УДК 512.643

Об аналитических семействах матриц, порождающих ограниченные полугруппы

П.А. Бахвалов, М.Д. Сурначёв

Институт прикладной математики им. М.В. Келдыша РАН, Миусская пл., 4, Москва, 125047
E-mails: bahvalo@mail.ru (Бахвалов П.А.), peitsche@yandex.ru (Сурначёв М.Д.)

Английская версия этой статьи печатается в журнале "Numerical Analysis and Applications" № 1, Vol. 14, 2021.

Бахвалов П.А., Сурначёв М.Д. Об аналитических семействах матриц, порождающих ограниченные полугруппы // Сиб. журн. вычисл. математики / РАН. Сиб. отд-ние.— Новосибирск, 2021.— Т. 24, № 1.— С. 3–16.

Рассматриваются полудискретные линейные разностные схемы с несколькими степенями свободы на одну ячейку для уравнения переноса с постоянным коэффициентом. Система уравнений, определяющих разностную схему, после преобразования Фурье распадается на системы обыкновенных дифференциальных уравнений, причём число уравнений в каждой такой системе равно числу переменных на ячейке. Матрица этих систем уравнений аналитически зависит от волнового вектора. В общем случае она не диагонализуема, а если и приводится к диагональному виду, то, вообще говоря, с потерей аналитической зависимости от волнового вектора. В настоящей работе показывается, что в одномерном случае для схем, устойчивых в L_2, эта матрица может быть локально преобразована к блочно-диагональному виду с сохранением аналитичности.

DOI: 10.15372/SJNM20210101

Ключевые слова: *спектральный анализ, разностная схема, проекторы Риса, преобразование матриц, блочная диагонализация.*

Bakhvalov P.A., Surnachev M.D. On analytical families of matrices generating bounded semigroups // Siberian J. Num. Math. / Sib. Branch of Russ. Acad. of Sci.— Novosibirsk, 2021.— Vol. 24, № 1.— P. 3–16.

We consider linear schemes with several degrees of freedom (DOFs) for the transport equation with a constant coefficient. The Fourier transform decomposes the scheme into a number of finite systems of ODEs, the number of equations in each system being equal to the number of DOFs. The matrix of these systems is an analytical function of the wave vector. Generally such a matrix is not diagonalizable and, if it is, the diagonal form can be non-smooth. We show that in a 1D case for L_2-stable schemes the matrix can be locally transformed to a block-diagonal form preserving the analytical dependence on the wave number.

Keywords: *spectral analysis, difference scheme, Riesz projection, matrix transform, block diagonalization.*

1. Введение

Рассмотрим начальную задачу для уравнения переноса

$$\begin{aligned} \frac{\partial v}{\partial t} + (\boldsymbol{\omega}\cdot\boldsymbol{\nabla})v = 0, \\ v(0,\boldsymbol{r}) = v_0(\boldsymbol{r}), \end{aligned} \tag{1.1}$$

с постоянной скоростью $\boldsymbol{\omega}$ в d-мерном единичном кубе, $d \in \mathbb{N}$, с периодическими граничными условиями. Предметом настоящего исследования являются полудискретные линейные разностные схемы с несколькими степенями свободы на одну ячейку для уравнения (1.1) вида

$$\sum_{\boldsymbol{\zeta}\in\mathcal{S}} Z_{\boldsymbol{\zeta}} \frac{du_{\boldsymbol{\eta}+\boldsymbol{\zeta}}}{dt}(t) + \frac{1}{h}\sum_{\boldsymbol{\zeta}\in\mathcal{S}} L_{\boldsymbol{\zeta}} u_{\boldsymbol{\eta}+\boldsymbol{\zeta}}(t) = 0, \quad \boldsymbol{\eta}\in\mathbb{Z}^d, \quad u_{\boldsymbol{\eta}}\in\mathbb{C}^{M^0}, \tag{1.2}$$

где $\mathcal{S}\subset\mathbb{Z}^d$ — конечное множество (шаблон схемы), M^0 — конечное множество (степеней свободы на ячейке), а $Z_{\boldsymbol{\zeta}}$ и $L_{\boldsymbol{\zeta}}$ — действительнозначные матрицы. Здесь и далее $\mathbb{Z}^d = \underbrace{\mathbb{Z}\times\cdots\times\mathbb{Z}}_{d}$, а символом $\mathbb{C}^{M^0}$ будем обозначать множество комплекснозначных векторов размерности $|M^0|$. Вопрос задания начальных данных для системы ОДУ (1.2) мы опустим, так как в дальнейшем изложении они не используются.

Будем рассматривать периодические по пространству решения (1.2). Пусть N — период решения, т. е. $u_{\boldsymbol{\eta}+N\boldsymbol{\zeta}} = u_{\boldsymbol{\eta}}$ для всех $\boldsymbol{\eta},\boldsymbol{\zeta}\in\mathbb{Z}^d$. Определим

$$\widehat{u}(t,\boldsymbol{\varphi}) = N^{-d/2}\sum_{\boldsymbol{\eta}\in I_N^d} u_{\boldsymbol{\eta}}(t)\exp(-i\boldsymbol{\varphi}\cdot\boldsymbol{\eta}),$$

где $I_N^d = (-\lfloor N/2\rfloor,\ldots,N-\lfloor N/2\rfloor)^d$. После преобразования Фурье система (1.2) распадается на системы ОДУ:

$$Z(\boldsymbol{\varphi})\frac{d\widehat{u}(t,\boldsymbol{\varphi})}{dt} + \frac{1}{h}L(\boldsymbol{\varphi})\widehat{u}(t,\boldsymbol{\varphi}) = 0,$$
$$Z(\boldsymbol{\varphi}) = \sum_{\boldsymbol{\eta}\in\mathcal{S}} Z_{\boldsymbol{\eta}}\exp\left(i\boldsymbol{\varphi}\cdot\boldsymbol{\eta}\right), \qquad L(\boldsymbol{\varphi}) = \sum_{\boldsymbol{\eta}\in\mathcal{S}} L_{\boldsymbol{\eta}}\exp\left(i\boldsymbol{\varphi}\cdot\boldsymbol{\eta}\right),$$

где $\boldsymbol{\varphi} = 2\pi\boldsymbol{k}/N$, $\boldsymbol{k}\in I_N^d$, $\boldsymbol{\varphi}\cdot\boldsymbol{\eta} = \varphi_1\eta_1+\cdots+\varphi_d\eta_d$. Число уравнений в одной такой системе равно числу переменных на ячейке. Предполагается, что для всех $\boldsymbol{\varphi}\in\mathbb{R}^d$ матрица $Z(\boldsymbol{\varphi})$ невырождена.

Для анализа точности схем вида (1.2) представляет интерес матрица

$$A(\boldsymbol{\varphi}) = i\boldsymbol{\omega}\cdot\boldsymbol{\varphi} I - Z^{-1}(\boldsymbol{\varphi})L(\boldsymbol{\varphi}), \tag{1.3}$$

которую мы назовём *матрицей ошибки*. Через неё выражается образ Фурье ошибки аппроксимации, а через матричную экспоненту $\exp(\nu A(\boldsymbol{\varphi}))$ — образ ошибки решения.

