УДК 519.63; 004.75; 536.252

С. В. Сирик (Нац. техн. ун-т Украины „КПИ”, Киев)

ОЦЕНКИ ТОЧНОСТИ КОНЕЧНОЭЛЕМЕНТНОГО МЕТОДА ПЕТРОВА – ГАЛЕРКИНА ПРИ ИНТЕГРИРОВАНИИ ОДНОМЕРНОГО СТАЦИОНАРНОГО УРАВНЕНИЯ КОНВЕКЦИИ-ДИФФУЗИИ-РЕАКЦИИ

The accuracy and convergence of the numerical solutions of the stationary one-dimensional linear convection-diffusion-reaction equation (with Dirichlet boundary conditions) by the Petrov – Galerkin finite-element method with piecewise-linear basis functions and piecewise-quadratic weighting functions are analyzed and the error estimates of the method are obtained in some norms depending on the choice of a collection of stabilization parameters.

Проаналізовано питання точності та збіжності числового розв’язку стаціонарного одновимірного лінійного рівняння конвекції-дифузії-реакції (з граничними умовами типу Діріхле) скінченноелементним методом Петрова – Гальоркіна з кусково-лінійними базисними та кусково-квадратичними ваговими функціями. Отримано оцінки точності методу в кількох нормах залежно від вибору набору стабілізуючих параметрів вагових функцій.

Введение. В настоящее время конечноэлементный метод Петрова – Галеркина (МПГ) является одним из наиболее успешных подходов к построению схем численного решения разнообразных задач математической физики, в особенности, задач моделирования процессов конвекции-диффузии-реакции (КДР) [1 – 5]. Отметим, что к этому классу относится большинство процессов, рассматриваемых в гидродинамике и магнитной гидродинамике [4, 5], а также встречающихся в химической промышленности [1, 2]. В ряде работ (см. [1 – 3, 6 – 9], а также обзор в [10]), посвященных численному решению задач конвекции-диффузии (КД), успешно применялись весовые (поверочные) функции МПГ вида

$$W_i(x) = N_i(x) + \alpha_i W_i^*(x), \tag{1}$$

где $N_i(x)$ — базисная (пробная) функция МПГ, соответствующая узлу сетки с индексом i, α_i — некоторый настроечный параметр (стабилизирующий параметр), а $W_i^*(x)$ выбирается таким образом, чтобы обеспечить стабилизирующий эффект (избавиться от ложных, не имеющих физического смысла осцилляций) в получаемых численных аппроксимациях.

Оценки точности МПГ с кусочно-линейными базисными и кусочно-квадратичными весовыми функциями вида (1) при решении одномерных линейных стационарных задач КД с граничными условиями первого рода [4] впервые были получены Гриффитсом и Лоренцом в [11], при этом коэффициенты уравнения КД предполагались постоянными и считалось, что $\alpha_i = \alpha = \text{const} \geq 0$ в (1) независимо от i (см. также [3, 10, 12, 13] по поводу обзоров данного случая и дальнейших исследований). Соответствующий случай с неодинаковыми параметрами α_i (зависящими от i) в (1) труднее для исследования и, возможно, поэтому намного меньше изучен (см. [3, 6, 10]). В статье [10] исследованы вопросы точности, устойчивости и сходимости МПГ при решении стационарных одномерных уравнений КД в зависимости от выбора параметров $\{\alpha_i\}$, обобщающие и уточняющие некоторые более ранние результаты в данном направлении (приведенные, в том числе, в [3, 6, 11 – 13]). В настоящей работе результаты [10] обобщены и распространены на случай стационарных одномерных уравнений КДР (в том числе, как отдельный частный случай, получены новые оценки точности МПГ для уравнений КД, уточняющие оценки из [10]). Конкретнее говоря, в работе поставлены следующие цели: 1) получение оценок констант Ладыженской – Бабушки – Брецци для МПГ при решении уравнения КДР (соответствующие определения приведены в п. 2); 2) получение оценок точности МПГ в пространствах H^1 и L_2 для уравнения КДР (основываясь на использовании результатов предыдущего пункта, теории соболевских пространств и теории сплайнов [14 – 18], а также общей теории конечноэлементных МПГ [19]). Достижение первой цели базируется на использовании подхода, предложенного изначально в [10] и существенно развиваемого в данной работе.

1. Постановка задачи, существование и единственность решения слабой формы краевой задачи. Рассмотрим краевую задачу для стационарного уравнения КДР

$$Lu \equiv k\frac{du}{dx} - \frac{d^2u}{dx^2} + cu = f \tag{2}$$

на отрезке $x \in [0;\, 1]$ ($u(x)$ является искомой неизвестной функцией, а коэффициенты k, c и $f(x)$ заданы) с однородными граничными условиями 1-го рода: $u(0) = u(1) = 0$ (неоднородные граничные условия можно привести к однородным [4]). Будем считать, что правая часть уравнения (2), т. е. f, принадлежит $L_2[0;\, 1]$, а коэффициенты k и c для упрощения выкладок положим постоянными. Для определенности положим, что $k > 0$ и $c \ge 0$ (последнее физически соответствует поглощению вещества вследствие процессов реакции, или же стоку тепла [1, 4, 5]).

Для скалярного произведения и нормы в $L_2[0;\, 1]$ используем обозначения $(\cdot\,, \cdot)_0$ и $\|\cdot\|_0$ соответственно [14, 15]. Также будем использовать пространства Соболева [14, 15] $H^m[0;\, 1]$ и $H_0^m[0;\, 1]$ — пополнения $C^{(\infty)}[0;\, 1]$ и $C_0^{(\infty)}[0;\, 1]$ (индекс 0 внизу означает, что носитель принадлежит интервалу $(0;\, 1)$) по норме $\|u\|_m \equiv \left(\int_0^1 \sum_{0\le i\le m} (d^i u/dx^i)^2\, dx\right)^{1/2}$. Для функций из $H^m[0;\, 1]$ будем также использовать полунормы $|u|_m \equiv$ $\equiv \left(\int_0^1 (d^m u/dx^m)^2\, dx\right)^{1/2}$. Известно (см. [14, 15]), что для функций из $H_0^1[0;\, 1]$ норма $\|\cdot\|_1$ эквивалентна полунорме $|\cdot|_1$.

Для записи слабой формы [2, 3, 6, 7, 11] рассматриваемой краевой задачи введем билинейную форму $a(u, v) \equiv a_1(u, v) + a_2(u, v)$, где $a_1(u, v) \equiv (u', v' + kv)_0$ (здесь $u' \equiv du/dx$), $a_2(u, v) \equiv (u, cv)_0$. Тогда слабую форму задачи можно представить в виде

$$\text{найти } u \in H_0^1[0;\, 1] \text{ такое, что для любого } v \in H_0^1[0;\, 1]$$
$$\text{выполняется } a(u, v) = (f, v)_0 . \tag{3}$$

Лемма 1. *Для билинейной формы* $a(u, v)$, *рассматриваемой на* $H_0^1[0;1] \times H_0^1[0;1]$, *справедливы следующие утверждения*:

1) *существует* $C_1 \equiv \sqrt{1 + k^2/(4\pi^2)} + c\pi^{-2} > 0$ *такое, что для любых* $u \in H_0^1[0;1]$ *и* $v \in H_0^1[0;1]$ *выполняется*

$$|a(u, v)| \le C_1 |u|_1 |v|_1; \tag{4}$$

2) $\inf_{u \in H_0^1[0;1]} \sup_{v \in H_0^1[0;1]} \dfrac{|a(u,v)|}{|u|_1 |v|_1} \ge 1$;

3) *для любого* $v \neq 0$ *справедливо* $\sup_{u \in H_0^1[0;1]} |a(u,v)| > 0$.

Доказательство. Докажем сначала пункт 1 леммы. В статье [11] доказано, что для любых $u \in H_0^1[0;1]$ и $v \in H_0^1[0;1]$ выполняется $|a_1(u, v)| \le \sqrt{1 + k^2/(4\pi^2)}\,|u|_1 |v|_1$, причем данная константа непрерывности является здесь неулучшаемой. Теперь из неравенства Коши – Буняковского и неравенства Фридрихса (с неулучшаемой константой) [17, 18]

$$||w||_0 \le \pi^{-1} |w|_1 \quad \forall w \in H_0^1[0;1] \tag{5}$$

имеем

$$|a(u, v)| \le |a_1(u, v)| + |a_2(u, v)| \le |a_1(u, v)| + c||u||_0 ||v||_0 \le$$

$$\le \left(\sqrt{1 + k^2/(4\pi^2)} + c\pi^{-2} \right) |u|_1 |v|_1 .$$

Докажем теперь пункты 2 и 3. Поскольку $a_1(u, u) = |u|_1^2$ и $a_2(u, u) = c||u||_0^2$, отсюда следует, что

$$\inf_{u \in H_0^1[0;1]} \sup_{v \in H_0^1[0;1]} \frac{|a(u,v)|}{|u|_1 |v|_1} \ge \inf_{u \in H_0^1[0;1]} \frac{|a(u,u)|}{|u|_1^2} = \inf_{u \in H_0^1[0;1]} \left(1 + c \frac{||u||_0^2}{|u|_1^2} \right) \ge 1 .$$

Далее,

$$\sup_{u \in H_0^1[0;1]} |a(u,v)| \ge |a(v,v)| = |v|_1^2 + c||v||_0^2 \ge |v|_1^2 > 0 \quad \forall v \neq 0$$

(достаточно воспользоваться эквивалентностью $||\cdot||_1$ и $|\cdot|_1$ на $H_0^1[0;1]$ или неравенством Фридрихса).

Лемма доказана.

Следствие 1. *Из леммы* 1, *а также линейности и ограниченности функционала* $F(v) = (f, v)_0$ *правой части задачи* (3), *в силу обобщенной теоремы Лакса – Мильграма* (*см.* [19, *с.* 112], *теорема* 5.2.1) *следует существование и единственность решения задачи* (3).

Лемма 1 также даст возможность в явном виде находить константы в априорных оценках при оценивании погрешности МПГ.

2. Аппроксимация краевой задачи МПГ. Считаем, что на отрезке $[0;1]$ задана система равномерно отстоящих точек (узлов) x_i, $i=0,\ldots,N+1$, с шагом $h=x_{i+1}-x_i$, $x_0=0$, $x_{N+1}=1$. С каждым узлом x_i свяжем непрерывную кусочно-линейную финитную базисную функцию $N_i(x)$; она равна нулю за пределами $x_{i-1}<x<x_{i+1}$, линейна на элементах $[x_{i-1};x_i]$, $[x_i;x_{i+1}]$ и равна единице в точке x_i. В качестве соответствующей узлу x_i сетки весовой функции $W_i(x)$ используем кусочно-квадратичные функции вида (1), где функция $W_i^*(x)$ определяется следующим образом [3, 10, 11, 13, 20]:

$$W_i^*(x) = \begin{cases} {}^2W((x_i-x)/h), & x\in[x_{i-1};x_i], \\ -{}^2W((x_{i+1}-x)/h), & x\in[x_i;x_{i+1}], \\ 0, & x\notin[x_{i-1};x_{i+1}], \end{cases}$$

а ${}^2W(\lambda)\equiv 3\lambda(1-\lambda)$. Особенности построения различных весовых функций и требования к ним приведены в [20, 21]. Обозначим символами Φ_h и Ψ_h конечномерные подпространства пространства $H_0^1[0;1]$, являющиеся линейными оболочками совокупностей базисных $\{N_i(x)\}_{i=1}^N$ и весовых $\{W_i(x)\}_{i=1}^N$ функций соответственно. Рассматриваемая здесь конечно-элементная аппроксимация Петрова – Галеркина решения задачи (3) имеет вид

$$\text{дано}\quad (\alpha_1;\alpha_2;\ldots;\alpha_N)^T\in R^N;\quad \text{найти}\quad u_h\in\Phi_h\quad \text{такое, что}$$

$$a(u_h, N_i+\alpha_i W_i^*) = (f, N_i+\alpha_i W_i^*)_0, \quad i=1,\ldots,N. \tag{6}$$

Основным средством при оценивании погрешности аппроксимации решения задачи (3) решением приближенной задачи (6) является следующая теорема [3, 11, 13, 22] (обобщенная лемма Сеа; доказательство приведено в [11, 13]; см. теорему 3.1 в статье [11]).

