ИПМ им.М.В.Келдыша РАН • Электронная библиотека

Препринты ИПМ • Препринт № 74 за 2019 г.

ISSN 2071-2898 (Print)
ISSN 2071-2901 (Online)

Бахвалов П.А., Сурначёв М.Д.

Линейные схемы с несколькими степенями свободы для многомерного уравнения переноса

Рекомендуемая форма библиографической ссылки: Бахвалов П.А., Сурначёв М.Д. Линейные схемы с несколькими степенями свободы для многомерного уравнения переноса // Препринты ИПМ им. М.В.Келдыша. 2019. № 74. 44 с. doi:10.20948/prepr-2019-74
URL: http://library.keldysh.ru/preprint.asp?id=2019-74

О р д е н а Л е н и н а
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
имени М.В.КЕЛДЫША
Р о с с и й с к о й а к а д е м и и н а у к

П. А. Бахвалов, М. Д. Сурначёв

Линейные схемы с несколькими степенями свободы для многомерного уравнения переноса

Москва — 2019

Бахвалов П. А., Сурначёв М. Д.

Линейные схемы с несколькими степенями свободы для многомерного уравнения переноса

Рассматриваются линейные схемы с несколькими степенями свободы на одну ячейку для многомерного уравнения переноса. Решение по такой схеме может обладать ошибкой $O(h^p+th^q)$, причём p совпадает с порядком аппроксимации или превосходит его на единицу, а $q \geqslant p$. Доказывается, что существует такое отображение гладких функций на сеточное пространство, отличающееся от обычного (например, L_2-проекции) на величину порядка h^p, в смысле которого схема будет обладать q-м порядком аппроксимации. В отличие от одномерного случая, локальное отображение с требуемыми свойствами может не существовать. Приводятся достаточные условия его существования.

Ключевые слова: аппроксимация и точность, суперсходимость

Pavel Alexeevisch Bakhvalov, Mikhail Dmitrievich Surnachev

Linear schemes with several degrees of freedom for the multidimensional transport equation

We consider linear schemes with several degrees of freedom for the transport equation. Solution error of these schemes can have estimate $O(h^p + th^q)$, where p is equal to or greater by one than approximation order and $q \geqslant p$. We prove the existence of a mapping of smooth functions on the mesh space providing the q-th order of the truncation error and deviating from the standard mapping (L_2-projection for example) by the order h^p. In contrast with 1D case local mapping with such properties generally does not exist. We prove sufficient existence conditions.

Key words: consistency and accuracy, superconvergence

1. Введение

Настоящая работа продолжает исследование линейных полудискретных схем с несколькими степенями свободы на одну ячейку для уравнения переноса с постоянным коэффициентом, начатое в [1–4]. Численная ошибка решения таких схем может обладать оценкой $O(h^p+th^q)$, причём p совпадает с порядком аппроксимации или превосходит его на единицу, а $q \geqslant p$. Например, для метода Галёркина с разрывными базисными функциями эта оценка наблюдается при $p = k+1$, $q = 2k+1$, где k – порядок используемых полиномов. В одномерном случае эта оценка доказана в [5], в двумерном случае на декартовой сетке – в [6]. На трансляционно-инвариантной треугольной сетке эта оценка также наблюдается, однако строгого доказательства этого факта авторам неизвестно.

Как ошибка аппроксимации, так и ошибка решения зависят не только от уравнений, определяющих разностную схему, но и от отображения функций из некоторого пространства решений дифференциального уравнения в пространство сеточных функций. Для того чтобы ошибка решения при использовании некоторого отображения Π_h (например, L_2-проекции) имела оценку $O(h^p+th^q)$, достаточно, чтобы существовало такое отображение $\tilde{\Pi}_h$, что:

1) $\|\tilde{\Pi}_h - \Pi_h\| = O(h^p)$;
2) в смысле $\tilde{\Pi}_h$ имеет место q-й порядок аппроксимации.

В [3] мы показали, что в одномерном случае существование такого отображения является не только достаточным, но и необходимым. При этом получение оптимальных значений p и q сводится к поиску отображения $\tilde{\Pi}_h$ определённого вида методом неопределённых коэффициентов.

Настоящая работа посвящена многомерному случаю. Как и в [3], в предположении, что ошибка решения обладает оценкой $O(h^p + th^q)$, доказывается существование отображения гладких функций на сеточное пространство, удовлетворяющего условиям 1) и 2). Однако в общем случае локального отображения, обладающего такими свойствами, не существует; приводятся достаточные условия его существования. Также изучается вопрос зависимости формального порядка точности схемы и порядка точности в длительном счёте от направления волнового вектора и направления скорости переноса.

2. Постановка задачи

Пусть $\boldsymbol{a}_j$, $j = 1, \ldots, d$, – линейно независимая система векторов, интерпретируемых как смещения сеточных блоков друг относительно друга при $h = 1$. Пусть $\boldsymbol{a}_j^*$, $j = 1, \ldots, d$, – взаимный базис к $\{\boldsymbol{a}_j\}$, то есть $\boldsymbol{a}_i^* \cdot \boldsymbol{a}_j = \delta_{ij}$. Обозначим через T линейный оператор в пространстве $\mathbb{R}^d$, сопоставляющий вектору $\boldsymbol{\eta}$ вектор $T\boldsymbol{\eta}$ с компонентами

$$T\boldsymbol{\eta} = \sum_{j=1}^{d} \eta_j \boldsymbol{a}_j. \tag{2.1}$$

Если $\boldsymbol{\eta} \in \mathbb{Z}^d$, то $T\boldsymbol{\eta}$ есть смещение сеточного блока $\boldsymbol{\eta}$ относительно нулевого. Обозначим $U = (T^*)^{-1}$. Имеем

$$(T^*\boldsymbol{x})_j = \boldsymbol{x} \cdot \boldsymbol{a}_j, \quad U\boldsymbol{y} = (T^*)^{-1}\boldsymbol{y} = \sum_{j=1}^{d} y_j \boldsymbol{a}_j^*.$$

Если $\boldsymbol{a}_j$ совпадают с векторами стандартного базиса, T и U тождественные.

Введём пространство $L_{2,per}(\mathbb{R}^d)$ как множество функций $f \in L_{2,loc}(\mathbb{R}^d)$, таких что при некотором $N_0 \in \mathbb{N}$ для всех $j = 1, \ldots, d$ и почти всех $\boldsymbol{r} \in \mathbb{R}^d$ выполняется $f(\boldsymbol{r} + N_0\boldsymbol{a}_j) = f(\boldsymbol{r})$. При этом будем называть N_0 *периодом* функции f. Для $f \in L_{2,per}(\mathbb{R}^d)$ определим норму

$$\|f\|^2 = \frac{1}{|\square|} \int\limits_{\square} |f(\boldsymbol{r})|^2 dV,$$

где $\square$ – параллелепипед, образованный векторами $N_0\boldsymbol{a}_1$, …, $N_0\boldsymbol{a}_d$.

Для $q \in \mathbb{N} \cup \{0\}$ определим пространство $H^q_{per}(\mathbb{R}^d) = L_{2,per}(\mathbb{R}^d) \cap W^q_{2,loc}(\mathbb{R}^d)$. Для $w \in H^q_{per}(\mathbb{R}^d)$, $q \in \mathbb{N} \cup \{0\}$, $r = 0, \ldots, q$, введём обозначение

$$\|\nabla^r w\|^2 = \sum_{|\boldsymbol{m}|=r} \frac{r!}{\boldsymbol{m}!} \|D^{\boldsymbol{m}} w\|^2, \quad D^{\boldsymbol{m}} = \frac{\partial^{|\boldsymbol{m}|}}{\partial x_1^{m_1} \ldots \partial x_d^{m_d}}. \tag{2.2}$$

Здесь $\boldsymbol{m} = (m_1, \ldots, m_d)$ – мультииндекс: $m_i \geqslant 0$, $|\boldsymbol{m}| = m_1 + \ldots + m_d$, $\boldsymbol{m}! = m_1! \ldots m_d!$. Введём обозначение $\boldsymbol{r}^m = x_1^{m_1} \ldots x_d^{m_d}$ для $\boldsymbol{r} = (x_1, \ldots, x_d)$. На $H^q_{per}(\mathbb{R}^d)$ будем использовать семейство норм $\|f\|^2_{(q,h)} = \sum\limits_{r=0}^{q} h^{2r} \|\nabla^r f\|^2$.

Также для $q \in \mathbb{N} \cup \{0\}$ введём $C^q_{per}(\mathbb{R}^d) = C^q(\mathbb{R}^d) \cup L_{2,per}(\mathbb{R}^d)$. На $C_{per}(\mathbb{R}^d) \equiv C^0_{per}(\mathbb{R}^d)$ будем использовать норму $\|f\|_\infty = \sup_{\boldsymbol{r}} |f(\boldsymbol{r})|$. На $C^q(\mathbb{R}^d)$ будем использовать семейство норм

$$\|f\|^2_{(\infty,q,h)} = \sum_{r=0}^{q} h^{2r} \|\nabla^r f\|^2_\infty, \quad \|\nabla^r f\|^2_\infty = \sum_{|\boldsymbol{m}|=r} \frac{r!}{\boldsymbol{m}!} \|D^{\boldsymbol{m}} w\|^2_\infty.$$

Такое введение норм обеспечивает выполнение условия $\|f\|_{(q,h)} \leqslant \|f\|_{(\infty,q,h)}$.

В настоящей работе рассматривается начальная задача для линейного уравнения переноса

$$\frac{\partial v}{\partial t} + \boldsymbol{\omega} \cdot \nabla v = 0, \quad \boldsymbol{r} \in \mathbb{R}^d, \tag{2.3}$$

$$v(0, \boldsymbol{r}) = v_0(\boldsymbol{r}) \in L_{2,per}(\mathbb{R}^d). \tag{2.4}$$

Скорость переноса $\boldsymbol{\omega}$ полагается постоянной во времени и пространстве.

Введём следующие обозначения. M^0 – конечное множество степеней свободы в одном сеточном блоке. $M = \mathbb{Z}^d \times M^0$ – общее множество степеней свободы. Если $f \in \mathbb{C}^M$, то $f_{\boldsymbol{\eta}} \in \mathbb{C}^{M^0}$ – часть вектора f в блоке $\boldsymbol{\eta} \in \mathbb{Z}^d$. V_{per}^N – множество последовательностей с периодом N:

$$V_{per}^N = \{f \in \mathbb{C}^M \ : \ \forall \boldsymbol{\eta}, \boldsymbol{\zeta} \in \mathbb{Z}^d \ \ f_{\boldsymbol{\eta}+N\boldsymbol{\zeta}} = f_{\boldsymbol{\eta}}\}.$$

$V_{per} = \bigcup\limits_{N \in \mathbb{N}} V_{per}^N$ – множество периодических последовательностей со скалярным произведением, определяемым для $f \in V_{per}^{N(f)}$, $g \in V_{per}^{N(g)}$ формулой

$$(f, g) = \frac{1}{N^d} \sum_{\boldsymbol{\eta}=(0,\ldots,N-1)^d} (f_{\boldsymbol{\eta}}, g_{\boldsymbol{\eta}}), \quad N = N(f)N(g).$$

Здесь $(f_{\boldsymbol{\eta}}, g_{\boldsymbol{\eta}})$ – некоторое скалярное произведение на $\mathbb{C}^{M^0}$. Если функция f имеет период $N(f)$, то она, очевидно, также имеет период $nN(f)$ для любого $n \in \mathbb{N}$, но на значение скалярного произведения замена $N(f)$ на $nN(f)$ не влияет. На $\mathbb{C}^{M^0}$ и V_{per} будем использовать нормы, порождённые этими скалярными произведениями: $\|f_{\boldsymbol{\eta}}\|^2 = (f_{\boldsymbol{\eta}}, f_{\boldsymbol{\eta}})$; $\|f\|^2 = (f, f)$.

Для аппроксимации (2.3) рассмотрим полудискретные схемы вида

$$\sum_{\boldsymbol{\zeta} \in \mathcal{S}} Z_{\boldsymbol{\zeta}} \frac{du_{\boldsymbol{\eta}+\boldsymbol{\zeta}}}{dt}(t) + \frac{1}{h} \sum_{\boldsymbol{\zeta} \in \mathcal{S}} L_{\boldsymbol{\zeta}} u_{\boldsymbol{\eta}+\boldsymbol{\zeta}}(t) = 0, \quad \boldsymbol{\eta} \in \mathbb{Z}^d, \quad u_{\boldsymbol{\eta}} \in \mathbb{C}^{M^0}, \tag{2.5}$$

где $\mathcal{S} \subset \mathbb{Z}^d$ – конечное множество, которое мы будем называть шаблоном схемы, а $Z_{\boldsymbol{\zeta}}$ и $L_{\boldsymbol{\zeta}}$ – действительнозначные матрицы. Для $\boldsymbol{\zeta} \notin \mathcal{S}$ будем полагать $Z_{\boldsymbol{\zeta}} = L_{\boldsymbol{\zeta}} = 0$.

Введём операторы $Z : \mathbb{C}^M \to \mathbb{C}^M$ и $L : \mathbb{C}^M \to \mathbb{C}^M$ равенствами

$$(Zu)_{\boldsymbol{\eta}} = \sum_{\boldsymbol{\zeta} \in \mathcal{S}} Z_{\boldsymbol{\zeta}} u_{\boldsymbol{\eta}+\boldsymbol{\zeta}}, \quad (Lu)_{\boldsymbol{\eta}} = \sum_{\boldsymbol{\zeta} \in \mathcal{S}} L_{\boldsymbol{\zeta}} u_{\boldsymbol{\eta}+\boldsymbol{\zeta}}. \tag{2.6}$$

Очевидно, $Z\,V_{per}^N \subseteq V_{per}^N$, $L\,V_{per}^N \subseteq V_{per}^N$, откуда $Z\,V_{per} \subseteq V_{per}$, $L\,V_{per} \subseteq V_{per}$.

Будем предполагать, что схема (2.5) является устойчивой, то есть существует $K > 0$, такая что для всех $u \in C^\infty([0,\infty), V_{per})$, являющихся решением (2.5), при всех $t \geqslant 0$ выполняется $\|u(t)\| \leqslant K\|u(0)\|$.

Всюду далее будем считать, что $1/h \in \mathbb{N}$. Для отображения данных на пространство сеточных функций используются операторы $\Pi_h, \mathcal{P}_h : L_{2,loc}(\mathbb{R}^d) \to \mathbb{C}^M$, задаваемые формулами

$$(\Pi_h f)_{\boldsymbol{\eta},\xi} = \int\limits_{\boldsymbol{r}\in G} \mu_\xi(\boldsymbol{r}) f(h(\boldsymbol{r}+T\boldsymbol{\eta}))d\boldsymbol{r}, \quad (\mathcal{P}_h f)_{\boldsymbol{\eta},\xi} = \int\limits_{\boldsymbol{r}\in G} \hat{\mu}_\xi(\boldsymbol{r}) f(h(\boldsymbol{r}+T\boldsymbol{\eta}))d\boldsymbol{r}, \tag{2.7}$$

где $\boldsymbol{\eta} \in \mathbb{Z}^d$ – индекс блока, $\xi \in M^0$ – индекс переменной внутри блока, $G \subset \mathbb{R}^d$ – некоторая ограниченная область, $\mu_\xi, \hat{\mu}_\xi \in L_2(G)$, $\int_G \mu_\xi(\boldsymbol{r}) = 1$. Если $f \in L_{2,per}(\mathbb{R}^d)$ имеет период N_0, то $\Pi_h f, \mathcal{P}_h f \in V_{per}^{(N_0/h)}$. Операторы Π_h и $\mathcal{P}_h$ как операторы из $L_{2,per}(\mathbb{R}^d)$ в V_{per} являются ограниченными равномерно по h.

Также будем рассматривать оператор $\tilde{\Pi}_h^{(p,q)} : W_{2,loc}^q(\mathbb{R}^d) \to \mathbb{C}^M$, задаваемый формулой

$$(\tilde{\Pi}_h^{(p,q)} f)_{\boldsymbol{\eta}} = (\Pi_h f)_{\boldsymbol{\eta}} + \sum_{0<|\boldsymbol{m}|\leqslant q} h^{|\boldsymbol{m}|} \mathfrak{C}^{(\boldsymbol{m})} \left(\mathcal{P}_h D^{\boldsymbol{m}} f\right)_{\boldsymbol{\eta}}, \tag{2.8}$$

где $\mathfrak{C}^{(\boldsymbol{m})}$ – некоторые диагональные матрицы размера $|M^0| \times |M^0|$.

Будем также рассматривать операторы $\mathring{\Pi}_h, \mathring{\mathcal{P}}_h : C_{loc} \to \mathbb{C}^M$, задаваемые формулами

$$(\mathring{\Pi}_h f)_{\boldsymbol{\eta},\xi} = f\left(h\left(\boldsymbol{\rho}_\xi + T\boldsymbol{\eta}\right)\right), \quad (\mathring{\mathcal{P}}_h f)_{\boldsymbol{\eta},\xi} = f\left(h\left(\hat{\boldsymbol{\rho}}_\xi + T\boldsymbol{\eta}\right)\right), \tag{2.9}$$

и оператор $\mathring{\tilde{\Pi}}_h^{(p,q)} : C_{loc}^q \to \mathbb{C}^M$, определяемый как

$$\left(\mathring{\tilde{\Pi}}_h^{(p,q)} f\right)_{\boldsymbol{\eta}} = \left(\mathring{\Pi}_h f\right)_{\boldsymbol{\eta}} + \sum_{p\leqslant|\boldsymbol{m}|\leqslant q} h^{|\boldsymbol{m}|} \mathfrak{C}^{(\boldsymbol{m})} (\mathring{\mathcal{P}}_h (D^{\boldsymbol{m}} f))_{\boldsymbol{\eta}}. \tag{2.10}$$

Оператор Π_h вида (2.7) является частным случаем (2.8) при $p = q = 0$, поэтому все утверждения, которые будут доказаны для $\tilde{\Pi}_h$, справедливы и для Π_h. Аналогично $\mathring{\Pi}_h$ вида (2.9) является частным случаем (2.10). Иногда для краткости верхние индексы p и q будем опускать.

Определение 1. *Пусть $\tilde{\Pi}_h$ – некоторый оператор вида* (2.8) *или* (2.10). *Ошибкой решения в смысле $\tilde{\Pi}_h$ с начальными данными v_0 будем называть величину*

$$\varepsilon_h(t, v_0, \tilde{\Pi}_h) = u(t) - \tilde{\Pi}_h v(t, \cdot), \tag{2.11}$$

где $u(t)$ – решение (2.5) *с условием $u(0) = \tilde{\Pi}_h v_0$, а $v(t,\boldsymbol{r}) = v_0(\boldsymbol{r} - \boldsymbol{\omega}t)$.*

Определение 2. *Пусть* $\mathbf{\Pi}$ *– оператор из* $H^q_{loc}(\mathbb{R}^d)$ *или* $C^q(\mathbb{R}^d)$ *в* $\mathbb{C}^M$*. Ошибкой аппроксимации на функции* f *в смысле* $\mathbf{\Pi}$ *будем называть величину* $\epsilon_h(f, \mathbf{\Pi}) \in \mathbb{C}^M$*, определяемую формулой*

$$\epsilon_h(f, \mathbf{\Pi}) = -Z\mathbf{\Pi}(\boldsymbol{\omega} \cdot \nabla f) + \frac{1}{h} L\mathbf{\Pi} f. \tag{2.12}$$

Если $\tilde{\Pi}^{(q)}_h$ – оператор вида (2.8), то для функций $f \in H^{q+1}_{per}(\mathbb{R}^d)$, имеющих период N_0, выполняется $\epsilon_h(f, \tilde{\Pi}^{(p,q)}_h) \in V^{(N_0/h)}_{per}$.

Определение порядка точности в длительном счёте, данное в [3], переносится на многомерный случай.

