

Университет Лимерика, Ирландия

На правах рукописи

Коптева Наталья Викторовна

**Апостериорные и априорные оценки
конечноэлементных решений некоторых
сингулярно возмущенных уравнений
на анизотропных сетках**

01.01.07 — вычислительная математика

ДИССЕРТАЦИЯ НА СОИСКАНИЕ УЧЕНОЙ СТЕПЕНИ
ДОКТОРА ФИЗИКО-МАТЕМАТИЧЕСКИХ НАУК

Научный консультант:
доктор физико-математических наук,
профессор В. Б. Андреев

ЛИМЕРИК — 2018

Содержание

Введение	6
----------	---

Глава 1

Апостериорные оценки для стационарных сингулярно возмущенных уравнений реакции-диффузии на локально квазиравномерных и анизотропных сетках	38
---------------------------------------------------------------------------------------------------------------------------------------------------	-----------

§1 Апостериорные оценки в норме максимума модуля на локально квазиравномерных сетках	38
---------------------------------------------------------------------------------------------	-----------

1.1 Введение	39
1.2 Существование и единственность решения исходной задачи. Оценки функции Грина	44
1.3 Вывод апостериорных оценок	52
1.4 Необходимость логарифма в апостериорных оценках в норме максимума модуля	68
1.5 Численные результаты	70

§2 Апостериорные оценки в норме максимума модуля на анизотропных сетках	77
--------------------------------------------------------------------------------	-----------

2.1 Введение	77
2.2 Основные предположения о триангуляции	82
2.3 Теорема о следах для анизотропных элементов	84
2.4 Представление ошибки через функцию Грина	86
2.5 Апостериорный анализ на частично структурированных анизотропных сетках	89
2.6 Апостериорный анализ на неструктурированных анизотропных сетках	99

2.7	Интерполяция функции Грина	117
2.8	Заключительные замечания	122

§3	Апостериорные оценки в энергетической норме на анизотропных сетках	123
3.1	Введение	123
3.2	Основные предположения о триангуляции	127
3.3	Теорема о следах для анизотропных элементов	129
3.4	Представление ошибки численного решения через невязки	130
3.5	Апостериорный анализ на частично структурированных анизотропных сетках	132
3.6	Апостериорный анализ на неструктурированных анизотропных сетках	135
3.7	Квазиинтерполяционный оператор P_h на анизотропных сетках	142
3.8	Численные результаты	148

Глава 2

	Апостериорные оценки в норме максимума модуля для полулинейных параболических уравнений с использованием эллиптических реконструкций	153
--	---------------------------------------------------------------------------------------------------------------------------------------------	------------

§4	Апостериорные оценки для полудискретных методов (без дискретизации по пространству)	153
4.1	Введение	154
4.2	Функция Грина для параболического оператора	157
4.3	Описание результатов для полудискретных методов (без дискретизации по пространству)	162
4.4	Полудискретный неявный метод Эйлера	163
4.5	Полудискретный метод Кранка-Николсон	165

4.6	Полудискретный разрывный метод Галеркина $dG(\mathbf{r})$ с квадратурой Радо	167
4.7	Доказательство Леммы 4.2	171
§5	Апостериорные оценки для полностью дискретных методов	177
5.1	Предположения о дискретизациях по пространству	178
5.2	Описание результатов для полностью дискретных методов	180
5.3	Полностью дискретный неявный метод Эйлера	188
5.4	Полностью дискретный метод Кранка-Николсон	195
5.5	Полностью дискретный разрывный метод Галеркина $dG(1)$	199
5.6	Эллиптические оцениватели	204
Глава 3		
Немонотонные эллиптические сингулярно возмущенные уравнения реакции-диффузии: асимптотический анализ и априорные оценки ошибки численного решения		
		208
§6	Априорные оценки ошибки численного решения в норме максимума модуля в гладкой области	208
6.1	Введение	209
6.2	Асимптотический анализ, верхние и нижние решения	214
6.3	Анализ численного метода	221
6.4	Численные результаты	232
6.5	Обобщения результатов параграфа	234
§7	Немонотонное полулинейное сингулярно возмущенное уравнение реакции-диффузии в многоугольной области	235
7.1	Введение	236
7.2	Пограничные и угловые функции	241

7.3	Существование и свойства угловых погранфункций	248
7.4	Возмущенное асимптотическое разложение, верхние и нижние решения, доказательство существования	271
7.5	Некоторые аспекты построения ε -равномерных численных ме- тодов	275
Заключение		277
Список литературы		279

Введение

Актуальность темы. Степень ее разработанности. Большое число задач физики и техники приводит к так называемым сингулярно возмущенным дифференциальным уравнениям, т. е. уравнениям, содержащим малый параметр в виде множителя при старших производных [119, 131, 132]. Структура решений для достаточно широких классов таких задач изучена с помощью асимптотических методов [16, 18, 19, 21, 30, 35, 36, 44, 45, 128, 147]. Хорошо известно, что для их решений характерны пограничные и внутренние слои; последние суть узкие участки области, в которых решение чрезвычайно быстро меняется, а его производные не являются ограниченными равномерно по малому параметру.

Кратко поясним механизм возникновения, например, пограничных слоев. Формально полагая малый параметр, присутствующий в уравнении, $\varepsilon = 0$, получаем так называемое вырожденное уравнение, которое имеет более низкий, по сравнению с исходным сингулярно возмущенным уравнением, порядок. Поэтому для вырожденного уравнения требуется меньшее, чем для исходного уравнения, число граничных/начальных условий. С одной стороны, при $\varepsilon \rightarrow 0$ решение исходной задачи, как правило, сходится к решению вырожденной задачи во внутренних точках области. С другой стороны, оно должно также удовлетворять неиспользованным в вырожденной задаче граничным условиям. Последнее приводит к образованию в окрестности точек границы, где заданы эти условия, пограничных слоев. Внутренние же слои (см. Рис. 1) порождаются аналогичным образом, с той лишь разницей, что они описывают переход между двумя частными решениями вырожденной задачи.

Также хорошо известно [28, 47, 119, 131, 132, 139], что точность классических численных методов, не учитывающих наличие в задаче малого параметра, может весьма существенно зависеть от значения данного параметра,

поэтому для достижения хорошей точности приходится использовать сетки с весьма большим числом узлов, тем большим, чем меньшие значения принимает этот параметр (см. Рис. 1). В связи с этим для сингулярно возмущенных задач разрабатываются специальные, так называемые робастные методы, точность которых не зависит от значений малых параметров.

Впервые робастность специальной схемы на равномерной сетке была установлена А. М. Ильиным [29] для одномерного сингулярно возмущенного уравнения типа конвекции-диффузии. Подобные специальные схемы, получившие название методов подгонки, были рассмотрены в ряде работ; см. [28, 102] и цитируемую там литературу. Аналогичный подход в рамках метода конечных элементов основан на использовании в качестве базисных кусочно-экспоненциальных функций; см., например, [40, 78, 140, 141, 142]. Обзор численных методов для задач данного типа, включая специальные варианты метода конечных элементов, приводится и в [6, 7]. Отметим также, что в [48] было доказано, что не существует схемы типа подгонки для

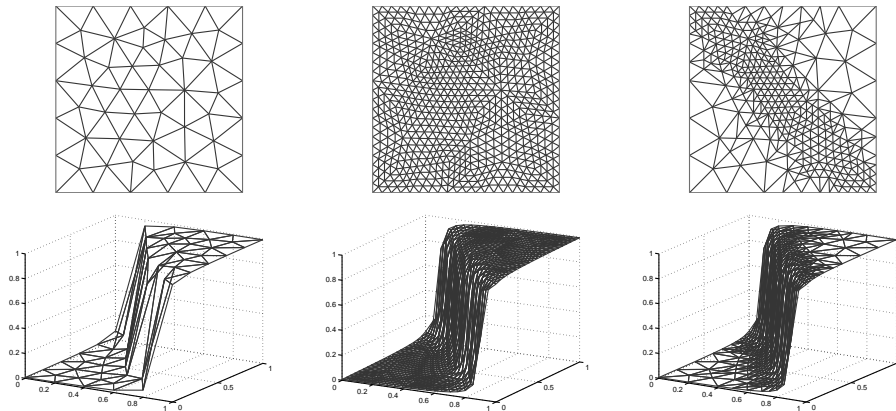


Рис. 1: Решение с внутренним слоем: низкая точность внутри слоя на стандартной квазиравномерной сетке (слева); высокая точность во всей области на квазиравномерной сетке с весьма большим числом узлов (в центре); локально квазиравномерное сгущение сетки приводит к существенной экономии числа степеней свободы (справа).

сингулярно возмущенного параболического уравнения, которая бы сходилась равномерно по малому параметру на равномерной сетке; см. также [118, §14].

Принципиально иной подход к построению робастных численных методов для сингулярно возмущенных задач, заключающийся в использовании локально сгущающихся сеток, был впервые теоретически обоснован Н. С. Бахваловым [8]. С точки зрения ошибки интерполяции, достаточно очевидно, что нет необходимости одинаково сгущать сетку во всей области, в то время как ее сгущение лишь в областях пограничных и внутренних слоев приводит к существенной экономии числа узлов сетки (см. Рис. 1, справа). Если же используются так называемые анизотропные сетки, элементы которых могут быть сколь угодно сильно сплющены, то высокая точность интерполяции во всей области может достигаться, даже если число узлов сетки не зависит от малого параметра (см. Рис. 2).

В то же время, хорошая точность интерполяции еще не гарантируют сходимость численных методов на соответствующих сетках. Построению специальных, сгущающихся в областях пограничных и внутренних слоев, сеток и их теоретическому исследованию для сингулярно возмущенных обыкновенных дифференциальных уравнений посвящена обширная литература; см.,

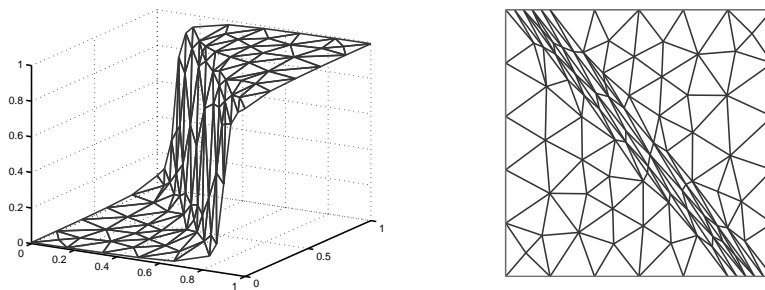


Рис. 2: Решение с внутренним слоем (слева); сетка, анизотропным образом сгущающаяся в области внутреннего слоя (справа); при этом высокая точность во всей области достигается при числе узлов, независимом от малого параметра.

например, [5, 9, 34, 91, 152, 154], а также монографию [111].

В контексте же сингулярно возмущенных уравнений в частных производных, построение специальных анизотропных сеток получило дальнейший импульс благодаря работам Г. И. Шишкина [49] (см. также [86, 118, 136, 178]), предложившего использовать специальные кусочно-равномерные сетки. Равномерные по малому параметру оценки ошибки решения на структурированных анизотропных сетках типа Бахвалова и Шишкина были установлены как самими этими авторами, так и в ряде последующих работ; см., например, [2, 3, 4, 10, 11, 12, 13, 66, 55, 155, 162], а также монографии [111, 132].

Таким образом, для ряда структурированных локально анизотропных сеток было теоретически установлено, что они позволяют получать надежные численные решения при использовании относительно небольших чисел степеней свободы, независящих от значений малых параметров. Отметим также, что построение таких сеток требует предварительного асимптотического анализа исходной задачи. Чтобы этого избежать, аналогичные сетки могут генерироваться с помощью соответствующих адаптивных процедур; см., например, [17, 98]. (В самом грубом варианте, например, градиенты численного решения могут использоваться при автоматической адаптации сетки.)

Заметим, что при теоретическом обосновании адаптивных алгоритмов генерации сеток, а также при выборе адаптационных критериев, как правило, используются так называемые апостериорные оценки ошибки численного решения. Последние суть теоретические оценки ошибки через само численное решение и локальные характеристики сетки, т. е. через величины, значения которых будут доступны в процессе численного решения задачи. Начало бурному развитию апостериорных оценок было положено новаторскими работами [57, 58]; см. монографии [52, 59, 148, 151].

Что касается апостериорных оценок для сингулярно возмущенных задач, то к ним также предъявляется требование робастности, под которой в данном

случае понимается то, что зависимость от малого параметра в таких оценках должна быть показана явно, при этом эффективность этих оценок не должна зависеть от малости данного параметра. Впервые робастные апостериорные оценки были получены Р. Ферфюртом [149] для сингулярно возмущенных уравнений реакции-диффузии, а в дальнейшем были им обобщены для уравнений конвекции-диффузии [146, 150]; см. также монографию [151, §§4.3–4.4]. Важно учитывать, что данные апостериорные оценки были получены в предположении о локальной квазировномерности сетки (см. Рис. 1, справа), а также в энергетических нормах.

С другой стороны, для таких уравнений норма максимума модуля представляется более подходящей, поскольку она является достаточно сильной (в отличие, например, от энергетической нормы) для того, чтобы обнаружить большую ошибку в зонах узких пограничных и внутренних слоев. Поэтому представляет большой интерес получение аналогичных робастных апостериорных оценок на локально квазиравномерных сетках, но в норме максимума модуля, что и будет сделано в §1.

Заметим, что апостериорные оценки получили развитие в рамках метода конечных элементов. При этом, несмотря на наличие большого количества работ по этой тематике, как правило, сетки предполагаются локально квазиравномерными. Это обусловлено тем, что теоретический аппарат метода конечных элементов плохо обобщается на случай неструктурированных анизотропных сеток, и большинство как априорных [43, 63], так и апостериорных [52, 59, 148, 151] оценок ошибки конечноэлементных решений неприменимы к случаю анизотропных сеток.

На анизотропных же сетках апостериорные оценки были установлены в работах [104, 105, 106, 107] (см. также [151, §4.5]). Отметим, что последние оценки были получены, во-первых, в энергетических нормах, а во-вторых, в них присутствуют так называемые функции согласования (*matching functions*). Последние являются функциями ошибки численного решения и, в общем

случае, могут принимать неограниченно большие значения на сильно анизотропных сетках (если только сетка не является правильно ориентированной по отношению к решению). Получение же робастных апостериорных оценок на анизотропных сетках, в которых функции согласования не присутствуют, ранее не рассматривалось. Этому вопросу и будут посвящены §2 и §3, где указанные оценки будут получены соответственно в норме максимума модуля и в энергетической норме.

Что касается апостериорных оценок для параболических уравнений в норме максимума модуля, то такие оценки были ранее получены главным образом для регулярных линейных задач [83, 62, 75, 77]; в случае же сингулярно возмущенных уравнений известные оценки не являются робастными, поскольку в них присутствуют нежелательные отрицательные степени малого параметра [62, 60]. Получению робастных апостериорных оценок в норме максимума модуля для сингулярно возмущенных уравнений будут посвящены §§4–5. При этом, наш подход будет применим к широкому классу временных дискретизаций, и в некоторых случаях будет приводить к новым результатам (или упрощать анализ) и для случая классических параболических уравнений.

Вернемся к обсуждению априорных оценок ошибки численных решений для случая анизотропных сеток. Как было отмечено выше, существует обширная литература по данной тематике, однако в большинстве случаев рассматриваются линейные уравнения. В случае же, например, полулинейных сингулярно возмущенных уравнений типа реакции-диффузии, как правило, предполагается их монотонность; см., например, [10, 14, 133]. Под немонотонными полулинейными эллиптическими уравнениями будут пониматься уравнения, в которых нелинейность по неизвестной функции не является монотонной, поэтому такие уравнения могут иметь несколько решений. Уравнения данного типа часто возникают при моделировании биологических и химических процессов [120, §14.7], [94, §2.3].

Отметим, что, в случае немонотонных сингулярно возмущенных уравнений стандартный технический аппарат, используемый при получении априорных оценок ошибки как для конечно-разностных, так и для конечноэлементных численных решений, неприменим. Поэтому при исследовании некоторых немонотонных сингулярно возмущенных обыкновенных дифференциальных уравнений в работах [70, 143, 182] был использован принципиально иной подход, основанный на построении дискретных верхних и нижних решений. Для случая же немонотонных уравнений в частных производных в некотором смысле аналогичный подход построения верхних и нижних решений, но при исследовании точности асимптотических разложений, был развит в работах Н. Н. Нефедова [37, 38]; см. также [20, 33, 39, 121, 122]. Объединению этих двух подходов с целью получения априорных оценок ошибки численных решений для немонотонных полулинейных эллиптических уравнений реакции-диффузии и будут посвящены §6 и §7.

При этом §6 будет посвящен случаю гладкой области, а §7 — существенно более сложному случаю выпуклой многоугольной области. В последнем случае открытым вопросом является даже построение и исследование точности асимптотического разложения для исходной задачи. Отметим, что в случае прямоугольной области последний вопрос был рассмотрен в работе [15] для линейного уравнения реакции-диффузии. Полулинейное же уравнение в прямоугольной области рассматривалась в работах [23, 24, 25, 26, 27], но при весьма серьезных ограничениях на данные в угловых точках области. В §7 данная задача будет рассмотрена, во-первых, в более общей многоугольной области, а во-вторых, при достаточно простых и, в некотором смысле, необходимых предположениях о данных задачи в углах области. Также будут рассмотрены аспекты построения робастных по отношению к малому параметру численных методов для данной задачи.

Таким образом, несмотря на наличие большого количества работ в области численных методов для сингулярно возмущенных уравнений, некоторые

важные вопросы остаются нерассмотренными. В частности, значительный интерес представляет развитие теоретического аппарата для получения робастных апостериорных оценок на анизотропных сетках. Также ранее не были получены робастные апостериорные оценки для сингулярно возмущенных параболических уравнений. Наконец, малоизученным является теоретическое обоснование точности численных методов для немонотонных сингулярно возмущенных уравнений, в особенности, в случае уравнений в частных производных.

Цели и задачи. Целью данной диссертации является теоретическое исследование конечноэлементных аппроксимаций некоторых полулинейных сингулярно возмущенных уравнений реакции-диффузии. Поскольку для таких задач характерны решения с резкими пограничными и внутренними слоями, особое внимание будет уделено анализу анизотропных сеток и оценкам в норме максимума модуля. Автором будет установлен ряд новых апостериорных и априорных оценок.

Научная новизна. В диссертации представлены новые более сильные и точные оценки точности численных решений сингулярно возмущенных задач, ряд из которых усиливает известные оценки и в регулярном случае. Получение этих результатов оказалось возможным исключительно благодаря разработанному автором новому теоретическому аппарату получения апостериорных оценок на анизотропных сетках, а также развитию автором аппарата теоретического исследования сходимости численных методов для немонотонных сингулярно возмущенных уравнений в частных производных.

Теоретическая и практическая значимость. Результаты диссертации имеют теоретическую направленность. Установленные апостериорные оценки могут быть использованы в адаптивных алгоритмах построения как локально квазиравномерных, так и анизотропных сеток. Теоретический ап-

парат, предложенный для получения апостериорных оценок на анизотропных сетках, может получить дальнейшее развитие при исследовании более сложных сингулярно возмущенных уравнений конвекции-диффузии. Теоретический же аппарат получения априорных оценок численного решения нелинейных многомерных уравнений реакции-диффузии, представленный в Главе 3, может быть использован при рассмотрении более сложных многомерных задач данного типа.

Методология и методы диссертационного исследования. При получении апостериорных оценок для конечноэлементных решений сингулярно возмущенных эллиптических уравнений в Главе 1 используется аппарат представления ошибки численного решения через невязки и функцию Грина исходного эллиптического оператора. Далее получены и использованы неулучшаемые оценки функции Грина, а также версии теоремы о следах для квазиравномерных и анизотропных элементов триангуляции. При получении оценок ошибки интерполяции для функции Грина используется теория квазиинтерполяционных операторов.

В Главе 2 при получении апостериорных оценок для численных решений параболических уравнений ключевой чертой используемого аппарата является интерполяция численного решения по времени, степень которой соответствует порядку метода (и является более низкой, по сравнению с типично используемыми в литературе интерполянтами). Также важную роль в представленном анализе играют оценки функции Грина параболического оператора. При рассмотрении полностью дискретных методов используется аппарат эллиптических реконструкций (аналогичный аппарату на основе проекции Рунца, часто используемому при получении априорных оценок для конечноэлементных решений параболических задач).

В Главе 3 при доказательстве существования и точности численных решений немонотонных эллиптических сингулярно возмущенным уравнений ис-

пользуется подход, основанный на построении дискретных верхних и нижних решений, при этом последние являются нелинейными модификациями асимптотического разложения для исходной задачи. При построении равномерно по параметру сходящихся численных методов используются сетки типа Бахвалова и Шишкина. Ключевым элементом анализа в §7 является исследование некоторого немонотонного полулинейного эллиптического уравнения в бесконечном секторе. При этом для оператора, возникающего при линеаризации данного уравнения и рассматриваемого в соответствующем секторе радиуса R , показано, что его главное собственное значение является положительным и отделено от нуля положительной постоянной равномерно по $R \rightarrow \infty$. Данный результат используется при построении асимптотических разложений, а также нижних и верхних решений в многоугольных областях.

Более подробно методы диссертационного исследования описаны в разделе введения, посвященном краткому описанию содержания работы.

Положения, выносимые на защиту:

1. Для полулинейных сингулярно возмущенных эллиптических уравнений реакции-диффузии в полигональных областях установлены явные апостериорные оценки ошибки численного решения на основе невязок в норме максимума модуля на локально квазиравномерных сетках. Постоянные в полученных оценках не зависят от диаметров элементов сетки и малого параметра.

2. Для полулинейных сингулярно возмущенных эллиптических уравнений реакции-диффузии в многоугольных областях получены явные апостериорные оценки ошибки численного решения в норме максимума модуля на неструктурированных анизотропных сетках. Постоянные в полученных оценках не зависят от диаметров элементов сетки, их аспектного соотношения (т. е. степени их сплюснутости) и малого параметра. Полученные апостериорные оценки являются новыми даже для уравнения Лапласа. Представ-

ленный подход обобщен для получения апостериорных оценок ошибки в энергетической норме.

3. Для полулинейных параболических уравнений второго порядка получены апостериорные оценки в норме максимума модуля для ошибки соответствующих численных решений. Рассмотрены временные полудискретизации и полностью дискретные методы на основе неявного метода Эйлера, метода Кранка-Николсон и разрывного метода Галеркина $dG(r)$ с квадратурой Радо.

4. Для немонотонных полулинейных сингулярно возмущенных эллиптических уравнений реакции-диффузии в гладких областях исследованы численные решения на сгущающихся в пограничных слоях сетках типа Бахвалова и Шишкина. Доказано существование решений соответствующих нелинейных дискретных задач и установлен второй порядок сходимости (с логарифмическим множителем в случае сетки Шишкина) в сеточной норме максимума модуля равномерно по малому параметру ε при условии $\varepsilon \leq Ch$, при этом как число степеней свободы не превосходит Ch^{-2} .

5. Для немонотонного полулинейного сингулярно возмущенного эллиптического уравнения реакции-диффузии в выпуклой многоугольной области построено асимптотическое разложение и установлено существование решения исходной задачи в окрестности построенного асимптотического разложения.

Личный вклад автора. Отметим, что вклад автора являлся определяющим при получении всех научных результатов, выносимых на защиту. Основные результаты, представленные в §2, §3 и §6, были получены и опубликованы автором без соавторов [168, 169, 163]. В остальных четырех параграфах представлены результаты, полученные совместно с А. Демловым [157], Р. Б. Келлоггом [161] и Т. Линсом [174, 175, 177] (последний был привлечен к данным исследованиям в рамках работы, под руководством автора,

по гранту Science Foundation Ireland). Несколько менее значительные результаты работы были получены автором как без соавторов [164, 165, 167, 172], так и совместно с и ее учениками [156, 179, 181], а также с М. Стайнзом [182, 183]. Что касается совместных публикаций, то личный вклад автора диссертации заключался как в постановке рассмотренных задач, так и в получении ключевых результатов указанных работ.

Степень достоверности результатов исследований, изложенных в диссертации. Апробация. Достоверность результатов диссертации подтверждается следующим:

1. Все основные результаты диссертации сформулированы в виде теорем и доказаны с использованием строгих математических обоснований.

2. Основные результаты диссертации изложены в **17 публикациях** в журналах, включенных в системы цитирования **Web of Science** и **Scopus** [156, 157, 161, 163–165, 167–169, 172, 174, 175, 177, 179, 181–183]. Таким образом, основные положения и их доказательства успешно прошли анонимное рецензирование, прежде чем они были опубликованы в международных рецензируемых журналах.

3. Результаты диссертации были представлены автором и получили одобрение специалистов на ряде **международных конференций**, включая следующие:

- Second International Conference “Multiscale methods and Large-scale Scientific Computing”, Marchuk Institute of Numerical Mathematics of Russian Academy of Sciences, Moscow, Russia, 2018 (приглашенный доклад)
- International Conference BAIL 2018 “Boundary and Interior Layers”, Glasgow, Scotland, UK, 2018
- “Adaptive Numerical Methods for Partial Differential Equations with Applications”, Banff International Research Station, Canada, 2018 (приглашенный доклад)

- 27th Biennial Conference on Numerical Analysis, Strathclyde University, UK, 2017
- International Conference “Contemporary Problems of Mathematical Physics and Computational Mathematics” dedicated to the 110th anniversary of academician A.N. Tikhonov, Moscow State University, Russia, 2016 (пленарный доклад)
- International Conference BAIL 2016 “Boundary and Interior Layers”, Beijing, China, 2016
- 2016 AARMS-CRM Workshop on “Numerical Analysis of Singularly Perturbed Differential Equations”, Halifax, Canada, 2016 (пленарный доклад)
- 6th Conference on “Numerical Analysis and Applications” NAA’16, Lozenetz, Bulgaria, 2016 (пленарный доклад)
- MAFELAP 2016 “The Mathematics of Finite Elements and Applications”, Brunel University, UK, 2016
- 13th Workshop on “Numerical Methods for Problems with Layer Phenomena” dedicated to the 90th birthday of A.B. Vasil’eva, Moscow State University, Russia, 2016
- International Workshop “Adaptive Algorithms for Computational PDEs”, University of Birmingham, UK, 2016 (приглашенный доклад)
- 26th Biennial Conference on Numerical Analysis, Strathclyde University, UK, 2015
- International Conference BAIL2014 “Boundary and Interior Layers”, Charles University in Prague, Czech Republic, 2014 (пленарный доклад)
- Sixth Conference “Finite Difference Methods: Theory and Applications”, Lozenetz, Bulgaria, 2014 (пленарный доклад)

- International Multidisciplinary Workshop MURPHYS-HSFS-2014, Weierstrass Institute (WIAS), Berlin, Germany, 2014
- British Computational PDEs Colloquium: New Trends, International Centre for Mathematical Sciences, Edinburgh, UK, 2014 (приглашенный доклад)
- International Workshop VMS2013 “Variational MultiScale and Stabilized Finite Elements”, Barcelona, Spain, 2013
- 25th Biennial Conference on Numerical Analysis, Strathclyde University, UK, 2013
- MAFELAP 2013 “The Mathematics of Finite Elements and Applications”, Brunel University, UK, 2013
- Fifth Conference on “Numerical Analysis and Applications” NAA’12, Lozenetz, Bulgaria, 2012 (пленарный доклад)
- International Conference FoCM’11 “Foundations of Computational Mathematics”, Budapest, Hungary, 2011 (приглашенный доклад)
- 24th Biennial Conference on Numerical Analysis, Strathclyde University, UK, 2011
- International Conference BAIL 2010 “Boundary and Interior Layers: Computational and Asymptotic Methods“, Zaragoza, Spain, 2010
- Fifth Conference “Finite Difference Methods: Theory and Applications”, Lozenetz, Bulgaria, 2010 (пленарный доклад)
- 23rd Biennial Conference on Numerical Analysis, Strathclyde University, UK, 2009
- International Conference “Contemporary Problems of Computational Mathematics and Mathematical Physics” in memory of the academician A.A. Samarskii, Moscow, Russia, 2009

- International Conference BAIL 2008 “Boundary and Interior Layers”, Limerick, Ireland, 2008
- Fourth Conference on “Numerical Analysis and Applications” NAA’08, Lozenetz, Bulgaria, 2008 (пленарный доклад)
- International Conference СМАМ-3 “Computational Methods in Applied Mathematics”, Minsk, Belarus, 2007.
- Second International Workshop on “Analysis and Numerical Approximation of Singular Problems” IWANASP’06, Karlovassi, Greece, 2006 (приглашенный доклад)
- Fourth Conference on “Finite Difference Methods: Theory and Applications”, Lozenetz, Bulgaria, 2006 (приглашенный доклад)
- Conference “Tikhonov and Contemporary Mathematics”, Moscow, Russia, 2006.
- Workshop on “Adaptive Method, Theory and Application”, Peking University, Beijing, China, 2005 (приглашенный доклад)
- 21st Biennial Conference on Numerical Analysis, Dundee, UK, 2005
- Conference MMA2005/СМАМ2, Trakai, Lithuania, 2005

Краткое содержание диссертации. Диссертация состоит из введения, трех глав, разделенных на 7 параграфов, заключения и списка цитированной литературы, содержащего 183 названия. Общий объем работы составляет 298 страниц.

Глава 1, включающая §§1–3, посвящена получению явных апостериорных оценок ошибки численного решения для следующего полулинейного син-

гулярно возмущенного эллиптического уравнения реакции-диффузии:

$$Lu := -\varepsilon^2 \Delta u + f(x, u) = 0 \quad \text{в } \Omega, \quad u = 0 \quad \text{на } \partial\Omega. \quad (1)$$

Предполагается, что $0 < \varepsilon \leq 1$, а f непрерывна в $\Omega \times \mathbb{R}$ и удовлетворяет условию $f(\cdot, s) \in L_\infty(\Omega)$ при любом $s \in \mathbb{R}$, а также одностороннему условию Липшица $f(x, u) - f(x, v) \geq C_f[u - v]$ при любых $u \geq v$ с некоторой постоянной $C_f \geq 0$. Отметим, что при выполнении последнего условия с $C_f = 0$ уравнения типа (1) в дальнейшем будем называть монотонными. Область $\Omega \subset \mathbb{R}^n$, $n = 2, 3$, является, возможно нелипшицевым, многоугольником или многогранником. Рассматриваемая задача имеет единственное решение $u \in H_0^1(\Omega) \cap C(\bar{\Omega})$ (см. Лемму 1.1 ниже). Также предполагается, что $C_f + \varepsilon^2 \geq 1$ (поскольку деление на $C_f + \varepsilon^2$ немедленно сводит (1) к данному случаю).

Будут рассмотрены стандартные конечноэлементные аппроксимации задачи (1). Пусть $S_h \subset H_0^1(\Omega)$ является стандартным лагранжевым конечноэлементным пространством непрерывных кусочно-полиномиальных функций степени r , соответствующих локально квазиравномерной триангуляции области \mathcal{T} , а $u_h \in S_h$ пусть удовлетворяет соотношению

$$\varepsilon^2 \langle \nabla u_h, \nabla v_h \rangle + \langle f(\cdot, u_h), v_h \rangle_h = 0 \quad \forall v_h \in S_h. \quad (2)$$

Здесь $\langle \cdot, \cdot \rangle$ обозначает скалярное произведение в пространстве $L_2(\Omega)$, а $\langle \cdot, \cdot \rangle_h$ — некоторое его приближение, получаемое с использованием квадратурных формул (более точные предположения для которого будут описаны ниже).

В §1 для задачи (1) будут установлены явные апостериорные оценки ошибки численного решения на основе невязок в норме максимума модуля на локально квазиравномерных сетках. Постоянные в полученных оценках не зависят от диаметров элементов сетки и малого параметра. При этом получены и использованы неупрощаемые оценки функции Грина линеаризованного оператора, которые даже в случае уравнения Лапласа несколько улучшают стандартные, и поэтому приводят к апостериорным оценкам с

более точной степенью логарифма. Доказано, что присутствующий в полученных апостериорных оценках логарифм является необходимым по крайней мере в случае линейных конечных элементов для уравнения Лапласа. (Ранее необходимость логарифма была доказана лишь для аналогичных априорных оценок.)

Полученные в параграфе апостериорные оценки представлены в Теореме 1.12. Для простоты изложения, предположим пока, что все интегралы в (2) вычисляются точно, т. е. $\langle \cdot, \cdot \rangle_h = \langle \cdot, \cdot \rangle$. Тогда верхняя апостериорная оценка их Теоремы 1.12 принимает вид

$$\|u - u_h\|_{\infty; \Omega} \leq C \max_{T \in \mathcal{T}} \left(\min\{1, \ell_h h_T^2 \varepsilon^{-2}\} \|\varepsilon^2 \Delta u_h - f(\cdot, u_h)\|_{\infty; T} + \min\{\varepsilon, \ell_h h_T\} \|\llbracket \nabla u_h \rrbracket\|_{\infty; \partial T} \right). \quad (3)$$

Здесь $h_T = \text{diam}(T)$, $\llbracket \nabla u_h \rrbracket$ обозначает стандартный скачок нормальной производной u_h при переходе через границу элемента сетки, а $\ell_h = \ln(2 + \varepsilon \underline{h}^{-1})$ с $\underline{h} = \min_{T \in \mathcal{T}} h_T$.

Заметим, что оценка (3) является робастной по ε в том смысле, что зависимость от ε в этой оценке показана явно. Также, аналогично работе [149] (в которой впервые были получены робастные апостериорные оценки для уравнений типа (1), но в энергетической норме), весовые коэффициенты в этой оценке принимают разную форму в зависимости от локального диаметра сетки h_T , принимающего значения $h_T < \varepsilon$ или $h_T \geq \varepsilon$.

Теорема 1.12 также описывает полученные нами робастные по ε нижние апостериорные оценки; последние теоретически обосновывают эффективность соответствующих верхних оценок.

Отметим также улучшение логарифмических коэффициентов в (3) по сравнению со стандартными оценками (см., например, [125, 74]) даже для случая уравнения Лапласа. Степень ℓ_h в (3) равна 1 при $n = 2$ и $n = 3$, в то время как степень аналогичного логарифмического множителя в [125] больше 1. Степень логарифма при $n = 3$ уже была улучшена до 1 в [74];

здесь же мы получаем аналогичное улучшение для $n = 2$. Заметим, что последнее улучшение будет установлено благодаря более точной, по сравнению со стандартными, оценке $|G(x, \cdot)|_{2,1,\Omega \setminus B(x,\rho)} \lesssim \varepsilon^{-2} \ell_\rho$ для функции Грина G эллиптического оператора типа L в (1) (см. Теорему 1.2). Здесь $B(x, \rho)$ — шар радиуса ρ с центром $x \in \Omega$, и, что весьма важно, логарифмический множитель $\ell_\rho := \ln(2 + \varepsilon \rho^{-1})$ имеет степень 1.

Более того, в §1.4 будет показано, что присутствующий в (3) логарифмический коэффициент является необходимым по крайней мере в случае линейных конечных элементов для уравнения Лапласа в некоторых выпуклых полигональных областях. Ранее необходимость логарифма была доказана лишь для аналогичных априорных оценок (см. [96, 61, 1] при $n = 2$ и [172] при $n = 3$).

Основной материал параграфа опубликован в работе [157].

В §2 для задачи (1) в многоугольных областях будут установлены явные апостериорные оценки ошибки численного решения на основе невязок в норме максимума модуля. По сравнению с §1, в данном параграфе, а также в §3, мы ограничим рассмотрение линейными конечными элементами в двумерной области, при этом наше внимание будет перенесено на существенно более сложный случай неструктурированных анизотропных сеток (т. е. элементы рассматриваемых здесь сеток могут быть сколь угодно сильно сплюснены). Постоянные в полученных оценках не зависят от диаметров элементов сетки, их аспектного соотношения (т. е. степени их сплюсненности) и малого параметра.

Даже для линейного уравнения Лапласа (получаемого из (1) при $\varepsilon = 1$ и $f_u = 0$) нам неизвестны какие-либо апостериорные оценки в норме максимума модуля при отсутствии предположения ограниченности аспектного соотношения элементов сетки (например, в работах [82, 123, 71, 125] предполагается локальная квазиравномерность сетки). Еще больший интерес анизотропные сетки представляют в контексте сингулярно возмущенных диффе-

ренциальных уравнений (таких, как (1) при $\varepsilon \ll 1$), поскольку для решений таких уравнений характерны пограничные и внутренние слои.

Как уже отмечалось, в §§2–3 для уравнения (1) будут рассмотрены стандартные линейные конечные элементы. Пусть $S_h \subset H_0^1(\Omega) \cap C(\bar{\Omega})$ является пространством непрерывных кусочно-линейных функций, ассоциируемым с триангуляцией \mathcal{T} , а численное решение $u_h \in S_h$ пусть удовлетворяет соотношению (2), в котором используется $\langle f(\cdot, u_h), v_h \rangle_h := \langle f_h^I, v_h \rangle$, где $f_h(\cdot) := f(\cdot; u_h)$, а f_h^I — стандартный кусочно-линейный интерполянт Лагранжа для f_h . Для простоты изложения, предположим пока, что углы всех анизотропных элементов близки к нетупым. Тогда наша первая оценка принимает вид

$$\begin{aligned} \|u - u_h\|_{\infty; \Omega} \leq C \ell_h \max_{z \in \mathcal{N}} \left(\min\{\varepsilon, H_z\} \|J_z\|_{\infty; \gamma_z} + \min\{\varepsilon^2, H_z^2\} \|\varepsilon^{-2} f_h^I\|_{\infty; \omega_z} \right) \\ + C \|f_h - f_h^I\|_{\infty; \Omega}, \end{aligned} \quad (4)$$

где постоянная C не зависит от диаметров элементов сетки и их аспектных соотношений (т. е. степени их сплюснутости), а также от малого параметра ε . (Данная оценка получается объединением (2.44), (2.47), (2.21) с Леммой 2.17.) Здесь \mathcal{N} является множеством вершин элементов \mathcal{T} , J_z обозначает стандартный скачок нормальной производной u_h через границу элемента, ω_z — множество элементов, окружающих вершину $z \in \mathcal{N}$, а γ_z — множество ребер внутри ω_z , $H_z = \text{diam}(\omega_z)$, $\ell_h = \ln(2 + \varepsilon \underline{h}^{-1})$, а \underline{h} — минимальный диаметр окружностей, вписанных в треугольники в \mathcal{T} .

Заметим, что при $\varepsilon = 1$ соотношение (4) дает стандартную апостериорную оценку, аналогичную результатам [82, 123, 125], с той существенной разницей, что здесь она будет установлена для случая анизотропных сеток. Отметим также, что оценка (4) практически идентична с апостериорной оценкой (3), полученной для сингулярно возмущенного случая $\varepsilon \in (0, 1]$, однако сейчас мы отказываемся от предположения локальной квазирваномерности триангуляции.

При внимательном рассмотрении стандартных доказательств для случая локально квазиравномерных сеток обнаруживается, что одним из серьезных препятствий при их обобщении на случай анизотропных сеток является применение теоремы о следах при оценке невязок через границы элементов. При этом в получающихся апостериорных оценках появляются коэффициенты, зависящие от аспектных соотношений элементов сетки, поэтому такие оценки чрезвычайно грубы и не могут быть использованы для практической оценки ошибки численного решения. Замечание 2.6 проливает некоторый свет на наш подход к преодолению этого теоретического препятствия.

Следует заметить, что, в отличие от других слагаемых в (4), поэлементные невязки $\|\varepsilon^{-2} f_h^I\|_{\infty; \omega_z}$ не отражают анизотропность соответствующих элементов сетки. С тем, чтобы получить более точную (и, в некотором смысле, анизотропную) оценку для поэлементных невязок, будут введены в рассмотрение цепочки коротких ребер, соединяющих анизотропные узлы (см. §2.6.4, а также Рис. 2.4, справа). При некоторых дополнительных ограничениях на такие цепочки будет получена более точная оценка

$$\begin{aligned} \|u - u_h\|_{\infty; \Omega} \leq & C \ell_h \left[\max_{z \in \mathcal{N}} \left(\min\{\varepsilon, H_z\} \|J_z\|_{\infty; \gamma_z} \right) \right. \\ & \left. + \max_{z \in \mathcal{N} \setminus \mathcal{N}_{\text{paths}}} \left(\min\{1, \varepsilon^{-2} H_z^2\} \|f_h^I\|_{\infty; \omega_z} \right) \right] \\ & + \max_{z \in \mathcal{N}_{\text{paths}}} \left(\min\{\varepsilon, H_z\} \min\{\varepsilon, h_z\} \|\varepsilon^{-2} f_h^I\|_{\infty; \omega_z} + \min\{1, \varepsilon^{-2} H_z^2\} \text{osc}(f_h^I; \omega_z) \right) \\ & + C \|f_h - f_h^I\|_{\infty; \Omega}, \end{aligned} \quad (5)$$

(которая получается объединением (2.44), (2.58), (2.21) с Леммой 2.17). Здесь $\mathcal{N}_{\text{paths}}$ обозначает множество узлов, которые задействованы хотя бы в одной цепочке, а $h_z \simeq H_z^{-1} |\omega_z|$. Поскольку для анизотропных узлов справедливо $h_z \ll H_z$, то очевидно, что оценка (5) является более точной по сравнению с (4). (Последнее также подтверждается результатами численных экспериментов в §2.5.4).

Основной материал параграфа опубликован в работе [168].

В §3, как и в §2, для задачи (1) в многоугольных областях будут установлены явные апостериорные оценки ошибки численного решения на основе невязок для случая неструктурированных анизотропных сеток, но, в отличие от §2, здесь оценки решения будут получены в так называемой энергетической норме $\|\cdot\|_{\varepsilon;\Omega}$, которая является нормализованным вариантом нормы в $W_2^1(\Omega)$ и определяется соотношением

$$\|v\|_{\varepsilon;\Omega} := \left\{ \varepsilon^2 \|\nabla v\|_{2;\Omega}^2 + \|v\|_{2;\Omega}^2 \right\}^{1/2}.$$

Постоянные в полученных оценках не зависят от диаметров элементов сетки, их аспектного соотношения и малого параметра. При этом для интерполяции функции Грина на анизотропных сетках предложен новый квазиинтерполяционный оператор. По сравнению с ранее известными апостериорными оценками данного типа, в полученных нами оценках отсутствуют нежелательные функции согласования (matching functions). Полученные результаты являются новыми даже для уравнения Лапласа.

В рамках предварительного описания основных результатов параграфа, наша первая оценка принимает форму

$$\begin{aligned} \|u_h - u\|_{\varepsilon;\Omega} \leq C \left\{ \sum_{z \in \mathcal{N}} \min\{h_z H_z, \varepsilon H_z\} \|\varepsilon J_z\|_{\infty;\gamma_z}^2 \right. \\ \left. + \sum_{z \in \mathcal{N}} \|\min\{1, H_z \varepsilon^{-1}\} f_h^I\|_{2;\omega_z}^2 + \|f_h - f_h^I\|_{2;\Omega}^2 \right\}^{1/2}, \quad (6) \end{aligned}$$

где постоянная C не зависит от диаметров элементов сетки и их аспектных соотношений (т. е. степени их сплюснутости), а также от малого параметра ε . (Данная оценка получается объединением (3.14), (3.15), (3.25), (3.26) с Замечанием 3.10 и Теоремой 3.18.) Здесь \mathcal{N} является множеством вершин элементов \mathcal{T} , J_z обозначает стандартный скачок нормальной производной u_h через границу элемента, ω_z — множество элементов, окружающих вершину $z \in \mathcal{N}$, γ_z — множество ребер внутри ω_z , $H_z = \text{diam}(\omega_z)$, а $h_z \simeq H_z^{-1}|\omega_z|$.

