Bergische Universität Wuppertal

Fachbereich Mathematik und Naturwissenschaften

Institute of Mathematical Modelling, Analysis and Computational Mathematics (IMACM)

Preprint BUW-IMACM 14/39

Pavel S. Petrov and Matthias Ehrhardt

On Mayfield's stability proof for the discretized transparent boundary condition for the parabolic equation

November 2014

http://www.math.uni-wuppertal.de

On Mayfield's stability proof for the discretized transparent boundary condition for the parabolic equation

P.S. Petrov

Il'ichev Pacific oceanological institute, 43 Baltiyskay St., Vladivostok, 41, 690041, Russia

Far Eastern Federal University,8 Sukhanova St., Vladivostok, 41, 690950, Russia

M. Ehrhardt

Bergische Universität Wuppertal, Gaußstrasse 20, D-42119 Wuppertal, Germany

Abstract

Mayfield's numerical implementation of transparent boundary condition for the Schrödinger-type parabolic equations is revisited. An inaccuracy in the original proof of the conditional stability for the resulting scheme is pointed out. The highly unusual and impressive original result is reestablished and a new proof is presented. Some further remarks and estimates on the instability which occurs when the Mayfield condition is violated are given.

Keywords: Transparent boundary conditions, Parabolic equation, Schrödinger equation, Numerical stability, Mayfield condition
PACS: 43.30.+m,
PACS: 43.20.Bi

1. Introduction

The methods for the artificial truncation of the computational domain for the numerical solution of the Schrödinger-type parabolic equations are being intensively developed for more than two decades. This domain truncation may be accomplished either by imposing the transparent boundary conditions

Email addresses: `petrov@poi.dvo.ru` (P.S. Petrov), `ehrhardt@uni-wuppertal.de` (M. Ehrhardt)

(TBC) or by extending the computational domain with so-called perfectly matching layers (or PMLs). For a comprehensive review with the comparison of different implementations of TBCs and PML we refer to [1]. Among the earliest works on nonlocal TBCs for the Schrödinger-type equations were probably the thesis of Mayfield [2] and the article of Baskakov and Popov [3]. Both works deal with the narrow-angle parabolic equations for the wave propagation in acoustical and optical applications respectively (these equations are mathematically equivalent to the Schrödinger equation). In the former thesis an implementation of the TBC in the Neumann-to-Dirichlet (NtD) form was proposed while in the latter paper a different implementation of the TBC in the Dirichlet-to-Neumann (DtN) form was obtained. Very recently it was shown that the numerical scheme incorporating the TBC of Baskakov and Popov is unconditionally stable [4]. Mayfield [2] has shown that her numerical scheme features only a conditional stability with a very unusual instability condition (cf. Lemma 3). We observed that her proof contains one inaccuracy and decided to fill this gap. Also we observed that although the stability criterion used by Mayfield [2] is more restrictive than is usually necessary, the numerical scheme fails to satisfy even a much less restrictive one under her instability condition.

2. TBC for the parabolic equation and Mayfield's implementation

Let us consider a narrow-angle parabolic equation (PE) describing sound propagation in the ocean [5]

$$2\mathrm{i}k_0 u_r + u_{zz} + k_0^2(\nu^2 - 1)u = 0\,, \tag{1}$$

where $u = u(r, z)$ is the acoustical pressure envelope function (i.e. acoustical pressure $p(r, z) = H_0^1(k_0 r)u(r, z)$), r is range variable, z is depth, k_0 is reference wavenumber and $\nu = \nu(r, z)$ is refractive index). Hereafter we use the acoustical notation following the original work of Mayfield [2], although it is straightforward to rewrite our results for the case of a Schrödinger equation. The equation (1) in computational acoustics is usually complemented by a pressure-release boundary condition $u(r, 0) = 0$ at the ocean surface $z = 0$ and by an initial condition $u(0, z) = u_0(z)$ modeling a point source field [5] to define an initial-boundary value problem (IBVP) in the domain $\Omega = \{(r, z)|r \geq 0,\ z \geq 0\}$. To solve such IBVP for (1) numerically, one must artificially truncate the the domain $z \geq 0$, $r \geq 0$ at a certain depth $z = z_b$, e.g. at the sea bottom.