Матрица $A(\boldsymbol{\varphi})$ аналитически зависит от параметра, поскольку такой же зависимостью обладают $Z(\boldsymbol{\varphi})$ и $L(\boldsymbol{\varphi})$. Работа с матричной экспонентой существенно упрощается, если матрица $A(\boldsymbol{\varphi})$ диагонализуема с сохранением аналитичности при $\boldsymbol{\varphi}\in\mathbb{R}^d$. Однако этот факт справедлив, вообще говоря, только в одномерном случае и только при выполнении условия $(A(\varphi))^* = -A(\varphi)$: матрица $iA(\varphi)$ является эрмитовой на действительной оси, что позволяет применить к ней известные результаты теории возмущений, см. [1, 2] и [3, § 3.5.5, следствие 3]. Отметим также работу [4], в которой рассматривается блочная диагонализация равномерно сильно гиперболических матриц.

Произвольную матрицу $A(\boldsymbol{\varphi})$, $\boldsymbol{\varphi}\in\mathbb{C}^d$, можно в окрестности $\boldsymbol{\varphi}=0$ с сохранением аналитичности привести к блочно-диагональному виду $A(\boldsymbol{\varphi}) = S(\boldsymbol{\varphi})M(\boldsymbol{\varphi})S^{-1}(\boldsymbol{\varphi})$ таким образом, чтобы каждый блок матрицы $M(0)$ имел только одно (возможно, кратное) собственное значение. Хотя все рассуждения, необходимые для его доказательства, в том или ином виде присутствуют в [1], само утверждение в ней не приведено. Мы приведём его доказательство, опуская доказательства лемм, в точности совпадающих с доказанными в [1]. Кроме того, мы покажем, что матрица $A(\boldsymbol{\varphi})$, действительнозначная при $i\boldsymbol{\varphi}\in\mathbb{R}^d$, может быть в окрестности $\boldsymbol{\varphi} = 0$ с сохранением аналитичности и действительнозначности приведена к блочно-диагональному виду, так чтобы каждый блок матрицы $M(0)$ либо имел одно действительное собственное значение, либо пару комплексно сопряжённых.

В одномерном случае можно получить более сильные результаты. Если матрица $A(\varphi)$ косоэрмитова, то каждый блок $M_j(\varphi)$ матрицы $M(\varphi)$ при $\varphi = 0$ оказывается равным $M_j(0) = \lambda_j I$, где I — единичная матрица. Поэтому можно записать $M_j(\varphi) = \lambda_j + \varphi \widetilde{M}_j(\varphi)$ и применить к матрице $\widetilde{M}_j(\varphi)$ предыдущее утверждение. Повторив эту операцию конечное число раз, матрицу $A(\varphi)$ можно с сохранением аналитичности привести к диагональному виду в окрестности $\varphi = 0$ (дополнительно можно показать существование разложения с аналитическими $M(\varphi)$ и $S(\varphi)$ на всей оси $\varphi \in \mathbb{R}$, см. [1, § 2.6.2] . В случае некосоэрмитовых матриц $A(\varphi)$ матрица $M_j(0)$ может не иметь простой структуры, из-за чего применительно к абстрактной матрице $A(\varphi)$ не удаётся сформулировать практически значимого утверждения. Однако в специальных случаях такое разложение возможно. Например, в [5] его существование было доказано для генератора возмущённых марковских цепей.

В настоящей работе показывается, что применительно к матрице ошибки (1.3) преодолеть эту трудность позволяет условие L_2-устойчивости схемы (1.2). В этом случае матрица $A(\varphi)$ может быть в некоторой окрестности $\varphi = 0$ с сохранением аналитичности представлена в виде $A(\varphi) = S(\varphi)M(\varphi)S^{-1}(\varphi)$ таким образом, чтобы матрица $M(\varphi)$ была блочно-диагональной, причём в каждом её блоке все собственные значения имели один и тот же порядок малости по φ при $\varphi \to 0$.

2. Спектральные проекторы матриц

Далее через $\mathbb{R}^{n\times m}$ и $\mathbb{C}^{n\times m}$ будем обозначать пространство действительно- и комплекснозначных матриц размера $n \times m$. Через $C(\mathbb{C}^d, \mathbb{C}^{n\times m})$ будем обозначать множество функций из $\mathbb{C}^d$ в $\mathbb{C}^{n\times m}$, непрерывных в окрестности $\boldsymbol{\varphi} = 0$, а через $\mathcal{A}(\mathbb{C}^d, \mathbb{C}^{n\times m})$ — множество функций из $\mathbb{C}^d$ в $\mathbb{C}^{n\times m}$, аналитических при $\boldsymbol{\varphi} = 0$. Через $\mathcal{A}(\mathbb{C}^d, \mathbb{R}^{n\times m}) \subset \mathcal{A}(\mathbb{C}^d, \mathbb{C}^{n\times m})$ будем обозначать подмножество $\mathcal{A}(\mathbb{C}^d, \mathbb{C}^{n\times m})$ функций, действительнозначных при $i\boldsymbol{\varphi} \in \mathbb{R}^d$. Через $\sigma(A)$ будем обозначать спектр матрицы A.

Лемма 1. *Пусть $A \in \mathbb{C}^{n\times n}$, $\lambda \in \sigma(A)$. Пусть $0 < \rho < \operatorname{dist}(\lambda, \sigma(A) \setminus \{\lambda\})$ (здесь и далее расстояние до пустого множества будем считать равным бесконечности). Тогда матрица*

$$P_\lambda(A) = \frac{1}{2\pi i} \oint_{|z-\lambda|=\rho} (zI - A)^{-1} dz \tag{2.1}$$

является проектором на корневое подпространство, соответствующее собственному значению λ. При этом для λ, $\mu \in \sigma(A)$ выполняется $P_\lambda(A)P_\mu(A) = \delta_{\lambda\mu} P_\lambda(A)$.

Доказательство леммы 1 см. в [1, § 1.5.3.] Матрицы $P_\lambda(A)$ называются спектральными проекторами или проекторами Риса.

Напомним, что спектр вещественной матрицы вместе с любым собственным значением содержит и комплексно сопряжённое к нему.

Лемма 2. *Пусть выполняются условия леммы 1, и $A \in \mathbb{R}^{n\times n}$. Пусть $\lambda \in \sigma(A)$. Тогда $\overline{P_\lambda}(A) = P_{\overline{\lambda}}(A)$. В частности, если $\lambda \in \mathbb{R}$, то $P_\lambda \in \mathbb{R}^{n\times n}$.*

Действительно, напрямую из (2.1) имеем

$$\overline{P_\lambda}(A) = -\frac{1}{2\pi i}\oint_{|z-\lambda|=\rho}(\overline{z}I - A)^{-1}d\overline{z} = \frac{1}{2\pi i}\oint_{|z-\overline{\lambda}|=\rho}(zI - A)^{-1}dz = P_{\overline{\lambda}}(A).$$

Лемма 3. *Пусть* $A(\boldsymbol{\varphi}) \in C(\mathbb{C}^d, \mathbb{C}^{n\times n})$. *Тогда собственные значения* $\lambda_j(\boldsymbol{\varphi})$, $j = 1, \dots, n$, *могут быть упорядочены таким образом, что* $\lambda_j(\boldsymbol{\varphi}) \to \lambda_j(0)$ *при* $\boldsymbol{\varphi} \to 0$.

Это утверждение следует из теоремы Руше для функций $f(z) = \det(A(0) - zI)$ и $g(z) = \det(A(\boldsymbol{\varphi}) - zI) - f(z)$.

