) into the third formula in Eqs. (10), we obtain a new representation of the matrix operator

$$ \mathbf{B} = \mathbf{I}_{\gamma p}^{\pm 1/2} \left(\boldsymbol{\varphi}_p, {}^{(\mp)}\boldsymbol{\varphi}_0^{\dagger} \right) \mathbf{I}_{\alpha} \left({}^{(\mp)}\boldsymbol{\varphi}_0, \boldsymbol{\varphi}_p^{T} \right) \mathbf{I}_{\gamma p}^{\pm 1/2} = {}^{(\pm)}\mathbf{F}_{p0} \mathbf{I}_{\alpha(\gamma p)} {}^{(\mp)}\mathbf{F}_{0p} . \tag{23} $$

Here the diagonal operator $\mathbf{I}_{\alpha(\gamma p)}$ is defined by the elements $\left\{ {}^{(\pm)}\alpha_m \left[\gamma_m^{(p)} \right]^{\pm 1} \right\}_{m=1}^{\infty} \in c_0$ and, consequently, is compact in ℓ_2. From (23) it then follows that operator $\mathbf{B}$ is compact as well.

When selecting the value of $K = N$ in expression (12) we get the following representation:

$$ \Lambda_{MN} = \mathbf{P}_M (\mathbf{B} - \mathbf{B}_{NN}) \mathbf{P}_M + \mathbf{P}_M \mathbf{B} \mathbf{Q}_M + \boldsymbol{\Xi}_{MN} , \tag{24} $$

where, according to (11), the matrix operator $\mathbf{B}_{NN}$ is originated by the difference of two degenerate-kernel integral operators. Based upon the general formula (12) for the operator difference under consideration we arrive at the following expression:

$$ \Lambda_{MN} = \mathbf{P}_M (\mathbf{B} - \mathbf{B}_{NN}) \mathbf{P}_M + \mathbf{P}_M \mathbf{B} \mathbf{Q}_M + \mathbf{P}_M \mathbf{Q}_N . $$

In view of Eq. (6) the following estimates are obtained:

$$ \begin{gathered} \|\Lambda_{MN}\| \le \|\Delta \mathbf{B}_{NN}\| + \|\mathbf{B}\mathbf{Q}_M\|, \quad \|\mathbf{P}_M \mathbf{Q}_N\| = 0, \quad M \le N; \\ \|\Lambda_{MN}\| \ge \left| 1 - \|\mathbf{P}_M \Delta \mathbf{B}_{NN} \mathbf{P}_M + \mathbf{P}_M \mathbf{B} \mathbf{Q}_M\| \right|, \quad \|\mathbf{P}_M \mathbf{Q}_N\| = 1, \quad M > N, \end{gathered} \tag{25} $$

where

$$ \|\mathbf{P}_M \Delta \mathbf{B}_{NN} \mathbf{P}_M + \mathbf{P}_M \mathbf{B} \mathbf{Q}_M\| \le \|\Delta \mathbf{B}_{NN}\| + \|\mathbf{B}\mathbf{Q}_M\| . $$

For individual terms in these inequalities we have (e.g., see [9,11])

$$\lim_{N\to\infty}\left\|\Delta\mathbf{B}_{NN}\right\|=0,\quad \lim_{M\to\infty}\left\|\mathbf{BQ}_M\right\|=0 .$$

Thus, we have the proof of

Theorem 3. For the case where $\Omega_1=\Omega_2$ the (relative) uniform P-convergence $\left\|\mathbf{P}_M\mathbf{R}-\tilde{\mathbf{R}}\right\|\to 0,\ M,N\to\infty$, is observed on condition that $M\le N$.

The numerical results illustrating the theorem are given in [1].

7. PROPERTIES OF THE CONDITION NUMBER

The computational stability required for practical purposes as well as the accuracy of numerical realization of an approximate solution is determined by the quantity $cond\left(\tilde{\mathbf{A}}\right)\equiv\left\|\tilde{\mathbf{A}}\right\|\cdot\left\|\tilde{\mathbf{A}}^{-1}\right\|$. It is easy to see that this condition number is bounded at any values of M and N. Indeed, $\left\{cond\left(\tilde{\mathbf{A}}\right)\right\}_{M,N=1}^{\infty}\in\ell_\infty$, because in view of the properties $\left(c'\right)$ and boundedness of the matrix operators $\mathbf{D}_0$ we have

$$cond\left(\tilde{\mathbf{A}}\right)<\left\|\tilde{\mathbf{A}}^{-1}\right\|\le 1+\left\|\mathbf{D}_0\right\|\cdot\left\|\mathbf{D}_0^T\right\|<\infty,\quad \forall M,N .$$

This general estimator may be updated when imposing auxiliary conditions. To this end we will introduce the quantity $cond_M\left(\mathbf{A}\right)\equiv\left\|\mathbf{P}_M\mathbf{A}\right\|\cdot\left\|\mathbf{A}^{-1}\mathbf{P}_M\right\|$, whose properties

$$1\le cond_M\left(\mathbf{A}\right)<\left\|\mathbf{A}^{-1}\right\|\le 1+\left\|\mathbf{D}\right\|<\infty$$

make it possible to state

Theorem 4. With the proviso that $\left\|\Lambda_{MN}\right\|<\left\|\mathbf{A}\right\|^{-1}$, the double-ended estimate

$$\frac{cond_M\left(\mathbf{A}\right)-\varepsilon_1}{1+\varepsilon_2}\le cond\left(\tilde{\mathbf{A}}\right)\le\frac{cond_M\left(\mathbf{A}\right)+\varepsilon_1}{1-\varepsilon_2}, \tag{26}$$

is valid for $\varepsilon_1=\left\|\mathbf{P}_M\mathbf{A}\right\|\left\|\mathbf{DP}_M-\tilde{\mathbf{D}}\right\|,\ \varepsilon_2=\left\|\mathbf{A}\right\|\left\|\mathbf{P}_M\mathbf{D}-\tilde{\mathbf{D}}\right\|$.