In the thesis of Mayfield [2] the following Neumann-to-Dirichlet TBC was used at the ocean bottom $z = z_b$ for the numerical solution of (1):

$$u(r, z_b) = \frac{-\exp(\mathrm{i}\pi/4)}{\sqrt{2k_0\pi}} \int_0^r \frac{\exp(\mathrm{i}k_0(\nu_b^2 - 1)(r - \xi)/2)}{\sqrt{r - \xi}} u_z(\xi, z_b) d\xi \,, \qquad (2)$$

where $\nu_b = \nu(z_b)$. Mayfield [2] discretized the equation (1) using the Crank-Nicholson scheme on the uniform grid $r^n = n\Delta r$, $z_m = m\Delta z$, $n = 0, 1, \dots, N$, $m = 0, 1, \dots, M$, where $\Delta z M = z_b$, $\Delta z N = r_{max}$

$$-\rho u_{m-1}^{n+1} + \sigma_m^n u_m^{n+1} - \rho u_{m+1}^{n+1} = \rho u_{m-1}^n - \bar{\sigma}_m^n u_m^n + \rho u_{m+1}^n \,, \qquad (3)$$

where $u_m^n = u(r^n, z_m)$, $\rho = \Delta r/(\Delta z)^2$, $\sigma_m^n = 2\rho + \zeta_m$, $\zeta_m^n = -4\mathrm{i}k_0 - k_0^2 \Delta r((\nu_m^n)^2 - 1)$. The discretized TBC (2) was written as

$$u_M^{n+1} = \epsilon u_{m-1}^{n+1} - \sqrt{\epsilon} S^n \,, \qquad (4)$$

where

$$\epsilon = \exp\left(\frac{-\sqrt{2\pi k_0}\Delta z}{2\sqrt{\mathrm{i}\Delta r}}\right) ,$$

$$S^p = \sum_{j=0}^{n-1} \exp(\mathrm{i}k_0(\nu_b^2 - 1)(j+1)\Delta r/2)(\sqrt{j+2} - \sqrt{j+1})(u_M^{N-j} - u_{M-1}^{N-j}) \,.$$

At the sea surface we employ a homogeneous Dirichlet BC $u_0^n = 0$ for all $n = 0, 1, 2, \dots$.

3. The Mayfield stability criterion

Now we combine the interior scheme (3) and the discretized TBC (4) into a marching matrix form for a column vector $\bar{u}^n = (u_1^n, u_2^n, \dots, u_M^n)^\top$, (where $\top$ stands for the transposition):

$$(Z + \rho T_\epsilon)\bar{u}^{n+1} = (-\bar{Z} - \rho T_\epsilon)\bar{u}^n + \bar{w}^n \,, \qquad (5)$$

where we introduced the (tri-)diagonal matrices

$$T_\epsilon = \begin{pmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & -1 & 2 & -1 \\ 0 & \dots & 0 & -1 & 2-\epsilon \end{pmatrix} , \quad Z = \begin{pmatrix} \zeta_1^n & 0 & \dots & 0 \\ 0 & \zeta_2^n & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \zeta_M^n \end{pmatrix}$$

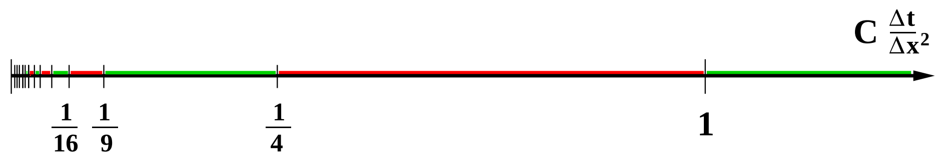

Figure 1: The Mayfield stability set determined by (6).

and a vector $\bar{w}^n = (0, \ldots, 0, -\rho\sqrt{\epsilon}(S^n + S^{n-1}))^\top$. In the subsequent analysis we assumed (following Mayfield) that $\nu \equiv 1$ (and thus $\zeta_m^n = -4\mathrm{i}k_0$ and Z is a scalar matrix).