Определение 1. Пусть $A(\boldsymbol{\varphi}) \in C(\mathbb{C}^d, \mathbb{C}^{n\times n})$. Пусть $\lambda \in \sigma(A(0))$. Совокупность собственных значений $\varkappa(\boldsymbol{\varphi})$ матрицы $A(\boldsymbol{\varphi})$, таких что $\varkappa(\boldsymbol{\varphi}) \to \lambda$ при $\boldsymbol{\varphi} \to 0$, называется λ-*группой*, а сумма соответствующих корневых подпространств — *тотальным собственным подпространством*. Оператор $\mathcal{P}_\lambda(\boldsymbol{\varphi}) = \sum P_{\varkappa(\boldsymbol{\varphi})}(A(\boldsymbol{\varphi}))$, где суммирование ведётся по всем собственным значениям $\varkappa(\boldsymbol{\varphi})$, входящим λ-группу, называется *тотальным проектором* λ-*группы*.

Поскольку $P_\lambda(A)P_\mu(A) = 0$ при $\lambda \neq \mu$, тотальный проектор $\mathcal{P}_\lambda(\boldsymbol{\varphi})$ есть проектор на сумму корневых подпространств, соответствующих собственным значениям, входящим в λ-группу. А поскольку сумма корневых подпространств совпадает со всем пространством, сумма тотальных проекторов всех λ-групп совпадает с тождественным оператором.

Лемма 4. *Пусть* $A(\boldsymbol{\varphi}) \in \mathcal{A}(\mathbb{C}^d, \mathbb{R}^{n\times n})$ *и* $\lambda \in \sigma(A(0))$. *Тогда тотальный проектор* λ-*группы* $\mathcal{P}_\lambda(\boldsymbol{\varphi}) \in \mathcal{A}(\mathbb{C}^d, \mathbb{C}^{n\times n})$. *Если дополнительно* $\lambda \in \mathbb{R}$, *то* $\mathcal{P}_\lambda(\boldsymbol{\varphi}) \in \mathcal{A}(\mathbb{C}^d, \mathbb{R}^{n\times n})$.

Действительно, пусть G_λ — λ-группа, соответствующая собственному значению λ. Пусть $\rho < \dfrac{1}{2}\mathrm{dist}(\lambda, \sigma(A(0)) \setminus \{\lambda\})$. По непрерывности в некоторой окрестности $\boldsymbol{\varphi} = 0$ для всех $\varkappa(\boldsymbol{\varphi}) \in G_\lambda$ выполняется $|\varkappa(\boldsymbol{\varphi}) - \lambda| < \rho$, а для всех $\varkappa(\boldsymbol{\varphi}) \notin G_\lambda$ выполняется $|\varkappa(\boldsymbol{\varphi}) - \lambda| > \rho$. Тогда

$$\mathcal{P}_\lambda(\boldsymbol{\varphi}) = \frac{1}{2\pi i}\oint_{|z-\lambda|=\rho}(zI - A(\boldsymbol{\varphi}))^{-1}dz. \tag{2.2}$$

Поскольку на контуре, где происходит интегрирование, подынтегральная функция аналитическая по $\boldsymbol{\varphi}$ и z, то $\mathcal{P}_\lambda(\boldsymbol{\varphi}) \in \mathcal{A}(\mathbb{C}^d, \mathbb{C}^{n\times n})$.

Пусть теперь $\lambda \in \mathbb{R}$. В силу леммы 2 для каждого $\varkappa(\boldsymbol{\varphi}) \in \sigma(A(\boldsymbol{\varphi}))$ справедливо $\overline{P}_{\varkappa(\boldsymbol{\varphi})}(A(\boldsymbol{\varphi})) = P_{\overline{\varkappa}(\boldsymbol{\varphi})}(A(\boldsymbol{\varphi}))$, поэтому $P_{\varkappa(\boldsymbol{\varphi})}(A(\boldsymbol{\varphi})) + P_{\overline{\varkappa}(\boldsymbol{\varphi})}(A(\boldsymbol{\varphi})) \in \mathbb{R}^{n\times n}$. Таким образом, $\mathcal{P}_\lambda(\varphi) \in \mathcal{A}(\mathbb{C}^d, \mathbb{R}^{n\times n})$.

3. Приведение матрицы к блочно-диагональному виду

Лемма 5. *Пусть* $A \in \mathbb{C}^{n\times n}$, λ_j, $j = 1, \dots, m$, — *все её собственные значения*, n_j — *их кратность*, K_j — *соответствующие корневые подпространства. Пусть* $\mu \in \mathbb{N}$, $\mu < m$. *Пусть* $I = K_1 \oplus \dots \oplus K_\mu$ *и* $J = K_{\mu+1} \oplus \dots \oplus K_m$. *Пусть* $e_k^I \in \mathbb{C}^n$, $k = 1, \dots, n_I$, *и* $e_k^J \in \mathbb{C}^n$, $k = 1, \dots, n_J$, — *базис в этих подпространствах. Тогда матрица* A *может быть представлена в виде* $A = SMS^{-1}$, *где*

$$S = \left(e_1^I, \ldots, e_{n_I}^I, e_1^J, \ldots, e_{n_J}^J\right), \quad M = \begin{pmatrix} M^{(I)} & 0 \\ 0 & M^{(J)} \end{pmatrix}, \tag{3.1}$$

причём $\sigma(M^{(I)}) = \{\lambda_j, j = 1, \ldots, \mu\}$ *и* $\sigma(M^{(J)}) = \{\lambda_j, j = \mu + 1, \ldots, m\}$.

Доказательство. Пусть $\mathcal{J}$ — нормальная форма Жордана матрицы A: $A = T\mathcal{J}T^{-1}$. Будем считать, что блоки, соответствующие собственным значениям λ_j, $j = 1, \ldots, \mu$, поставлены в начало; в противном случае домножим матрицу T на такую матрицу перестановок, чтобы это свойство выполнилось. Векторы e_k^I, $k = 1, \ldots, n_I$, и первые n_I столбцов матрицы T образуют два базиса в одном и том же пространстве I. Аналогично векторы e_k^J, $k = 1, \ldots, n_J$, и последние n_J столбцов матрицы T образуют два базиса в одном и том же пространстве J. Поэтому существует блочно-диагональная матрица преобразования

$$U = \begin{pmatrix} U^{(I)} & 0 \\ 0 & U^{(J)} \end{pmatrix},$$

такая что матрица S, определённая (3.1), представляется в виде $S = TU$. Отсюда $A = T\mathcal{J}T^{-1} = S(U^{-1}\mathcal{J}U)S^{-1}$. Матрица $M = U^{-1}\mathcal{J}U$ является блочно-диагональной как произведение трёх блочно-диагональных матриц, а спектры её блоков совпадают со спектрами соответствующих блоков $\mathcal{J}$. □

Теорема 1. *Пусть* $A \in \mathcal{A}(\mathbb{C}^d, \mathbb{C}^{n\times n})$. *Пусть* $\sigma(A(0)) = G \cup H$, $G \cap H = \varnothing$. *Тогда в некоторой окрестности* $\boldsymbol{\varphi} = 0$ *справедливо представление* $A(\boldsymbol{\varphi}) = S(\boldsymbol{\varphi})M(\boldsymbol{\varphi})S^{-1}(\boldsymbol{\varphi})$, *где* $S, M, S^{-1} \in \mathcal{A}(\mathbb{C}^d, \mathbb{C}^{n\times n})$,

$$M(\boldsymbol{\varphi}) = \begin{pmatrix} M^{(G)}(\boldsymbol{\varphi}) & 0 \\ 0 & M^{(H)}(\boldsymbol{\varphi}) \end{pmatrix}, \tag{3.2}$$

причём $\sigma(M^{(G)}(0)) = G$, $\sigma(M^{(H)}(0)) = H$, *а размер* n_G *матрицы* $M^{(G)}(\boldsymbol{\varphi})$ *равен суммарной кратности собственных значений, лежащих в* G. *Если дополнительно* $A \in \mathcal{A}(\mathbb{C}^d, \mathbb{R}^{n\times n})$ *и* $\overline{G} = G$, *то* $S, M, S^{-1} \in \mathcal{A}(\mathbb{C}^d, \mathbb{R}^{n\times n})$.