Proof. The triangle inequality enables to find the estimates:

$$\left|\,\left\|\mathbf{A}^{-1}\mathbf{P}_M\right\|-\left\|\tilde{\mathbf{A}}^{-1}\right\|\,\right|\le\varepsilon_1\left\|\mathbf{P}_M\mathbf{A}\right\|^{-1},\quad \left|\,\left\|\mathbf{P}_M\mathbf{A}\right\|-\left\|\tilde{\mathbf{A}}\right\|\,\right|\le\varepsilon_2\left\|\tilde{\mathbf{A}}\right\|,$$

from which the formula (26) immediately follows under the condition that $\varepsilon_2 < 1$.

Corollary. In case of the (relative) uniform convergence, i.e., $\|\Lambda_{MN}\| \to 0$, $M, N \to \infty$, the numbers M_0 and N_0 are present, such that

$$\left| \frac{cond\left(\tilde{\mathbf{A}}\right)}{cond_M\left(\mathbf{A}\right)} - 1 \right| < \|\Lambda_{MN}\|$$

for $M > M_0$ and $N > N_0$.

The results from the numerical investigations into the behavior of the condition number in case of the uniform convergence are cited in [1].

8. CONCLUSIONS

The applicability of the truncation procedure to matrix models of the mode-matching technique has been substantiated rigorously for the problem on mode diffraction by an abrupt discontinuity in a rectangular waveguide. Making use of the projection procedure, the finite-dimensional approximations to the operator-based Fresnel formulae have been constructed and a study has been performed of the qualitative characteristics of their convergence. Based upon the fundamental properties of Cayley transform the problem of estimating of the found approximations has been reduced to investigating the projection convergence (P-convergence) of the finite matrices to a given matrix operator.

It has been proved that for the mathematical model of the mode-matching technique, which is represented as the operator-based Fresnel formulae, the obtained projection approximations are unconditionally and strongly P-convergent to a true solution. In other words, for the matrix model under consideration the phenomenon of relative strong convergence is absent.

It has been found that for the canonical problem on a step in a rectangular waveguide the observance of the "Mittra rule" for the reduced field expansions signifies the fastest strong convergence of the matrix operator $\Delta\mathbf{B}_{NM}$ to zero; this operator is generated by the difference of the Green's function traces on the aperture and it determines a part of approximation error in accordance with the estimate (20).

For the canonical problem on the jump (or break) of the rectangular waveguide curvature, the relative uniform P-convergence of projection approximations takes place in accordance with formula (25). The condition for this convergence states that the number of modes being allowed for should smaller just for the waveguide port for which the reflection operator is being sought.

Finally, we have derived the estimates for the condition number of the truncated matrix equation, which ensure the stability of practical computations.

The developed approach and the results thus achieved allow one to rigorously justify the truncation procedure in the numerical realization of the matrix-operator models of the mode-matching technique in solving the problems of mode diffraction on abrupt discontinuity in the waveguide.

REFERENCES

1. Petrusenko, I.V., (2004), Analytic – numerical analysis of waveguide bends, *Electromagnetics*, **24**:237-254.
2. Petrusenko, I.V., (2004), Matrix operator technique for analysis of wave transformers, *Proc. 10th Int. Conf. on Mathematical Methods in Electromagnetic Theory (MMET'04)*, Dnipropetrovs'k, pp. 118-120.
3. Petrusenko, I.V. and Sirenko, Yu.K., (2011), Fresnel formulae for scattering operators, *Telecommunications and Radio Engineering*, **70**(9):749-758.
4. Petrusenko, I.V. and Sirenko, Yu.K., (2009), Generalization of the power conservation law for scalar mode-diffraction problems, *Telecommunications and Radio Engineering*, **68**(16):1399-1410.
5. Mittra, R. and Lee, S.W., (1971), *Analytical techniques in the theory of guided waves*, New York, The Macmillan Company, - 327 p.
6. Gohberg, I.C. and Fel'dman, I.A., (2006), *Convolution equations and projection methods for their solution (Translations of mathematical monographs)*, Providence, American Mathematical Society, - 261 p. (Translated from Russian).
7. Trenogin, V.A., (2002), *Functional analysis*, Phizmathlit, Moscow: 488 p. (in Russian).
8. Weyl, H., (1997), *The classical groups: Their invariants and representations*, Chichester, Princeton University Press, - 316 p.
9. Richtmyer, R.D., (1978), *Principles of advanced mathematical physics. Vol. 1*, New York - Heidelberg - Berlin, Springer-Verlag, - 422 p.
10. Petrusenko, I.V. and Dmitryuk, S.G., (1986), The method of solution of the generalized waveguide bifurcation problem, *Padiotekhnika i Elektronika*, **31**(7):1285-1293 (in Russian).
11. Gribanov, Yu.I., (1959), On the complete continuity of matrix operators in vector spaces, *Doklady Akademii Nauk SSSR*, **129**(5):975-978 (in Russian).