The stability investigation by Mayfield [2] for the numerical scheme (5) is based on the eigenvalue analysis for the error-propagating matrix $P = P(\epsilon) = (Z + \rho T)^{-1}(-\bar{Z} - \rho T)$. Next we briefly summarize the results from [2].

Lemma 1. *([2], Lemma 6.3) The eigenvalues λ_j of the error-propagating matrix $P(\epsilon)$ satisfy the condition $|\lambda_j| \leq 1$ if and only if eigenvalues μ_j of the matrix T_ϵ are such that* $\mathrm{Im}(\mu_j) \leq 0$.

Lemma 2. *([2], Lemma 6.4) All eigenvalues μ_j of the matrix T_ϵ satisfy* $\mathrm{Im}(\mu_j) \leq 0$ *if and only if* $\mathrm{Im}(\epsilon) \geq 0$.

Lemma 3. *([2], Lemma 6.5) The condition* $\mathrm{Im}(\epsilon) \geq 0$ *is fulfilled if and only if the mesh ratio $\rho = \Delta r/(\Delta z)^2$ satisfies the following inequality*

$$\frac{k_0}{4\pi(2n+1)^2} \leq \frac{\Delta r}{(\Delta z)^2} \leq \frac{k_0}{4\pi(2n)^2}\,, \text{ where } n = 0, 1, 2, \ldots\,. \tag{6}$$

The set of ρ values described by the Mayfield condition (6) is shown in Fig. 1.

If by the definition we accept the inequalities $|\lambda_j| \leq 1$ for all j as the stability criterion for the scheme (5), then it is stable whenever the meshsizes are chosen to satisfy (6). This result is very elegant and unusual, however the proof of Lemma 2 given in [2] is inaccurate (and this lemma is the least trivial part of the criterion proof). More precisely, the eigenvalues $\mu_j(\epsilon)$ were represented in a perturbative form (with ϵ acting as a small parameter):

$$\mu_j(\epsilon) = \mu_j(0) - \epsilon(x_m^j)^2 + \alpha\,, \quad j = 1, 2, \ldots, M\,, \tag{7}$$

where $\mu_j(0)$ denotes an eigenvalue of the matrix T_0, x_m^j is m-th coordinate of the corresponding eigenvector $\bar{x}^j$ and where α stands for higher-order terms in ϵ. Mayfield [2] proved that the first order perturbation theory actually provides the exact values for the eigenvalues $\mu_j(\epsilon)$, i.e. that $\alpha = 0$. To this end the equations (7) were summed up for all j to obtain

$$\sum_{j=1}^{M} \mu_j(\epsilon) = \sum_{j=1}^{M} \mu_j(0) - \sum_{j=1}^{M} \epsilon (x_m^j)^2 + M\alpha \,. \tag{8}$$

Note that $\sum_j (x_m^j)^2 = 1$, since the matrix T_0 is symmetric and thus has the complete orthonormal set of eigenvectors. Since $\sum_j \mu_j(\epsilon) = \mathrm{tr}(T_\epsilon) = 2m - \epsilon$ and $\sum_j \mu_j(0) = \mathrm{tr}(T_0) = 2m$, this equation reduces to the $m\alpha = 0$ which implies $\alpha = 0$. Note however that one must assume α in (7) to be the same for all j to obtain (8). This assumption was not justified in [2] and in fact it is wrong. It is therefore necessary to provide an alternative proof of Lemma 2.