Доказательство. Обозначим через $P_\lambda(\boldsymbol{\varphi})$ тотальный проектор λ-группы, соответствующей собственному значению $\lambda \in \sigma(A(0))$. Введём

$$P_G(\boldsymbol{\varphi}) = \sum_{\lambda\in G} P_\lambda(\boldsymbol{\varphi}), \quad P_H(\boldsymbol{\varphi}) = \sum_{\lambda\in H} P_\lambda(\boldsymbol{\varphi}).$$

Обозначим через $p_j^{(G)}(\boldsymbol{\varphi})$ и $p_j^{(H)}(\boldsymbol{\varphi})$ столбцы матриц $P_G(\boldsymbol{\varphi})$ и $P_H(\boldsymbol{\varphi})$ соответственно.

Поскольку матрица $P_G(0)$ есть проектор на n_G-мерное собственное подпространство матрицы $A(0)$, её ранг равен n_G, следовательно, у $P_G(0)$ есть ненулевой минор порядка n_G. В силу непрерывности $P_G(\boldsymbol{\varphi})$ этот минор в некоторой окрестности $\boldsymbol{\varphi} = 0$ остаётся ненулевым. Следовательно, столбцы $p_{l_k}^{(G)}(\boldsymbol{\varphi})$, $k = 1, \ldots, n_G$, входящие в этот минор, образуют базис в сумме тотальных собственных подпространств матрицы $A(\boldsymbol{\varphi})$, соответствующих $\lambda \in G$. Аналогично существуют столбцы $p_{l_k}^{(H)}(\boldsymbol{\varphi})$, $k = n_G + 1, \ldots, n$, образующие базис в сумме тотальных собственных подпространств матрицы $A(\boldsymbol{\varphi})$, соответствующих $\lambda \in H$.

Воспользуемся теперь леммой 5 для матрицы $A(\boldsymbol{\varphi})$ и базисов $p_{l_k}^{(G)}(\boldsymbol{\varphi})$, $k = 1, \ldots, n_G$ и $p_{l_k}^{(H)}(\boldsymbol{\varphi})$, $k = n_G + 1, \ldots, n$. В результате имеем представление $A(\boldsymbol{\varphi}) = S(\boldsymbol{\varphi})M(\boldsymbol{\varphi})S^{-1}(\boldsymbol{\varphi})$, где

$$S(\boldsymbol{\varphi}) = \left(p_{l_1}^{(G)}(\boldsymbol{\varphi}), \ldots, p_{l_{n_G}}^{(G)}(\boldsymbol{\varphi}), p_{l_{n_G+1}}^{(H)}(\boldsymbol{\varphi}), \ldots, p_{l_n}^{(H)}(\boldsymbol{\varphi})\right),$$

$M(\boldsymbol{\varphi})$ имеет блочно-диагональный вид (3.2), $\sigma(M^{(G)}(0)) = G$ и $\sigma(M^{(H)}(0)) = H$. Матрица $S(\boldsymbol{\varphi})$ будет аналитической по $\boldsymbol{\varphi}$ в окрестности $\boldsymbol{\varphi} = 0$, так как в силу леммы 4 каждый её столбец аналитический. Поскольку корневые подпространства имеют нулевое пересечение, справедливо $\det S(0) \neq 0$ и, следовательно, $\det S(\boldsymbol{\varphi}) \neq 0$ в некоторой окрестности $\boldsymbol{\varphi} = 0$. Следовательно, аналитической будет и $S^{-1}(\boldsymbol{\varphi})$, поскольку её компоненты вычисляются через произведения элементов $S(\boldsymbol{\varphi})$, делённые на $\det S(\boldsymbol{\varphi})$. Значит, $M(\boldsymbol{\varphi}) = S^{-1}(\boldsymbol{\varphi})A(\boldsymbol{\varphi})S(\boldsymbol{\varphi})$ также будет аналитической в окрестности $\boldsymbol{\varphi} = 0$.

Пусть теперь $A \in \mathcal{A}(\mathbb{C}^d, \mathbb{R}^{n\times n})$ и $G = \overline{G}$. Возьмём произвольное $i\boldsymbol{\varphi} \in \mathbb{R}^d$. В силу леммы 2 имеем $P_G(\boldsymbol{\varphi}) \in \mathbb{R}^{n\times n}$. Поэтому $S(\boldsymbol{\varphi}) \in \mathbb{R}^{n\times n}$, а, следовательно, $M(\boldsymbol{\varphi}) \in \mathbb{R}^{n\times n}$. Таким образом, $S, M, S^{-1} \in \mathcal{A}(\mathbb{C}^d, \mathbb{R}^{n\times n})$. □

Применив теорему 1 несколько раз, легко добиться того, чтобы в матрице $M(\boldsymbol{\varphi})$ каждый блок соответствовал одному собственному значению.

Следствие 1. *Пусть $A(\boldsymbol{\varphi}) \in \mathcal{A}(\mathbb{C}^d, \mathbb{C}^{n\times n})$. Пусть $A(0)$ имеет m различных собственных значений λ_k кратности n_k, $k = 1, \ldots, m$. Тогда справедливо представление $A(\boldsymbol{\varphi}) = S(\boldsymbol{\varphi})M(\boldsymbol{\varphi})S^{-1}(\boldsymbol{\varphi})$, где $S, M, S^{-1} \in \mathcal{A}(\mathbb{C}^d, \mathbb{C}^{n\times n})$,*

$$M(\boldsymbol{\varphi}) = \begin{pmatrix} M^{(1)}(\boldsymbol{\varphi}) & 0 & \ldots & 0 \\ 0 & M^{(2)}(\boldsymbol{\varphi}) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & M^{(m)}(\boldsymbol{\varphi}) \end{pmatrix}, \tag{3.3}$$

где матрицы $M^{(k)}(\boldsymbol{\varphi})$ имеют размер n_k и спектр $M^{(k)}(0)$ состоит из одной точки λ_k.

Если $A \in \mathcal{A}(\mathbb{C}^d, \mathbb{R}^{n\times n})$ и у $A(0)$ были комплексные собственные значения, то матрицы $S(\boldsymbol{\varphi})$ и $M(\boldsymbol{\varphi})$, даваемые следствием 2, вообще говоря, не являются действительнозначными при $i\boldsymbol{\varphi} \in \mathbb{R}^d$. Можно сохранить действительнозначность, если в один блок относить комплексно сопряжённые собственные значения.