Теорема 1. *Пусть выполняются следующие условия*: 1) Φ_h *и* Ψ_h *являются конечномерными подпространствами* $H_0^1[0;1]$ *такими, что* $c_2(h)\equiv\inf_{u\in\Phi_h}\sup_{v\in\Psi_h}\dfrac{|a(u,v)|}{|u|_1|v|_1}>0$ *и для любых* $v\in\Psi_h$, $v\neq 0$ $\sup_{u\in\Phi_h}|a(u,v)|>0$; 2) $u\in H_0^1[0;1]$ *— единственное решение задачи* (3). *Тогда задача* (6) *имеет единственное решение* u_h, *которое удовлетворяет оценке*

$$|u-u_h|_1 \le \left(1+\frac{C_1}{c_2(h)}\right)\inf_{w_h\in\Phi_h}|u-w_h|_1, \tag{7}$$

где C_1 *— константа непрерывности для формы* $a(u,v)$ (*см. лемму* 1).

В оценке (7) величина $\inf_{w_h \in \Phi_h} |u - w_h|_1$ фиксирована при выбранном пространстве Φ_h и зависит лишь от аппроксимирующих свойств Φ_h (см. ниже доказательство теоремы 3), но не зависит от $\{\alpha_i\}_{i=1}^{N}$. Величина $c_2(h)$, называемая константой Ладыженской – Бабушки – Бреззи [13, 3, 6] (в некоторых источниках этим термином называют также и любые оценки снизу для $c_2(h)$, см. [13]), оказывает существенное влияние на точность и сходимость МПГ в H^1 и L_2, ее роль при исследовании точности и сходимости МПГ обсуждается в [3, 10, 11, 13]. Из определения величин $c_2(h)$ и C_1 следует неравенство $c_2(h) \le C_1$. Отметим (см. введение), что впервые оценки снизу для $c_2(h)$ (для задач КД и при условии $\alpha_i = \alpha = \text{const} \ge 0$) были получены в [11] (см. также [13]), причем найденные там оценки являются „квазиоптимальными" (неулучшаемыми в некоторых случаях [11, 13]). Константа C_1 найдена в лемме 1 в явном виде, что, в свою очередь, дает возможность в явном виде (а не только асимптотически при $h \to 0$) оценивать сверху погрешность МПГ в полунорме $|\cdot|_1$. В [10] были установлены оценки снизу для $c_2(h)$, включающие в себя (как частные случаи) известные автору оценки, полученные ранее другими исследователями (как для случаев с неодинаковыми α_i, зависящими от индекса i, так и, в том числе, дающие „квазиоптимальные" оценки [11, 13] в частном случае при $\alpha_i = \alpha = \text{const} \ge 0$). Чем большей будет величина $c_2(h)$, тем меньшей будет погрешность $|u - u_h|_1$ в (7), поэтому при получении оценок снизу для $c_2(h)$ чем большей по величине является такая оценка, тем она будет сильнее и „лучше" в указанном смысле (по отношению к другим оценкам снизу, меньшим по величине, чем данная оценка).

Определение 1. *Будем говорить, что величина $c_2(h)$ равномерно отделена от нуля, если существуют $\varepsilon > 0$ и $h_0 > 0$ такие, что для любого $h \in (0; h_0]$ выполняется $c_2(h) \ge \varepsilon$.*

Лемма 2. *Для произвольных*

$$u(x) = \sum_{j=1}^{N} u_j N_j(x) \in \Phi_h, \qquad v(x) = \sum_{i=1}^{N} v_i W_i(x) \in \Psi_h$$

справедливо $a(u, v) = \vec{v}^T A_h \vec{u}$, $|u|_1^2 = \vec{u}^T B_h \vec{u}$, $|v|_1^2 = \vec{v}^T C_h \vec{v}$, где $\vec{u} \equiv (u_1; u_2; \ldots; u_N)^T \in R^N$, $\vec{v} \equiv (v_1; v_2; \ldots; v_N)^T \in R^N$ (индекс T — знак транспонирования), а матрицы A_h, B_h и C_h размера $N \times N$ представляются поэлементно следующим образом ($1 \le i \le N$, $1 \le j \le N$):

$$(A_h)_{i,j} = \begin{cases} -1/h - (1+\alpha_i)k/2 + ch(1/6 + \alpha_i/4), & j = i-1, \\ 2/h + k\alpha_i + 2ch/3, & j = i, \\ -1/h + (1-\alpha_i)k/2 + ch(1/6 - \alpha_i/4), & j = i+1, \\ 0, & |i-j| > 1, \end{cases} \tag{8}$$

$$(B_h)_{i,j} = \begin{cases} -1/h, & |i-j| = 1, \\ 2/h, & j = i, \\ 0, & |i-j| > 1, \end{cases} \tag{9}$$

$$(C_h)_{i,j} = \begin{cases} -(1+3\alpha_i\alpha_j)/h, & |i-j| = 1, \\ (2+6\alpha_i^2)/h, & j = i, \\ 0, & |i-j| > 1. \end{cases} \tag{10}$$

Доказательство данной леммы получается путем непосредственного подсчета интегралов, входящих в выражения для $a(u, v)$, $|u|_1^2$ и $|v|_1^2$.

Следствие 2. *Задача* (6) *сводится к решению системы линейных уравнений с матрицей* A_h *и правой частью* $\vec{f} \equiv (f_1; f_2; \ldots; f_N)^T$, *где* $f_i \equiv (f, W_i)_0$.

3. Получение основных соотношений для оценивания погрешности МПГ. Пусть $\|\cdot\|$ — стандартная евклидова векторная норма на R^N (т. е. $\|\vec{x}\| = \sqrt{(\vec{x}, \vec{x})}$ при $\vec{x} \in R^N$, где $(\vec{x}, \vec{y})$ — евклидово скалярное произведение). Этим же символом $\|\cdot\|$ обозначаем соответствующую подчиненную матричную (операторную) норму произвольной матрицы A размера $N \times N$: $\|A\| \equiv \sup_{\|\vec{x}\|=1}\|A\vec{x}\|$. Единичную матрицу обозначаем через E. Тогда *логарифмической нормой* (ЛН) матрицы A называют число [24, 25] $\mu(A) = \lim_{t\to 0+0}\left(\|E+tA\|-1\right)/t$. В дальнейшем будем пользоваться следующими свойствами ЛН произвольной матрицы A:

1) для любого $\vec{x} \in R^N$ выполняется неравенство $-\mu(-A)\|\vec{x}\|^2 \le \vec{x}^T A\vec{x} = (A\vec{x}, \vec{x}) \le \le \mu(A)\|\vec{x}\|^2$, причем верхняя и нижняя границы неравенства достигаются на некоторых векторах [24, 25];

2) для произвольной матрицы $A = \{a_{i,j}\}_{i,j=1}^N$ справедлива оценка

$$\mu(A) \le \max_{1\le i\le N}\left(a_{i,i} + \frac{1}{2}\sum_{j\ne i}\left| a_{i,j} + a_{j,i}\right| \right)$$

(свойство доказано в лемме 2 работы [10]);

3) для любого $a \ge 0$ справедливо $\mu(aA) = a\mu(A)$ (см. [25]);

4) $\mu(A+B) \le \mu(A) + \mu(B)$, где B — любая матрица одинакового с A размера [25];

5) $\mu(A) = \max_i\left(\lambda_i\left((A+A^T)/2\right)\right)$, где $\lambda_i\left((A+A^T)/2\right)$ — i-е собственное число матрицы $\left(A+A^T\right)/2$ (см. [24, 25]).

Введем вспомогательные матрицы D, B, A_0, M и A размера $N \times N$, которые нам понадобятся в дальнейшем при нахождении оценок для величины $c_2(h)$:

$$D_{i,j} = \begin{cases} -\alpha_i\alpha_j, & |i-j|=1, \\ 2\alpha_i^2 & j=i, \\ 0, & |i-j|>1, \end{cases} \qquad (A_0)_{i,j} = \begin{cases} -\alpha_i, & |i-j|=1, \\ 2\alpha_i, & j=i, \\ 0, & |i-j|>1, \end{cases} \qquad M_{i,j} = \begin{cases} 2/3 \mp \alpha_i, & j=i\pm 1, \\ 8/3, & j=i, \\ 0, & |i-j|>1, \end{cases}$$

$B \equiv hB_h$ (см. выражение (9), т. е. B — это матрица B_h, но без множителя $1/h$), $A \equiv B + \frac{kh}{2}A_0 + \frac{ch^2}{4}M$. Тогда матрицу A_h из (8) можно представить как $A_h = (A_1 + A)/h$, где матрица A_1 определяется данным соотношением (она кососимметрична, $A_1 = -A_1^T$, на главной диагонали у нее расположены нули, на верхней и нижней поддиагоналях — величины $\pm kh/2$ соответственно). Мотивация при введении данного представления состоит в том, что в дальнейшем мы столкнемся с необходимостью вычисления и оценивания выражений типа $(A_h\vec{x},\vec{x})$, но в силу $A_1 = -A_1^T$ имеем $(A_h\vec{x},\vec{x}) = (A\vec{x},\vec{x})/h$ $\forall\vec{x}\in R^N$, что упрощает эту задачу. Отметим также, что для любого $\vec{x}\in R^N$ справедливо $\vec{x}^T B\vec{x} = (B\vec{x},\vec{x}) > 0$ при $\vec{x}\neq 0$, т. е. квадратичная форма $(B\vec{x},\vec{x})$ положительно определена (это следует из леммы 2, см. также [26]). Спектр матрицы B состоит из точек [26] $\lambda_j^{(N)} \equiv 4\sin^2(\pi j/(2N+2))$, где $1\le j\le N$.