4. The Proof of Lemma 2

Let $D_M = \det(T_0(M) - \mu E(M))$ be the characteristic polynomial of the matrix T_0 of dimension M. $E(M)$ stands for the identity matrix of the same size. Using the row decomposition of the determinant it is easy to check that

$$\det(T_\epsilon(M) - \mu E) = D_M - \epsilon D_{M-1} \,.$$

It is also easy to see that the following recurrence relation holds for D_n:

$$D_M = a D_M - D_{M-1} \,,$$

where for convenience $a = 2 - \mu$. This relation may be resolved in the form

$$D_M = C_1 \kappa_1^M + C_2 \kappa_2^M \,,$$

where κ_1 and κ_2 are the roots of the polynomial $\kappa^2 - a\kappa + 1$ and the constants C_1 and C_2 may be determined from the conditions $D_0 = 1$ and $D_1 = a$. After some straightforward calculations we arrive at

$$D_M = \frac{1}{\sqrt{a^2-4}} \left(\frac{a + \sqrt{a^2-4}}{2} \right) - \frac{1}{\sqrt{a^2-4}} \left(\frac{a - \sqrt{a^2-4}}{2} \right) .$$

The eigenvalues μ of the matrix $T_\epsilon(M)$ (or the corresponding values of a) may be determined from the equation

$$\det(T_\epsilon(M) - \mu E) = D_M - \epsilon D_{M-1} = 0 \,,$$

which may now be rewritten as

$$(a+\sqrt{a^2-4})^{M-1}(a+\sqrt{a^2-4}-2\epsilon) = (a-\sqrt{a^2-4})^{M-1}(a-\sqrt{a^2-4}-2\epsilon)\,.$$

In order to further simplify this equation we introduce the new variable $b = a/2 = 1-\mu/2$ to obtain

$$(b+\sqrt{b^2-1})^{M-1}(b+\sqrt{b^2-1}-\epsilon) = (b-\sqrt{b^2-1})^{M-1}(b-\sqrt{b^2-1}-\epsilon)\,. \quad (9)$$

It is clear that all eigenvalues of $T_\epsilon(M)$ are non-positive if and only if all solutions b_j of this equation satisfy $\mathrm{Im}(b) \geq 0$. Also note that the values $z_{1,2} = b \pm \sqrt{b^2-1}$ are the roots of the polynomial $z^2 - 2zb + 1$. Using these definitions we now reformulate Lemma 2:

Lemma 4. *Let z_1 and z_2 be the roots of polynomial $z^2 - 2zb + 1$. Then the quantity $\epsilon = \frac{z_1^M - z_2^M}{z_1^{M-1} - z_2^{M-1}}$ satisfies the condition $\mathrm{Im}(\epsilon) \geq 0$ if and only if $\mathrm{Im}(b) \geq 0$.*

Note that the equivalence of Lemma 2 and Lemma 4 follows from (9) and the relation $b = 1 - \mu/2$.

Proof. We use mathematical induction on M. We establish a basis for $M = 2$:

$$\epsilon = \frac{z_1^2 - z_2^2}{z_1 - z_2} = z_1 + z_2 = 2b\,,$$

and $\mathrm{Im}(\epsilon) = 2\,\mathrm{Im}(b)$.

To prove the inductive step assume that $\mathrm{sgn}\left(\mathrm{Im}\left(\frac{z_1^k - z_2^k}{z_1^{k-1} - z_2^{k-1}}\right)\right) = \mathrm{sgn}(b)$ for $k < M$. Observe that

$$\frac{z_1^k - z_2^k}{z_1^{k-1} - z_2^{k-1}} = z_1 + z_2 - \frac{z_1^{k-2} - z_2^{k-2}}{z_1^{k-1} - z_2^{k-1}}\,.$$