Следствие 2. *Пусть $A(\boldsymbol{\varphi}) \in \mathcal{A}(\mathbb{C}^d, \mathbb{R}^{n\times n})$. Тогда справедливо представление $A(\boldsymbol{\varphi}) = S(\boldsymbol{\varphi})M(\boldsymbol{\varphi})S^{-1}(\boldsymbol{\varphi})$, где $S, M, S^{-1} \in \mathcal{A}(\mathbb{C}^d, \mathbb{R}^{n\times n})$, $M(\boldsymbol{\varphi})$ имеет вид (3.3), причём спектр $M^{(k)}(0)$ состоит из одной точки λ_k, если $\lambda_k \in \mathbb{R}$, и двух точек λ_k и $\overline{\lambda}_k$ иначе.*

4. Основная теорема

Лемма 6. *Пусть $\widetilde{A} \in \mathbb{C}^{n\times n}$ имеет собственное значение $\lambda = 0$, которому соответствует жорданова клетка размера больше 1. Тогда* $\sup_{\nu>0} \|\exp(\nu\widetilde{A})\| = +\infty$.

Доказательство. Представим матрицу $\widetilde{A}$ в виде $\widetilde{A} = SJS^{-1}$, где J — её жорданова нормальная форма. Тогда $e^{\nu\widetilde{A}} = Se^{\nu J}S^{-1}$. Если привести явный вид для $e^{\nu J}$, видно, что $\|e^{\nu J}\| \to \infty$ при $\nu \to \infty$. Отсюда следует доказательство леммы. □

Теорема 2. *Пусть $A \in \mathcal{A}(\mathbb{C}, \mathbb{C}^{n\times n})$. Пусть существует $K > 0$, такое что для всех $\varphi \in \mathbb{R}$ и всех $\nu > 0$ справедливо $\|\exp(\nu A(\varphi))\| \leqslant K$. Тогда в окрестности $\varphi = 0$ матрица $A(\varphi)$ представима в блочно-диагональном виде:*

$$A(\varphi) = S(\varphi)M(\varphi)S^{-1}(\varphi), \tag{4.1}$$

$$M(\varphi) = \begin{pmatrix} M_0(\varphi) & 0 & \dots & 0 & 0 \\ 0 & \varphi M_1(\varphi) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \varphi^m M_m(\varphi) & 0 \\ 0 & 0 & \dots & 0 & M_\infty(\varphi) \equiv 0 \end{pmatrix}, \tag{4.2}$$

где $S, M, S^{-1} \in \mathcal{A}(\mathbb{C}, \mathbb{C}^{n\times n})$, блоки $M_k(\varphi)$, $k \in \mathbb{N}\cup\{0,\infty\}$, — квадратные, невырожденные при $\varphi = 0$ (за исключением $k = \infty$), и некоторые их них могут отсутствовать (иметь нулевой размер). Если дополнительно $A \in \mathcal{A}(\mathbb{C}, \mathbb{R}^{n\times n})$, то $S, M, S^{-1} \in \mathcal{A}(\mathbb{C}, \mathbb{R}^{n\times n})$.

Доказательство. Доказательство проведём индукцией по размерности матрицы A. При $n = 1$, полагая $S(\varphi) = I$, $M(\varphi) = A(\varphi)$, получаем представление (4.1), (4.2).

Теперь предположим, что для всех матриц размерности меньше n утверждение теоремы 2 выполняется. Рассмотрим матрицу $A(\varphi)$. Если $A(0)$ невырождена или $A(\varphi) \equiv 0$, полагая $S(\varphi) = I$, $M(\varphi) = A(\varphi)$, получаем представление (4.1), (4.2). Иначе существует $r \in \mathbb{N} \cup \{0\}$, такое что $A(\varphi) = \varphi^r \widehat{A}(\varphi)$, где $\widehat{A}(\varphi)$ — аналитическая и $\widehat{A}(0) \neq 0$. Если $\widehat{A}(0)$ невырождена, полагая $S(\varphi) = I$, $M(\varphi) = A(\varphi)$, снова получаем представление (4.1), (4.2).

Допустим теперь, что $\widehat{A}(0)$ имеет собственное значение $\lambda = 0$, которому соответствует жорданова клетка размера больше 1. В силу леммы 6 найдётся такое $\nu_A > 0$, что $\|\exp(\nu_A \widehat{A}(0))\| \geqslant K + 1$. Следовательно, для всех $\varphi > 0$ выполняется

$$\sup_{\nu>0} \|\exp(\nu A(\varphi))\| = \sup_{\nu>0} \|\exp(\nu\varphi^r \widehat{A}(\varphi))\| = \sup_{\nu>0} \|\exp(\nu \widehat{A}(\varphi))\| \geqslant \|\exp(\nu_A \widehat{A}(\varphi))\|.$$

В силу непрерывности матричной экспоненты правая часть в некоторой окрестности $\varphi = 0$ строго больше K, что противоречит условию теоремы.

Таким образом, остаётся рассмотреть случай, когда собственное значение $\lambda = 0$ для матрицы $\widehat{A}(0)$ существует и является полупростым (т. е. его геометрическая кратность совпадает с алгебраической и отлична от нуля). Воспользуемся для матрицы $\widehat{A}(\varphi)$ теоремой 1. Имеем

$$A(\varphi) = \varphi^r S_1(\varphi) \begin{pmatrix} M^{(*)}(\varphi) & 0 \\ & \\ 0 & M^{(1)}(\varphi) \end{pmatrix} (S_1(\varphi))^{-1},$$

где $M^{(*)}(0)$ невырождена, а матрица $M^{(1)}(0)$ нильпотентна; если $A \in \mathcal{A}(\mathbb{C}, \mathbb{R}^{n\times n})$, то $S_1, S_1^{-1} \in \mathcal{A}(\mathbb{C}, \mathbb{R}^{n\times n})$. Поскольку у матрицы $\widehat{A}(\varphi)$ собственное значение $\lambda = 0$ было полупростым, таким же является собственное значение $\lambda = 0$ матрицы $M^{(1)}(0)$. Следовательно, $M^{(1)}(0) = 0$ и $M^{(1)}(\varphi) = \varphi \widehat{M}(\varphi)$, где $\widehat{M}(\varphi)$ — аналитическая функция в окрестности $\varphi = 0$.

По построению размерности матриц $M^{(*)}(\varphi)$ и $M^{(1)}(\varphi)$ отличны от нуля. Значит, размерность матрицы $\widehat{M}(\varphi)$ меньше n, и поэтому для матрицы $i^{r+1}\widehat{M}(\varphi)$ верно утверждение индукции. Значит, $A(\varphi)$ представляется в виде

$$\varphi^r S_1(\varphi)\begin{pmatrix} M^{(*)}(\varphi) & 0 \\ 0 & S_2(\varphi)\varphi\begin{pmatrix} M_0(\varphi) & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & \varphi^m M_m(\varphi) & 0 \\ 0 & \dots & 0 & 0 \end{pmatrix} S_2^{-1}(\varphi) \end{pmatrix} S_1^{-1}(\varphi),$$

причём все матрицы $M_k(\varphi)$ при $k \in \mathbb{N}\cup\{0\}$, если присутствуют, то невырождены при $\varphi = 0$, а если $A \in \mathcal{A}(\mathbb{C}, \mathbb{R}^{n\times n})$, то $S_2, S_2^{-1} \in \mathcal{A}(\mathbb{C}, \mathbb{R}^{n\times n})$. Полагая

$$S(\varphi) = S_1(\varphi)\begin{pmatrix} I & 0 \\ 0 & S_2(\varphi) \end{pmatrix},$$

$$M(\varphi) = \begin{pmatrix} \varphi^r M^{(*)}(\varphi) & 0 & \dots & 0 & 0 \\ 0 & \varphi^{r+1}M_0(\varphi) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \varphi^{r+m+1}M_m(\varphi) & 0 \\ 0 & 0 & \dots & 0 & M_\infty(\varphi)\equiv 0 \end{pmatrix},$$

получаем представление (4.1), (4.2). Теорема доказана. □

5. О точности схем в длительном счёте

Будем говорить, что $A \in \mathcal{A}(\mathbb{C}^d, \mathbb{C}^{n\times n})$ устойчива с константой K, если при всех $\boldsymbol{\varphi} \in \mathbb{R}^d$ выполняется

$$\sup_{\nu\geqslant 0} \|\exp(\nu A(\boldsymbol{\varphi}))\| \leqslant K. \tag{5.1}$$

Критерии устойчивости схемы даются теоремой Крайса о матрицах [6].