Определение 3. *Предположим, что существуют такие константы* C_1 *и* C_2*, что при всех начальных условиях* $v_0 \in H^r_{per}(\mathbb{R}^d)$*,* $r \geqslant Q + 1$*, и при всех* t *и* h *схема обладает оценкой ошибки*

$$\|\varepsilon_h(t, v_0, \tilde{\Pi}_h)\| \leqslant C_1 h^P \|\nabla^P v_0\| + C_2 (t + h) h^Q \|\nabla^{Q+1} v_0\|, \tag{2.13}$$

где $+\infty \geqslant Q \geqslant P > 0$*. Тогда будем говорить, что схема обладает формальным порядком точности* P *и порядком точности в длительном счёте* Q *в смысле* $\tilde{\Pi}_h$ *на* $H^r_{per}(\mathbb{R}^d)$*. Если выполняется оценка* $\|\varepsilon_h(t, v_0, \tilde{\Pi}_h)\| \leqslant C_1 h^P \|\nabla^P v_0\|$ *при* $r \geqslant P > 0$*, будем говорить, что* $Q = \infty$*, а если* $\varepsilon_h(t, v_0, \tilde{\Pi}_h) \equiv 0$*, будем говорить, что* $P = Q = \infty$*. Если для всех* $P' > P$*,* $Q' \geqslant P'$ *и для всех* $P' = P$*,* $Q' > Q$ *не существует оценки вида* (2.13)*, то такие параметры* P *и* Q *будем называть оптимальными.*

Определение порядков на $C^r_{per}(\mathbb{R}^d)$ отличается от определения 3 заменой $H^r_{per}(\mathbb{R}^d)$ на $C^r_{per}(\mathbb{R}^d)$ и норм $\|\ \cdot\ \|$ в правой части (2.13) на $\|\ \cdot\ \|_\infty$.

3. Спектральное представление схемы

Для $\boldsymbol{\phi} \in \mathbb{C}^d$ введём матрицы

$$Z(\boldsymbol{\phi}) = \sum_{\boldsymbol{\eta} \in \mathcal{S} \subset \mathbb{Z}^d} Z_{\boldsymbol{\eta}} \exp\left(i\boldsymbol{\phi} \cdot \boldsymbol{\eta}\right), \quad L(\boldsymbol{\phi}) = \sum_{\boldsymbol{\eta} \in \mathcal{S} \subset \mathbb{Z}^d} L_{\boldsymbol{\eta}} \exp\left(i\boldsymbol{\phi} \cdot \boldsymbol{\eta}\right), \tag{3.1}$$

$$A(\boldsymbol{\phi}) = -Z^{-1}(\boldsymbol{\phi}) L(\boldsymbol{\phi}) + i(T^{-1}\boldsymbol{\omega}) \cdot \boldsymbol{\phi} I. \tag{3.2}$$

Здесь $\boldsymbol{\phi} \cdot \boldsymbol{\eta} = \phi_1\eta_1 + \ldots + \phi_d\eta_d$. Для $\boldsymbol{\phi} \in \mathbb{R}^d$ детерминант $Z(\boldsymbol{\phi})$ отделён от нуля. Схема (2.5) устойчива с константой K тогда и только тогда, когда при всех $\boldsymbol{\phi} \in \mathbb{R}^d$ и $\nu > 0$ выполняется

$$\|\exp\left(A(\boldsymbol{\phi})\nu\right)\| \leqslant K. \tag{3.3}$$

В частности, если $A(\boldsymbol{\phi}) = -A^*(\boldsymbol{\phi})$, то схема устойчива с $K = 1$.

Обозначим через $\mathfrak{e}$ вектор размерности M^0, состоящий из единиц. Будем отождествлять его с вектором $\mathfrak{e} \in V_{per}$, все блочные компоненты которого равны $\mathfrak{e}$. Всюду далее будем предполагать, что схема точна на константе, то есть $L(0)\mathfrak{e} = 0$.

Функции аппроксимационной ошибки и ошибки решения в образах Фурье имеют вид

$$\hat{\epsilon}(\boldsymbol{\phi}, \tilde{\Pi}_h) = A(\boldsymbol{\phi})(\tilde{\Pi}_1 e^{i(U\boldsymbol{\phi})\cdot\boldsymbol{r}})_0, \tag{3.4}$$

$$\hat{\varepsilon}(\boldsymbol{\phi}, \nu, \tilde{\Pi}_h) = \left(e^{\nu A(\boldsymbol{\phi})} - I\right)(\tilde{\Pi}_1 e^{i(U\boldsymbol{\phi})\cdot\boldsymbol{r}})_0. \tag{3.5}$$

Здесь и ниже индекс 1 означает подстановку $h = 1$, а индекс 0 – взятие блочной компоненты в блоке $\boldsymbol{\eta} = 0$.

Для $\boldsymbol{k} \in \mathbb{Z}^d$ определим волновой вектор

$$\boldsymbol{\alpha}(\boldsymbol{k}) = \frac{2\pi}{Nh} U\boldsymbol{k} = \frac{2\pi}{Nh} \sum_{j=1}^{d} k_j \boldsymbol{a}_j^*. \tag{3.6}$$

Для всех $\boldsymbol{k} \in \mathbb{Z}^d$ выполняется

$$\|\varepsilon_h(t, e^{i\boldsymbol{\alpha}(\boldsymbol{k})\cdot\boldsymbol{r}}, \tilde{\Pi}_h)\| = \|\hat{\varepsilon}(T^*\boldsymbol{\alpha}(\boldsymbol{k})h, t/h, \tilde{\Pi}_h)\|. \tag{3.7}$$

Следующие три утверждения позволяют переносить оценки функций (3.4) и (3.5) на произвольные достаточно гладкие периодические функции.

Утверждение 3.1. *Пусть* $\tilde{\Pi}_h^{(p,q)}$ *– оператор вида* (2.8)*. Пусть* $P_A \in \mathbb{N} \cup \{0\}$*. Тогда следующие утверждения эквивалентны:*

- *для некоторого* $C > 0$ *в некоторой окрестности* $\boldsymbol{\phi} = 0$ *справедливо* $\|\hat{\epsilon}(\boldsymbol{\phi}, \tilde{\Pi}_h^{(p,q)})\| \leqslant C|\boldsymbol{\phi}|^{P_A+1}$*;*
- *существуют такие* $C_1, C_2 \geqslant 0$*, что для всех* $v_0 \in H^r_{per}(\mathbb{R}^d)$*,* $r = \max\{P_A, q\} + 1$*, справедливо*

$$\|\epsilon_h(v_0, \tilde{\Pi}_h^{(p,q)})\| \leqslant C_1\|\nabla^{P_A+1} v_0\| h^{P_A} + C_2\|\nabla^r v_0\| h^{r-1}; \tag{3.8}$$

- *для всех мультииндексов* $\boldsymbol{m}$*, таких что* $|\boldsymbol{m}| \leqslant P_A$*, выполняется*

$$\epsilon_1(\boldsymbol{r}^{\boldsymbol{m}}/\boldsymbol{m}!, \tilde{\Pi}_1^{(p,q)}) = 0. \tag{3.9}$$

Утверждение 3.2. *Пусть* $\tilde{\Pi}_h^{(p,q)}$ *– некоторый оператор вида* (2.8)*. Пусть* $P, Q \in \mathbb{N}$*,* $Q \geqslant \max\{P, q\}$*. Тогда следующие утверждения эквивалентны:*

- *для некоторых* $C_1, C_2 \geqslant 0$ *в некоторой окрестности* $\boldsymbol{\phi} = 0$ *справедливо*

$$\|\hat{\varepsilon}(\boldsymbol{\phi}, \nu, \tilde{\Pi}_h^{(p,q)})\| \leqslant C_1|\boldsymbol{\phi}|^P + C_2\nu|\boldsymbol{\phi}|^{Q+1}; \tag{3.10}$$

- *для некоторых* $C_1, C_2 \geqslant 0$*, при любых начальных данных* $v_0 \in H^{Q+1}_{per}(\mathbb{R}^d)$ *для ошибки решения справедлива оценка* (2.13)*, то есть схема обладает порядком точности* P *и порядком точности в длительном счёте* Q*.*

Утверждение 3.3. *Пусть* $\mathring{\tilde{\Pi}}_h^{(p,q)}$ *– некоторый оператор вида* (2.10)*. Пусть* $P, Q \in \mathbb{N}$*,* $Q \geqslant \max\{P, q\}$*. Тогда следующие утверждения эквивалентны:*

- *для некоторых* $C_1, C_2 \geqslant 0$ *в некоторой окрестности* $\boldsymbol{\phi} = 0$ *справедливо*

$$\|\hat{\varepsilon}(\boldsymbol{\phi}, \nu, \mathring{\tilde{\Pi}}_h^{(p,q)})\| \leqslant C_1|\boldsymbol{\phi}|^P + C_2\nu|\boldsymbol{\phi}|^{Q+1}; \tag{3.11}$$

- *схема обладает порядком точности* P *и порядком точности в длительном счёте* Q *на* $C^{Q+1}_{per}(\mathbb{R}^d)$*.*

Пусть $\mathbb{R}^{n\times n}$ и $\mathbb{C}^{n\times n}$ – пространства действительно- и комплекснозначных матриц размера n. Через $\mathcal{A}(\ ,\mathbb{R}^{n\times n})$ будем обозначать множество функций из $\mathbb{C}^d$ в $\mathbb{C}^{n\times n}$, аналитических при $\boldsymbol{\phi} = 0$, таких что при $i\boldsymbol{\phi} \in \mathbb{R}^d$ выполняется $A(\boldsymbol{\phi}) \in \mathbb{R}^{n\times n}$. Проводимый ниже анализ опирается на следующие две теоремы, доказательство которых основано на теории возмущений (см. [7]) и приведено в [2].

Теорема 3.4. *Пусть* $A \in \mathcal{A}(\ ,\mathbb{R}^{n\times n})$*. Пусть* $A(0)$ *имеет нулевое собственное значение кратности* n_0*,* $0 < n_0 < n$*. Тогда в некоторой окрестности* $\boldsymbol{\phi} = 0$ *справедливо представление* $A(\boldsymbol{\phi}) = S(\boldsymbol{\phi})M(\boldsymbol{\phi})S^{-1}(\boldsymbol{\phi})$*, где* $S, M, S^{-1} \in \mathcal{A}(\ ,\mathbb{R}^{n\times n})$*,*

$$M(\boldsymbol{\phi}) = \begin{pmatrix} M^{(*)}(\boldsymbol{\phi}) & 0 \\ 0 & M^{(0)}(\boldsymbol{\phi}) \end{pmatrix}, \tag{3.12}$$

причём $M^{(0)}(0)$ *– нильпотентная матрица размера* n_0*, а* $M^{(*)}(0)$ *– невырожденная матрица размера* $n - n_0$*.*

Теорема 3.5. *Пусть* $d = 1$ *и* $A \in \mathcal{A}(\ ,\mathbb{R}^{n\times n})$*. Пусть существует* $K > 0$*, такое что для всех* $\phi \in \mathbb{R}$ *и всех* $\nu > 0$ *справедливо* $\|\exp(\nu A(\phi))\| \leqslant K$*. Тогда в окрестности* $\phi = 0$ *матрица* $A(\phi)$ *представима в блочно-диагональном виде*

$$A(\phi) = S(\phi)M(\phi)S^{-1}(\phi), \tag{3.13}$$

$$M(\phi) = \begin{pmatrix} M_0(\phi) & 0 & \dots & 0 & 0 \\ 0 & \phi M_1(\phi) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \phi^m M_m(\phi) & 0 \\ 0 & 0 & \dots & 0 & M_\infty(\phi) \equiv 0 \end{pmatrix}, \qquad (3.14)$$

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

4. Общие замечания

Утверждение 4.1. *Существование такого* K*, что* $\|e^{\nu A}\| \leqslant K$ *для всех* $\nu > 0$*, равносильно одновременному выполнению двух условий:*

- *все собственные числа* λ *матрицы* A *таковы, что* $\operatorname{Re}\lambda \leqslant 0$*;*
- *собственные числа* λ *матрицы* A*, такие что* $\operatorname{Re}\lambda = 0$*, полупростые.*

Действительно, представим матрицу A в виде $A = SJS^{-1}$, где J – её жорданова нормальная форма. Тогда $e^{\nu A} = Se^{\nu J}S^{-1}$. Доказываемое утверждение становится очевидным, если привести явный вид для $e^{\nu J}$.

Утверждение 4.2. *Пусть матрица* $A(\phi)$ *определена* (3.2)*. Тогда собственное значение* $\lambda = 0$ *у матрицы* $A(0)$ *полупростое, то есть не может быть жордановых клеток размера больше 1, соответствующих* $\lambda = 0$*.*

Действительно, в силу точности на константе $A(0)\mathfrak{e} = 0$, поэтому $\lambda = 0$ – собственное значение $A(0)$. Из устойчивости схемы следует выполнение (3.3), в частности, для $\phi = 0$. В силу утверждения 4.1 собственное значения $\lambda = 0$ полупростое.

Утверждение 4.3. *Пусть* $\breve{L}$ *– ограничение* L *на* V_{per}*. Пусть* $\operatorname{Ker}\breve{L} = \operatorname{span}\{\mathfrak{e}\}$*. Тогда собственное значение* $\lambda = 0$ *матрицы* $A(0)$ *простое.*

По определению (3.1)–(3.2) имеем $A(0) = -Z^{-1}(0)L(0)$. Рассмотрим произвольный вектор $w_0 \in \mathbb{C}^{M^0}$, такой что $w_0 \notin \operatorname{span}\{\mathfrak{e}\}$. Пусть $w \in V_{per}$ – вектор, все блочные компоненты которого одинаковы и равны w_0. Тогда $\breve{L}w \neq 0$. Но поскольку у $\breve{L}w$ все блочные компоненты одинаковы и равны $L(0)w_0$, справедливо $L(0)w_0 \neq 0$. Отсюда $A(0)w_0 = -Z^{-1}(0)L(0)w_0 \neq 0$. Значит, $w_0 = \mathfrak{e}$ является единственным (с точностью до множителя) решением $A(0)w_0 = 0$. Отсюда с учётом утверждения 4.2 следует, что $\lambda = 0$ является простым собственным значением $A(0)$.

Обратное утверждение, вообще говоря, неверно: если собственное значение $A(0)$ простое, то это не гарантирует отсутствия неконстантных решений уравнения $\breve{L}w = 0$. Действительно, рассмотрим одномерное уравнение переноса с $\omega = 1$ и схему с направленной разностью, выписанной независимо для двух переменных внутри блока. После этого в каждом втором блоке компоненты вектора внутри блока поменяем местами. Получаем схему с блочными компонентами

$$Z_0 = I, \quad L_0 = I, \quad L_{-1} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}.$$

У матрицы $L(0)$ есть два простых собственных значения: $\lambda = 0$ и $\lambda = 2$. Однако вектор $w \in V_{per}$, имеющий период $N = 2$, с компонентами $w_\eta = (-1)^\eta (1, -1)^T$ лежит в ядре $\breve{L}$, так же как и $\mathfrak{e}$.

Утверждение 4.4. *Пусть $A \neq 0$, $\|A\| \leqslant 1$ и*

$$f(A) = \sum_{k=1}^{\infty} \frac{A^{k-1}}{k!}. \tag{4.1}$$

Тогда $\|(f(A))^{-1}\| \leqslant 4$. Если дополнительно A не вырождена, то

$$\|(e^A - I)^{-1}\| \leqslant 4\|A^{-1}\|. \tag{4.2}$$

Доказательство см. в [3].

По определению $\hat{\varepsilon}(\boldsymbol{\phi}, \nu, \tilde{\Pi}_h)$ имеет вид (3.5). В силу теоремы 3.4 в некоторой окрестности $\boldsymbol{\phi} = 0$ справедливо представление $A(\boldsymbol{\phi}) = S(\boldsymbol{\phi})M(\boldsymbol{\phi})S^{-1}(\boldsymbol{\phi})$, где $M(\boldsymbol{\phi})$ имеет вид (3.12), откуда

$$\hat{\varepsilon}(\boldsymbol{\phi}, \nu, \tilde{\Pi}_h) = S(\boldsymbol{\phi}) \begin{bmatrix} e^{M^{(*)}(\boldsymbol{\phi})\nu} - I & 0 \\ 0 & e^{M^{(0)}(\boldsymbol{\phi})\nu} - I \end{bmatrix} S^{-1}(\boldsymbol{\phi})(\tilde{\Pi}_1 e^{i(U\boldsymbol{\phi})\cdot\boldsymbol{r}})_0, \tag{4.3}$$

причём $M^{(*)}(\boldsymbol{\phi})$ и $M^{(0)}(\boldsymbol{\phi})$ – аналитические функции $\boldsymbol{\phi}$, $M^{(*)}(0)$ невырожденная, а $M^{(0)}(0)$ – нильпотентная.

Введём векторы $v^{(*)}(\boldsymbol{\phi}, \tilde{\Pi}_h)$ и $v^{(0)}(\boldsymbol{\phi}, \tilde{\Pi}_h)$, размерности которых соответствуют блокам в представлении (4.3), так чтобы

$$\begin{pmatrix} v^{(*)}(\boldsymbol{\phi}, \tilde{\Pi}_h) \\ v^{(0)}(\boldsymbol{\phi}, \tilde{\Pi}_h) \end{pmatrix} = S^{-1}(\boldsymbol{\phi})(\tilde{\Pi}_1 e^{i(U\boldsymbol{\phi})\cdot\boldsymbol{r}})_0. \tag{4.4}$$

Обозначим

$$\hat{\varepsilon}^{(1)}(\boldsymbol{\phi},\nu,\tilde{\Pi}_h) = S(\boldsymbol{\phi}) \begin{pmatrix} \left[\exp\left(M^{(*)}(\boldsymbol{\phi})\nu\right) - I\right] v^{(*)}(\boldsymbol{\phi},\tilde{\Pi}_h) \\ 0 \end{pmatrix}, \qquad (4.5)$$

$$\hat{\varepsilon}^{(2)}(\boldsymbol{\phi},\nu,\tilde{\Pi}_h) = S(\boldsymbol{\phi}) \begin{pmatrix} 0 \\ \left[\exp\left(M^{(0)}(\boldsymbol{\phi})\nu\right) - I\right] v^{(0)}(\boldsymbol{\phi},\tilde{\Pi}_h) \end{pmatrix}.$$

Очевидно, выполняется

$$\hat{\varepsilon}(\boldsymbol{\phi},\nu,\tilde{\Pi}_h) = \hat{\varepsilon}^{(1)}(\boldsymbol{\phi},\nu,\tilde{\Pi}_h) + \hat{\varepsilon}^{(2)}(\boldsymbol{\phi},\nu,\tilde{\Pi}_h). \qquad (4.6)$$

Спектр матрицы $M^{(0)}(0)$ состоит из одной точки $\lambda = 0$, откуда в силу утверждения 4.2 следует, что $M^{(0)}(0) = 0$, и $M^{(0)}(\boldsymbol{\phi}) = \sum_j M_j^{(0)}(\boldsymbol{\phi})\phi_j$. Поэтому

$$M^{(0)}(T^*\boldsymbol{\alpha}h)\frac{t}{h} = \sum_{j=1}^{d} M_j^{(0)}(T^*\boldsymbol{\alpha}h)(T^*\boldsymbol{\alpha})_j t$$

является аналитической функцией t, h и $\boldsymbol{\alpha}$. Следовательно, таковой является и $\hat{\varepsilon}^{(2)}(T^*\boldsymbol{\alpha}h, t/h, \tilde{\Pi}_h)$.

В частности, если $L(0) = 0$, то $A(0) = 0$ и $\hat{\varepsilon}^{(1)}(\boldsymbol{\phi},\nu,\tilde{\Pi}_h) = 0$. Таким образом, имеем следующий результат.

Утверждение 4.5. *Если $L(0) = 0$, то $\hat{\varepsilon}(\boldsymbol{\alpha}h, t/h, \mathring{\Pi}_h)$ является аналитической функцией t, h и $\boldsymbol{\alpha}$.*

Утверждение 4.6. *Пусть схема* (2.5) *обладает оценкой* (2.13) *в смысле Π_h. Тогда в некоторой окрестности $\boldsymbol{\phi} = 0$ при всех $\nu > 0$ выполняется*

$$\|v^{(*)}(\boldsymbol{\phi},\Pi_h)\| \leqslant C|\boldsymbol{\phi}|^P, \quad \|\hat{\varepsilon}^{(1)}(\boldsymbol{\phi},\nu,\Pi_h)\| \leqslant (K+1)C|\boldsymbol{\phi}|^P. \qquad (4.7)$$

По утверждению 3.2 из оценки (2.13) следует (3.10). В силу аналитичности матрицы $S^{-1}(\boldsymbol{\phi})$ в некоторой окрестности $\boldsymbol{\phi} = 0$, отсюда получаем

$$\|\hat{\varepsilon}^{(m)}(\boldsymbol{\phi},\nu,\Pi_h)\| \leqslant C_1|\boldsymbol{\phi}|^P + C_2\nu|\boldsymbol{\phi}|^{Q+1},$$

для $m = 1, 2$. То есть для

$$E = \left[\exp\left(M^{(*)}(\boldsymbol{\phi})\nu\right) - I\right] v^{(*)}(\boldsymbol{\phi})$$

имеем

$$\|E\| \leqslant C_1'|\boldsymbol{\phi}|^P + C_2'\nu|\boldsymbol{\phi}|^{Q+1}.$$

Матрица $M^{(*)}(0)$ невырождена, поэтому для достаточно малых $\boldsymbol{\phi}$ можно считать, что $\|(M^{(*)}(\boldsymbol{\phi}))^{-1}\| \leqslant 2\|(M^{(*)}(0))^{-1}\|$ и $\|M^{(*)}(\boldsymbol{\phi})\| \leqslant 2\|M^{(*)}(0)\|$. Положим $\nu = 1/(2\|M^{(*)}(0)\|)$. Тогда в силу (4.2) имеем

$$\|v^{(*)}(\boldsymbol{\phi})\| \leqslant \left\| \left[\exp\left(M^{(*)}(\boldsymbol{\phi})\nu \right) - I \right]^{-1} \right\| \|E\| \leqslant$$
$$\leqslant 8\|(M^{(*)}(0))^{-1}\| \left(C_1 \, |\boldsymbol{\phi}|^P + C_2 \, \frac{1}{2\|M^{(*)}(0)\|} |\boldsymbol{\phi}|^{Q+1} \right).$$

Отсюда получаем первое неравенство в (4.7). Второе неравенство следует из него в силу устойчивости.