Indeed, multiplying both sides of this equality by $z_1^{k-1} - z_2^{k-1}$ we obtain on the left side $z_1^k - z_2^k$ while the expression on the right side becomes

$$\begin{aligned} &(z_1+z_2)(z_1^{k-1} - z_2^{k-1}) - z_1^{k-2} - z_2^{k-2} = \\ &z_1^k - z_2^k + z_2 z_1^{k-1} - z_1 z_2^{k-1} - z_1^{k-2} + z_2^{k-2} = z_1^k - z_2^k\,, \end{aligned} \quad (10)$$

(since $z_1 z_2 = 1$). We then have

$$\frac{z_1^M - z_2^M}{z_1^{M-1} - z_2^{M-1}} = 2b - \left(\frac{z_1^{M-1} - z_2^{M-1}}{z_1^{M-2} - z_2^{M-2}}\right)^{-1}\,. \quad (11)$$

By induction $\operatorname{sgn}\left(\operatorname{Im}\left(\frac{z_1^{M-1}-z_2^{M-1}}{z_1^{M-2}-z_2^{M-2}}\right)\right) = \operatorname{sgn}(\operatorname{Im}(b))$ and this implies

$$\operatorname{sgn}\left(\operatorname{Im}\left(\left(\frac{z_1^{M-1}-z_2^{M-1}}{z_1^{M-2}-z_2^{M-2}}\right)^{-1}\right)\right) = -\operatorname{sgn}(\operatorname{Im}(b)) .$$

Consequently, the imaginary part of right-hand side of (11) has the same sign as $\operatorname{Im}(b)$, and this completes the proof of the inductive step. □

5. Uniform estimate on error growth

Now we estimate the rate of error growth in the case when the stability criterion (6) is violated. To accomplish this we need first to obtain a different estimation of the perturbed eigenvalues λ_j involving only ϵ and M (but not the eigenvector $\bar{x}^j$ as in (7)). Since the eigenvalues of the symmetric matrix are differentiable functions of its elements, we assume that every solution b_j of the equation (9) may be written in the form

$$b_j(\epsilon) = b_j(0) + \epsilon t_j + O(\epsilon^2) , \qquad (12)$$

where $b_j(0)$ is some solution of (9) with $\epsilon = 0$:

$$(b_j(0) + \sqrt{b_j(0)^2 - 1})^M = (b_j(0) - \sqrt{b_j(0)^2 - 1})^M . \qquad (13)$$

Note that all these solutions are related to the eigenvalues $\mu_j(0) \equiv \mu_j^0$ of the unperturbed matrix T_0 via $\mu_j^0 = 2 - 2b_j(0)$ and therefore are all distinct. Substituting b_j from (12) into the equation (9), we then combine the terms of the same order of ϵ. For ϵ^0 we obtain the equation (13). For ϵ^1 we have the equation which may be solved for t to obtain $t_j = (1 - b_j(0)^2)/M$. From the perturbative formula (12) we may then easily express the first-order perturbation for $\mu_j(\epsilon)$ and then the perturbation for the eigenvalues $\lambda_j(\epsilon)$ of the error-propagating matrix $P(\epsilon)$ reads:

$$\lambda_j(\epsilon) = \frac{\zeta + \rho\mu_j(\epsilon)}{-\bar{\zeta} - \rho\mu_j(\epsilon)} = \lambda_j^0 + \frac{-8\mathrm{i}k_0\mu_j^0(\mu_j^0 - 4)}{(4\mathrm{i}k_0 + \rho\mu_j^0)^2}\frac{\epsilon}{M} + O(\epsilon^2) , \qquad (14)$$

where $\mu_j^0 = 4\mathrm{i}k_0(1 - \lambda_j^0)/(1 + \lambda_j^0)$, and $\lambda_j^0 \equiv \lambda_j(0)$ is an eigenvalue of the matrix $P(0)$.