При некоторых дополнительных предположениях о триангуляции будет также установлена и более точная (хотя и несколько более сложная) апо-

априорная оценка, в которой, по сравнению с (6), грубо говоря, слагаемые $\min\{1, H_z \varepsilon^{-1}\}$ будут заменены на $\min\{1, h_z \varepsilon^{-1}\}$, при этом в оценке появятся некоторые новые слагаемые, связанные с поэлементными невязками (см. Теорему 3.11).

Важно отметить, что в апостериорных оценках на основе невязок всегда присутствует некоторая постоянная, аналогичная C в правых частях оценок (4), (5) и (6), в то время как во всех более ранних апостериорных оценках в энергетической норме на неструктурированных анизотропных сетках из [104, 105, 107] такие постоянные определяются через так называемые функции согласования (matching functions). Последние же являются функциями ошибки численного решения и принимают умеренные значения, если либо сетка локально квазиравномерна, либо если сетка, будучи анизотропной, правильно ориентирована по отношению к решению. В общем же случае функции согласования зависят от аспектного соотношения элементов сетки, и, таким образом, постоянные в оценках [104, 105, 107] (аналогичные C в (6)) могут принимать неограниченно большие значения на сильно анизотропных сетках (если только сетка не является правильно ориентированной по отношению к решению).

Достаточно очевидно, что присутствие подобных функций согласования в апостериорных оценках весьма нежелательно. С другой стороны, в §2 получены апостериорные оценки в норме максимума модуля, в которых функции согласования вообще отсутствуют. Поэтому совершенно естественно, что в большей степени в §3 мы будем опираться на теоретический аппарат, описанный в §2. В то же время будет преодолен ряд дополнительных технических трудностей. Например, модифицированный интерполянт Лагранжа, который используется при интерполяции функции Грина в §2, становится неприменимым при получении оценок в энергетической норме даже в случае уравнения Лапласа. Поэтому для интерполяции функции Грина на анизотропных сетках построен новый квазиинтерполяционный оператор, кото-

рой может представлять независимый интерес. Также нам потребуется более тонкая версия теоремы о следах для анизотропных элементов, и будут ослаблены некоторые предположения о триангуляции, сделанные в §2 для упрощения изложения.

Основной материал параграфа опубликован в работе [169].

Глава 2, включающая §§4–5, посвящена полулинейным параболическим уравнениям второго порядка. Для данных уравнений рассматриваются временные полудискретизации и полностью дискретные методы на основе неявного метода Эйлера, метода Кранка-Николсон и разрывного метода Галеркина $dG(r)$ с квадратурой Радо. Основными результатами главы являются апостериорные оценки в норме максимума модуля для ошибки соответствующих численных решений.

Полученные результаты являются новыми и для классических, и для сингулярно возмущенных уравнений. Простота предложенного теоретического подхода связана с использованием интерполянтов численного решения по времени, степень которых соответствует порядку метода и является более низкой по сравнению с типично используемыми в литературе интерполянтами (например, при анализе неявного метода Эйлера будет использоваться кусочно-постоянная интерполяция). Также важную роль в представленном анализе играют оценки для функции Грина параболического оператора.

Рассматривается полулинейное параболическое уравнение в виде

$$\mathcal{M}u := \partial_t u + \mathcal{L}u + f(x, t, u) = 0 \quad \text{при } (x, t) \in Q := \Omega \times (0, T],$$

где $\mathcal{L} = \mathcal{L}(t)$ является линейным эллиптическим оператором второго порядка, а пространственная область $\Omega \subset \mathbb{R}^n$ имеет липшицеву границу. При этом накладываются следующие начальное и граничное условия:

$$u(x, 0) = \varphi(x) \quad \text{при } x \in \bar{\Omega}, \quad u(x, t) = 0 \quad \text{при } (x, t) \in \partial\Omega \times [0, T].$$

Предполагается, что функция f непрерывна в $\bar{\Omega} \times [0, T] \times \mathbb{R}$, дифферен-

цируема по третьему аргументу, а также при некоторых неотрицательных постоянных γ и $\bar{\gamma}$ удовлетворяет соотношениям

$$0 \leq \gamma^2 \leq \partial_z f(x, t, z) \leq \bar{\gamma}^2 \quad \text{при } (x, t, z) \in \bar{\Omega} \times [0, T] \times \mathbb{R}.$$

В качестве модельного уравнения используется вариант с $\mathcal{L} := -\varepsilon^2 \Delta = -\varepsilon^2 \sum_{i=1}^n \partial_{x_i}^2$:

$$\mathcal{M}u := \partial_t u - \varepsilon^2 \Delta u + f(x, t, u) = 0 \quad (7)$$

в ограниченной полигональной области $\Omega \subset \mathbb{R}^n$ при $n = 1, 2, 3$. Последнее уравнение будет рассмотрено в двух режимах:

$$(i) \quad \varepsilon = 1, \quad \gamma \geq 0; \quad (ii) \quad \varepsilon \ll 1, \quad \gamma > 0.$$

Заметим, что режим (ii) соответствует сингулярно возмущенному уравнению реакции-диффузии, для решений которого характерны резкие пограничные и внутренние слои. Весьма важно, чтобы апостериорные оценки, полученные для данного случая, были робастными в том смысле, что любая зависимость от малого параметра ε должна быть представлена явным образом [112, 132], а также эффективность таких оценок не должна ухудшаться при уменьшении ε .

Основной материал главы опубликован в работе [175].

В §4 представлены апостериорные оценки для полудискретных методов (без дискретизации по пространству). Основные результаты параграфа, апостериорные оценки для неявного метода Эйлера, метода Кранка-Николсон и разрывного метода Галеркина, представлены соответственно в Теоремах 4.8, 4.10 и 4.13. Эти апостериорные оценки объединены в виде более общей оценки (4.10), присутствующие в которой величины описаны в Таблице 4.1.

Например, в случае независящего от t оператора \mathcal{L} неявной метод Эйлера на произвольной неравномерной сетке по времени $\{t_j\}$ принимает вид

$$\frac{U^j - U^{j-1}}{t_j - t_{j-1}} + \mathcal{L}U^j + f(\cdot, t_j, U^j) = 0 \quad \text{в } \Omega, \quad j = 1, \dots, M; \quad U^0 = \varphi.$$

Здесь $U^j \in H_0^1(\Omega) \cap C(\bar{\Omega})$ — численное решение, ассоциируемое с временным уровнем t_j . А оценка из Теоремы 4.8 для этого случая принимает вид

$$\begin{aligned} \|U^m - u(\cdot, t_m)\|_{\infty, \Omega} &\leq (\kappa_1 \ell_m + \kappa_2) \max_{j=1, \dots, m-1} \|U^j - U^{j-1}\|_{\infty, \Omega} + 2 \kappa_0 \|U^m - U^{m-1}\|_{\infty, \Omega} \\ &\quad + \kappa_0 \sum_{j=1}^m \int_{t_{j-1}}^{t_j} e^{-\gamma^2(t_m-s)} \|f(\cdot, s, U^j) - f(\cdot, t_j, U^j)\|_{\infty, \Omega} ds, \end{aligned} \quad (8)$$

где $m = 1, \dots, M$, $\ell_m = \ell_m(\gamma) := \int_{\tau_m}^{t_m} s^{-1} e^{-\frac{1}{2}\gamma^2 s} ds \leq \ln(t_m/\tau_m)$.

Здесь, а также в более общей оценке (4.10) присутствуют постоянные κ_0 , $\kappa_1 > 0$ и $\kappa_2 \geq 0$ из следующих оценок на функцию Грина \mathcal{G} соответствующего линеаризованного параболического оператора (которая определена в (4.5)):

$$\|\mathcal{G}(x, t; \cdot, s)\|_{1, \Omega} \leq \kappa_0 e^{-\gamma^2(t-s)}, \quad \int_0^{t-\tau} \|\partial_s \mathcal{G}(x, t; \cdot, s)\|_{1, \Omega} ds \leq \kappa_1 \ell(\tau, t) + \kappa_2,$$

где $x \in \Omega$, $\tau \in (0, t]$, $t \in (0, T]$, а $\ell(\tau, t) := \int_{\tau}^t s^{-1} e^{-\frac{1}{2}\gamma^2 s} ds \leq \ln(t/\tau)$. Заметим, что данные постоянные для модельной задачи (7) получены в Лемме 4.2.

§5 посвящен обобщению результатов, полученных в §4, на случай полностью дискретных методов. В качестве дискретизаций по пространству используются стандартные методы конечных элементов. Основные результаты параграфа, апостериорные оценки в норме максимума модуля для полностью дискретных неявного метода Эйлера, метода Кранка-Николсон и разрывного метода Галеркина, представлены соответственно в Теоремах 5.9, 5.17 and 5.22. Эти апостериорные оценки объединены в виде более общей оценки (5.8), которая в большой степени имитирует оценку (4.10) для соответствующего полудискретного метода.

Например, в случае независящего от t оператора \mathcal{L} для полностью дискретного неявного метода Эйлера на произвольной неравномерной сетке по

времени $\{t_j\}$ оценка из Теоремы 5.9 принимает вид

$$\begin{aligned}
\|u_h^m - u(\cdot, t_m)\|_{\infty, \Omega} &\leq \kappa_0 e^{-\gamma^2 t_m} \|u_h^0 - \varphi\|_{\infty, \Omega} \\
&+ (\kappa_1 \ell_m + \kappa_2) \max_{j=1, \dots, m-1} \left\{ \|u_h^j - u_h^{j-1}\|_{\infty, \Omega} + \eta^j \right\} \\
&+ 2\kappa_0 \|u_h^m - u_h^{m-1}\|_{\infty, \Omega} + (\kappa_0 + 1) \eta^m \\
&+ \kappa_0 \sum_{j=1}^m \int_{t_{j-1}}^{t_j} e^{-\gamma^2(t_m-s)} \|\theta_h(\cdot, s)\|_{\infty, \Omega} ds. \tag{9}
\end{aligned}$$

где $m = 1, \dots, M$. Здесь $u_h^j \in V_h \cap H_0^1(\Omega)$ — численное решение, ассоциируемое с временным уровнем t_j . Также предполагается, что сетка по пространству и соответствующее конечноэлементное пространство V_h не зависят от j . (Теорема 5.9 дает апостериорную оценку и для более общего случая.) Величина θ_h в (9) в некотором смысле аналогична $f(\cdot, s, U^j) - f(\cdot, t_j, U^j)$ в (8) (см. Замечание 5.6).

Наконец, новые, по сравнению с (8), величины η^j в (9) описывают ошибку, порождаемую дискретизацией по пространству. При обобщении результатов предыдущего параграфа на случай полностью дискретных методов мы опираемся на так называемые эллиптические реконструкции (elliptic reconstructions). Данный технический аппарат был предложен в работах [114, 108, 75] и является аналогом техники на основе проекции Ритца (Ritz projection) [153], часто используемой при получении априорных оценок ошибки для конечноэлементных решений параболических задач. Данный подход позволяет представить η^j в терминах явных эллиптических апостериорных оценивателей, соответствующих используемой пространственной дискретизации. Например, для метода Эйлера величина η^j явным образом выражается через численное решение u_h^j и соответствующую функцию $\frac{u_h^j - u_h^{j-1}}{t_j - t_{j-1}} + f(\cdot, t_j, u_h^j)$ (см. также §5.2.1).

Глава 3, включающая §§6–7, посвящена немонотонным эллиптическим сингулярно возмущенным уравнениям реакции-диффузии. Под немонотон-

ными полулинейными эллиптическими уравнениями будут пониматься уравнения, в которых нелинейность по неизвестной функции не является монотонной, поэтому такие уравнения могут иметь несколько решений.

Рассматривается краевая задача для полулинейного сингулярно возмущенного уравнения реакции-диффузии в виде

$$Fu \equiv -\varepsilon^2 \Delta u + b(x, u) = 0, \quad x = (x_1, x_2) \in \Omega \subset \mathbb{R}^2, \quad (10a)$$

$$u(x) = g(x), \quad x \in \partial\Omega, \quad (10b)$$

где ε — положительный параметр, принимающий сколь угодно малые значения, $\Delta = \partial_{x_1}^2 + \partial_{x_2}^2$ обозначает оператор Лапласа, а Ω является ограниченной двумерной областью.

Полагая $\varepsilon = 0$ в (10a), получаем вырожденную версию (10):

$$b(x, u_0(x)) = 0 \quad \text{для } x \in \Omega. \quad (11)$$

Заметим, что решение u_0 уравнения (11), вообще говоря, не удовлетворяет граничному условию в (10b). Это приводит к тому, что для решений (10) характерны пограничные слои. Решения (10) могут также иметь внутренние переходные слои [19, 38, 147, 183].

Отметим, что функция b не предполагается монотонной по u , поэтому и вырожденное уравнение (11), и исходная задача (10) могут иметь несколько решений. В литературе же по вычислительной математике нередко предполагается (см., например, [133, 10]), что уравнение (10a) монотонно по u и, более того, $b_u(x, u) > \gamma^2 > 0$ при любых $(x, u) \in \Omega \times \mathbb{R}^1$ для некоторой положительной постоянной γ . При данном условии вырожденная задача имеет единственное решение u_0 , которое является достаточно гладким в $\bar{\Omega}$ (что, в силу компактности $\bar{\Omega}$, следует из применения теоремы о неявной функции). Тем не менее, данное глобальное условие является довольно серьезным ограничением. Например, математические модели возникающие при моделировании биологических и химических процессов нередко используют уравнения типа (10), в которых $b(x, u)$ немонотонна по u [120, §14.7], [94, §2.3].

В §6 будут рассмотрены погранслойные решения задачи (10) в случае гладкой области Ω . Целью параграфа является доказательство существования численных решений, а также исследование их точности на сгущающихся в пограничных слоях сетках типа Бахвалова и Шишкина. Установлен второй порядок сходимости (с логарифмическим множителем в случае сетки Шишкина) в сеточной норме максимума модуля равномерно по малому параметру ε при условии $\varepsilon \leq Ch$. Здесь $h > 0$ является максимальным диаметром элементов сетки, в то время как число степеней свободы не превосходит Ch^{-2} . Полученные теоретические оценки сходимости будут проиллюстрированы результатами численных экспериментов.

Отметим, что, в силу немонотонности исходного уравнения в сочетании с наличием малого параметра при старших производных, стандартный технический аппарат, используемый при получении априорных оценок ошибки как для конечно-разностных, так и для конечноэлементных численных решений, становится неприменимым. В данном параграфе, при доказательстве существования и точности численных решений используется принципиально иной подход, основанный на построении дискретных верхних и нижних решений, при этом последние являются нелинейными модификациями асимптотического разложения для исходной задачи. Данный подход ранее использовался в работах [70, 143, 182] для обыкновенных дифференциальных уравнений, и впервые был обобщен автором диссертации на случаи эллиптических и параболических уравнений в [163, 181].

Задача (10) будет рассмотрена в гладкой области Ω при следующих предположениях [88, 38]:

A1 Существует устойчивое вырожденное решение, т. е. вырожденная задача (11) имеет достаточно гладкое решение u_0 , удовлетворяющее при некоторой постоянной $\gamma > 0$ условию

$$b_u(x, u_0) > \gamma^2 > 0 \quad \text{для любого } x \in \Omega.$$

A2 Граничное условие $g(x)$ удовлетворяет

$$\int_{u_0(x)}^v b(x, s) ds > 0 \quad \text{для любого } v \in (u_0(x), g(x)]', \quad x \in \partial\Omega.$$

Используемое здесь обозначение $(a, b]'$ совпадает с $(a, b]$ при $a < b$ и $[b, a)$ при $a > b$, в то время как $(a, b]' = \emptyset$ при $a = b$.

Условия A1 и A2 естественным образом возникают при построении асимптотического разложения для задачи (10) и гарантируют, что существует погранслоное решение u , для которого справедливо соотношение $u \approx u_0$ во внутренней части области, в то время как вдоль границы области возникает пограничный слой ширины $O(\varepsilon |\ln \varepsilon|)$ [88, 38]; см. также Теорему 6.2.

Рассматриваемый нами численный метод использует дискретизацию области с использованием сгущающихся в окрестности погранслоя криволинейных прямоугольных сеток типа Бахвалова и Шишкина, число узлов которых не превосходит Ch^{-2} (более подробное описание дано в §6.3.1). Здесь $h > 0$ — максимальный диаметр элементов сетки. Что касается дискретизации уравнения (10а), то оно аппроксимируется с использованием метода конечных разностей на криволинейной прямоугольной части сетки и линейного метода конечных элементов с сосредоточенными массами на квазиравномерной триангуляции Делоне (Delaunay) во внутренней части области. С точки зрения используемого здесь технического аппарата, ключевым свойством рассматриваемого численного метода является то, что дискретизация оператора $-\Delta$ соответствует M -матрице. Последнее позволяет использовать дискретные нижние и верхние решения при доказательстве существования численного решения и исследовании его точности.

Основной результат параграфа, Теорема 6.28, устанавливает второй порядок сходимости (с логарифмическим множителем в случае сетки Шишкина) в сеточной норме максимума модуля равномерно по малому параметру ε .

Основной материал параграфа опубликован в работе [163], а обобщения данного подхода кратко описаны в §6.5 и опубликованы в работах [179, 180,

181, 183].

В §7 задача (10) рассматривается в выпуклой многоугольной области. Наличие угловых точек на границе $\partial\Omega$ приводит к дополнительным сложностям в асимптотическом анализе. В дополнение к достаточно стандартным погранслойным функциям, в асимптотическом разложении появляются угловые пограничные функции. Эти угловые функции определяются как решения некоторых нелинейных краевых задач в бесконечном выпуклом секторе, что серьезно усложняет анализ, поскольку даже существование решений таких задач далеко неочевидно.

Отметим, что наш интерес к асимптотическому анализу задачи (10) обусловлен тем, что построенное асимптотическое разложение и соответствующие верхние и нижние решения могут быть использованы при построении и теоретическом исследовании ε -равномерных численных методов для рассматриваемой задачи (аналогично §6). Некоторые аспекты построения таких численных методов для (10) кратко обсуждаются в §7.5.

Будут также сделаны те же предположения A1 и A2, что и в §6, а также следующее дополнительное предположение.

A3 (*Угловые условия*) Для каждой угловой точки P_j многоугольной области Ω , при $g(P_j) \neq u_0(P_j)$ предполагается, что

$$\frac{b(P_j, g(P_j))}{g(P_j) - u_0(P_j)} > 0.$$

Обсудим предположение A3. Ключевым элементом анализа, изложению которого посвящен данный параграф, будет исследование некоторых решений полулинейного эллиптического уравнения

$$-\Delta z + f(z) = 0 \tag{12}$$

в бесконечном секторе. Уравнение типа (12) представляет интерес, поскольку угловые погранфункции, ассоциируемые с вершиной P_j могут быть описаны с использованием некоторого решения уравнения (12) при $f(z) = b(P_j, z)$ с

условием $z = g(P_j)$ на границе сектора (см. описание задачи (7.12)). Как будет показано, предположение А3 не только является достаточным для существования решения z . Результат из [144] показывает, что А3 также является в некотором смысле необходимым для существования решения z определенного типа (см. Замечание 7.7). Более того, с использованием А3 будет установлена устойчивость решений (12), понимаемая в том смысле, что главное собственное значение оператора, возникающего при линеаризации (12) около решения z и рассматриваемого в соответствующем секторе радиуса R , является положительным и отделено от нуля положительной постоянной равномерно по $R \rightarrow \infty$ (см. §7.3.4). Раздел, посвященный задаче (12) и соответствующему линеаризованному оператору, является ключевым в §7, и может также представлять независимый интерес.

Основным результатом параграфа является построение асимптотического разложения первого порядка u_{as} для задачи (10) и доказательство существования решения этой задачи $u(x)$, для которого справедлива оценка $|u - u_{\text{as}}| \leq C\varepsilon^2$. Также будут установлены поточечные оценки производных для компонент построенного асимптотического разложения. Последние могут быть полезными при построении и теоретическом исследовании сходимости численных методов для рассматриваемой задачи. Аналогичные результаты для гладких областей были получены в работах [88, 38].

Результаты данного параграфа являются принципиально новыми для многоугольных областей. Отметим, что в случае прямоугольной области результаты, аналогичные нашим, были получены в работе [15] для линейного варианта задачи (10). Нелинейная же задача (10) в прямоугольной области рассматривалась в работах [23, 24, 25, 26, 27], но при весьма более сильных, по сравнению с А3, ограничениях в угловых точках области. (Например, в [25] предполагается, что $b(P_j, \cdot)$ монотонна между $g(P_j)$ и $u_0(P_j)$, а в [23, 26] $b(P_j, \cdot)$ предполагается квадратичной.) Следуя работам [97, 38], для доказательства существования решения будет использована теория верхних и ниж-

них решений. При этом последние будут построены с помощью нелинейной модификации асимптотического разложения.

Основной материал параграфа опубликован в работе [161], а некоторые технические вычисления и доказательства помещены в [160].

В **Заключении** будут перечислены основные результаты работы.

Начиная со следующего параграфа, в диссертации будет использована двухпозиционная нумерация утверждений, формул, рисунков и таблиц, при этом первое число указывает на номер параграфа.

Глава 1

Апостериорные оценки для стационарных сингулярно возмущенных уравнений реакции-диффузии на локально квазиравномерных и анизотропных сетках

§1 Апостериорные оценки в норме максимума модуля на локально квазиравномерных сет- ках

Настоящий параграф посвящен получению явных апостериорных оценок ошибки численного решения на основе невязок в норме максимума модуля для сингулярно возмущенного полулинейного уравнения реакции-диффузии в полигональных областях. Рассматривается стандартный метод конечных элементов. Постоянные в полученных оценках не зависят от диаметров элементов сетки и малого параметра. Одним из ключевых элементов представленного анализа являются неулучшаемые оценки функции Грина линеаризованного оператора. Даже в случае уравнения Лапласа используемые нами оценки функции Грина несколько улучшают стандартные, и поэтому приводят к апостериорным оценкам с более точной степенью логарифма. Будет также показано, что присутствующий в полученных оценках логарифм является необходимым по крайней мере в случае линейных конечных элементов

для уравнения Лапласа. Теоретические выводы иллюстрируются результатами численных экспериментов. Основной материал параграфа опубликован в работе [157].

1.1 Введение

Целью настоящего параграфа является получение явных апостериорных оценок ошибки численного решения в норме максимума модуля для следующего сингулярно возмущенного уравнения реакции-диффузии:

$$Lu := -\varepsilon^2 \Delta u + f(x, u) = 0 \quad \text{в } \Omega, \quad u = 0 \quad \text{на } \partial\Omega. \quad (1.1)$$

Предполагается, что $0 < \varepsilon \leq 1$, а f непрерывна в $\Omega \times \mathbb{R}$ и удовлетворяет условию $f(\cdot, s) \in L_\infty(\Omega)$ при любом $s \in \mathbb{R}$, а также одностороннему условию Липшица $f(x, u) - f(x, v) \geq C_f[u - v]$ при любых $u \geq v$ с некоторой постоянной $C_f \geq 0$. Отметим, что при выполнении последнего условия с $C_f = 0$ уравнения типа (1.1) в дальнейшем будем называть монотонными. Область $\Omega \subset \mathbb{R}^n$, $n = 2, 3$, является, возможно нелипшицевым, многоугольником или многогранником. Рассматриваемая задача имеет единственное решение $u \in H_0^1(\Omega) \cap C(\bar{\Omega})$ (см. Лемму 1.1 ниже). Также предполагается, что $C_f + \varepsilon^2 \geq 1$ (поскольку деление на $C_f + \varepsilon^2$ немедленно сводит (1.1) к данному случаю).

Будут рассмотрены стандартные конечноэлементные аппроксимации задачи (1.1). Пусть $S_h \subset H_0^1(\Omega)$ является стандартным лагранжевым конечноэлементным пространством непрерывных кусочно-полиномиальных функций степени r , соответствующих локально квазиравномерной триангуляции области \mathcal{T} , а $u_h \in S_h$ пусть удовлетворяет соотношению

$$\varepsilon^2 (\nabla u_h, \nabla v_h) + (f(\cdot, u_h), v_h)_h = 0 \quad \forall v_h \in S_h. \quad (1.2)$$

Здесь (\cdot, \cdot) обозначает скалярное произведение в пространстве $L_2(\Omega)$, а $(\cdot, \cdot)_h$ — некоторое его приближение, получаемое с использованием квадратурных формул (более точные предположения для которого будут описаны ниже).

Уравнения типа (1.1) и его параболический аналог $\partial_t u + Lu = 0$ возникают при моделировании тонких пластин, а также в биологических, химических и инженерных приложениях. Отметим, что полезность результатов данной главы не исчерпывается стационарным случаем, поскольку они могут быть использованы как ингредиенты в апостериорных оценках для соответствующих параболических уравнений [175] (см. также Гл. 2).

Явные апостериорные оценки ошибки численного решения на основе невязок в норме максимума модуля для метода конечных элементов были рассмотрены в ряде публикаций. Из них наиболее ранними представляются работы [82, 123], в которых получены апостериорные оценки в норме L_∞ для линейных эллиптических уравнений в двумерных областях. Подход, использованный в [123], был обобщен для трехмерного случая в [71], а в [126, 127, 125] рассматриваются эллиптические задачи с препятствием и монотонные полулинейные задачи. Отметим также работу [74], содержащую апостериорные оценки в норме максимума модуля для разрывного метода Галеркина в применении к уравнению Лапласа, а также улучшенные оценки для непрерывного метода Галеркина. Подход, используемый в настоящем параграфе, наиболее близок к работам [74] и [125]. В частности, использование теоретического аппарата из [74] позволяет рассматривать произвольные многоугольные и многогранные области, в то время как результаты [125] применимы лишь для липшицевых областей. При выводе же апостериорных оценок, в которых учитываются ошибки, возникающие в результате применения квадратурных формул, мы будем опираться на теоретический аппарат [125]. Отметим также, что сходимость адаптивных алгоритмов на основе явных апостериорных оценок была доказана в работах [138, 103].

Что касается апостериорных оценок для сингулярно возмущенных уравнений реакции-диффузии, этому вопросу также посвящен ряд публикаций, при этом ошибка решения, как правило, оценивается в энергетической норме. Работа [149] представляется первой, в которой получены явные апосте-

риорные оценки конечноэлементных решений на основе невязок для стационарных уравнений реакции-диффузии, являющиеся робастными по малому параметру. В [100] представлены похожие результаты, которые затем обобщены для задачи Бринкмана в [101]. В случае анизотропных сеток явные апостериорные оценки для конечно-элементных решений сингулярно возмущенных уравнений реакции-диффузии получены в [105, 106]. Двумя основными чертами всех этих работ является то, что весовые коэффициенты в полученных апостериорных оценках принимают разную форму в зависимости от локального диаметра сетки h_T , принимающего значения $h_T < \varepsilon$ или $h_T \geq \varepsilon$, а также то, что зависимость от ε во всех этих оценках показана явно. Были получены и другие типы апостериорных оценок, являющиеся робастными по ε ; см, например, [51, 53], где в апостериорных оценках используются решения вспомогательных локальных задач.

Отметим, что в контексте сингулярно возмущенных уравнений реакции-диффузии, энергетическая норма представляется недостаточно сильной, поскольку присутствие в ней малого параметра ε приводит к тому, что в практических ситуациях, эта норма не сильнее нормы $L_2(\Omega)$; этот вопрос подробно обсуждается, например, в [110]. Норма максимума модуля, напротив, позволяет идентифицировать большую ошибку решения в пограничных слоях, поэтому представляется более подходящей для оценки ошибки в пограничных решениях. В норме максимума модуля апостериорные оценки для уравнений типа (1.1) получены в [165, 156]; представленные в этих работах результаты применимы только к прямоугольным сеткам (но, что весьма важно, не зависят от аспектного соотношения элементов сетки, т. е. применимы и в случае анизотропных сеток). Что же касается априорных оценок ошибок численных решений в норме максимума модуля для подобных уравнений, то этому вопросу посвящен целый ряд работ. В [133, 109] такие оценки даны для конечноэлементных решений на глобально квазиравномерных сетках, в то время как в случае локально анизотропных сеток (как для

конечно-разностных, так и для конечноэлементных решений) отметим работы [10, 49, 66, 155, 165], а также приведенные в них ссылки.

Основным результатом настоящего параграфа является получение апостериорных оценок в норме максимума модуля, являющихся робастными по малому параметру ε , подобно описанным выше апостериорным оценкам в энергетической норме. Отметим также, что даже в случае уравнения Лапласа используемый нами теоретический аппарат несколько улучшает стандартный, поскольку приводит к оценкам с более точной степенью логарифма (как правило, присутствующего в оценках в норме максимума модуля).

Далее мы приводим краткое описание результатов параграфа с тем, чтобы более подробно проиллюстрировать их новизну. Для простоты изложения, предположим пока, что все интегралы вычисляются точно, т. е. $(\cdot, \cdot)_h = (\cdot, \cdot)$. (В более детальной версии будут также присутствовать индикаторы квадратурных ошибок в духе работы [125]). Ниже будет доказано, что

$$\begin{aligned} \|u - u_h\|_{\infty; \Omega} \leq C \max_{T \in \mathcal{T}} \left(\min\{1, \ell_h h_T^2 \varepsilon^{-2}\} \|\varepsilon^2 \Delta u_h - f(\cdot, u_h)\|_{\infty; T} \right. \\ \left. + \min\{\varepsilon, \ell_h h_T\} \|\llbracket \nabla u_h \rrbracket\|_{\infty; \partial T} \right). \end{aligned} \quad (1.3)$$

Здесь $h_T = \text{diam}(T)$, $\llbracket \nabla u_h \rrbracket$ обозначает стандартный скачок нормальной производной u_h при переходе через границу элемента сетки, а $\ell_h = \ln(2 + \varepsilon \underline{h}^{-1})$ с $\underline{h} = \min_{T \in \mathcal{T}} h_T$. Мы также получим робастные по ε нижние апостериорные оценки, показывающие эффективность соответствующих верхних оценок. Для сравнения, заметим, что апостериорный анализ [125], строго говоря, применим к (1.1), хотя робастность по ε не являлась целью этой работы. Оценки в [125] получены аналогично нашим, но C_f предполагается равным 0, что приводит к следующему результату:

$$\|u - u_h\|_{\infty; \Omega} \leq C \tilde{\ell}_h^{\alpha_n} \max_{T \in \mathcal{T}} \left(h_T^2 \varepsilon^{-2} \|\varepsilon^2 \Delta u_h - f(\cdot, u_h)\|_{\infty; T} + h_T \|\llbracket \nabla u_h \rrbracket\|_{\infty; \partial T} \right). \quad (1.4)$$

Здесь $\tilde{\ell}_h = \ln(1/h)$, а $\alpha_2 = 2$ и $\alpha_3 = 4/3$. В обеих вышеприведенных оценках постоянная C не зависит от ε . Основным улучшением в (1.3) по сравнению с (1.4) является значение весовых коэффициентов в случае $\varepsilon^2 \ll C_f$, т.е. когда задача является сингулярно возмущенной. Действительно, в этом случае оценка (1.3) является существенно более точной в части сетки с $h_T \gg \varepsilon$. При $\varepsilon = 1$, две вышеприведенные оценки эквивалентны, не считая различий в значениях и степенях присутствующих логарифмических коэффициентов. Численные результаты, приведенные ниже в §1.5, показывают, что оценка (1.4) является равномерно по ε робастной в том смысле, что ее индекс эффективности (оценка ошибки деленная на реальную ошибку) бесконечно растет при фиксированном $\max h_T$ и $\varepsilon \rightarrow 0$. Приведенные вычислительные эксперименты также показывают, что использование индикаторов ошибки, получаемых из (1.4), при маркировке элементов сетки приводит к худшей работе соответствующих адаптивных алгоритмов.

Наша апостериорная оценка (1.3) почти эквивалентна (1.4) при $C_f \lesssim \varepsilon$, т.е. если задача не является сингулярно возмущенной; (в этом случае нормализующее условие $C_f + \varepsilon^2 \geq 1$ приводит к $\varepsilon \simeq 1$). Отметим, что в этом случае в одностороннем условии Липшица можно положить $C_f = 0$, т.е. оно сводится к требованию монотонности уравнения, которое и предполагается в [125]. Таким образом наша оценка обеспечивает единое рассмотрение как сингулярно возмущенного, так и эллиптического режимов, а также перехода между этими двумя режимами.

Отметим также улучшение логарифмических коэффициентов в (1.3) по сравнению с (1.4). Во-первых, ℓ_h в (1.3) меньше, чем $\tilde{\ell}_h$ в (1.4) при $\varepsilon \ll 1$. (Заметим, что в априорных оценках [133] также присутствует ℓ_h .) Далее, степень ℓ_h в (1.3) равна 1 при $n = 2$ и $n = 3$, в то время как степень $\tilde{\ell}_h$ в (1.4) больше 1. Степень логарифма при $n = 3$ уже была улучшена до 1 в [74]; здесь же мы получаем аналогичное улучшение для $n = 2$. Более того, ниже будет показано, что присутствующий в (1.3) логарифмический коэффициент явля-

ется необходимым по крайней мере в случае линейных конечных элементов для уравнения Лапласа в некоторых выпуклых полигональных областях.

Как и в процитированных работах по апостериорному анализу эллиптических задач в норме максимума, поточечная ошибка решения будет выражена через функцию Грина, поэтому оценки функции Грина будут играть ключевую роль в представленных доказательствах. Такие оценки хорошо известны для оператора Лапласа. В [125], где рассматривается полулинейное монотонное уравнение, при получении оценки (1.4) авторы воспользовались барьерными функциями с тем, чтобы затем применить оценки для регуляризованной функции Грина оператора Лапласа. Мы воспользуемся несколько упрощенной версией этого подхода, при этом, как и в [74], будет использоваться не регуляризованная, а стандартная функция Грина. Отметим, что в качестве упрощенного оператора будет использоваться $-\varepsilon^2 \Delta + C_f$, поскольку последний сохраняет ключевые свойства сингулярно возмущенного уравнения из (1.1).

Настоящий параграф организован следующим образом. В §1.2 будут установлены существование и единственность решения исходной задачи, а также будут получены некоторые оценки функции Грина, которые будут играть ключевую роль в последующем анализе. Апостериорные оценки ошибки решения, являющиеся робастными по ε , будут получены в §1.3. Далее, в §1.4 будет показано, что логарифмический коэффициент является необходимым в апостериорных оценках в норме максимума модуля. Наконец, результаты некоторых численных экспериментов будут представлены в §1.5.

1.2 Существование и единственность решения исходной задачи. Оценки функции Грина

Поскольку нам неизвестны доказательства существования и единственности исходной задачи в точности при сделанных нами предположениях, сначала мы установим данный результат, а затем выведем ряд ключевых для

последующих доказательств оценок функции Грина.

Л е м м а 1.1. Пусть $f \in C(\Omega \times \mathbb{R})$, $f(\cdot, s) \in L_\infty(\Omega)$ при любом $s \in \mathbb{R}$, и $f(x, u) - f(x, v) \geq 0$ при любых $x \in \Omega$ и $u \geq v$, а также Ω является полигональной областью в \mathbb{R}^n , при $n = 2$ или $n = 3$. Тогда (1.1) имеет единственное решение $u \in H_0^1(\Omega)$, а также это решение $u \in W_l^2(\Omega) \subseteq W_q^1 \subset C(\bar{\Omega})$ при некоторых $l > \frac{n}{2}$ и $q > n$.

Д о к а з а т е л ь с т в о. См. доказательство Леммы 1 в [157]. Отметим лишь, что последнее основывается на оценках из [69, 72, 116]. \square

Как и в ряде других работ по апостериорным оценкам в норме максимума модуля, поточечная ошибка решения будет представлена через функцию Грина. Отметим, что ниже, в §1.3.1, будет показано, что для этих целей подойдет функция Грина упрощенного оператора $\bar{L} := -\varepsilon^2 \Delta + C_f$. Полученные нами оценки остаются справедливыми и для функции Грина линейризованного оператора при дополнительном предположении $C_f \leq f_u \lesssim C_f$.

Существует функция Грина $G(x, \xi) : \Omega \times \Omega \rightarrow \mathbb{R}$ такая, что при любом $v \in H_0^1(\Omega) \cap W_1^q(\Omega)$ с $q > n$, справедливо представление

$$v(x) = \varepsilon^2(\nabla v, \nabla G(x, \cdot)) + C_f(v, G(x, \cdot)). \quad (1.5)$$

При любом $x \in \Omega$, эта функция G удовлетворяет соотношениям

$$\begin{aligned} \bar{L}G &= -\varepsilon^2 \Delta_\xi G + C_f G = \delta(x - \xi), & \xi \in \Omega, \\ G(x; \xi) &= 0, & \xi \in \partial\Omega. \end{aligned} \quad (1.6)$$

Здесь $\delta(\cdot)$ является n -мерной функцией Дирака.

Нам потребуются некоторые обозначения. В дальнейшем будем писать $a \simeq b$, если $a \lesssim b$ и $a \gtrsim b$, в то время как $a \lesssim b$ будет обозначать $a \leq Cb$ с некоторой постоянной C , зависящей от Ω , r , и регулярности элементов локально квазиравномерной сетки \mathcal{T} . Отметим, что C не зависит от диаметров элементов \mathcal{T} и ε . Также, для $\mathcal{D} \subseteq \bar{\Omega}$, $1 \leq p \leq \infty$ и $k \geq 0$, будут использоваться нормы $\|v\|_{p; \mathcal{D}} = \|v\|_{L_p(\mathcal{D})}$ и $|v|_{k; p; \mathcal{D}} = |v|_{W_p^k(\mathcal{D})}$, где $|\cdot|_{W_p^k(\mathcal{D})}$ обозначает

стандартную полунорму в пространстве Соболева функций, интегрируемых с p -ой степенью их модуля вместе со своими обобщенными производными до порядка k .

Т е о р е м а 1.2. *Для G из (1.6) при любом $x \in \Omega$ справедлива оценка*

$$\|G(x, \cdot)\|_{1;\Omega} + \varepsilon \|G(x, \cdot)\|_{\frac{n}{n-1};\Omega} + \varepsilon |G(x, \cdot)|_{1,1;\Omega} \lesssim 1. \quad (1.7)$$

Также, для шара $B(x, \rho)$ радиуса ρ с центром $x \in \Omega$, пусть $\ell_\rho := \ln(2 + \varepsilon\rho^{-1})$. Тогда

$$\|G(x, \cdot)\|_{1, B(x, \rho) \cap \Omega} \lesssim \varepsilon^{-2} \rho^2 \ell_\rho^{k_2}, \quad k_2 = 1 \text{ and } k_3 = 0, \quad (1.8a)$$

$$\|G(x, \cdot)\|_{\frac{n}{n-2}, \Omega \setminus B(x, \rho)} \lesssim \varepsilon^{-2} \ell_\rho, \quad (1.8b)$$

$$|G(x, \cdot)|_{1, \frac{n}{n-1}; \Omega \setminus B(x, \rho)} \lesssim \varepsilon^{-2} \ell_\rho, \quad (1.8c)$$

$$|G(x, \cdot)|_{1,1, B(x, \rho) \cap \Omega} \lesssim \varepsilon^{-2} \rho, \quad (1.8d)$$

$$|G(x, \cdot)|_{2,1, \Omega \setminus B(x, \rho)} \lesssim \varepsilon^{-2} \ell_\rho. \quad (1.8e)$$

З а м е ч а н и е 1.3. *В работе [74] получены аналогичные оценки для случая $\varepsilon = 1$, $C_f = 0$. При $n = 2$, оценка (1.8e) является более точной по сравнению с [74, (5.21)], поскольку логарифмический коэффициент $\ln^2(1/\underline{h})$ в последней улучшен до $\ln(1/\underline{h})$. Следовательно, аналогичное улучшение справедливо также и для всех апостериорных оценок, полученных в [74].*

З а м е ч а н и е 1.4. *Аналогичные оценки для функции Грина, но в существенно более простых квадратных и кубических областях, получены в [156, 165]. Доказательства в этих работах являются в некотором смысле конструктивными, поэтому позволяют сделать вывод, что все оценки в Теореме 1.2 неулучшаемы по ε , ρ и ℓ_ρ . Отметим также работы [158, 159], в которых получены неулучшаемые оценки для функции Грина существенно более сложного сингулярно возмущенного уравнения конвекции-диффузии.*

Прежде чем мы приступим к доказательству Теоремы 1.2, нам потребуются ряд вспомогательных результатов. Начнем с несколько видоизмененных оценок, полученных в [74] для функции Грина оператора Лапласа.

Л е м м а 1.5. При $\varepsilon = 1$ и $C_f = 0$, для функции Грина G из (1.6) справедливы оценки (1.7), (1.8d) и (1.8e).

Д о к а з а т е л ь с т в о. В случае $\varepsilon = 1$ и $C_f = 0$ оценка для $|G(x, \cdot)|_{1,1;\Omega}$ в (1.7) немедленно следует из (1.8d) при $\varepsilon = 1$, $\rho = \text{diam}(\Omega)$, в то время как остальные результаты в (1.7) могут быть легко получены из поточечных оценок G из [74, (2.6)].

Если $n = 3$, оценки (1.8d) и (1.8e) при $\varepsilon = 1$ немедленно следуют из [74, (5.23) и (5.26)].

Если же $n = 2$, необходимо улучшить степень логарифма в [74, (5.23) и (5.27)], с тем чтобы получить более точные оценки (1.8d) и (1.8e), что может быть сделано следующим образом. Заметим, что первая строка в [74, (5.23)], а также в [74, (5.25)], остается справедливой, если ∇G заменена в каждой из рассмотренных подобластей Ω_j разностью $\nabla(G - \min_{\Omega_j} G)$. Воспользовавшись этим наблюдением, можно несколько модифицировать доказательства оценок [74, (5.23) и (5.27)] с тем, чтобы получить более точные (1.8d) и (1.8e). При этом существенным является применение следующей леммы. \square

Л е м м а 1.6. Пусть $n = 2$, $\varepsilon = 1$, $C_f = 0$, и $\Omega_\rho = [B(x, \rho) \setminus B(x, \frac{1}{2}\rho)] \cap \Omega$ при любых $\rho > 0$ и $x \in \Omega$. Тогда функция Грина G из (1.6) удовлетворяет соотношению

$$\sup_{\Omega_\rho} G(x, \cdot) - \inf_{\Omega_\rho} G(x, \cdot) \leq C,$$

где C не зависит от ρ или x .

Д о к а з а т е л ь с т в о. Зафиксируем $x \in \Omega$, и пусть $r_0 = \text{dist}(x, \partial\Omega)$. Отметим, что достаточно показать, что

$$\max\left\{0, \frac{1}{2\pi} \ln \frac{r_0}{|\xi - x|}\right\} \leq G(x, \xi) \leq \max\left\{0, \frac{1}{2\pi} \ln \frac{r_0}{|\xi - x|}\right\} + C, \quad \xi \in \Omega. \quad (1.9)$$

Здесь нижняя оценка может быть легко получена с помощью фундаментального решения $\Gamma(x, \xi) = \frac{1}{2\pi} \ln |\xi - x|^{-1}$ в \mathbb{R}^2 , в силу принципа максимума.