Утверждение 4.7. *Пусть схема* (2.5) *обладает оценкой* (2.13) *в смысле* Π_h. *Тогда в некоторой окрестности* $\boldsymbol{\phi} = 0$ *выполняется*

$$\|M^{(0)}(\boldsymbol{\phi}) v^{(0)}(\boldsymbol{\phi}, \Pi_h)\| \leqslant C|\boldsymbol{\phi}|^{P+1}. \tag{4.8}$$

Аналогично предыдущему утверждению, для

$$E = \left[\exp\left(M^{(0)}(\boldsymbol{\phi})\nu \right) - I \right] v^{(0)}(\boldsymbol{\phi}, \Pi_h)$$

имеем

$$\|E\| \leqslant C_1'|\boldsymbol{\phi}|^P + C_2'\nu|\boldsymbol{\phi}|^{Q+1}.$$

Рассмотрим некоторое $\boldsymbol{\phi}$. Если $M^{(0)}(\boldsymbol{\phi}) = 0$, то искомое утверждение очевидно. Далее будем считать, что $M^{(0)}(\boldsymbol{\phi}) \neq 0$. Преобразуем E следующим образом:

$$E = f\left(M^{(0)}(\boldsymbol{\phi})\nu \right) M^{(0)}(\boldsymbol{\phi})\nu v^{(0)}(\boldsymbol{\phi}, \Pi_h),$$

где f определено (4.1). Положим $\nu = 1/\|M^{(0)}(\boldsymbol{\phi})\|$. Тогда в силу утверждения 4.4 имеем

$$\left\| M^{(0)}(\boldsymbol{\phi})\nu v^{(0)}(\boldsymbol{\phi}, \Pi_h) \right\| = \left\| \left[f(M^{(0)}(\boldsymbol{\phi})\nu) \right]^{-1} E \right\| \leqslant 4\|E\|$$

и, следовательно,

$$\left\| M^{(0)}(\boldsymbol{\phi}) v^{(0)}(\boldsymbol{\phi}, \Pi_h) \right\| \leqslant C_1' \frac{1}{\nu} |\boldsymbol{\phi}|^P + C_2'|\boldsymbol{\phi}|^{Q+1}.$$

Поскольку $1/\nu = \|M^{(0)}(\boldsymbol{\phi})\|$, а $M^{(0)}(0) = 0$, имеем $1/\nu \leqslant \tilde{C}|\boldsymbol{\phi}|$. Таким образом, получаем (4.8).

Утверждение 4.8. *Пусть схема* (2.5) *обладает оценкой* (2.13) *в смысле* Π_h. *Тогда для любого* $q \in \mathbb{N} \cup \{0\}$ *существует оператор* $\tilde{\Pi}_h^{(P,q)}$ *вида* (2.8)*, такой что величины* $v^{(*)}(\boldsymbol{\phi}, \tilde{\Pi}_h^{(P,q)})$ *и* $v^{(0)}(\boldsymbol{\phi}, \tilde{\Pi}_h^{(P,q)})$*, определённые* (4.4)*, удовлетворяют оценкам*

$$\begin{gathered} \|v^{(*)}(\boldsymbol{\phi}, \tilde{\Pi}_h^{(P,q)})\| \leqslant C|\boldsymbol{\phi}|^{q+1}, \\ \|v^{(0)}(\boldsymbol{\phi}, \tilde{\Pi}_h^{(P,q)}) - v^{(0)}(\boldsymbol{\phi}, \Pi_h)\| \leqslant C|\boldsymbol{\phi}|^{q+1}. \end{gathered} \tag{4.9}$$

При $q < P$ выполняется $\tilde{\Pi}_h^{(P,q)} \equiv \Pi_h$, и искомая оценка следует из утверждения 4.6. Пусть $q \geqslant P$. Введём в окрестности $\boldsymbol{\phi} = 0$ аналитическую функцию

$$V(\boldsymbol{\phi}) = S(\boldsymbol{\phi}) \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} S^{-1}(\boldsymbol{\phi}) (\Pi_1 e^{i(U\boldsymbol{\phi})\cdot \boldsymbol{r}})_0. \tag{4.10}$$

По построению

$$\|V(\boldsymbol{\phi}) - (\Pi_1 e^{i(U\boldsymbol{\phi})\cdot \boldsymbol{r}})_0\| \leqslant C\|v^{(*)}(\boldsymbol{\phi})\| \leqslant \tilde{C}|\boldsymbol{\phi}|^P.$$

Определим диагональную матрицу $\mathfrak{C}(\boldsymbol{\phi})$ равенством

$$V(\boldsymbol{\phi}) - (\Pi_1 e^{i(U\boldsymbol{\phi})\cdot \boldsymbol{r}})_0 = \mathfrak{C}(\boldsymbol{\phi}) (\mathcal{P}_1 e^{i(U\boldsymbol{\phi})\cdot \boldsymbol{r}})_0 \tag{4.11}$$

и введём диагональные матрицы

$$\mathfrak{C}^{(\boldsymbol{m})} = \frac{1}{\boldsymbol{m}!} \frac{1}{i^{|\boldsymbol{m}|}} \left. \frac{\partial^{|\boldsymbol{m}|} \mathfrak{C}(T^* \boldsymbol{\psi})}{\partial \psi_1^{m_1} \dots \partial \psi_d^{m_d}} \right|_{\boldsymbol{\psi}=0}. \tag{4.12}$$

Покажем, что $\mathfrak{C}^{(\boldsymbol{m})} = 0$ при $|\boldsymbol{m}| < P$. Поскольку матрица $\mathfrak{C}(\boldsymbol{\phi})$ диагональная, а $(\mathcal{P}_1 e^{i(U\boldsymbol{\phi})\cdot \boldsymbol{r}})_0$ стремится к $\mathfrak{e}$ при $\boldsymbol{\phi} \to 0$, порядок малости $\mathfrak{C}(\boldsymbol{\phi})$ по $\boldsymbol{\phi}$ совпадает с порядком малости правой части (4.11). Ввиду невырожденности $S(0)$ он совпадает с порядком малости вектора $v^{(*)}(\boldsymbol{\phi})$, который имеет величину не больше $O(|\boldsymbol{\phi}|^P)$. Следовательно, $\mathfrak{C}(\boldsymbol{\phi}) = O(|\boldsymbol{\phi}|^P)$, а это и означает, что её первые $P-1$ производные нулевые.

Введём оператор $\tilde{\Pi}_h^{(P,q)}$ равенством (2.8). Вычисляя $(\tilde{\Pi}_1^{(P,q)} e^{i(U\boldsymbol{\phi})\cdot \boldsymbol{r}})_0$ напрямую из определения (2.8), получаем

$$(\tilde{\Pi}_1^{(P,q)} e^{i(U\boldsymbol{\phi})\cdot \boldsymbol{r}})_0 = (\Pi_1 e^{i(U\boldsymbol{\phi})\cdot \boldsymbol{r}})_0 + \left[\sum_{\boldsymbol{m}: P \leqslant |\boldsymbol{m}| \leqslant q} \mathfrak{C}^{(\boldsymbol{m})} (iU\boldsymbol{\phi})^{\boldsymbol{m}} \right] (\mathcal{P}_1 e^{i(U\boldsymbol{\phi})\cdot \boldsymbol{r}})_0.$$

Так как $\mathfrak{C}^{(\boldsymbol{m})} = 0$ при $|\boldsymbol{m}| < P$, выражение в квадратных скобках совпадает с тейлоровским разложением функции $\mathfrak{C}(T^* \boldsymbol{\psi})$ по степеням $\boldsymbol{\psi}$ до порядка q (см. (4.12)) при $\boldsymbol{\psi} = U\boldsymbol{\phi}$. Так как $T^* U\boldsymbol{\phi} = \boldsymbol{\phi}$, получаем

$$(\tilde{\Pi}_1^{(P,q)} e^{i(U\boldsymbol{\phi})\cdot \boldsymbol{r}})_0 = (\Pi_1 e^{i(U\boldsymbol{\phi})\cdot \boldsymbol{r}})_0 + \mathfrak{C}(\boldsymbol{\phi}) (\mathcal{P}_1 e^{i(U\boldsymbol{\phi})\cdot \boldsymbol{r}})_0 + O(|\boldsymbol{\phi}|^{q+1}).$$

В силу (4.11) отсюда следует

$$(\tilde{\Pi}_1^{(P,q)} e^{i(U\boldsymbol{\phi})\cdot\boldsymbol{r}})_0 = V(\boldsymbol{\phi}) + O(|\boldsymbol{\phi}|^{q+1}).$$

Домножая это равенство слева на $S^{-1}(\boldsymbol{\phi})$, получаем искомые оценки.

5. Определение формального порядка

В [3] был предложен алгоритм установления оптимальных значений формального порядка точности P и порядка точности в длительном счёте Q в одномерном случае. Покажем, что его часть, связанная с определением P, переносится на многомерный случай.

Алгоритм 1 (определения формального порядка точности)**.**

1. *Определяем* $P_A = \max\{m \in \mathbb{N} \cup \{0\} \,:\, (\epsilon_1(\boldsymbol{r}^{\boldsymbol{m}}, \Pi_1))_0 = 0 \;\forall \boldsymbol{m} : |\boldsymbol{m}| \leqslant m\}$.
2. *Вычисляем* $f^{\boldsymbol{m}} = -(\epsilon_1(\boldsymbol{r}^{\boldsymbol{m}}/\boldsymbol{m}!, \Pi_1))_0$ *для всех* $\boldsymbol{m} : |\boldsymbol{m}| = P_A + 1$.
3. *Если для всех* $\boldsymbol{m} : |\boldsymbol{m}| = P_A + 1$ *выполняется* $f^{\boldsymbol{m}} \in \mathrm{Im}L(0)$*, то полагаем* $P' = P_A + 1$*, иначе полагаем* $P' = P_A$.

Утверждение 5.1. *Значение P', найденное алгоритмом 1, совпадает с оптимальным значением порядка точности P.*

Если $P' = P_A + 1$, то системы $L(0)\mathfrak{C}^{(\boldsymbol{m})}\mathfrak{e} = f^{(\boldsymbol{m})}$, $|\boldsymbol{m}| = P_A + 1$, совместны. Определим отображение $\tilde{\Pi}_h^{(P_A+1,P_A+1)}$, в которое подставим решения этих систем. В силу устойчивости и неравенства треугольника имеем оценку (2.13) при $P = Q = P_A + 1$ (доказательство повторяет одномерный случай).

Если $P = P_A + 1$, то по утверждению 4.8 найдутся такие коэффициенты $\mathfrak{C}^{(\boldsymbol{m})}$, $|\boldsymbol{m}| = P$, что для отображения $\tilde{\Pi}_h^{(P,P)}$, определённого (2.8), будут выполняться неравенства (4.9) для $q = P$. Первое из них даёт

$$\|M^{(*)}(\boldsymbol{\phi})v^{(*)}(\boldsymbol{\phi}, \tilde{\Pi}_h^{(P,P)})\| \leqslant C|\boldsymbol{\phi}|^{P+1}.$$

Второе с учётом утверждения 4.7 даёт

$$\|M^{(0)}(\boldsymbol{\phi})v^{(0)}(\boldsymbol{\phi}, \tilde{\Pi}_h^{(P,P)})\| \leqslant C|\boldsymbol{\phi}|^{P+1}.$$

Таким образом, имеем

$$\|A(\boldsymbol{\phi})(\tilde{\Pi}_h^{(P,P)} e^{i(U\boldsymbol{\phi})\cdot\boldsymbol{r}})_0\| \leqslant C|\boldsymbol{\phi}|^{P+1},$$

и по утверждению 3.2, в смысле $\tilde{\Pi}_h^{(P,P)}$ схема обладает порядком аппроксимации P. Следовательно, коэффициенты $\mathfrak{C}^{(\boldsymbol{m})}$, $|\boldsymbol{m}| = P$, этого оператора удовлетворяют системе

$$L(0)\mathfrak{C}^{(\boldsymbol{m})}\mathfrak{e} = -(\epsilon_1(\boldsymbol{r}^{\boldsymbol{m}}/\boldsymbol{m}!, \tilde{\Pi}_1))_0,$$

поэтому правая часть этого равенства для всех $\boldsymbol{m}$, $|\boldsymbol{m}| = P$, лежит в образе $L(0)$, и алгоритм выдаст значение $P' = P_A + 1$.

Следствие 5.2. *Если $L(0) = 0$, то порядок точности схемы не может превосходить её порядок аппроксимации.*

Действительно, если схема обладает порядком аппроксимации P_A, то для какого-то $\boldsymbol{m}$: $|\boldsymbol{m}| = P_A$ выполняется $f^{\boldsymbol{m}} \neq 0$. Поскольку $L(0) = 0$, то $f^{\boldsymbol{m}} \notin \mathrm{Im}L(0)$, и алгоритм выдаст значение $P' = P_A$. По утверждению получаем $P = P_A$.

6. Существование отображения

Перейдём к рассмотрению вопроса, переносятся ли на многомерный случай результаты, полученные в [3], применительно к порядку точности в длительном счёте.

Утверждение 6.1. *Пусть $A \in \mathbb{C}^{n\times n}$ и $v \in \mathbb{C}^n$. Пусть при всех $\nu \geqslant 0$ справедливы неравенства $\|e^{\nu A}\| \leqslant K$ и*

$$\|(e^{\nu A} - I)v\| \leqslant (\tilde{C}_1 + \tilde{C}_2\nu)\|v\|. \tag{6.1}$$

Тогда для $w = (A^*A + \varepsilon^2)^{-1}\varepsilon^2 v$, где $\varepsilon = \tilde{C}_2/\tilde{C}_1$ и при $\tilde{C}_1 = 0$ или $\tilde{C}_2 = 0$ под w понимается соответствующее предельное значение, справедливы оценки

$$\|v - w\| \leqslant \delta\tilde{C}_1\|v\|, \quad \|Aw\| \leqslant \delta\tilde{C}_2\|v\|, \tag{6.2}$$

где δ зависит только от n и K.

Это утверждение доказано в [4].

Теорема 6.2. *Пусть Π_h определён* (2.7)*. Пусть схема* (2.5) *в смысле Π_h на $H^{Q+1}_{per}(\mathbb{R}^d)$ обладает формальным порядком точности P и порядком точности в длительном счёте Q. Тогда существует оператор $\tilde{\Pi}_h : L_{2,per}(\mathbb{R}^d) \to V_{per}$, такой что для некоторого C выполняются условия*

$$\|\tilde{\Pi}_h f - \Pi_h f\| \leqslant Ch^P\|f\|, \tag{6.3}$$

$$\|\epsilon_h(f, \tilde{\Pi}_h)\| \leqslant Ch^Q\|f\| \tag{6.4}$$

для всех h и $f \in H^{Q+1}_{per}(\mathbb{R}^d)$.

Обозначим $v(\boldsymbol{\phi}) = (\Pi_1 e^{i(U\boldsymbol{\phi})\cdot\boldsymbol{r}})_0$. Поскольку норма экспоненты равна единице, имеем $\|v(\boldsymbol{\phi})\| \leqslant \|\Pi_1\|$. В силу утверждения (3.10) имеем

$$\left\| \left(e^{\nu A(\boldsymbol{\phi})} - I\right) v(\boldsymbol{\phi}) \right\| \leqslant C_1|\boldsymbol{\phi}|^P + C_2\nu|\boldsymbol{\phi}|^{Q+1}.$$

Это неравенство совпадает с (6.1) при подстановке $A = A(\boldsymbol{\phi})$, $v = v(\boldsymbol{\phi})$, $\tilde{C}_1 = C_1|\boldsymbol{\phi}|^P$, $\tilde{C}_2 = C_2|\boldsymbol{\phi}|^{Q+1}$. В силу утверждения 6.1 существует $w(\boldsymbol{\phi})$, такой что выполняется

$$\|v(\boldsymbol{\phi}) - w(\boldsymbol{\phi})\| \leqslant \delta C_1|\boldsymbol{\phi}|^P\|\Pi_1\|, \quad \|A(\boldsymbol{\phi})w(\boldsymbol{\phi})\| \leqslant \delta C_2|\boldsymbol{\phi}|^{Q+1}\|\Pi_1\|.$$

Введём $\omega'(\boldsymbol{\phi}) = \omega(\boldsymbol{\phi})$ при $|\boldsymbol{\phi}| \leqslant 1$ и $\omega'(\boldsymbol{\phi}) = 0$ при $\|\boldsymbol{\phi}\| > 1$. Тогда получаем

$$\|v(\boldsymbol{\phi}) - w'(\boldsymbol{\phi})\| \leqslant \max\{1, \delta C_1\}|\boldsymbol{\phi}|^P\|\Pi_1\|, \quad \|A(\boldsymbol{\phi})w'(\boldsymbol{\phi})\| \leqslant \delta C_2|\boldsymbol{\phi}|^{Q+1}\|\Pi_1\|.$$

Определим вначале действие $\tilde{\Pi}_h$ на функцию вида $f(\boldsymbol{r}) = \exp(i\boldsymbol{\alpha}(\boldsymbol{k}) \cdot \boldsymbol{r})$, где $\boldsymbol{k} \in \mathbb{Z}^d$, а $\boldsymbol{\alpha}(\boldsymbol{k})$ определена (3.6). Положим

$$\left(\tilde{\Pi}_h f\right)_{\boldsymbol{\eta}} = \exp(i\boldsymbol{\phi} \cdot \boldsymbol{\eta}/h)w'(\boldsymbol{\phi}). \tag{6.5}$$

Тогда будет обеспечиваться

$$\|\tilde{\Pi}_h f - \Pi_h f\| = \|(\tilde{\Pi}_h f)_0 - (\Pi_h f)_0\| = \|w'(\boldsymbol{\phi}) - v(\boldsymbol{\phi})\| \leqslant \max\{1, \delta C_1\}|\boldsymbol{\phi}|^P\|\Pi_1\|, \tag{6.6}$$

$$\|\epsilon_h(f, \tilde{\Pi}_h)\| = \|(\epsilon_h(f, \tilde{\Pi}_h))_0\| = \frac{1}{h}\left\|\sum_{\boldsymbol{\eta}\in\mathcal{S}}(-ih\boldsymbol{\omega} \cdot \boldsymbol{\alpha}(\boldsymbol{k})Z_{\boldsymbol{\eta}} + L_{\boldsymbol{\eta}})(\tilde{\Pi}_h f)_{\boldsymbol{\eta}}\right\| = \tag{6.7}$$

$$= \frac{1}{h}\|Z(\boldsymbol{\phi})Aw'(\boldsymbol{\phi})\| \leqslant \frac{1}{h}\|Z(\boldsymbol{\phi})\|\delta C_2|T^*\boldsymbol{\alpha}(\boldsymbol{k})h|^{Q+1} \leqslant \tilde{C}h^Q|\boldsymbol{\alpha}(\boldsymbol{k})|^{Q+1}.$$

Далее, если $f \in L_{2,per}(\mathbb{R}^d)$, то она может быть представлена своим сходящимся в L_2 рядом Фурье:

$$f = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} U_{\boldsymbol{k}} \exp(i\boldsymbol{\alpha}(\boldsymbol{k}) \cdot \boldsymbol{r}),$$

причём $\|f\|^2 = \sum |U_{\boldsymbol{k}}|^2|\boldsymbol{k}|^2$. Действие $\tilde{\Pi}_h$ на каждую из гармоник определено (6.5), а действие $\tilde{\Pi}_h$ на f определим исходя из линейности $\tilde{\Pi}_h$. Покажем, что $\tilde{\Pi}_h$ является ограниченным. Действительно, в силу ортогональности образов $\exp(i\boldsymbol{\alpha}(\boldsymbol{k}) \cdot \boldsymbol{r})$ для разных $\boldsymbol{k} \in I_N^d$ под действием $\tilde{\Pi}_h$ имеем

$$\|\tilde{\Pi}_h f\|^2 = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} |U_{\boldsymbol{k}}|^2 \left|w'(2\pi\boldsymbol{k}/N)\right|^2 \leqslant \|f\|^2 \sup_{\boldsymbol{\phi}\in\mathbb{R}^d} \|w(\boldsymbol{\phi})\|^2 \leqslant$$

$$\leqslant \|f\|^2 \sup_{\boldsymbol{\phi}\in\mathbb{R}^d} (\|v(\boldsymbol{\phi})\| + \|w'(\boldsymbol{\phi}) - v(\boldsymbol{\phi})\|)^2 \leqslant \|f\|^2(2 + \delta C_1)^2\|\Pi_h\|^2.$$

По теореме 15 в [1] из ограниченности $\tilde{\Pi}_h$ из неравенств (6.6) и (6.7) следует (6.3) и (6.4) для произвольной функции $f \in H_{per}^{Q+1}(\mathbb{R}^d)$.