Лемма 3. *Пусть* ξ *— произвольное (но фиксированное) действительное число. Тогда*:

1) *неравенство* $(A\vec{x},\vec{x}) \ge \xi(B\vec{x},\vec{x})$ *имеет место (при любом* $\vec{x}\in R^N$*) тогда и только тогда, когда выполняется условие* $-\mu(-(A-\xi B)) \ge 0$;

2) *если выполнено условие* $-\mu(-(A-\xi B)) \ge 0$, *то*

$$(A\vec{x},\vec{x}) \ \ge\ \xi(B\vec{x},\vec{x}) - \mu(-(A-\xi B))(B\vec{x},\vec{x})/\lambda_N^{(N)} \qquad \forall\vec{x}\in R^N; \tag{11}$$

3) *выполняется неравенство* $-\mu(-(A-\xi B)) \ge \tilde{C}_{AB}$, *где* $\tilde{C}_{AB}$ *определяется как*

$$\tilde{C}_{AB} \ \equiv\ \min\left\{ 2-2\xi+kh\alpha_1+\frac{2ch^2}{3}-\frac{1}{2}\left|2\xi-2-\frac{kh}{2}(\alpha_1+\alpha_2)+ch^2\left(\frac{1}{3}+\frac{\alpha_2-\alpha_1}{4}\right)\right|,\right.$$

$$-\min_{1<i<N}\left(2-2\xi+kh\alpha_i+\frac{2ch^2}{3}-\frac{1}{2}\left|2\xi-2-\frac{kh}{2}(\alpha_{i-1}+\alpha_i)+ch^2\left(\frac{1}{3}+\frac{\alpha_i-\alpha_{i-1}}{4}\right)\right| -\right.$$

$$\left.-\frac{1}{2}\left|2\xi-2-\frac{kh}{2}(\alpha_i+\alpha_{i+1})+ch^2\left(\frac{1}{3}+\frac{\alpha_{i+1}-\alpha_i}{4}\right)\right|\right),$$

$$\left. 2-2\xi+kh\alpha_N+\frac{2ch^2}{3}-\frac{1}{2}\left|2\xi-2-\frac{kh}{2}(\alpha_{N-1}+\alpha_N)+ch^2\left(\frac{1}{3}+\frac{\alpha_N-\alpha_{N-1}}{4}\right)\right|\right\};$$

при этом если $-\mu(-(A-\xi B))\ge 0$ *(для чего достаточно, чтобы* $\tilde{C}_{AB}\ge 0$*), то*

$$(A\vec{x},\vec{x}) \ \ge\ \xi(B\vec{x},\vec{x})+\tilde{C}_{AB}(B\vec{x},\vec{x})/\lambda_N^{(N)} \qquad \forall\vec{x}\in R^N\,; \tag{12}$$

4) *если выполнено условие* $-\mu(-(A-\xi B))<0$, *то*

$$(A\vec{x},\vec{x}) \ \ge\ \xi(B\vec{x},\vec{x})-\mu(-(A-\xi B))(B\vec{x},\vec{x})/\lambda_1^{(N)} \ \ge\ \xi(B\vec{x},\vec{x})+\tilde{C}_{AB}(B\vec{x},\vec{x})/\lambda_1^{(N)} \qquad \forall\vec{x}\in R^N\,.$$

Доказательство. Пункт 1 леммы следует из свойства 1 ЛН, примененного к матрице $A-\xi B$. Докажем пункт 2 леммы. В силу свойства 1 ЛН $((A-\xi B)\vec{x},\vec{x})\ge -\mu(-(A-\xi B))\|\vec{x}\|^2$. Оценим $\|\vec{x}\|$ в данном неравенстве. В силу неравенства Рэлея – Ритца [23] имеем

$$\lambda_1^{(N)}\|\vec{x}\|^2 \ \le\ (B\vec{x},\vec{x}) \ \le\ \lambda_N^{(N)}\|\vec{x}\|^2 \qquad \forall\vec{x}\in R^N\,. \tag{13}$$

Отсюда получаем оценку $(B\vec{x},\vec{x})/\lambda_N^{(N)}\le\|\vec{x}\|^2$ и, как следствие, неравенство (11). Для доказательства пункта 3 достаточно применить свойство 2 ЛН к матрице $A-\xi B$, выполнив при этом несложные, но довольно громоздкие выкладки по приведению подобных членов в соответствующих выражениях (см. также доказательство лемм 3 и 4 в работе [10]). Неравенство (12) при этом следует из неравенства (11). Утверждение пункта 4 следует из неравенств $((A-\xi B)\vec{x},\vec{x})\ge -\mu(-(A-\xi B))\|\vec{x}\|^2$ и $\|\vec{x}\|^2\le(B\vec{x},\vec{x})/\lambda_1^{(N)}$.

Лемма доказана.

***Следствие* 3.** *Пусть* ξ *имеет вид*

$$\xi \ =\ 1+\frac{kh\delta}{2}+\frac{ch^2\phi}{4}\,, \tag{14}$$

где δ *и* ϕ *— некоторые действительные числа. Тогда для выполнения условия* $-\mu(-(A-\xi B))\ge 0$ *из пункта* 2 *леммы* 3 *и неравенства*

$$(A\vec{x},\vec{x}) \ \ge\ \xi(B\vec{x},\vec{x})-\frac{kh}{2}\mu(-(A_0-\delta B))\frac{(B\vec{x},\vec{x})}{\lambda_N^{(N)}}-\frac{ch^2}{4}\mu(-(M-\phi B))\frac{(B\vec{x},\vec{x})}{\lambda_N^{(N)}} \tag{15}$$

достаточно выполнения условий $-\mu(-(A_0-\delta B))\ge 0$ *и* $-\mu(-(M-\phi B))\ge 0$.

Доказательство. Из свойств 3 и 4 ЛН получаем неравенства

$$-\mu(-(A-\xi B)) \ =\ -\mu(-kh(A_0-\delta B)/2-ch^2(M-\phi B)/4) \ \ge$$

$$\ge\ -(kh/2)\mu(-(A_0-\delta B))-(ch^2/4)\mu(-(M-\phi B)) \ \ge\ 0\,.$$

Неравенство (15) при этом следует из неравенства (11).

***Замечание* 1.** Оценки снизу для выражений $-\mu(-(A_0-\delta B))$ и $-\mu(-(M-\phi B))$ (и соответствующие достаточные условия для выполнения неравенств $-\mu(-(A_0-\delta B))\ge 0$, $-\mu(-(M-$

$-\phi B)) \geq 0$) можно получить из пункта 3 леммы 3, взяв ξ в виде (14) и формально положив там $c = 0$ или $k = 0$ соответственно (тогда выражение $\tilde{C}_{AB}$ даст искомую оценку снизу). Отметим, что частный случай подобного рода достаточных условий для матрицы A_0 (который также можно получить описанным выше способом из леммы 3, положив в $\tilde{C}_{AB}$ $c = 0$ и раскрыв модули со знаком минус) был выведен в работе [10] (см. лемму 4 в [10]).

Лемма 4. *Пусть* χ — *произвольное* (*но фиксированное*) *действительное число*. *Тогда*:

1) *неравенство* $(D\vec{x},\vec{x}) \leq \chi(B\vec{x},\vec{x})$ *имеет место* (*при любом* $\vec{x} \in R^N$) *тогда и только тогда*, *когда выполняется условие* $-\mu(-(\chi B - D)) \geq 0$;

2) *если выполнено условие* $-\mu(-(\chi B - D)) \geq 0$, *то*

$$0 \;\leq\; (D\vec{x},\vec{x}) \;\leq\; \chi(B\vec{x},\vec{x}) + \mu(-(\chi B - D))(B\vec{x},\vec{x})/\lambda_N^{(N)} \qquad \forall \vec{x} \in R^N; \tag{16}$$

3) *выполняется неравенство* $-\mu(-(\chi B - D)) \geq \tilde{C}_{DB}$, *где* $\tilde{C}_{DB}$ *определяется как*

$$\tilde{C}_{DB} \;\equiv\; \min\Big\{ 2\chi - 2\alpha_1^2 - |\chi - \alpha_1\alpha_2|,\;\; 2\chi - 2\alpha_N^2 - |\chi - \alpha_{N-1}\alpha_N|,$$

$$\min_{1<i<N}\Big(2\chi - 2\alpha_i^2 - |\chi - \alpha_{i-1}\alpha_i| - |\chi - \alpha_i\alpha_{i+1}| \Big)\Big\};$$

при этом если $-\mu(-(\chi B - D)) \geq 0$ (*для чего достаточно*, *чтобы* $\tilde{C}_{DB} \geq 0$), *то*

$$0 \;\leq\; (D\vec{x},\vec{x}) \;\leq\; \chi(B\vec{x},\vec{x}) - \tilde{C}_{DB}(B\vec{x},\vec{x})/\lambda_N^{(N)} \qquad \forall \vec{x} \in R^N; \tag{17}$$

4) *если выполнено условие* $-\mu(-(\chi B - D)) < 0$, *то*

$$0 \;\leq\; (D\vec{x},\vec{x}) \;\leq\; \chi(B\vec{x},\vec{x}) + \mu(-(\chi B - D))(B\vec{x},\vec{x})/\lambda_1^{(N)} \;\leq$$

$$\leq\; \chi(B\vec{x},\vec{x}) - \tilde{C}_{DB}(B\vec{x},\vec{x})/\lambda_1^{(N)} \qquad \forall \vec{x} \in R^N.$$

Доказательство. Пункт 1 леммы следует из свойства 1 ЛН. Неравенство $0 \leq (D\vec{x},\vec{x})$ доказано в [10] (см. лемму 3 в [10]). Пункт 2 следует из неравенства

$$((\chi B - D)\vec{x},\vec{x}) \;\geq\; -\mu(-(\chi B - D))\|\vec{x}\|^2 \;\geq\; -\mu(-(\chi B - D))(B\vec{x},\vec{x})/\lambda_N^{(N)}$$

(см. доказательство леммы 3). Для доказательства пункта 3 достаточно применить свойство 2 ЛН к матрице $\chi B - D$, неравенство (17) при этом следует из (16). Пункт 4 следует из неравенства $-\mu(-(\chi B - D))\|\vec{x}\|^2 \geq -\mu(-(\chi B - D))(B\vec{x},\vec{x})/\lambda_1^{(N)}$ (см. (13)) при $\mu(-(\chi B - D)) > 0$.

Лемма доказана.

***Замечание* 2.** Лемма 4 обобщает лемму 3 из статьи [10]. В [10] также изложены соображения по поводу выбора допустимых значений числа χ, при котором не нарушается неравенство $(D\vec{x},\vec{x}) \leq \chi(B\vec{x},\vec{x})$ (в частности, там показано, что для выполнения данного неравенства необходимо, чтобы $\chi \geq \max_{1\leq i\leq N} \alpha_i^2$).

4. Погрешность и сходимость МПГ. Докажем следующую теорему, дающую оценку величины $c_2(h)$ в теореме 1.

Теорема 2. *Пусть выполняются условия пунктов* 2 *в леммах* 3 *и* 4. *Тогда*

$$c_2(h) \;\geq\; C_0(h) \;\equiv\; \frac{\xi - \mu(-(A-\xi B))/\lambda_N^{(N)}}{\sqrt{1+3\left(\chi+\mu\left(-(\chi B-D)\right)/\lambda_N^{(N)}\right)}}. \tag{18}$$

Доказательство. В силу леммы 2 имеем $c_2(h) = \min_{\vec{u}\in R^N} \max_{\vec{v}\in R^N} \dfrac{\left|\vec{v}^T A_h\vec{u}\right|}{\sqrt{\vec{u}^T B_h\vec{u}}\sqrt{\vec{v}^T C_h\vec{v}}}$, откуда получаем неравенство $c_2(h) \geq \min_{\vec{u}\in R^N} \dfrac{\vec{u}^T A_h\vec{u}}{\sqrt{\vec{u}^T B_h\vec{u}}\sqrt{\vec{u}^T C_h\vec{u}}}$. Далее, используя равенства $(A_h\vec{u},\vec{u}) = (A\vec{u},\vec{u})h^{-1}$, $B_h = h^{-1}B$, $C_h = h^{-1}(B+3D)$ и неравенства (11) и (16), получаем

$$c_2(h) \;\geq\; \min_{\vec{u}\in R^N} \frac{\vec{u}^T A\vec{u}}{\sqrt{\vec{u}^T B\vec{u}}\sqrt{\vec{u}^T (B+3D)\vec{u}}} \;\geq$$

$$\geq\; \min_{\vec{u}\in R^N} \frac{(\xi - \mu(-(A-\xi B))/\lambda_N^{(N)})\vec{u}^T B\vec{u}}{\sqrt{\vec{u}^T B\vec{u}}\sqrt{\vec{u}^T B\vec{u} + 3(\chi+\mu(-(\chi B-D))/\lambda_N^{(N)})\vec{u}^T B\vec{u}}}\,.$$

Сокращая здесь числитель и знаменатель на $\vec{u}^T B\vec{u}$, получаем оценку (18).

Теорема доказана.