Чтобы доказать верхнюю оценку, предположим, без потери общности, что ближайшей точкой к x на $\partial\Omega$ является O , а также что $\Omega \subset S$, где область S является либо (i) $S = \mathbb{R}^2 \setminus \{(\xi_1, 0), \xi_1 \geq 0\}$, либо, в случае более сложного многоугольника Ω , (ii) $S = \{|\xi - x| < \text{diam}(\Omega)\} \setminus \{(\xi_1, 0), 0 \leq \xi_1 \leq C_S\}$ с некоторой постоянной $C_S \geq 1$. Поскольку из $\Omega \subset S$ следует $G(x, \xi) \leq G_S(x, \xi)$, где G_S — функция Грина для области S , то для получения верхней оценки в (1.9) достаточно показать, что

$$G_S(x, \xi) \leq \max\left\{0, \frac{1}{2\pi} \ln \frac{5r_0}{|\xi - x|}\right\} + C. \quad (1.10)$$

Чтобы завершить доказательство, оценка (1.10) будет установлена сначала для случая (i), а затем для случая (ii).

(i) Явное представление для функции Грина в $S = \mathbb{R}^2 \setminus \{(\xi_1, 0), \xi_1 \geq 0\}$ получено в [22, с. 143, (16.55)]:

$$G_S(x, \xi) = \frac{1}{4\pi} \ln \left(\frac{t^2 - 2t \cos\left(\frac{1}{2}[\theta + \theta_0]\right) + 1}{t^2 - 2t \cos\left(\frac{1}{2}[\theta - \theta_0]\right) + 1} \right), \quad t = \sqrt{\frac{r}{r_0}},$$

где (r_0, θ_0) и (r, θ) обозначают полярные координаты соответственно для x и ξ . При $r \geq 4r_0$, получаем $t \geq 2$, и, следовательно, $G_S \leq \frac{1}{2\pi} \ln \left| \frac{t+1}{t-1} \right| \leq \frac{1}{2\pi} \ln 3$. Последняя оценка справедлива в подобласти $\{|\xi - x| \geq 5r_0\} \subset \{r \geq 4r_0\}$. Далее, для оставшейся подобласти $\{|\xi - x| \leq 5r_0\}$, воспользовавшись принципом максимума, получаем $G_S \leq \frac{1}{2\pi} \ln \frac{5r_0}{|\xi - x|} + \frac{1}{2\pi} \ln 3$. Это завершает доказательство (1.10) с постоянной $C = \frac{1}{2\pi} \ln 3$ для случая (i).

(ii) Пусть теперь $S = \{|\xi - x| \leq \text{diam}(\Omega)\} \setminus \{(\xi_1, 0), 0 \leq \xi_1 \leq C_S\}$. Во-первых, заметим, что $G_S(x, \xi) \leq \frac{1}{2\pi} \ln \frac{\text{diam}(\Omega)}{C_S}$ при условии $|\xi - x| \geq C_S$. Далее, пусть G'_S обозначает функцию Грина для случая (i). Тогда, воспользовавшись принципом максимума для функции $G_S - G'_S$ в области $|\xi - x| \leq C_S$, получаем $|G_S - G'_S| \leq C$. Теперь оценка (1.10) в этой области вытекает из соответствующей оценки для G'_S , полученной для случая (i). \square

Л е м м а 1.7. Пусть $D \subset D' \subseteq \hat{\Omega} := \varepsilon^{-1}\Omega$, и выполняются условия

$\text{dist}\{\partial D \setminus \partial \hat{\Omega}, \partial D' \setminus \partial \hat{\Omega}\} \gtrsim 1$ и $\text{diam}(D') \simeq d$. Тогда для любой $v \in L_2(\Omega)$, удовлетворяющей $\Delta v \in L_2(\Omega)$, справедливо соотношение

$$\|v\|_{2,1;D} \lesssim d^{n/2} (\|\Delta v\|_{2;D'} + \|v\|_{2;D'}). \quad (1.11)$$

Д о к а з а т е л ь с т в о. Положим $\alpha \in (1, \frac{4}{3})$. Заметим, что $|v|_{2,\alpha;\Omega} \leq C_\alpha \|\Delta v\|_{\alpha;\Omega}$ в области Ω [74, Лемма 2.1], где постоянная $C_\alpha = C_\alpha(\Omega)$ принимает фиксированное значение. Отсюда получаем, что $|v|_{2,\alpha;\hat{\Omega}} \leq C_\alpha \|\Delta v\|_{\alpha;\hat{\Omega}}$ в растянутой области $\hat{\Omega}$. Также справедливо, что $|\omega v|_{2,\alpha;\hat{\Omega}} \leq C_\alpha \|\Delta(\omega v)\|_{\alpha;\hat{\Omega}}$ для некоторой гладкой функции-срезки ω , равной 1 в D и исчезающей в $\hat{\Omega} \setminus D'$. Отсюда получаем соотношение

$$|v|_{2,\alpha;D} \lesssim \|\Delta v\|_{\alpha;D'} + \|\nabla v\|_{\alpha;D'} + \|v\|_{\alpha;D'},$$

где мы воспользовались тем, что $\text{dist}\{\partial D \setminus \partial \hat{\Omega}, \partial D' \setminus \partial \hat{\Omega}\} \gtrsim 1$. Далее, в силу $|D| \leq |D'| \lesssim d^n$, получаем $|\cdot|_{2,1;D} \leq |\cdot|_{2,\alpha;D} \cdot |D'|^{1-1/\alpha}$ и $\|\cdot\|_{\alpha;D'} \leq \|\cdot\|_{2;D'} \cdot |D'|^{1/\alpha-1/2}$, поэтому

$$|v|_{2,1;D} \lesssim d^{n/2} (\|\Delta v\|_{2;D'} + \|\nabla v\|_{2;D'} + \|v\|_{2;D'}).$$

Наконец, объединим этот результат с $\|\nabla v\|_{2;D'} \leq C(\|\Delta v\|_{2;D''} + \|v\|_{2;D''})$, где область D'' определяется для D' аналогично D' для D (а постоянная C не зависит от размера области). Остается ввести замену обозначений $D'' =: D'$.

□

Д о к а з а т е л ь с т в о Т е о р е м ы 1.2. Рассмотрим два случая.

С л у ч а й 1: $0 < \varepsilon^2 \leq C_f \simeq 1$. Начнем с (1.8а). Воспользовавшись принципом максимума, легко показать, что $0 \leq G(x; \xi) \leq g_n(x; \xi)$, где g_n — функция Грина для оператора $-\varepsilon^2 \Delta + C_f$ в \mathbb{R}^n . Вспомним, что [46]

$$g_2 = \frac{1}{2\pi\varepsilon^2} K_0(\sqrt{C_f} r/\varepsilon), \quad g_3 = \frac{1}{4\pi\varepsilon^3} \frac{e^{-\sqrt{C_f} r/\varepsilon}}{r/\varepsilon}, \quad r = |\xi - x|. \quad (1.12)$$

Здесь K_0 — модифицированная функция Бесселя второго рода нулевого порядка, для которой справедливы оценки [50]

$$\begin{aligned} K_0(s) &\lesssim \ln(2 + s^{-1}), \quad s > 0, \\ K_0(s) &\lesssim s^{-1/2}e^{-s}, \quad s \gtrsim 1. \end{aligned} \tag{1.13}$$

Теперь соотношение (1.8a) легко выводится с помощью аналогичной оценки для $\|g_n(x, \cdot)\|_{1, B(x, \rho)}$.

Далее, (1.8b) и оценки для $\|G(x, \cdot)\|_{1, \Omega}$ и $\|G(x, \cdot)\|_{\frac{n}{n-1}; \Omega}$ в (1.7) могут быть получены аналогично, в силу (1.12) и (1.13).

Заметим, что (1.8c) следует из (1.8a), (1.8d) и (1.8e). Чтобы в этом убедиться, воспользуемся функцией-срезкой ω , равной 1 в $\Omega \setminus B(x, \rho)$ и исчезающей в $B(x, \frac{1}{2}\rho) \cap \Omega$. Тогда из соболевского вложения $W_1^2(\Omega) \hookrightarrow W_{\frac{n}{n-1}}^1(\Omega)$ следует, что

$$\begin{aligned} \|\nabla G\|_{\frac{n}{n-1}; \Omega \setminus B(x, \rho)} &\lesssim \|\nabla(\omega G)\|_{\frac{n}{n-1}; \Omega} \\ &\lesssim |G|_{2,1; \Omega \setminus B(x, \frac{1}{2}\rho)} + \rho^{-1}\|\nabla G\|_{1; B(x, \rho) \cap \Omega} + \rho^{-2}\|G\|_{1; B(x, \rho) \cap \Omega}. \end{aligned} \tag{1.14}$$

Отсюда, действительно, получаем (1.8c), в силу (1.8a), (1.8d) и (1.8e).

Чтобы установить оставшиеся оценки, введем в рассмотрение вспомогательную функцию Грина \bar{G} для оператора $-\varepsilon^2\Delta$ в области $B(x; 2\varepsilon) \cap \Omega$. Очевидно, что \bar{G} является нормализованной функцией Грина для оператора $-\Delta$ в растянутых переменных, для которой справедлива Лемма 1.5. Точнее, $\bar{G}(x, \xi) = \varepsilon^{-n}G_0(x/\varepsilon, \xi/\varepsilon)$, где G_0 — функция Грина для $-\Delta$ в области $\varepsilon^{-1}[B(x; 2\varepsilon) \cap \Omega]$, поэтому Лемма 1.5 для G_0 приводит к оценкам (1.8d) и (1.8e) для \bar{G} с Ω замененной на $B(x; 2\varepsilon) \cap \Omega$.

Следовательно для завершения доказательства достаточно установить

следующие оценки:

$$\varepsilon^2 |(\bar{G} - G)(x; \cdot)|_{2,1;B(x;\varepsilon)\cap\Omega} + \varepsilon |(\bar{G} - G)(x; \cdot)|_{1,1;B(x;\varepsilon)\cap\Omega} \lesssim 1, \quad (1.15a)$$

$$|(\bar{G} - G)(x; \cdot)|_{1,1;B(x;\rho)\cap\Omega} \lesssim \varepsilon^{-2}\rho, \quad \rho \leq \varepsilon, \quad (1.15b)$$

$$\varepsilon^2 |G(x; \cdot)|_{2,1;\Omega \setminus B(x;\varepsilon)} + \varepsilon |G(x; \cdot)|_{1,1;\Omega \setminus B(x;\varepsilon)} \lesssim 1. \quad (1.15c)$$

Действительно, оценка для $|G(x, \cdot)|_{1,1;\Omega}$ в (1.7) следует из (1.15a), (1.15c) и версии (1.7) для \bar{G} . Заметим, что (1.7) приводит к (1.8d) при $\rho \geq \varepsilon$. При $\rho \leq \varepsilon$ оценка (1.8d) следуют из (1.15b), (1.15c) и версии (1.8d) для \bar{G} . Наконец, оценка (1.8e) следует из (1.15a), (1.15c) и версии (1.8e) для \bar{G} .

Приступим к доказательству оценок (1.15).

Для (1.15a), положим $w(\xi) := \bar{G} - G$ при $\xi \in B(x; 2\varepsilon) \cap \Omega$. Отметим, что из (1.6) следует, что $-\varepsilon^2 \Delta_\xi w = C_f G$. Далее, воспользовавшись растянутой переменной $\hat{\xi} = \xi/\varepsilon$ и обозначением $\hat{v}(\hat{\xi}) := v(\xi)$ для произвольной функции v , а также $\hat{D} := \varepsilon^{-1}D$ для произвольной области D , получаем $-\Delta \hat{w} = C_f \hat{G}$ в $\hat{B}(x; 2\varepsilon) \cap \hat{\Omega}$, и затем $|\Delta \hat{w}| + |\hat{w}| \lesssim \hat{G} + \hat{\hat{G}}$. Отсюда, применив (1.11) с $d = 1$, получаем

$$\begin{aligned} \|\hat{w}\|_{2,1;\hat{B}(x;\varepsilon)\cap\hat{\Omega}} &\lesssim \|\Delta \hat{w}\|_{2;\hat{B}(x;2\varepsilon)\cap\hat{\Omega}} + \|\hat{w}\|_{2;\hat{B}(x;2\varepsilon)\cap\hat{\Omega}} \\ &\lesssim \|\hat{G} + \hat{\hat{G}}\|_{2;\hat{B}(x;2\varepsilon)\cap\hat{\Omega}}. \end{aligned}$$

Перепишав этот результат в терминах переменной ξ , приходим к соотношению

$$\varepsilon^{-n} \left\{ \varepsilon^2 |w|_{2,1;B(x;\varepsilon)\cap\Omega} + \varepsilon |w|_{1,1;B(x;\varepsilon)\cap\Omega} \right\} \lesssim \|\hat{G} + \hat{\hat{G}}\|_{2;\hat{B}(x;2\varepsilon)\cap\hat{\Omega}} \lesssim \varepsilon^{-n},$$

где мы также воспользовались $G + \bar{G} \leq g_n$ и (1.12). Отсюда немедленно следует (1.15a).

Чтобы установить (1.15b), мы частично воспроизведем доказательство (1.15a) с областями $B(x; \varepsilon)$ и $B(x; 2\varepsilon)$ соответственно замененными на $B(x; \rho)$ и $B(x; \rho + \varepsilon)$. В частности,

$$\varepsilon^{-n} \left\{ \varepsilon |w|_{1;B(x;\rho)\cap\Omega} \right\} = \|\nabla \hat{w}\|_{1;\hat{B}(x;\rho)\cap\hat{\Omega}} \lesssim (\rho/\varepsilon)^{n/2} \|\nabla \hat{w}\|_{2;\hat{B}(x;\rho)\cap\hat{\Omega}},$$

в то время как $-\Delta\hat{w} = C_f \hat{G}$ приводит к

$$\|\nabla\hat{w}\|_{2;\hat{B}(x;\rho)\cap\hat{\Omega}} \lesssim \|\hat{G} + \hat{\tilde{G}}\|_{2;\hat{B}(x;\rho+\varepsilon)\cap\hat{\Omega}} \lesssim \varepsilon^{-n}.$$

Отсюда получаем (1.15b), поскольку $(\rho/\varepsilon)^{n/2} \leq \rho/\varepsilon$ при $\rho \leq \varepsilon$ и $n = 2, 3$.

Наконец, для доказательства (1.15c), положим $\rho_j := 2^j$ и разделим область $\Omega \setminus B(x; \varepsilon)$ на непересекающиеся области $\mathcal{D}_j := [B(x, \varepsilon\rho_{j+1}) \setminus B(x, \varepsilon\rho_j)] \cap \Omega$ где $j = 0, 1, \dots$. Также, $\mathcal{D}_j \subset \mathcal{D}'_j := \mathcal{D}_{j-1} \cup \bar{\mathcal{D}}_j \cup \mathcal{D}_{j+1}$ с тем, чтобы $\text{dist}(\partial\mathcal{D}'_j \setminus \partial\Omega, \partial\mathcal{D}_j \setminus \partial\Omega) \geq \varepsilon/2$. Уравнение из (1.6) переписывается как $-\Delta\hat{G} + C_f\hat{G} = 0$ в каждой \mathcal{D}'_j , поэтому, воспользовавшись (1.11) с $d = \rho_{j-1} \geq \frac{1}{2}$, получаем

$$\|\hat{G}\|_{2,1;\hat{\mathcal{D}}_j} \lesssim \rho_j^{n/2} \|\hat{G}\|_{2;\hat{\mathcal{D}}'_j} \lesssim \rho_j^n \|G\|_{\infty;\mathcal{D}'_j}.$$

В силу $G \leq g_n$ and (1.12), справедливо $\rho_j^n \|G\|_{\infty;\mathcal{D}'_j} \lesssim \rho_j^{\mu_n} \varepsilon^{-n} e^{-c\rho_j}$, где, из (1.12) и (1.13), $\mu_2 = 3/2$ и $\mu_3 = 2$. Отсюда, в терминах исходной переменной ξ ,

$$\varepsilon^{-n} \left\{ \varepsilon^2 |G(x; \cdot)|_{2,1;\Omega \setminus B(x;\varepsilon)} + \varepsilon |G(x; \cdot)|_{1,1;\Omega \setminus B(x;\varepsilon)} \right\} \lesssim C \varepsilon^{-n} \sum_{j=1}^{\infty} \rho_j^{\mu_n} e^{-c\rho_j} \lesssim \varepsilon^{-n}.$$

Это немедленно приводит к последней оценке (1.15c) в (1.15) при $0 \leq \varepsilon^2 \leq C_f \simeq 1$.

С л у ч а й 2: $\varepsilon^2 \simeq 1$, $C_f \geq 0$. Доказательство для этого более простого случая, не являющегося сингулярно возмущенным, здесь опущено; см. полное доказательство [157, Теоремы 1]. \square

1.3 Вывод апостериорных оценок

1.3.1 Представление ошибки решения через функцию Грина

В [125, §4.1], при рассмотрении монотонного полулинейного уравнения авторы воспользовались барьерными функциями с тем, чтобы представить ошибку решения через функцию Грина оператора Лапласа (и таким образом избежать рассмотрения функции Грина более сложного линеаризованного оператора). Мы воспользуемся во многом упрощенной версией этого подхода, при

этом в качестве упрощенного оператора будет использоваться $-\varepsilon^2\Delta + C_f$, поскольку последний сохраняет ключевые свойства сингулярно возмущенного уравнения из (1.1).

Для произвольных $u, v \in C(\bar{\Omega})$, определим вспомогательную функцию w , удовлетворяющую соотношениям

$$-\varepsilon^2\Delta w + C_f w = [f(\cdot, v) - f(\cdot, u)] - C_f[v - u] \quad \text{в } \Omega, \quad w = 0 \quad \text{на } \partial\Omega. \quad (1.16)$$

В следующей лемме будет установлено весьма полезное представление для разности $v - u$ (где можно полагать, что $v = u_h$) через функцию Грина оператора $-\varepsilon^2\Delta + C_f$.

Л е м м а 1.8. Пусть $e = [v - u] + w$, где w удовлетворяет (1.16), а $C_f \geq 0$. Тогда

$$\|v - u\|_{\infty; \Omega} \leq 2\|e\|_{\infty; \Omega}, \quad (1.17a)$$

$$e(x) = \varepsilon^2(\nabla v, \nabla G(x, \cdot)) + (f(\cdot, v), G(x, \cdot)), \quad (1.17b)$$

где G удовлетворяет (1.6).

Д о к а з а т е л ь с т в о. При любом фиксированном $\theta > 0$, положим $\Omega' = \{|u - v| > \theta\}$. Область Ω' является корректно определенной подобластью Ω , поскольку $u, v \in C(\bar{\Omega})$. Тогда $|w| \leq \|e\|_{\infty; \Omega} + \theta$ в $\Omega \setminus \Omega'$, включая $\partial\Omega'$. Далее, в Ω' введем $p(x) := \frac{f(\cdot, v) - f(\cdot, u)}{v - u} \geq C_f$ и заметим, что $p \in L_\infty(\Omega')$. Уравнение (1.16) для w в Ω' может быть переписано как $-\varepsilon^2\Delta w + pw = (p - C_f)e$. Положим $w^\pm := \|e\|_{\infty; \Omega} + \theta \pm w$. Легко показать, что $[-\varepsilon^2\Delta + p]w^\pm \geq p\|e\|_{\infty; \Omega} \pm (p - C_f)e \geq 0$ в Ω' , а $w^\pm \geq 0$ на $\partial\Omega'$. Поэтому, воспользовавшись слабым принципом максимума (см. Теорему 8.1 в [92]), получаем $w^\pm \geq 0$, т. е. $|w| \leq \|e\|_{\infty; \Omega} + \theta$ в Ω' , а следовательно и в Ω . Поскольку это справедливо для любого $\theta > 0$, заключаем, что $|w| \leq \|e\|_{\infty; \Omega}$ в Ω . Это немедленно ведет к (1.17a). Чтобы установить (1.17b), заметим, что определение G приводит к представлению

$$e(x) = \varepsilon^2(\nabla e, \nabla G(x, \cdot)) + (C_f e, G(x, \cdot)).$$

Отсюда, в силу (1.16) и (1.1), получаем (1.17b). \square

Наконец, положим $v := u_h$ и приведем соответствующее представление для $e(x)$, которое будет использоваться при получении апостериорных оценок. Зафиксируем $x \in \Omega$, например, полагая, что $|e(x)|$ принимает свое максимальное значение в Ω в этой точке x , и будем писать $G = G(x, \cdot)$ для функции Грина из (1.6). Поскольку $v = u_h$, соотношения (1.17b) и (3.2) приводят при любом $G_h \in S_h$ к представлению

$$\begin{aligned} e(x) &= \varepsilon^2(\nabla u_h, \nabla G) + (f(\cdot, u_h), G) \\ &= \varepsilon^2(\nabla u_h, \nabla(G - G_h)) + (f_h, G - G_h) \\ &\quad + (f_h, G_h) - (f_h, G_h)_h, \quad \text{где } f_h := f(\cdot, u_h). \end{aligned} \quad (1.18)$$

1.3.2 Вывод апостериорных оценок без учета квадратур

Введем некоторый непрерывный кусочно-линейный квазиинтерполянт G_h для функции $G = G(x, \cdot)$, который будет описан ниже. Здесь $x \in \Omega$ остается фиксированным, и, таким образом, интерполяция осуществляется по второму аргументу G . Отметим, что $G_h \in S_h$ при любом $r \geq 1$. Воспользуемся достаточно стандартным определением для G_h [67, 135]. Пусть \mathcal{N} обозначает множество вершин сетки \mathcal{T} , а ϕ_z является стандартной кусочно-линейной базисной функцией, соответствующей $z \in \mathcal{N}$. Если $z \in \Omega$, то положим F_z равным любому элементу сетки $T \in \mathcal{T}$, для которого $z \in \bar{T}$. Если же $z \in \partial\Omega$, то положим F_z равным грани (т. е. $n - 1$ -симплексу) некоторого $T \in \mathcal{T}$, для которой $z \in \bar{F}_z \subset \partial\Omega$. Далее, пусть функция $\psi_z \in \mathbb{P}_1(F_z)$ будет дуальной для $\{\phi_{z'}\}$ на F_z в том смысле, что $\int_{F_z} \psi_z \phi_{z'} = 1$ при $z = z'$ и 0 в противном случае. Здесь \mathbb{P}_m обозначает пространство полиномов степени, не превышающей m . Положим $G_h := \sum_{z \in \mathcal{N}} \phi_z \int_{F_z} G \psi_z = \sum_{z \in \mathcal{N}_I} \phi_z \int_{F_z} G \psi_z$, где \mathcal{N}_I — множество внутренних узлов сетки. Все F_z в последней сумме являются n -симплексами. При таком определении G_h удовлетворяет следующим локальным свойствам:

$$|G - G_h|_{k,p;T} \lesssim h_T^{j-k} |G|_{j,p;\omega_T} \quad \text{для } T \in \mathcal{T}, \quad (1.19)$$

при любых $0 \leq k \leq j \leq 2$ и $1 \leq p \leq \infty$, при условии, что правая часть в (1.19) определена. Здесь обозначение ω_T используется для области, включающей T и все элементы в \mathcal{T} , касающиеся T .

Справедлива следующая лемма.

Л е м м а 1.9. Пусть x является произвольной точкой в Ω , $G = G(x, \cdot)$, а G_h — вышеопределенный кусочно-линейный квазиинтерполант. Тогда справедлива оценка

$$\begin{aligned} |\varepsilon^2(\nabla u_h, \nabla(G - G_h)) + (f_h, G - G_h)| \lesssim \\ \max_{T \in \mathcal{T}} \left[\min\{1, \ell_{h,x} h_T^2 \varepsilon^{-2}\} \|\varepsilon^2 \Delta u_h - f(\cdot, u_h)\|_{L^\infty(T)} \right. \\ \left. + \min\{\varepsilon, \ell_{h,x} h_T\} \|\llbracket \nabla u_h \rrbracket\|_{\infty; \partial T} \right]. \end{aligned} \quad (1.20)$$

Здесь $\llbracket \nabla u_h \rrbracket$ обозначает стандартный скачок нормальной производной u_h при переходе через границу элемента сетки, а

$$\ell_{h,x} := \ln(2 + \varepsilon h_{T_0}^{-1}) \quad \text{где } T_0 \ni x. \quad (1.21)$$

Д о к а з а т е л ь с т в о. Поскольку $(f_h, G_h) = (f_h, G_h)_h$ в (1.18), несложно убедиться, что

$$\begin{aligned} e(x) &= \frac{1}{2} \sum_{T \in \mathcal{T}_h} \varepsilon^2 \int_{\partial T} (G - G_h) \llbracket \nabla u_h \rrbracket \cdot \nu + \sum_{T \in \mathcal{T}_h} \int_T (f_h - \varepsilon^2 \Delta u_h) (G - G_h) \\ &=: I + II. \end{aligned}$$

Далее,

$$|II| \lesssim \max_{T \in \mathcal{T}_h} \alpha_T \|f_h - \varepsilon^2 \Delta_h u_h\|_{\infty, \Omega} \sum_{T \in \mathcal{T}_h} \alpha_T^{-1} \|G - G_h\|_{1; T},$$

$$\alpha_T = \min\{\varepsilon^2, \ell_{h,x} h_T^2\}.$$

В силу (1.19),

$$\|G - G_h\|_{1; T} \lesssim \min\{\|G\|_{1; \omega_T}, h_T^2 \|D^2 G\|_{1; \omega_T}\}.$$

Поскольку $\alpha_T^{-1} \leq \varepsilon^{-2} + \ell_{h,x}^{-1} h_T^{-2}$, получаем оценку

$$\alpha_T^{-1} \|G - G_h\|_{1; \omega_T} \lesssim \min\{\varepsilon^{-2} \|G\|_{1; \omega_T} + \ell_{h,x}^{-1} \|D^2 G\|_{1; \omega_T}, (\varepsilon^{-2} + \ell_{h,x}^{-1} h_T^{-2}) \|G\|_{1; \omega_T}\}.$$

Для любого $T \in \mathcal{T}$, пусть ω'_T обозначает область, включающую ω_T и все элементы сетки, касающиеся ω_T . Пусть также $x \in T_0$. Тогда

$$|II| \lesssim \max_{T \in \mathcal{T}} \alpha_T \|f_h - \varepsilon^2 \Delta u_h\|_{\infty, T} S_{II},$$

где, в силу (1.7), (1.8a) и (1.8e),

$$\begin{aligned} S_{II} &\lesssim \sum_{T: T \notin \omega'_{T_0}} (\varepsilon^{-2} \|G\|_{1; \omega_T} + \ell_{h,x}^{-1} \|D^2 G\|_{1; \omega_T}) + (\varepsilon^{-2} + \ell_{h,x}^{-1} h_T^{-2}) \|G\|_{1; \omega'_T} \\ &\lesssim \varepsilon^{-2} \|G\|_{1; \Omega} + \ell_{h,x}^{-1} \|D^2 G\|_{1; \Omega \setminus B(x; ch_{T_0})} + (\varepsilon^{-2} + \ell_{h,x}^{-1} h_T^{-2}) \|G\|_{1; B(x; Ch_{T_0})} \\ &\lesssim \varepsilon^{-2}. \end{aligned}$$

Таким образом,

$$|II| \lesssim \max_T (\min\{1, \ell_{h,x} h_T^2 \varepsilon^{-2}\} \|f_h - \varepsilon^2 \Delta u_h\|_{\infty, T}).$$

Далее, для I , получаем оценку

$$|I| \lesssim \varepsilon^2 \max_{T \in \mathcal{T}} \beta_T \|\llbracket \nabla u_h \rrbracket\|_{\infty, \partial T} \sum_{T \in \mathcal{T}} \beta_T^{-1} \|G - G_h\|_{1; \partial T}, \quad \beta_T = \min\{\varepsilon, \ell_{h,x} h_T\}.$$

В силу теоремы о следах и (1.19), справедливо соотношение

$$\|G - G_h\|_{1; \partial T} \lesssim \|\nabla(G - G_h)\|_{1; T} + h_T^{-1} \|G - G_h\|_{1; T} \lesssim \min\{\|\nabla G\|_{1; \omega_T}, h_T \|D^2 G\|_{1; \omega_T}\}.$$

Заметим, что $\beta_T^{-1} \leq \varepsilon^{-1} + (\ell_{h,x} h_T)^{-1}$, а $\ell_{h,x}^{-1} \lesssim 1$, поэтому

$$\beta_T^{-1} \|G - G_h\|_{1; \partial T} \lesssim \min\{\varepsilon^{-1} \|\nabla G\|_{1; \omega_T} + \ell_{h,x}^{-1} \|D^2 G\|_{1; \omega_T}, (\varepsilon^{-1} + h_T^{-1}) \|\nabla G\|_{1; \omega_T}\}.$$

Следовательно

$$|I| \lesssim \varepsilon^2 \max_{T \in \mathcal{T}} \beta_T \|\llbracket \nabla u_h \rrbracket\|_{\infty, \partial T} S_I,$$

где, в силу (1.7), (1.8d) и (1.8e),

$$\begin{aligned} S_I &\lesssim \sum_{T: T \notin \omega'_{T_0}} (\varepsilon^{-1} \|\nabla G\|_{1; \omega_T} + \ell_{h,x}^{-1} \|D^2 G\|_{1; \omega_T}) + (\varepsilon^{-1} + h_{T_0}^{-1}) \|\nabla G\|_{1; \omega'_T} \\ &\lesssim \varepsilon^{-1} \|\nabla G\|_{1; \Omega} + \ell_{h,x}^{-1} \|D^2 G\|_{1; \Omega \setminus B(x; ch_{T_0})} + (\varepsilon^{-1} + h_{T_0}^{-1}) \|\nabla G\|_{1; B(x; Ch_{T_0})} \\ &\lesssim \varepsilon^{-2}. \end{aligned} \tag{1.22}$$

Наконец,

$$|I| \lesssim \max_T (\min\{\varepsilon, \ell_{h,x} h_T\} \|[\nabla u_h]\|_{\infty, \partial T}).$$

Объединяя полученные оценки, завершаем доказательство леммы. \square

1.3.3 Апостериорные оценки с учетом квадратур

Поскольку мы уже получили оценку Леммы 1.9, остается оценить ошибку квадратуры $(f_h, G_h) - (f_h, G_h)_h$ в (1.18), чему и посвящен настоящий раздел. Подход, который будет здесь использован, в большой степени имитирует доказательство Леммы 3.2 в [125]. В то же время, по сравнению с [125], потребуются серьезные изменения в доказательстве, с тем чтобы адекватно отразить свойства рассматриваемого сингулярно возмущенного уравнения. Пусть $E_T(g) = \int_T g \, dx - (g, 1)_{h,T}$ обозначает квадратурную ошибку на элементе T . Аналогично [125], будем предполагать, что используемая квадратурная формула является точной для многочленов степени q :

$$E_T(\psi) = 0 \quad \text{для } \psi \in \mathbb{P}_q, \quad (1.23a)$$

а также устойчивой в норме L_∞ в следующем смысле:

$$|E_T(\psi)| \lesssim |T| \|\psi\|_{\infty; T} \quad \text{для } \psi \in C(\bar{T}). \quad (1.23b)$$

Дополнительно, будем предполагать, что квадратурная формула линейна. Легко проверить, что данные предположения удовлетворяются стандартными квадратурами, используемыми в конечноэлементных кодах.

Л е м м а 1.10. Пусть I_h^j является стандартным интерполяционным оператором Лагранжа степени j , и пусть μ^j и λ обозначают следующие кусочно-постоянные функции: $\mu^j = \mu_T^j := \|f_h - I_h^j f_h\|_{\infty; T}$ и $\lambda = \lambda_T := \min\{1, \varepsilon^{-1} h_T\}$ на любом T . Введем также два произвольных разбиения множества \mathcal{T} на непересекающиеся подмножества: $\mathcal{T}_1 \cup \mathcal{T}'_1 = \mathcal{T}$ и $\mathcal{T}_2 \cup \mathcal{T}'_2 = \mathcal{T}$.

Тогда при выполнении условий (1.23) справедлива оценка

$$(f_h, G_h) - (f_h, G_h)_h \lesssim \mu_{\text{quad}} := \|\mu^q\|_{\infty; \mathcal{T}_1} + \varepsilon^{-2} \ell_{h,x} \|\mu^q\|_{\frac{n}{2}; \mathcal{T}'_1} + \|\lambda \mu^{q-1}\|_{\infty; \mathcal{T}_2} + \varepsilon^{-1} \ell_{h,x} \|\lambda \mu^{q-1}\|_{n; \mathcal{T}'_2}. \quad (1.24)$$

Более того, $\mathcal{T}_i, \mathcal{T}'_i$, $i = 1, 2$, могут быть выбраны таким образом, что

$$\begin{aligned} \mu_{\text{quad}} &\lesssim \mu_{\Sigma}^q + \mu_{\Sigma}^{q-1} \\ &:= \|\min\{h_T^{-2}, \varepsilon^{-2} \ell_{h,x}\} \mu^q\|_{\frac{n}{2}; \mathcal{T}} + \|\min\{h_T^{-1}, h_T \varepsilon^{-2} \ell_{h,x}\} \mu^{q-1}\|_{n; \mathcal{T}}. \end{aligned} \quad (1.25)$$

Д о к а з а т е л ь с т в о. Заметим, что $(f_h, G_h) - (f_h, G_h)_h = E_T(f_h G_h)$. Положим $G_{h,T} = \frac{1}{|T|} \int_T G_h \, dx$. Тогда для любого $T \in \mathcal{T}$

$$\begin{aligned} E_T(f_h G_h) &= E_T(f_h G_{h,T}) + E_T(f_h [G_h - G_{h,T}]) \\ &= E_T([f_h - I_h^q f_h] G_{h,T}) + E_T([f_h - I_h^{q-1} f_h] [G_h - G_{h,T}]), \end{aligned} \quad (1.26)$$

где мы воспользовались (1.23а), а также $[I_h^q f_h] G_{h,T} \in \mathbb{P}_q$ и $I_h^{q-1} f_h [G_h - G_{h,T}] \in \mathbb{P}_q$ (поскольку на каждом элементе $G_{h,T}$ является постоянной, а G_h — линейной функцией).

Для первого слагаемого в (1.26), воспользовавшись (1.23b) и определением $G_{h,T}$, получаем оценку

$$|E_T([f_h - I_h^q f_h] G_{h,T})| \lesssim |T| \mu_T^q |G_{h,T}| \lesssim \mu_T^q \|G_h\|_{1;T} = (\mu^q, |G_h|)_T.$$

Пусть T_0 является элементом, содержащим точку x в (1.18), а ω'_{T_0} содержит ω_{T_0} и элементы, касающиеся ω_{T_0} , в то время как ω''_{T_0} содержит ω'_{T_0} и элементы, касающиеся ω'_{T_0} . Для произвольного разбиения множества $\mathcal{T} = \mathcal{T}_1 \cup \mathcal{T}'_1$ получаем

$$\begin{aligned} \sum_{T \in \mathcal{T}} |E_T([f_h - I_h^q f_h] G_{h,T})| &\lesssim (\mu^q, |G_h|) \\ &\lesssim \|\mu^q\|_{\infty; \mathcal{T}_1} \|G_h\|_{1; \mathcal{T}_1} + \|\mu^q\|_{\infty; \omega'_{T_0} \cap \mathcal{T}'_1} \|G_h\|_{1; \omega'_{T_0} \cap \mathcal{T}'_1} \\ &\quad + \|\mu^q\|_{\frac{n}{2}; \mathcal{T}_1 \setminus \omega'_{T_0}} \|G_h\|_{\frac{n}{n-2}; \mathcal{T}'_1 \setminus \omega'_{T_0}}. \end{aligned}$$

Далее, воспользовавшись (1.19), а затем (1.7), (1.8a) и (1.8b), получаем

$$\begin{aligned} \|G_h\|_{1;\mathcal{T}_1} &\lesssim \|G\|_{1;\Omega} \lesssim 1, \\ \|G_h\|_{1;\omega'_{T_0} \cap \mathcal{T}'_1} &\lesssim \|G\|_{1;\omega''_{T_0} \cap \Omega} \lesssim \varepsilon^{-2} h_{T_0}^2 \ell_{h,x}, \\ \|G_h\|_{\frac{n}{n-2};\mathcal{T}'_1 \setminus \omega'_{T_0}} &\lesssim \|G\|_{\frac{n}{n-2};\Omega \setminus \omega_{T_0}} \lesssim \varepsilon^{-2} \ell_{h,x}. \end{aligned}$$

Здесь также было принято во внимание, что $\omega''_{T_0} \subset B(x, ch_{T_0})$ и $\omega_{T_0} \supset B(x, c'h_{T_0})$ для некоторых c и c' . Следовательно приходим к оценке

$$\begin{aligned} \sum_{T \in \mathcal{T}} |E_T([f_h - I_h^q f_h] G_{h,T})| \\ \lesssim \|\mu^q\|_{\infty;\mathcal{T}_1} + \varepsilon^{-2} \ell_{h,x} (h_{T_0}^2 \|\mu^q\|_{\infty;\omega'_{T_0} \cap \mathcal{T}'_1} + \|\mu^q\|_{\frac{n}{2};\mathcal{T}'_1 \setminus \omega'_{T_0}}). \end{aligned}$$

Заметим, что $h_T^2 \|\mu^q\|_{\infty;T} \lesssim \|\mu^q\|_{\frac{n}{2};T}$. Это наблюдение будет полезным при рассмотрении $T \in \omega'_{T_0} \cap \mathcal{T}'_1$. Поскольку таких элементов T конечное число, и для каждого их них $h_T \simeq h_{T_0}$, немедленно заключаем, что $h_T^2 \|\mu^q\|_{\infty;\omega'_{T_0} \cap \mathcal{T}'_1} \lesssim \|\mu^q\|_{\frac{n}{2};\omega'_{T_0} \cap \mathcal{T}'_1}$. Поэтому для первого слагаемого в (1.26) наконец получаем оценку

$$\sum_{T \in \mathcal{T}} |E_T([f_h - I_h^q f_h] G_{h,T})| \lesssim \|\mu^q\|_{\infty;\mathcal{T}_1} + \varepsilon^{-2} \ell_{h,x} \|\mu^q\|_{\frac{n}{2};\mathcal{T}'_1}. \quad (1.27)$$

Второе слагаемое в (1.26) оценивается аналогичным образом. Сначала воспользуемся (1.23b), а затем соответствующим обратным неравенством:

$$|E_T([f_h - I_h^{q-1} f_h] [G_h - G_{h,T}])| \lesssim \mu_T^{q-1} \|G_h - G_{h,T}\|_{1;T} = (\lambda \mu^{q-1}, z_h)_T.$$

Здесь используется вспомогательная функция $z_h := \lambda_T^{-1} |G_h - G_{h,T}|$ на любом T . Таким образом, для произвольного разбиения множества $\mathcal{T} = \mathcal{T}_2 \cup \mathcal{T}'_2$ получаем

$$\begin{aligned} \sum_{T \in \mathcal{T}} |E_T([f_h - I_h^{q-1} f_h] [G_h - G_{h,T}])| &\lesssim (\lambda \mu^{q-1}, |z_h|) \\ &\lesssim \|\lambda \mu^{q-1}\|_{\infty;\mathcal{T}_2} \|z_h\|_{1;\mathcal{T}_2} + \|\lambda \mu^{q-1}\|_{\infty;\omega'_{T_0} \cap \mathcal{T}'_2} \|z_h\|_{1;\omega'_{T_0} \cap \mathcal{T}'_2} \\ &\quad + \|\lambda \mu^{q-1}\|_{n;\mathcal{T}'_2 \setminus \omega'_{T_0}} \|z_h\|_{\frac{n}{n-1};\mathcal{T}'_2 \setminus \omega'_{T_0}}. \end{aligned}$$

Отметим, что $\lambda_T = \min\{1, \varepsilon^{-1}h_T\}$ ведет к $\lambda_T^{-1} \leq 1 + \varepsilon h_T^{-1}$. Воспользовавшись также определением и аппроксимационными свойствами $G_{h,T}$, а затем (1.19) с $k = j = 0, 1$ и $p = 1, \frac{n}{n-1}$, приходим к оценке

$$\|z_h\|_{p;T}^p = \lambda_T^{-p} \|G_h - G_{h,T}\|_{p;T}^p \lesssim \|G_h\|_{p;T}^p + |\varepsilon G_h|_{1,p;T}^p \lesssim \|G\|_{p;\omega_T}^p + |\varepsilon G|_{1,p;\omega_T}^p.$$

Объединяя ее с (1.7), (1.8a), (1.8c) и (1.8d), заключаем, что

$$\|z_h\|_{1;\mathcal{T}_2} \lesssim \|G\|_{1;\Omega} + \varepsilon |G|_{1,1;\Omega} \lesssim 1,$$

$$\|z_h\|_{1;\omega'_{T_0} \cap \mathcal{T}'_2} \lesssim \|G\|_{1;\omega''_{T_0} \cap \Omega} + \varepsilon |G|_{1,1;\omega''_{T_0} \cap \Omega} \lesssim \min\left\{\frac{h_{T_0}^2}{\varepsilon^2} \ell_{h,x} + \frac{h_{T_0}}{\varepsilon}, 1\right\} \lesssim \varepsilon^{-1} h_{T_0},$$

$$\|z_h\|_{\frac{n}{n-1};\mathcal{T}'_2 \setminus \omega'_{T_0}} \lesssim \|G\|_{\frac{n}{n-1};\Omega \setminus \omega_{T_0}} + \varepsilon |G|_{1,\frac{n}{n-1};\Omega \setminus \omega_{T_0}} \lesssim \varepsilon^{-1} \ell_{h,x}.$$

Здесь мы вновь воспользовались тем, что $\omega''_{T_0} \subset B(x, ch_{T_0})$, а $\omega_{T_0} \supset B(x, c'h_{T_0})$.

Следовательно

$$\begin{aligned} & \sum_{T \in \mathcal{T}} |E_T([f_h - I_h^{q-1} f_h][G_h - G_{h,T}])| \\ & \lesssim \|\lambda \mu^{q-1}\|_{\infty;\mathcal{T}_2} + \varepsilon^{-1} h_{T_0} \|\lambda \mu^{q-1}\|_{\infty;\omega'_{T_0} \cap \mathcal{T}'_2} + \varepsilon^{-1} \ell_{h,x} \|\lambda \mu^{q-1}\|_{n;\mathcal{T}'_2 \setminus \omega'_{T_0}}. \end{aligned}$$

Заметим, что $h_T \|\lambda \mu^{q-1}\|_{\infty;T} \lesssim \|\lambda \mu^{q-1}\|_{n;T}$. Поскольку найдется лишь конечное число таких элементов T , что $T \in \omega'_{T_0} \cap \mathcal{T}'_2$, и для любого из них справедливо $h_T \simeq h_{T_0}$, то $h_{T_0} \|\lambda \mu^{q-1}\|_{\infty;\omega'_{T_0} \cap \mathcal{T}'_2} \lesssim \|\lambda \mu^{q-1}\|_{n;\omega'_{T_0} \cap \mathcal{T}'_2}$. Следовательно для второго слагаемого в (1.26) наконец получаем оценку

$$\sum_{T \in \mathcal{T}} |E_T([f_h - I_h^{q-1} f_h][G_h - G_{h,T}])| \lesssim \|\lambda \mu^{q-1}\|_{\infty;\mathcal{T}_2} + \varepsilon^{-1} \ell_{h,x} \|\lambda \mu^{q-1}\|_{n;\mathcal{T}'_2}.$$

Объединяя ее с (1.26) и (1.27), получаем желаемое соотношение (1.24).