В одномерном случае $w(\phi)$, даваемое утверждением 6.1, является аналитической функцией ϕ. Эту функцию можно приблизить суммой первых $Q+1$ членов ряда Тейлора. Зная коэффициенты при ϕ^k, $k = P, \ldots, Q$, легко построить такое отображение $\tilde{\Pi}_h^{(P,Q)}$ вида (2.10), в смысле которого будет иметь место Q-й порядок аппроксимации (подробно построение этого отображения описано в [1]). В многомерном случае функция $w(\phi)$, как правило, не является аналитической, и оператор $\tilde{\Pi}_h^{(P,Q)}$ вида (2.10), доставляющий Q-й порядок аппроксимации, вообще говоря, не существует. Соответствующий пример будет приведён ниже, а пока рассмотрим два частных случая, когда оператор $\tilde{\Pi}_h^{(P,Q)}$ искомого вида существует и, следовательно, оптимальное значение порядка точности в длительном счёте удаётся установить алгоритмически.

7. Простой случай

Начнём со случая, когда собственное значение $\lambda = 0$ матрицы $A(0)$ является простым. Будем называть такой случай *простым*.

Утверждение 7.1. *Пусть схема* (2.5) *обладает оценкой* (2.13) *в смысле* Π_h. *Пусть размер матрицы* $M^{(0)}(\boldsymbol{\phi})$ *равен 1. Тогда в некоторой окрестности* $\boldsymbol{\phi} = 0$ *выполняется*

$$\|M^{(0)}(\boldsymbol{\phi})\| \leqslant C|\boldsymbol{\phi}|^{Q+1}. \tag{7.1}$$

Обозначим единственный элемент матрицы $M^{(0)}(\boldsymbol{\phi})$ через $\lambda(\boldsymbol{\phi})$. Он является аналитической функцией $\boldsymbol{\phi}$. Пусть j_0 – минимальный порядок малости по $\boldsymbol{\phi}$ членов в разложении $\lambda(\boldsymbol{\phi})$ по степеням $\phi_1, \ldots, \phi_d$. Если $\lambda(\boldsymbol{\phi}) \equiv 0$, будем считать $j_0 = \infty$.

В некоторой окрестности $\boldsymbol{\phi} = 0$ имеем $\|v^{(*)}(\boldsymbol{\phi})\| \leqslant C|\boldsymbol{\phi}|^P$. Следовательно, $v^{(0)}(0) \neq 0$, иначе было бы $v(0) = 0$, откуда из (4.4) следовало бы $(\Pi_1 1)_0 = 0$. Напомним, что $v^{(0)}(\boldsymbol{\phi})$ – это однокомпонентная величина. Отсюда для достаточно малых $|\boldsymbol{\phi}|$ можно записать

$$\left|e^{\nu\lambda(\boldsymbol{\phi})} - 1\right| \leqslant \frac{2}{|v^{(0)}(0)|}(C_1|\boldsymbol{\phi}|^P + C_2|\boldsymbol{\phi}|^{Q+1}(\nu+1)).$$

Положим $\nu = |\boldsymbol{\phi}|^{-Q-1/2}$, тогда правая часть будет стремиться к нулю при $\boldsymbol{\phi} \to 0$. Следовательно, $\lambda(\boldsymbol{\phi})|\boldsymbol{\phi}|^{-Q-1/2} \to 0$. Но поскольку $\lambda(\boldsymbol{\phi})$ – аналитическая функция $\boldsymbol{\phi}$ при $\boldsymbol{\phi} = 0$, имеем $j_0 \geqslant Q+1$, то есть $\lambda(\boldsymbol{\phi}) \leqslant C|\boldsymbol{\phi}|^{Q+1}$.

Утверждение 7.2. *Пусть* Π_h *и* $\mathcal{P}_h$ *определены* (2.7). *Пусть схема* (2.5) *в смысле* Π_h *на* $H^{Q+1}_{per}(\mathbb{R}^d)$ *обладает формальным порядком точности* P *и порядком*

точности в длительном счёте Q. Пусть собственное значение $\lambda = 0$ матрицы $A(0)$, определённой (3.2)*, является простым. Тогда существуют такие диагональные матрицы $\mathfrak{C}^{(\boldsymbol{m})} \in \mathbb{R}^{M^0}$, что схема обладает порядком аппроксимации Q в смысле модифицированного отображения $\tilde{\Pi}_h^{(P,Q)} : H_{per}^{Q+1}(\mathbb{R}^d) \to V_{per}$, определяемого* (2.8)*.*

Будем пользоваться введёнными выше обозначениями. Поскольку по условию $\lambda = 0$ является простым, то размер блока $M^{(0)}(\boldsymbol{\phi})$ равен 1. Обозначим единственный элемент этой матрицы через $\lambda(\boldsymbol{\phi})$. По утверждению 4.8 существуют такие коэффициенты $\mathfrak{C}^{(\boldsymbol{m})}$, $P \leqslant |\boldsymbol{m}| \leqslant Q$, что выполняется неравенство

$$\|v^{(*)}(\boldsymbol{\phi}, \tilde{\Pi}_h^{(P,Q)})\| \leqslant C|\boldsymbol{\phi}|^{Q+1}.$$

Отсюда получаем

$$\|M^{(*)}(\boldsymbol{\phi})v^{(*)}(\boldsymbol{\phi}, \tilde{\Pi}_h^{(P,Q)})\| \leqslant C|\boldsymbol{\phi}|^{Q+1}.$$

В силу утверждения 7.1 имеем $\|M^{(0)}(\boldsymbol{\phi})\| \leqslant C|\boldsymbol{\phi}|^{Q+1}$, откуда

$$\|M^{(0)}(\boldsymbol{\phi})v^{(0)}(\boldsymbol{\phi}, \tilde{\Pi}_h^{(P,Q)})\| \leqslant C|\boldsymbol{\phi}|^{Q+1}.$$

Таким образом,

$$\|A(\boldsymbol{\phi})(\tilde{\Pi}_h^{(P,Q)} e^{i(U\boldsymbol{\phi})\cdot\boldsymbol{r}})_0\| \leqslant C|\boldsymbol{\phi}|^{Q+1},$$

и по утверждению 3.2 в смысле $\tilde{\Pi}_h^{(P,Q)}$ схема обладает порядком аппроксимации Q.

Утверждение 7.3. *Пусть $\mathring{\Pi}_h$ и $\mathring{\mathcal{P}}_h$ – операторы вида* (2.9)*. Пусть схема* (2.5) *в смысле $\mathring{\Pi}_h$ на $C_{per}^{Q+1}(\mathbb{R}^d)$ обладает формальным порядком точности P и порядком точности в длительном счёте Q. Пусть собственное значение $\lambda = 0$ матрицы $A(0)$, определённой* (3.2)*, является простым. Тогда существуют такие диагональные матрицы $\mathfrak{C}^{(\boldsymbol{m})} \in \mathbb{R}^{M^0}$, что схема обладает порядком аппроксимации Q в смысле модифицированного отображения $\mathring{\tilde{\Pi}}_h : C_{per}^{Q+1}(\mathbb{R}^d) \to V_{per}$, определяемого* (2.10)*.*

Это утверждение следует из предыдущего в силу возможности приближения операторов $\mathring{\Pi}_h$ и $\mathring{\mathcal{P}}_h$ операторами вида (2.7) с любым наперёд заданным порядком точности. Строгое доказательство повторяет доказательство для одномерного случая, проведённое в [3].

Утверждения 7.2 и 7.3 позволяют обобщить алгоритм 2 определения оптимальных значений P и Q, приведённый в [3], на многомерный случай. Уравнения на нахождение коэффициентов $\mathfrak{C}^{(\boldsymbol{m})}$ записываются в виде

$$L(0)\mathfrak{C}^{(\boldsymbol{m})}\mathfrak{e} = f^{m}, \quad f^{m} = -\left(\epsilon_1\left(\frac{\boldsymbol{r}^{\boldsymbol{m}}}{\boldsymbol{m}!}, \tilde{\Pi}_1^{(p,|\boldsymbol{m}|-1)}\right)\right)_0, \tag{7.2}$$

где $\tilde{\Pi}_h^{(p,m-1)}$ – оператор, определённый равенством

$$\left(\tilde{\Pi}_h^{(p,m-1)} f\right)_{\boldsymbol{\eta}} = (\Pi_h f)_{\boldsymbol{\eta}} + \sum_{p \leqslant |\boldsymbol{n}| \leqslant m-1} h^{|\boldsymbol{n}|}\mathfrak{C}^{(\boldsymbol{n})}\left(\mathcal{P}_h(D^{\boldsymbol{n}} f)\right)_{\boldsymbol{\eta}}. \tag{7.3}$$

Алгоритм 2 (определения формального порядка точности и порядка точности в длительном счёте для простого случая)**.**

1. *Определяем* $P_A = \max\{m \in \mathbb{N} \cup \{0\} \,:\, (\epsilon_1(\boldsymbol{r}^{\boldsymbol{m}}, \Pi_1))_0 = 0 \ \forall \boldsymbol{m} : |\boldsymbol{m}| \leqslant m\}$.
2. *Полагаем* $m = P_A + 1$.
3. *Вычисляем* $f^{\boldsymbol{m}} = -(\epsilon_1(\boldsymbol{r}^{\boldsymbol{m}}/\boldsymbol{m}!, \tilde{\Pi}_1^{(P_A+1,m-1)}))_0$ *для всех* $\boldsymbol{m} : |\boldsymbol{m}| = m$, *подставляя в* $\tilde{\Pi}_h^{(P_A+1,m-1)}$ *ранее найденные коэффициенты* $\mathfrak{C}^{(\boldsymbol{n})}$, $P_A + 1 \leqslant |\boldsymbol{n}| \leqslant m - 1$.
4. *Если для всех* $\boldsymbol{m} : |\boldsymbol{m}| = m$ *выполняется* $f^{\boldsymbol{m}} \in \mathrm{Im}L(0)$, *то:*
   - *находим диагональные матрицы* $\mathfrak{C}^{(\boldsymbol{m})}$ *из систем* $L(0)\mathfrak{C}^{(\boldsymbol{m})}\mathfrak{e} = f^{\boldsymbol{m}}$ *(с точностью до* $c_{\boldsymbol{m}}I$, $c_{\boldsymbol{m}} \in \mathbb{R}$*);*
   - *увеличиваем значение* m *на единицу;*
   - *возвращаемся к п. 3.*
5. *Полагаем* $Q' = m - 1$.
6. *Если* $Q' = P_A$, *полагаем* $P' = P_A$, *иначе полагаем* $P' = P_A + 1$.

Доказательство корректности этого алгоритма полностью повторяет доказательство корректности алгоритма 2 в [3].

8. Квазиодномерный случай

Будем называть схему *квазиодномерной*, если матрица $A(\boldsymbol{\phi})$ зависит только от проекции $\boldsymbol{\phi}$ на некоторое выделенное направление. Примером такой схемы является метод Галёркина с разрывными базисными функциями на двумерной регулярной сетке (то есть либо состоящей из одинаковых параллелепипедов, либо полученной из неё разбиением этих параллелепипедов диагоналями однородным образом) в случае, когда скорость переноса направлена вдоль одной из сеточных линий.

Утверждение 8.1. *Пусть* Π_h *и* $\mathcal{P}_h$ *– отображения вида* (2.7). *Пусть схема* (2.5) *в смысле* Π_h *обладает формальным порядком точности* P *и порядком*

точности в длительном счёте Q. Пусть шаблон схемы $\mathcal{S}$ лежит в некотором одномерном подмножестве $\mathbb{Z}^d$, то есть существует такой $\boldsymbol{\eta} \in \mathbb{Z}^d$, что $\mathcal{S} \subset \{m\boldsymbol{\eta}, m = -M, \ldots, M\}$. Тогда существуют такие диагональные матрицы $\mathfrak{C}^{(m)} \in \mathbb{R}^{M^0}$, что схема обладает порядком аппроксимации Q в смысле модифицированного отображения $\tilde{\Pi}_h : H^{Q+1}_{per}(\mathbb{R}^d) \to V_{per}$, определяемого (2.8).

Доказательство этого утверждения в основном повторяет одномерный случай (см. [3]), поэтому приведём только его схему. Без ограничения общности будем считать, что выделенное направление совпадает с осью x.

Зависимость матрицы только от ϕ_x позволяет получить для неё представление (3.13)–(3.14). По аналогии с (4.4) введём вектор $v(\boldsymbol{\phi})$ следующим образом:

$$v(\boldsymbol{\phi}) = S^{-1}(\phi_x) \left(\Pi_1 e^{i(U\boldsymbol{\phi})\cdot\boldsymbol{r}}\right)_0, \tag{8.1}$$

и через $v_j(\boldsymbol{\phi})$ будем обозначать его компоненты, размерность которых соответствует размерам блоков матриц $M_j(\phi_x)$ в представлении (3.14). Введём целые числа p_j условиями

$$c_j\psi^{p_j} \leqslant \sup_{\boldsymbol{\phi}:|\boldsymbol{\phi}|=\psi} \|v_j(\boldsymbol{\phi})\| \leqslant 2c_j\psi^{p_j}, \tag{8.2}$$

где $c_j \neq 0$, и $p_j = \infty$, если $v_j(\boldsymbol{\phi}) \equiv 0$. Повторяя доказательство для одномерного случая, получаем $p_j \geqslant \min\{P, Q+1-j\}$. Определим в окрестности $\boldsymbol{\phi} = 0$ аналитическую функцию

$$V(\boldsymbol{\phi}) = S(\phi_x) \begin{pmatrix} \delta_0 I & 0 & \ldots & 0 & 0 \\ 0 & \delta_1 I & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & \delta_m I & 0 \\ 0 & 0 & \ldots & 0 & \delta_\infty I \end{pmatrix} S^{-1}(\phi_x)(\Pi_1 e^{i(U\boldsymbol{\phi})\cdot\boldsymbol{r}})_0, \tag{8.3}$$

где размеры блоков соответствуют представлению (3.14) матрицы M, а $\delta_j = 0$, если $p_j \geqslant P$, и $\delta_j = 1$ при $p_j < P$. Определим коэффициенты $\mathfrak{C}^{(m)}$ равенствами (4.11), (4.12). Как и в доказательстве утверждения 7.2, получаем $(\tilde{\Pi}_1^{(P,Q)} e^{i(U\boldsymbol{\phi})\cdot\boldsymbol{r}})_0 = V(\boldsymbol{\phi}) + O(|\boldsymbol{\phi}|^{Q+1})$ и $A(\boldsymbol{\phi})(\tilde{\Pi}_1^{(P,Q)} e^{i(U\boldsymbol{\phi})\cdot\boldsymbol{r}})_0 = O(|\boldsymbol{\phi}|^{Q+1})$, что в силу утверждения 3.1 влечёт Q-й порядок аппроксимации в смысле $\tilde{\Pi}_h^{(P,Q)}$.

Таким образом, в квазиодномерном случае анализ структуры ошибки решения по схеме (2.5) и получение оптимальной оценки точности может быть проведено путём поиска модифицированного отображения $\tilde{\Pi}_h$ в виде (2.8).

9. Контрпример

Условия, накладываемые в утверждениях 7.2 и 8.1, являются существенными. Обоснуем это примером схемы, которая обладает оценкой ошибки $\|\varepsilon_h(v_0,t,\mathring{\Pi}_h)\| \leqslant C_1 h\|\nabla v_0\|_\infty + C_2 h^2(h+t)\|\nabla^3 v_0\|_\infty$, но для которой не существует отображение $\mathring{\tilde{\Pi}}_h$ вида (2.10), в смысле которого схема обладает вторым порядком аппроксимации.

Рассмотрим уравнение переноса (2.3)–(2.4) в $\mathbb{R}^2$ со скоростью переноса $\boldsymbol{\omega} = 0$. Пусть $\boldsymbol{a}_1 = (1,0)^T$, $\boldsymbol{a}_2 = (0,1)^T$, так что $T = U = I$. Рассмотрим схему вида (2.5) при $|M^0| = 3$, ненулевые коэффициенты которой равны $Z_{(0,0)} = I$,

$$\begin{gathered} L_{(-1,-1)} = W, \quad L_{(0,-1)} = -F, \quad L_{(1,-1)} = -W, \\ L_{(-1,0)} = -E, \quad L_{(0,0)} = 0, \quad L_{(1,0)} = E, \\ L_{(-1,1)} = -W, \quad L_{(0,1)} = F, \quad L_{(1,1)} = W, \end{gathered}$$

где

$$W = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \; F = \begin{pmatrix} 1 & -2 & 1 \\ -2 & 1 & 1 \\ 1 & 1 & -2 \end{pmatrix}, \; E = \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ -1 & 1 & 0 \end{pmatrix}.$$

Поскольку $L_{\boldsymbol{\eta}} = -L^*_{-\boldsymbol{\eta}}$, выполняется $A(\boldsymbol{\phi})^* = -A(\boldsymbol{\phi})$ и схема является устойчивой. Будем считать, что $\mathring{\Pi}_h$ определяется как $(\mathring{\Pi}_h f)_{\boldsymbol{\eta},\xi} = f(\eta_1 h, \eta_2 h)$.

Вначале проведём анализ аппроксимационной ошибки, вычисляя величины $(\epsilon_1(f,\mathring{\Pi}_1))_0 = (L\mathring{\Pi}_1 f)_0 = \sum_{\boldsymbol{\eta}} L_{\boldsymbol{\eta}} (\mathring{\Pi}_1 f)_{\boldsymbol{\eta}}$ для $f = \boldsymbol{r}^{\boldsymbol{m}}$. Имеем $(\mathring{\Pi}_1 \boldsymbol{r}^{\boldsymbol{m}})_{\boldsymbol{\eta}} = \boldsymbol{\eta}^{\boldsymbol{m}} \mathfrak{e}$,

$$\sum_{\boldsymbol{\eta}} L_{\boldsymbol{\eta}} = 0, \quad \sum_{\boldsymbol{\eta}} \eta_1 L_{\boldsymbol{\eta}} = 2E, \quad \sum_{\boldsymbol{\eta}} \eta_2 L_{\boldsymbol{\eta}} = 2F, \tag{9.1}$$

$$\sum_{\boldsymbol{\eta}} \eta_1^2 L_{\boldsymbol{\eta}} = 0, \quad \sum_{\boldsymbol{\eta}} \eta_1 \eta_2 L_{\boldsymbol{\eta}} = 4W, \quad \sum_{\boldsymbol{\eta}} \eta_2^2 L_{\boldsymbol{\eta}} = 0. \tag{9.2}$$

Поскольку $E\mathfrak{e} = F\mathfrak{e} = 0$, схема обладает первым порядком аппроксимации, однако, так как $W\mathfrak{e} \neq 0$, вторым порядком аппроксимации схема не обладает.