***Следствие* 4.** 1. *Пусть выполняются условия теоремы* 2. *Используя пункты* 3 *лемм* 3 *и* 4, *оценку* (18) *можно ослабить, получив для* $c_2(h)$ *следующие оценки снизу*:

$$c_2(h) \;\geq\; \frac{\xi + \tilde{C}_{AB}/\lambda_N^{(N)}}{\sqrt{1+3\left(\chi+\mu\left(-(\chi B-D)\right)/\lambda_N^{(N)}\right)}}, \tag{19}$$

$$c_2(h) \;\geq\; \frac{\xi - \mu(-(A-\xi B))/\lambda_N^{(N)}}{\sqrt{1+3\left(\chi - \tilde{C}_{DB}/\lambda_N^{(N)}\right)}}, \tag{20}$$

$$c_2(h) \;\geq\; \frac{\xi + \tilde{C}_{AB}/\lambda_N^{(N)}}{\sqrt{1+3\left(\chi - \tilde{C}_{DB}/\lambda_N^{(N)}\right)}}. \tag{21}$$

2. *Требование выполнения пункта* 2 *леммы* 3 *в формулировке теоремы* 2 *можно заменить требованием выполнения пункта* 4 *леммы* 3; *при этом будет справедлива оценка* (18), *где в числителе правой части взято выражение* $\xi - \mu(-(A-\xi B))/\lambda_1^{(N)}$ (*или его оценка снизу* $\xi + \tilde{C}_{AB}/\lambda_1^{(N)}$).

3. *Вместо выполнения пункта* 2 *леммы* 4 *в условиях теоремы* 2 *можно потребовать выполнения пункта* 4 *леммы* 4; *при этом справедлива оценка* (18), *где в знаменателе правой части взято выражение* $\sqrt{1+3\left(\chi+\mu(-(\chi B-D))/\lambda_1^{(N)}\right)}$ $\Big($*или его оценка сверху* $\sqrt{1+3\left(\chi-\tilde{C}_{DB}/\lambda_1^{(N)}\right)}\Big)$.

Доказательство. Справедливость оценок (19) – (21) непосредственно следует из оценки (18) с учетом пунктов 3 лемм 3 и 4. Отметим также, что оценки (19) – (21) можно установить независимо от (18), следуя доказательству теоремы 2 (но вместо неравенств (11) и (16) используя при доказательстве неравенства (12) и (17) соответственно). Для доказательства пунктов 2 и 3 следствия достаточно повторить схему доказательства теоремы 2 (применяя при этом неравенства из пунктов 4 лемм 3 и 4). Отметим, что пункты 2 и 3 следствия 4 не зависят один от другого, поэтому если их условия выполняются одновременно, для величины $c_2(h)$ имеют место соответствующие оценки.

***Замечание* 3.** Если ξ определяется выражением (14) и справедливы условия следствия 3, то оценки (18) и (20) можно ослабить, заменив (оценив снизу) числитель выражением $\xi-kh\,\mu(-(A_0-\delta B))/(2\lambda_N^{(N)})-ch^2\mu(-(M-\phi B))/(4\lambda_N^{(N)})$. В свою очередь, ЛН в последнем выражении также можно оценить снизу (см. замечание 1), получив таким образом для $c_2(h)$ еще более слабые оценки. Отсюда, как частный случай, положив $c=0$ и отбросив слагаемые с ЛН в числителе и знаменателе (что приведет лишь к огрублению оценки), получим оценку $c_2(h)\ge(1+kh\delta/2)/\sqrt{1+3\chi}$ для уравнения КД, выведенную в [10] (см. теорему 2 в [10]).

Ранее было установлено (см. следствие 1), что решение задачи (3) существует и единственно. Докажем теперь теорему, которая дает оценку точности МПГ и на основании которой устанавливается его сходимость.

Теорема 3. *Пусть* $u\in H_0^1[0;\,1]$ *является решением задачи* (3). *Предположим дополнительно*, *что* u *также принадлежит пространству* $H^2[0;\,1]$. *Пусть выполнены условия теоремы* 2 *и* $C_0(h)>0$. *Тогда решение* $u_h\in\Phi_h$ *задачи* (6) *существует*, *единственно и удовлетворяет оценкам*

$$|u-u_h|_1 \le \left(1+\frac{C_1}{c_2(h)}\right)h|u|_2 \le \left(1+\frac{C_1}{C_0(h)}\right)h|u|_2 . \tag{22}$$

Доказательство. Покажем, что $\sup_{u\in\Phi_h}|a(u,v)|>0 \quad \forall v\in\Psi_h,\ \ v\ne 0$. Действительно, в силу условия $C_0(h)>0$, леммы 2 и пункта 2 леммы 3 имеем неравенства

$$\sup_{u\in\Phi_h}|a(u,v)| = \sup_{\vec u\in R^N}\left|\vec v^T A_h\vec u\right| \ge \left|\vec v^T A_h\vec v\right| \ge \vec v^T A_h\vec v \ge$$
$$\ge h^{-1}\left(\xi-\mu(-(A-\xi B))/\lambda_N^{(N)}\right)(B\vec v,\vec v) > 0 .$$

Таким образом, все условия теоремы 1 теперь выполнены. Отсюда получаем, что решение $u_h\in\Phi_h$ задачи (6) существует, единственно и удовлетворяет оценке (7). Для оценки величины $\inf_{w_h\in\Phi_h}|u-w_h|_1$ воспользуемся тем, что пространство Φ_h является пространством интерполяционных сплайнов первой степени, для которого при $u \in H^2[0;\,1]$ справедлива оценка $\inf_{w_h \in \Phi_h} |u - w_h|_1 \le h|u|_2$ (см. [26]). Отсюда получаем оценку (22).

Теорема доказана.

***Следствие* 5.** *Пусть величина* $c_2(h)$ *равномерно отделена от нуля* (*для этого достаточно*, *чтобы* $C_0(h)$ *была равномерно отделена от нуля*). *Тогда при выполнении условий теоремы* 3 *из* (22) *получаем* $|u - u_h|_1 = O(h)$ *и* $\|u - u_h\|_1 = O(h)$ (*соответственно*, *и* $\|u - u_h\|_0 = O(h)$) *при* $h \to 0$. *Использовав теорему Соболева о вложении* [14, 15], *также получим сходимость* u_h *к* u *со скоростью* $O(h)$ *и в пространстве* $C[0;\,1]$.

5. Случай с одинаковыми параметрами α_i**.** Рассмотрим случай, когда все $\alpha_i = \alpha = \mathrm{const}$. Тогда матрицы $A_0 = \alpha B$, $A = (1 + kh\alpha/2)B + (ch^2/4)M$, поэтому, беря ξ в виде (14) (при $\delta = \alpha$), получаем, что неравенство $(A\vec{x}, \vec{x}) \ge \xi(B\vec{x}, \vec{x})$ сводится к неравенству $(M\vec{x}, \vec{x}) \ge \phi(B\vec{x}, \vec{x})$, а условие $-\mu(-(A - \xi B)) \ge 0$ — соответственно (используя свойство 3 ЛН) к условию $-\mu(-(M - \phi B)) \ge 0$.

Утверждение 1. *Число* $\phi = \phi_* \equiv \dfrac{1 + 2\sin^2(\pi h/2)}{3 - 3\sin^2(\pi h/2)}$ *является максимальным неотрицательным числом*, *при котором для любого* $\vec{x} \in R^N$ *выполняется неравенство* $(M\vec{x}, \vec{x}) \ge \phi(B\vec{x}, \vec{x})$. *При этом* $\mu(-(M - \phi B)) = 0$.

Доказательство. Обозначим $G \equiv M - \phi B$, тогда $(G\vec{x}, \vec{x}) = \left(((G + G^T)/2)\vec{x}, \vec{x}\right) \ge 0$ эквивалентно (в силу неравенства Рэлея – Ритца [23]) тому, что все собственные числа матрицы $(G + G^T)/2$ неотрицательны. Но матрица $(G + G^T)/2$ является трехдиагональной симметричной теплицевой матрицей [27]. Для таких матриц спектр можно описать в явном виде (см. [27, с. 304]). В результате получим, что собственные числа λ_j для $(G + G^T)/2$ имеют вид

$$\lambda_j = \frac{8}{3} - 2\phi + 2\left(\frac{2}{3} + \phi\right)\cos\left(\frac{j\pi}{N+1}\right), \quad 1 \le j \le N.$$

Неравенство $\lambda_N \ge 0$ для минимального собственного числа $\lambda_N = \dfrac{4}{3} - 4\phi + \left(\dfrac{8}{3} + 4\phi\right)\sin^2\left(\dfrac{\pi h}{2}\right)$ эквивалентно $\phi \le \dfrac{1 + 2\sin^2(\pi h/2)}{3 - 3\sin^2(\pi h/2)}$, откуда, п оскольку $\mu(-(M - \phi B))$ равно максимальному собственному числу $(-G - G^T)/2$ (свойство 5 ЛН), получим $\mu(-(M - \phi B)) = 0$ при $\phi = \phi_*$.

Утверждение доказано.

Аналогично, поскольку $D = \alpha^2 B$, из леммы 4 получаем, что в качестве χ можно взять $\chi = \alpha^2$, причем данное значение будет минимальным, при котором выполняется неравенство $(D\vec{x}, \vec{x}) \le \chi(B\vec{x}, \vec{x})$. Опираясь на это, а также на утверждение 1 и то, что ЛН нулевой матрицы равна нулю, из оценки (18) получаем следующее утверждение.

Утверждение 2. *При* $\alpha_i = \alpha = \text{const}$ *справедлива оценка*

$$c_2(h) \;\geq\; C_0(h) \;=\; \frac{1 + kh\alpha/2 + ch^2\phi_*/4}{\sqrt{1+3\alpha^2}}\,. \tag{23}$$

***Замечание* 4.** Отметим, что для уравнения КД при $c = 0$ оценка (23) совпадает с оценкой из работ [11, 13], которая, как показано в [11, 13], является „квазиоптимальной” в том смысле, что при нечетных N имеет место точное равенство $c_2(h) = (1 + kh\alpha/2)(1+3\alpha^2)^{-1/2}$. В случае уравнения КДР это, вообще говоря, не будет иметь места, однако можно доказать следующее утверждение (которое, тем не менее, для уравнения КД при $c = 0$ все еще будет свидетельствовать о „квазиоптимальности” оценки (23), обобщая таким образом соответствующие результаты [11, 13]).

Утверждение 3. *При нечетном* N *и* $\alpha_i = \alpha = \text{const}$ *справедлива оценка*

$$c_2(h) \;\leq\; \frac{\left|1 + kh\alpha/2 + ch^2/3\right|}{\sqrt{1+3\alpha^2}}\,. \tag{24}$$

Доказательство. Рассмотрим матрицу $F \equiv A_h - (1 + kh\alpha/2 + ch^2\beta)B_h$, где β — некоторое (подлежащее определению) действительное число. Матрица F является трехдиагональной теплицевой матрицей, спектр которой имеет вид (см. доказательство утверждения 1)

$$\lambda_j \;=\; \frac{2ch}{3}(1-3\beta) + 2\sqrt{\left(\frac{k}{2} + \frac{ch}{6} - \frac{ch\alpha}{4} + ch\beta\right)\left(-\frac{k}{2} + \frac{ch}{6} + \frac{ch\alpha}{4} + ch\beta\right)} \;\times$$

$$\times\; \cos\frac{j\pi}{N+1}\,, \quad 1 \leq j \leq N\,.$$

Отсюда видим, что при нечетном N, $j = (N+1)/2$ и $\beta = 1/3$ $\lambda_j = 0$, откуда следует, что $\det F = 0$ и, соответственно, существует $\vec{u}_0 \in R^N$, $\vec{u}_0 \neq 0$, для которого $F\vec{u}_0 = 0$. В силу леммы 2 имеем $c_2(h) = \min_{\vec{u}\in R^N} \max_{\vec{v}\in R^N} \dfrac{\left|\vec{v}^T A_h \vec{u}\right|}{\sqrt{\vec{u}^T B_h \vec{u}}\sqrt{\vec{v}^T C_h \vec{v}}}$, откуда для любого $\vec{u} \in R^N$, $\vec{u} \neq 0$ получаем оценку $c_2(h) \leq \max_{\vec{v}\in R^N} \dfrac{\left|\vec{v}^T A_h \vec{u}\right|}{\sqrt{\vec{u}^T B_h \vec{u}}\sqrt{\vec{v}^T C_h \vec{v}}}$. Взяв $\vec{u} = \vec{u}_0$ и использовав равенство $C_h = h^{-1}(B + 3\alpha^2 B)$, получим

$$c_2(h) \;\leq\; \frac{\left|1 + kh\alpha/2 + ch^2/3\right|}{\sqrt{1+3\alpha^2}} \max_{\vec{v}\in R^N} \frac{\left|\vec{v}^T B \vec{u}_0\right|}{\sqrt{\vec{u}_0^T B \vec{u}_0}\sqrt{\vec{v}^T B \vec{v}}}\,.$$

Для получения (24) осталось показать, что максимум в последнем выражении равен единице. Для этого заметим, что матрица B является симметричной и положительно определенной (см. (13)), поэтому можно ввести скалярное произведение [26] $(\vec{x}, \vec{y})_B \equiv (B\vec{x}, \vec{y})$ и норму $||\vec{x}||_B^2 = (\vec{x}, \vec{x})_B$. Из неравенства Коши – Буняковского получаем

$$\left|\vec{v}^T B\vec{u}_0\right| = \left|(\vec{u}_0, \vec{v})_B\right| \le \|\vec{u}_0\|_B \|\vec{v}\|_B = \sqrt{\vec{u}_0^T B\vec{u}_0}\sqrt{\vec{v}^T B\vec{v}},$$

причем равенство достигается при $\vec{v} = \vec{u}_0$.