Для получения же оценки (1.25) заметим, что $\|\mu^q\|_{\infty;T} \lesssim h_T^{-2} \|\mu^q\|_{\frac{n}{2};T}$, следовательно

$$\|\mu^q\|_{\infty;\mathcal{T}_1} \lesssim \left(\sum_{T \in \mathcal{T}_1} |T| (h_T^{-2} \mu_T^q)^{n/2} \right)^{2/n} = \|h_T^{-2} \mu^q\|_{\frac{n}{2};\mathcal{T}_1}. \quad (1.28)$$

Выбирая \mathcal{T}_1 содержащим элементы, удовлетворяющие $h_T^{-2} \lesssim \varepsilon^{-2} \ell_{h,x}$, а также применяя аналогичные вычисления для слагаемого $\|\lambda \mu^{q-1}\|_{\infty;\mathcal{T}_2}$, получаем (1.25). \square

1.3.4 Эффективность апостериорных оценок

Нам понадобятся некоторые новые определения. Пусть $\ell_h = \max_{x \in \Omega} \ell_{h,x}$, и

$$\eta_\infty(T) = \min\{1, \ell_h h_T^2 \varepsilon^{-2}\} \|\varepsilon^2 \Delta u_h - f_h\|_{\infty;T} + \min\{\varepsilon, \ell_h h_T\} \|\llbracket \nabla u_h \rrbracket\|_{\infty; \partial T}. \quad (1.29)$$

Для функции $f_h(x) = f(x, u_h)$, пусть $f_{h,T}$ является ее L_2 проекцией на $\mathbb{P}_{r-1}(T)$ при $T \in \mathcal{T}$. В дополнение, определим осцилляционные величины:

$$\begin{aligned} \text{osc}(T) &= \min\{1, \ell_h h_T^2 \varepsilon^{-2}\} \|f_h - f_{h,T}\|_{\infty;T}, \\ \text{osc}(\omega_T) &= \max_{T' \subset \omega_T} \text{osc}(T'). \end{aligned} \quad (1.30)$$

Введем также нормы

$$\|w\|_{2,1,\varepsilon;\omega} = \sum_{i=0}^2 \tilde{\varepsilon}^i |w|_{i,1;\omega}, \quad \omega \subset \Omega, \quad (1.31a)$$

$$\|w\|_{-2,\infty,\varepsilon;\omega} = \sup_{v \in H_0^1(\omega) \cap W_1^2(\omega), \|v\|_{2,1,\varepsilon;\omega} = 1} \int_{\omega} wv \, dx, \quad \omega \subset \Omega. \quad (1.31b)$$

Л е м м а 1.11. *Для любого $T \in \mathcal{T}$ справедлива оценка*

$$\begin{aligned} \eta_\infty(T) &\lesssim \ell_h \|u - u_h\|_{\infty;\omega_T} + \text{osc}(\omega_T) \\ &\quad + \min\{\min\{1, \ell_h h_T^2 \varepsilon^{-2}\} \|f - f_h\|_{\infty;\omega_T}, \ell_h \|f - f_h\|_{-2,\infty,\varepsilon;\omega_T}\}. \end{aligned} \quad (1.32)$$

Здесь $f = f(\cdot, u)$. Также, если $q \geq r - 1$, то

$$\begin{aligned} \|\mu^q\|_{\infty;T} + \|\lambda \mu^{q-1}\|_{\infty;T} &\lesssim \varepsilon^2 h_T^{-2} \|u - u_h\|_{\infty;T} + \|f_h - f_{h,T}\|_{\infty;T} \\ &\quad + \min\{\|f - f_h\|_{\infty;T}, (1 + \varepsilon^2 h_T^{-2}) \|f - f_h\|_{-2,\infty,\varepsilon;T}\}, \end{aligned} \quad (1.33a)$$

$$\begin{aligned} \varepsilon^{-2} \ell_h \|\mu^q\|_{\frac{n}{2};T} + \varepsilon^{-1} \ell_h \|\lambda \mu^{q-1}\|_{n;T} &\lesssim \ell_h \|u - u_h\|_{\infty;T} + \ell_h h_T^2 \varepsilon^{-2} \|f_h - f_{h,T}\|_{\infty;T} \\ &\quad + \min\{h_T^2 \varepsilon^{-2} \ell_h \|f - f_h\|_{\infty;T}, \ell_h (1 + h_T^2 \varepsilon^{-2}) \|f - f_h\|_{-2,\infty,\varepsilon;T}\}, \end{aligned} \quad (1.33b)$$

$$\begin{aligned} &\|\min\{h_T^{-2}, \varepsilon^{-2} \ell_h\} \mu^q\|_{\frac{n}{2};T} + \|\min\{h_T^{-1}, h_T \varepsilon^{-2} \ell_{h,x}\} \mu^{q-1}\|_{n;T} \\ &\lesssim \ell_h \|u - u_h\|_{\infty;T} + \text{osc}(T) \\ &\quad + \min\{\min\{1, \ell_h h_T^2 \varepsilon^{-2}\} \|f - f_h\|_{\infty;T}, \ell_h \|f - f_h\|_{-2,\infty,\varepsilon;T}\}. \end{aligned} \quad (1.33c)$$

Д о к а з а т е л ь с т в о. Заметим, что

$$\begin{aligned} & \varepsilon^2 \int_{\Omega} \nabla(u - u_h) \nabla v \, dx + \int_{\Omega} (f - f_h) v \, dx \\ &= \int_{\Omega} (\varepsilon^2 \Delta u_h - f_h) v \, dx + \frac{\varepsilon^2}{2} \sum_{T \in \mathcal{T}_h} \int_{\partial T} \llbracket \nabla u_h \rrbracket v \, ds, \quad v \in H_0^1(\Omega). \end{aligned} \quad (1.34)$$

Здесь мы несколько злоупотребляем обозначениями и понимаем функцию Δu_h как определенную поэлементно и, следовательно, принадлежащую $L_{\infty}(\Omega)$.

Сначала рассмотрим поэлементную невязку $\min\{1, \ell_h h_T^2 \varepsilon^{-2}\} \|\varepsilon^2 \Delta u_h - f_h\|_{\infty; T}$. Найдется такая функция $b_T \in \mathbb{P}_{2n+r+1}$, что $b_T = 0$ и $\nabla b_T = 0$ на ∂T , а также $\|b_T\|_{1; T} \lesssim 1$ и

$$\begin{aligned} \|\varepsilon^2 \Delta u_h - f_{h,T}\|_{\infty; T} &\lesssim \int_T (\varepsilon^2 \Delta u_h - f_{h,T}) b_T \, dx \\ &\lesssim \left| \int_T (\varepsilon^2 \Delta u_h - f_h) b_T \, dx \right| + \|f_h - f_{h,T}\|_{\infty; T}. \end{aligned} \quad (1.35)$$

Вычитая $\varepsilon^2 \Delta u - f = 0$ из $\varepsilon^2 \Delta u_h - f_h$, затем воспользовавшись (1.34), далее интегрируя по частям с учетом $\nabla b_T = 0$ на ∂T и наконец воспользовавшись обратными неравенствами и $\|b_T\|_1 \lesssim 1$, получаем

$$\begin{aligned} \int_T (\varepsilon^2 \Delta u_h - f_h) b_T \, dx &= \int_T \varepsilon^2 \nabla(u - u_h) \nabla b_T \, dx + \int_T (f - f_h) b_T \, dx \\ &= - \int_T \varepsilon^2 (u - u_h) \Delta b_T \, dx + \int_T (f - f_h) b_T \, dx \\ &\lesssim \varepsilon^2 h_T^{-2} \|u - u_h\|_{\infty; T} + \min\{\|f - f_h\|_{\infty; T}, \|b_T\|_{2,1,\varepsilon; T} \|f - f_h\|_{-2,\infty,\varepsilon; T}\}. \end{aligned} \quad (1.36)$$

В силу $\|\varepsilon^2 \Delta u_h - f_h\|_{\infty; T} \leq \|\varepsilon^2 \Delta u_h - f_{h,T}\|_{\infty; T} + \|f_h - f_{h,T}\|_{\infty; T}$, а также воспользовавшись (1.35) и (1.36) и убедившись, что $\min\{1, \ell_h h_T^2 \varepsilon^{-2}\} \|b_T\|_{2,1,\varepsilon; T} \lesssim \ell_h$, приходим к соотношению

$$\begin{aligned} \min\{1, \ell_h h_T^2 \varepsilon^{-2}\} \|\varepsilon^2 \Delta u_h - f_h\|_{\infty; T} &\lesssim \ell_h \|u - u_h\|_{\infty; T} + \text{osc}(T) \\ &+ \min\{\min\{1, \ell_h h_T^2 \varepsilon^{-2}\} \|f - f_h\|_{\infty; T}, \ell_h \|f - f_h\|_{-2,\infty,\varepsilon; T}\}, \end{aligned} \quad (1.37)$$

в котором правая часть не превосходит правую часть в (1.32), что и требовалось доказать.

Рассмотрим теперь $\min\{\varepsilon, \ell_h h_{\tilde{e}}\} \|\llbracket \nabla u_h \rrbracket\|_{\infty; \tilde{e}}$, т. е. невязку через границы элементов сетки, где $\tilde{e} = \tilde{T}_1 \cap \tilde{T}_2$, для некоторых $\tilde{T}_1, \tilde{T}_2 \in \mathcal{T}$, является $(n-1)$ -симплексом (т.е., например, внутренним ребром сетки при $n=2$), а $h_{\tilde{e}} := \text{diam } \tilde{e}$. Здесь мы несколько модифицируем стандартное доказательство для случая малого ε . При $h_{\tilde{e}} \leq \varepsilon$, положим $e := \tilde{e}$ и $T_i := \tilde{T}_i$, $i=1,2$. В противном случае выберем такую точку $x \in e$, что $\|\llbracket \nabla u_h \rrbracket\|_{\infty; \tilde{e}} = \llbracket \nabla u_h \rrbracket(x)$, а $e \subset \tilde{e}$ является $(n-1)$ -симплексом диаметра ε . Также, выберем такие n -симплексы $T_i \subset \tilde{T}_i$, $i=1,2$, что $e = T_1 \cap T_2$. Далее, положим $\alpha = \text{diam}(e) = \min\{h_{\tilde{e}}, \varepsilon\}$.

Найдется такая функция $b_e \in \mathbb{P}_{4n+r-4}(T_1 \cup T_2)$, для которой справедливы $\|b_e\|_{1;e} \lesssim 1$ и $\|b_e\|_{1;T_1 \cap T_2} \lesssim \alpha$, а также

$$\|\llbracket \nabla u_h \rrbracket\|_{\infty; \tilde{e}} = \|\llbracket \nabla u_h \rrbracket\|_{\infty; e} \lesssim \int_e \llbracket \nabla u_h \rrbracket b_e \, ds. \quad (1.38)$$

Воспользовавшись (1.34), проинтегрировав по частям и далее заметив, что $\varepsilon^2 \Delta u - f = 0$, получаем оценку

$$\begin{aligned} \frac{1}{2} \int_e \llbracket \nabla u_h \rrbracket b_e \, ds &= - \int_{T_1 \cup T_2} (u - u_h) \Delta b_e \, dx \\ &\quad + \varepsilon^{-2} \left(\int_{T_1 \cup T_2} (f - f_h) b_e \, dx - \int_{T_1 \cup T_2} (\varepsilon^2 \Delta u_h - f_h) b_e \, dx \right) \\ &\lesssim \|u - u_h\|_{\infty; T_1 \cup T_2} \|\Delta b_e\|_{1; T_1 \cup T_2} + \varepsilon^{-2} \|b_e\|_{1; T_1 \cup T_2} \|\varepsilon^2 \Delta u_h - f_h\|_{\infty; T_1 \cup T_2} \\ &\quad + \varepsilon^{-2} \min\{\|b_e\|_{1; T_1 \cup T_2} \|f - f_h\|_{\infty; T}, \sum_{i=1}^2 \|f - f_h\|_{-2, \infty, \varepsilon; T_i} \|b_e\|_{2, 1, \varepsilon; T_i}\} \\ &\lesssim \alpha^{-1} \|u - u_h\|_{\infty; T_1 \cup T_2} + \alpha \varepsilon^{-2} \|\varepsilon^2 \Delta u_h - f_h\|_{\infty; T_1 \cup T_2} + \\ &\quad + \varepsilon^{-2} \min\{\alpha \|f - f_h\|_{\infty; T}, \sum_{i=1}^2 \|f - f_h\|_{-2, \infty, \varepsilon; T_i} \|b_e\|_{2, 1, \varepsilon; T_i}\}. \end{aligned} \quad (1.39)$$

Несложно убедиться, что $\min\{\varepsilon, h_e \ell_h\} \|b_e\|_{2,1,\varepsilon;T_i} \lesssim \varepsilon^2 \ell_h$, следовательно

$$\begin{aligned} & \min\{\varepsilon, h_e \ell_h\} \|\llbracket \nabla u_h \rrbracket\|_{\infty;\tilde{e}} \lesssim \ell_h \|u - u_h\|_{\infty;\tilde{T}_1 \cup \tilde{T}_2} \\ & + \min\{1, \ell_h h_e^2 \varepsilon^{-2}\} \|\varepsilon^2 \Delta u_h - f_h\|_{\infty;\tilde{T}_1 \cup \tilde{T}_2} \\ & + \min\{\min\{1, \ell_h h_e^2 \varepsilon^{-2}\} \|f - f_h\|_{\infty;T_1 \cup T_2}, \ell_h \sum_{i=1}^2 \|f - f_h\|_{-2,\infty,\varepsilon;T_i}\}. \end{aligned} \quad (1.40)$$

Объединяя (1.40) с (1.37), получаем (1.32).

Наконец, оценим эффективность оценки для ошибки квадратуры. Отметим, что при $q \geq r - 1$ на любом элементе T справедливо $I_h^q \Delta u_h = I_h^{q-1} \Delta u_h = \Delta u_h$, и следовательно $f_h - I_h^j f_h = (Id - I_h^j)(f_h - \varepsilon^2 \Delta u_h)$ при $j = q, q - 1$, где Id является единичным оператором. Поскольку интерполяционный оператор Лагранжа I_h^j является L_∞ -стабильным, то при $q \geq r - 1$ получаем

$$\|\mu^q\|_{\infty;T} + \|\lambda \mu^{q-1}\|_{\infty;T} \lesssim \|f_h - \varepsilon^2 \Delta u_h\|_{\infty;T}. \quad (1.41)$$

Далее,

$$\begin{aligned} \varepsilon^{-2} \ell_h \|f_h - I_h^q f_h\|_{\frac{n}{2};T} & \lesssim h_T^2 \varepsilon^{-2} \ell_h \|f_h - I_h^q f_h\|_{\infty;T} \\ & \lesssim h_T^2 \varepsilon^{-2} \ell_h \|f_h - \varepsilon^2 \Delta u_h\|_{\infty;T}. \end{aligned} \quad (1.42)$$

Аналогично,

$$\varepsilon^{-1} \ell_h \|\lambda(f_h - I_h^q f_h)\|_{n;T} \lesssim h_T^2 \varepsilon^{-2} \ell_h \|f_h - \varepsilon^2 \Delta u_h\|_{\infty;T}. \quad (1.43)$$

Объединяя (1.35) и (1.36) с (1.41), а затем с (1.42) и (1.43), получаем соответственно (1.33a) и (1.33b), где мы также воспользовались соотношением $\|b_T\|_{2,1,\varepsilon;T} \lesssim 1 + \varepsilon^2 h_T^{-2}$. Оценка (1.33c) выводится аналогично. \square

1.3.5 Квазиоптимальный выбор разбиений сетки в оценках ошибки квадратуры

Обсудим практический выбор непересекающихся разбиений сетки $\mathcal{T}_1 \cup \mathcal{T}'_1$ и $\mathcal{T}_2 \cup \mathcal{T}'_2$, которые присутствуют в апостериорных оценках ошибки квадратуры

в Лемме 1.10. Веса в оценках ошибки квадратуры в (1.25) аналогичны весам в η_∞ , поэтому и имеет место оценка эффективности (1.33с). Однако, как уже отмечалось в [125], данная оценка эффективности для ошибки квадратуры не может быть использована для получения содержательной глобальной нижней оценки ошибки, поскольку ошибка квадратуры аккумулируется по сетке несколько иначе, чем невязки в оценке ошибки без учета квадратур. С другой стороны, здесь мы опишем алгоритм выбора непересекающихся разбиений сетки, которые являются квазиоптимальными в том смысле, что любой другой выбор \mathcal{T}_i и \mathcal{T}'_i не может улучшить получаемую верхнюю апостериорную оценку более, чем в 2 раза. Численные эксперименты также подтверждают, что оценка ошибки квадратуры при данном выборе $\mathcal{T}_i, \mathcal{T}'_i$ существенно ниже, чем при соответствующем выборе разбиений, приводящем к (1.25). Таким образом, оценка (1.25) не имеет большого практического смысла и приводится здесь, лишь поскольку для нее может быть получена в некотором смысле аналогичная оценка эффективности.

Итак, опишем алгоритм выбора непересекающихся разбиений сетки. Выберем $\mathcal{T}_1, \mathcal{T}'_1$ следующим образом ($\mathcal{T}_2, \mathcal{T}'_2$ выбираются аналогично). Пронумеруем элементы \mathcal{T} таким образом, что $\mu_{\mathcal{T}_1}^q \geq \mu_{\mathcal{T}_2}^q \geq \dots \geq \mu_{\mathcal{T}_N}^q$, где $N = \#\mathcal{T}$. Далее, пусть $\mathcal{T}_1 = \{T_i\}_{j \leq i \leq N}$ и $\mathcal{T}'_1 = \mathcal{T} \setminus \mathcal{T}_1$, где номер j является максимальным, при котором таким образом выбранные $\mathcal{T}_1, \mathcal{T}'_1$ удовлетворяют $\varepsilon^{-2} \ell_h \|\mu^q\|_{\frac{n}{2}; \mathcal{T}'_1} = \varepsilon^{-2} \ell_h (\sum_{i=1}^{j-1} |T_i| \mu_{T_i}^q)^{2/n} < \mu_{\mathcal{T}_{j-1}}^q$. Далее, пусть

$$\begin{aligned} \mu_{\mathcal{T}_1}^q &= \|\mu^q\|_{\infty; \mathcal{T}_1}, & \mu_{\mathcal{T}'_1}^q &= \varepsilon^{-2} \ell_h \|\mu^q\|_{\frac{n}{2}; \mathcal{T}'_1} \quad \text{где } \mathcal{T}_1, \mathcal{T}'_1 \text{ выбраны выше,} \\ \mu_{\mathcal{T}}^q &= \mu_{\mathcal{T}_1}^q + \mu_{\mathcal{T}'_1}^q. \end{aligned} \tag{1.44}$$

Вышеописанный выбор \mathcal{T}_1 и \mathcal{T}'_1 является квазиоптимальным в том смысле, что $\|\mu^q\|_{\infty; \mathcal{T}_1} + \varepsilon^{-2} \ell_h \|\mu^q\|_{\frac{n}{2}; \mathcal{T}'_1} \leq 2(\|\mu^q\|_{\infty; \tilde{\mathcal{T}}_1} + \varepsilon^{-2} \ell_h \|\mu^q\|_{\frac{n}{2}; \tilde{\mathcal{T}}'_1})$ для любого другого разбиения $\mathcal{T} = \tilde{\mathcal{T}}_1 \cup \tilde{\mathcal{T}}'_1$. Чтобы в этом убедиться, заметим, что поскольку μ^q аккумулируется на \mathcal{T}_1 в норме максимума модуля, то при оптимальном выборе \mathcal{T}_1 должно выполняться условие $\mu_{T_i}^q \in \mathcal{T}_1 \Rightarrow \mu_{T_l}^q \in \mathcal{T}_1$ при $l \geq i$. Далее, при вышеописанном выборе j , рассмотрим сначала $k < j$, для которого

получаем

$$\begin{aligned} \max_{j \leq i \leq N} \mu_{T_i}^q + \varepsilon^{-2} \ell_h \left(\sum_{i=1}^{j-1} |T| \mu_{T_i}^{q \ n/2} \right)^{2/n} &\leq \mu_{T_j}^q + \mu_{T_{j-1}}^q \leq 2\mu_{T_k}^q \\ &\leq 2 \left[\mu_{T_{j-1}}^q + \varepsilon^{-2} \ell_h \left(\sum_{i=1}^{k-1} |T| \mu_{T_i}^{q \ n/2} \right)^{2/n} \right]. \end{aligned} \quad (1.45)$$

Если же $k > j$, несложно убедиться, что

$$\begin{aligned} \max_{j \leq i \leq N} \mu_{T_i}^q + \varepsilon^{-2} \ell_h \left(\sum_{i=1}^{j-1} |T| \mu_{T_i}^{q \ n/2} \right)^{2/n} &\leq \varepsilon^{-2} \ell_h \left[\left(\sum_{i=1}^j |T| \mu_{T_i}^{q \ n/2} \right)^{2/n} + \left(\sum_{i=1}^{j-1} |T| \mu_{T_i}^{q \ n/2} \right)^{2/n} \right] \\ &\leq 2\varepsilon^{-2} \ell_h \left(\sum_{i=1}^{k-1} |T| \mu_{T_i}^{q \ n/2} \right)^{2/n} \\ &\leq 2 \left[\max_{k \leq i \leq N} \mu_{T_i}^q + \varepsilon^{-2} \ell_h \left(\sum_{i=1}^{k-1} |T| \mu_{T_i}^{q \ n/2} \right)^{2/n} \right]. \end{aligned} \quad (1.46)$$

Таким образом, наше утверждение о квазиоптимальности вышеописанного выбора разбиений сетки доказано.

1.3.6 Результаты параграфа и дискуссия

Введем глобальный оценитель ошибки на основе невязок

$$\eta_{\mathcal{T}}^{\infty} = \max_{T \in \mathcal{T}} \eta_{\infty}(T). \quad (1.47)$$

Также основные обозначения описаны в Таблице 1.1.

Объединяя полученные выше результаты, получаем следующую теорему.

Т е о р е м а 1.12. *Для произвольных непересекающихся разбиений $\mathcal{T} = \mathcal{T}_1 \cup \mathcal{T}_2$ и $\mathcal{T} = \mathcal{T}'_1 \cup \mathcal{T}'_2$, справедлива оценка*

$$\|u - u_h\|_{\infty; \Omega} \lesssim \eta_{\mathcal{T}}^{\infty} + \mu_{\text{quad}}. \quad (1.48)$$

Дополнительно, $\mathcal{T}_i, \mathcal{T}'_i$, $i = 1, 2$, могут быть выбраны таким образом, что

$$\|u - u_h\|_{\infty; \Omega} \lesssim \eta_{\mathcal{T}}^{\infty} + \mu_{\Sigma}^q + \mu_{\Sigma}^{q-1}. \quad (1.49)$$

Альтернативно, при квазиоптимальном выборе $\mathcal{T}_i, \mathcal{T}'_i$, описанном в §1.3.5, получаем

$$\|u - u_h\|_{\infty; \Omega} \lesssim \eta_{\mathcal{T}}^{\infty} + \mu_{\mathcal{T}}^q + \mu_{\mathcal{T}}^{q-1}, \quad (1.50)$$

при этом любой другой выбор $\mathcal{T}_i, \mathcal{T}'_i$ не улучшает оценку квадратурной ошибки более, чем в 2 раза.

Для $T \in \mathcal{T}$ справедлива также оценка эффективности

$$\begin{aligned} \eta_{\infty}(T) &\lesssim \ell_h \|u - u_h\|_{\infty; \omega_T} + \text{osc}(T) \\ &+ \min\{\min\{1, \ell_h h_T^2 \varepsilon^{-2}\} \|f - f_h\|_{\infty; \omega_T}, \ell_h \|f - f_h\|_{-2, \infty, \varepsilon; \omega_T}\}. \end{aligned} \quad (1.51)$$

Дополнительно, при $q \geq r - 1$ имеем

$$\begin{aligned} \|\mu^q\|_{\infty; T} + \|\lambda \mu^{q-1}\|_{\infty; T} &\lesssim \varepsilon^2 h_T^{-2} \|u - u_h\|_{\infty; T} + \|f_h - f_{h,T}\|_{\infty; T} \\ &+ \min\{\|f - f_h\|_{\infty; T}, (1 + \varepsilon^2 h_T^{-2}) \|f - f_h\|_{-2, \infty, \varepsilon; T}\}, \end{aligned} \quad (1.52a)$$

$$\begin{aligned} \varepsilon^{-2} \ell_h \|\mu^q\|_{\frac{n}{2}; T} + \varepsilon^{-1} \ell_h \|\lambda \mu^{q-1}\|_{n; T} &\lesssim \ell_h \|u - u_h\|_{\infty; T} \\ &+ \ell_h h_T^2 \varepsilon^{-2} \|f_h - f_{h,T}\|_{\infty; T} \\ &+ \min\{h_T^2 \varepsilon^{-2} \ell_h \|f - f_h\|_{\infty; T}, \ell_h (1 + h_T^2 \varepsilon^{-2}) \|f - f_h\|_{-2, \infty, \varepsilon; T}\}, \end{aligned} \quad (1.52b)$$

$$\begin{aligned} &\| \min\{h_T^{-2}, \varepsilon^{-2} \ell_h\} \mu^q \|_{\frac{n}{2}; T} + \| \min\{h_T^{-1}, h_T \varepsilon^{-2} \ell_h\} \mu^{q-1} \|_{n; T} \\ &\lesssim \ell_h \|u - u_h\|_{\infty; T} + \text{osc}(T) \\ &+ \min\{\min\{1, \ell_h h_T^2 \varepsilon^{-2}\} \|f - f_h\|_{\infty; T}, \ell_h \|f - f_h\|_{-2, \infty, \varepsilon; T}\}. \end{aligned} \quad (1.52c)$$

В контексте Теоремы 1.12, обсудим отношение между невязкой и ошибкой решения. Невязка \mathcal{R}_h определяется соотношением

$$\langle \mathcal{R}_h, v \rangle = \varepsilon^2 \int_{\Omega} \nabla(u - u_h) \nabla v \, dx + \int_{\Omega} [f(x, u) - f(x, u_h)] v \, dx. \quad (1.53)$$

Лемма 1.8 может быть переформулирована как $\|u - u_h\|_{\infty;\Omega} \lesssim |\langle \mathcal{R}_h, G \rangle|$, а Лемма 1.9 и Лемма 1.10 вместе дают вычисляемую оценку для $\langle \mathcal{R}_h, G \rangle$. Обычно при получении апостериорных оценок на основе невязок ошибка оценивается через норму в сопряженном пространстве Соболева, например, через $\|\mathcal{R}_h\|_{H^{-1}(\Omega)}$ при оценке ошибки в норме $H^1(\Omega)$. При получении же оценок ошибки в норме максимума модуля ситуация несколько более сложная. В [125] удается ассоциировать оценки ошибки конечноэлементных решений уравнения $-\Delta u + f(x, u) = 0$ в норме максимума модуля с оценками для $\|\mathcal{R}_h\|_{-2,\infty,1;\Omega}$ благодаря использованию регуляризованной функции Грана, которая лежит в пространстве $W_1^2(\Omega)$. Однако данный метод применим лишь к липшицевым областям, а также приводит к несколько менее точным апостериорным оценкам (в смысле степеней логарифмов). Поскольку, как и [74], мы используем классическую функцию Грина, отношение между ошибкой и \mathcal{R}_h становится несколько более сложным.

1.4 Необходимость логарифма в апостериорных оценках в норме максимума модуля

Убедимся, что логарифмический коэффициент необходимо присутствует в апостериорной оценке (1.3). Для этого используется идея из [81] и доказываются верхние априорные оценки [157, Лемма 10], а также верхние и нижние апостериорные оценки [157, Лемма 11] для $u - u_h$ в модифицированной норме ВМО $\|\cdot\|_{\text{bmo}_\Delta(\Omega)}$ [79, 157] при условии, что Ω — выпуклая полигональная область. Данные оценки по существу совпадают со стандартными оценками в норме L_∞ , с той лишь разницей, что в них отсутствует логарифмический коэффициент. Объединяя описанные результаты для $\|u - u_h\|_{\text{bmo}_\Delta(\Omega)}$, немедленно получаем следующую лемму.

Л е м м а 1.13 ([157, Леммы 10 и 11]). Пусть $\Omega \subset \mathbb{R}^n$ является выпуклой полигональной областью, а $u \in W_\infty^2(\Omega)$ удовлетворяет $-\Delta u = f$. Также

Символ	Определение	Применение
$\ell_{h,x}, \ell_h$	(1.21), (1.29)	Логарифмические коэффициенты
μ^j, μ_T^j	Лемма 1.10	Поэлементный индикатор ошибки квадратуры
λ	Лемма 1.10	Вес квадратурного индикатора
μ_{quad}	(1.24)	Глобальный квадратурный оценщик, произвольное разбиение
μ_{Σ}^q	(1.25)	Неоптимальный квадратурный оценщик
$\eta_{\infty}(T)$	(1.29)	Поэлементный индикатор ошибки без учета квадратур
$f_h, f_{h,T}$	(1.30)	$f(\cdot, u_h)$ и ее проекция на \mathbb{P}_{T-1}
$\text{osc}(T)$	(1.30)	Осцилляция данных
$\mu_{\mathcal{T}_1}^q, \mu_{\mathcal{T}'_1}^q$	(1.44)	Квадратурный оценщик, квазиоптимальное разбиение
$\mu_{\mathcal{T}}^q$	(1.44)	Глобальный квадратурный оценщик, квазиоптимальное разбиение
$\eta_{\mathcal{T}}^{\infty}$	(1.47)	Глобальный оценщик без учета квадратур

Таблица 1.1: Основные обозначения

пусть u_h является аппроксимацией u , полученной с помощью линейных конечных элементов на квазиравномерной сетке диаметра h . Тогда

$$\max_{T \in \mathcal{T}_h} [h_T^2 \|f + \Delta u_h\|_{\infty; T} + h_T \|\llbracket \nabla u_h \rrbracket\|_{\infty; \partial T}] \lesssim h^2 |u|_{2, \infty; \Omega}. \quad (1.54)$$

Объединяя (1.54) с (1.3), заключаем, что

$$\begin{aligned} \|u - u_h\|_{\infty; \Omega} &\lesssim \log h^{-1} \max_{T \in \mathcal{T}_h} [h_T^2 \|f + \Delta u_h\|_{\infty; T} + h_T \|\llbracket \nabla u_h \rrbracket\|_{\infty; \partial T}] \\ &\lesssim \log h^{-1} h^2 |u|_{2, \infty; \Omega}. \end{aligned}$$

Остается воспользоваться контрпримерами, описанными в [96, 61, 1] при $n = 2$ и [172] при $n = 3$, показывающими, что логарифм необходимо присутствует в стандартной априорной оценке $\|u - u_h\|_{\infty; \Omega} \lesssim \log h^{-1} h^2 |u|_{2, \infty; \Omega}$ для

линейных конечных элементов. В процитированных работах показано, что при определенном выборе квазиравномерных сеток (и определенном выборе u в [172]) справедливо соотношение $\|u - u_h\|_{\infty; \Omega} \gtrsim \log(h^{-1})h^2|u|_{2, \infty; \Omega}$

Таким образом, приходим к следующему результату.

Л е м м а 1.14. При условиях Леммы 1.13, в оценке

$$\|u - u_h\|_{\infty; \Omega} \lesssim \log h^{-1} \max_{T \in \mathcal{T}_h} [h_T^2 \|f + \Delta u_h\|_{\infty; T} + h_T \|[\nabla u_h]\|_{\infty; \partial T}]$$

необходимо присутствует логарифмический коэффициент $\log h^{-1}$.

1.5 Численные результаты

Все численные эксперименты были выполнены с использованием пакета MATLAB и библиотеки iFEM [65]. Во всех тестах использовались линейные конечные элементы и стандартная адаптация сетки. При численном решении нелинейных задач применялся "damped" метод Ньютона. С использованием обозначений Таблицы 1.1 полученный нами глобальный оценщик ошибки принимает форму $\eta = \eta_{\mathcal{T}}^{\infty} + \mu_{\mathcal{T}_1}^q + \mu_{\mathcal{T}'_1}^q + \mu_{\mathcal{T}_2}^{q-1} + \mu_{\mathcal{T}'_2}^{q-1}$. Здесь η представлена в виде суммы пяти компонент, которые несколько по-разному аккумулируются по сетке, поэтому общая стратегия маркировки элементов на основе единого поэлементного индикатора ошибки неочевидна. Вместо этого, на каждой итерации элементы маркировались для каждой компоненты оценщика отдельно при условии, что рассматриваемая компонента отвечала по крайней мере за 10% глобального оценщика η . Для сравнения, мы использовали аналогичную стратегию с оценщиком из [125], только вместо 5 компонент последний оценщик включал 3. Во всех экспериментах применялась стандартная квадратурная формула Гаусса степени $q = 3$ (т. е. использовались квадратурные точки с барицентрическими координатами $(1/3, 1/3, 1/3)$, $(0.6, 0.2, 0.2)$, $(0.2, 0.6, 0.2)$ и $(0.2, 0.2, 0.6)$ и соответствующими весами $-27/48$, $25/48$, $25/48$ и $25/48$), которая очевидно удовлетворяет предположениям, описанным в §1.3.3.

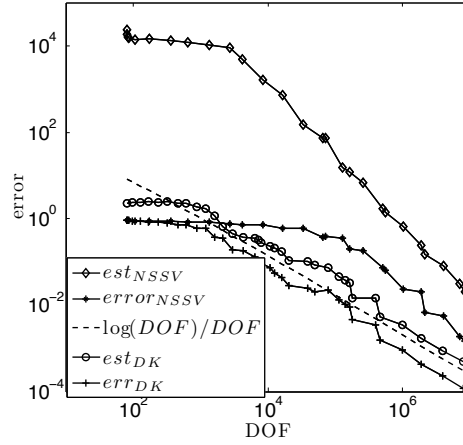


Рис. 1.1: Эксперимент 1: сравнение убывания ошибки в норме максимума модуля и соответствующего оценителя при адаптивном измельчении сетки с использованием полученного нами оценителя (помеченного индексом DK) и оценителя из [125] (помеченного индексом NSSV); DOF обозначает число степеней свободы.

1.5.1 Эксперимент 1: Преимущества ε -робастных оценителей

Чтобы проиллюстрировать преимущества ε -робастных оценителей, пусть Ω будет единичным квадратом, а

$$u_1(x, y) = \left[\cos(\pi x/2) - \frac{e^{-x/\varepsilon} - e^{-1/\varepsilon}}{1 - e^{-\varepsilon}} \right] \left[1 - y - \frac{e^{-y/\varepsilon} - e^{-1/\varepsilon}}{1 - e^{-1/\varepsilon}} \right]. \quad (1.55)$$

Функция u_1 имеет пограничные слои вдоль частей границы Ω , совпадающих с координатными осями. Введем также $u_2(x, y) = 0.01 \sin(100\pi x) \sin(100\pi y)$, а затем точное решение $u = u_1 + u_2$, и будем численно решать уравнение $-\varepsilon^2 \Delta u + u - g = 0$ с соответствующей g . Будет рассмотрен случай $\varepsilon^2 = 10^{-6}$. На Рис. 1.1 показано, как убывает ошибка и соответствующий оценитель при маркировке с использованием неробастного оценителя (1.4) из [125], а затем ε -робастного оценителя, полученного в (1.48). Отметим, что соответствующие оценители квадратурных ошибок не играли большой роли при маркировке и адаптивном измельчении сетки.

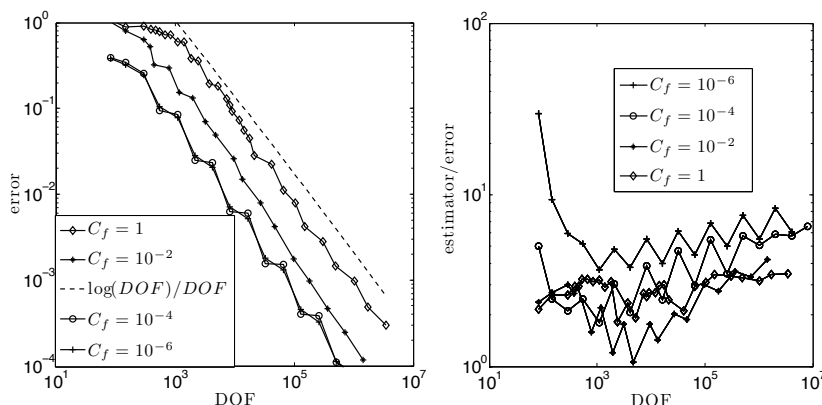


Рис. 1.2: Эксперимент 2: сравнение убывания ошибки в норме максимума модуля при $\varepsilon^2 = 10^{-6}/C_f$ и различных C_f (слева); соответствующие индексы эффективности (справа).

Рис. 1.1 показывает, что неробастный оценщик переоценивает ошибку приблизительно в 10^4 раз на ранних этапах вычислений, причем это переоценивание зависит от ε и становится более сильным при $\varepsilon \rightarrow 0$. Также отметим, что при адаптивных вычислениях с использованием неробастного оценщика ошибка убывает значительно медленнее, чем при использовании робастного оценщика. Это объясняется тем, что на начальном этапе оценщик из (1.4) приводит к слишком сильному измельчению сетки вне пограничных слоев.

1.5.2 Эксперимент 2: Робастность по ε

С целью проиллюстрировать робастность полученных апостериорных оценок по ε рассмотрим тестовую задачу из §1.5.1 с точным решением $u = u_1 + u_3$, где u_1 определено в (1.55), а $u_3(x, y) = 2 \sin(4\pi x) \sin(4\pi y)$, и соответствующей g . Рассмотрим $\varepsilon^2 = 10^{-6}/C_f$ с $C_f = 1, 10^{-2}, 10^{-4}, 10^{-6}$. Для указанных значений C_f на Рис. 1.2 показаны значения $\|u - u_h\|_{\infty; \Omega}$ по отношению к соответствующим числам степеней свободы (обозначенным DOF). Также приведены значения индексов эффективности, получаемых по формуле $\eta/\|u - u_h\|_{\infty; \Omega}$.

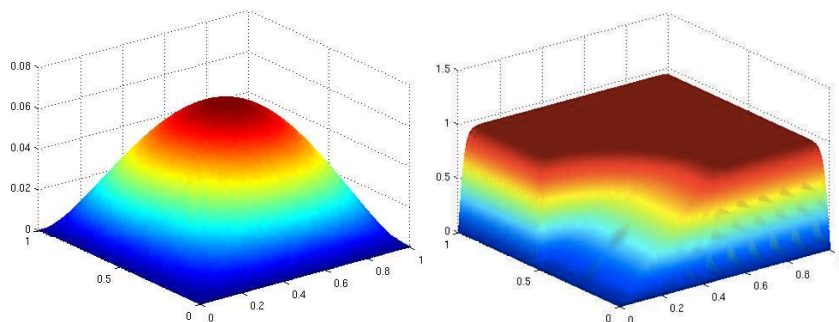


Рис. 1.3: Эксперимент 3: Адаптивные численные решения при $\varepsilon^2 = 1$ и $DOF=4536$ (слева), $\varepsilon^2 = 10^{-4}$ и $DOF=4236$ (справа).

Полученные результаты показывают, что адаптивные алгоритмы с использованием полученных индикаторов ошибки эффективно работают при различных значениях ε , а индексы эффективности стабилизируются при $\varepsilon \rightarrow 0$.

1.5.3 Эксперимент 3: Индикаторы квадратурных ошибок

Рассмотрим тестовую задачу $-\varepsilon^2 \Delta u + u = g$ в $\Omega = (0, 1) \times (0, 1)$, с правой частью $g(x, y) = 2x$ при $x^2 + y^2 < 1/4$ и $g(x, y) = 1$ в противном случае, т. е. g имеет разрыв вдоль кривой $x^2 + y^2 = 1/4$ кроме точки $(x, y) = (1/2, 0)$. Точное решение u неизвестно, но при $\varepsilon \ll 1$ оно имеет внутренний слой вдоль $x^2 + y^2 = 1/4$, а также пограничные слои вдоль границы области, что подтверждается численными решениями при $\varepsilon^2 = 1$ и $\varepsilon^2 = 10^{-4}$, показанными на Рис. 1.3.

Заметим, что некоторые треугольные элементы пересекают кривую $x^2 + y^2 = 1/4$, где f разрывна. Следовательно $\|\mu^q\|_{\infty; \mathcal{T}}$ равномерно отделена от 0, поскольку разрывная g аппроксимируется в L_∞ непрерывными функциями. Заметим, что на треугольниках, непересекающихся с $x^2 + y^2 = 1/4$ квадратурная ошибка и квадратурные индикаторы равны 0. На Рис. 1.4 показано убывание различных оценщиков при $\varepsilon^2 = 10^{-4}$. Слева показано поведение оценщика без учета квадратур $\eta_{\mathcal{T}}^\infty$, а также компонент квадратурного оценщика $\mu_{\mathcal{T}}^q$ и $\mu_{\mathcal{T}}^{q-1}$. Заметим, что $\mu_{\mathcal{T}}^q$ играет доминирующую

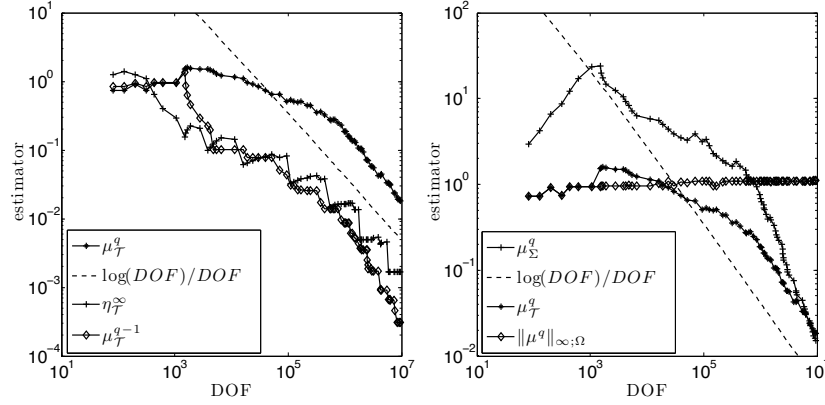


Рис. 1.4: Эксперимент 3: убывание квадратурных оценщиков и оценщика без учета квадратур в случае разрывной g при $\varepsilon^2 = 10^{-4}$ (слева); сравнение различных квадратурных оценщиков (справа).