Применим теорему 2 для изучения эффектов суперсходимости и повышенной точности в длительном счёте, присущих некоторым схемам вида (1.2). А именно, покажем, что они всегда объясняются наличием такого отображения функции на сеточное пространство, в смысле которого имеется более высокий порядок аппроксимации. Этот факт можно сформулировать в виде следующей теоремы.

Теорема 3. *Пусть* $A \in \mathcal{A}(\mathbb{C}, \mathbb{R}^{n\times n})$ *устойчива с константой* K, $V\in\mathcal{A}(\mathbb{C}, \mathbb{R}^{n\times 1})$, $V(0)\neq 0$. *Пусть существуют такие* $C \geqslant 0$ *и* $Q \geqslant P \geqslant 1$, *что при всех* $\nu \geqslant 0$ *и* $\varphi \in \mathbb{R}$ *справедлива оценка*

$$\|(e^{\nu A(\varphi)} - I)V(\varphi)\| \leqslant C(|\varphi|^P + |\varphi|^{Q+1}\nu). \tag{5.2}$$

Тогда существуют такие $W(\varphi) \in \mathcal{A}(\mathbb{C}, \mathbb{R}^{n\times 1})$ *и* $\widetilde{C} > 0$, *что в некоторой окрестности* $\varphi = 0$ *справедливы оценки*

$$\|V(\varphi) - W(\varphi)\| \leqslant \widetilde{C}|\varphi|^P, \quad \|AW(\varphi)\| \leqslant \widetilde{C}|\varphi|^{Q+1}. \tag{5.3}$$

Для доказательства нам понадобятся два свойства матричной экспоненты.

Лемма 7. *Для любых матриц* X *и* Y *верно*

$$\|e^{X+Y} - e^X\| \leqslant e^{\|X\|} \left(e^{\|Y\|} - 1\right). \tag{5.4}$$

Доказательство см. в [7].

Лемма 8. *Пусть A — невырожденная матрица и $\|A\| \leqslant 1$. Тогда*

$$\|(e^A - I)^{-1}\| \leqslant 4\|A^{-1}\|. \tag{5.5}$$

Доказательство. Введём

$$f(A) = \sum_{k=1}^{\infty} \frac{A^{k-1}}{k!}.$$

Тогда

$$\|f(A) - I\| = \left\| \sum_{k=2}^{\infty} \frac{A^{k-1}}{k!} \right\| \leqslant \sum_{k=2}^{\infty} \frac{\|A\|^{k-1}}{k!} \leqslant \sum_{k=2}^{\infty} \frac{1}{k!} = \left(\sum_{k=0}^{\infty} \frac{1}{k!} \right) - 2 = e - 2,$$

$$\|(e^A - I)^{-1}\| = \|A^{-1}(f(A))^{-1}\| \leqslant \|A^{-1}\| \|(I + (f(A) - I))^{-1}\| \leqslant \frac{\|A^{-1}\|}{1 - \|f(A) - I\|} \leqslant 4\|A^{-1}\|,$$

что и требовалось доказать. □

Перейдём к доказательству теоремы 3.

Доказательство.

Шаг 1. Задание искомой функции $W(\varphi)$. Воспользуемся представлением (4.1), (4.2) матричной функции A, даваемым теоремой 2. Символом $\overline{\aleph}$ будем обозначать множество $j \in \mathbb{N} \cup \{0, \infty\}$, таких что в матрице $M(\varphi)$ присутствует блок $\varphi^j M_j(\varphi)$. Введём обозначение $v(\varphi) = S^{-1}(\varphi)V(\varphi)$. Для $j \in \overline{\aleph}$ символом $v_j(\varphi)$ обозначим j-й блок $v(\varphi)$ (размерность вектора v_j соответствует размерности M_j). Функция $v_j(\varphi)$ является аналитической функцией φ в некоторой окрестности $\varphi = 0$, поскольку таковой является $S^{-1}(\varphi)$. Поэтому либо найдётся такое $p_j \in \mathbb{N} \cup \{0\}$, что для достаточно малых φ имеет место

$$\frac{1}{2} c_j |\varphi|^{p_j} \leqslant \|v_j(\varphi)\| \leqslant c_j |\varphi|^{p_j}, \tag{5.6}$$

либо в некоторой окрестности нуля выполняется $v_j(\varphi) \equiv 0$. В последнем случае положим $p_j = +\infty$.

Определим в окрестности $\varphi = 0$ аналитические функции $w_j(\varphi) = 0$, если $p_j \geqslant P$, и $w_j(\varphi) = v_j(\varphi)$, если $p_j < P$. Составим из компонент $w_j(\varphi)$ вектор $w(\varphi)$ и положим $W(\varphi) = S(\varphi)w(\varphi)$. Или, иначе,

$$W(\varphi) = S(\varphi) \begin{pmatrix} \delta_0 I & 0 & \dots & 0 & 0 \\ 0 & \delta_1 I & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \delta_m I & 0 \\ 0 & 0 & \dots & 0 & \delta_\infty I \end{pmatrix} S^{-1}(\varphi)V(\varphi), \tag{5.7}$$

где $\delta_j = 0$, если $p_j \geqslant P$, и $\delta_j = 1$, если $p_j < P$.

Шаг 2. Проверка свойств $W(\varphi)$. Вектор $W(\varphi)$ является действительнозначным при $i\varphi \in \mathbb{R}$, поскольку таковыми являются вектор $V(\varphi)$ и матрицы $S(\varphi)$ и $S^{-1}(\varphi)$. Левое неравенство в (5.3) для некоторой константы $\widetilde{C}$ выполняется по построению. Рассмотрим теперь правое неравенство в (5.3). Выражение $A(\varphi)W(\varphi)$ отличается от правой части (5.7) заменой δ_j на $\delta_j \varphi^j M_j(\varphi)$, и поэтому в некоторой окрестности $\varphi = 0$ справедливо

$$\|A(\varphi)V(\varphi)\| \leqslant \widetilde{C} \max_{j:p_j<P} |\varphi|^j \|M_j(\varphi)\| \|v_j(\varphi)\| \leqslant \widetilde{C}' |\varphi|^{\widetilde{Q}+1},$$

где

$$\widetilde{Q} = \min_{j:p_j<P} \{j + p_j - 1\}, \tag{5.8}$$

а $\widetilde{C}'$ не зависит от φ. Таким образом, для доказательства правого неравенства в (5.3) остаётся показать, что при условии (5.2) величина $\widetilde{Q}$, определённая (5.8), удовлетворяет неравенству $\widetilde{Q} \geqslant Q$. Покажем это.