Покажем, что выполняется оценка

$$\|\varepsilon_h(t,v_0,\mathring{\Pi}_h)\| \leqslant C_1 h\|\nabla v_0\|_\infty + C_2 h^2(t+h)\|\nabla^3 v_0\|_\infty. \tag{9.3}$$

Достаточно показать эту оценку для функций вида $v_0 = e^{i\boldsymbol{\alpha}(\boldsymbol{k})\cdot\boldsymbol{r}}$, $\boldsymbol{k} \in \mathbb{Z}^2$. Действительно, если для $v_0 = e^{i\boldsymbol{\alpha}(\boldsymbol{k})\cdot\boldsymbol{r}}$ выполняется (9.3), то для $\boldsymbol{\phi} = h\boldsymbol{\alpha}(\boldsymbol{k})$ имеем

$$\left\|\hat{\varepsilon}\left(\nu,\boldsymbol{\phi},\mathring{\Pi}_h\right)\right\| = \left\|\varepsilon_h\left(\nu h, e^{i\boldsymbol{\phi}\cdot\boldsymbol{r}/h}, \mathring{\Pi}_h\right)\right\| \leqslant C_1|\boldsymbol{\phi}| + C_2(\nu+1)|\boldsymbol{\phi}|^3,$$

откуда в силу утверждения 3.3 следует (9.3) для всех $v_0 \in C^3_{per}(\mathbb{R}^2)$.

Показать (9.3) для $v_0 = e^{i\boldsymbol{\alpha}(\boldsymbol{k})\cdot\boldsymbol{r}}$, $\boldsymbol{k} \in \mathbb{Z}^2$, можно двумя способами. Первым способом является введение модифицированного отображения $\mathring{\tilde{\Pi}}_h$. Действительно, зафиксируем направление $\boldsymbol{e} = \boldsymbol{\alpha}(\boldsymbol{k})/|\boldsymbol{\alpha}(\boldsymbol{k})|$. Будем искать отображение $\tilde{\Pi}_h$ в виде

$$\mathring{\tilde{\Pi}}_{h,\boldsymbol{e}} f = \mathring{\Pi}_h f + h\mathfrak{C}_{\boldsymbol{e}} \mathring{\Pi}_h \frac{\partial f}{\partial \boldsymbol{e}} \tag{9.4}$$

исходя из условии 2-го порядка аппроксимации рассматриваемой схемы на функциях вида $v(\boldsymbol{r}) = \hat{v}(\boldsymbol{e}\cdot\boldsymbol{r})$. Поскольку схема обладает 1-м порядком аппроксимации на любых гладких функциях, достаточно рассмотреть функцию $v(\boldsymbol{r}) = (\boldsymbol{e}\cdot\boldsymbol{r})^2/2$. Имеем аппроксимационную ошибку

$$\begin{gathered}\left(\epsilon_1(v, \mathring{\tilde{\Pi}}_{1,\boldsymbol{e}})\right)_0 = \sum_{\boldsymbol{\eta}} L_{\boldsymbol{\eta}} \left(\frac{(\boldsymbol{e}\cdot\boldsymbol{\eta})^2}{2}\mathfrak{e} + (\boldsymbol{e}\cdot\boldsymbol{\eta})\mathfrak{C}_{\boldsymbol{e}}\mathfrak{e} \right) = \\ = 4e_x e_y W\mathfrak{e} + 2(e_x E + e_y F)\mathfrak{C}_{\boldsymbol{e}}\mathfrak{e}.\end{gathered} \tag{9.5}$$

Подставляя теперь выражения для E, F и W и приравнивая $(\epsilon_1(v, \mathring{\tilde{\Pi}}_{1,\boldsymbol{e}}))_0$ к нулю, получаем

$$\begin{pmatrix} e_x + e_y & -2e_y & -e_x + e_y \\ -2e_y & -e_x + e_y & e_x + e_y \\ -e_x + e_y & e_x + e_y & -2e_y \end{pmatrix} \mathfrak{C}_{\boldsymbol{e}}\mathfrak{e} = -2e_x e_y \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}. \tag{9.6}$$

Матрица $e_x E + e_y F$ симметрическая. Поскольку сумма столбцов равна нулю, у неё есть собственное значение $\lambda = 0$, которому соответствует собственный вектор $\mathfrak{e}$. С другой стороны, главный минор размера 2 равен $(e_x + e_y)(-e_x + e_y) - 4e_y^2 = -3e_y^2 - e_x^2$ и отличен от нуля при $\boldsymbol{e} \neq 0$. Следовательно, критерием совместности этой системы является ортогональность правой части вектору $\mathfrak{e}$, и он, очевидно, выполняется. Можно показать, что решением этой системы является диагональная матрица

$$\mathfrak{C}_{\boldsymbol{e}} = \frac{2e_x e_y}{e_x^2 + 3e_y^2} \begin{pmatrix} e_y & 0 & 0 \\ 0 & -e_x & 0 \\ 0 & 0 & -e_y \end{pmatrix} + c_{\boldsymbol{e}} I. \tag{9.7}$$

Положим $c_{\boldsymbol{e}} = 0$. Таким образом, для любого направления $\boldsymbol{e}$ существует отображение $\mathring{\tilde{\Pi}}_{h,\boldsymbol{e}}$ вида (9.4), в смысле которого схема обладает вторым порядком аппроксимации на плоских волнах, таких что волновой вектор направлен по $\boldsymbol{e}$.

По неравенству треугольника и в силу устойчивости с $K = 1$ имеем

$$\begin{gathered}\|\varepsilon_h(t, e^{i\boldsymbol{\alpha}(\boldsymbol{k})\cdot\boldsymbol{r}}, \mathring{\Pi}_h)\| \leqslant \|\varepsilon_h(t, e^{i\boldsymbol{\alpha}(\boldsymbol{k})\cdot\boldsymbol{r}}, \mathring{\tilde{\Pi}}_{h,\boldsymbol{e}})\| + 2\left\|(\mathring{\tilde{\Pi}}_{h,\boldsymbol{e}} - \mathring{\Pi}_h)e^{i\boldsymbol{\alpha}(\boldsymbol{k})\cdot\boldsymbol{r}}\right\| \leqslant \\ \leqslant t\left\|\epsilon_h(e^{i\boldsymbol{\alpha}(\boldsymbol{k})\cdot\boldsymbol{r}}, \mathring{\tilde{\Pi}}_{h,\boldsymbol{e}})\right\| + 2h\|\mathfrak{C}_{\boldsymbol{e}}\||\boldsymbol{\alpha}(\boldsymbol{k})|.\end{gathered} \tag{9.8}$$

Оценим первое слагаемое в правой части. Очевидно,

$$\left\|\epsilon_h(e^{i\boldsymbol{\alpha}(\boldsymbol{k})\cdot\boldsymbol{r}},\mathring{\tilde{\Pi}}_{h,\boldsymbol{e}})\right\| = \left\|\left(\epsilon_h(e^{i\boldsymbol{\alpha}(\boldsymbol{k})\cdot\boldsymbol{r}},\mathring{\tilde{\Pi}}_{h,\boldsymbol{e}})\right)_0\right\|.$$

Далее, по определению

$$\left(\epsilon_h(e^{i\boldsymbol{\alpha}(\boldsymbol{k})\cdot\boldsymbol{r}},\mathring{\tilde{\Pi}}_{h,\boldsymbol{e}})\right)_0 = -\sum_{\boldsymbol{\eta}\in\mathcal{S}} L_{\boldsymbol{\eta}}\left(\mathring{\tilde{\Pi}}_h e^{i\boldsymbol{\alpha}(\boldsymbol{k})\cdot\boldsymbol{r}}\right)_{\boldsymbol{\eta}}.$$

Введём $g(x) = 1 + ix - x^2/2$. В силу точности на многочленах 2-го порядка от $(\boldsymbol{\alpha}(\boldsymbol{k})\cdot\boldsymbol{r})$ получаем

$$\left(\epsilon_h(e^{i\boldsymbol{\alpha}(\boldsymbol{k})\cdot\boldsymbol{r}},\mathring{\tilde{\Pi}}_{h,\boldsymbol{e}})\right)_0 = -\frac{1}{h}\sum_{\boldsymbol{\eta}\in\mathcal{S}} L_{\boldsymbol{\eta}}\left(\mathring{\tilde{\Pi}}_{h,\boldsymbol{e}}\left[e^{i\boldsymbol{\alpha}(\boldsymbol{k})\cdot\boldsymbol{r}} - g(\boldsymbol{\alpha}(\boldsymbol{k})\cdot\boldsymbol{r})\right]\right)_{\boldsymbol{\eta}} =$$

$$= -\frac{1}{h}\sum_{\boldsymbol{\eta}\in\mathcal{S}} L_{\boldsymbol{\eta}}\left[e^{i\boldsymbol{\alpha}(\boldsymbol{k})\cdot h\boldsymbol{\eta}} - g(h\boldsymbol{\alpha}(\boldsymbol{k})\cdot\boldsymbol{\eta})\right]\mathfrak{e} -$$

$$-\sum_{\boldsymbol{\eta}\in\mathcal{S}} L_{\boldsymbol{\eta}}(i\boldsymbol{\alpha}(\boldsymbol{k})\cdot\boldsymbol{e})\left[e^{i\boldsymbol{\alpha}(\boldsymbol{k})\cdot h\boldsymbol{\eta}} - (1+ih\boldsymbol{\alpha}(\boldsymbol{k})\cdot\boldsymbol{\eta})\right]\mathfrak{C}_{\boldsymbol{e}}\mathfrak{e}.$$

Видно, что эта величина имеет оценку $O(|\boldsymbol{\alpha}(\boldsymbol{k})|^3h^2)$ с константой, не зависящей от направления $\boldsymbol{\alpha}$. Подставляя эту оценку в (9.8), получаем (9.3) с подстановкой $e^{i\boldsymbol{\alpha}(\boldsymbol{k})\cdot\boldsymbol{r}}$ вместо v_0, причём константы C_1 и C_2 не зависят от $\boldsymbol{\alpha}(\boldsymbol{k})$, что и требовалось доказать.

Оценку (9.3) можно показать и при помощи спектрального анализа. Мы не будем проводить рассуждение строго, ограничившись общими соображениями. Имеем

$$A(\boldsymbol{\phi}) = -L(\boldsymbol{\phi}) = W\left(e^{i(\phi_1+\phi_2)} + e^{-i(\phi_1+\phi_2)} - e^{i(\phi_1-\phi_2)} - e^{i(-\phi_1+\phi_2)}\right) +$$

$$+E\left(e^{i\phi_1} - e^{-i\phi_1}\right) + F\left(e^{i\phi_1} - e^{-i\phi_1}\right) =$$

$$= -4W\sin(\phi_1)\sin(\phi_2) + 2iE\sin(\phi_1) + 2iF\sin(\phi_2).$$

Введём обозначения $u = \sin(\phi_1)$, $v = \sin(\phi_2)$. Тогда

$$A(\boldsymbol{\phi}) = 2i\begin{pmatrix} v+u & -2v+2iuv & v-u \\ -2v-2iuv & v-u & v+u+2iuv \\ v-u & v+u-2iuv & -2v \end{pmatrix}.$$

Характеристическое уравнение для матрицы $A(\boldsymbol{\phi})/2i$ имеет вид

$$\lambda^3 - (9v^2 + 3u^2 + 8v^2u^2)\lambda + 4v^2u^2(v-u) = 0.$$

После замены переменных $\lambda = \mu(9v^2 + 3u^2 + 8v^2u^2)^{1/2}$ получаем

$$\mu^3 - \mu + \omega = 0,$$

где

$$\omega = 4v^2u^2(v-u)(9v^2 + 3u^2 + 8v^2u^2)^{-3/2}.$$

При $(u,v) \to (0,0)$ выполняется $\omega \to 0$. При $\omega = 0$ имеем три корня: $\mu = 0$, $\mu = \pm 1$; им соответствует корень $\lambda = 0$ кратности 3. При малых ω имеем $\mu_1(\omega) \approx \omega$, что соответствует $|\lambda_1(\boldsymbol{\phi})| \sim |\boldsymbol{\phi}|^3$, а остальные собственные значения $|\lambda_2(\boldsymbol{\phi})|, |\lambda_3(\boldsymbol{\phi})| \sim |\boldsymbol{\phi}|$.

Пусть $\Lambda(\boldsymbol{\phi}) = \mathrm{diag}\{\lambda_1(\boldsymbol{\phi}), \lambda_2(\boldsymbol{\phi}), \lambda_3(\boldsymbol{\phi})\}$, $A(\boldsymbol{\phi}) = S(\boldsymbol{\phi})\Lambda(\boldsymbol{\phi})S^{-1}(\boldsymbol{\phi})$. Компоненты вектора $v(\boldsymbol{\phi}) = S^{-1}(\boldsymbol{\phi})(\mathring{\Pi}_1 e^{i\boldsymbol{\phi}\cdot\boldsymbol{r}})_0$, соответствующие второму и третьему собственным значениям, имеют величину не больше чем порядка $O(|\boldsymbol{\phi}|)$, поскольку в противном случае не может быть сеточной сходимости. Значит, компонента, соответствующая первому собственному значению, имеет величину порядка $O(1)$, поскольку хотя бы одна компонента должна иметь величину порядка $O(1)$. Отсюда

$$\hat{\varepsilon}(\boldsymbol{\phi}, \nu, \mathring{\Pi}_h) = S(\boldsymbol{\phi})\left(e^{\nu\Lambda(\phi)} - 1\right)v(\boldsymbol{\phi}) = O(|\boldsymbol{\phi}| + \nu|\boldsymbol{\phi}|^3).$$

В силу утверждения 3.2 для ошибки решения по рассматриваемой схеме имеет место оценка $O(h + h^2 t)$.

Наконец, продемонстрируем эту оценку в численном счёте. Положим начальные данные равными $v_0(\boldsymbol{r}) = e^{i\boldsymbol{\alpha}\cdot\boldsymbol{r}}$ при $\boldsymbol{\alpha} = (2\pi, -2\pi)$. Расчёты будем проводить на сетках с $h = 1/60$, $h = 1/120$ и $h = 1/240$. Зависимость нормы ошибки решения от времени приведена на рис. 1. Видно, что при малых временах ошибка имеет величину порядка $O(h)$, тогда как растущая со временем компонента ошибки – величину порядка $O(h^2)$. Интегрирование по времени осуществлялось явным 3-стадийным методом Рунге–Кутты 3-го порядка. Шаг по времени был выбран равным $h/20$; его дальнейшее уменьшение визуально не отражалось на величине ошибки.

Теперь, покажем, что не существует отображения $\mathring{\tilde{\Pi}}_h$ вида

$$\begin{aligned}\left(\mathring{\tilde{\Pi}}_h f\right)_{\boldsymbol{\eta}} &= \left(\mathring{\Pi}_h f\right)_{\boldsymbol{\eta}} + h\left(\mathfrak{C}^{(x)}\left(\mathring{\Pi}_h f_x\right)_{\boldsymbol{\eta}} + \mathfrak{C}^{(y)}\left(\mathring{\Pi}_h f_y\right)_{\boldsymbol{\eta}}\right) + \\ &+ h^2\left(\mathfrak{C}^{(xx)}\left(\mathring{\Pi}_h f_{xx}\right)_{\boldsymbol{\eta}} + \mathfrak{C}^{(xy)}\left(\mathring{\Pi}_h f_{xy}\right)_{\boldsymbol{\eta}} + \mathfrak{C}^{(yy)}\left(\mathring{\Pi}_h f_{yy}\right)_{\boldsymbol{\eta}}\right),\end{aligned} \tag{9.9}$$

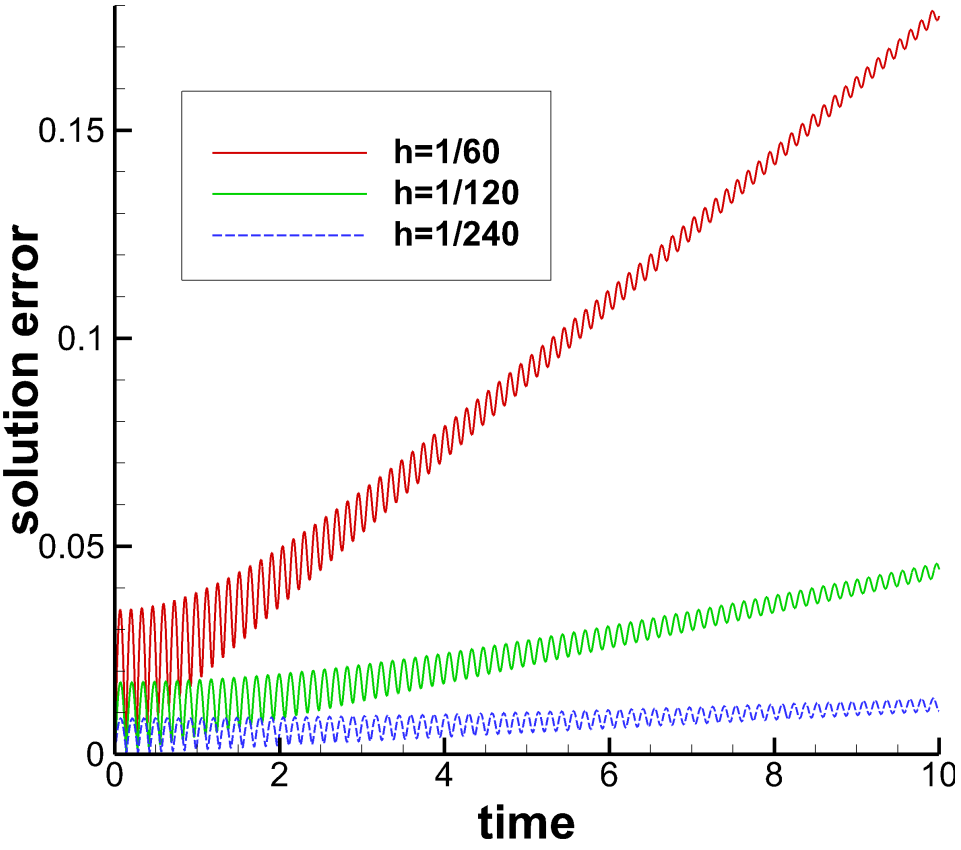

Рис. 1. Зависимость ошибки решения от времени

в смысле которого схема обладает 2-м порядком аппроксимации. Действительно, поскольку $A(0) = 0$, аппроксимационная ошибка в смысле отображения (9.9) на квадратичном полиноме не зависит от коэффициентов $\mathfrak{C}^{(xx)}$, $\mathfrak{C}^{(xy)}$, $\mathfrak{C}^{(yy)}$. Поэтому отображение, доставляющее 2-й порядок аппроксимации, нужно искать в виде

$$\left(\mathring{\tilde{\Pi}}_h f\right)_{\boldsymbol{\eta}} = \left(\mathring{\Pi}_h f\right)_{\boldsymbol{\eta}} + h\left(\mathfrak{C}^{(x)}\left(\mathring{\Pi}_h f_x\right)_{\boldsymbol{\eta}} + \mathfrak{C}^{(y)}\left(\mathring{\Pi}_h f_y\right)_{\boldsymbol{\eta}}\right).$$

Аппроксимационная ошибка должна быть равна нулю на любой квадратичной функции; достаточно проверять функции $f = x^2$, $f = y^2$, $f = xy$. На функции $f = x^2$ образ под действием этого проектора совпадает с (9.4) при $\boldsymbol{e} = (1,0)^T$, где $\mathfrak{C}_{\boldsymbol{e}} \equiv \mathfrak{C}^{(x)}$. Система (9.6) на нахождение этих коэффициентов при подстановке $\alpha_1 \neq 0$, $\alpha_2 = 0$ имеет решение $\mathfrak{C}^{(x)} = \kappa_x I$, где κ_x – любое. Аналогичным образом получаем $\mathfrak{C}^{(y)} = \kappa_y I$, где κ_y – любое. Из двух уравнений, таким образом, мы однозначно (с точностью до κI, не влияющих на аппроксимационную ошибку) определили коэффициенты отображения. Однако легко удостовериться, что аппроксимационная ошибка на функции xy не зависит от κ_x и κ_y и в силу (9.2) равна $4Wh$. Таким образом, искомого отображения не существует, что и требовалось доказать.

10. Особые направления волнового вектора

С формальной точки зрения, мы можем рассматривать только волны с волновыми векторами $\boldsymbol{\alpha}(\boldsymbol{k})$, $\boldsymbol{k} \in \mathbb{Z}^d$. То есть число направлений, по которым могут распространяться плоские волны, счётно. Более того, при разных N векторы $\varepsilon_h(t, e^{i\boldsymbol{\alpha}(\boldsymbol{k})\cdot\boldsymbol{r}}, \Pi_h)$ имеют разные размерности. Но ввиду равенства (3.7) величина $\|\varepsilon_h(t, e^{i\boldsymbol{\alpha}(\boldsymbol{k})\cdot\boldsymbol{r}}, \Pi_h)\|^2$ является аналитической функцией $\boldsymbol{\alpha}(\boldsymbol{k})$, и это позволяет по непрерывности доопределить её на некоторую окрестность нуля $\boldsymbol{\alpha} \in \tilde{G} \subset \mathbb{R}^d$.