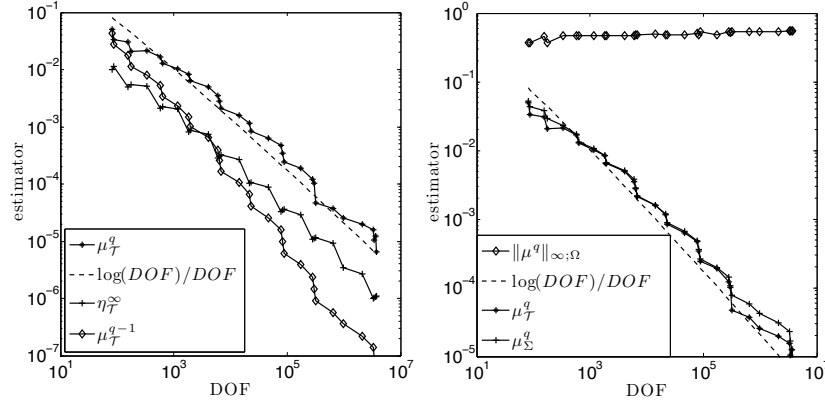


Рис. 1.5: Эксперимент 3: убывание квадратурных оценщиков и оценщика без учета квадратур в случае разрывной g при $\varepsilon^2 = 1$ (слева); сравнение различных квадратурных оценщиков (справа).

роль в глобальном оценщике и, следовательно, в адаптивном измельчении сетки. Справа мы сравниваем поведение различных квадратурных оценщиков. Здесь $\mu_\Sigma^q = \|\min\{h_T^{-2}, \varepsilon^{-2}\ell_h\}\mu^q\|_{\frac{n}{2};\mathcal{T}}$ из (1.25). Заметим, что сначала $\mu_\Sigma^q = \|\mu^q\|_{\infty;\Omega}$, т. е. $\mathcal{T}_1 = \mathcal{T}$ в определении μ_Σ^q . Далее, при добавлении элементов к \mathcal{T}'_1 , сначала $\|\mu^q\|_{\infty;\Omega} < \mu_\Sigma^q \leq 2\|\mu^q\|_{\infty;\Omega}$, а затем, при приблизительно

$10^5 < DOF < 10^6$, квадратурный оценитель $\mu_{\mathcal{T}}^q$ становится меньше, чем любой из μ_{Σ}^q и $\|\mu^q\|_{\infty;\Omega}$. Асимптотически же $\mu_{\mathcal{T}}^q = \mu_{\Sigma}^q$, т. е. $\mathcal{T}' = \mathcal{T}$. Соответствующие графики для $\varepsilon^2 = 1$ представлены на Рис. 1.5. Объединяя результаты для этих двух случаев, можно заключить, что используемый нами квадратурный оценитель с разбиением сетки является удобным в использовании и робастным.

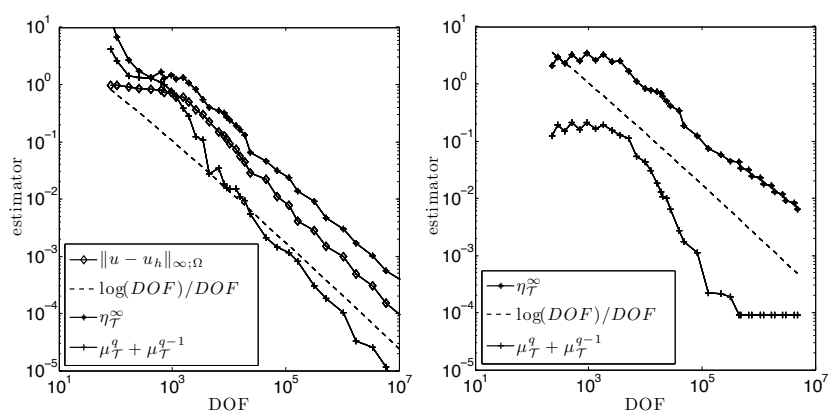


Рис. 1.6: Эксперимент 4: убывание ошибки решения и оценителей в случае известного точного решения и в квадратной области Ω (слева); компоненты оценителя в случае неизвестного и в L-образной области Ω (справа).

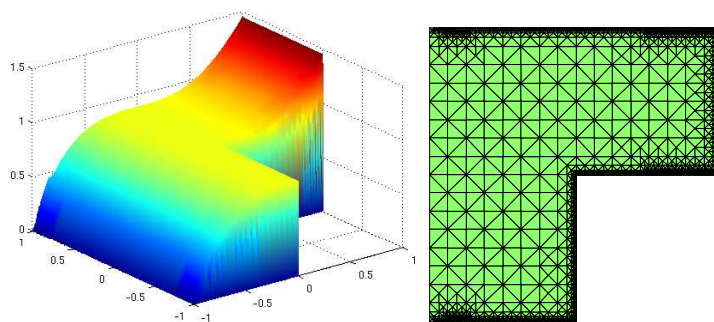


Рис. 1.7: Эксперимент 4: адаптивное численное решение (слева); адаптивно сгенерированная сетка с $DOF=13787$ (справа).

1.5.4 Эксперимент 4: Нелинейность типа Пуассона-Больцмана

Сингулярно возмущенные уравнения типа Пуассона-Больцмана рассматривались в ряде публикаций; см., например, [87]. Как простой прототип, рассмотрим уравнение $-\varepsilon^2 \Delta u + \sinh u = g(x, y)$ с $\varepsilon^2 = 10^{-6}$. Сначала выберем Ω единичным квадратом, а точное решение $u = u_1 + u_3$ из §1.5.2. Наш адаптивный метод достаточно хорошо работает для этого примера, что демонстрирует Рис. 1.6 (слева). Далее рассмотрим L-образную область Ω , решения в которой имеют угловую сингулярность, и положим $g(x, y) = 1 + x^3$. Убывание оценителя для этого случая показано на Рис. 1.6 (справа), а численное решение и адаптивно сгенерированная сетка представлены на Рис. 1.7.

§2 Апостериорные оценки в норме максимума модуля на анизотропных сетках

Для сингулярно возмущенных полулинейных уравнений реакции-диффузии в многоугольных областях получены явные апостериорные оценки ошибки численного решения в норме максимума модуля. По сравнению с §1, рассматривается существенно более сложный случай неструктурированных анизотропных сеток (т. е. элементы рассматриваемых здесь сеток могут быть сколь угодно сильно сплющены). Постоянные в полученных оценках не зависят от диаметров элементов сетки, их аспектного соотношения (т. е. степени их сплющенности) и малого параметра. Основной материал параграфа опубликован в работе [168].

2.1 Введение

В настоящем параграфе рассматриваются конечноэлементные аппроксимации для сингулярно возмущенного уравнения реакции-диффузии

$$Lu := -\varepsilon^2 \Delta u + f(x, y; u) = 0 \quad \text{при } (x, y) \in \Omega, \quad u = 0 \quad \text{на } \partial\Omega, \quad (2.1)$$

в многоугольной, возможно нелипшицевой, области $\Omega \subset \mathbb{R}^2$. Предполагается, что $0 < \varepsilon \leq 1$, функция f непрерывна в $\Omega \times \mathbb{R}$ и удовлетворяет условию $f(\cdot; s) \in L_\infty(\Omega)$ при любом $s \in \mathbb{R}$, а также одностороннему условию Липшица $f(x, y; u) - f(x, y; v) \geq C_f[u - v]$ при всех $u \geq v$ с некоторой постоянной $C_f \geq 0$. Отметим, что эта задача имеет единственное решение $u \in W_\ell^2(\Omega) \subseteq W_q^1 \subset C(\bar{\Omega})$ для некоторых $\ell > 1$ и $q > 2$ (см Лемму 1.1 выше). Также предполагается, что $C_f + \varepsilon^2 \geq 1$ (поскольку деление на $C_f + \varepsilon^2$ немедленно сводит (2.1) к данному случаю).

Апостериорные оценки ошибки на основе невязок численного решения в норме максимума модуля были рассмотрены в §1 для случая локально квазиравномерных триангуляций. В настоящем параграфе, а также в §3, мы

ограничим рассмотрение линейными конечными элементами в двумерной области, при этом наше внимание будет перенесено на существенно более сложный случай неструктурированных анизотропных сеток (т. е. элементы рассматриваемых здесь сеток могут быть сильно сплющены). Рис. 2.1 иллюстрирует типы узлов сетки, к которым применим наш анализ.

Даже для линейного уравнения Лапласа (получаемого из (2.1) при $\varepsilon = 1$ и $f_u = 0$) нам неизвестны какие-либо апостериорные оценки в норме максимума модуля при отсутствии предположения ограниченности аспектного соотношения элементов сетки (например, в работах [82, 123, 71, 125] предполагается локальная квазиравномерность сетки). Еще больший интерес анизотропные сетки представляют в контексте сингулярно возмущенных дифференциальных уравнений (таких, как (2.1) при $\varepsilon \ll 1$). Для таких уравнений норма максимума модуля является достаточно сильной (в отличие, например, от нормы L_2) для того, чтобы обнаружить большую ошибку в зоне узких пограничных и внутренних слоев.

Отметим также, что для ряда структурированных локально анизотропных сеток (т. е. анизотропных в областях пограничных и внутренних слоев, и стандартных за их пределами) было теоретически установлено, что они позволяют получать надежные численные решения при использовании относительно небольших чисел степеней свободы (последние обычно не зависят от значений малых параметров); см., например, работы [66, 163, 178, 132], а также приведенные в них ссылки. Отметим также, что такие сетки, как правило, конструируются априори, либо же генерируются адаптивно с использованием эвристических методов.

В настоящем параграфе для уравнения (2.1) будут рассмотрены стандартные линейные конечные элементы. Пусть $S_h \subset H_0^1(\Omega) \cap C(\bar{\Omega})$ является пространством непрерывных кусочно-линейных функций, ассоциируемым с триангуляцией \mathcal{T} , а численное решение $u_h \in S_h$ пусть удовлетворяет соотно-

шению

$$\varepsilon^2 \langle \nabla u_h, \nabla v_h \rangle + \langle f_h^I, v_h \rangle = 0 \quad \forall v_h \in S_h, \quad f_h(\cdot) := f(\cdot; u_h). \quad (2.2)$$

Здесь $\langle \cdot, \cdot \rangle$ обозначает скалярное произведение в $L_2(\Omega)$, а f_h^I — стандартный кусочно-линейный интерполянт Лагранжа для f_h .

В рамках предварительного описания основных результатов параграфа, предположим, что углы всех анизотропных элементов близки к нетупым. Тогда наша первая оценка принимает вид

$$\begin{aligned} \|u - u_h\|_{\infty; \Omega} \leq C \ell_h \max_{z \in \mathcal{N}} \left(\min\{\varepsilon, H_z\} \|J_z\|_{\infty; \gamma_z} + \min\{\varepsilon^2, H_z^2\} \|\varepsilon^{-2} f_h^I\|_{\infty; \omega_z} \right) \\ + C \|f_h - f_h^I\|_{\infty; \Omega}, \end{aligned} \quad (2.3)$$

где постоянная C не зависит от диаметров элементов сетки и их аспектных соотношений (т. е. степени их сплюсненности), а также от малого параметра ε . (Данная оценка получается объединением (2.44), (2.47), (2.21) с Леммой 2.17.) Здесь \mathcal{N} является множеством вершин элементов \mathcal{T} , J_z обозначает стандартный скачок нормальной производной u_h через границу элемента, ω_z — множество элементов, окружающих вершину $z \in \mathcal{N}$, а γ_z — множество ребер внутри ω_z , $H_z = \text{diam}(\omega_z)$, $\ell_h = \ln(2 + \varepsilon \underline{h}^{-1})$, а \underline{h} — минимальный диаметр окружностей, вписанных в треугольники в \mathcal{T} .

Заметим, что при $\varepsilon = 1$ соотношение (2.3) дает стандартную апостериорную оценку, аналогичную результатам [82, 123, 125], с той существенной

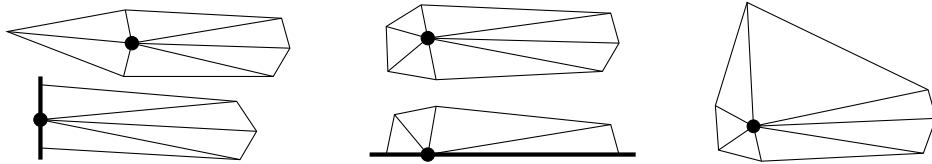


Рис. 2.1: Примеры анизотропных узлов $z \in \mathcal{N}_{\text{ani}}$ (слева), полу-анизотропных узлов $z \in \mathcal{N}_{\text{s,ani}}$ (центр), изотропного узла $z \in \mathcal{N}_{\text{iso}}$ (справа) и узла $z \in \mathcal{N}_{\text{ani}} \cap \mathcal{N}_{\partial\Omega}^*$ (слева снизу).

разницей, что здесь она будет установлена для случая анизотропных сеток. Отметим также, что оценка (2.3) практически идентична с апостериорными оценками [157] (см. также (1.3) в §1), полученными для сингулярно возмущенного случая $\varepsilon \in (0, 1]$, однако здесь мы отказываемся от предположения локальной квазиравномерности триангуляции.

При внимательном рассмотрении стандартных доказательств для случая локально квазиравномерных сеток обнаруживается, что одним из серьезных препятствий при их обобщении на случай анизотропных сеток является применение теоремы о следах при оценке невязок через границы элементов. При этом в получающихся апостериорных оценках появляются коэффициенты, зависящие от аспектных соотношений элементов сетки, поэтому такие оценки чрезвычайно грубы и не могут быть использованы для практической оценки ошибки численного решения. Замечание 2.6 проливает некоторый свет на наш подход к преодолению этого теоретического препятствия.

Следует заметить, что, в отличие от других слагаемых в (2.3), поэлементные невязки $\|\varepsilon^{-2} f_h^I\|_{\infty; \omega_z}$ не отражают анизотропность соответствующих элементов сетки. С тем, чтобы получить более точную (и, в некотором смысле, анизотропную) оценку для поэлементных невязок, будут введены в рассмотрение цепочки коротких ребер, соединяющих анизотропные узлы (см. Рис. 2.4, справа). При некоторых дополнительных ограничениях на такие цепочки будет получена более точная оценка

$$\begin{aligned} \|u - u_h\|_{\infty; \Omega} &\leq C \ell_h \left[\max_{z \in \mathcal{N}} \left(\min\{\varepsilon, H_z\} \|J_z\|_{\infty; \gamma_z} \right) \right. \\ &\quad \left. + \max_{z \in \mathcal{N} \setminus \mathcal{N}_{\text{paths}}} \left(\min\{1, \varepsilon^{-2} H_z^2\} \|f_h^I\|_{\infty; \omega_z} \right) \right] \\ &\quad + \max_{z \in \mathcal{N}_{\text{paths}}} \left(\min\{\varepsilon, H_z\} \min\{\varepsilon, h_z\} \|\varepsilon^{-2} f_h^I\|_{\infty; \omega_z} + \min\{1, \varepsilon^{-2} H_z^2\} \text{osc}(f_h^I; \omega_z) \right) \\ &\quad + C \|f_h - f_h^I\|_{\infty; \Omega}, \end{aligned} \tag{2.4}$$

(которая получается объединением (2.44), (2.58), (2.21) с Леммой 2.17). Здесь $\mathcal{N}_{\text{paths}}$ обозначает множество узлов, которые задействованы хотя бы в одной

цепочке, а $h_z \simeq H_z^{-1}|\omega_z|$. Поскольку для анизотропных узлов справедливо $h_z \ll H_z$, то очевидно, что оценка (2.4) является более точной по сравнению с (2.3). (Последнее также подтверждается результатами численных экспериментов в §2.5.4).

Отметим, что полезность результатов данной главы не исчерпывается стационарным случаем, поскольку они могут быть использованы как ингредиенты в апостериорных оценках для соответствующих параболических уравнений [175] (см. также Гл. 2).

Ранее апостериорные оценки для задач типа (2.1) на анизотропных сетках были получены в работах [156, 165, 105, 107]. В [165, 156] ошибка численного решения оценивается в норме максимума модуля, но рассмотрены только прямоугольные анизотропные сетки, в то время как в [107, 105] рассмотрены более общие анизотропные сетки, но ошибка оценивается в более слабой энергетической норме. Более подробно работы [107, 105] будут прокомментированы в следующем §3.

Параграф организован следующим образом. В §2.2 описываются основные предположения о триангуляции \mathcal{T} , а также допустимые типы анизотропных узлов. Для последних в §2.3 получена наилучшая версия теоремы о следах. В §2.4 ошибка численного решения представлена через функцию Грина линейризованной задачи. Далее в §2.5 упрощенная версия представленного здесь апостериорного анализа представлена для случая частично структурированных анизотропных сеток, в то время как в §2.6 апостериорные оценки выводятся для случая более общих анизотропных сеток. Параграф завершается получением оценок ошибки интерполяции для функции Грина в §2.7, а также заключительными замечаниями в §2.8.

Обозначения. Будем писать $a \simeq b$, если $a \lesssim b$ и $a \gtrsim b$, в то время как $a \lesssim b$ будет обозначать $a \leq Cb$ с некоторой постоянной C , зависящей от Ω и f . Отметим, что C не зависит от диаметров и аспектных соотношений элементов \mathcal{T} , а также от ε . Также, для $\mathcal{D} \subseteq \bar{\Omega}$, $1 \leq p \leq \infty$ и $k \geq 0$, будут использо-

ваться нормы $\|v\|_{p;\mathcal{D}} = \|v\|_{L_p(\mathcal{D})}$ и $|v|_{k,p;\mathcal{D}} = |v|_{W_p^k(\mathcal{D})}$, где $|\cdot|_{W_p^k(\mathcal{D})}$ обозначает стандартную полунорму в пространстве Соболева функций, интегрируемых с p -ой степенью их модуля вместе со своими обобщенными производными до порядка k .

2.2 Основные предположения о триангуляции

Будем использовать обозначения $z = (x_z, y_z)$, S и T соответственно для узлов, ребер и элементов сетки, в то время как \mathcal{N} , \mathcal{S} и \mathcal{T} обозначают их соответствующие множества. Для любого $T \in \mathcal{T}$ пусть H_T обозначает максимальную длину ребра, а $h_T := 2H_T^{-1}|T|$ — минимальную высоту в треугольнике T . Для любого $z \in \mathcal{N}$ обозначение ω_z будет использоваться для области, состоящей из элементов, окружающих z , а \mathcal{S}_z — для множества ребер, исходящих из z , а также

$$H_z := \text{diam}(\omega_z), \quad h_z := \max_{T \subset \omega_z} h_T, \quad \gamma_z := \mathcal{S}_z \setminus \partial\Omega, \quad \dot{\gamma}_z := \{S \subset \gamma_z : |S| \lesssim h_z\}. \quad (2.5)$$

В дальнейшем всегда будет предполагаться, что триангуляция удовлетворяет следующим условиям.

- *Максимальный угол.* Максимальный внутренний угол в любом треугольнике $T \in \mathcal{T}$ является равномерно ограниченным некоторой постоянной $\alpha_0 < \pi$.
- *Локальная ориентация элементов.* Для любого узла $z \in \mathcal{N}$, найдется такой прямоугольник $R_z \supset \omega_z$, что $|R_z| \simeq |\omega_z|$.
- Число треугольников, содержащих любой узел сетки, является равномерно ограниченным.

Заметим, что вышеописанные условия автоматически выполняются на локально квазиравномерных триангуляциях.

Дополнительно будем предполагать, что любой узел сетки $z \in \mathcal{N}$ принадлежит к одному из следующих *Типов Узлов* (см. также Рис. 2.1), которые здесь описываются с использованием фиксированной достаточно малой постоянной c_0 (последняя позволяет различать анизотропные и неанизотропные элементы сетки).

(1) *Анизотропными узлами*, множество которых будет обозначаться \mathcal{N}_{ani} , назовем узлы, которые удовлетворяют соотношениям

$$h_z < c_0 H_z, \quad h_T \simeq h_z \text{ и } H_T \simeq H_z \quad \forall T \subset \omega_z. \quad (2.6)$$

Заметим, что данные условия подразумевают, что \mathcal{S}_z содержит не более двух ребер длины $\lesssim h_z$.

(2) *Полу-анизотропными узлами*, множество которых будет обозначаться $\mathcal{N}_{\text{s,ani}}$, назовем узлы, которые удовлетворяют соотношениям $z \notin \mathcal{N}_{\text{ani}}$ и

$$h_z < c_0 H_z, \quad H_T \simeq H_z \text{ и } h_T \simeq h_z \quad \text{или} \quad H_T \simeq h_T \simeq h_z \quad \forall T \subset \omega_z. \quad (2.7)$$

Также предполагается, что все ребра S в \mathcal{S}_z , имеющие длину $|S| \simeq H_z$, расположены внутри некоторого сектора с вершиной z и внутренним углом при этой вершине $\simeq \frac{h_z}{H_z}$.

(3) *Изотропными узлами*, множество которых будет обозначаться \mathcal{N}_{iso} , назовем узлы, которые удовлетворяют соотношениям

$$c_0 H_z \leq h_z \leq H_z, \quad h_T \simeq H_T \text{ или } H_T \simeq H_z \quad \forall T \subset \omega_z, \quad (2.8)$$

$$\forall S \in \mathcal{S}_z \quad \exists \tilde{T} \subset \omega_z : S \subset \partial \tilde{T}, \quad h_{\tilde{T}} \simeq H_{\tilde{T}},$$

где S является ребром некоторого треугольника \tilde{T} , который не обязательно является элементом \mathcal{T} , но может быть выбран, как описано в (2.8). В силу данного определения, по крайней мере один T in ω_z не является анизотропным и удовлетворяет $h_T \simeq H_T \simeq H_z$, в то время как остальные T в ω_z могут быть анизотропными, или, не будучи анизотропными, могут удовлетворять $H_T \ll h_z \simeq H_z$ (см. Рис. 2.1, справа). Отметим, что в простейшем случае

изотропный узел z окружен лишь неанизотропными элементами, удовлетворяющими $h_T \simeq H_T \simeq H_z$.

(1*) Можно ожидать, что анизотропные элементы, расположенные около границы, будут ориентированы вдоль границы. Чтобы специально выделить узлы, для которых данное свойство не выполняется, введем в рассмотрение следующее подмножество анизотропных граничных узлов $\mathcal{N}_{\partial\Omega}^*$:

$$\mathcal{N}_{\partial\Omega}^* := \left\{ z \in \partial\Omega \cap \mathcal{M} \setminus \mathcal{N}_{\text{iso}} : |\mathcal{S}_z \cap \partial\Omega| \lesssim h_z \text{ или } z \text{ — угол области } \Omega \right\}. \quad (2.9)$$

Будем предполагать, что $\mathcal{N}_{\partial\Omega}^* \cap \mathcal{N}_{\text{s.ani}} = \emptyset$, поэтому $\mathcal{N}_{\partial\Omega}^* \subset \mathcal{N}_{\text{ani}}$.

З а м е ч а н и е 2.1. Отметим, что хотя типы узлов (1)–(3) описывают большинство интересных, с точки зрения практики, ситуаций, но, конечно, они ни в коей мере не исчерпывают все возможные конфигурации, в которых могут появляться анизотропные элементы.

2.3 Теорема о следах для анизотропных элементов

Нам потребуется теорема о следах для анизотропных элементов, которая будет здесь сформулирована с использованием нормализованной нормы пространства $W_1^1(D)$, определяемой соотношением

$$\|v\|_D := (\text{diam}D)^{-1} \|v\|_{1;D} + \|\nabla v\|_{1;D}. \quad (2.10)$$

Отметим, что поскольку $\text{diam}(\omega_z) = H_z$, а $\text{diam}(T) \simeq H_T$, то

$$\|v\|_{\omega_z} = H_z^{-1} \|v\|_{1;\omega_z} + \|\nabla v\|_{1;\omega_z}, \quad \|v\|_T \simeq H_T^{-1} \|v\|_{1;T} + \|\nabla v\|_{1;T}.$$

Поскольку же $\text{diam}(T) \leq \text{diam}(\omega_z)$ для любого $T \subset \omega_z$, то

$$\|v\|_{\omega_z} \leq \sum_{T \subset \omega_z} \|v\|_T, \quad (2.11)$$

в то время как для анизотропного узла, в силу (2.6), получаем

$$\|v\|_{\omega_z} \simeq \sum_{T \subset \omega_z} \|v\|_T. \quad (2.12)$$

Л е м м а 2.2. Для любого узла $z \in \mathcal{N}$, принадлежащего одному из типов (2.6), (2.7) или (2.8), и любой функции $v \in W_1^1(\omega_z)$ справедливо соотношение

$$\|v\|_{1; \dot{\gamma}_z} + \frac{h_z}{H_z} \|v\|_{1; \gamma_z \setminus \dot{\gamma}_z} \lesssim \sum_{T \subset \omega_z} \|v\|_T, \quad (2.13)$$

где используются γ_z и $\dot{\gamma}_z$ из (2.5). Также, для любого отрезка $\bar{S}_z \subset \omega_z$, исходящего из z и удовлетворяющего $|\bar{S}_z| \simeq H_z$, справедливо соотношение

$$\frac{h_z}{H_z} \|v\|_{1; \bar{S}_z} \lesssim \|v\|_{\omega_z}. \quad (2.14)$$

Д о к а з а т е л ь с т в о. Заметим, что для любого ребра S в любом треугольнике $T \in \mathcal{T}$ справедливо соотношение

$$\frac{1}{|S|} \|v\|_{1; S} \lesssim \frac{1}{|T|} \|v\|_{1; T} + \frac{\text{diam}(T)}{|T|} \|\nabla v\|_{1; T}.$$

Здесь $\text{diam}(T) \simeq H_T$, поэтому $\frac{1}{|S|} \|v\|_{1; S} \lesssim \frac{H_T}{|T|} \|v\|_T$, в то время как $|T| \simeq h_T H_T$ приводит к

$$\frac{h_T}{|S|} \|v\|_{1; S} \lesssim \|v\|_T \quad \text{при } S \subset \partial T. \quad (2.15)$$

Далее рассмотрим каждый допустимый тип узла отдельно. В случае анизотропного узла z (2.6), для любого ребра S в любом $T \subset \omega_z$ имеем $h_T \simeq h_z$. Воспользовавшись этим свойством, из (2.15) немедленно получаем (2.13), поскольку $|S| \simeq h_z$ для любого $S \in \dot{\gamma}_z$, а $|S| \simeq H_z$ для любого $S \in \gamma_z \setminus \dot{\gamma}_z$.

Пусть теперь узел z является полу-анизотропным (2.7). Если $S \in \gamma_z$ является стороной такого $T \subset \omega_z$, что $H_T \simeq H_z$ и $h_T \simeq h_z$, то $\|v\|_{1; S}$ оценивается аналогично случаю анизотропных узлов типа (2.6); см. выше. В противном случае $S \in \dot{\gamma}_z$, а также S является стороной некоторого неанизотропного элемента $T' \subset \omega_z$, при этом $h_{T'} \simeq H_{T'} \simeq |S|$, поэтому в силу (2.15) имеем

$\|v\|_{1;S} \lesssim \|v\|_{T'}$. Объединяя эти два вывода, получаем желаемое соотношение (2.13).

Наконец рассмотрим изотропный узел z (2.8), при этом $\overset{\circ}{\gamma}_z = \gamma_z$ поскольку $h_z \simeq H_z$. Если $S \in \gamma_z$ является стороной неанизотропного треугольника $T' \subset \omega_z$, т. е. $h_{T'} \simeq H_{T'} \simeq |S|$, то в силу (2.15) получаем $\|v\|_{1;S} \lesssim \|v\|_{T'}$. В противном случае S является стороной некоторого анизотропного треугольника T'' , при этом либо $|S| \simeq h_{T''}$, либо $|S| \simeq H_{T''} \simeq H_z$. При $|S| \simeq h_{T''}$, применение (2.15) приводит к $\|v\|_{1;S} \lesssim \|v\|_{T''}$. Если же $|S| \simeq H_z$, то найдется такой неанизотропный треугольник $\tilde{T} \subset \omega_z$ (который может и не принадлежать \mathcal{T}), что S является одной из его сторон; и опять применяя (2.15), получаем $\|v\|_{1;S} \lesssim \|v\|_{\tilde{T}} \lesssim \|v\|_{\omega_z}$. Объединяя все эти выводы с (2.11) получаем (2.13) для любого изотропного узла z .

Остается установить (2.14). Заметим, что для любого исходящего из z отрезка $\bar{S}_z \subset \omega_z$, удовлетворяющего $|\bar{S}_z| \simeq H_z$, найдется такой треугольник $\tilde{T} \subset \omega_z$ (который может и не принадлежать \mathcal{T}), что \bar{S}_z является одной из его сторон, при этом $H_{\tilde{T}} = |\bar{S}_z| \simeq H_z$ и $h_{\tilde{T}} \simeq h_z$. Теперь применение (2.15) приводит к $\frac{h_z}{H_z} \|v\|_{1;\bar{S}_z} \lesssim \|v\|_{\tilde{T}} \lesssim \|v\|_{\omega_z}$. Таким образом мы установили оставшуюся оценку (2.14). \square

2.4 Представление ошибки через функцию Грина

З а м е ч а н и е 2.3. С тем, чтобы несколько упростить изложение, в дальнейшем будем предполагать, что f дифференцируема по u , при этом $f_u(x, y; u) \leq \bar{C}_f$ при любых x, y, u . Заметим, что все последующие результаты параграфа могут быть установлены и без этого дополнительного предположения с помощью Леммы 5 из [157] (см. также аналогичную Лемму 1.8).

Для оценки ошибки численного решения, последняя будет представлена через функцию Грина линейного оператора, получающегося линеаризацией

$Lu_h - Lu$. При любом фиксированном $(x', y') \in \Omega$ данная функция G удовлетворяет соотношениям

$$\begin{aligned} -\varepsilon^2 \Delta_{(x,y)} G + p(x, y) G &= \delta(x', y'; x, y), & (x, y) \in \Omega, \\ G(x', y'; x, y) &= 0, & (x, y) \in \partial\Omega. \end{aligned} \quad (2.16)$$

Здесь коэффициент $p = \int_0^1 f_u(\cdot, u + [v - u]s) ds$ получен с помощью стандартной линеаризации $f(x, y; v) - f(x, y; u) = p(x, y)[v - u]$, и, в силу Замечания 2.3, удовлетворяет $C_f \leq p \leq \bar{C}_f$, а δ обозначает двумерную функцию Дирака.

Воспользовавшись данной функцией Грина $G(x', y'; x, y) : \Omega \times \Omega \rightarrow \mathbb{R}$, несложно показать, что для решения u задачи (2.1) и любой функции $v \in \mathring{W}_q^1(\Omega) \subset C(\bar{\Omega})$ with $q > 2$ справедливо представление

$$(v - u)(x', y') = \varepsilon^2 \langle \nabla v, \nabla G(x', y'; \cdot) \rangle + \langle f(\cdot; v), G(x', y'; \cdot) \rangle. \quad (2.17)$$

Нам потребуются следующие оценки, аналогичные [157] (см. также Теорему 1.2).

Л е м м а 2.4. Пусть G является функцией Грина, удовлетворяющей (2.16) при условиях $0 \leq C_f \leq p \leq \bar{C}_f$ и $C_f + \varepsilon^2 \geq 1$. Тогда при любом $(x', y') \in \Omega$, справедливо соотношение

$$\|G(x', y'; \cdot)\|_{1;\Omega} + \varepsilon |G(x', y'; \cdot)|_{1,1;\Omega} \lesssim 1. \quad (2.18)$$

При этом для любого шара $B(x', y'; \rho)$ радиуса ρ с центром в (x', y') , и $\ell_\rho := \ln(2 + \varepsilon\rho^{-1})$,

$$\|G(x', y'; \cdot)\|_{1;B(x', y'; \rho) \cap \Omega} \lesssim \varepsilon^{-2} \rho^2 \ell_\rho, \quad (2.19a)$$

$$|G(x', y'; \cdot)|_{1,1;B(x', y'; \rho) \cap \Omega} \lesssim \varepsilon^{-2} \rho, \quad (2.19b)$$

$$|G(x', y'; \cdot)|_{2,1;\Omega \setminus B(x', y'; \rho)} \lesssim \varepsilon^{-2} \ell_\rho. \quad (2.19c)$$

Д о к а з а т е л ь с т в о. Все результаты леммы представлены в Теореме 1 из [157] (а также в аналогичной Теореме 1.2) для случая $p = C_f$. При этом доказательство применимо и к более общему случаю $C_f \leq p \leq \bar{C}_f$. \square

Воспользовавшись (2.17) при $v = u_h$, а также (2.2), несложно убедиться, что $\forall v_h \in S_h$ выполняется соотношение

$$(u_h - u)(x', y') = \varepsilon^2 \langle \nabla u_h, \nabla(G - v_h) \rangle + \langle f_h^I, G - v_h \rangle + \langle f_h - f_h^I, G \rangle, \quad (2.20)$$

где, несколько злоупотребляя обозначениями, $G = G(x', y'; \cdot)$. Для квадратурной ошибки $\langle f_h - f_h^I, G \rangle =: \mathcal{E}_{\text{quad}}$, в силу (2.18), получаем

$$|\mathcal{E}_{\text{quad}}| \lesssim \|f_h - f_h^I\|_{\infty; \Omega}. \quad (2.21)$$

Заметим, что хотя данная оценка квадратурной ошибки достаточно груба по сравнению с полученной в §1.3.3, наше основное внимание будет сосредоточено на получении адекватных апостериорных оценок ошибки без учета квадратур, что является достаточно сложной задачей в случае анизотропных сеток.

Далее введем стандартный базис, состоящий из функций $\phi_z \in S_h$, соответствующих узлу $z \in \mathcal{N}$, равных 1 в этом узле и 0 в остальных узлах сетки. Положим $v_h := G_h + \sum_{z \in \mathcal{N}} \bar{g}_z \phi_z \in S_h$, где $G_h \in S_h$ является некоторым интерполянтном для G , а значение \bar{g}_z выбирается равным некоторому усреднению величины $G - G_h$ в окрестности z (которое будет определено ниже), при этом $\bar{g}_z = 0$ для $z \in \partial\Omega$ (с тем, чтобы $v_h \in S_h$). Далее введем функцию $g := G - G_h$, с помощью которой получаем $G - v_h = g - \sum_{z \in \mathcal{N}} \bar{g}_z \phi_z = \sum_{z \in \mathcal{N}} (g - \bar{g}_z) \phi_z$. Объединяя это с (2.20), приходим к достаточно стандартному представлению ошибки

$$\begin{aligned} (u_h - u)(x', y') &= \sum_{z \in \mathcal{N}} \varepsilon^2 \int_{\gamma_z} (g - \bar{g}_z) \phi_z \llbracket \nabla u_h \rrbracket \cdot \nu + \sum_{z \in \mathcal{N}} \int_{\omega_z} f_h^I (g - \bar{g}_z) \phi_z + \mathcal{E}_{\text{quad}} \\ &=: I + II + \mathcal{E}_{\text{quad}}, \end{aligned} \quad (2.22)$$

которое справедливо при любом $G_h \in S_h$ и любом наборе $\{\bar{g}_z\}_{z \in \mathcal{N}}$, удовлетворяющем $\bar{g}_z = 0$ при всех $z \in \partial\Omega$.

В (2.22), величина $\llbracket \nabla u_h \rrbracket$ обозначает стандартный скачок нормальной производной при переходе через внутреннее ребро. Для определенности будем считать, что единичная нормаль ν к любому ребру из γ_z направлена по

часовой стрелке по отношению к z , в то время как $[[w]]$ для любой функции w понимается как скачок w при переходе через любое ребро из γ_z , вычисляемый в направлении против часовой стрелки по отношению к z . Таким образом, $[[\nabla u_h]] \cdot \nu =: J_z$ — скачок, в направлении против часовой стрелки по отношению к z , нормальной производной u_h по направлению против часовой стрелки. Очевидно, что $[[\nabla u_h]] = |J_z|$. При использовании обозначения $[[\nabla u_h]]$ на ребрах, расположенных на границе области, будем считать, что $u_h = 0$ в $\mathcal{R}^2 \setminus \Omega$.

2.5 Апостериорный анализ на частично структурированных анизотропных сетках

Чтобы проиллюстрировать основные новые идеи, начнем с изложения несколько упрощенной версии представляемого здесь апостериорного анализа для случая частично структурированных анизотропных сеток в квадратной области $\Omega = (0, 1)^2$. Будет предполагаться, что триангуляция удовлетворяет следующим условиям.

- A1. Пусть $\{x_i\}_{i=0}^n$ является произвольной сеткой по переменной x на отрезке $(0, 1)$. При этом для любого элемента $T \in \mathcal{T}$, при некотором i ,
- (i) его самое короткое ребро расположено на линии $x = x_i$;
 - (ii) противоположная же этому ребру вершина расположена на одной из линий $x = x_{i+1}$ и $x = x_{i-1}$ (см. Рис. 2.2, слева).
- A2. Пусть $\mathcal{N} = \mathcal{N}_{\text{ani}}$, т. е. любой узел $z \in \mathcal{N}$ является анизотропным в смысле (2.6).
- A3. Пусть все углы элементов триангуляции близки к нетупым в том смысле, что в любом треугольнике $T \in \mathcal{T}$ максимальный угол ограничен величиной $\frac{\pi}{2} + \alpha_1 \frac{h_T}{H_T}$ с некоторой положительной постоянной α_1 .

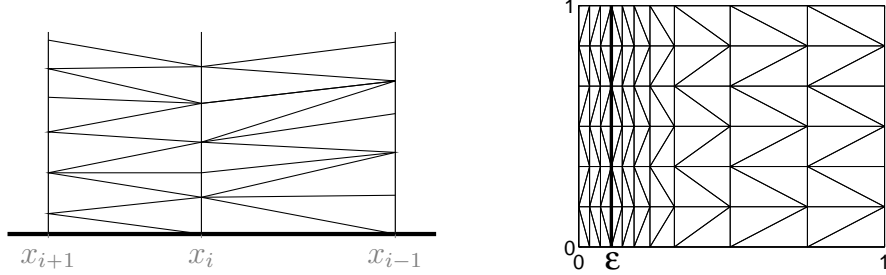


Рис. 2.2: Частично структурированная анизотропная сетка (слева); триангуляция, используемая в §2.5.4 (справа).

Вышеописанные условия подразумевают, что все элементы сетки являются анизотропными и ориентированы по направлению оси x . Также условие А3 предполагает, что при $x_z = x_i$

$$\omega_z \subseteq \omega_z^* := (x_{i-1}, x_{i+1}) \times (y_z^-, y_z^+), \quad y_z^+ - y_z^- \simeq h_z, \quad \text{diam } \omega_z^* \simeq H_z, \quad (2.23)$$

где (y_z^-, y_z^+) — множество значений y в пределах ω_z , при этом будем считать, что $x_{-1} := x_0$, а $x_{n+1} := x_n$.

З а м е ч а н и е 2.5. Из вышеописанных условий (в особенности А3) следует, что найдется такое $J \lesssim 1$, что $\omega_z^* \subset \omega_z^{(J)}$ при любом $z \in \mathcal{N}$, где используются обозначения $\omega_z^{(0)} := \omega_z$ и $\omega_z^{(j+1)}$ для элементов ω или касающихся $\omega_z^{(j)}$. (Отметим также, что $J = 1$ в случае триангуляции с нетупыми углами, т. е. при $\alpha_1 = 0$ в условии А3.)

2.5.1 Выбор \bar{g}_z . Основные результаты

Выбор значений \bar{g}_z в (2.22) будет связан с локальной ориентацией анизотропных элементов, и является ключевым для представляемого здесь анализа. Пусть $\bar{g}_z = 0$ при $z \in \partial\Omega$, а в противном случае, при $x_z = x_i$ с некоторым $1 \leq i \leq n-1$, положим

$$\int_{x_{i-1}}^{x_{i+1}} (g(x, y_z) - \bar{g}_z) \varphi_i(x) dx = 0. \quad (2.24)$$

Здесь $\varphi_i(x)$ — стандартная одномерная кусочно-линейная базисная функция, соответствующая x_i (т. е. имеющая носитель (x_{i-1}, x_{i+1}) , равная 1 в $x = x_i$, и линейная на (x_{i-1}, x_i) и (x_i, x_{i+1})).

З а м е ч а н и е 2.6. При $x_z = x_i$, пусть $\bar{S}_z \subset \omega_z^*$ будет отрезком, соединяющим точки (x_{i-1}, y_z) и (x_{i+1}, y_z) . Тогда определение (2.24) для \bar{g}_z идентично следующему:

$$\int_{\bar{S}_z} (g - \bar{g}_z) \varphi_i = 0 \quad \text{при } x_z = x_i, \quad 1 \leq i \leq n-1. \quad (2.25)$$

В случае же триангуляции без тупых углов данное определение эквивалентно $\int_{\bar{S}_z} (g - \bar{g}_z) \phi_z = 0$. Отметим, что описанный здесь выбор \bar{g}_z принципиально отличается от более стандартного и достаточно часто используемого (в случае локально квазиравномерных триангуляций), которой обозначен здесь \bar{g}'_z и определяется соотношением $\int_{\omega_z} (g - \bar{g}'_z) \phi_z = 0$ (см., например, Лекцию 5 в [124]).

Т е о р е м а 2.7. Пусть $\lambda_z = \min\{\varepsilon, H_z\}$, $g = G(x', y'; \cdot) - G_h$ с любой $G_h \in S_h$, а

$$\Theta := \varepsilon^2 \sum_{z \in \mathcal{N}} \lambda_z^{-1} \|g\|_{\omega_z^*}, \quad \Theta' := \varepsilon^2 \sum_{z \in \mathcal{N}} \lambda_z^{-2} \|g\|_{1, \omega_z^*}. \quad (2.26)$$

Тогда $(u_h - u)(x', y') = I + II + \mathcal{E}_{\text{quad}}$ в любой точке $(x', y') \in \Omega$, при этом для $\mathcal{E}_{\text{quad}}$ справедлива оценка (2.21), а при условиях А1–А3 также справедливы соотношения

$$|I| \lesssim \Theta \max_{z \in \mathcal{N}} \{ \lambda_z \|J_z\|_{\infty; \gamma_z} \}, \quad (2.27)$$

$$|II| \lesssim \Theta \max_{z \in \mathcal{N}} \{ \varepsilon^{-2} \lambda_z H_z \|f_h^I\|_{\infty; \omega_z} \}, \quad (2.28)$$

$$|II| \lesssim \Theta' \max_{z \in \mathcal{N}} \{ \varepsilon^{-2} \lambda_z^2 \|f_h^I\|_{\infty; \omega_z} \}. \quad (2.29)$$

Также для Π справедлива альтернативная оценка

$$|\Pi| \lesssim \max_{z \in \mathcal{N} \setminus \mathcal{N}_{\partial\Omega}^*} \{ (\Theta + \Theta') \varepsilon^{-2} \lambda_z \min\{\varepsilon, h_z\} \|f_h^I\|_{\infty; \omega_z} + \Theta' \varepsilon^{-2} \lambda_z^2 \operatorname{osc}(f_h^I; \omega_z) \} \\ + \Theta' \max_{z \in \mathcal{N}_{\partial\Omega}^*} \{ \varepsilon^{-2} \lambda_z^2 \|f_h^I\|_{\infty; \omega_z} \}, \quad (2.30)$$

где $\mathcal{N}_{\partial\Omega}^* = \{z \in \mathcal{N} : x_z = 0 \text{ или } x_z = 1\}$ (в соответствии с (2.9)).