Note that under normal circumstances (i.e. when the parameter values are typical for the acoustical problems) ϵ is indeed very small as compared to the elements of T. This is probably the reason why Mayfield [2] was able to obtain a good agreement of the eigenvalues of T_ϵ computed directly and estimated via (7) assuming α to be 0. Although the perturbative estimation (7) in her example is not exact, it is still very accurate. The requirement for ϵ to be small however obviously puts an additional restriction on the values of the mesh ratio ρ (since actually $\epsilon = \epsilon(\rho)$). Our proof removes this restriction.

Note that sometimes the requirement $|\lambda_j| \leq 1$ is not necessary for the stability. According to a more general definition [6], a numerical scheme

$$\bar{u}^{n+1} = P\bar{u}^n$$

is stable is there exist non-negative numbers K and β such that

$$\|P\|^N \leq K\mathrm{e}^{\beta r_{max}} \tag{15}$$

for any sufficiently small values Δr and Δz such that $N\Delta r = r_{max}$. We now estimate the norm of the error-propagating matrix for the scheme (5). Let λ_{max} be the eigenvalue of $P(\epsilon)$ with the maximal magnitude. If the condition (6) is violated then $|\lambda_{max}| = 1 + C\epsilon/M + O(\epsilon^2)$ (C is a certain positive constant determined from (14)). We have therefore the following asymptotic equivalence for $\|P\|^N$ (for small ϵ)

$$\|P\|^N = |\lambda_{max}|^N \sim (1 + C\epsilon/M)^N\,, \quad \epsilon \to 0$$

Now let M tend to infinity while keeping the value of $\rho = \rho_0$ constant (to investigate what happens with (5) for very fine meshes). We have then the following relation $N = \frac{r_{max}}{z_b^2\rho_0}M^2$ and estimate $\|P\|^N$ for large M and constant $\rho = \rho_0$ as follows:

$$\|P\|^N \sim (1 + C\epsilon/M)^N \sim \exp\left(\frac{r_{max}C\epsilon}{z_b^2\rho}M\right)\,, \quad M \to \infty\,. \tag{16}$$

This expression for $\|P\|^N$ clearly does not satisfy the stability criterion (15) which is less restrictive than the one of Mayfield [2].

6. Conclusion

In this work we reestablished the classical result of Mayfield [2] on the conditional stability of the numerical scheme (5). We pointed out an inaccuracy in the original proof of the crucial Lemma 2 and presented an alternative one. It was shown that even under the relaxed stability condition (15) the scheme (5) is only conditionally stable when the Mayfield condition (6) is violated.

7. Acknowledgements

The reported study was accomplished during P.S. Petrov's visit to the Bergische Universität Wuppertal under the DAAD program "Forschungsaufenthalte für Hochschullehrer und Wissenschaftler". P.S. Petrov also acknowledges RFFI foundation for the partial support under the contract No. 14-05-3148614_mol_a.

References

[1] X. Antoine, A. Arnold, C. Besse, M. Ehrhardt, A. Schädle, A review of transparent and artificial boundary condition techniques for linear and nonlinear Schrödinger equations, Commun. Comp. Phys. 4 (2008) 729–796.

[2] B. Mayfield, Nonlocal boundary conditions for the Schrödinger equation, Ph.D. thesis, University of Rhode Island, Providence, RI (1989).

[3] V. A. Baskakov, A. V. Popov, Implementation of transparent boundaries for numerical solution of the Schrödinger equation, Wave Motion 14 (1991) 123–128.

[4] Z. Z. Sun, X. Wu, The stability and convergence of a difference scheme for the Schrödinger equation on an infinite domain using artificial boundary conditions, J. Comput. Phys. 214 (2006) 209–223.

[5] F. B. Jensen, M. B. Porter, W. A. Kuperman, H. Schmidt, Computational ocean acoustics, Springer, 2011.

[6] J. W. Thomas, Numerical partial differential equations: finite difference methods, Springer, 1995.