Пусть Ω – единичная сфера в $\mathbb{R}^d$. Будем называть направление $\boldsymbol{e} \in \Omega$ *минус-особым* по аппроксимации (точности, точности в длительном счёте), если существует последовательность $\boldsymbol{e}_n \to \boldsymbol{e}$, такая что при всех n порядок аппроксимации (формальный порядок точности, порядок точности в длительном счёте) для волн, распространяющихся в направлении $\boldsymbol{e}_n$, равен A, но для волн, распространяющихся в направлении $\boldsymbol{e}$, он равен $\tilde{A} < A$. Аналогично будем называть направление *плюс-особым*, еслии $\tilde{A} > A$. Рассмотрим вопрос, какими могут быть особые направления.

Пример схемы, для которой существует плюс-особое направление, построить легко. Рассмотрим конечно-разностную схему 1-го порядка на равномерной сетке для аппроксимации уравнения переноса $\partial u/\partial t + \partial u/\partial x = 0$ в двумерном случае ($\boldsymbol{r} = (x, y)^T \in \mathbb{R}^2$, $\boldsymbol{\omega} = (1, 0)^T$):

$$\frac{du_{i,j}}{dt} + \frac{u_{i,j} - u_{i-1,j}}{h} = 0.$$

Рассмотрим ошибку аппроксимации в смысле некоторого локального отображения Π_h вида (2.7). Для решений, зависящих только от y, схема будет точна, то есть порядки аппроксимации, точности и точности в длительном счёте формально будут бесконечными. Для всех остальных решений схема будет обладать первым порядком аппроксимации, первым порядком точности и первым порядком точности в длительном счёте.

Теперь рассмотрим вопрос, возможны ли минус-особые направления.

Утверждение 10.1. *Минус-особых направлений по аппроксимации не существует.*

Порядок аппроксимации P_A в смысле отображений $\hat{\Pi}_h = \Pi_h$ или $\hat{\Pi}_h = \mathring{\Pi}_h$ для плоских волн, распространяющихся по некоторому направлению $\tilde{\boldsymbol{e}}$, равносилен выполнению равенств $(\epsilon_1((\boldsymbol{r}\cdot\tilde{\boldsymbol{e}})^m, \hat{\Pi}_1))_0 = 0$ для $m = 1, \ldots, P_A$, где $\epsilon_h(f, \hat{\Pi}_h)$ определено (2.12). Выражения $(\epsilon_1((\boldsymbol{r}\cdot\tilde{\boldsymbol{e}})^m, \hat{\Pi}_1))_0$ являются аналитическими функциями $\tilde{\boldsymbol{e}}$, поэтому если они равны нулю для некоторой последовательности $\boldsymbol{e}_n \to \boldsymbol{e}$, то они равны нулю и для $\boldsymbol{e}$. Таким образом, минус-особых направлений по аппроксимации не существует.

Утверждение 10.2. *Минус-особых направлений по точности не существует.*

Аналогичное утверждение можно получить для порядка точности. Если для волн, распространяющихся по направлениям $\boldsymbol{e}_n$, имеет место порядок точности P, то величины $\hat{\varepsilon}^{(1)}(T^*\boldsymbol{\alpha}h, t/h, \Pi_h)$ и $\hat{\varepsilon}^{(2)}(T^*\boldsymbol{\alpha}h, t/h, \Pi_h)$ являются величинами P-го порядка малости по h. Поскольку $\hat{\varepsilon}^{(2)}(T^*\boldsymbol{\alpha}h, t/h, \Pi_h)$ – аналитическая по h, это означает, что члены в её разложении слагаемые при h^p, $p < P$, отсутствуют. Поскольку эта функция также аналитическая по $\boldsymbol{\alpha}$, свойство отсутствия членов h^p, $p < P$, переносится с направлений $\boldsymbol{e}_n$ на направление $\boldsymbol{e}$. Величина $\hat{\varepsilon}^{(1)}(T^*\boldsymbol{\alpha}h, t/h, \Pi_h)$ имеет представление (4.5), причём верно (4.7). Следовательно, $\|v^{(*)}(\boldsymbol{e}_n|\boldsymbol{\alpha}|h)\| \leqslant \hat{C}h^P|\boldsymbol{\alpha}|^P$. Но $v^{(*)}(\boldsymbol{\alpha}h)$ по построению является аналитической функцией $\boldsymbol{\alpha}$ и h, поэтому $\|v^{(*)}(\boldsymbol{e}|\boldsymbol{\alpha}|h)\| \leqslant \hat{C}h^P|\boldsymbol{\alpha}|^P$, и в силу устойчивости получаем, что $\hat{\varepsilon}^{(1)}(T^*\boldsymbol{\alpha}h, t/h, \Pi_h)$ имеет P-й порядок малости по h. Таким образом, минус-особых направлений по точности не существует.

Утверждение 10.3. *Пусть собственное значение $\lambda = 0$ матрицы $A(0)$ является простым. Тогда минус-особых направлений по точности в длительном счёте не существует.*

Действительно, пусть схема имеет некий формальный порядок P. Порядок точности в длительном счёте Q определяется только его “физическим” собственным значением $\lambda(\boldsymbol{\phi})$. Напомним, что $\lambda(\boldsymbol{\phi})$ является аналитической функцией в окрестности $\boldsymbol{\phi} = 0$. Порядок точности в длительном счёте Q на $v_0 \in H^{Q+1}_{per}(\mathbb{R}^d)$ имеется тогда и только тогда, когда в некоторой окрестности нуля выполняется $|\lambda(\boldsymbol{\phi})| \leqslant C|\boldsymbol{\phi}|^{Q+1}$. Аналогично, порядок точности в длительном счёте Q при начальных данных вида $v_0(\boldsymbol{r}) = f(\boldsymbol{r} \cdot \boldsymbol{e})$, имеется тогда и только тогда, когда $|\lambda(\boldsymbol{e}\phi)| \leqslant C|\phi|^{Q+1}$. Отсюда видно, что в простом случае минус-особых направлений по точности в длительном счёте не бывает.

Условие простоты собственного значения $\lambda = 0$ матрицы $A(0)$ является существенным. Пример схемы, обладающей минус-особым направлением по точности в длительном счёте, будет приведён в следующей работе.

Замечание

В [4] доказано следующее утверждение.

Утверждение 10.4. *Пусть* $A \in \mathbb{C}^{n\times n}$ *удовлетворяет* (3.3)*,* $v \in \mathbb{C}^n$*. Пусть существуют такие* $C_1, C_2 \geqslant 0$*, что при всех* $\nu \geqslant 0$ *справедлива оценка*

$$\|(e^{\nu A} - I)v\| \leqslant (C_1 + C_2\nu)\|v\|. \tag{10.1}$$

Тогда

$$v^* \left(A^*A + \frac{C_2^2}{C_1^2}\right)^{-1} A^*Av \leqslant \tilde{\delta}^2 C_1^2 \|v\|^2, \tag{10.2}$$

где $\tilde{\delta}$ *зависит только от* n *и* K*. Обратно, если выполняется* (10.2)*, то*

$$\|(e^{\nu A} - I)v\| \leqslant \tilde{\delta}(K+1)(C_1 + C_2\nu)\|v\|. \tag{10.3}$$

Покажем, что из него следует 1) сходимость с порядком аппроксимации (устойчивость схемы предполагается) и 2) утверждение 5.2. Действительно, пусть схема обладает порядком аппроксимации P_A. Тогда по утверждению 3.1 имеем $\|Av\| = O(|\boldsymbol{\phi}|^{P_A+1})$. Значит, $v^*A^*Av = O(|\boldsymbol{\phi}|^{2(P_A+1)})$, то есть

$$v^* \left(|\boldsymbol{\phi}|^2\right)^{-1} A^*Av = O(|\boldsymbol{\phi}|^{2P_A}).$$

Добавлением положительно полуопределённой эрмитовой матрицы под минус первую степень мы не можем увеличить значение выражения, поэтому

$$v^* \left(A^*A + |\boldsymbol{\phi}|^2\right)^{-1} A^*Av = O(|\boldsymbol{\phi}|^{2P_A}),$$

что по утверждению 10.4 влечёт неравенство (10.3) с $C_1 = O(|\boldsymbol{\phi}|^{P_A})$ и $C_2 = O(|\boldsymbol{\phi}|^{P_A+1})$.

Пусть теперь $A(0) = 0$ и схема обладает порядком точности P. Тогда $A(\boldsymbol{\phi})$ выражается в виде произведения $|\boldsymbol{\phi}|$ на некоторую функцию (не обязательно аналитическую), ограниченную в нуле, и, следовательно, $A^*A = |\boldsymbol{\phi}|^2 F(\boldsymbol{\phi})$, где F ограничено в нуле. Так как схема обладает порядком точности P, по утверждению 10.4 выполняется 10.2 с подстановкой $C_1 = c_1|\boldsymbol{\phi}|^P$, $C_2 = c_2|\boldsymbol{\phi}|^{P+1}$, то есть

$$v^* \left(|\boldsymbol{\phi}|^2(I + F(\boldsymbol{\phi}))\right)^{-1} A^*Av = O(|\boldsymbol{\phi}|^{2P}).$$

Ввиду ограниченности F в нуле отсюда следует

$$v^*A^*Av = O(|\boldsymbol{\phi}|^{2P+2}),$$

то есть $\|Av\| = O(|\boldsymbol{\phi}|^{P+1})$. В силу утверждения 3.1 схема обладает P-м порядком аппроксимации.

Минус-особое направление по точности в длительном счёте

Выше было показано, что минус-особых направлений по аппроксимации и точности не бывает. Приведём пример схемы, для которой есть минус-особые направления по точности в длительном счёте. А именно, обладающей 1-м порядком точности в длительном счёте на волнах, распространяющихся в горизонтальном направлении, и 2-м порядком точности на остальных плоских волнах.

Рассмотрим уравнение (2.3) с $\boldsymbol{\omega} = 0$, и пусть $\boldsymbol{a}_1 = (1,0)^T$, $\boldsymbol{a}_2 = (0,1)^T$. Положим $M^0 = \{L, R\}$ и определим матрицы

$$W = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad E = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Ненулевые коэффициенты схемы (2.5) выберем равными $Z_{0,0} = I$,

$$L_{0,0} = -2W, \quad L_{\pm 1,0} = \pm E, \quad L_{0,\pm 1} = W.$$

Оператор $\mathring{\Pi}_h$ зададим равенствами $(\mathring{\Pi}_h f)_{\boldsymbol{\eta},L} = (\mathring{\Pi}_h f)_{\boldsymbol{\eta},R} = f(\eta_1 h, \eta_2 h)$. Матрица $A(\boldsymbol{\phi}) = -L(\boldsymbol{\phi})$ имеет вид

$$A(\phi_1, \phi_2) = \begin{pmatrix} 2i \sin\phi_1 & -2i \sin\phi_1 - 4\sin^2(\phi_2/2) \\ -2i \sin\phi_1 + 4\sin^2(\phi_2/2) & 2i \sin\phi_1 \end{pmatrix}. \quad (11.4)$$

Поскольку $(A(\boldsymbol{\phi}))^* = -(A(\boldsymbol{\phi}))$, схема устойчива.

Как и в предыдущем примере, будем искать оператор $\mathring{\tilde{\Pi}}_h$ вида (9.4), доставляющий второй порядок аппроксимации. Имеем аппроксимационную ошибку

$$\left(\epsilon_1(v, \mathring{\tilde{\Pi}}_{1,\boldsymbol{e}})\right)_0 = \sum_{\boldsymbol{\eta}} L_{\boldsymbol{\eta}} \left(\frac{(\boldsymbol{e}\cdot\boldsymbol{\eta})^2}{2} \mathfrak{e} + (\boldsymbol{e}\cdot\boldsymbol{\eta}) \mathfrak{C}_{\boldsymbol{e}} \mathfrak{e} \right). \quad (11.5)$$

Подставляя в (11.5) значения коэффициентов $L_{\boldsymbol{\eta}}$, получаем

$$e_y^2 W \mathfrak{e} + 2e_x E |\boldsymbol{\alpha}| \mathfrak{C}_{\boldsymbol{e}} \mathfrak{e} = 0.$$

Подставляя далее коэффициенты W и E,

$$2e_x \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} (\mathfrak{C}_{\boldsymbol{e}} \mathfrak{e}) + e_y^2 \begin{pmatrix} -1 \\ 1 \end{pmatrix} = 0.$$

Видно, что при $e_x \neq 0$ система совместна и имеет решение

$$\mathfrak{C}_{\boldsymbol{e}} = -\frac{e_y^2}{2e_x} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + c_{\boldsymbol{e}} I,$$

тогда как при $e_x = 0$ система несовместна. Таким образом, для всех плоских волн с волновым вектором $\boldsymbol{\alpha}$, за исключением распространяющихся в горизонтальном направлении, схема обладает 2-м порядком точности в длительном счёте. Можно записать оценку

$$\|\varepsilon_h(t, e^{i\boldsymbol{\alpha}\cdot\boldsymbol{r}}, \mathring{\Pi}_h)\| \leqslant \min\left\{\frac{\alpha_y^2}{2|\alpha_x|}h + C_2|\boldsymbol{\alpha}|^3h^2t, C_3|\boldsymbol{\alpha}|^2ht\right\}.$$

Интересно также выписать для построенной схемы ошибку решения в явном виде, вычислив матричную экспоненту. Введём обозначения $a = 2\sin\phi_1$, $b = 4\sin^2(\phi_2/2)$, $m = \sqrt{a^2+b^2}$, $v = a + bi$. Тогда матрица $A(\boldsymbol{\phi})$, заданная (11.4) примет вид

$$A(\boldsymbol{\phi}) = \begin{pmatrix} ia & -ia-b \\ -ia+b & ia \end{pmatrix} = S\Lambda S^{-1}, \tag{11.6}$$

где

$$S = \begin{pmatrix} -1 & 1 \\ v/m & v/m \end{pmatrix}, \quad \Lambda = \begin{pmatrix} ia+im & 0 \\ 0 & ia-im \end{pmatrix}, \quad S^{-1} = \frac{1}{2}\begin{pmatrix} -1 & m/v \\ 1 & m/v \end{pmatrix}.$$

Матричная экспонента равна

$$\exp(\nu A(\boldsymbol{\phi})) = e^{i\nu a}\begin{pmatrix} \cos(\nu m) & -i\frac{m}{v}\sin(\nu m) \\ -i\frac{v}{m}\sin(\nu m) & \cos(\nu m) \end{pmatrix}. \tag{11.7}$$

Поскольку $A(\boldsymbol{\phi})$ была косоэрмитовой, матричная экспонента получилась унитарной матрицей (отметим, что $|v| = m$). Функция ошибки решения (3.5) имеет вид

$$\hat{\varepsilon}(\boldsymbol{\phi}, \nu, \mathring{\Pi}_h) = \left[e^{\nu A(\boldsymbol{\phi})} - I\right](\mathring{\Pi}_1 e^{i\boldsymbol{\phi}\cdot\boldsymbol{r}})_0. \tag{11.8}$$

Поскольку $A(0) = 0$, в силу утверждения 5.2 функция $\hat{\varepsilon}(\boldsymbol{\alpha}h, t/h, \mathring{\Pi}_h)$ является аналитической функцией $t, h, \boldsymbol{\alpha}$. То есть она представима в виде ряда по степеням h. Опуская выкладки ввиду их громоздкости (для этого нужно вернуться к "физическим" переменным h, $t = h\nu$, $\alpha_x = \phi_1/h$, $\alpha_y = \phi_2/h$), приведём первый член этого ряда:

$$\hat{\varepsilon}(\boldsymbol{\alpha}h, t/h, \mathring{\Pi}_h) = -\frac{\alpha_y^2}{4\alpha_x}(1 - \exp(-4t\alpha_x))h + O(h^2). \tag{11.9}$$

При любом фиксированном $\alpha_x \neq 0$ слагаемое при h является ограниченной функцией времени и в длительном счёте имеет место второй порядок точности. Покажем теперь, что при $\alpha_x = 0$ второй порядок точности в длительном счёте не имеет места. Из неограниченности по времени члена при h^1 в (11.9)

это, вероятно, не следует. Поэтому рассмотрим выражение для матричной экспоненты (11.7) целиком. Подставляя в него $a = 0$, $v = ib$, $m = b$, получаем

$$\exp(\nu A(\boldsymbol{\phi})) = \begin{pmatrix} \cos(\nu b) & -\sin(\nu b) \\ \sin(\nu b) & \cos(\nu b) \end{pmatrix},$$

откуда

$$\hat{\varepsilon}(\boldsymbol{\alpha} h, t/h, \mathring{\Pi}_h) = \begin{pmatrix} \cos(\nu b) - 1 - \sin(\nu b) \\ \cos(\nu b) - 1 + \sin(\nu b) \end{pmatrix},$$

где $\nu b = 4t \sin^2(\alpha_y h/2)/h$. Очевидно, что полученное выражение для ошибки решения оценкой $O(h + h^2 t)$ не обладает.

Вспомогательное утверждение

Утверждение 11.5. *Пусть $\boldsymbol{m}, \boldsymbol{n}_j \in \mathbb{R}^d_+$, $j = 1, \ldots, N$. Следующие два утверждения эквивалентны.*

1. $\exists C \ \forall \boldsymbol{\phi} \in [0,1]^d \ \ \boldsymbol{\phi}^{\boldsymbol{m}} \leqslant C \sum\limits_j \boldsymbol{\phi}^{\boldsymbol{n}_j}$;
2. $\exists \delta_j \geqslant 0 \ \ \sum \delta_j = 1, \ \sum \delta_j \boldsymbol{n}_j \leqslant \boldsymbol{m}$.

Действительно, пусть выполняется условие 2. Тогда при $\boldsymbol{\phi} \in [0,1]^d$ имеем

$$\boldsymbol{\phi}^{\boldsymbol{m}} \leqslant \boldsymbol{\phi}^{\sum \delta_j \boldsymbol{n}_j} = \prod_j (\boldsymbol{\phi}^{\boldsymbol{n}_j})^{\delta_j} \leqslant \sum_j \boldsymbol{\phi}^{\boldsymbol{n}_j}.$$

Последнее неравенство написано в силу неравенства Юнга. Таким образом, имеем условие 1.

Обратно, пусть условие 2 не выполняется. Пусть $M = \mathrm{Conv}\{\boldsymbol{n}_j\} + \mathbb{R}^d_+$, где плюсом обозначена сумма Минковского. Очевидно, M выпуклое и $\boldsymbol{m} \notin M$. По теореме отделимости найдётся такое $\boldsymbol{q} \in \mathbb{R}^d$, что для всех $\boldsymbol{a} \in M$ выполняется $\boldsymbol{q} \cdot \boldsymbol{a} > C > \boldsymbol{q} \cdot \boldsymbol{m}$. В частности, подставив $\boldsymbol{a} = \boldsymbol{n}_j$ и взяв минимум по j, получаем

$$\min_j \{\boldsymbol{q} \cdot \boldsymbol{n}_j\} > C > \boldsymbol{q} \cdot \boldsymbol{m}. \tag{11.10}$$

Покажем, что $\boldsymbol{q} \in \mathbb{R}^d_+$. Допустим противное, что какая-то компонента вектора $\boldsymbol{q}$ отрицательная, и обозначим её номер через k. Выберем любой $\boldsymbol{a} \in M$ и рассмотрим последовательность

$$\boldsymbol{a}_l = \boldsymbol{a} + l(0, \ldots, 1, \ldots, 0)^T,$$

где единица стоит в k-й позиции. По построению такой вектор лежит в M, но $\boldsymbol{q} \cdot \boldsymbol{a}_l \to -\infty$ и поэтому не может превосходить C для всех l. Следовательно, $\boldsymbol{q} \in \mathbb{R}^d_+$.

Допустим теперь, что выполняется условие 1. Подставим в него $\phi_1 = \varepsilon^{q_1}$, …, $\phi_d = \varepsilon^{q_d}$ и устремим ε к нулю. Сравнивая степени ε в левой и правой частях неравенства, получаем

$$\boldsymbol{q} \cdot \boldsymbol{m} \geqslant \min_j \{\boldsymbol{q} \cdot \boldsymbol{n}_j\},$$

что противоречит (11.10).