2.5.2 Невязки, ассоциируемые с ребрами. Доказательство (2.27)

Доказательство (2.27). Заметим, что $[[\nabla u_h]] \cdot \nu = [[\partial_x u_h]] \nu_x + [[\partial_y u_h]] \nu_y$. Поэтому можем разбить I из (2.22) на три компоненты:

$$I = I' + I'' + I''' := \sum_{z \in \mathcal{N}} \varepsilon^2 \int_{\gamma_z} (g - \bar{g}_z) \phi_z [[\partial_x u_h]] \nu_x \\ + \sum_{z \in \mathcal{N}} \varepsilon^2 \int_{\gamma_z} [g - g(x, y_z)] \phi_z [[\partial_y u_h]] \nu_y \\ + \sum_{z \in \mathcal{N}} \varepsilon^2 \int_{\gamma_z} [g(x, y_z) - \bar{g}_z] \phi_z [[\partial_y u_h]] \nu_y. \quad (2.31)$$

В силу (2.26), для вывода желаемой оценки (2.27) достаточно убедиться, что

$$|I'| + |I''| \lesssim \sum_{z \in \mathcal{N}} \varepsilon^2 \|g\|_{\omega_z^*} \|J_z\|_{\infty; \gamma_z}, \quad I''' = 0. \quad (2.32)$$

Причем приведенная здесь оценка для I'' легко получается объединением соотношений $|\phi_z [[\partial_y u_h]]| \leq |J_z|$ и $|\nu_y| \leq 1$, а также $\int_{\gamma_z} |g(x, y) - g(x, y_z)| \lesssim \|\partial_y g\|_{1; \omega_z^*} \lesssim \|g\|_{\omega_z^*}$.

Получение оценки для I' в (2.32) несколько более трудоемко. Вновь воспользуемся $|\phi_z [[\partial_x u_h]]| \leq |J_z|$, поэтому остается показать, что

$$\int_{\gamma_z} (|g \nu_x| + |\bar{g}_z \nu_x|) \lesssim \|g\|_{\omega_z^*}. \quad (2.33)$$

Здесь весьма важно, что $|\nu_x| \lesssim \frac{h_z}{|S|}$ при любом $S \in \gamma_z$ (что следует из $|\nu_x| = |\sin \angle(S, i_x)|$ и А3). При этом при $S \subset \dot{\gamma}_z$ имеем $|S| \simeq h_z$, поэтому $|\nu_x| \leq 1$,

в то время как при $S \subset \gamma_z \setminus \overset{\circ}{\gamma}_z$ получаем $|S| \simeq H_z$, и поэтому $|\nu_x| \leq \frac{h_z}{H_z}$. Далее применение оценки (2.13), объединенной с (2.12), приводит к $\int_{\gamma_z} |g \nu_x| \lesssim \|g\|_{\omega_z}$. Отсюда $\int_{\gamma_z} |g \nu_x| \lesssim \|g\|_{\omega_z^*}$, поскольку $\text{diam } \omega_z \simeq \text{diam } \omega_z^*$. Заметим, что $\int_{\gamma_z} |\bar{g}_z \nu_x| \lesssim h_z |\bar{g}_z| \lesssim \frac{h_z}{H_z} \|g\|_{1; \bar{S}_z}$, для чего мы воспользовались оценкой (2.25), в которой $\bar{S}_z \subset \omega_z^*$ и $|\bar{S}_z| \simeq H_z$. Применение же (2.14) приводит к $\int_{\gamma_z} |\bar{g}_z \nu_x| \lesssim \|g\|_{\omega_z^*}$. Отметим также, что $\bar{g}_z = 0$ при $x_z = x_0$ и $x_z = x_n$, поэтому соотношение (2.33) остается справедливым. Таким образом, оценка для I' в (2.32) установлена.

Остается оценить $I''' = \sum_{z \in \mathcal{N}} I_z'''$, где

$$I_z''' := \varepsilon^2 \int_{\gamma_z} [g(x, y_z) - \bar{g}_z] \phi_z \llbracket \partial_y u_h \rrbracket \nu_y. \quad (2.34)$$

Сначала рассмотрим $z \in \mathcal{N} \setminus \partial\Omega$, т. е. $x_z = x_i$ при некотором $i = 1, \dots, n-1$, а $y_z \neq 0, 1$. Структура рассматриваемой сетки такова, что $\phi_z = \varphi_i(x)$ на $\gamma_z \setminus \overset{\circ}{\gamma}_z$, в то время как $\nu_y |d\nu| = -\text{sgn}(x - x_i) dx$, а функция $[g(x, y_z) - \bar{g}_z]$ зависит только от переменной x . Отсюда получаем

$$\begin{aligned} I_z''' &= \varepsilon^2 \left(\sum_{S \in \gamma_z^-} \llbracket \partial_y u_h \rrbracket \right) \int_{x_{i-1}}^{x_i} [g(x, y_z) - \bar{g}_z] \varphi_i(x) dx \\ &\quad - \varepsilon^2 \left(\sum_{S \in \gamma_z^+} \llbracket \partial_y u_h \rrbracket \right) \int_{x_i}^{x_{i+1}} [g(x, y_z) - \bar{g}_z] \varphi_i(x) dx. \end{aligned} \quad (2.35)$$

Здесь используются $\gamma_z^+ = \{S \in \gamma_z : \text{proj}_x S = (x_i, x_{i+1})\}$ и $\gamma_z^- = \{S \in \gamma_z : \text{proj}_x S = (x_{i-1}, x_i)\}$, где $\text{proj}_x(\cdot)$ обозначает проекцию на ось x . При этом $\gamma_z = \overset{\circ}{\gamma}_z \cup \gamma_z^+ \cup \gamma_z^-$, где $\overset{\circ}{\gamma}_z = \gamma_z \cap \{x = x_i\}$ содержит два коротких ребра, на которых $\nu_y = 0$, и поэтому $\int_{\overset{\circ}{\gamma}_z}$ не присутствует в (2.35). Объединяя (2.35) с определением (2.24) для \bar{g}_z , приходим к представлению

$$I_z''' = \varepsilon^2 \left(\sum_{S \in \gamma_z^- \cup \gamma_z^+} \llbracket \partial_y u_h \rrbracket \right) \int_{x_{i-1}}^{x_i} [g(x, y_z) - \bar{g}_z] \varphi_i(x) dx. \quad (2.36)$$

Здесь $\sum_{S \in \gamma_z^- \cup \gamma_z^+} \llbracket \partial_y u_h \rrbracket = \sum_{S \in \mathcal{S}_z} \llbracket \partial_y u_h \rrbracket = 0$ (поскольку $\gamma_z = \mathcal{S}_z$, а $\llbracket \partial_y u_h \rrbracket|_{\overset{\circ}{\gamma}_z} = 0$). Отсюда $I_z''' = 0$ при $z \in \mathcal{N} \setminus \partial\Omega$.

Если же $z \in \mathcal{N} \cap \partial\Omega$, и при этом $y_z = 0$ или $y_z = 1$, то $\bar{g}_z = 0$, а также $g(x, y_z) = 0$ (поскольку $g = 0$ на $\partial\Omega$), поэтому, в силу (2.34), опять получаем $I_z''' = 0$. Наконец рассмотрим I_z''' в случае $z \in \mathcal{N} \cap \partial\Omega$, но при этом $y_z \neq 0, 1$ и $x_z = x_n$ (случай $x_z = x_0$ рассматривается аналогично). Тогда $\bar{g}_z = 0$, но при этом (2.35) и (2.36) остаются справедливыми, а $\gamma_z^+ = \emptyset$. Также в этом случае имеем $\sum_{\mathcal{S} \in \gamma_z^-} \llbracket \partial_y u_h \rrbracket = \sum_{\mathcal{S} \in \mathcal{S}_z} \llbracket \partial_y u_h \rrbracket = 0$ (поскольку $\gamma_z^- = \mathcal{S}_z \setminus \partial\Omega$, а $\llbracket \partial_y u_h \rrbracket \big|_{\mathcal{S}_z \cap \partial\Omega} = 0$). Отсюда вновь получаем $I_z''' = 0$. \square

З а м е ч а н и е 2.8. *Внимательное изучение вышеизложенного доказательства показывает, что оно остается справедливым, если $\{\bar{g}_z\}_{z \in \mathcal{N}}$, определяемые в (2.24), будут заменены значениями $\{\bar{g}_z^*\}_{z \in \mathcal{N}}$ такими, что $\bar{g}_z^* = 0$ при $z \in \partial\Omega$, а в противном случае справедливы соотношения*

$$H_z |\bar{g}_z - \bar{g}_z^*| \lesssim \|g\|_{\omega_z^*}, \quad h_z H_z |\bar{g}_z^*| \lesssim \|g\|_{1; \omega_z^*}. \quad (2.37)$$

Действительно, тогда в I нужно учесть дополнительное слагаемое $I^* := \sum_{z \in \mathcal{N}} \varepsilon^2 \int_{\gamma_z} (\bar{g}_z - \bar{g}_z^*) \phi_z \llbracket \nabla u_h \rrbracket \cdot \nu$, для которого несложно получить оценку $|I^*| \leq \sum_{z \in \mathcal{N}} \varepsilon^2 H_z |\bar{g}_z - \bar{g}_z^*| \|J_z\|_{\infty; \gamma_z}$. Из последней вытекает (в силу первого соотношения в (2.37)) такая же оценка, что и для $|I'| + |I''|$ в (2.32).

Заметим, что соотношения (2.37) справедливы, если \bar{g}_z^* , при $z \in \mathcal{N} \setminus \partial\Omega$ и $x_z = x_i$, определяется как

$$\int_{\omega_z^*} [g(x, y) - \bar{g}_z^*] \varphi_i(x) = 0, \quad (2.38)$$

где используется $\omega_z^* = (x_{i-1}, x_{i+1}) \times (y_z^-, y_z^+)$, при этом $|\omega_z^*| \simeq h_z H_z$. Тогда второе соотношение в (2.37) очевидно выполняется, в то время как $\int_{\omega_z^*} [g(x, y_z) - g(x, y)] \varphi_i(x) \simeq h_z H_z (\bar{g}_z - \bar{g}_z^*)$ приводит к $H_z |\bar{g}_z - \bar{g}_z^*| \lesssim \|\partial_y g\|_{1; \omega_z^*}$, и таким образом к первому соотношению в (2.37).

2.5.3 Поэлементные невязки. Доказательство оценок (2.28)–(2.30)

Далее рассмотрим компоненту II в ошибке решения (2.22).

Доказательство (2.28). В силу (2.26), достаточно показать, что

$$|II| \lesssim \sum_{z \in \mathcal{N}} H_z \|g\|_{\omega_z^*} \|f_h^I\|_{\infty; \omega_z}. \quad (2.39)$$

Заметим, что при оценке I' в §2.5.2, было установлено, что $h_z |\bar{g}_z| \lesssim \frac{h_z}{H_z} \|g\|_{1; \bar{S}_z} \lesssim \|g\|_{\omega_z^*}$ (при этом $\bar{g}_z = 0$ при $z \in \partial\Omega$). Следовательно $\|g - \bar{g}_z\|_{1; \omega_z} \lesssim \|g\|_{1; \omega_z} + h_z H_z |\bar{g}_z| \lesssim H_z \|g\|_{\omega_z^*}$. Объединяя это с определением II в (2.22), приходим к оценке (2.39), из которой вытекает (2.28). \square

Доказательство (2.29). Данная оценка устанавливается аналогично (2.28), с той лишь разницей, что определение Θ' объединяется с несколько иной оценкой $|II| \lesssim \sum_{z \in \mathcal{N}} \|g\|_{1; \omega_z^*} \|f_h^I\|_{\infty; \omega_z}$. Чтобы получить последнюю, воспользовавшись Замечанием 2.8, заменим $\{\bar{g}_z\}$ в (2.22) значениями $\{\bar{g}_z^*\}$ из (2.38). При этом, в силу (2.37), справедлива оценка $\|g - \bar{g}_z^*\|_{1; \omega_z} \lesssim \|g\|_{1; \omega_z^*}$. \square

Доказательство (2.30). Воспользовавшись Замечанием 2.8, мы вновь заменим $\{\bar{g}_z\}$ в (2.22) значениями $\{\bar{g}_z^*\}$ из (2.38). Введем разбиение $\mathcal{N} = \cup_{i=0}^n \mathcal{N}_i$, где $\mathcal{N}_i := \{z : x_z = x_i\}$. Заметим, что $\mathcal{N}_0 \cup \mathcal{N}_n = \mathcal{N}_{\partial\Omega}^*$. Далее разобьем II из (2.22) в виде $II = \sum_{i=1}^{n-1} II_i + II_{\text{osc}} + II_{\partial\Omega}^*$ следующим образом:

$$\begin{aligned} II_i &:= \sum_{z \in \mathcal{N}_i} \int_{\omega_z} f_h^I(x_i, y) (g - \bar{g}_z^*) \phi_z, \\ II_{\text{osc}} &:= \sum_{z \in \mathcal{N} \setminus \mathcal{N}_{\partial\Omega}^*} \int_{\omega_z} [f_h^I - f_h^I(x_z, y)] (g - \bar{g}_z^*) \phi_z, \\ II_{\partial\Omega}^* &:= \sum_{z \in \mathcal{N}_{\partial\Omega}^*} \int_{\omega_z} f_h^I (g - \bar{g}_z^*) \phi_z. \end{aligned}$$

Здесь для II_{osc} и $II_{\partial\Omega}^*$ мы немедленно получаем оценки, аналогичные (2.29):

$$|II_{\text{osc}}| \lesssim \Theta' \max_{z \in \mathcal{N} \setminus \mathcal{N}_{\partial\Omega}^*} \left\{ \varepsilon^{-2} \lambda_z^2 \|f_h^I - f_h^I(x_z, y)\|_{\infty; \omega_z} \right\}, \quad (2.40)$$

$$|II_{\partial\Omega}^*| \lesssim \Theta' \max_{z \in \mathcal{N}_{\partial\Omega}^*} \left\{ \varepsilon^{-2} \lambda_z^2 \|f_h^I\|_{\infty; \omega_z} \right\}. \quad (2.41)$$

Таким образом, остается оценить компоненты Π_i при $1 \leq i \leq n-1$, которые могут быть представлены в виде

$$\Pi_i = \sum_{z \in \mathcal{N}_i} \int_0^1 f_h^I(x_i, y) \int_{x_{i-1}}^{x_{i+1}} (g - \bar{g}_z^*) \phi_z dx dy.$$

Поскольку $\sum_{z \in \mathcal{N}_i} \phi_z = \varphi_i(x)$, то получаем

$$\sum_{z \in \mathcal{N}_i} \int_{x_{i-1}}^{x_{i+1}} g \phi_z dx = \int_{x_{i-1}}^{x_{i+1}} g \varphi_i dx =: \hat{g}_i(y) \int_{x_{i-1}}^{x_{i+1}} \varphi_i dx = \sum_{z \in \mathcal{N}_i} \int_{x_{i-1}}^{x_{i+1}} \hat{g}_i(y) \phi_z dx.$$

Здесь \hat{g}_i определена аналогично \bar{g}_z в (2.24). Далее,

$$\Pi_i = \sum_{z \in \mathcal{N}_i} \int_0^1 f_h^I(x_i, y) \int_{x_{i-1}}^{x_{i+1}} (\hat{g}_i(y) - \bar{g}_z^*) \phi_z dx dy = \sum_{z \in \mathcal{N}_i} \int_{\omega_z} f_h^I(x_i, y) (\hat{g}_i(y) - \bar{g}_z^*) \phi_z.$$

Сделовательно

$$|\Pi_i| \lesssim \sum_{z \in \mathcal{N}_i} \|\hat{g}_i(y) - \bar{g}_z^*\|_{1; \omega_z} \|f_h^I\|_{\infty; \omega_z^*} \lesssim \sum_{z \in \mathcal{N}_i} \min\{\|g\|_{1; \omega_z^*}, h_z \|g\|_{\omega_z^*}\} \|f_h^I\|_{\infty; \omega_z^*}. \quad (2.42)$$

Здесь вы воспользовались соотношением $\|\hat{g}_i(y) - \bar{g}_z^*\|_{1; \omega_z} \lesssim \|g\|_{1; \omega_z^*}$ (которое следует из определения \hat{g}_i и (2.38)), а также $|\omega_z| \simeq h_z H_z$, объединенным с $h_z H_z \|\hat{g}_i(y) - \bar{g}_z^*\|_{\infty, \omega_z} \lesssim h_z \|g\|_{\omega_z^*}$. Чтобы получить последнее соотношение, зафиксируем $y' \in (y_z^-, y_z^+)$ и заметим, что определение \hat{g}_i и (2.38) приводят к $\int_{\omega_z^*} [g(x, y') - g(x, y)] \varphi_i(x) \simeq h_z H_z [\hat{g}_i(y') - \bar{g}_z^*]$, где $\omega_z^* = (x_{i-1}, x_{i+1}) \times (y_z^-, y_z^+)$, при этом $|\omega_z^*| \simeq h_z H_z$, поэтому действительно $H_z |\hat{g}_i(y') - \bar{g}_z^*| \lesssim \|\partial_y g\|_{1; \omega_z^*} \lesssim \|g\|_{\omega_z^*}$.

Наконец, объединяя (2.42) со справедливым при любых $a, b > 0$ соотношением $\min\{a, h_z b\} \lesssim (\lambda_z^{-2} a + \lambda_z^{-1} b) \min\{\lambda_z^2, \lambda_z h_z\}$ и $\min\{\lambda_z^2, \lambda_z h_z\} = \lambda_z \min\{\varepsilon, h_z\}$, а затем с (2.40), (2.41) и $\|f_h^I - f_h^I(x_z, y)\|_{\infty; \omega_z} \leq \text{osc}(f_h^I; \omega_z^*)$, а также с Замечанием 2.5, приходим к желаемой оценке (2.30). \square

2.5.4 Численные результаты

Прежде чем приступить к теоретическому анализу для более общих анизотропных сеток, протестируем оценители Теоремы 2.7 на простой тестовой

Таблица 2.1: Сетка Бахвалова (см. Рис. 2.2, справа), $M = \frac{1}{2}N$: ошибки в норме максимума модуля и соответствующие оценщики \mathcal{E} и $\mathcal{E}_{(2.29)}$.

N	$\varepsilon = 1$	$\varepsilon = 2^{-5}$	$\varepsilon = 2^{-10}$	$\varepsilon = 2^{-15}$	$\varepsilon = 2^{-20}$	$\varepsilon = 2^{-25}$	$\varepsilon = 2^{-30}$
	Ошибки (нечетные ряды) & Порядки Сходимости (четные ряды)						
64	3.373e-4	3.723e-3	8.952e-3	8.973e-3	8.973e-3	8.973e-3	8.973e-3
	2.00	1.91	1.01	1.00	1.00	1.00	1.00
128	8.445e-5	9.935e-4	4.446e-3	4.484e-3	4.484e-3	4.484e-3	4.484e-3
	2.00	1.98	1.04	1.00	1.00	1.00	1.00
256	2.112e-5	2.523e-4	2.165e-3	2.236e-3	2.236e-3	2.236e-3	2.236e-3
	Оценщик $\mathcal{E} = \max\{\mathcal{E}_{(2.27)}, \mathcal{E}_{(2.30)}\}$ (нечетные ряды) & Индексы Эффективности (четные ряды)						
64	7.353e-3	1.204e-1	1.224e-1	1.230e-1	1.302e-1	1.302e-1	1.302e-1
	21.80	32.33	13.68	14.48	14.51	14.51	14.51
128	1.885e-3	3.212e-2	6.005e-2	6.621e-2	6.646e-2	6.647e-2	6.647e-2
	22.32	32.33	13.51	14.77	14.82	14.82	14.82
256	4.771e-4	8.268e-3	3.073e-2	3.328e-2	3.354e-2	3.354e-2	3.354e-2
	22.59	32.77	14.20	14.89	15.00	15.00	15.00
	Оценщик $\mathcal{E}_{(2.29)}$ (нечетные ряды) & Индексы Эффективности (четные ряды)						
64	6.810e-3	2.516e-1	9.403e-1	9.981e-1	9.999e-1	1.000e+0	1.000e+0
	20.19	67.59	105.04	111.23	111.44	111.45	111.45
128	1.761e-3	1.120e-1	8.858e-1	9.961e-1	9.999e-1	1.000e+0	1.000e+0
	20.86	112.72	199.26	222.15	222.98	223.01	223.01
256	4.480e-4	4.036e-2	7.901e-1	9.922e-1	9.998e-1	1.000e+0	1.000e+0
	21.21	159.97	365.01	443.82	447.17	447.27	447.28

задаче, которая получается из (2.1) при $\Omega = (0, 1)^2$ и $f = u - F(x, y)$, и где F соответствует точному решению $u = 4y(1-y)[1-x^2 - (e^{-x/\varepsilon} - e^{-1/\varepsilon})/(1 - e^{-x/\varepsilon})]$ (последнее имеет пограничный слой вдоль $x = 0$). Будет рассмотрена одна априорно построенная адаптивная сетка, полученная из прямоугольной сетки вычерчиванием диагоналей так, как показано на Рис. 2.2 (справа). При этом прямоугольная сетка определяется как тензорное произведение сетки Бахвалова $\{\chi(\frac{i}{N})\}_{i=1}^N$ по переменной x [8] и равномерной сетки $\{\frac{j}{M}\}_{j=0}^M$ по пе-

ременной y . Сетка Бахвалова определяется с помощью непрерывной на $[0, 1]$ функции χ ; при этом $\chi(t) = t$ при $\varepsilon > \frac{1}{6}$; в противном случае $\chi(t) = 3\varepsilon \ln \frac{1}{1-2t}$ при $t \in (0, \frac{1}{2} - 3\varepsilon)$, а при других значениях t функция χ линейна и удовлетворяет $\chi(1) = 1$.

В Таблице 2.1 представлены значения ошибки численного решения в норме максимума модуля (при этом максимум вычислялся по узлам сетки), соответствующие вычислительные порядки сходимости, а также оцениватели $\mathcal{E} := \max\{\mathcal{E}_{(2.27)}, \mathcal{E}_{(2.30)}\}$ и $\mathcal{E}_{(2.29)}$ с соответствующими индексами эффективности (вычисляемыми как отношение оценивателя к ошибке). При этом $\mathcal{E}_{(2.\cdot)}$ понимается как правая часть оценки (2. \cdot), в которой полагается $\Theta = \Theta' = 1$ (в то время как, в силу Леммы 2.17 ниже, $\Theta + \Theta' \lesssim \ell_h$). Также величины типа $\min\{1, \varepsilon^{-1}a\}$ заменены соответствующими более гладкими значениями $\frac{a}{\varepsilon+a}$, например, $\mathcal{E}_{(2.29)} = \max_{z \in \mathcal{N}} \left\{ \left(\frac{H_z}{\varepsilon + H_z} \right)^2 \|f_h^I\|_{\infty; \omega_z} \right\}$. Заметим, что \mathcal{E} определяется как максимум, а не соответствующая сумма $\mathcal{E}_{(2.27)} + \mathcal{E}_{(2.30)} \simeq \mathcal{E}$, поскольку при этом сравнение с оценивателем $\mathcal{E}_{(2.29)}$ становится более сбалансированным. Заметим также, что оцениватель \mathcal{E} аналогичен (2.4), в то время как $\mathcal{E}_{(2.29)}$ дает лишь несколько более точную оценку для поэлементных невязок по сравнению с (2.3).

Отметим, что используемая здесь сетка обеспечивает ошибку итерполяции второго порядка, т. е. $\|u - u^I\|_{\infty; \infty} \lesssim N^{-2}$; однако порядки сходимости ошибки численного решения ухудшаются с 2 до 1 при $\varepsilon \rightarrow 0$ (данный феномен был описан и теоретически объяснен в работе [167]).

При рассмотренных значениях ε и N аспектные соотношения элементов сетки принимают значения от 1 до $3.6e+8$. При этом оцениватель \mathcal{E} работает достаточно хорошо, а его индексы эффективности стабилизируются при $\varepsilon \rightarrow 0$. Если же используется компонента $\mathcal{E}_{(2.29)}$ оценивателя из (2.3), мы наблюдаем, что она достаточно хорошо работает при $\varepsilon \simeq 1$, однако в сингулярно возмущенном режиме индексы эффективности для $\mathcal{E}_{(2.29)}$ весьма заметно ухудшаются, т. е. использование этого более простого оценивателя для по-

элементных невязок может приводить к значительной переоценке ошибки численного решения.

2.6 Апостериорный анализ на неструктурированных анизотропных сетках

2.6.1 Первый оценитель

Приступим к рассмотрению неструктурированных анизотропных сеток, и начнем с получения оценителя, аналогичного (2.27), (2.28), (2.21). Для упрощения изложения сделаем следующее предположение.

- $\mathcal{A}1$. Если $z \in \mathcal{N}$ является углом многоугольника Ω , то $z \in \mathcal{N}_{\text{iso}}$ (т. е. все углы области являются изотропными узлами, и ни один из них не принадлежит $\mathcal{N}_{\partial\Omega}^*$).

Данное предположение разумно, поскольку типичные угловые особенности не являются анизотропными. Иногда будет сделано еще одно предположение.

- $\mathcal{A}2$. Максимальный угол при любом анизотропном узле z не превосходит $\frac{\pi}{2} + \alpha_1 \frac{h_z}{H_z}$ для некоторой положительной постоянной α_1 .

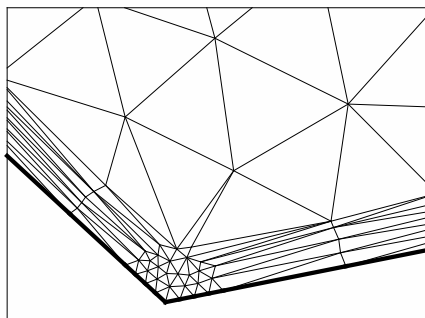


Рис. 2.3: Пример триангуляции, рассматриваемой в §2.6, которая удовлетворяет $\mathcal{A}1$, $\mathcal{A}2$ и $\mathcal{A}3$.

Что касается последнего предположения, то оно используется лишь при доказательстве (2.44) и также несколько упрощает изложение. При этом (2.44) может быть получено и без этого предположения (как это сделано при доказательстве (3.25) в §3.6.2).

Рис. 2.3 показывает пример сетки, которая удовлетворяет всем предположениям, сделанным в данном параграфе (см. Теоремы 2.9 и 2.15).

Т е о р е м а 2.9. Пусть $\lambda_z = \min\{\varepsilon, H_z\}$, при этом $g = G(x', y'; \cdot) - G_h$ при любом $G_h \in S_h$, а

$$\Theta := \varepsilon^2 \sum_{z \in \mathcal{N}} \lambda_z^{-1} \sum_{T \subset \omega_z} \|g\|_T, \quad \Theta' := \varepsilon^2 \sum_{z \in \mathcal{N}} \lambda_z^{-2} \|g\|_{1, \omega_z}. \quad (2.43)$$

Тогда $(u_h - u)(x', y') = I + II + \mathcal{E}_{\text{quad}}$ в любой точке $(x', y') \in \Omega$, где $\mathcal{E}_{\text{quad}}$ оценивается в (2.21), при этом при выполнении условий $\mathcal{A}1$ и $\mathcal{A}2$, справедлива оценка

$$|I| \lesssim \Theta \max_{z \in \mathcal{N}} \{ \lambda_z \|J_z\|_{\infty; \gamma_z} \}. \quad (2.44)$$

Если же выполняется только $\mathcal{A}1$ (а выполнение $\mathcal{A}2$ не предполагается), то

$$|I| \lesssim \Theta \left(\max_{z \in \mathcal{N}} \{ \lambda_z \|J_z\|_{\infty; \gamma_z} \} + \max_{z \in \tilde{\mathcal{N}}_{\text{ani}}} \{ \lambda_z H_z h_z^{-1} |\mathcal{J}_z| \} \right), \quad \mathcal{J}_z = \sum_{S \in \mathring{\mathcal{S}}_z} \sigma_z J_z, \quad (2.45)$$

где $\tilde{\mathcal{N}}_{\text{ani}} = [\mathcal{N}_{\text{ani}} \setminus \partial\Omega] \cup \mathcal{N}_{\partial\Omega}^* \subset \mathcal{N}_{\text{ani}}$, и используются обозначения $\sigma_z|_S = \cos \angle(S, \hat{S}_z)$, $\mathring{\mathcal{S}}_z = \mathring{\gamma}_z$ при $z \in \mathcal{N}_{\text{ani}} \setminus \partial\Omega$, а также $\mathring{\mathcal{S}}_z = \mathcal{S}_z \cap \partial\Omega$ при $z \in \mathcal{N}_{\partial\Omega}^*$; при этом также справедливы оценки

$$|II| \lesssim \Theta \max_{z \in \mathcal{N}} \{ \varepsilon^{-2} \lambda_z H_z \|f_h^I\|_{\infty; \omega_z} \}, \quad (2.46)$$

$$|II| \lesssim \Theta' \max_{z \in \mathcal{N}} \{ \varepsilon^{-2} \lambda_z^2 \|f_h^I\|_{\infty; \omega_z} \}. \quad (2.47)$$

З а м е ч а н и е 2.10. Отметим, что в определении \mathcal{J}_z в (2.45) используются скачки J_z нормальной производной численного решения при переходе через ребра, принадлежащие \mathring{S}_z , которые вычисляются в направлении против часовой стрелки по отношению к z . При этом \mathring{S}_z включает ровно 2 ребра длины $\simeq h_z$. Если же $\mathring{S}_z \subset \partial\Omega$, то используется $u_h = 0$ в $\mathcal{R}^2 \setminus \Omega$, при этом $|\sigma_z| = |\cos \angle(\partial\Omega, \hat{S}_z)|$ не меняется на \mathring{S}_z , поэтому в этом случае справедливо $|\mathcal{J}_z| = |\sigma_z| |\sum_{S \in \mathring{S}_z} [\nabla u_h]|$.

2.6.2 Невязки, ассоциируемые с ребрами. Доказательство оценок (2.44) и (2.45) на неструктурированных сетках

Д о к а з а т е л ь с т в о (2.44). Достаточно показать, что $|\mathcal{J}_z| \lesssim \frac{h_z}{H_z} \|J_z\|_{\infty; \gamma_z}$; тогда оценка (2.44) немедленно следует из (2.45), что мы сейчас и покажем. Действительно, при любом $z \in \tilde{\mathcal{N}}_{\text{ani}}$ любое ребро $S \in \mathring{S}_z$ имеет длину $\simeq h_z$, поэтому $\mathcal{A}2$, объединенное с условием локальной ориентации элементов сетки, а также с (2.6), приводит к $|\angle(S, \hat{S}_z) - \frac{\pi}{2}| \lesssim \frac{h_z}{H_z}$. Отсюда $|\sigma_z| = |\cos \angle(S, \hat{S}_z)| \lesssim \frac{h_z}{H_z}$. Поэтому при $\mathring{S}_z = \dot{\gamma}_z \subset \gamma_z$ необходимая оценка для $|\mathcal{J}_z|$ очевидна. В противном случае (см. Замечание 2.10) $\mathring{S}_z \subset \partial\Omega$, при этом $|\mathcal{J}_z| \lesssim \frac{h_z}{H_z} |\sum_{S \in \gamma_z} [\nabla u_h]|$ (в силу $\mathcal{S}_z = \mathring{S}_z \cup \gamma_z$, а также $\sum_{S \in \mathcal{S}_z} [\nabla u_h] = 0$). И мы опять приходим к необходимой оценке для $|\mathcal{J}_z|$. \square

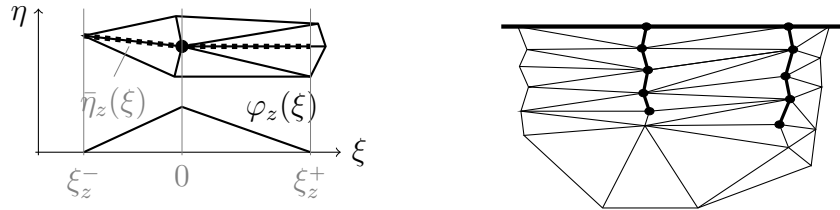


Рис. 2.4: Локальные обозначения, ассоциируемые с узлом z (слева); выделены цепочки коротких ребер, соответственно соединяющих анизотропные и полу-анизотропные узлы (справа).

Доказательство (2.45). Для любого фиксированного узла $z \in \mathcal{N}$, введем следующие локальные обозначения. Выберем локальные декартовы координаты (ξ, η) таким образом, что $z = (0, 0)$, а единичный вектор i_ξ в направлении оси ξ расположен вдоль самого длинного ребра $\hat{S}_z \in \mathcal{S}_z$ (см. Рис. 2.4 (слева)).

Далее введем разбиение $\mathcal{S}_z = \mathring{\mathcal{S}}_z \cup \mathcal{S}_z^+ \cup \mathcal{S}_z^-$, где $\mathring{\mathcal{S}}_z = \{S \subset \mathcal{S}_z : |S| \lesssim h_z\}$ (таким образом $\mathring{\gamma}_z = \mathring{\mathcal{S}}_z \setminus \partial\Omega$). Здесь также используется $\mathcal{S}_z^+ := \{S \subset \mathcal{S}_z \setminus \mathring{\mathcal{S}}_z : S_\xi \subset \mathbb{R}_+\}$, где $S_\xi = \text{proj}_\xi(S)$ является проекцией ребра S на ось ξ . Пусть $(\xi_z^-, \xi_z^+) \ni 0$ будет максимальным интервалом, удовлетворяющим $(\xi_z^-, 0) \subset S_\xi$ при любом $S \in \mathcal{S}_z^-$ и $(0, \xi_z^+) \subset S_\xi$ при любом $S \in \mathcal{S}_z^+$. Также пусть $\varphi_z(\xi)$ обозначает стандартную кусочно-линейную базисную функцию с носителем (ξ_z^-, ξ_z^+) , которая равна 1 при $\xi = 0$. Заметим, что если $\mathcal{S}_z^- = \emptyset$ (и $\mathcal{S}_z^+ = \emptyset$), то будем считать, что $\xi_z^- = 0$ (и $\xi_z^+ = 0$), при этом φ_z при $\xi < 0$ (и $\xi > 0$) не будет использоваться в наших вычислениях.

Некоторые наблюдения, связанные с вышеизложенными определениями, для разных типов узлов представлены в следующей таблице (пока актуальны строки, описывающие \mathcal{S}_z^\pm и ξ_z^\pm).

	$z \in \mathcal{N}_{\text{ani}} \setminus \mathcal{N}_{\partial\Omega}^*$	$z \in \mathcal{N}_{\partial\Omega}^* \subset \mathcal{N}_{\text{ani}}$	$z \in \mathcal{N}_{\text{s,ani}}$	$z \in \mathcal{N}_{\text{iso}}$
\mathcal{S}_z^\pm	$\mathcal{S}_z^-, \mathcal{S}_z^+ \neq \emptyset$	$\mathcal{S}_z^- = \emptyset, \mathcal{S}_z^+ \neq \emptyset$		$\mathcal{S}_z^- = \mathcal{S}_z^+ = \emptyset$
ξ_z^\pm	$ \xi_z^- \simeq \xi_z^+ \simeq H_z$	$\xi_z^- = 0, \xi_z^+ \simeq H_z$		$\xi_z^- = \xi_z^+ = 0$
\bar{g}_z	(2.48) \Rightarrow (2.49)	$\bar{g}_z = 0$	(2.48) \Rightarrow (2.49) $_{ \mathcal{S}_z^- = \emptyset}$	$\bar{g}_z = 0$
I_z'''	(2.54) $\forall z \in \mathcal{N}_{\text{ani}}$		$I_z''' = 0$	

Далее, для $\xi \in [\xi_z^-, \xi_z^+]$ введем непрерывную функцию $\bar{\eta}_z(\xi)$, удовлетворяющую следующим условиям: (i) $\bar{\eta}_z(0) = 0$; (ii) $(\xi, \bar{\eta}_z(\xi)) \in \omega_z$ при любом $\xi \in (\xi_z^-, \xi_z^+)$; (iii) $\bar{\eta}_z(\xi)$ линейна на $[\xi_z^-, 0]$ и $[0, \xi_z^+]$. (Заметим, что, например, можно выбрать $\bar{\eta}_z(\xi)$ таким образом, что отрезок $\{(\xi, \bar{\eta}_z(\xi)) : \xi \in (\xi_z^-, 0)\}$ расположен на любом ребре из \mathcal{S}_z^- , а отрезок $\{(\xi, \bar{\eta}_z(\xi)) : \xi \in (0, \xi_z^+)\}$ расположен на любом ребре из \mathcal{S}_z^+ ; см. Рис. 2.4 (слева).)

Наконец мы можем определить \bar{g}_z (см. также строку для \bar{g}_z в таблице выше). Положим $\bar{g}_z := 0$, если $z \in \mathcal{N}_{\text{iso}}$ является изотропным узлом или $z \in \partial\Omega$. В противном случае, пусть

$$\int_{\xi_z^-}^{\xi_z^+} [g(\xi, \bar{\eta}_z(\xi)) - \bar{g}_z] \varphi_z(\xi) d\xi = 0. \quad (2.48)$$

Также введем $\bar{S}_z^- := \{(\xi, \bar{\eta}_z(\xi)) : \xi \in (\xi_z^-, 0)\}$ и $\bar{S}_z^+ := \{(\xi, \bar{\eta}_z(\xi)) : \xi \in (0, \xi_z^+)\}$, т. е. \bar{S}_z^\pm является отрезком, соединяющим $(0, 0)$ и $(\xi_z^\pm, \bar{\eta}_z(\xi_z^\pm))$. Поэтому, воспользовавшись (2.48), а затем (2.14) и (2.11), получаем

$$h_z |\bar{g}_z| \lesssim \frac{h_z}{H_z} \|g\|_{1; \bar{S}_z^- \cup \bar{S}_z^+} \lesssim \|g\|_{\omega_z} \lesssim \sum_{T \subset \omega_z} \|g\|_T. \quad (2.49)$$

В случае узлов $z \in [\mathcal{N}_{\text{ani}} \setminus \mathcal{N}_{\partial\Omega}^*] \cap \partial\Omega$ и $z \in \mathcal{N}_{\text{s,ani}} \cap \partial\Omega$ формулы (2.48) и (2.49) также приводят к $\bar{g}_z = 0$; для этого $\bar{\eta}_z$ для этих узлов выбирается таким образом, что отрезок $\{(\xi, \bar{\eta}_z(\xi)) : \xi \in (\xi_z^-, \xi_z^+)\}$ расположен на границе области (т. е. $\bar{S}_z^\pm \subset \mathcal{S}_z \cap \partial\Omega$).

Перейдем теперь к оценке I из (2.22). Для этого представим I в виде суммы $I = \sum_{z \in \mathcal{N}} I_z$, где I_z , аналогично (2.31), допускает представление

$$\begin{aligned} I_z &:= \varepsilon^2 \int_{\gamma_z} (g - \bar{g}_z) \phi_z [\nabla u_h] \cdot \nu = I'_z + I''_z + I'''_z + I''''_z \\ &:= \varepsilon^2 \int_{\check{\gamma}_z} (g - \bar{g}_z) \phi_z [\nabla u_h] \cdot \nu + \varepsilon^2 \int_{\gamma_z \setminus \check{\gamma}_z} (g - \bar{g}_z) \phi_z [\partial_\xi u_h] \nu_\xi \\ &\quad + \varepsilon^2 \int_{\gamma_z \setminus \check{\gamma}_z} [g - g(\xi, \bar{\eta}_z(\xi))] \varphi_z [\partial_\eta u_h] \nu_\eta \\ &\quad + \varepsilon^2 \int_{\gamma_z \setminus \check{\gamma}_z} [g(\xi, \bar{\eta}_z(\xi)) - \bar{g}_z] \varphi_z [\partial_\eta u_h] \nu_\eta \\ &\quad + \varepsilon^2 \int_{\gamma_z \setminus \check{\gamma}_z} (g - \bar{g}_z) \{\phi_z - \varphi_z\} [\partial_\eta u_h] \nu_\eta, \end{aligned} \quad (2.50)$$

где, несколько злоупотребляя обозначениями, $g = g(\xi, \eta)$.

Для получения оценки (2.45) начнем с получения с следующего соотношения:

$$|I'_z| + |I''_z| + |I'''_z| \lesssim \left(\varepsilon^2 \sum_{T \subset \omega_z} \|g\|_T \right) \|J_z\|_{\infty; \gamma_z} \quad \text{при } z \in \mathcal{N}. \quad (2.51)$$

Здесь при получении оценки для I_z'' воспользуемся $|\partial_\eta u_h| \leq |J_z|$ и $|\nu_\eta| \leq 1$, а также соотношением $\int_{\gamma_z \setminus \dot{\gamma}_z} |g(\xi, \eta) - g(\xi, \bar{\eta}_z(\xi))| \varphi_z \lesssim \|\partial_\eta g\|_{1; \omega_z} \lesssim \|g\|_{\omega_z}$, объединенным с (2.11).

Получение оценки для I_z' в (2.51) оказывается несколько более сложным. Поскольку $|\phi_z[\nabla u_h]| + |\phi_z[\partial_\xi u_h]| \leq |J_z|$, а $|\nu| \leq 1$, остается показать, что

$$\int_{\dot{\gamma}_z} (|g| + |\bar{g}_z|) + \int_{\gamma_z \setminus \dot{\gamma}_z} (|g \nu_\xi| + |\bar{g}_z \nu_\xi|) \lesssim \sum_{T \subset \omega_z} \|g\|_T. \quad (2.52)$$

Весьма важно, что здесь $|\nu_\xi| \lesssim \frac{h_z}{|S|}$ для любого ребра $S \in \gamma_z \setminus \dot{\gamma}_z$. (Последнее следует из $|\nu_\xi| = |\sin \angle(S, i_\xi)|$ в силу условия локальной ориентации элементов.) Далее, если $S \subset \gamma_z \setminus \dot{\gamma}_z$, а $z \in \mathcal{N}_{\text{ani}} \cup \mathcal{N}_{\text{s,ani}}$, то $|S| \simeq H_z$, и поэтому $|\nu_\xi| \leq \frac{h_z}{H_z}$ (в то время, как при $z \in \mathcal{N}_{\text{iso}}$ оценивать $|\nu_\xi|$ не нужно, поскольку тогда $\gamma_z \setminus \dot{\gamma}_z = \emptyset$). Теперь, воспользовавшись (2.13), получаем приведенные в (2.52) оценки для $|g|$ и $|g \nu_\xi|$. Далее, исключая случай $\bar{g}_z = 0$, получаем $\int_{\dot{\gamma}_z} |\bar{g}_z| + \int_{\gamma_z \setminus \dot{\gamma}_z} |\bar{g}_z \nu_\xi| \lesssim h_z |\bar{g}_z|$. Объединяя это соотношение с (2.49), немедленно получаем оставшиеся оценки для $|\bar{g}_z|$ и $|\bar{g}_z \nu_\xi|$ в (2.52). Наконец, заметим что последние оценки являются тривиальными при $\bar{g}_z = 0$. Таким образом, оценка (2.52) установлена, а следовательно и оценка для I_z' в (2.51).