Утверждение доказано.

***Замечание* 5.** 1. Взяв производную по α от правой части (23), получим, что максимум правой части будет достигнут при

$$\alpha = \alpha_{\max} \equiv \frac{2kh}{12 + 3ch^2\phi_*}. \tag{25}$$

Для уравнения КД ($c = 0$) формула (25) дает известный результат $\alpha = kh/6$ [11, 13, 21].

2. Отметим, что величина $C_1/C_0(h)$ (и $C_1/c_2(h)$) в (22), (23) является ограниченной при $k \to +\infty$ или $c \to +\infty$, если $\alpha \ne 0$ не зависит от k и c, а h — фиксированное. Действительно, существуют пределы $\lim_{k\to+\infty} \dfrac{C_1}{C_0(h)} = \dfrac{\sqrt{1+3\alpha^2}}{\pi h\alpha}$ и $\lim_{c\to+\infty} \dfrac{C_1}{C_0(h)} = \dfrac{4\sqrt{1+3\alpha^2}}{\pi^2 h^2\phi_*}$. Это также справедливо, когда α определяется через (25). В этом случае существуют пределы $\lim_{k\to+\infty} \dfrac{C_1}{C_0(h)} = \dfrac{\sqrt{3}}{\pi h}$ и $\lim_{c\to+\infty} \dfrac{C_1}{C_0(h)} = \dfrac{4}{\pi^2 h^2\phi_*}$.

6. Среднеквадратические оценки и сходимость МПГ. С помощью неравенства Фридрихса (5) из (22) непосредственно получаем соответствующую оценку $\|u - u_h\|_0$. Отметим также, что следствие 5 (когда выполняются его условия) может предоставить лишь первый порядок точности МПГ по h в $L_2[0; 1]$, т. е. $\|u - u_h\|_0 = O(h)$ при $h \to 0$. В данном пункте представлены уточненные среднеквадратические оценки для МПГ. Докажем сначала лемму, характеризирующую интерполяционные свойства пространства Ψ_h весовых функций.

Лемма 5. *Для произвольной функции* $v \in H_0^1[0; 1] \cap H^2[0; 1]$ *существует единственная функция* $v_h \in \Psi_h$, *интерполирующая* v *в узлах* x_i, $0 \le i \le N+1$. *При этом если выполняются условия пункта* 2 *леммы* 4, *то справедлива оценка*

$$|v - v_h|_1 \le \left(1 + \sqrt{3\left(\chi + \mu(-(\chi B - D))/\lambda_N^{(N)}\right)}\right) h|v|_2 + |v|_1\sqrt{3\left(\chi + \mu(-(\chi B - D))/\lambda_N^{(N)}\right)}. \tag{26}$$

Доказательство. Первый пункт леммы непосредственно следует из определения базисных и весовых функций; при этом $v_h(x) = \sum_{j=1}^N v(x_j)W_j(x)$. В силу (1) имеем представление $v_h(x) = \tilde{u}_h(x) + e_h(x)$, где $\tilde{u}_h(x) \equiv \sum_{j=1}^N v(x_j)N_j(x)$, $e_h(x) \equiv \sum_{j=1}^N v(x_j)\alpha_j W_j^*(x)$. Вычисляя непосредственно интегралы, как в лемме 2, получаем равенство $|e_h|_1^2 = \vec{v}^T(3/h)D\vec{v}$, в котором вектор $\vec{v} \equiv (v(x_1); \ldots; v(x_N))^T$. Теперь, применяя леммы 4 и 2, получаем неравенство

$$|e_h|_1^2 = \vec{v}^T(3/h)D\vec{v} \le 3\Big(\chi+\mu(-(\chi B-D))/\lambda_N^{(N)}\Big)\vec{v}^T B_h\vec{v} = 3\Big(\chi+\mu(-(\chi B-D))/\lambda_N^{(N)}\Big)|\tilde{u}_h|_1^2.$$

Поскольку $\tilde{u}_h(x)$ является кусочно-линейным интерполяционным сплайном для $v(x)$, справедлива оценка $|v-\tilde{u}_h|_1 \le h|v|_2$ (см. [26]). Применяя теперь данное неравенство, неравенство треугольника и полученное выше неравенство между $|e_h|_1^2$ и $|\tilde{u}_h|_1^2$, получаем неравенства

$$|v-v_h|_1 \le |v-\tilde{u}_h|_1+|e_h|_1 \le$$

$$\le |v-\tilde{u}_h|_1+\sqrt{3\Big(\chi+\mu(-(\chi B-D))/\lambda_N^{(N)}\Big)}|\tilde{u}_h|_1 \le$$

$$\le |v-\tilde{u}_h|_1+\Big(|v-\tilde{u}_h|_1+|v|_1\Big)\sqrt{3\Big(\chi+\mu(-(\chi B-D))/\lambda_N^{(N)}\Big)} =$$

$$= \left(1+\sqrt{3\Big(\chi+\mu(-(\chi B-D))/\lambda_N^{(N)}\Big)}\right)|v-\tilde{u}_h|_1 +$$

$$+\sqrt{3\Big(\chi+\mu(-(\chi B-D))/\lambda_N^{(N)}\Big)}|v|_1 \le$$

$$\le \left(1+\sqrt{3\Big(\chi+\mu(-(\chi B-D))/\lambda_N^{(N)}\Big)}\right)h|v|_2 + \sqrt{3\Big(\chi+\mu(-(\chi B-D))/\lambda_N^{(N)}\Big)}|v|_1.$$

***Следствие* 6.** 1. *Если в условиях леммы* 5 *потребовать выполнения условий пункта* 3 *леммы* 4 ($c\ \ -\mu(-(\chi B-D))\ge 0$), *то из неравенства* (26) *можно получить неравенство*

$$|v-v_h|_1 \le \left(1+\sqrt{3\Big(\chi-\tilde{C}_{DB}/\lambda_N^{(N)}\Big)}\right)h|v|_2+|v|_1\sqrt{3\Big(\chi-\tilde{C}_{DB}/\lambda_N^{(N)}\Big)}.$$

2. *Если в лемме* 5 *потребовать выполнения условий пункта* 4 *леммы* 4, *то выполняется*

$$|v-v_h|_1 \le \left(1+\sqrt{3\Big(\chi+\mu(-(\chi B-D))/\lambda_1^{(N)}\Big)}\right)h|v|_2 +$$

$$+|v|_1\sqrt{3\Big(\chi+\mu(-(\chi B-D))/\lambda_1^{(N)}\Big)} \le$$

$$\le \left(1+\sqrt{3\Big(\chi-\tilde{C}_{DB}/\lambda_1^{(N)}\Big)}\right)h|v|_2+|v|_1\sqrt{3\Big(\chi-\tilde{C}_{DB}/\lambda_1^{(N)}\Big)}.$$

Данное следствие можно доказать, следуя схеме доказательства леммы 5 и используя при этом соответствующий пункт леммы 4 для установления неравенства между $(\vec{v}^T D\vec{v})/h$ и $\vec{v}^T B_h\vec{v}$. Используем лемму 5 для оценки точности МПГ в L_2.

Теорема 4. *Пусть выполнены условия теоремы* 3. *Тогда справедливы оценки*

$$\|u-u_h\|_0 \le C_1\left(1+\frac{C_1}{c_2(h)}\right)\left(\tilde{C}_2\left(1+\sqrt{3\left(\chi+\mu(-(\chi B-D))/\lambda_N^{(N)}\right)}\right)h\,+\right.$$

$$\left.+\,\tilde{C}_1\sqrt{3\left(\chi+\mu(-(\chi B-D))/\lambda_N^{(N)}\right)}\right)h|u|_2 \le$$

$$\le C_1\left(1+\frac{C_1}{C_0(h)}\right)h|u|_2\,\times$$

$$\times\left(\tilde{C}_2\left(1+\sqrt{3\left(\chi+\mu(-(\chi B-D))/\lambda_N^{(N)}\right)}\right)h+\tilde{C}_1\sqrt{3\left(\chi+\mu(-(\chi B-D))/\lambda_N^{(N)}\right)}\right), \qquad (28)$$

где $\tilde{C}_1>0$ *и* $\tilde{C}_2>0$ *— некоторые константы, не зависящие от* h, *и* $\{\alpha_i\}_{i=1}^N$.

Доказательство. Воспользуемся идеей приема Обэна – Нитче [28, 13]. Рассмотрим задачу $L^*v=u-u_h$ с граничными условиями $v(0)=v(1)=0$ (здесь $L^*v=-kv'-v''+cv$ — формально сопряженный к L оператор, а погрешность МПГ $u-u_h\in H_0^1[0;1]$ выступает в роли правой части уравнения). Ее обобщенное (слабое) решение v, как следует из общих теорем о разрешимости и регулярности в теории эллиптических граничных задач (см. гл. 8 в [14], главы 11 и 12 в [29], а также [15, 18]), существует, единственно и $v\in H_0^1[0;1]\bigcap H^2[0;1]$, причем существуют такие $\tilde{C}_1>0$ и $\tilde{C}_2>0$ (не зависящие от решения v и правой части $u-u_h$), что

$$|v|_1 \le \tilde{C}_1\|u-u_h\|_0 \quad \text{и} \quad |v|_2 \le \tilde{C}_2\|u-u_h\|_0. \qquad (29)$$

Умножая уравнение $L^*v=u-u_h$ скалярно в $L_2[0;1]$ на любое $w\in H_0^1[0;1]$, получаем соотношение $a(w,v)=(u-u_h,w)_0$. Отсюда, в частности, при $w=u-u_h$ имеем

$$a(u-u_h,v) = \|u-u_h\|_0^2. \qquad (30)$$

Из (3) и (6) для любого $v_h\in\Psi_h$ получаем

$$a(u-u_h,v_h) = a(u,v_h)-a(u_h,v_h) = (f,v_h)_0-(f,v_h)_0=0.$$

Поэтому (30) можно записать следующим образом: $a(u-u_h,v-v_h) = \|u-u_h\|_0^2 \quad \forall v_h\in\Psi_h$. Воспользовавшись непрерывностью билинейной формы a (см. (4)), отсюда получим

$$\|u-u_h\|_0^2 \le C_1|u-u_h|_1|v-v_h|_1. \qquad (31)$$

В качестве $v_h\in\Psi_h$ выберем функцию, которая интерполирует v в узлах. Тогда из леммы 5, использовав неравенства (29), будем иметь

$$|v - v_h|_1 \le \left(1 + \sqrt{3\left(\chi + \mu(-(\chi B - D))/\lambda_N^{(N)}\right)}\right) h |v|_2 +$$

$$+ |v|_1 \sqrt{3\left(\chi + \mu(-(\chi B - D))/\lambda_N^{(N)}\right)} \le$$

$$\le \left(\tilde{C}_2 \left(1 + \sqrt{3\left(\chi + \mu(-(\chi B - D))/\lambda_N^{(N)}\right)}\right) h +\right.$$

$$\left. + \tilde{C}_1 \sqrt{3\left(\chi + \mu(-(\chi B - D))/\lambda_N^{(N)}\right)} \right) ||u - u_h||_0. \tag{32}$$

Использовав теперь вместе (31) неравенства (32) и (22), получим оценку (28).