12. Странное особое направление

В рассмотренном выше примере для волн, распространяющихся вдоль особого направления, порядок точности в длительном счёте был на единицу меньше, чем для окрестных направлений. Но возможен и такой случай, когда направление $\boldsymbol{e}$ не является минус-особым, но при приближении к нему константы C_1 и/или C_2 в оценке вида (2.13) для $v_0(\boldsymbol{r}) = e^{i|\boldsymbol{\alpha}|\boldsymbol{e}_n \cdot \boldsymbol{r}}$ неограниченно растут.

Как и в предыдущих двух примерах, рассмотрим уравнение (2.3)–(2.4) в $\mathbb{R}^2$ со скоростью переноса $\boldsymbol{\omega} = 0$, и пусть $\boldsymbol{a}_1$ и $\boldsymbol{a}_2$ совпадают с единичными ортами. Положим $M^0 = \{1, 2, 3, 4, 5\}$ и определим матрицы

$$W = \begin{pmatrix} 0 & 1 & -1 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad E = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$R = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}, \quad Q = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 \end{pmatrix}.$$

Ненулевые коэффициенты схемы (2.5) выберем равными $Z_{0,0} = I$,

$$L_{0,0} = -2W + 6Q, \quad L_{\pm1,0} = W - 4Q \mp (2E - R), \quad L_{\pm2,0} = Q,$$
$$L_{\pm1,1} = L_{\pm1,-1} = \pm E.$$

Оператор $\mathring{\Pi}_h$ зададим равенствами

$$(\mathring{\Pi}_h f)_{\boldsymbol{\eta},\xi} = f\left(\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} h\right), \quad \xi = 1, \ldots, 4; \quad (\mathring{\Pi}_h f)_{\boldsymbol{\eta},5} = f\left(\begin{pmatrix} \eta_1 + 1 \\ \eta_2 \end{pmatrix} h\right).$$

Матрица $A(\boldsymbol{\phi}) = -L(\boldsymbol{\phi})$ имеет вид

$$\begin{aligned} A(\phi_1, \phi_2) = 4\sin^2\left(\frac{\phi_1}{2}\right) W + 8i \sin\phi_1 \sin^2\left(\frac{\phi_2}{2}\right) E - \\ -2i \sin\left(\frac{\phi_1}{2}\right) R - 16\sin^4\left(\frac{\phi_1}{2}\right) Q. \end{aligned} \tag{12.1}$$

Поскольку $(A(\boldsymbol{\phi}))^* = -(A(\boldsymbol{\phi}))$, схема устойчива.

Прежде всего, определим порядок аппроксимации. Имеем

$$\hat{\epsilon}(\boldsymbol{\phi}, \mathring{\Pi}_h) = A(\phi_1, \phi_2) \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ e^{i\phi_1} \end{pmatrix} = -2i \sin\left(\frac{\phi_1}{2}\right)(1 - e^{i\phi_1}) \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} + O(|\boldsymbol{\phi}|^3).$$

Таким образом, $\hat{\epsilon}(\boldsymbol{\phi}, \mathring{\Pi}_h)$ имеет второй порядок малости по $|\boldsymbol{\phi}|$ и, следовательно, схема имеет 1-й порядок аппроксимации. Далее, поскольку $A(0) = 0$, в силу следствия 5.2 схема обладает 1-м порядком точности и не обладает вторым.

Теперь займёмся исследованием этой схемы на порядок точности в длительном счёте. Уравнения для первых трёх компонент численного решения не связаны с уравнениями для последних двух компонент. Поэтому рассмотрим их в отдельности.

Вначале рассмотрим систему для первых трёх компонент сеточной функции. Соответствующее ограничение матрицы $A(\phi_1, \phi_2)$ имеет вид

$$A(\phi_1, \phi_2) = \begin{pmatrix} iy & x & -x \\ -x & 0 & x \\ x & -x & -iy \end{pmatrix}, \tag{12.2}$$

где $x = 4\sin^2(\phi_1/2)$, $y = 8\sin(\phi_1)\sin^2(\phi_2/2)$. Собственными числами матрицы $A(\phi_1, \phi_2)$ являются $\pm ir$ и 0, где $r = \sqrt{3x^2 + y^2}$. Будем считать, что $\sin\phi_1 \geqslant 0$; случай $\sin\phi_1 < 0$ может быть рассмотрен аналогично. Также исключим случай $\phi_1 = \phi_2 = 0$, где матрица нулевая. Матрица правых собственных векторов имеет вид

$$S = \begin{pmatrix} 3(a-i)(1+b) & 3a(a+i) & a \\ 3a(1+b+3ia) & (1+b)(1-b-3ia) & a-ib \\ -3a(2-b) & -(b+1)(b+2) & a \end{pmatrix},$$

где $a = x/r$, $b = y/r$. Детерминант этой матрицы равен $\det S = -18(b+1)(2ia+b)$ и не обращается в 0, поскольку $a \geqslant 0$, $b \geqslant 0$ и одновременно не обращаются в ноль. Обратная матрица имеет вид

$$S^{-1} = \frac{1}{\det S} \begin{pmatrix} -(1+b)(i-3a+2ib) & -3a(b+ia+1) & 3a(b+2ia) \\ 3a(2ib-3a-i) & -3a(ib-3a+i) & 3(b+1)(2ia+b) \\ -18a(b+1)(2ia+b) & 18(b+1)(ia^2+ab-i) & -18a(b+1)(2ia+b) \end{pmatrix}.$$

Если вычислить действие матрицы S^{-1} на образ начальных данных, получаем

$$S^{-1}\mathfrak{e} = \frac{1}{\det S} \begin{pmatrix} -3ib(1+ia+b) \\ 3b(1+3ia+b) \\ -18(b+1)(ia^2+ab-i) \end{pmatrix}.$$

Последняя компонента не вносит вклад в ошибку, так как соответствующее собственное значение равно 0. Первые две компоненты ведут себя одинаковым образом, поэтому достаточно рассмотреть одну, например первую. Имеем

$$(S^{-1}\mathfrak{e})_1 = \frac{ib(1+ia+b)}{6(b+1)(2ia+b)} = \frac{iy(r+ix+y)}{6(y+r)(2ix+y)} = \frac{iy}{6(2ix+y)} - \frac{xy}{6(y+r)(2ix+y)}.$$

Подставим выражения для x и y. Пренебрегая членами высоких порядков малости, получаем $x \approx \phi_1^2$ и $y \approx 2\phi_1\phi_2^2$. Таким образом,

$$(S^{-1}\mathfrak{e})_1 \approx \frac{i\phi_2^2}{6(i\phi_1+\phi_2^2)} - \frac{\phi_1\phi_2^2}{6(2\phi_2^2+\sqrt{3\phi_1^2+4\phi_2^4})(i\phi_1+\phi_2^2)}.$$

При любом линейном соотношении между ϕ_1 и ϕ_2, исключая $\phi_1 = 0$, величина $(S^{-1}\mathfrak{e})_1$ имеет первый порядок малости по $|\boldsymbol{\phi}|$. В то же время, если положить $\phi_1 = 0$, то $x = y = 0$, поэтому все три собственные значения $A(\phi_1, \phi_2)$ обращаются в ноль и, следовательно, ошибка решения тождественно равна нулю. Таким образом, какое бы мы ни взяли направление, для волн, распространяющихся по этому направлению, имеет место оценка ошибки $O(h)$, и можно говорить, что порядок точности в длительном счёте равен бесконечности. Однако константа в этой оценке существенным образом зависит от направления. Действительно, если положить, например, $\phi_1 = \phi_2^2$, то $(S^{-1}\mathfrak{e})_1 \to (i - 1/\sqrt{7})(1-i)/12$ при $\boldsymbol{\phi} = (\phi_2^2, \phi_2) \to 0$. Соответствующее собственное значение $\lambda_1(\phi_2^2, \phi_2) = ir \approx i\sqrt{7}\phi_2^4$. Таким образом, при $\boldsymbol{\phi} = (\phi_2^2, \phi_2)$ в проколотой окрестности нуля выполняется

$$\sup_{\nu \geqslant 0} \|\hat{\varepsilon}(\boldsymbol{\phi}, \nu, \mathring{\Pi}_h)\|_3 \geqslant C \sup_{\nu \geqslant 0} |e^{\nu\lambda_1(\boldsymbol{\phi})}(S^{-1}\mathfrak{e})_1| = 2C|(S^{-1}\mathfrak{e})_1| \geqslant \tilde{C} > 0.$$

Здесь $\| \cdot \|_{(3)}$ – полунорма на V_{per}, учитывающая только первые три компоненты решения в каждом сеточном блоке. Следовательно, оценки вида $\|\varepsilon_h(t, e^{i\boldsymbol{\alpha}\cdot\boldsymbol{r}}, \mathring{\Pi}_h)\| \leqslant Ch$, где C не зависит от t и h, для рассматриваемой схемы не существует.

Графики $\|\varepsilon_h(t, e^{i\boldsymbol{\alpha}\cdot\boldsymbol{r}}, \mathring{\Pi}_h)\|_{(3)}$ в зависимости от времени на сетках с $h = 1/60$, $h = 1/120$, $h = 1/240$ при $\alpha_x = 4\pi$, $\alpha_y = -2\pi$ приведены на рис. 2.

Рассмотрим теперь эту подсистему методом введения вспомогательного отображения. Понимая под W и E ограничения соответствующих операторов на трёхмерное векторное подпространство, имеем

$$\sum_{\boldsymbol{\eta}} L_{\boldsymbol{\eta}} = 0, \quad \sum_{\boldsymbol{\eta}} L_{\boldsymbol{\eta}}(\boldsymbol{e}\cdot\boldsymbol{\eta}) = 0, \tag{12.3}$$

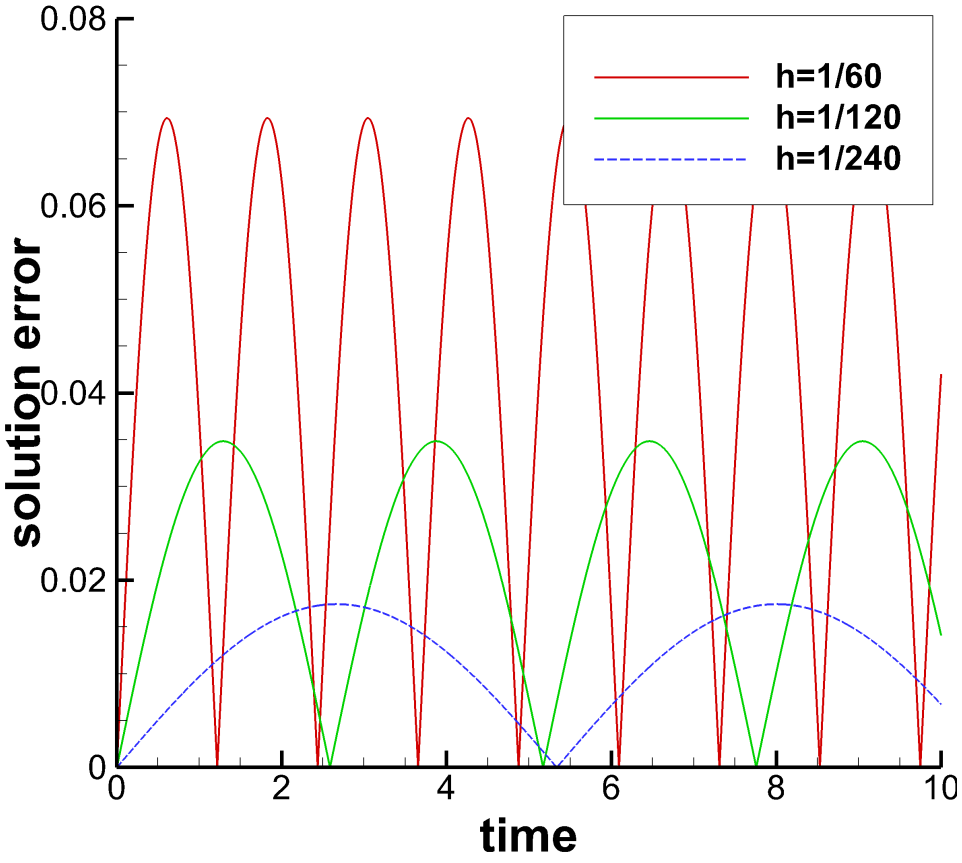

Рис. 2. Зависимость ошибки решения от времени

$$\sum_{\eta} L_{\eta} \frac{(\boldsymbol{e}\cdot\boldsymbol{\eta})^2}{2} = e_x^2 W, \quad \sum_{\eta} L_{\eta} \frac{(\boldsymbol{e}\cdot\boldsymbol{\eta})^3}{6} = \sum_{\eta} L_{\eta} \frac{(e_x\eta_1 + e_y\eta_2)^3}{6} = 2e_x e_y^2 E.$$

$$\sum_{\eta} L_{\eta} \frac{(\boldsymbol{e}\cdot\boldsymbol{\eta})^4}{24} = \sum_{\eta} L_{\eta} \frac{(e_x\eta_1 + e_y\eta_2)^4}{24} = \frac{e_x^4}{12} W. \qquad (12.4)$$

Поскольку

$$\sum_{\eta} L_{\eta}\mathfrak{e} = 0, \quad \sum_{\eta} L_{\eta}\boldsymbol{\eta}\mathfrak{e} = 0, \quad \sum_{\eta} L_{\eta}\boldsymbol{\eta}\otimes\boldsymbol{\eta}\mathfrak{e} = 0,$$

порядок аппроксимации в смысле $\mathring{\Pi}_h$ равен 2.

Выпишем уравнение на нахождение коэффициентов оператора $\mathring{\tilde{\Pi}}_{h,\boldsymbol{e}}$, в смысле которого может быть обеспечен 3-й порядок аппроксимации на плоских волнах, распространяющихся по направлению $\boldsymbol{e}$. Поскольку выполняется (12.3), в уравнение будут входить только коэффициенты 1-го порядка $\mathfrak{C}_{\boldsymbol{e}}$, но не

2-го и не 3-го. Имеем

$$\left(\epsilon_1((\boldsymbol{e}\cdot\boldsymbol{r})^3/6, \mathring{\tilde{\Pi}}_{1,\boldsymbol{e}})\right)_0 = \sum_{\boldsymbol{\eta}} L_{\boldsymbol{\eta}} \left(\frac{(\boldsymbol{e}\cdot\boldsymbol{\eta})^3}{6}\mathfrak{e} + \frac{(\boldsymbol{e}\cdot\boldsymbol{\eta})^2}{2}\mathfrak{C}_{\boldsymbol{e}}\mathfrak{e}\right) = \\ = 2e_x e_y^2 E\mathfrak{e} + e_x^2 W\mathfrak{C}_{\boldsymbol{e}}\mathfrak{e}.$$

Приравнивая к нулю, получаем

$$e_x \left[e_x \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} (\mathfrak{C}_{\boldsymbol{e}}\mathfrak{e}) + 2e_y^2 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right] = 0.$$

Видно, что при всех $\boldsymbol{e}$ система совместна. При $e_x \neq 0$ она имеет решение

$$\mathfrak{C}_{\boldsymbol{e}} = -2\frac{e_y^2}{e_x} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + c_{\boldsymbol{e}} I, \tag{12.5}$$

тогда как при $e_x = 0$ ей удовлетворяют любые значения $\mathfrak{C}_{\boldsymbol{e}}$. Положим $c_{\boldsymbol{e}} = 0$. Рассмотрим аппроксимационную ошибку на полиноме 4-го порядка в смысле полученного оператора $\mathring{\tilde{\Pi}}_{h,\boldsymbol{e}}$:

$$\left(\epsilon_1 \left(\frac{(\boldsymbol{e}\cdot\boldsymbol{r})^4}{24}, \mathring{\tilde{\Pi}}_{1,\boldsymbol{e}}\right)\right)_0 = \sum_{\boldsymbol{\eta}} L_{\boldsymbol{\eta}} \left(\frac{(\boldsymbol{e}\cdot\boldsymbol{\eta})^4}{24}\mathfrak{e} + \frac{(\boldsymbol{e}\cdot\boldsymbol{\eta})^3}{6}\mathfrak{C}_{\boldsymbol{e}}\mathfrak{e}\right).$$

Используя (12.4) и заметив, что $W\mathfrak{e} = 0$ и $E\mathfrak{C}_{\boldsymbol{e}}\mathfrak{e} = 0$, получаем, что $(\epsilon_1((\boldsymbol{e}\cdot\boldsymbol{r})^4, \tilde{\Pi}_{1,\boldsymbol{e}}))_0 = 0$, то есть в смысле отображения $\tilde{\Pi}_{h,\boldsymbol{e}}$ схема обладает 4-м порядком аппроксимации на волнах, распространяющихся по $\boldsymbol{e}$.

Таким образом, какое бы ни было выбрано направление, для плоских волн, распространяющихся в этом направлении, можно записать оценку ошибки

$$\|\varepsilon_h(t, e^{i\boldsymbol{\alpha}\cdot\boldsymbol{r}}, \mathring{\Pi}_h)\|_{(3)} \leqslant C_1\left(\frac{\boldsymbol{\alpha}}{|\boldsymbol{\alpha}|}\right) h|\boldsymbol{\alpha}| + C_2 h^4(t+h)|\boldsymbol{\alpha}|^5.$$

Коэффициент $C_1(\,\cdot\,)$ не является ограниченной функцией на единичной окружности. Коэффициент C_2 же на единичной окружности ограничен, поскольку множитель $1/e_x = |\boldsymbol{\phi}|/\phi_1$ в (12.5) компенсируется тем, что $A(\phi_1, \phi_2) = O(\phi_1)$ при $\phi_1 \to 0$.

Проведённые рассуждения не дали ответ на вопрос, каково оптимальное значение порядка точности в длительном счёте при формальном порядке точности, равном 1. Получим этот ответ при помощи утверждения 10.4. Пусть матрица A определена (12.2). Выпишем критерий (10.2) с подстановкой $C_1 = |\boldsymbol{\phi}|$,

$C_2 = |\boldsymbol{\phi}|^{Q+1}$, $v = \mathfrak{e}$. Имеем

$$\exists\, \tilde{\delta} \;:\quad \mathfrak{e}^*(A^*A + |\boldsymbol{\phi}|^{2Q})^{-1}A^*A\mathfrak{e} \leqslant \tilde{\delta}|\boldsymbol{\phi}|^2\|\mathfrak{e}\|^2.$$

Подставляя выражение для матрицы A и упрощая, получаем

$$\exists\, c \;:\quad 2y^2 \leqslant c|\boldsymbol{\phi}|^2(3x^2 + y^2 + |\boldsymbol{\phi}|^{2Q}).$$

Подставляя выражения для x и y, отбрасывая члены, заведомо мажорируемые соседними, и опуская несущественные коэффициенты, получаем

$$\exists\, c \;:\quad \phi_1^2\phi_2^4 \leqslant c(\phi_1^2 + \phi_2^2)(\phi_1^4 + \phi_1^2\phi_2^4 + (\phi_1^2 + \phi_2^2)^Q).$$

Выражение $(x+y)^Q$ для неотрицательных x и y эквивалентно $x^Q + y^Q$, поэтому можно переписать критерий в виде

$$\exists\, c \;:\quad \phi_1^2\phi_2^4 \leqslant c(\phi_1^2 + \phi_2^2)(\phi_1^4 + \phi_1^2\phi_2^4 + \phi_1^{2Q} + \phi_2^{2Q}). \tag{12.6}$$

Нас интересует наибольшее значение Q, при котором это условие выполняется. При $Q = 2$ это верно: раскрыв скобки в правой части, можно увидеть в ней слагаемое $\phi_1^2\phi_2^4$, совпадающее с левой частью. Покажем при помощи утверждения 11.5, что неравенство (12.6) не выполняется при $Q > 2$. Действительно, в левой части неравенства находится $\boldsymbol{\phi}^{\boldsymbol{m}}$, где $\boldsymbol{m} = (2{,}4)$, а в его правой части находятся слагаемые

$$\boldsymbol{n}_1 = (6,0), \quad \boldsymbol{n}_2 = (4,2), \quad \boldsymbol{n}_3 = (2,2Q), \quad \boldsymbol{n}_4 = (2,6), \quad \boldsymbol{n}_5 = (0,2+2Q).$$

Остальные слагаемые мажорируются перечисленными. Изобразив точку $\boldsymbol{m}$ и точки $\boldsymbol{n}_j$ на координатной плоскости, можно заметить, что при всех $Q > 2$ точка $\boldsymbol{m}$ лежит ниже выпуклой оболочки точек $\boldsymbol{n}$. Следовательно, неравенство (12.6) не выполняется.