Оценка же для I_z'''' в (2.51) может быть установлена аналогично оценке для I_z' , при этом воспользуемся тем, что $|\nu_\eta| \leq 1$, а также важным наблюдением, что $0 \leq \phi_z - \varphi_z \lesssim \frac{h_z}{H_z}$ на любом ребре $S \in \gamma_z \setminus \dot{\gamma}_z \subset \mathcal{S}_z^+ \cup \mathcal{S}_z^-$. Установим последнее свойство для случая $S \in \mathcal{S}_z^+$ (поскольку случай $S \in \mathcal{S}_z^-$ рассматривается аналогично). Рассмотрим любой треугольник $T \subset \omega_z$ с двумя ребрами $S', S'' \subset \mathcal{S}_z^+$, т. е. $|S'| \simeq |S''| \simeq H_z$. Тогда, в силу либо (2.6), либо (2.7), получаем $h_T \simeq h_z$, и далее, в силу условия ограниченности максимального угла, $||S'| - |S''|| \lesssim h_z$, и следовательно $||S'_\xi| - |S''_\xi|| \lesssim h_z$. Поскольку аналогичное соотношение может быть получено для любых двух ребер из \mathcal{S}_z^+ , то $|S'_\xi \setminus (0, \xi_z^+)| \lesssim h_z \forall S \in \mathcal{S}_z^+$. Отсюда получаем $0 \leq \phi_z - \varphi_z \lesssim \frac{h_z}{H_z}$ на любом ребре $S \in \mathcal{S}_z^+$. Таким образом, оценка для I_z'''' в (2.51) также установлена.

Остается получить оценку для I_z'''' в (2.50), в которой $[g(\xi, \bar{\eta}_z(\xi)) - \bar{g}_z] \varphi_z$

является функцией ξ , в то время как $\nu_\eta|d\nu| = -\text{sgn}(\xi) d\xi$. Поэтому получаем

$$\begin{aligned} I_z''' &= \varepsilon^2 \left(\sum_{S \in \mathcal{S}_z^- \setminus \partial\Omega} \llbracket \partial_\eta u_h \rrbracket \right) \int_{\xi_z^-}^0 [g(\xi, \bar{\eta}_z(\xi)) - \bar{g}_z] \varphi_z(\xi) d\xi \\ &\quad - \varepsilon^2 \left(\sum_{S \in \mathcal{S}_z^+ \setminus \partial\Omega} \llbracket \partial_\eta u_h \rrbracket \right) \int_0^{\xi_z^+} [g(\xi, \bar{\eta}_z(\xi)) - \bar{g}_z] \varphi_z(\xi) d\xi. \end{aligned} \quad (2.53)$$

Рассмотрим различные типы узлов отдельно (см. также таблицу выше).

Начнем со случая $z \in \mathcal{N}_{\text{iso}}$, для которого $\xi^\pm = 0$ и $\mathcal{S}_z^\pm = \emptyset$, поэтому $I_z''' = 0$. Далее, в случае $z \in \mathcal{N}_{\text{s.ani}}$ имеем $\xi_z^- = 0$ и $\mathcal{S}_z^- = \emptyset$, поэтому объединяя (2.53) с (2.48), немедленно получаем $I_z''' = 0$. В случае же $z \in [\mathcal{N}_{\text{ani}} \setminus \mathcal{N}_{\partial\Omega}^*] \cap \partial\Omega$ справедливо $\bar{g}_z = 0$, в то время как $\bar{\eta}_z$ было определено таким образом, что $(\xi, \bar{\eta}_z(\xi)) \in \partial\Omega$, поэтому $g(\xi, \bar{\eta}_z(\xi)) = 0$; отсюда опять приходим к $I_z''' = 0$.

Для оставшихся узлов $z \in \tilde{\mathcal{N}}_{\text{ani}} = [\mathcal{N}_{\text{ani}} \setminus \partial\Omega] \cup \mathcal{N}_{\partial\Omega}^*$ мы утверждаем, что

$$|I_z'''| \lesssim \left(\varepsilon^2 \sum_{T \subset \omega_z} \|g\|_T \right) H_z h_z^{-1} |\mathcal{J}_z| \quad \text{при } z \in \tilde{\mathcal{N}}_{\text{ani}}. \quad (2.54)$$

Действительно, рассмотрим $z \in \mathcal{N}_{\text{ani}} \setminus \partial\Omega$. Тогда $\mathcal{S}_z^\pm \setminus \partial\Omega = \mathcal{S}_z^\pm$, а $\hat{\mathcal{S}}_z = \hat{\gamma}_z$, поэтому, объединяя (2.53) с определением (2.48) для \bar{g}_z , а также $\sum_{S \in \mathcal{S}_z^- \cup \mathcal{S}_z^+} \llbracket \partial_\eta u_h \rrbracket = -\sum_{S \in \hat{\mathcal{S}}_z} \llbracket \partial_\eta u_h \rrbracket$, получаем

$$I_z''' = \varepsilon^2 \left(\sum_{S \in \hat{\mathcal{S}}_z} \llbracket \partial_\eta u_h \rrbracket \right) \int_0^{\xi_z^+} [g(\xi, \bar{\eta}_z(\xi)) - \bar{g}_z] \varphi_z(\xi) d\xi. \quad (2.55)$$

Поскольку в силу (2.48) и (2.49) справедливо $\int_0^{\xi_z^+} |g(\xi, \bar{\eta}_z(\xi)) - \bar{g}_z| \varphi_z d\xi \lesssim \|g\|_{\hat{\mathcal{S}}_z^+} \lesssim H_z h_z^{-1} \sum_{T \subset \omega_z} \|g\|_T$, то

$$|I_z'''| \lesssim \left(\varepsilon^2 \sum_{T \subset \omega_z} \|g\|_T \right) H_z h_z^{-1} \left| \sum_{S \in \hat{\mathcal{S}}_z} \llbracket \partial_\eta u_h \rrbracket \right|.$$

Здесь $\llbracket \partial_\eta u_h \rrbracket = -J_z \cos \angle(S, i_\xi)$, где $\cos \angle(S, i_\xi) = \cos \angle(S, \hat{S}_z) = \sigma_z|_S$, поэтому получаем (2.54) при $z \in \mathcal{N}_{\text{ani}} \setminus \partial\Omega$. При $z \in \mathcal{N}_{\partial\Omega}^* \subset \mathcal{N}_{\text{ani}}$ оценка (2.54) может быть получена аналогично, только теперь в (2.53) используются $\mathcal{S}_z^- = \emptyset$ и

$\xi_z^- = 0$, а также $\mathcal{S}_z^+ \setminus \partial\Omega = \mathcal{S}_z^+$, что приводит к (2.55) (в котором $\bar{g}_z = 0$); последнее соотношение опять приводит к (2.54), где теперь $\hat{\mathcal{S}}_z = \mathcal{S}_z \cap \partial\Omega$. Таким образом, мы установили (2.54) для любого $z \in \tilde{\mathcal{N}}_{\text{ani}}$.

Объединяя (2.50), (2.51), (2.54) с $I_z''' = 0$ при $z \in \mathcal{N} \setminus \tilde{\mathcal{N}}_{\text{ani}}$, получим желаемую оценку (2.45). \square

З а м е ч а н и е 2.11. Внимательное изучение доказательства оценки (2.45) показывает, что Теорема 2.9 остается справедливой, если локальные декартовы координаты (ξ, η) удовлетворяют $|\angle(\hat{S}_z, i_\xi)| \lesssim \frac{h_z}{H_z}$ (вместо $\angle(\hat{S}_z, i_\xi) = 0$), поскольку в этом случае $\cos \angle(S, i_\xi) = \sigma_z|_S + O(\frac{h_z}{H_z})$ опять приводит к (2.45). Очевидно также, что требование $z = (0, 0)$ в координатах (ξ, η) лишь упрощает изложение, и может быть опущено.

З а м е ч а н и е 2.12. Изучение вышеизложенного доказательства также показывает, что оно остается справедливым, если величины $\{\bar{g}_z\}_{z \in \mathcal{N}}$, которые были определены в (2.48), будут заменены на $\{\bar{g}_z^*\}_{z \in \mathcal{N}}$ при выполнении $\bar{g}_z^* = \bar{g}_z = 0$ при $z \in \mathcal{N}_{\text{iso}} \cup \partial\Omega$, а также при выполнении соотношений (2.37) с $\omega_z^* := \omega_z$. Действительно, тогда I_z будет включать дополнительное слагаемое $I_z^* := \varepsilon^2 \int_{\gamma_z} (\bar{g}_z - \bar{g}_z^*) \phi_z \llbracket \nabla u_h \rrbracket \cdot \nu$, для которого, аналогично Замечанию 2.8, имеем $|I_z^*| \leq \varepsilon^2 \|g\|_{\omega_z} \|J_z\|_{\infty; \gamma_z}$. Объединяя последнее соотношение с (2.11), заключаем, что $|I_z^*|$ ограничено правой частью из (2.51).

З а м е ч а н и е 2.13. В некоторых случаях (например, при $z \in \mathcal{N}_{\text{s,ani}} \setminus \partial\Omega$ при условии, что \mathcal{S}_z^+ содержит не меньше двух ребер, или при $z \in \mathcal{N}_{\text{ani}} \setminus \partial\Omega$ при условии, что каждое из множеств \mathcal{S}_z^\pm содержит не меньше двух ребер) можно выбрать $\bar{\eta}_z(\xi)$ в (2.48) таким образом, что криволинейный прямоугольник $\tilde{\omega}_z := \{\xi \in (\xi_z^-, \xi_z^+), |\eta - \bar{\eta}_z(\xi)| \leq \theta h_z\}$ является подмножеством ω_z . Здесь $\theta > 0$ является фиксированной достаточно малой постоянной. Тогда соотношения (2.37) с $\omega_z^* := \omega_z$ будут выполняться, в частности, при следующем выборе \bar{g}_z^* :

$$\int_{\tilde{\omega}_z} [g(\xi, \eta) - \bar{g}_z^*] \varphi_z(\xi) = 0. \quad (2.56a)$$

Второе соотношение в (2.37) следует из $|\tilde{\omega}_z| = 2\theta h_z(\xi_z^+ - \xi_z^-) \simeq h_z H_z$ и $\tilde{\omega}_z \subset \omega_z$, поэтому действительно получаем $h_z H_z |\bar{g}_z^*| \lesssim \|g\|_{1;\omega_z}$. Далее заметим, что $\int_{\tilde{\omega}_z} [g(\xi, \bar{\eta}_z(\xi)) - g(\xi, \eta)] \varphi_z \simeq h_z H_z (\bar{g}_z - \bar{g}_z^*)$ приводит к $H_z |\bar{g}_z - \bar{g}_z^*| \lesssim \|\partial_\eta g\|_{1;\omega_z} \leq \|g\|_{\omega_z}$, т. е. первое соотношение из (2.37) также выполняется.

В более общем случае, т. е. при $z \in \mathcal{N} \setminus [\mathcal{N}_{\text{iso}} \cup \partial\Omega] = [\mathcal{N}_{\text{ani}} \cup \mathcal{N}_{\text{s,ani}}] \setminus \partial\Omega$, можно определить \bar{g}_z^* , которое будет удовлетворять (2.37) с $\omega_z^* := \omega_z$, следующим образом:

$$\int_{\tilde{\omega}_z} [g(\xi, \eta) - \bar{g}_z^*] = 0, \quad (2.56b)$$

где $\tilde{\omega}_z := \{\xi \in (\xi_z^-, \xi_z^+), |\eta - \bar{\eta}_z(\xi)| \leq \theta h_z \varphi_z(\xi)\}$. При этом используется фиксированная достаточно малая постоянная $\theta > 0$ с тем, чтобы $\tilde{\omega}_z \subset \omega_z \forall z$. При этом первое соотношение из (2.37) устанавливается с использованием $\int_{\tilde{\omega}_z} [g(\xi, \bar{\eta}_z(\xi)) - g(\xi, \eta)] = |\tilde{\omega}_z| (\bar{g}_z - \bar{g}_z^*)$.

З а м е ч а н и е 2.14 (Конечные элементы более высокого порядка). В случае конечных элементов более высокого, чем первый, порядка оценка, аналогичная (2.45), принимает несколько более сложный вид, и требует дополнительных обозначений. Положим $\mathcal{J}_z^\pm(\xi) := \sum_{S \in \mathcal{S}_z^\pm \setminus \partial\Omega} [\partial_\eta u_h]$. В отличие от случая линейных элементов, величина $\mathcal{J}_z^\pm(\xi)$ не является постоянной, поэтому не может быть вынесена за пределы соответствующих интегралов в (2.53). Поэтому мы получаем вариант оценки (2.45), в котором $|\mathcal{J}_z|$ заменяется величиной $|\mathcal{J}_z^-(0) + \mathcal{J}_z^+(0)| + \text{osc}(\mathcal{J}_z^\pm; \omega_z)$ при $z \in \tilde{\mathcal{N}}_{\text{ani}}$ и $\text{osc}(\mathcal{J}_z^\pm; \omega_z)$ при $z \in \mathcal{N}_{\text{s,ani}}$. Заметим также, что вместо поэлементной невязки, равной f_h^I для линейных элементов, теперь используется более общая поэлементная невязка $-\varepsilon^2 \Delta u_h + f_h^I$; в остальном все результаты из §2.6 остаются справедливыми и для случая элементов более высокого порядка.

2.6.3 Поэлементные невязки. Доказательство оценок (2.46) и (2.47)

Д о к а з а т е л ь с т в о (2.46). Данная оценка устанавливается аналогично (2.28) в §2.5.3. Единственное отличие заключается в том, что, если $\bar{g}_z \neq 0$,

то используется соотношение $h_z |\bar{g}_z| \lesssim \|g\|_{\omega_z}$ из (2.49), и поэтому получаем $\|g - \bar{g}_z\|_{1;\omega_z} \lesssim H_z \|g\|_{\omega_z}$. Объединяя последнее с (2.11) приходим к (2.46). \square

Д о к а з а т е л ь с т в о (2.47). Данная оценка устанавливается аналогично (2.28), только определение Θ' объединяется с оценкой $|II| \lesssim \sum_{z \in \mathcal{N}} \|g\|_{1;\omega_z} \|f_h^I\|_{\infty;\omega_z}$. Для последнего используются Замечания 2.12 и 2.13, позволяющие заменить $\{\bar{g}_z\}$ в (2.22) величиной $\{\bar{g}_z^*\}$ либо из (2.56a), либо, в более общем случае, из (2.56b), что, в силу (2.37), приводит к $\|g - \bar{g}_z^*\|_{1;\omega_z} \lesssim \|g\|_{1;\omega_z}$. \square

2.6.4 Дальнейшие предположения о триангуляции. Более точный оценитель

С тем, чтобы получить оценители аналогичные (2.27), (2.30), (2.21), но для более общих триангуляций, будут сделаны некоторые дополнительные предположения. Весьма существенным оказывается, что поэлементные невязки будут оцениваться не для отдельных узлов, а для цепочек анизотропных или полу-анизотропных узлов, объединенных короткими ребрами. Для этого воспользуемся следующим определением (см. также Рис. 2.4, справа).

О п р е д е л е н и е. *(Полу-)Анизотропной Цепочкой*, или просто *Цепочкой*, будем называть упорядоченную последовательность узлов $\{z_j\}_{j=1}^k$ из $\mathcal{N}_{\text{ani}} \setminus \mathcal{N}_{\partial\Omega}^*$ (либо же из $\mathcal{N}_{\text{s,ani}}$), где каждый узел z_j соединяется с узлом z_{j+1} , при $j = 1, \dots, k-1$, ребром длины $\simeq h_{z_j} \simeq h_{z_{j+1}}$, а любой конечный узел цепочки (т. е. z_l при $l = 1, k$) либо расположен на $\partial\Omega$, либо же соединен ребром длины $\simeq h_{z_l}$ с изотропным узлом z для которого $H_z \simeq H_{z_l}$. (Например, если узел $z_1 \notin \partial\Omega$, то он соединен с $z_0 \in \mathcal{N}_{\text{iso}}$ ребром длины $\simeq h_{z_1}$ и $H_{z_0} \simeq H_{z_1}$.)

Пусть $\mathcal{N}_i \subset \mathcal{N}_{\text{ani}} \setminus \mathcal{N}_{\partial\Omega}^*$ является анизотропной цепочкой при $i = 1, \dots, n_{\text{ani}}$, а $\mathcal{N}_i \subset \mathcal{N}_{\text{s,ani}}$, является семи-анизотропной цепочкой при $i = n_{\text{ani}} + 1, \dots, n_{\text{ani}} + n_{\text{s,ani}}$ (без верхних ограничений на n_{ani} или $n_{\text{s,ani}}$). Также введем подмножество узлов $\mathcal{N}_{\text{paths}} := \cup_{i=1}^{n_{\text{ani}} + n_{\text{s,ani}}} \mathcal{N}_i$. Заметим, что $\mathcal{N} \setminus \mathcal{N}_{\text{paths}}$ может включать

(полу-)анизотропные узлы, не принадлежащие ни одной цепочке.

А3. *Локальная ориентация цепочки.* Для каждой (полу-)анизотропной цепочки \mathcal{N}_i , $i = 1, \dots, n_{\text{ani}} + n_{\text{s,ani}}$, можно ввести декартовы координаты $(\xi, \eta) = (\xi_i, \eta_i)$, при которых справедливо $|\sin(\angle(S, i_\xi))| \lesssim \frac{h_z}{|S|}$ для любого ребра $S \subset \mathcal{S}_z$, исходящего из любого узла $z \in \mathcal{N}_i$.

Заметим, что из условия **А3** следует, что для любого $z \in \mathcal{N}_i$, $i \leq n_{\text{ani}} + n_{\text{s,ani}}$, с использованием обозначения $\Omega_i := \cup_{z \in \mathcal{N}_i} \omega_z$, справедливы соотношения

$$\omega_z \subseteq \omega_z^* := \Omega_i \cap \{\eta_i \in (\eta_z^-, \eta_z^+)\}, \quad \eta_z^+ - \eta_z^- \simeq h_z, \quad \text{diam } \omega_z \simeq \text{diam } \omega_z^* \simeq H_z, \quad (2.57)$$

где (η_z^-, η_z^+) — область значений η в ω_z . Отметим также, что в силу условия ограниченности максимального угла справедливо $\omega_z^* \subset \omega_z^{(J)}$ при некотором $J \lesssim 1$ (где $\omega_z^{(J)}$ определено в §2.5).

Т е о р е м а 2.15. *При выполнении условий Теоремы 2.9, а также условия **А3** (при этом **А2** не предполагается), при любом $(x', y') \in \Omega$ справедливо представление $(u_h - u)(x', y') = I + II + \mathcal{E}_{\text{quad}}$, где I удовлетворяет (2.45), $\mathcal{E}_{\text{quad}}$ удовлетворяет (2.21), а*

$$|II| \lesssim \max_{z \in \mathcal{N}_{\text{paths}}} \{ (\Theta + \Theta') \varepsilon^{-2} \lambda_z \min\{\varepsilon, h_z\} \|f_h^I\|_{\infty; \omega_z} + \Theta' \varepsilon^{-2} \lambda_z^2 \text{osc}(f_h^I; \omega_z) \} + \Theta' \max_{z \in \mathcal{N} \setminus \mathcal{N}_{\text{paths}}} \{ \varepsilon^{-2} \lambda_z^2 \|f_h^I\|_{\infty; \omega_z} \}. \quad (2.58)$$

При выполнении дополнительного условия **А2** для I также справедлива оценка (2.44).

Доказательству данной теоремы посвящены §2.6.5 и §2.6.6.

2.6.5 Выбор \bar{g}_z для более точного оценителя. Невязки, ассоциируемые с ребрами

При выводе более точного оценителя (2.58) для поэлементных невязок, нам придется несколько изменить определение (2.48) для величин \bar{g}_z , соот-

ветсвующих узлам некоторой цепочки $\mathcal{N}_i = \{z_j\}_{j=1}^k$ (где $k = k(i)$) следующим образом. Для каждого узла $z_j \in \mathcal{N}_i$ в качестве локальных декартовых координат (ξ, η) , описанных в §2.6.2 и затем используемых в (2.48), теперь будем использовать (ξ_i, η_i) . (Другими словами, раньше координаты (ξ, η) выбирались независимо для каждого узла z , теперь же $(\xi, \eta) = (\xi_i, \eta_i)$ — одни и те же для всех узлов цепочки).

В силу Замечания 2.11 и условия $\mathcal{A}3$, все выводы Теоремы 2.9, за исключением (2.47), остаются справедливыми при вышеописанном несколько измененном определении (ξ, η) .

З а м е ч а н и е 2.16. Для того, чтобы выполнялись все оценки Теоремы 2.9, включая (2.47), достаточно заменить величины $\{\bar{g}_z\}$ в (2.22) на $\{\bar{g}_z^*\}$, удовлетворяющие (2.37), где ω_z^* выбирается в соответствии с (2.57) и удовлетворяет $\omega_z^* \subset \omega_z^{(J)}$ for $z \in \mathcal{N}_{\text{paths}}$, в то время, как $\omega_z^* := \omega_z$, а величины \bar{g}_z^* описаны в Замечании 2.12 при $z \in \mathcal{N} \setminus \mathcal{N}_{\text{paths}}$. Таким образом, при $z \in \mathcal{N} \setminus \mathcal{N}_{\text{paths}}$ пусть $\{\bar{g}_z^*\}$ выбираются как описано в Замечании 2.13 (т. е. $\bar{g}_z^* = \bar{g}_z = 0$ при $z \in \mathcal{N}_{\text{iso}} \cup \partial\Omega$, а в противном случае используется либо (2.56a), либо (2.56b)). При $z \in \mathcal{N}_{\text{paths}}$ необходимые нам величины \bar{g}_z^* будут описаны в §2.6.6.

2.6.6 Поэлементные невязки. Доказательство оценки (2.58)

Д о к а з а т е л ь с т в о. Начнем с замены величин $\{\bar{g}_z\}$ в (2.22) на $\{\bar{g}_z^*\}$, описанные в Замечании 2.16 (при этом \bar{g}_z^* для узлов $z \in \mathcal{N}_{\text{paths}}$ будут выбраны ниже). Далее, поскольку $\mathcal{N} = \left(\bigcup_{i=1}^{n_{\text{ani}}+n_{\text{s,ani}}} \mathcal{N}_i\right) \cup (\mathcal{N} \setminus \mathcal{N}_{\text{paths}})$, представим Π из

(2.22) в виде суммы $\Pi = \sum_i \Pi_i + \Pi_{\text{osc}} + \Pi_{\setminus \text{paths}}$, где

$$\begin{aligned}\Pi_i &:= \sum_{z \in \mathcal{N}_i} \int_{\omega_z} f_h^I(X_z, Y_z) (g - \bar{g}_z^*) \phi_z, \\ \Pi_{\text{osc}} &:= \sum_{z \in \mathcal{N}_{\text{paths}}} \int_{\omega_z} [f_h^I - f_h^I(X_z, Y_z)] (g - \bar{g}_z^*) \phi_z, \\ \Pi_{\setminus \text{paths}} &:= \sum_{z \in \mathcal{N} \setminus \mathcal{N}_{\text{paths}}} \int_{\omega_z} f_h^I (g - \bar{g}_z^*) \phi_z.\end{aligned}$$

Здесь все интегралы понимаются в исходных переменных (x, y) , а также используются некоторые $(X_z, Y_z) = (X_z(x, y), Y_z(x, y)) \in \omega_z^*$, которые будут выбраны ниже для $(x, y) \in \omega_z$ при $z \in \mathcal{N}_{\text{paths}}$. Для Π_{osc} и $\Pi_{\setminus \text{paths}}$, мы немедленно получаем оценки, аналогичные (2.47) (в силу Замечания 2.16):

$$|\Pi_{\text{osc}}| \lesssim \Theta' \max_{z \in \mathcal{N}_{\text{paths}}} \left\{ \varepsilon^{-2} \lambda_z^2 \text{osc}(f_h^I; \omega_z^*) \right\}, \quad (2.59)$$

$$|\Pi_{\setminus \text{paths}}| \lesssim \Theta' \max_{z \in \mathcal{N} \setminus \mathcal{N}_{\text{paths}}} \left\{ \varepsilon^{-2} \lambda_z^2 \|f_h^I\|_{\infty; \omega_z} \right\}. \quad (2.60)$$

Чтобы завершить доказательство, достаточно показать, что при $i = 1, \dots, n_{\text{ani}}$

$$\begin{aligned}|\Pi_i| &\lesssim \left(\sum_{z \in \mathcal{N}_i} \lambda_z^{-1} \|g\|_{\omega_z^*} + \sum_{z \in \mathcal{N}_i} \lambda_z^{-2} \|g\|_{1; \omega_z^*} \right) \max_{z \in \mathcal{N}_i} \left\{ \min\{\lambda_z^2, \lambda_z h_z\} \|f_h^I\|_{\infty; \omega_z} \right\} \\ &\quad + \left(\sum_{z \in \mathcal{N}_i} \lambda_z^{-2} \|g\|_{1; \omega_z^*} \right) \left[\max_{z \in \mathcal{N}_i} \left\{ \lambda_z^2 \text{osc}(f_h^I; \omega_z^*) \right\} + \max_{z \in \mathcal{N} \setminus \mathcal{N}_{\text{paths}}} \left\{ \lambda_z^2 \|f_h^I\|_{\infty; \omega_z} \right\} \right],\end{aligned} \quad (2.61)$$

а при $i = n_{\text{ani}} + 1, \dots, n_{\text{ani}} + n_{\text{s,ani}}$ справедлив вариант оценки (2.61), в котором $\|g\|_{\omega_z^*}$ заменено на $\|g\|_{\omega_z^*} + \sum_{T \subset \omega_z} \|g\|_T$. Действительно, желаемая оценка (2.58) следует из (2.59) и (2.60), объединенных с наблюдением, что сумма $\sum_i |\Pi_i|$ ограничена правой частью из (2.58). Последнее соображение может быть установлено объединением оценок (2.61) при всех i с соотношением $\min\{\lambda_z^2, \lambda_z h_z\} = \lambda_z \min\{\varepsilon, h_z\}$, а затем с (2.11) и (2.43). Также используется $\omega_z^* \subset \omega_z^{(J)}$, что позволяет заменить ω_z^* в (2.61) на ω_z .

Перейдем к доказательству оценки (2.61). Рассмотрим Π_i при некотором $i = 1, \dots, n_{\text{ani}}$, что соответствует анизотропной цепочке $\mathcal{N}_i = \{z_j\}_{j=1}^k$ (с

$k = k(i)$). Вспомним, что $\Omega_i := \cup_{z \in \mathcal{N}_i} \omega_z$, и положим $\Omega_i^* := \Omega_i \setminus (\omega_{z_1}^* \cup \omega_{z_k}^*)$. В пределах Ω_i^* будут использоваться локальные координаты $(\xi, \eta) = (\xi_i, \eta_i)$, ассоциируемые с данной цепочкой, в соответствии с описанием в §2.6.5. Заметим, что в силу условия \mathcal{A}_3 , а также условия ограниченности максимального угла, зигзагообразная кривая, состоящая из ребер, соединяющих узлы в цепочке \mathcal{N}_i , может быть описана как $\xi = \kappa_i(\eta)$, а зигзагообразные кривые, описывающие $\partial\Omega_i \setminus [\partial\omega_{z_1} \cup \partial\omega_{z_k}]$, могут быть описаны как $\xi = \kappa_i^\pm(\eta)$. Таким образом, Ω_i^* является криволинейным прямоугольником, ограниченным кривыми $\xi = \kappa_i^\pm(\eta)$ и линиями $\eta = \eta_{z_1}^+$ и $\eta = \eta_{z_k}^-$ (предполагая, без ограничения общности, что η возрастает по мере продвижения по цепочке \mathcal{N}_i из z_1 к z_k). Заметим также, что $|\frac{d}{d\eta}\kappa_i| + |\frac{d}{d\eta}\kappa_i^\pm| \lesssim 1$, и поэтому

$$\begin{aligned} \pm[\kappa_i^\pm(\eta) - \kappa_i(\eta)] &\simeq H_z \quad \text{при } \eta \in (\eta_{z_1}^-, \eta_{z_1}^+) \cup (\eta_{z_k}^+, \eta_{z_k}^-), \\ \text{osc}(\kappa(\eta); \omega_z^*) + \text{osc}(\kappa^\pm(\eta); \omega_z^*) &\lesssim h_z, \end{aligned} \quad (2.62)$$

где также было принято во внимание (2.57).

Поскольку $\omega_z = (\omega_z \cap \omega_{z_1}^*) \cup (\omega_z \cap \omega_{z_k}^*) \cup (\omega_z \cap \Omega_i^*)$, воспользуемся представлением $\Pi_i = \Pi_i^{(1)} + \Pi_i^{(k)} + \Pi_i^*$, где слагаемые $\Pi_i^{(1)}$ и $\Pi_i^{(k)}$ соответствуют областям $\omega_z \cap \omega_{z_1}^*$ и $\omega_z \cap \omega_{z_k}^*$, а Π_i^* соответствует $\omega_z \cap \Omega_i^*$. Далее, в каждой области $\omega_z \cap \omega_{z_l}^*$ определим $(X_z(x, y), Y_z(x, y)) := (x, y)$, что приводит к

$$\Pi_i^{(l)} = \sum_{z \in \mathcal{N}_i} \int_{\omega_z \cap \omega_{z_l}^*} f_h^l(g - \bar{g}_z^*) \phi_z, \quad \text{при } l = 1, k. \quad (2.63)$$

Слагаемое же Π_i^* перепишем в переменных (ξ, η) . При этом $(X_z(x, y), Y_z(x, y))$ становится некоторым $(\Psi_z(\xi, \eta), \Upsilon_z(\xi, \eta))$, которое полагается в каждой области $\omega_z \cap \Omega_i^*$ равным $(\Psi_z, \Upsilon_z) := (\kappa_i(\eta), \eta) \in \omega_z^*$. Таким образом, несколько злоупотребляя обозначениями, получаем

$$\Pi_i^* = \sum_{z \in \mathcal{N}_i} \int_{\eta_{z_1}^+}^{\eta_{z_k}^-} f_h^I(\kappa_i(\eta), \eta) \int_{\kappa_i^-(\eta)}^{\kappa_i^+(\eta)} (g - \bar{g}_z^*) \phi_z d\xi d\eta, \quad (2.64)$$

где $g = g(\xi, \eta)$ и $\phi_z = \phi_z(\xi, \eta)$.

Без потери общности, пусть $z_k \in \partial\Omega$, а z_1 соединяется с некоторым $z_0 \in \mathcal{N}_{\text{iso}}$ ребром длины $\simeq h_{z_1}$ with $H_{z_0} \simeq H_{z_1}$. Введем $\mathcal{N}_i^{(l)} := \{z \in \mathcal{N}_i : \omega_z \cap \omega_{z_l}^* \neq \emptyset\}$ при $l = 1, k$, а затем введем $\omega_z^{**} := \omega_z^* \cup \omega_{z_k}$ при $z \in \mathcal{N}_i^{(k)}$, и $\omega_z^{**} := \omega_z^* \cup \omega_{z_1}$ при $z \in \mathcal{N}_i^{(1)} \setminus \mathcal{N}_i^{(k)}$. Заметим, что ω_z^{**} удовлетворяет (2.57), а также $\omega_z^{**} \subset \omega_z^{(2J)}$. Следовательно, переопределение $\omega_z^* := \omega_z^{**}$ при $z \in \mathcal{N}_i^{(1)} \cup \mathcal{N}_i^{(k)}$ не повлияет на вышеприведенные доказательства. Наконец, определим величины \bar{g}_z^* при $z \in \mathcal{N}_i$ следующим образом. Положим $\bar{g}_z^* = 0$ при $z \in \mathcal{N}_i^{(k)}$ (при этом, как будет показано ниже, (2.37) остается справедливым). В противном случае, если $z \in \mathcal{N}_i \setminus [\mathcal{N}_i^{(1)} \cup \mathcal{N}_i^{(k)}]$, то воспользуемся определением (2.56a) для \bar{g}_z^* , в котором $\tilde{\omega}_z := (\xi_z^-, \xi_z^+) \times (\eta_z^-, \eta_z^+)$; если же $z \in \mathcal{N}_i^{(1)} \setminus \mathcal{N}_i^{(k)}$, то воспользуемся определением (2.56b) для \bar{g}_z^* , в котором $\tilde{\omega}_z := \tilde{\omega}_{z_1}$, где $\tilde{\omega}_{z_1} \subset \omega_{z_1}$ выбирается, как описано в Замечании 2.13. Заметим, что при $z \in \mathcal{N}_i \setminus \mathcal{N}_i^{(k)}$ второе соотношение в (2.37) следует из $\tilde{\omega}_z \subset \omega_z^*$, в то время как первое соотношение может быть получено с помощью соображений, аналогичных использованным в Замечании 2.13.

Далее рассмотрим $\Pi_i^{(1)}$ и $\Pi_i^{(k)}$ из (2.63). Для $\Pi_i^{(1)}$, воспользовавшись вторым соотношением из (2.37), несложно убедиться, что

$$|\Pi_i^{(1)}| \lesssim \left(\sum_{z \in \mathcal{N}_i^{(1)}} \lambda_z^{-2} \|g\|_{\omega_z^*} \right) \max_{z \in \mathcal{N}_i} \left\{ \lambda_z^2 \|f_h^I\|_{\infty; \omega_z \cap \omega_{z_1}^*} \right\}. \quad (2.65)$$

Здесь $\|f_h^I\|_{\infty; \omega_{z_1}^*} \leq \|f_h^I\|_{\infty; \omega_{z_0}} + \text{osc}(f_h^I; \omega_{z_1}^*)$, в то время как $H_z \simeq H_{z_1} \simeq H_{z_0}$, и поэтому $\lambda_z \simeq \lambda_{z_1} \simeq \lambda_{z_0}$ в пределах $\omega_{z_1}^*$. Таким образом,

$$\max_{z \in \mathcal{N}_i} \left\{ \lambda_z^2 \|f_h^I\|_{\infty; \omega_z \cap \omega_{z_1}^*} \right\} \lesssim \lambda_{z_0}^2 \|f_h^I\|_{\infty; \omega_{z_0}} + \lambda_{z_1}^2 \text{osc}(f_h^I; \omega_{z_1}^*). \quad (2.66)$$

Объединяя (2.66) с (2.65), а также в силу $z_0 \in \mathcal{N}_{\text{iso}} \subset \mathcal{N} \setminus \mathcal{N}_{\text{paths}}$ и $z_1 \in \mathcal{N}_i$, заключаем, что $|\Pi_i^{(1)}|$ ограничено правой частью из (2.61).

Перейдем к получению аналогичной оценки (2.61) для $\Pi_i^{(k)}$. Заметим, что (2.63) приводит к

$$|\Pi_i^{(k)}| \lesssim \left[\sum_{z \in \mathcal{N}_i^{(k)}} (\lambda_z^{-2} + \lambda_z^{-1} h_z^{-1}) \|g - \bar{g}_z^*\|_{1; \omega_z \cap \omega_{z_k}^*} \right] \max_{z \in \mathcal{N}_i} \left\{ \min\{\lambda_z^2, \lambda_z h_z\} \|f_h^I\|_{\infty; \omega_z} \right\}, \quad (2.67)$$

где $\bar{g}_z^* = 0$. Поскольку $h_z \simeq h_{z_k}$, $H_z \simeq H_{z_k}$ и $\lambda_z \simeq \lambda_{z_k}$ в пределах $\omega_{z_k}^*$, то сумма \sum в (2.67) является $\lesssim (\lambda_{z_k}^{-2} + \lambda_{z_k}^{-1} h_{z_k}^{-1}) \|g\|_{1; \omega_{z_k}^*}$. Объединяя это с важным наблюдением (которое будет объяснено ниже), что

$$\|g\|_{1; \omega_{z_k}^*} \lesssim h_{z_k} \|\nabla g\|_{1; \omega_{z_k}^*} \lesssim h_{z_k} \|g\|_{\omega_{z_k}^*}, \quad (2.68)$$

приходим к заключению, что сумма \sum в (2.67) является $\lesssim \lambda_{z_k}^{-2} \|g\|_{1; \omega_{z_k}^*} + \lambda_{z_k}^{-1} \|g\|_{\omega_{z_k}^*}$. Последнее соображение приводит к желаемой оценке (2.61) для $II_i^{(k)}$. Обсудим теперь (2.68), в котором второе неравенство очевидно. Первое же неравенство в (2.68) справедливо в силу $z_k \in \partial\Omega$, а также $|\mathcal{S}_{z_k} \cap \partial\Omega| \simeq H_{z_k}$. Более подробно, здесь используется тот факт, что $g = 0$ на $\partial\Omega$, а также, что все точки области $\omega_{z_k}^*$ находятся в пределах $\lesssim h_{z_k}$ от $\mathcal{S}_{z_k} \cap \partial\Omega$. Воспользовавшись (2.49), аналогичным образом можно также показать, что $H_z |\bar{g}_z| \lesssim \|g\|_{1; \bar{\mathcal{S}}_z^- \cup \bar{\mathcal{S}}_z^+} \lesssim \|\nabla g\|_{1; \omega_z^*}$ при $z \in \mathcal{N}_i^{(k)}$; последнее обозначает, что наш выбор $\bar{g}_z^* = 0$ при $z \in \mathcal{N}_i^{(k)}$ удовлетворяет (2.37). Таким образом, мы установили, что $|II_i^{(k)}|$ также ограничено правой частью из (2.61).

Для завершения доказательства оценки (2.61) для II_i остается показать, что аналогичная оценка верна и для II_i^* , описанного в (2.64). Введем функцию $\varphi_i(\xi, \eta) := \sum_{z \in \mathcal{N}_i} \phi_z$ и заметим, что при любом $\eta \in (\eta_{z_1}^+, \eta_{z_k}^-)$

$$\sum_{z \in \mathcal{N}_i} \int_{\kappa_i^-(\eta)}^{\kappa_i^+(\eta)} g \phi_z d\xi = \int_{\kappa_i^-(\eta)}^{\kappa_i^+(\eta)} g \varphi_i d\xi =: \hat{g}_i(\eta) \int_{\kappa_i^-(\eta)}^{\kappa_i^+(\eta)} \varphi_i d\xi = \sum_{z \in \mathcal{N}_i} \int_{\kappa_i^-(\eta)}^{\kappa_i^+(\eta)} \hat{g}_i(\eta) \phi_z d\xi. \quad (2.69)$$

Здесь центральное соотношение служит определением для $\hat{g}_i(\eta)$. Далее, можем переписать (2.64) в виде

$$\begin{aligned} II_i^* &= \sum_{z \in \mathcal{N}_i} \int_{\eta_{z_1}^+}^{\eta_{z_k}^-} f_h^I(\kappa_i(\eta), \eta) \int_{\kappa_i^-(\eta)}^{\kappa_i^+(\eta)} (\hat{g}_i(\eta) - \bar{g}_z^*) \phi_z d\xi d\eta \\ &= \sum_{z \in \mathcal{N}_i} \int_{\omega_z \cap \Omega_i^*} f_h^I(\kappa_i(\eta), \eta) (\hat{g}_i(\eta) - \bar{g}_z^*) \phi_z. \end{aligned}$$

Отсюда, после некоторых вычислений, которые будут прокомментированы

ниже, приходим к оценке

$$\begin{aligned}
|II_i^*| &\lesssim \sum_{z \in \mathcal{N}_i} \|\hat{g}_i(\eta) - \bar{g}_z^*\|_{1; \omega_z \cap \Omega_i^*} \|f_h^I\|_{\infty; \omega_z^*} \lesssim \sum_{z \in \mathcal{N}_i} \min\left\{\|g\|_{1; \omega_z^*}, h_z \|g\|_{\omega_z^*}\right\} \|f_h^I\|_{\infty; \omega_z^*} \\
&\lesssim \left(\sum_{z \in \mathcal{N}_i} \lambda_z^{-2} \|g\|_{1; \omega_z^*} + \sum_{z \in \mathcal{N}_i} \lambda_z^{-1} \|g\|_{\omega_z^*} \right) \max_{z \in \mathcal{N}_i} \left\{ \min\{\lambda_z^2, \lambda_z h_z\} \|f_h^I\|_{\infty; \omega_z^*} \right\}.
\end{aligned} \tag{2.70}$$

Здесь была использована оценка $\|\hat{g}_i(\eta) - \bar{g}_z^*\|_{1; \omega_z \cap \Omega_i^*} \lesssim \|g\|_{1; \omega_z^*}$ (которая немедленно следует из определения \hat{g}_i , а также из второго соотношения в (2.37)). Также мы воспользовались оценкой $\|\hat{g}_i(\eta) - \bar{g}_z^*\|_{1; \omega_z \cap \Omega_i^*} \lesssim h_z \|g\|_{\omega_z^*}$, которая может быть получена объединением $|\omega_z| \simeq h_z H_z$ с соотношением $|\bar{g}_z - \bar{g}_z^*| \lesssim H_z^{-1} \|g\|_{\omega_z^*}$ из (2.37), а также с важной оценкой

$$\|\hat{g}_i(\eta) - \bar{g}_z\|_{\infty; \omega_z \cap \Omega_i^*} \lesssim H_z^{-1} \|g\|_{\omega_z^*}. \tag{2.71}$$

Чтобы получить (2.71), достаточно показать, что $|\hat{g}_i(\eta) - \bar{g}_z| \lesssim H_z^{-1} \|g\|_{\omega_z^*}$ при $\eta \in (\eta_z^-, \eta_z^+) \cap (\eta_{z_1}^+, \eta_{z_k}^-)$. Последнее устанавливается аналогично оценке для $\hat{g}_i(y) - \bar{g}_z^*$ в доказательстве (2.30) в §2.5.3 (см. также более подробное изложение в [168, §7.6]). Наконец, (2.70) приводит к желаемой оценке типа (2.61) для II_i^* .

Заметим, что оценка для II_i при $i = n_{\text{ani}} + 1, \dots, n_{\text{ani}} + n_{\text{s,ani}}$, что соответствует полу-анизотропным цепочкам \mathcal{N}_i , получается аналогичным образом (что более подробно изложено в [168, §7.6]). Таким образом, необходимые оценки типа (2.58) получены, что завершает доказательство. \square

2.6.7 Численные результаты

Воспользуемся той же тестовой задачей, что и в §2.5.4. Элементы сетки маркируются с использованием варианта оценщика из Теоремы 2.15, который определяется как $\mathcal{E} := \max\{\mathcal{E}_{(2.44)}, \mathcal{E}_{(2.58)}, \mathcal{E}_{(2.21)}\}$. Здесь $\mathcal{E}_{(2.\cdot)}$ соответствуют правой части оценки (2. \cdot), в которой мы полагаем $\Theta = \Theta' = 1$ (в то время

как, в Лемме 2.17 будет установлено, что $\Theta + \Theta' \lesssim \ell_h$). Также все величины вида $\min\{1, \varepsilon^{-1}a\}$ заменены их более гладкими версиями $\frac{a}{\varepsilon+a}$. Таким образом, например, $\mathcal{E}_{(2.44)} := \max_{T \in \mathcal{T}} \left\{ \max_{z \in \mathcal{N}} \left[\frac{\varepsilon H_z}{\varepsilon + H_z} \|J_z\|_{\infty; \gamma_z \cap \partial T} \right] \right\}$. Аналогично, $\mathcal{E}'_{(2.58)} := \max_{T \in \mathcal{T}} \left\{ \frac{h_T H_T}{(\varepsilon + h_T)(\varepsilon + H_T)} \|f_h^I\|_{\infty; T} \right\}$ и $\mathcal{E}''_{(2.58)} := \max_{T \in \mathcal{T}} \left\{ \left(\frac{H_T}{\varepsilon + H_T} \right)^2 \text{osc}(f_h^I; T) \right\}$ формируют оценитель $\mathcal{E}_{(2.58)} := \max\{\mathcal{E}'_{(2.58)}, \mathcal{E}''_{(2.58)}\}$. (Заметим, что в наших экспериментах нет необходимости различать анизотропные и неанизотропные элементы сетки явным образом в используемых оценителях, поскольку для неанизотропных элементов справедливо $h_T \simeq H_T$, а также $\text{osc}(f_h^I; T) \leq \|f_h^I\|_{T, \infty}$.) В используемый оценитель мы также включили $\mathcal{E}_{(2.21)} := \max_{T \in \mathcal{T}} H_T^2$ в качестве несколько грубого варианта для квадратурной ошибки (2.21) (заметим, что это компонента оценителя не играет существенной роли в процессе адаптивного измельчения сетки).