Теорема доказана.

***Следствие* 7.** *Если справедливы условия пункта* 1 *следствия* 6, *то* $\sqrt{3\left(\chi + \mu(-(\chi B - D))/\lambda_N^{(N)}\right)}$ *в оценке* (28) *можно заменить выражением* $\sqrt{3\left(\chi - \tilde{C}_{DB}/\lambda_N^{(N)}\right)}$.

***Следствие* 8.** *Если* $\chi + \mu(-(\chi B - D))/\lambda_N^{(N)} = O(h^2)$ *при* $h \to 0$, *выполняются условия теоремы* 4 *и* $C_0(h)$ *равномерно отделена от нуля*, *то из* (28) *следует сходимость МПГ в* $L_2[0;\,1]$ *со вторым порядком точности*, *т. е.* $||u - u_h||_0 = O(h^2)$ *при* $h \to 0$.

Выполнение условия $\chi + \mu(-(\chi B - D))/\lambda_N^{(N)} = O(h^2)$ можно обеспечить, если стабилизирующие параметры α_i будут величинами $O(h)$ (см. также замечания к лемме 3 в работе [10]). При проведении практических расчетов обычно выполняется $\alpha_i = O(h)$, иначе уровень вносимой МПГ искусственной вязкости может оказаться слишком большим (в особенности при больших значениях k, см. ниже формулу (33)), что, как правило, приводит к существенным потерям точности [7, 21, 28]. Кроме того, как видно из (26), стремление $\chi + \mu(-(\chi B - D))/\lambda_N^{(N)}$ к нулю при $h \to 0$ обеспечивает сходимость аппроксимаций функциями из Ψ_h. В частности, для $\alpha_{\max}$, определяемого формулой (25), справедливо $\alpha_{\max} = O(h)$. Следовательно, при одинаковых $\alpha_i = \alpha_{\max}$ и $\chi = \alpha^2$ (см. пункт 5) получаем $||u - u_h||_0 = O(h^2)$ при $h \to 0$.

***Следствие* 9.** 1. *Теоремы* 3 *и* 4 (*и*, *соответственно*, *оценки* (22), (28)) *остаются справедливыми*, *если при выполнении их условий в качестве величины* $C_0(h)$ *в них будет взята правая часть неравенства* (19), (20) *или* (21) (*см. пункт* 1 *следствия* 4). *Аналогично*, *при выполнении условий замечания* 3 *в качестве* $C_0(h)$ *могут быть взяты правые части выражений* (18) *и* (20), *где числитель заменен выражением*

$$\xi - kh\,\mu(-(A_0 - \delta B))/(2\lambda_N^{(N)}) - ch^2\mu(-(M - \phi B))/(4\lambda_N^{(N)}).$$

2. *Требования выполнения условий пунктов* 2 *лемм* 3 *и* 4 *в теореме* 3 *можно заменить* (*по отдельности или одновременно*) *требованиями выполнения условий пунктов* 4 *данных лемм*, *при этом величина* $C_0(h)$ *в* (22) *определяется выражением* (18) *с учетом пунктов* 2 *и* 3 *следствия* 4 *соответственно*. *Теорема* 3 *в таком случае остается в силе*.

3. *Требование выполнения условий пункта* 2 *леммы* 3 *в теореме* 4 *можно заменить требованием выполнения условий пункта* 4 *леммы* 3; *при этом величина* $C_0(h)$ *определяется через* (18) *с учетом пункта* 2 *следствия* 4. *Если, кроме того, дополнительно потребовать выполнения условий пункта* 4 *леммы* 4, *где величина* $C_0(h)$ *при этом определяется через* (18) *с учетом пункта* 3 *следствия* 4 (*см. пункт* 2 *следствия* 9), *а выражение* $\sqrt{3\left(\chi+\mu(-(\chi B-D))/\lambda_N^{(N)}\right)}$ *в* (28) *заменено на* $\sqrt{3\left(\chi+\mu(-(\chi B-D))/\lambda_1^{(N)}\right)}$ *или* $\sqrt{3\left(\chi-\tilde{C}_{DB}/\lambda_1^{(N)}\right)}$ (*см. пункт* 2 *следствия* 6), *то теорема* 4 *в таком случае остается в силе*.

***Замечание* 6.** Требование $u\in H_0^1[0;1]\bigcap H^2[0;1]$ в условиях теорем 3 и 4 не является ограничительным, поскольку данное условие, как следует из известных результатов о разрешимости и регулярности в теории эллиптических граничных задач (см. доказательство теоремы 4), при $f\in L_2[0;1]$ выполнено автоматически. Данные результаты также устанавливают существование таких констант $\tilde{C}_1$ и $\tilde{C}_2$, при которых справедливы соотношения (29). Отметим, что константой $\tilde{C}_2$ также можно оценить решение u через правую часть f: $|u|_2\le\tilde{C}_2\|f\|_0$. Данные константы для рассматриваемых в работе задач несложно найти в явном виде. Действительно, из (3) при $v=u$ и неравенства Коши – Буняковского получаем $|u|_1^2+c\|u\|_0^2=(f,u)_0\le\|f\|_0\|u\|_0$. Используя неравенство Фридрихса (5), отсюда получаем $\|u\|_0\le(\pi^2+c)^{-1}\|f\|_0$. Далее, $|u|_1^2\le\|f\|_0\|u\|_0\le(\pi^2+c)^{-1}\|f\|_0^2$, откуда имеем неравенство $|u|_1\le(\pi^2+c)^{-1/2}\|f\|_0$. Наконец, из уравнения (2) получаем неравенства

$$|u|_2=\|ku'+cu-f\|_0\le k|u|_1+c\|u\|_0+\|f\|_0\le(1+k(\pi^2+c)^{-1/2}+c(\pi^2+c)^{-1})\|f\|_0.$$

Данные оценки (как видно из самого их вывода) имеют место и для задачи с оператором L^*. Отсюда получаем $\tilde{C}_1=(\pi^2+c)^{-1/2}$ и $\tilde{C}_2=1+k(\pi^2+c)^{-1/2}+c(\pi^2+c)^{-1}$. Таким образом, поскольку можно явно оценить $|u|_2$ через L_2- норму правой части f, в правых частях априорных оценок (22) и (28) не остается неопределенных (неизвестных) величин.

***Замечание* 7.** Отметим, что одномерность задачи и предположение о постоянстве коэффициентов k и c позволили получить выведенные в работе оценки в относительно законченном и явном аналитическом виде и, таким образом, обобщить более ранние результаты в данном направлении. Но в более сложных случаях (например, для многомерных задач, особенно с переменными коэффициентами, задач со сложной геометрией области) проведение подобного анализа весьма затруднительно. Поэтому актуальной становится разработка численных процедур нахождения приближенных оценок (особенно, численной проверки того, удовлетворяется ли ключевое для стабилизации метода условие $c_2(h)>0$ при выбранной паре пространств базисных и весовых функций). Подобные процедуры, базирующиеся на решении обобщенных задач на собственные значения, разработаны в [30 – 33].

7. Численные примеры. Рассмотрим уравнение (2) при $f(x)=-k-cx$. Для оценки уклонения численного решения u_h от точного u используем величины $\mathrm{err}_{L_2}\equiv\|u-u_h\|_0$,

$\mathrm{err}_{H^1} \equiv |u - u_h|_1$ и $\mathrm{err}_C \equiv \max_{1\le i\le N} |u(x_i) - u_h(x_i)|$ и для простоты положим $\alpha_i = \alpha = \mathrm{const}$. Будем сравнивать два варианта выбора α_i: 1) $\alpha_i = kh/6$; как показано в [11, 13], такой выбор для задач КД (без реактивного члена) является H^1-оптимальным в том смысле, что минимизирует величину err_{H^1} (и, кроме того, для задач КД дает 4-й порядок локальной погрешности аппроксимации; см. ниже приложение); 2) $\alpha_i = \alpha_{\max}$, определяемое выражением (25). Для значений параметров $k = 10^3$, $c = 10^4$ (реактивная составляющая преобладает над конвективной), $k = 10^4$, $c = 10^4$ (реактивная и конвективная составляющие равнозначны) и $k = 10^4$, $c = 10^3$ (конвективная составляющая преобладает над реактивной) результаты расчетов приведены в таблицах 1, 2 и 3 соответственно. Расчеты показывают полное преимущество (в смысле величин погрешностей, характеризуемых err_{H^1}, err_{L_2} и err_C) выбора $\alpha_i = \alpha_{\max}$ перед соответствующим выбором параметров α_i без учета реакционных процессов, что является численным подтверждением результатов пункта 5. Отметим, что точное решение данной задачи (когда k и c принимают заданные выше значения) формирует пограничный слой при $x = 1$. Эмпирические порядки сходимости P_{H^1} и P_{L_2} определялись как $\ln\left(\left|u - u_{h_{i-1}}\right|_1 / \left|u - u_{h_i}\right|_1\right)/\ln(h_{i-1}/h_i)$ и $\ln\left(\left\|u - u_{h_{i-1}}\right\|_0 / \left\|u - u_{h_i}\right\|_0\right)/\ln(h_{i-1}/h_i)$ соответственно (см. работы [34, 35]), где h_i — шаг сетки при i-м расчете (i-я строка таблицы). Видно, что величины P_{H^1} и P_{L_2} при уменьшении шага монотонно приближаются слева к теоретическим порядкам скоростей сходимости в нормах $|\cdot|_1$ и $\|\cdot\|_0$ в рассматриваемом случае — числам 1 и 2 (см. следствия 5 и 8). Возникающие системы линейных алгебраических уравнений решались методом трехточечной прогонки [26].