Таким образом, оптимальное значение порядка точности в длительном счёте при 1-м формальном порядке равно 2. Можно записать оценку

$$\|\varepsilon_h(t, e^{i\boldsymbol{\alpha}\cdot\boldsymbol{r}}, \mathring{\Pi}_h)\|_{(3)} \leqslant \min\left\{ C_1\left(\frac{\boldsymbol{\alpha}}{|\boldsymbol{\alpha}|}\right)h|\boldsymbol{\alpha}| + C_2h^4(t+h)|\boldsymbol{\alpha}|^5, Ch^2(t+h)|\boldsymbol{\alpha}|^2 \right\}.$$

Теперь перейдём к рассмотрению системы для последних двух компонент сеточной функции. Блок матрицы $A(\boldsymbol{\phi})$, определяющий поведение последних двух компонент решения, зависит только от ϕ_1. Этот блок имеет вид (11.6), где $a = 2\sin(\phi_1/2)$, $b = -16\sin^4(\phi_1/2)$. Выпишем выражение (11.8) для функции численной ошибки. Имеем

$$\hat{\varepsilon}(\boldsymbol{\phi}, \nu, \mathring{\Pi}_h) = S(e^{\nu\Lambda} - I)S^{-1}(\mathring{\Pi}_1 \exp(i\boldsymbol{\phi}\cdot\boldsymbol{r}))_0 =$$
$$= S\begin{pmatrix} e^{\nu(ia+im)} - 1 & 0 \\ 0 & e^{\nu(ia-im)} - 1 \end{pmatrix}\frac{1}{2}\begin{pmatrix} -1 & m/v \\ 1 & m/v \end{pmatrix}\begin{pmatrix} 1 \\ \exp(i\alpha_1 h) \end{pmatrix},$$

где $\nu = t/h$, $v = a + ib$, $\boldsymbol{\phi} = \boldsymbol{\alpha} h$, $m = |v|$. Поскольку S унитарная, домножение на S не оказывает влияния на норму ошибки. Вычислим произведение последних двух сомножителей. Имеем

$$2S^{-1}(\mathring{\Pi}_h \exp(i\boldsymbol{\alpha} \cdot \boldsymbol{r}))_0 = \begin{pmatrix} -1 + e^{i\alpha_1 h} m/v \\ 1 + e^{i\alpha_1 h} m/v \end{pmatrix}. \tag{12.7}$$

Далее без ограничения общности будем считать, что $\phi_1 > 0$; если $\phi_1 < 0$, то все рассуждения проводятся аналогично с тем отличием, что две компоненты (12.7) обмениваются свойствами. При $\phi_1 > 0$ имеем $a/m = v/m = 1 + O(h^2)$, поэтому в некоторой окрестности $h = 0$ первая компонента (12.7) по модулю не превосходит $2\alpha_1 h$. При этом $|\exp(\nu(ia + im)) - 1| \leqslant 2$, следовательно, произведение не превосходит $4\alpha_1 h$. Вторая компонента (12.7) по модулю не превосходит 2, а

$$\begin{aligned} |\exp(\nu(ia - im)) - 1| = |\exp(it(a-m)/h) - 1| \leqslant t|a-m|/h = \\ = \frac{t}{h}\left|a - \sqrt{a^2+b^2}\right| = \frac{t}{h}\frac{b^2}{a+\sqrt{a^2+b^2}} \leqslant \frac{t}{h}\frac{b^2}{a} = \\ = \frac{t}{h}\frac{256\sin^8(\phi_1/2)}{2\sin(\phi_1/2)} = 128\frac{t}{h}\sin^7(\alpha_1 h/2) \leqslant t h^6. \end{aligned}$$

Таким образом, 4-я и 5-я компоненты ошибки удовлетворяют оценке $O(h+th^6)$.

Складывая оценки ошибки для первых трёх и последних двух компонент решения, получаем оценку ошибки вида $O(h+th^4)$, зависящую от направления. Одновременно с этим, оценивая ошибку по первым трём компонентам напрямую исходя из свойств аппроксимации и устойчивости, получаем суммарную оценку $O(h + th^2)$, от направления не зависящую.

Разобранный пример показывает возможность такого случая, что на всех решениях в виде плоских волн имеется порядок точности P и порядок точности в длительном счёте Q, но при этом на $H^{Q+1}_{per}(\mathbb{R}^d)$ порядок точности в длительном счёте Q не имеет места.

13. Зависимость от скорости переноса

Выше мы полагали скорость переноса $\boldsymbol{\omega}$ в уравнении (2.3) и схему (2.5) для аппроксимации этого уравнения заданными. Теперь зададимся вопросом, как свойства схемы зависят от $\boldsymbol{\omega}$ (по-прежнему полагаем $\boldsymbol{\omega}$ постоянным во времени и пространстве). При этом будем считать, что коэффициенты схемы линейно зависят от вектора $\boldsymbol{\omega}$, и, следовательно, рассматривать схемы вида

$$\sum_{\boldsymbol{\zeta}\in\mathcal{S}} Z_{\boldsymbol{\zeta}} \frac{du_{\boldsymbol{\eta}+\boldsymbol{\zeta}}}{dt}(t) + \frac{1}{h}\sum_{l=1}^{d}\omega_l \sum_{\boldsymbol{\zeta}\in\mathcal{S}} L^{(l)}_{\boldsymbol{\zeta}} u_{\boldsymbol{\eta}+\boldsymbol{\zeta}}(t) = 0. \tag{13.1}$$

Можно показать, что схема вида (13.1), устойчивая для всех $\boldsymbol{\omega} \in \mathbb{R}^d$, удовлетворяет условию $\sum\limits_{\boldsymbol{\eta}} L_{\boldsymbol{\zeta}}^{(l)} = 0$. В силу следствия 5.2, для такой схемы порядок точности не может превосходить порядок аппроксимации. В большинстве же используемых на практике схем используется информация о направлении характеристики, что делает зависимость схемы от $\boldsymbol{\omega}$ кусочно-линейной. Коэффициенты $L_{\boldsymbol{\zeta}}$ в (2.5) линейно зависят не только от компонент вектора $\boldsymbol{\omega}$, но и от величин вида $|\boldsymbol{\omega} \cdot \boldsymbol{n}_m|$, где $\boldsymbol{n}_m$ – некоторые векторы. При каждом фиксированном направлении $\boldsymbol{\omega}/|\boldsymbol{\omega}|$ зависимость коэффициентов от $|\boldsymbol{\omega}|$ остаётся линейной.

Легко построить схему вида (13.1), такую что для некоторых направлений переноса имеют место более высокие порядки аппроксимации и точности, чем для остальных направлений. Для этого достаточно взять равномерную сетку, размер блока $|M^0| = 1$ и записать конечно-разностную схему, аппроксимировав производную по x с порядком Q_x, а производную по y – с порядком $Q_y > Q_x$. Тогда для всех направлений, кроме совпадающего с осью OY и противоположному ему, схема будет обладать порядком аппроксимации и порядком точности Q_x, а для двух выделенных направлений – порядком Q_y.

Более интересно, что возможна и обратная ситуация: для некоторых выделенных направлений порядок может быть меньшим, чем для остальных. Такой эффект ранее упоминался для метода Галёркина с разрывными базисными функциями, но не в точности схемы, а лишь в её оценках (см [8]), не являющихся оптимальными. В то же время этот эффект нами наблюдался на практике при оценке точности метода спектральных разностей [9]. Приведём простой пример, иллюстрирующий такую возможность.

Рассмотрим уравнение переноса (2.3) при $d = 2$ со скоростью переноса $(\omega_x, \omega_y)^T$, где $\omega_x, \omega_y \geqslant 0$. Пусть векторы $\boldsymbol{a}_j$ совпадают с единичными ортами. Положим $M^0 = \{L, R\}$ и определим отображение

$$(\Pi_h f)_{\boldsymbol{\eta},L} = f(h\eta_x, h(\eta_y + 1)), \quad (\Pi_h f)_{\boldsymbol{\eta},R} = f(h(\eta_x + 1), h\eta_y).$$

Далее блоки будем нумеровать привычными индексами $i \equiv \eta_x$, $j \equiv \eta_y$. Производную для степени свободы R будем определять центральной разностной производной:

$$\frac{du_{i,j}^R}{dt} + \omega_x \frac{u_{i+1,j}^R - u_{i-1,j}^R}{2h} + \omega_y \frac{u_{i,j+1}^R - u_{i,j-1}^R}{h} = 0.$$

Производную для степени свободы L будем определять при помощи направленных разностей. Воспользуемся тем, что точка, где определяется значение L на ячейке $i, j-1$, совпадает с точкой определения значения R на ячейке $i-1, j$, и "завяжем" эти компоненты друг на друга:

$$\frac{du_{i,j}^L}{dt} + \omega_x \frac{u_{i,j}^L - u_{i-1,j}^L}{h} + \omega_y \frac{u_{i,j}^L - u_{i-1,j}^R}{2h} = 0.$$

В матричном виде эта схема примет вид (2.5), где

$$L_{-1,0} = \begin{pmatrix} -\omega_x & -\omega_y \\ 0 & -\omega_x/2 \end{pmatrix}, \quad L_{0,0} = \begin{pmatrix} \omega_x + \omega_y & 0 \\ 0 & 0 \end{pmatrix},$$
$$L_{1,0} = \begin{pmatrix} 0 & 0 \\ 0 & \omega_x/2 \end{pmatrix}, \quad L_{0,1} = \begin{pmatrix} 0 & 0 \\ 0 & \omega_y/2 \end{pmatrix}, \quad L_{0,-1} = \begin{pmatrix} 0 & 0 \\ 0 & -\omega_y/2 \end{pmatrix}.$$

Если обозначить $\boldsymbol{\phi} = (\phi, \psi)$, то матрица $L(\boldsymbol{\phi})$ примет вид

$$L(\boldsymbol{\phi}) = \begin{pmatrix} \omega_x(1 - e^{-i\phi}) + \omega_y & -\omega_y e^{-i\phi} \\ 0 & i(\omega_x \sin\phi + \omega_y \sin\psi) \end{pmatrix},$$

и $A(\boldsymbol{\phi}) = i(\omega_x\phi + \omega_y\psi) - L(\boldsymbol{\phi})$.

Покажем, что эта схема устойчива при $\omega_x, \omega_y \geqslant 0$, то есть что $\| \exp(\nu A(\boldsymbol{\phi}))\|$ ограничено по $\nu \geqslant 0$ равномерно по $\boldsymbol{\phi}$. Матрица $A(\boldsymbol{\phi})$ имеет вид

$$A(\boldsymbol{\phi}) = \begin{pmatrix} \lambda_1(\boldsymbol{\phi}) & a(\boldsymbol{\phi}) \\ 0 & \lambda_2(\boldsymbol{\phi}) \end{pmatrix},$$

при

$$\lambda_1(\boldsymbol{\phi}) = i(\omega_x\phi + \omega_y\psi) - (\omega_x(1 - e^{-i\phi}) + \omega_y),$$
$$\lambda_2(\boldsymbol{\phi}) = i(\omega_x\phi + \omega_y\psi) - i(\omega_x \sin\phi + \omega_y \sin\psi),$$

$a(\boldsymbol{\phi}) = -\omega_y e^{-i\phi}$. Можно показать, что

$$e^{\nu A(\phi)} = \exp\left[\nu \begin{pmatrix} \lambda_1(\boldsymbol{\phi}) & a(\boldsymbol{\phi}) \\ 0 & \lambda_2(\boldsymbol{\phi}) \end{pmatrix}\right] = \begin{pmatrix} e^{\nu\lambda_1(\boldsymbol{\phi})} & \frac{a(\boldsymbol{\phi})}{\lambda_2(\boldsymbol{\phi}) - \lambda_1(\boldsymbol{\phi})}\left[e^{\nu\lambda_2(\boldsymbol{\phi})} - e^{\nu\lambda_1(\boldsymbol{\phi})}\right] \\ 0 & e^{\nu\lambda_2(\boldsymbol{\phi})} \end{pmatrix}.$$

При $\omega_x \geqslant 0$, $\omega_y \geqslant 0$ верно $\operatorname{Re}\lambda_2(\boldsymbol{\phi}) = 0$ и $\operatorname{Re}\lambda_1(\boldsymbol{\phi}) \leqslant 0$, поэтому диагональные компоненты $e^{\nu A(\phi)}$ ограничены. Величина в квадратных скобках при больших ν стремится к единице и, следовательно, тоже ограничена. Осталось показать, что ограниченной является величина

$$Y(\boldsymbol{\phi}) = a(\boldsymbol{\phi})/(\lambda_2(\boldsymbol{\phi}) - \lambda_1(\boldsymbol{\phi})).$$

Для проверки этого условия запишем

$$|Y(\boldsymbol{\phi})|^2 = \left|\frac{\omega_y}{i(\omega_x \sin\phi + \omega_y \sin\psi) - (\omega_x(1 - e^{-i\phi}) + \omega_y)}\right|^2 =$$

$$= \frac{\omega_y^2}{(\omega_x(1 - \cos\phi) + \omega_y)^2 + \omega_y^2 \sin^2\psi} \leqslant 1,$$

поскольку в силу неотрицательности ω_x и ω_y минимум знаменателя достигается при $\phi = \psi = 0$. Таким образом, записанная схема при $\omega_x \geqslant 0$, $\omega_y \geqslant 0$ является устойчивой.

Выпишем теперь ошибку аппроксимации. Подставим в схему (2.5) функции $x^2/2$, xy и $y^2/2$. Поскольку для компонент R используется схема 2-го порядка, производная на них точна и ошибка аппроксимации равна 0. Для компоненты L вычислим её явно. Для $f^{xx} \equiv -(\epsilon_1(x^2/2, \Pi_1))_0$, $f^{xy} \equiv -(\epsilon_1(xy, \Pi_1))_0$, $f^{yy} \equiv -(\epsilon_1(y^2/2, \Pi_1))_0$, получим

$$f^{xx} = h \begin{pmatrix} \omega_x/2 \\ 0 \end{pmatrix}, \quad f^{xy} = h \begin{pmatrix} -\omega_x/2 \\ 0 \end{pmatrix}, \quad f^{yy} = h \begin{pmatrix} -\omega_y/2 \\ 0 \end{pmatrix}.$$

Получим теперь уравнение на корректор. Заметим, что

$$L(0) = \begin{pmatrix} \omega_y & -\omega_y \\ 0 & 0 \end{pmatrix}.$$

При $\omega_y \neq 0$ матрица $L(0)$ имеет два собственных значения: 0 и $\omega_y \neq 0$. Поскольку нулевое собственное значение является простым, диагональные матрицы $\mathfrak{C}^{(m)}$ отображения $\tilde{\Pi}_h$ определяются однозначно с точностью до cI, где $c \in \mathbb{R}$, а I – единичная матрица. Отсюда следует, что $\mathfrak{C}^{(\boldsymbol{m})}$, $|\boldsymbol{m}| = 1$, можно полагать равными нулю. Коэффициенты второго порядка находятся из соотношений

$$L(0)\mathfrak{C}^{xx} = f^{xx}, \quad L(0)\mathfrak{C}^{xy} = f^{xy}, \quad L(0)\mathfrak{C}^{yy} = f^{yy}.$$

Все эти системы, очевидно, совместны. Таким образом, отображение $\tilde{\Pi}_h$ построено, что доказывает 2-й порядок точности рассматриваемой схемы (при $\omega_x \geqslant 0$, $\omega_y > 0$).

Если же $\omega_y = 0$, то $L(0) = 0$. Применяя следствие 5.2, получаем, что схема не может обладать порядком точности, превосходящим порядок аппроксимации, то есть её порядок точности равен 1.

Таким образом, построен пример схемы для уравнения (2.3) при $\omega_x \geqslant 0$, $\omega_y \geqslant 0$, такой что при $\omega_y \neq 0$ она обладает 1-м порядком аппроксимации и 2-м порядком точности, а при $\omega_y = 0$ – 1-м порядком аппроксимации и 1-м порядком точности.

14. Нерёшенные задачи

Мы привели ряд примеров, показывающих, что в многомерном случае схемы могут обладать неожиданным поведением ошибки или странным образом откликаться на попытки исследовать это поведение методом введения вспомогательного отображения. Тем не менее, несколько вопросов пока остались неотвеченными. Приведём некоторые из них.

1. Может ли оптимальное значение порядка точности в длительном счёте для схемы вида (2.5) в смысле $\mathring{\Pi}_h$ вида (2.9) не быть целым числом?
2. Пусть схема обладает порядком точности P и порядком точности в длительном счёте Q. Может ли существовать такое направление $\boldsymbol{e}_\infty$, что при $\boldsymbol{e} \to \boldsymbol{e}_\infty$ коэффициенты $\mathfrak{C}_{\boldsymbol{e}}^{(Q)}$, вычисленные исходя из условия Q-го порядка аппроксимации на волнах вида $f(\boldsymbol{r} \cdot \boldsymbol{e})$, неограниченно растут?
3. Множество направлений, по которым могут распространяться волны, не покрывает всю единичную сферу. Например, при смещениях $\boldsymbol{a}_1 = (1,0)^T$, $\boldsymbol{a}_2 = (0,1)^T$ справедливо $|e_y|/|e_x| \leqslant N_0/h$, где N_0 – период решения. Каковы могут быть оценки ошибки, учитывающие зависимость от N_0?
4. Рассмотрим разрывный метод Галёркина на тетраэдральной трансляционно-инвариантной сетке и предположим, что он обладает оценкой $O(h^P + th^Q)$. Вопрос: найдётся ли оператор вида (2.10), доставляющий Q-й порядок аппроксимации? Если вектор переноса совпадает с вектором одного из рёбер сетки, имеем квазиодномерный случай, и ответ положительный. Если он не лежит ни в одной из плоскостей граней сетки, имеем простой случай, и ответ также положительный. А если вектор переноса лежит вдоль грани, но не вдоль ребра?
5. И, наконец, главный вопрос, мотивировавший авторов на проведение настоящего исследования: существует ли метод, позволяющий установить порядок точности в длительном счёте для заданной схемы?

15. Заключение

Было показано, что для схемы, обдалающей p-м порядком точности и q-м порядком точности в длительном счёте, найдётся отображение, отличающееся от исходного на величину $O(h^p)$ и доставляющее q-й порядок аппроксимации. Однако это отображение, вообще говоря, нелокально; были доказаны два достаточных условия, при которых оно является локальным. Был приведён ряд искусственных примеров схем, демонстрирующих неожиданное поведение в длительном счёте.

Список литературы

1. Бахвалов П. А., Сурначёв М. Д. О спектральном анализе схем для линейного уравнения переноса // Препринты ИПМ им. М.В.Келдыша. 2019. № 70. 28 с.

2. Бахвалов П. А., Сурначёв М. Д. Об аналитических семействах матриц, порождающих ограниченные полугруппы // Сибирский журнал вычислительной математики. (представлено в редакцию).

3. Бахвалов П. А., Сурначёв М. Д. Линейные схемы с несколькими степенями свободы для одномерного уравнения переноса // Препринты ИПМ им. М.В.Келдыша. 2019. № 73. 40 с.

4. Бахвалов П. А., Сурначёв М. Д. О приведении устойчивых матриц к блочно-диагональному виду // Препринты ИПМ им. М.В.Келдыша. 2019. №71. 15 с.

5. Cao W., Zhang Z., Zou Q. Superconvergence of discontinuous Galerkin methods for linear hyperbolic equations // SIAM Journal on Numerical Analysis. 2014. Vol. 52, no. 5. P. 2555–2573.

6. Superconvergence of discontinuous Galerkin methods for two-dimensional hyperbolic equations / Cao W., Shu C.-W., Yang Y. et al. // SIAM Journal on Numerical Analysis. 2015. Vol. 53, no. 4. P. 1651–1671.

7. Kato T. Perturbation theory for linear operators. Grund. math. Wiss., B. 132, Springer, 1966. P. XIX, 592.

8. Richter G. An optimal-order error estimate for the discontinious Galerkin method // Mathematics of Computation. 1988. Vol. 50. P. 75–88.

9. Balan A., May G., Schöberl J. A stable high-order Spectral Difference method for hyperbolic conservation laws on triangular elements // J. Comput. Phys. 2012. Vol. 231. P. 2359–2375.

Оглавление

1 Введение 3
2 Постановка задачи 4
3 Спектральное представление схемы 7
4 Общие замечания 10
5 Определение формального порядка 15
6 Существование отображения 16
7 Простой случай 18
8 Квазиодномерный случай 20
9 Контрпример 22
10 Особые направления волнового вектора 27
11 Минус-особое направление по точности в длительном счёте 30
12 Странное особое направление 33
13 Зависимость от скорости переноса 39
14 Нерёшенные задачи 42
15 Заключение 43
Список литературы 43