Шаг 3. Доказательство $\widetilde{Q} \geqslant Q$. Пользуясь представлением (4.1), (4.2), запишем

$$\begin{aligned}(\exp(\nu A(\varphi)) - I)V(\varphi) &= S(\varphi)\Big(\exp(\nu M(\varphi)) - I\Big)v(\varphi) \\ &= S(\varphi) \begin{pmatrix} \vdots \\ \big(\exp(\nu\varphi^j M_j(\varphi)) - I\big)\, v_j(\varphi) \\ \vdots \end{pmatrix},\end{aligned}$$

Пусть m_j — размерность блока M_j матрицы M, данной в (4.2). Введём вектор $E_j(\varphi, \nu)$ размерности m_j равенством:

$$E_j(\varphi, \nu) = (\exp(\nu\varphi^j M_j(\varphi)) - I)v_j(\varphi). \tag{5.9}$$

Из (5.2) для некоторой $C' > 0$ при всех φ в некоторой окрестности нуля и всех $\nu \geqslant 0$ имеем

$$\|E_j(\varphi, \nu)\| \leqslant C'(\varphi^P + \varphi^{Q+1}\nu). \tag{5.10}$$

Покажем, что оценка (5.10) влечёт неравенство

$$p_j \geqslant \min\{P, Q + 1 - j\}, \tag{5.11}$$

где p_j определено (5.6). В случае $j = +\infty$ или $p_j = +\infty$ это неравенство очевидно, поэтому будем предполагать $j < +\infty$ и $p_j < +\infty$.

Далее будем рассматривать такие пары значений φ и ν, что $\varphi > 0$ и $\nu = \varphi^{-j}\|M_j(0)\|^{-1}$. Рассмотрим вначале величину

$$\overline{E}_j(\varphi, \nu) = (\exp(\varphi^j \nu M_j(0)) - I)v_j(\varphi). \tag{5.12}$$

По определению матричной нормы имеем

$$\big\|\overline{E}_j(\varphi, \varphi^{-j}\|M_j(0)\|^{-1})\big\| \geqslant \big\|\big[\exp(M_j(0)\|M_j(0)\|^{-1}) - I\big]^{-1}\big\|^{-1} \|v_j(\varphi)\|.$$

С использованием (5.5) и (5.6) получаем

$$\left\|\overline{E}_j\big(\varphi,\varphi^{-j}\|M_j(0)\|^{-1}\big)\right\| \geqslant \frac{1}{8}\|(M_j(0))^{-1}\|^{-1}\|M_j(0)\|^{-1}c_j\varphi^{p_j} = \frac{1}{8}c_j\mu_j^{-1}\varphi^{p_j}, \quad (5.13)$$

где μ_j — число обусловленности матрицы $M_j(0)$. Матрица $M_j(0)$ невырожденная (см. теорему 2), поэтому $\mu_j < \infty$.

Получим теперь оценку на $E_j(\varphi,\nu)$. Представим $M_j(\varphi) = M_j(0) + \varphi\varkappa(\varphi)$, где $\varkappa \in \mathcal{A}(\mathbb{C},\mathbb{C}^{n\times n})$. Тогда (5.9) переписывается в виде

$$E_j(\varphi,\nu) = \left[\exp\left(\varphi^j\nu M_j(0) + \varphi^{j+1}\nu\varkappa(\varphi)\right) - I\right] v_j(\varphi).$$

Рассмотрим разность между E_j и $\overline{E}_j$:

$$E_j(\varphi,\nu) - \overline{E}_j(\varphi,\nu) = \left[\exp\left(\varphi^j\nu M_j(0) + \varphi^{j+1}\nu\varkappa(\varphi)\right) - \exp\left(\varphi^j\nu M_j(0)\right)\right] v(\varphi).$$

Оценим её по формуле (5.4) при рассматриваемых ограничениях на φ и ν:

$$\begin{aligned}
&\left\|E_j(\varphi,\varphi^{-j}\|M_j(0)\|^{-1}) - \overline{E}_j(\varphi,\varphi^{-j}\|M_j(0)\|^{-1})\right\| \\
&\qquad\qquad \leqslant e\left[\exp(\varphi\|\varkappa(\varphi)\|\|M_j(0)\|^{-1}) - 1\right] c_j\varphi^{p_j} \\
&\qquad\qquad \leqslant e\varphi\|\varkappa(\varphi)\|\|M_j(0)\|^{-1}\exp(\varphi\|\varkappa(\varphi)\|\|M_j(0)\|^{-1})\, c_j\varphi^{p_j}.
\end{aligned}$$

При написании последнего неравенства использовалось $e^y - 1 \leqslant ye^y$ при $y \geqslant 0$. Отсюда по неравенству треугольника из (5.13) в некоторой окрестности $\varphi = 0$ получаем

$$\left\|E_j(\varphi,\varphi^{-j}\|M_j(0)\|^{-1})\right\| \geqslant \frac{1}{9}c_j\mu_j^{-1}\varphi^{p_j}. \quad (5.14)$$

С другой стороны, из (5.10) имеем

$$\left\|E_j(\varphi,\varphi^{-j}\|M_j(0)\|^{-1})\right\| \leqslant C'\big(\varphi^P + \varphi^{Q+1-j}\|M_j(0)\|^{-1}\big). \quad (5.15)$$

Сопоставим (5.14) с (5.15), поделив на φ^{p_j}:

$$\frac{1}{9}c_j\mu_j^{-1} \leqslant C'\big(\varphi^{P-p_j} + \varphi^{Q+1-j-p_j}\|M_j(0)\|^{-1}\big).$$

Перейдём в этом неравенстве к пределу при $\varphi \to 0$. Поскольку левая часть ненулевая, правая часть не должна стремиться к нулю, т. е. должно выполняться либо $P \leqslant p_j$, либо $Q + 1 - j \leqslant p_j$. Это доказывает (5.11). Подставляя (5.11) в (5.8), получаем $\widetilde{Q} \geqslant Q$. Это заканчивает доказательство теоремы. □

6. Примеры

Может сложиться ошибочное впечатление, что наличие жордановых клеток в матрице $A(\varphi)$ свидетельствует о неустойчивости схемы. Приведём два примера устойчивых схем, аппроксимирующих уравнение переноса $\partial u/\partial t + \partial u/\partial x = 0$, в матрице которых при всех $\varphi \neq 0$ имеются жордановы клетки.

Пусть на одной ячейке определены две переменные: "L" и "R". Пусть начальные данные отображаются на сетку оператором $(\Pi_h f)_{j,L} = (\Pi_h f)_{j,R} = f(jh)$. Для каждой производной запишем схему с направленной разностью, а для степеней свободы "L" добавим к ней дополнительную диссипацию, посчитанную по переменным "R":

$$\frac{du_j^L}{dt} + \frac{u_j^L - u_{j-1}^L}{h} - \frac{u_{j+1}^R - 2u_j^R + u_{j-1}^R}{h} = 0, \quad \frac{du_j^R}{dt} + \frac{u_j^R - u_{j-1}^R}{h} = 0.$$

При этом матрица $A(\varphi) = i\varphi - L(\varphi)$ имеет вид:

$$A(\varphi) = \begin{pmatrix} \lambda(\varphi) & -2a(\varphi) \\ 0 & \lambda(\varphi) \end{pmatrix}, \tag{6.1}$$

где $\lambda(\varphi) = i\varphi - 1 + e^{-i\varphi}$ и $a(\varphi) = 1 - \cos\varphi$. При всех φ матрица $A(\varphi)$ имеет собственное значение $\lambda(\varphi)$ кратности 2, причём при $\varphi \neq 0$ оно соответствует жордановой клетке размера 2. Тем не менее схема является устойчивой:

$$\exp(\nu A(\varphi)) = \begin{pmatrix} e^{\nu\lambda(\varphi)} & -2\nu a(\varphi) e^{\nu\lambda(\varphi)} \\ 0 & e^{\nu\lambda(\varphi)} \end{pmatrix}.$$

Поскольку $\operatorname{Re}\lambda(\varphi) = -a(\varphi) < 0$, выполняется $|e^{\nu\lambda(\varphi)}| \leqslant 1$ и

$$|-2\nu a(\varphi) e^{\nu\lambda(\varphi)}| = 2\nu |a(\varphi)| \exp(\nu \operatorname{Re}\lambda(\varphi)) = 2\nu a(\varphi) \exp(-\nu a(\varphi)),$$

что в силу положительности $a(\varphi)$ не превосходит $\sup_{x>0} 2xe^{-x} = 2/e$.