Таблица 1. Значения погрешностей при $k = 10^3$, $c = 10^4$

Число узлов сетки	$\alpha_i = kh/6$			$\alpha_i = \alpha_{\max}$				
	err_{H^1}	err_{L_2}	err_C	err_{H^1}	P_{H^1}	err_{L_2}	P_{L_2}	err_C
50	21,8033877	0,1299759	0,5554216	21,6694486	—	0,1123325	—	0,4485942
100	20,4965453	0,0620007	0,3054484	20,4329138	0,084	0,0596124	0,901	0,2723953
200	17,6589109	0,0273827	0,0975694	17,6462063	0,210	0,0271254	1,128	0,0906496
300	15,0455446	0,0157007	0,0354665	15,0428464	0,392	0,0156379	1,353	0,0332544
400	12,8419161	0,0100862	0,0146000	12,8412559	0,548	0,0100641	1,528	0,0137153
500	11,0667703	0,0069638	0,0066201	11,0665836	0,665	0,0069541	1,653	0,0062107
600	9,6547301	0,0050662	0,0032208	9,6546704	0,747	0,0050613	1,739	0,0030105
700	8,5264575	0,0038363	0,0016444	8,5264364	0,805	0,0038336	1,799	0,0015274
800	7,6145006	0,0029983	0,0008634	7,6144927	0,846	0,0029967	1,842	0,0007942
900	6,8672685	0,0024039	0,0004563	6,8672654	0,876	0,0024028	1,873	0,0004132
1000	6,2465928	0,0019680	0,0002358	6,2465916	0,898	0,0019674	1,896	0,0002079
2000	3,2361556	0,0005098	$4{,}54\cdot10^{-5}$	3,2361555	0,948	0,0005097	1,947	$4{,}38\cdot10^{-5}$
4000	1,6329823	$1{,}2861\cdot10^{-4}$	$1{,}71\cdot10^{-5}$	1,6329822	0,986	$1{,}2860\cdot10^{-4}$	1,986	$1{,}70\cdot10^{-5}$

Таблица 2. Значения погрешностей при $k = 10^4$, $c = 10^4$

Число узлов сетки	$\alpha_i = kh/6$			$\alpha_i = \alpha_{\max}$				
	err_{H^1}	err_{L_2}	err_C	err_{H^1}	P_{H^1}	err_{L_2}	P_{L_2}	err_C
50	70,6844912	0,40291860	0,9412584	70,6749471	—	0,35463746	—	0,9230437
100	70,5988295	0,20606444	0,8879850	70,5898238	0,002	0,19782213	0,830	0,8790427
200	70,2823043	0,10233396	0,7875731	70,2745933	0,006	0,10127405	0,959	0,7836055
500	68,5273444	0,04032875	0,5495296	68,5225320	0,028	0,04026159	1,003	0,5483965
1000	64,3588196	0,01939313	0,3026330	64,3566624	0,090	0,01938518	1,053	0,3022944
2000	55,3408874	0,00858221	0,0978308	55,3404638	0,218	0,00858138	1,175	0,0977611
4000	40,1171762	0,00315686	0,0156596	40,1171529	0,464	0,00315678	1,442	0,0156508
8000	23,7380785	$9{,}3744 \cdot 10^{-4}$	$1{,}325 \cdot 10^{-3}$	23,7380781	0,757	$9{,}3743 \cdot 10^{-4}$	1,751	$1{,}324 \cdot 10^{-3}$
12000	16,4529008	$4{,}33401 \cdot 10^{-4}$	$2{,}506 \cdot 10^{-4}$	16,4529006	0,904	$4{,}3340 \cdot 10^{-4}$	1,903	$2{,}505 \cdot 10^{-4}$
16000	12,5183505	$2{,}47361 \cdot 10^{-4}$	$7{,}652 \cdot 10^{-5}$	12,5183504	0,950	$2{,}4736 \cdot 10^{-4}$	1,949	$7{,}648 \cdot 10^{-5}$
20000	10,0829896	$1{,}59403 \cdot 10^{-4}$	$3{,}165 \cdot 10^{-5}$	10,0829896	0,969	$1{,}59403 \cdot 10^{-4}$	1,969	$3{,}163 \cdot 10^{-5}$

Таблица 3. Значения погрешностей при $k = 10^4$, $c = 10^3$

Число узлов сетки	$\alpha_i = kh/6$			$\alpha_i = \alpha_{\max}$				
	err_{H^1}	err_{L_2}	err_C	err_{H^1}	P_{H^1}	err_{L_2}	P_{L_2}	err_C
50	70,6806045	0,39637422	0,9397343	70,6796952	—	0,39144786	—	0,9380290
100	70,5956424	0,20606364	0,8879846	70,5947364	0,002	0,20519266	0,918	0,8870864
200	70,2791033	0,10233391	0,7875732	70,2783298	0,006	0,10222640	0,998	0,7871756
500	68,5240651	0,04032864	0,5495305	68,5235833	0,028	0,04032191	1,012	0,5494171
1000	64,3553520	0,01939295	0,3026373	64,3551361	0,090	0,01939215	1,055	0,3026035
2000	55,3369974	0,00858196	0,0978422	55,3369550	0,218	0,00858188	1,175	0,0978352
4000	40,1131675	0,00315666	0,0156705	40,1131652	0,464	0,00315665	1,442	0,0156696
8000	23,7351649	$9{,}37355 \cdot 10^{-4}$	$13{,}292 \cdot 10^{-4}$	23,7351648	0,757	$9{,}37354 \cdot 10^{-4}$	1,751	$13{,}292 \cdot 10^{-4}$
12000	16,4507766	$4{,}33361 \cdot 10^{-4}$	$25{,}253 \cdot 10^{-5}$	16,4507765	0,904	$4{,}3336 \cdot 10^{-4}$	1,903	$25{,}252 \cdot 10^{-5}$
16000	12,5167033	$2{,}47339 \cdot 10^{-4}$	$77{,}573 \cdot 10^{-6}$	12,5167032	0,950	$2{,}47338 \cdot 10^{-4}$	1,949	$77{,}569 \cdot 10^{-6}$
20000	10,0816509	$1{,}59388 \cdot 10^{-4}$	$32{,}340 \cdot 10^{-6}$	10,0816509	0,969	$1{,}59388 \cdot 10^{-4}$	1,969	$32{,}338 \cdot 10^{-6}$

Приложение. Локальная погрешность и монотонность разностных схем МПГ. Определим локальную погрешность [26] разностной схемы

$$(L_h y)(x_i) \equiv k\,\frac{y_{i+1} - y_{i-1}}{2h} - \left(1 + \frac{kh\alpha_i}{2}\right)\frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} +$$

$$+ c\left(\left(\frac{1}{6} + \frac{\alpha_i}{4}\right)y_{i-1} + \frac{2}{3}y_i + \left(\frac{1}{6} - \frac{\alpha_i}{4}\right)y_{i+1}\right) =$$

$$= \frac{1}{h}(f, W_i)_0\,, \qquad 1 \le i \le N\,, \quad y_0 = y_{N+1} = 0\,, \tag{33}$$

которая получается после применения МПГ (см. следствие 2), предположив при этом (для возможности использования рядов Тейлора), что $u \in C^{(6)}(0; 1)$ и $f \in C^{(4)}(0; 1)$ (такую высокую гладкость мы требуем только в данном пункте для обеспечения возможности выписывания членов $O(h^4)$). Здесь L_h — разностный оператор, определенный соотношением (33) на пространстве сеточных функций [26]. Под выражением $L_h u$ будем понимать действие данного оператора на (точечную) сеточную проекцию решения u. Воспользовавшись разложениями в ряды Тейлора и теоремой о среднем для оценки интегралов $(f, W_i)_0$, после приведения всех подобных членов получим локальную погрешность аппроксимации $\psi(x_i) \equiv$ $\equiv (f, W_i)_0/h - (L_h u)(x_i)$ в узле x_i разностной схемы (33) на решении $u(x)$ исходной дифференциальной задачи

$$\psi(x_i) \;=\; \left(\frac{kh^2}{12} - \frac{\alpha_i h}{2}\right) f'(x_i) + \left(\frac{ch\alpha_i}{2} - \frac{ckh^2}{12}\right) u'(x_i) \;+$$

$$+ \left(\frac{kh\alpha_i}{2} - \frac{ch^2}{12} - \frac{k^2h^2}{12}\right) u''(x_i) + O(\alpha_i h^3 + h^4)\,.$$

Из полученного выражения для $\psi(x_i)$ видим, что при $\alpha_i = O(h)$ (см. следствие 8) справедливо $\psi(x_i) = O(h^2)$ (в том числе это будет, когда $\alpha_i = \alpha_{\max}$). Однако в частном случае при $c = 0$ и $\alpha_i = \alpha_{\max} = kh/6$ получаем $\psi(x_i) = O(h^4)$. Последний факт (для $c = 0$) был установлен еще Гриффитсом и Лоренцом в [11] (см. также [13] и комментарии в [21] по поводу подобного выбора α_i, но уже для случая нестационарных уравнений). Докажем лемму, из которой будут следовать важные утверждения о качественном поведении аппроксимации МПГ.

Лемма 6. *При выполнении неравенств*

$$-1 - (1+\alpha_i)\frac{kh}{2} + ch^2\left(\frac{1}{6} + \frac{\alpha_i}{4}\right) \;\le\; 0\,, \quad -1 + (1-\alpha_i)\frac{kh}{2} + ch^2\left(\frac{1}{6} - \frac{\alpha_i}{4}\right) \;\le\; 0 \tag{34}$$

(*при* $1 \le i \le N$) *матрица* A_h, *определяемая выражением* (8), *является* M*-матрицей* [3, 36].

Доказательство. При $c = 0$ из неравенств (34) следует результирующее условие $\alpha_i \ge 1 - 2/(kh)$, при выполнении которого, как показано в [10] (см. лемму 5 в [10]), матрица A_h является M-матрицей. Значит, для завершения доказательства остается лишь рассмотреть случай, когда $c > 0$. Отметим, что неравенства (34) обеспечивают выполнение условий $(A_h)_{i,j} \le 0$ при $i \ne j$ и, таким образом, необходимые условия, чтобы A_h была M-матрицей, выполнены. Далее воспользуемся следующим критерием для M-матриц [36]: некая матрица $A = \{a_{i,j}\}_{i,j=1}^N$ с $a_{i,j} \le 0$ при $i \ne j$ является M-матрицей тогда и только тогда, когда существует вектор $\vec{e}$ с положительными компонентами, для которого вектор $A\vec{e}$ также является вектором с положительными компонентами. Легко видеть, что в данном случае в качестве такого вектора $\vec{e}$ можно взять вектор $\vec{e} = (1, \ldots, 1)^T$. Действительно,

$$(A_h\vec{e})_1 = \frac{1}{h} + \frac{k}{2} + \frac{k\alpha_1}{2} + \frac{5ch}{6} - \frac{ch\alpha_1}{4} =$$

$$= \left(\frac{1}{h} + \frac{k}{2} + \frac{k\alpha_1}{2} - \frac{ch}{6} - \frac{ch\alpha_1}{4}\right) + ch \geq 0 + ch > 0,$$

$$(A_h\vec{e})_N = \frac{1}{h} - \frac{k}{2} + \frac{k\alpha_N}{2} + \frac{5ch}{6} + \frac{ch\alpha_N}{4} =$$

$$= \left(\frac{1}{h} - \frac{k}{2} + \frac{k\alpha_N}{2} - \frac{ch}{6} + \frac{ch\alpha_N}{4}\right) + ch \geq 0 + ch > 0,$$

$(A_h\vec{e})_i = ch > 0$ при $1 < i < N$. Значит, A_h является M-матрицей.

Лемма доказана.

***Замечание* 8.** Неравенства (34) гарантируют устойчивость решения (в смысле отсутствия ложных нефизических осцилляций) разностной схемы (33), обеспечивая для нее выполнение условий разностного принципа максимума (и, соответственно, монотонность) [26]. Отметим, что при $\alpha_i = 0$ и $c = 0$ условия (34) приводят к известному условию $kh \leq 2$ устойчивости классического метода Галеркина (см. главу 7 в [37]). Нарушение условий (34) может привести к возникновению в численном решении больших погрешностей и ложных (нефизических) осцилляций (см. [10], где подробно описан смысл данных условий применительно к уравнению КД и возникающие последствия при их нарушении).

***Замечание* 9.** С помощью леммы 3 можно получить простую оценку возмущения решения МПГ в узлах сетки в случае, когда правая часть $\vec{f}$ (см. следствие 2) вычислена не точно, а с некоторой погрешностью $\vec{\varepsilon}_f$ (например, вследствие применения каких-нибудь квадратурных формул при вычислении интегралов $(f, W_i)_0$). Действительно, тогда евклидова норма $\|\cdot\|$ вектора возмущения не превышает [23, 36] величину $\left\|A_h^{-1}\right\| \cdot \left\|\vec{\varepsilon}_f\right\|$. Если $\mu(-A_h) < 0$, то $\left\|A_h^{-1}\right\| \leq -1/\mu(-A_h)$ (см. [25]). Теперь если выполняется неравенство $\tilde{C}_{AB} > 0$ (при $\xi = 0$ и фиксированном h), то из свойств 3 и 5 ЛН, равенства $A_h = (A_1 + A)/h$ и пункта 3 леммы 3 получаем $-\mu(-A_h) = -\mu(-A)/h \geq \tilde{C}_{AB}/h > 0$, откуда $\left\|A_h^{-1}\right\| \leq -h/\mu(-A) \leq h/\tilde{C}_{AB}$.

***Замечание* 10.** Пусть при всех допустимых k и c (h считаем зафиксированным) выполнены условия теоремы 2, тогда $C_0(h) \geq \tilde{C}_0(h) \equiv \xi/\sqrt{1+3\chi}$. Допустим также, что $\tilde{C}_0(h) > 0$. Тогда если $(\tilde{C}_0(h))^{-1} = O(k^{s_1})$ (при $k \to +\infty$) и $(\tilde{C}_0(h))^{-1} = O(c^{s_2})$ (при $c \to +\infty$), то из соотношений $C_1 = O(k)$, $C_1 = O(c)$ получаем $C_1/\tilde{C}_0(h) = O(k^{s_1+1})$ и $C_1/\tilde{C}_0(h) = O(c^{s_2+1})$. Отсюда следует, что если $s_1 + 1 \leq 0$ и $s_2 + 1 \leq 0$, то в силу очевидного неравенства $1 \le C_1/c_2(h) \le C_1/C_0(h) \le C_1/\tilde{C}_0(h)$ (следствие из $C_1 \ge c_2(h) \ge C_0(h) \ge \tilde{C}_0(h)$) множители $1 + C_1/c_2(h)$ и $1 + C_1/C_0(h)$ в оценках теорем 3 и 4 являются ограниченными при неограниченном возрастании k и c (этот же результат был получен в частном случае для одинаковых α_i, см. пункт 2 замечания 5; при этом для s_1 и s_2 будет $s_1 = s_2 = -1$).

Заключение. В данной работе рассматриваются вопросы точности (в пространстве L_2 и соболевском пространстве H^1) численного интегрирования стационарного одномерного уравнения КДР конечноэлементным МПГ с кусочно-линейными базисными функциями и кусочно-квадратичными весовыми функциями типа (1), где каждая функция W_i имеет индивидуальный стабилизирующий параметр α_i. Отметим, что результаты и выкладки статьи без существенных затруднений могут быть перенесены и на другие весовые функции и классы таких функций (например, практически без изменений переносятся на классы весовых функций, введенные в работах [20, 21]). На основании развиваемого в статье подхода, основанного на исследовании с помощью аппарата логарифмических норм матричных неравенств специального вида (которые составляются для матриц, порожденных билинейной формой задачи и полунормой $|\cdot|_1$ в пространстве весовых функций (лемма 2), относительно матрицы, порожденной полунормой $|\cdot|_1$ в пространстве базисных функций), для величины $c_2(h)$ получены оценки снизу (теорема 2) в зависимости от выбора набора параметров $\{\alpha_i\}$, которыми, главным образом, и определяется погрешность метода при фиксированном пространстве базисных функций. Выведены оценки точности МПГ в пространствах H^1 (теорема 3) и L_2 (теорема 4). Подробно исследован случай с $\alpha_i = \alpha = \mathrm{const}$ для уравнения КДР (пункт 5) и результаты исследований подтверждены расчетными данными (пункт 7).

В настоящей работе коэффициенты k и c в исследуемом уравнении КДР для упрощения выкладок были положены постоянными (что также позволило получить достаточно „тонкие" априорные оценки точности МПГ в относительно законченном и явном аналитическом виде и найти явные выражения для констант, входящих в эти оценки). Соответствующий одномерный случай с переменными коэффициентами является технически более сложным, однако развиваемый в работе подход может быть без затруднений принципиального характера применен к его исследованию (в этом случае при нахождении оценок для $c_2(h)$ изменится только конкретный вид матрицы A_h). При этом схема рассуждений и доказательств, а также неравенства (и соответствующие условия теорем), записанные в общем (абстрактном) виде через логарифмические нормы, останутся в силе. Это же касается и случая использования в МПГ неравномерных сеток, а также решения многомерных стационарных задач КДР.

Автор выражает глубокую благодарность кандидату техн. наук Сальникову Н. Н. и профессору Молчанову А. А. за плодотворное обсуждение данной работы и ценные замечания к ней, а также профессору Чертову О. Р. и чл.-кор. НАН Украины Губареву В. Ф. за внимание к исследованиям автора.

1. *Finlayson B. A.* Numerical methods for problems with moving fronts. – Seattle, Washington USA: Ravenna Park Publ., Inc., 1992. – 613 p.

2. *Дейнека В. С., Сергиенко И. В., Скопецкий В. В.* Математические модели и методы расчета задач с разрывными решениями. – Киев: Наук. думка, 1995. – 262 с.
3. *Roos H.-G., Stynes M., Tobiska L.* Robust numerical methods for singularly perturbed differential equations. – Berlin; Heidelberg: Springer-Verlag, 2008. – 604 p.
4. *Тихонов А. Н., Самарский А. А.* Уравнения математической физики. – М.: Наука, 1972. – 736 с.
5. *Самарский А. А., Вабищевич П. Н.* Численные методы решения задач конвекции-диффузии. – М.: Либроком, 2009. – 248 с.
6. *Grossmann C., Roos H.-G., Stynes M.* Numerical treatment of partial differential equations. – Berlin; Heidelberg: Springer-Verlag, 2007. – 596 p.
7. *Fries T. P., Matthies H. G.* A review of Petrov – Galerkin stabilization approaches and an extension to Meshfree methods. – Germany; Brunswick: Techn. Univ. Braunschweig, Informatikbericht-Nr., 2004. – 71 p.
8. *Hughes T. J. R., Scovazzi G., Tezduyar T. E.* Stabilized methods for compressible flows // J. Sci. Comput. – 2010. – **43**. – P. 343 – 368.
9. *John V., Schmeyer E.* Finite element methods for time-dependent convection-diffusion-reaction equations with small diffusion // Comput. Meth. Appl. Mech. Eng. – 2008. – **198**. – P. 475 – 494.
10. *Сирик С. В.* Точность и устойчивость метода Петрова – Галеркина при интегрировании стационарного уравнения конвекции-диффузии // Кибернетика и систем. анализ. – 2014. – **50**, № 2. – С. 132 – 143.
11. *Griffiths D. F., Lorenz J.* An analysis of the Petrov – Galerkin finite element method // Comput. Meth. Appl. Mech. Eng. – 1978. – **14**. – P. 39 – 64.
12. *Morton K. W.* Finite element methods for non-self-adjoint problems // Lect. Notes Math. / Ed. P. R. Turner: Proc. SERC Summer School, Lancaster (1981). – Berlin: Springer-Verlag, 1982. – **965**. – P. 113 – 148.
13. *Griffiths D. F.* Discretised eigenvalue problems, LBB constants and stabilization // Numer. Anal. – Edinburgh: Longman, 1996. – P. 57 – 75.
14. *Гилбарг Д., Трудингер М.* Эллиптические дифференциальные уравнения с частными производными второго порядка. – М.: Наука, 1989. – 464 с.
15. *Agmon S.* Lectures on elliptic boundary value problems. – Princeton, NJ: Van Nostrand, 1965. – 300 p.
16. *Митчелл Э., Уэйт Р.* Метод конечных элементов для уравнений с частными производными. – М.: Мир, 1981. – 216 с.
17. *Ректорис К.* Вариационные методы в математической физике и технике. – М.: Мир, 1985. – 590 с.
18. *Ладыженская О. А.* Краевые задачи математической физики. – М.: Наука, 1973. – 408 с.
19. *Babuska I., Aziz A. K.* Survey lectures on the mathematical foundations of the finite element method // The Mathematical Foundations of the Finite Element Method with Applications to Partial / Ed. A. K. Aziz. – New York: Acad. Press, 1972. – P. 2 – 363.
20. *Сальников Н. Н., Сирик С. В., Терещенко И. А.* О построении конечномерной математической модели процесса конвекции-диффузии с использованием метода Петрова – Галеркина // Проблемы управления и информатики. – 2010. – № 3. – С. 94 – 109.
21. *Сирик С. В., Сальников Н. Н., Белошапкин В. К.* Выбор весовых функций в методе Петрова – Галеркина для интегрирования линейных одномерных уравнений конвекции-диффузии // Управляющие системы и машины. – 2014. – № 1. – С. 38 – 47.
22. *Xu J., Zikatanov L.* Some observations on Babuska and Brezzi theories // Numer. Math. – 2003. – **94**. – P. 195 – 202.
23. *Хорн Р., Джонсон Ч.* Матричный анализ: Пер. с англ. – М.: Мир, 1989. – 655 с.
24. *Деккер К., Вервер Я.* Устойчивость методов Рунге – Кутты для жестких нелинейных дифференциальных уравнений: Пер. с англ. – М.: Мир, 1988. – 334 с.
25. *Desoer C. A., Haneda H.* The measure of a matrix as a tool to analyze computer algorithms for circuit analysis // IEEE Trans. Circ. Theory. – 1972. – **19**, № 5. – P. 480 – 486.
26. *Самарский А. А., Гулин А. В.* Численные методы математической физики. – 2-е изд. – М.: Науч. мир, 2003. – 316 с.
27. *Noschese S., Pasquini L., Reichel L.* Tridiagonal Toeplitz matrices: properties and novel applications // Numer. Linear Algebra and Appl. – 2013. – **20**, № 2. – P. 302 – 326.
28. *Brenner S. C., Scott R.* The mathematical theory of finite element methods. – Berlin; Heidelberg: Springer-Verlag, 2007. – 404 p.
29. *Хатсон В., Пим Дж.* Приложения функционального анализа и теории операторов: Пер. с англ. – М.: Мир, 1983. – 432 с.
30. *Bathe K.-J., Hendriana D., Brezzi F., Sangalli G.* Inf-sup testing of upwind methods // Int. J. Numer. Meth. Eng. – 2000. – **48**. – P. 745 – 760.
31. *Sangalli G.* Numerical evaluation of finite element methods in convection-diffusion problems // Calcolo. – 2000. – **37**. – P. 233 – 251.

32. *Sangalli G.* Numerical evaluation of FEM with application to the 1-D advection-diffusion problem // Math. Models and Meth. Appl. Sci. – 2002. – **12**, № 2. – P. 205 – 228.
33. *Buffa A., de Falco C., Sangalli G.* Isogeometric analysis: stable elements for the 2D Stokes equation // Int. J. Numer. Meth. Fluids. – 2011. – **65**. – P. 1407 – 1422.
34. *Абрамов Є., Квасниця Г., Шинкаренко Г.* Частинами квадратичні та кубічні апроксимації h -адаптивного МСЕ для одновимірних крайових задач // Вісн. Львів. ун-ту. Сер. прикл. математика та інформатика. – 2011. – Вип. 17. – С. 47 – 61.
35 *Трушевський В. М., Шинкаренко Г. А.* Розпаралелена апроксимація еліптичних крайових задач штучною нейромережею з радіально-базисними функціями // Вісн. Львів. ун-ту. Сер. прикл. математика та інформатика. – 2014. – Вип. 22. – С. 108 – 117.
36. *Berman A., Plemmons R. J.* Nonnegative matrices in the mathematical sciences. – Philadelphia: SIAM, 1994. – 340 p.
37. *Флетчер К.* Численные методы на основе метода Галеркина: Пер. с англ. – М.: Мир, 1988. – 352 с.

Получено 07.11.14,
после доработки — 06.05.15