Пусть теперь на одной ячейке определены три переменные: "L", "R" и "M". Пусть начальные данные отображаются на сетку оператором $(\Pi_h f)_{j,L} = (\Pi_h f)_{j,R} = (\Pi_h f)_{j,M} = f(jh)$. Запишем схему:

$$\frac{du_j^L}{dt} + \frac{u_j^L - u_{j-1}^L}{h} = 0, \qquad \frac{du_j^R}{dt} + \frac{u_j^R - u_{j-1}^L}{h} = 0,$$

$$\frac{du_j^M}{dt} + \frac{u_{j+1}^R - u_j^R}{h} + \frac{u_j^M - u_j^L}{h} = 0.$$

При этом

$$A(\varphi) = i\varphi - \begin{pmatrix} 1 - e^{-i\varphi} & 0 & 0 \\ -e^{-i\varphi} & 1 & 0 \\ -1 & e^{i\varphi} - 1 & 1 \end{pmatrix}.$$

У матрицы $A(\varphi)$ есть простое собственное значение $i\varphi - (1 - e^{-i\varphi})$ и собственное значение $\lambda = i\varphi - 1$, которому соответствует жорданова клетка размера 2. Она не мешает устойчивости, так как соответствует значению с отрицательной действительной частью, отделённой от нуля.

7. Многомерный случай

В многомерном случае легко построить контрпример, показывающий, что матрицу $A \in \mathcal{A}(\mathbb{C}^d, \mathbb{R}^{n\times n})$ нельзя привести к блочно-диагональному виду, отнеся в каждый блок собственные значения одного порядка по $|\boldsymbol{\varphi}|$.

Пусть $\boldsymbol{\varphi} = (\varphi_1, \varphi_2)$ и

$$A(\varphi_1, \varphi_2) = \begin{pmatrix} 2i\sin\varphi_1 & -2i\sin\varphi_1 - 4\sin^2(\varphi_2/2) \\ -2i\sin\varphi_1 + 4\sin^2(\varphi_2/2) & 2i\sin\varphi_1 \end{pmatrix}. \tag{7.1}$$

Поскольку $(A(\boldsymbol{\varphi}))^* = -A(\boldsymbol{\varphi})$, она устойчива с константой $K = 1$. Собственные значения этой матрицы имеют вид $\lambda_\pm = 2i\sin\varphi_1 \pm i\sqrt{4\sin^2\varphi_1 + 16\sin^4(\varphi_2/2)}$. Одно из собственных значений ограничено величиной порядка $O(|\boldsymbol{\varphi}|)$, а второе — величиной порядка $O(|\boldsymbol{\varphi}|^2)$. Следовательно, если мы хотим привести матрицу $A(\boldsymbol{\varphi})$ к искомому виду, то $M(\boldsymbol{\varphi})$ должна быть диагональной матрицей из собственных значений. Но $\lambda_\pm(\boldsymbol{\varphi})$ не являются аналитическими функциями в окрестности $\boldsymbol{\varphi} = 0$.

Некоторые результаты о преобразовании матриц, аналитически зависящих от нескольких переменных, можно найти в [8, 9] и др.

Литература

1. **Kato T.** Perturbation theory for linear operators // Grund. math. Wiss. — Springer, 1966. — Vol. 132. — P. 477–513.
2. **Като Т.** Теория возмущений линейных операторов. — М.: Мир, 1972.
3. **Baumgärtel H.** Analytic perturbation theory for matrices and operators. — Basel; Boston; Stuttgart: Birkhäuser Verlag, 1985.
4. **Volevich L.R., Shirikyan A.R.** Remarks on Strongly Hyperbolic matrices. — Moscow, 1999. — (Preprint / Keldysh Institute of Applied Mathematics; 009 (Mi ipmp1238); — P. 1–38).
5. **Delebecque F.** A reduction process for perturbed Markov chains // SIAM J. Appl. Math. — 1983. — Vol. 43, № 2. — P. 325–350.
6. **Kreiss H.O.** Über Matrizen die beshränkte Halbgruppen erzeugen // Mathematica Scandinavica. — 1959. — Vol. 7.— P. 71–80.
7. **Levis A.H.** Some computational aspects of the matrix exponential // IEEE Trans. Automatic Control, AC-14. — 1969. — Vol. 14, iss. 4. — P. 410–411.
8. **Kurdyka K., Paunescu L.** Hyperbolic polynomials and multiparameter real analytic perturbation theory // Duke Mathematical J. — 2008. — Vol. 141, № 1. — P. 123–149.
9. **Kirillov O.N.** Eigenvalue bifurcation in multiparameter families of non-self-adjoint operator matrices // Zeitschrift für angewandte Mathematik und Physik. — 2010. — Vol. 61, № 2. — P. 221–234.

Поступила в редакцию 8 июля 2019 г.
После исправления 4 сентября 2019 г.
Принята к печати 21 октября 2020 г.

Литература в транслитерации

1. **Kato T.** Perturbation theory for linear operators // Grund. math. Wiss. — Springer, 1966. — Vol. 132. — P. 477–513.

2. **Kato T.** Teoriya vozmuschenii lineinykh operatorov. — M.: Mir, 1972.

3. **Baumgärtel H.** Analytic perturbation theory for matrices and operators. — Basel; Boston; Stuttgart: Birkhäuser Verlag, 1985.

4. **Volevich L.R., Shirikyan A.R.** Remarks on Strongly Hyperbolic matrices. — Moscow, 1999. — (Preprint / Keldysh Institute of Applied Mathematics; 009 (Mi ipmp1238); — P. 1–38).

5. **Delebecque F.** A reduction process for perturbed Markov chains // SIAM J. Appl. Math. — 1983. — Vol. 43, № 2. — P. 325–350.

6. **Kreiss H.O.** Über Matrizen die beshränkte Halbgruppen erzeugen // Mathematica Scandinavica. — 1959. — Vol. 7.— P. 71–80.

7. **Levis A.H.** Some computational aspects of the matrix exponential // IEEE Trans. Automatic Control, AC-14. — 1969. — Vol. 14, iss. 4. — P. 410–411.

8. **Kurdyka K., Paunescu L.** Hyperbolic polynomials and multiparameter real analytic perturbation theory // Duke Mathematical J. — 2008. — Vol. 141, № 1. — P. 123–149.

9. **Kirillov O.N.** Eigenvalue bifurcation in multiparameter families of non-self-adjoint operator matrices // Zeitschrift für angewandte Mathematik und Physik. — 2010. — Vol. 61, № 2. — P. 221–234.