В каждом эксперименте мы начинали с равномерной сетки, состоящей из прямоугольных треугольников диаметра $H_T = 2^{-8}, 2^{-16}, 2^{-32}$ с аспектным соотношением $\frac{H_T}{h_T} = 2$. На каждой итерации маркировались элементы, отвечающие по крайней мере за 5% глобального оценителя, при этом маркировалось не более 15% всех элементов. Промаркированные элементы измельчались только в направлении оси x с использованием так называемого зеленого

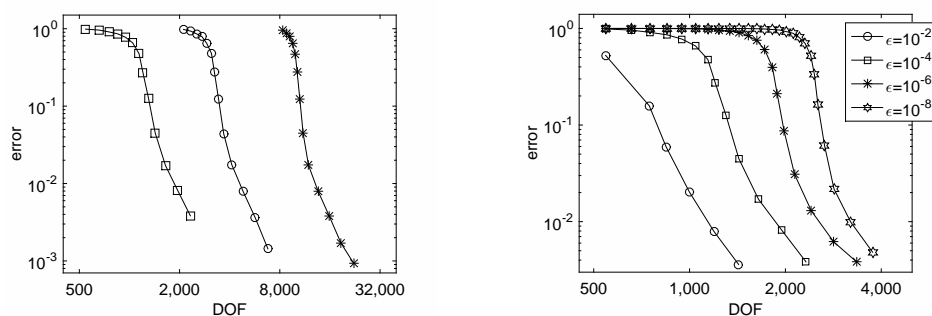


Рис. 2.5: Ошибка численного решения в норме максимума модуля при $\varepsilon = 10^{-4}$ и различных начальных значениях DOF (слева), и при различных ε (справа).

(green) измельчения. Для улучшения геометрических свойств сетки, а также с целью уменьшения величины $\max_{T \in \mathcal{T}} \{\text{osc}(f_h^I; T)\}$, для пар треугольников с общим ребром применялась процедура перебрасыванию диагоналей (edge swapping). Итерации продолжались до тех пор, пока не выполнялось условие $\text{osc}(f_h^I; T) \leq 4H_T$ для всех треугольников T . На Рис. 2.5 приведены значения ошибки в норме максимума модуля $\|u - u_h\|_{\infty; \Omega}$ по отношению к соответствующим числам степеней свободы (обозначенным DOF) при фиксированном $\varepsilon = 10^{-4}$ (слева) и при нескольких значениях ε (справа). Мы наблюдаем весьма быстрое уменьшение ошибки в процессе анизотропного измельчения сетки (по сравнению с локально квазиравномерным измельчением сетки в §1.5). Заметим также, что аспектные соотношения для элементов сетки принимали значения между 2 и $3.35e+7$, в то время как индексы эффективности не превосходили 85 во всех экспериментах. Принимая все вышесказанное во внимание (а также наблюдения по сходимости в норме максимума модуля на анизотропных сетках, описанные в [167]), заключаем, что полученный нами оценщик работает достаточно хорошо при весьма высоких аспектных соотношениях элементов сетки.

2.7 Интерполяция функции Грина

Мы получили ряд апостериорных оценок ошибки численного решения, но пока в всех этих оценках присутствуют величины Θ и Θ' , которые связаны с интерполяционной ошибкой для функции Грина, и которые нам еще предстоит оценить. Заметим, что хотя мы использовали несколько разные определения (2.26) и (2.43) для Θ и Θ' , но присутствуют ли эти величины в Теореме 2.7, или же в Теоремах 2.9 и 2.15, для них справедлива оценка

$$\Theta + \Theta' \lesssim \bar{\Theta} := \varepsilon^2 \sum_{T \in \mathcal{T}} \left(\lambda_T^{-1} \|\nabla g\|_{1;T} + \lambda_T^{-2} \|g\|_{1;T} \right), \quad \lambda_T := \min\{\varepsilon, H_T\}, \quad (2.72)$$

(в силу (2.11), а также Замечания 2.5). Вспомним также, что здесь используется $g = G(x', y'; \cdot) - G_h$ с любым интерполянтом $G_h \in S_h$.

Для оценки $g = G - G_h$ в $\bar{\Theta}$ было бы очень удобно воспользоваться стандартными квазиинтерполянтами, удовлетворяющими $|G - G_h|_{k,p;T} \lesssim H_T^{j-k} |G|_{j,p;\omega_T}$ при любых $0 \leq k \leq j \leq 2$, $1 \leq p \leq \infty$ (как это сделано в §1 для случая локально квазирваномерных сеток). Однако в случае неструктурированных анизотропных сеток построение таких интерполянтов является весьма сложной задачей, и потребовало бы дополнительных предположений о триангуляции. (Этот вопрос обсуждается, например, в Главе 3 монографии [56] для случая прямоугольных анизотропных сеток). Поэтому мы подойдем к построению G_h с несколько другой стороны и воспользуемся менее стандартным интерполянтом G_h , который, грубо говоря, будет совпадать с интерполянтом Лагранжа при $H_T \lesssim \varepsilon$ и будет равным 0 при $H_T \gtrsim \varepsilon$. Заметим, что построение данного интерполянта потребует дополнительного, хотя и не слишком жесткого, предположения о триангуляции; см. Замечание 2.18 ниже.

Для любой точки $(x', y') \in \Omega$, положим $\omega_{(x',y')} := \omega_{z'}$, либо $\omega_{(x',y')} := \omega_{S'}$, либо же $\omega_{(x',y')} := \omega_{T'}$, если, соответственно, $(x', y') = z'$ при некотором $z' \in \mathcal{N}$, либо $(x', y') \in S'$ при некотором $S' \in \mathcal{S}$, либо же $(x', y') \in T'$ при некотором $T' \in \mathcal{T}$. Здесь мы используем достаточно стандартные обозначения ω_z , ω_S и ω_T для множеств элементов сетки, соответственно окружающих $z \in \mathcal{N}$, $S \in \mathcal{S}$ и $T \in \mathcal{T}$. Также, пусть

$$\bar{h}_{(x',y')} := \min_{T \subset \omega_{(x',y')}} h_T, \quad \ell_{h,(x',y')} := \ln(2 + \varepsilon^{-1} \bar{h}_{(x',y')}). \quad (2.73)$$

Очевидно, что при любом $(x', y') \in \Omega$ справедливо $\bar{h}_{(x',y')} \geq \underline{h} = \min_{T \in \mathcal{T}} h_T$, и поэтому $\ell_{h,(x',y')} \leq \ell_h = \ln(2 + \varepsilon^{-1} \underline{h})$.

Также будет использовано обозначение $\text{dist}_\Omega(\cdot, \cdot)$ для дистанции, понимаемой как длина минимального пути внутри Ω (т. е. $\text{dist}_\Omega = \text{dist}$, только если область Ω выпукла).

Л е м м а 2.17. Пусть $\mathcal{T}_1 := \{T : H_T \geq c_1 \varepsilon\}$ и $\mathcal{T}_2 := \{T : H_T \leq c_2 \varepsilon\}$, где используются произвольные фиксированные постоянные $0 < c_1 < c_2$. Существует интерполянт $G_h \in S_h$, при котором для $\bar{\Theta}$ из (2.72) справедливо

соотношение

$$\bar{\Theta} \lesssim \ell_{h,(x',y')},$$

если либо $\mathcal{T} = \mathcal{T}_1$, либо $\mathcal{T} = \mathcal{T}_2$, либо же $\text{dist}_\Omega(\mathcal{T}_1 \setminus \mathcal{T}_2, \mathcal{T}_2 \setminus \mathcal{T}_1) \gtrsim \varepsilon$.

З а м е ч а н и е 2.18. Для выполнения условия $\text{dist}_\Omega(\mathcal{T}_1 \setminus \mathcal{T}_2, \mathcal{T}_2 \setminus \mathcal{T}_1) \gtrsim \varepsilon$ в Лемме 2.17 требуется, чтобы H_T не переключалось слишком резко между $H_T \leq c_1\varepsilon$ и $H_T \geq c_2\varepsilon$, т. е., грубо говоря, переходная подобласть $\{c_1\varepsilon \leq H_T \leq c_2\varepsilon\}$ должна иметь внутренний диаметр $\gtrsim \varepsilon$.

Д о к а з а т е л ь с т в о. На протяжении доказательства (x', y') будет оставаться фиксированным, поэтому будем писать просто G для $G(x', y'; \cdot)$, а также \hbar для $\hbar_{(x',y')}$ и ℓ'_h для $\ell_{h,(x',y')}$. Начнем с доказательства соотношения

$$\bar{\Theta}_1 := \varepsilon^2 \sum_{T \in \mathcal{T}_1} \left(\lambda_T^{-1} \|\nabla G\|_{1;T} + \lambda_T^{-2} \|G\|_{1;T} \right) \lesssim 1. \quad (2.74)$$

Действительно, поскольку $T \in \mathcal{T}_1$, то $\lambda_T \simeq \varepsilon$, поэтому $\bar{\Theta}_1 \lesssim \varepsilon \|\nabla G\|_{1;\Omega} + \|G\|_{1;\Omega}$. Объединяя это с (2.18) получаем (2.74).

Далее, построим интерполянт \tilde{G}^I типа Лагранжа для G , который будет удовлетворять

$$\bar{\Theta}_2 := \varepsilon^2 \sum_{T \in \mathcal{T}_2} \left(\lambda_T^{-1} \|\nabla(G - \tilde{G}^I)\|_{1;T} + \lambda_T^{-2} \|G - \tilde{G}^I\|_{1;T} \right) \lesssim \ell'_h. \quad (2.75)$$

Заметим, что здесь $T \in \mathcal{T}_2$, и поэтому $\lambda_T \simeq H_T$. Рассмотрим функцию $\tilde{G} := \min\{G, \varepsilon^{-2} c_3 \ell'_h\}$ при некоторой постоянной $c_3 > 0$, которая является непрерывной аппроксимацией для G . Наконец, определим \tilde{G}^I как стандартный кусочно-линейный интерполянт Лагранжа для \tilde{G} . Заметим, что в случае $C_f \simeq 1$, справедлива верхняя оценка $G \leq G_0 := \varepsilon^{-2} (2\pi)^{-1} K_0(\sqrt{C_f} r/\varepsilon)$, где $r = [(x - x')^2 + (y - y')^2]^{1/2}$, а K_0 — модифицированная функция Бесселя второго рода нулевого порядка. (Если же $C_f < \frac{1}{2}$, то $\varepsilon^2 \geq \frac{1}{2}$, и можно воспользоваться аналогичной оценкой $G_0 := \varepsilon^{-2} (2\pi)^{-1} \ln(r^{-1} \text{diam } \Omega)$.) Вспомнив, что $K_0(s) \lesssim \ln(2 + s^{-1})$ при $s > 0$, заключаем, что найдется такая

достаточно малая постоянная c_3 , что $\{r \geq \hbar\} \subseteq \{G_0 \leq \varepsilon^{-2}c_3 \ell'_h\}$, и поэтому $\Omega \setminus B(x', y'; \hbar) \subseteq \{G = \tilde{G}\}$ независимо от ε или \hbar . Далее положим $\mathcal{T}_2^* := \{T \in \mathcal{T}_2 : T \cap B(x', y'; \hbar) = \emptyset\}$, а $\bar{\Theta}_2^*$ пусть будет определено как $\bar{\Theta}_2$ в (2.75), только с \mathcal{T}_2^* вместо \mathcal{T}_2 . Поскольку $G = \tilde{G}$ на любом $T \in \mathcal{T}_2^*$, то $\tilde{G}^I = G^I$, и поэтому

$$\bar{\Theta}_2^* \simeq \varepsilon^2 \sum_{T \in \mathcal{T}_2^*} \left(H_T^{-1} \|\nabla(G - G^I)\|_{1;T} + H_T^{-2} \|G - G^I\|_{1;T} \right) \lesssim \varepsilon^2 |G|_{2,1;\Omega \setminus B(x', y'; \hbar)},$$

где мы воспользовались Следствием 2.1 из монографии [56]. Объединяя последнее соотношение с (2.19с), приходим к $\bar{\Theta}_2^* \lesssim \ell'_h$.

При рассмотрении $\mathcal{T}_2 \setminus \mathcal{T}_2^* = \{T \in \mathcal{T}_2 : T \cap B(x', y'; \hbar) \neq \emptyset\}$ заметим, что, в силу (2.73), справедливо

$$B(x', y'; \hbar) \subset \omega_{(x', y')}, \quad T \cap B(x', y'; \hbar) \neq \emptyset \Rightarrow T \subset B(x', y'; 2H_T). \quad (2.76)$$

Здесь первое соотношение очевидно, а для получения второго заметим, что из $T \cap B((x', y'); \hbar) \neq \emptyset$ следует, что (i) T расположен на расстоянии не больше \hbar от (x', y') , (ii) $T \subset \omega_{(x', y')}$ (также в силу первого соотношения в (2.76)), и поэтому $\hbar = \hbar_{(x', y')} \leq h_T \leq H_T$. Наконец оценим

$$\bar{\Theta}_2 - \bar{\Theta}_2^* \simeq \varepsilon^2 \sum_{T \in \mathcal{T}_2 \setminus \mathcal{T}_2^*} \left(H_T^{-1} \|\nabla(G - \tilde{G}^I)\|_{1;T} + H_T^{-2} \|G - \tilde{G}^I\|_{1;T} \right).$$

Для функции Грина G , в силу (2.76), (2.19а) и (2.19б), справедлива оценка $\varepsilon^2 (H_T^{-1} \|\nabla G\|_{1;T} + H_T^{-2} \|G\|_{1;T}) \lesssim \ell_\rho$, где $\rho = 2H_T$, и поэтому $\ell_\rho \leq \ell'_h$. Аналогичная оценка для \tilde{G}^I справедлива в силу $\varepsilon^2 |\tilde{G}^I| \leq c_3 \ell'_h$, а также $\varepsilon^2 |\nabla \tilde{G}^I| \leq c_3 \ell'_h h_T^{-1}$, объединенного с $|T| \simeq h_T H_T$. Из (2.76) также следует, что число треугольников $T \in \mathcal{T}_2 \setminus \mathcal{T}_2^*$ является $\simeq 1$. Поэтому заключаем, что $\bar{\Theta}_2 - \bar{\Theta}_2^* \leq \ell'_h$. Поскольку также $\bar{\Theta}_2^* \leq \ell'_h$, желаемая оценка (2.75) установлена.

Далее, в силу (2.74) и (2.75), если либо $\mathcal{T} = \mathcal{T}_1$, либо же $\mathcal{T} = \mathcal{T}_2$, то соответственно определяя либо $G_h = 0 \in S_h$, либо же $G_h = \tilde{G}^I \in S_h$, получаем утверждение леммы.

Осталось рассмотреть случай $\text{dist}_\Omega(\mathcal{T}_1 \setminus \mathcal{T}_2, \mathcal{T}_2 \setminus \mathcal{T}_1) \gtrsim \varepsilon$. Положим $G_h := (\mu^I \tilde{G})^I \in S_h$, где μ^I является кусочно-линейным интерполянтом Лагранжа для функции-веса μ , определяемой следующим образом. Для любой точки $(x, y) \in \Omega$ положим $\mu := 0$ в $\mathcal{T}_1 \setminus \mathcal{T}_2$, $\mu := 1$ в $\mathcal{T}_2 \setminus \mathcal{T}_1$, и

$$\mu := \min\{1; \text{dist}_\Omega((x, y), \mathcal{T}_1 \setminus \mathcal{T}_2) / \text{dist}_\Omega(\mathcal{T}_1 \setminus \mathcal{T}_2, \mathcal{T}_2 \setminus \mathcal{T}_1)\}$$

при $(x, y) \in \mathcal{T}_1 \cap \mathcal{T}_2$. Заметим, что функция μ непрерывна, а для ее интерполянта Лагранжа μ^I с помощью оценок типа неравенства треугольника можно получить

$$\|\nabla \mu^I\|_{\infty; \Omega} \leq 1 / \text{dist}_\Omega(\mathcal{T}_1 \setminus \mathcal{T}_2, \mathcal{T}_2 \setminus \mathcal{T}_1) \lesssim \varepsilon^{-1}. \quad (2.77)$$

Достаточно оценить градиент $\hat{\mu}(x, y) := \text{dist}_\Omega((x, y), \mathcal{T}_1 \setminus \mathcal{T}_2)$ в $\mathcal{T}_1 \cap \mathcal{T}_2$. Пусть $\hat{\mu}(P_1) = \text{dist}_\Omega(P_1, Q_1) \leq \hat{\mu}(P_2) = \text{dist}_\Omega(P_2, Q_2)$, где $Q_1, Q_2 \in \partial(\mathcal{T}_1 \setminus \mathcal{T}_2)$. Поскольку $\hat{\mu}(P_2) \leq \text{dist}_\Omega(P_2, Q_1)$, то $0 \leq \hat{\mu}(P_2) - \hat{\mu}(P_1) \leq \text{dist}_\Omega(P_2, Q_1) - \text{dist}_\Omega(P_1, Q_1) \leq \text{dist}_\Omega(P_1, P_2)$. Если же P_1 и P_2 являются узлами сетки с общим ребром, то $\text{dist}_\Omega(P_1, P_2) = |P_1 P_2|$, поэтому производная $\hat{\mu}^I$ вдоль этого ребра равна $[\hat{\mu}(P_2) - \hat{\mu}(P_1)] / |P_1 P_2| \in [0, 1]$. Отсюда получаем (2.77).

Представим $g = G - G_h$ из (2.72) в виде $g = (1 - \mu^I)G + [\mu^I G - (\mu^I \tilde{G})^I]$ и заметим, что $1 - \mu^I = 0$ в $\mathcal{T} \setminus \mathcal{T}_1$, в то время как $\mu^I = 0$ в $\mathcal{T} \setminus \mathcal{T}_2$. Следовательно $\bar{\Theta} \leq \bar{\Theta}_1^\mu + \bar{\Theta}_2^\mu$, где $\bar{\Theta}_1^\mu$ определяем как $\bar{\Theta}_1$ в (2.74), только с $(1 - \mu^I)G$ вместо G , а $\bar{\Theta}_2^\mu$ определяем как $\bar{\Theta}_2$ в (2.75), только с $\mu^I G - (\mu^I \tilde{G})^I$ вместо $G - \tilde{G}^I$. Аналогично доказательству оценки для $\bar{\Theta}_1$, приходим к $\bar{\Theta}_1^\mu \lesssim \varepsilon \|\nabla\{(1 - \mu^I)G\}\|_{1; \Omega} + \|(1 - \mu^I)G\|_{1; \Omega} \lesssim 1$. Аналогично же доказательству оценки для $\bar{\Theta}_2$, приходим к $\bar{\Theta}_2^\mu \lesssim \ell'_h$. В частности, чтобы получить оценку того же типа, что и для $\bar{\Theta}_2^*$, заметим, что $\bar{\Theta}_2^{\mu, *} \lesssim \varepsilon^2 |\mu^I G|_{2,1; \Omega \setminus B(x', y'; h)} \lesssim \varepsilon^2 |G|_{2,1; \Omega \setminus B(x', y'; h)} + \varepsilon |G|_{1,1; \Omega}$; объединяя последнее соотношение с (2.19с) и (2.18), действительно приходим к $\bar{\Theta}_2^{\mu, *} \lesssim \ell'_h$. Разность же $\bar{\Theta}_2^\mu - \bar{\Theta}_2^{\mu, *}$ оценивается аналогично $\bar{\Theta}_2 - \bar{\Theta}_2^*$; для этого используется соотношение $\varepsilon^2 (H_T^{-1} \|\nabla(\mu^I G)\|_{1; T} + H_T^{-2} \|\mu^I G\|_{1; T}) \lesssim \ell'_h$, которое следует из $\varepsilon H_T^{-1} \|G\|_{1; T} \lesssim \sqrt{\ell'_h}$, в силу (2.19а) и (2.18). Объединяя вышеизложенные наблюдения, получаем желаемую оценку $\bar{\Theta} \leq \bar{\Theta}_1^\mu + \bar{\Theta}_2^\mu \lesssim \ell'_h$. \square

2.8 Заключительные замечания

Основные результаты параграфа, изложенные в Теоремах 2.7, 2.9 и 2.15, представляют ряд апостериорных оценок для ошибки конечноэлементных решений в норме максимума модуля на достаточно общих анизотропных триангуляциях. Точнее говоря, ошибка представляется в виде $(u_h - u)(x', y') = I + II + \mathcal{E}_{\text{quad}}$, где для квадратурной ошибки $\mathcal{E}_{\text{quad}}$ справедлива оценка (2.21), в то время как вышеуказанные теоремы описывают оценки для I и II . При этом, последние оценки приобретают смысл лишь после объединения с оценкой $\Theta + \Theta' \lesssim \ell_{h,(x',y')}$, которая представлена в Лемме 2.17.

§3 Апостериорные оценки в энергетической норме на анизотропных сетках

Для сингулярно возмущенных полулинейных уравнений реакции-диффузии в многоугольных областях получены явные апостериорные оценки ошибки численного решения в энергетической норме. Как и в §2, рассматривается случай неструктурированных анизотропных сеток (т. е. элементы рассматриваемых здесь сеток могут быть сколь угодно сильно сплющены). Постоянные в полученных оценках не зависят от диаметров элементов сетки, их аспектного соотношения (т. е. степени их сплюсненности) и малого параметра. Для интерполяции функции Грина на анизотропных сетках построен новый квазиинтерполяционный оператор, который может представлять независимый интерес. Основной материал параграфа опубликован в работе [169], а некоторые обобщения сделаны в работах [171, 173].

3.1 Введение

В настоящем параграфе будет продолжено изучение конечноэлементных аппроксимаций для сингулярно возмущенного уравнения реакции-диффузии

$$Lu := -\varepsilon^2 \Delta u + f(x, y; u) = 0 \quad \text{for } (x, y) \in \Omega, \quad u = 0 \quad \text{on } \partial\Omega, \quad (3.1)$$

в многоугольной, возможно нелипшицевой, области $\Omega \subset \mathbb{R}^2$, которое будет рассматриваться при тех же условиях, что и (2.1) в §2. В частности, будет предполагаться, что $0 < \varepsilon \leq 1$, что функция f удовлетворяет одностороннему условию Липшица $f(x, y; u) - f(x, y; v) \geq C_f[u - v]$ при всех $u \geq v$ с некоторой постоянной $C_f \geq 0$, а также что $C_f + \varepsilon^2 \geq 1$.

Нашей целью вновь является получение апостериорных оценок ошибки на основе невязок численного решения в случае неструктурированных анизотропных сеток (т. е. элементы рассматриваемых здесь сеток могут быть сильно сплющены). В отличие от §2, здесь оценки решения будут получены

в так называемой энергетической норме $\|\cdot\|_{\varepsilon;\Omega}$, которая является нормализованным вариантом нормы в $W_2^1(\Omega)$ и определяется для любой области $\mathcal{D} \subseteq \Omega$ соотношением

$$\|v\|_{\varepsilon;\mathcal{D}} := \left\{ \varepsilon^2 \|\nabla v\|_{2;\mathcal{D}}^2 + \|v\|_{2;\mathcal{D}}^2 \right\}^{1/2}.$$

В настоящем параграфе для уравнения (3.1) будут вновь рассмотрены стандартные линейные конечные элементы. Пусть $S_h \subset H_0^1(\Omega) \cap C(\bar{\Omega})$ является пространством непрерывных кусочно-линейных функций, ассоциируемым с триангуляцией \mathcal{T} , а численное решение $u_h \in S_h$ пусть удовлетворяет соотношению

$$\varepsilon^2 \langle \nabla u_h, \nabla v_h \rangle + \langle f_h^I, v_h \rangle = 0 \quad \forall v_h \in S_h, \quad f_h(\cdot) := f(\cdot; u_h). \quad (3.2)$$

Здесь $\langle \cdot, \cdot \rangle$ обозначает скалярное произведение в $L_2(\Omega)$, а f_h^I — стандартный кусочно-линейный интерполянт Лагранжа для f_h .

В рамках предварительного описания основных результатов параграфа, наша первая оценка принимает вид

$$\begin{aligned} \|u_h - u\|_{\varepsilon;\Omega} \leq C \left\{ \sum_{z \in \mathcal{N}} \min\{h_z H_z, \varepsilon H_z\} \|\varepsilon J_z\|_{\infty;\gamma_z}^2 \right. \\ \left. + \sum_{z \in \mathcal{N}} \|\min\{1, H_z \varepsilon^{-1}\} f_h^I\|_{2;\omega_z}^2 + \|f_h - f_h^I\|_{2;\Omega}^2 \right\}^{1/2}, \end{aligned} \quad (3.3)$$

где постоянная C не зависит от диаметров элементов сетки и их аспектных соотношений (т. е. степени их сплюснутости), а также от малого параметра ε . (Данная оценка получается объединением (3.14), (3.15), (3.25), (3.26) с Замечанием 3.10 и Теоремой 3.18.) Здесь \mathcal{N} является множеством вершин элементов \mathcal{T} , J_z обозначает стандартный скачок нормальной производной u_h через границу элемента, ω_z — множество элементов, окружающих вершину $z \in \mathcal{N}$, γ_z — множество ребер внутри ω_z , $H_z = \text{diam}(\omega_z)$, а $h_z \simeq H_z^{-1} |\omega_z|$.

Чтобы показать связь (3.3) с оценками для интерполяционной ошибки, а также с возможными адаптивными процедурам, заметим, что величина

$|J_z|$ может интерпретироваться как аппроксимирующая диаметр области ω_z в метрике, порожденной квадратным корнем гессиана точного решения (в то время как f_h^I аппроксимирует $\varepsilon^2 \Delta u$).

При некоторых дополнительных предположениях о триангуляции будет также установлена и более точная (хотя и несколько более сложная) апостериорная оценка, в которой, по сравнению с (3.3), грубо говоря, слагаемые $\min\{1, H_z \varepsilon^{-1}\}$ будут заменены на $\min\{1, h_z \varepsilon^{-1}\}$, при этом в оценке появятся некоторые новые слагаемые, связанные с поэлементными невязками.

Ряд работ, посвященных априорным и апостериорным оценкам для уравнений типа (3.1), уже был процитирован в §§1–2. Здесь мы более подробно остановимся на работах, в которых рассмотрены анизотропные сетки, при этом ошибки конечноэлементных решений оцениваются в энергетической норме или же в аналогичной норме H^1 . Прежде всего отметим работу [137], в которой рассматривались прямоугольные анизотропные сетки, и для уравнения Лапласа были получены апостериорные оценки в норме H^1 . Случаю же неструктурированных анизотропных сеток посвящены работы [104, 107], где получены апостериорные оценки в норме H^1 для уравнения Лапласа, и работы [105, 107], где для линейного уравнения типа (3.1) с постоянными коэффициентами выведены апостериорные оценки в энергетической норме.

Важно отметить, что в апостериорных оценках на основе невязок всегда присутствует некоторая постоянная, аналогичная C в правой части оценки (3.3), в то время как во всех апостериорных оценках из [104, 105, 107] такие постоянные определяются через так называемые функции согласования (matching functions). Последние же являются функциями ошибки численного решения и принимают умеренные значения, если либо сетка локально квазиравномерна, либо если сетка, будучи анизотропной, правильно ориентирована по отношению к решению. В общем же случае функции согласования зависят от аспектного соотношения элементов сетки, и, таким образом, постоянные в оценках [104, 105, 107] (аналогичные C в (3.3)) могут принимать

неограниченно большие значения на сильно анизотропных сетках (если только сетка не является правильно ориентированной по отношению к решению).

Достаточно очевидно, что присутствие подобных функций согласования в апостериорных оценках весьма нежелательно. С другой стороны, в более поздней работе [168] (см. также §2) были получены апостериорные оценки в норме максимума модуля, в которых функции согласования вообще отсутствуют (благодаря использованию нового технического приема, описанного в Замечании 2.6). Таким образом, целью настоящего параграфа является обобщение теоретического аппарата из [168] для получения апостериорных оценок ошибки конечноэлементных решений в энергетической норме, в которых нежелательные, как было описано выше, функции согласования присутствовать не будут (см., например, (3.3)).

Совершенно естественно, что в большой степени мы будем опираться на теоретический аппарат, описанный в §2. В то же время будет преодолен ряд дополнительных технических трудностей. Прежде всего, поскольку ошибка численного решения будет оцениваться в другой норме, нам потребуются более тонкая версия теоремы о следах для анизотропных элементов из §2.3. Далее, модифицированный интерполянт Лагранжа, который использовался при интерполяции функции Грина в §2, становится неприменимым при получении оценок в энергетической норме даже в случае уравнения Лапласа. Поэтому для интерполяции функции Грина на анизотропных сетках построен новый квазиинтерполяционный оператор, который может представлять независимый интерес. Наконец, здесь нам удастся упростить некоторые предположения о триангуляции: в частности, будет опущено условие $\mathcal{A}2$ из §2.6.1 (в котором максимальный угол при любом анизотропном узле z предполагается не превосходящим $\frac{\pi}{2} + \alpha_1 \frac{h_z}{H_z}$). Последнее упрощение может быть использовано и при анализе ошибки в норме максимума модуля (см. Замечание 3.8).

Параграф организован следующим образом. В §3.2 описываются основные предположения о триангуляции \mathcal{T} , а также допустимые типы анизотроп-

ных узлов. В §3.3 рассматриваются варианты теоремы о следах для анизотропных элементов. В §3.4 ошибка численного решения представлена через невязки, что затем используется в §3.5 при получении апостериорных оценок на частично структурированных анизотропных сетках. Основные результаты параграфа для случая неструктурированных анизотропных сеток изложены в §3.6, а §3.7 посвящен построению квазиинтерполяционных операторов для анизотропных элементов. Завершается параграф результатами численных экспериментов, представленными в §3.8.

Обозначения. В настоящем параграфе будут вновь использованы основные обозначения, описанные в последнем абзаце в §2.1.

3.2 Основные предположения о триангуляции

Будем использовать обозначения $z = (x_z, y_z)$, S и T соответственно для узлов, ребер и элементов сетки, в то время как \mathcal{N} , \mathcal{S} и \mathcal{T} обозначают их соответствующие множества. Для любого $T \in \mathcal{T}$ пусть H_T обозначает максимальную длину ребра, Для любого $z \in \mathcal{N}$ обозначение ω_z будет использоваться для области, состоящей из элементов, окружающих z , а \mathcal{S}_z — для множества ребер, исходящих из z , а также

$$H_z := \text{diam}(\omega_z), \quad h_z := H_z^{-1}|\omega_z|, \quad \gamma_z := \mathcal{S}_z \setminus \partial\Omega, \quad \mathring{\gamma}_z := \{S \subset \gamma_z : |S| \lesssim h_z\}. \quad (3.4)$$

В дальнейшем всегда будет предполагаться, что триангуляция удовлетворяет следующим условиям.

- *Максимальный угол.* Максимальный внутренний угол в любом треугольнике $T \in \mathcal{T}$ является равномерно ограниченным некоторой постоянной $\alpha_0 < \pi$.
- *Локальная ориентация элементов.* Для любого узла $z \in \mathcal{N}$, найдется такой прямоугольник $R_z \supset \omega_z$, что $|R_z| \simeq |\omega_z|$.

- Число треугольников, содержащих любой узел сетки, является равномерно ограниченным.

Дополнительно будем предполагать, что любой узел сетки $z \in \mathcal{N}$ принадлежит к одному из следующих *Типов Узлов*, которые здесь описываются с использованием фиксированной достаточно малой постоянной c_0 (последняя позволяет различать анизотропные и неанизотропные элементы сетки), при этом будем писать $a \ll b$, если $a < c_0 b$.

(1) *Анизотропными узлами*, множество которых будет обозначаться \mathcal{N}_{ani} , назовем узлы, которые удовлетворяют соотношениям

$$h_z < c_0 H_z, \quad h_T \simeq h_z \text{ и } H_T \simeq H_z \quad \forall T \subset \omega_z. \quad (3.5)$$

Заметим, что данные условия подразумевают, что \mathcal{S}_z содержит не более двух ребер длины $\lesssim h_z$.

(2) *Регулярными узлами*, множество которых будет обозначаться \mathcal{N}_{reg} , назовем узлы, окруженные квазиравномерными элементами (т. е. элементами, для которых $h_T \simeq H_T$).

(1*) Можно ожидать, что анизотропные элементы, расположенные около границы, будут ориентированы вдоль границы. Чтобы специально выделить узлы, для которых данное свойство не выполняется, введем в рассмотрение следующее подмножество анизотропных граничных узлов $\mathcal{N}_{\partial\Omega}^*$:

$$\mathcal{N}_{\partial\Omega}^* := \{ z \in \partial\Omega \cap \mathcal{N}_{\text{ani}} : |\mathcal{S}_z \cap \partial\Omega| \lesssim h_z \text{ или } z \text{ — угол области } \Omega \}. \quad (3.6)$$

Заметим, что определение (3.5) совпадает с (2.6), а регулярные узлы являются более простым частным случаем изотропных узлов (2.8). Результаты настоящего параграфа могут быть обобщены для всех типов узлов, описанных в §2.2 [169]. В настоящем же параграфе рассмотрение ограничено анизотропными и регулярными узлами лишь для упрощения изложения.

3.3 Теорема о следах для анизотропных элементов

Нам потребуется теорема о следах для анизотропных элементов, которая будет здесь сформулирована с использованием нормализованной нормы пространства $W_p^1(\mathcal{D})$, где $p = 1, 2$, определяемой соотношением

$$\|v\|_{p;\mathcal{D}} := (\text{diam}\mathcal{D})^{-1} \|v\|_{p;\mathcal{D}} + \|\nabla v\|_{p;\mathcal{D}}.$$

В частности, в силу $\text{diam}(\omega_z) = H_z$ и $\text{diam}(T) \simeq H_T$, имеем

$$\|v\|_{p;\omega_z} = H_z^{-1} \|v\|_{p;\omega_z} + \|\nabla v\|_{p;\omega_z}, \quad \|v\|_{p;T} \simeq H_T^{-1} \|v\|_{p;T} + \|\nabla v\|_{p;T}. \quad (3.7)$$

Отметим, что данные обозначения несколько отличаются от обозначения (2.10) для аналогичной нормы, которое использовалось в §2.

Л е м м а 3.1. Для любого узла $z \in \mathcal{N} = \mathcal{N}_{\text{ani}} \cup \mathcal{N}_{\text{reg}}$ и любой функции $v \in W_1^1(\omega_z)$ справедливо соотношение

$$\|v\|_{1;\dot{\gamma}_z} + \frac{h_z}{H_z} \|v\|_{1;\gamma_z \setminus \dot{\gamma}_z} + \frac{h_z}{H_z} \|v\|_{1;\bar{S}_z} \lesssim \|v\|_{1;\omega_z}, \quad (3.8)$$

где используются γ_z и $\dot{\gamma}_z$ из (3.4), а $\bar{S}_z \subset \omega_z$ является произвольным отрезком, исходящим из z и удовлетворяющим $|\bar{S}_z| \simeq H_z$.

Д о к а з а т е л ь с т в о. Данная лемма является вариантом Леммы 3.1 из §2.3, переписанной в обозначениях (3.7). \square

С л е д с т в и е 3.2. При выполнении условий Леммы 3.1, справедливо соотношение

$$\|v\|_{1;\dot{\gamma}_z} + \frac{h_z}{H_z} \|v\|_{1;\gamma_z \setminus \dot{\gamma}_z} + \frac{h_z}{H_z} \|v\|_{1;\bar{S}_z} \lesssim \left\{ h_z \|v\|_{2;\omega_z} \|v\|_{2;\omega_z} \right\}^{1/2}. \quad (3.9)$$

Д о к а з а т е л ь с т в о. Заметим, что $|\dot{\gamma}_z| \lesssim h_z$, а также $|\gamma_z \setminus \dot{\gamma}_z| \lesssim H_z$ и $|\bar{S}_z| \lesssim H_z$ приводят к

$$\|v\|_{1;\dot{\gamma}_z} + \frac{h_z}{H_z} \|v\|_{1;(\gamma_z \setminus \dot{\gamma}_z) \cup \bar{S}_z} \lesssim h_z^{1/2} \left\{ \|v\|_{2;\dot{\gamma}_z}^2 + \frac{h_z}{H_z} \|v\|_{2;\gamma_z \setminus \dot{\gamma}_z \cup \bar{S}_z}^2 \right\}^{1/2}.$$

Далее, заменяя v на v^2 в (3.8), немедленно получаем

$$\|v\|_{2;\dot{\gamma}_z}^2 + \frac{h_z}{H_z} \|v\|_{2;(\gamma_z \setminus \dot{\gamma}_z) \cup \bar{S}_z}^2 \lesssim \|\!\|v^2\|\!\|_{1;\omega_z},$$

в то время как, в силу (3.7),

$$\|\!\|v^2\|\!\|_{1;\omega_z} \lesssim H_z^{-1} \|v\|_{2;\omega_z}^2 + \|v\|_{2;\omega_z} \|\nabla v\|_{2;\omega_z} \simeq \|v\|_{2;\omega_z} \|\!\|v\|\!\|_{2;\omega_z}.$$

Объединение вышеизложенных наблюдений приводит к (3.9). \square

З а м е ч а н и е 3.3. *Ниже, в Замечаниях 3.6 и 3.10, будут описаны несколько улучшенные, по сравнению с представленными в настоящем параграфе, оценки из [173] для компоненты ошибки I из (3.14). Для получения данных более точных оценок используется следующая более тонкая теорема о следах для анизотропных элементов из Леммы 3.1 работы [173]:*

$$|S|^{-1} \|v\phi_z\|_{1;S}^2 \lesssim \|v\|_{2;\omega_z} \|\nabla v\|_{2;\omega_z} + \|v\|_{2;\omega_z}^2 \begin{cases} H_z^{-1} & \text{if } S \subset \dot{\gamma}_z, \\ h_z^{-1} & \text{if } S \subset \gamma_z \setminus \dot{\gamma}_z, \end{cases} \quad (3.10)$$

где $\phi_z \in S_h$ — стандартная базисная функция, ассоциируемая с z . Отметим, что (3.9) станет идентична с оценкой (3.10), если в правую часть последней включить вес $H_z/h_z \gg 1$ при $\|\nabla v\|_{2;\omega_z}$. Отметим также, что более тонкая оценка (3.10) остается справедливой лишь если оценивается $\|v\phi_z\|_{2;S}$ вместо $\|v\|_{2;S}$, оцениваемого в (3.9).

3.4 Представление ошибки численного решения через невязки

Воспользовавшись монотонностью f , а также $C_f + \varepsilon^2 \geq 1$, несложно убедиться, что

$$\begin{aligned} \|u_h - u\|_{\varepsilon;\Omega}^2 &\lesssim \varepsilon^2 \langle \nabla(u_h - u), \nabla(u_h - u) \rangle + \langle f(\cdot; u_h) - f(\cdot; u), u_h - u \rangle \\ &= \varepsilon^2 \langle \nabla u_h, \nabla(u_h - u) \rangle + \langle f(\cdot; u_h), u_h - u \rangle, \end{aligned}$$

где также использовалось (3.1). Далее, в предположении $\|u_h - u\|_{\varepsilon;\Omega} > 0$, положим

$$G := \frac{u_h - u}{\|u_h - u\|_{\varepsilon;\Omega}} \quad \Rightarrow \quad \|G\|_{\varepsilon;\Omega} = 1. \quad (3.11)$$

Тогда $\|u_h - u\|_{\varepsilon;\Omega} \lesssim \varepsilon^2 \langle \nabla u_h, \nabla G \rangle + \langle f(\cdot; u_h), G \rangle$. Поэтому (3.2) приводит, $\forall v_h \in S_h$, к оценке

$$\|u_h - u\|_{\varepsilon;\Omega} \lesssim \varepsilon^2 \langle \nabla u_h, \nabla(G - v_h) \rangle + \langle f_h^I, G - v_h \rangle + \langle f_h - f_h^I, G \rangle. \quad (3.12)$$

Здесь $\langle f_h - f_h^I, G \rangle =: \mathcal{E}_{\text{quad}}$ описывает квадратурную ошибку, для которой, в силу (3.11), получаем

$$|\mathcal{E}_{\text{quad}}| \leq \|f_h - f_h^I\|_{\varepsilon;\Omega}^*, \quad (3.13)$$

где норма $\|\cdot\|_{\varepsilon;\Omega}^*$ является сопряженной норме $\|\cdot\|_{\varepsilon;\Omega}$; см. также Замечание 3.4 ниже.

Пусть $\phi_z \in S_h$ опять обозначает стандартную базисную функцию, соответствующую узлу $z \in \mathcal{N}$, равную 1 в этом узле и 0 в остальных узлах сетки. Введем $v_h := G_h + \sum_{z \in \mathcal{N}} \bar{g}_z \phi_z \in S_h$, где $G_h \in S_h$ будет интерполировать G , а \bar{g}_z выбирается равным некоторому усреднению величины $G - G_h$ в окрестности z (которое будет определено ниже), при этом $\bar{g}_z = 0$ для $z \in \partial\Omega$ (с тем, чтобы $v_h \in S_h$). Далее введем функцию $g := G - G_h$, с помощью которой получаем $G - v_h = g - \sum_{z \in \mathcal{N}} \bar{g}_z \phi_z = \sum_{z \in \mathcal{N}} (g - \bar{g}_z) \phi_z$. Объединяя это с (3.12), приходим к достаточно стандартному представлению ошибки

$$\begin{aligned} \|u_h - u\|_{\varepsilon;\Omega} &\lesssim \sum_{z \in \mathcal{N}} \varepsilon^2 \int_{\gamma_z} (g - \bar{g}_z) \phi_z \llbracket \nabla u_h \rrbracket \cdot \nu + \sum_{z \in \mathcal{N}} \int_{\omega_z} f_h^I (g - \bar{g}_z) \phi_z + \mathcal{E}_{\text{quad}} \\ &=: I + II + \mathcal{E}_{\text{quad}}, \end{aligned} \quad (3.14)$$

которое справедливо при любом $G_h \in S_h$ и любом наборе $\{\bar{g}_z\}_{z \in \mathcal{N}}$, удовлетворяющем $\bar{g}_z = 0$ при всех $z \in \partial\Omega$.

В (3.14), величина $\llbracket \nabla u_h \rrbracket$ обозначает стандартный скачок нормальной производной при переходе через внутреннее ребро. Для определенности будем считать, что единичная нормаль ν к любому ребру из γ_z направлена по

часовой стрелке по отношению к z , в то время как $\llbracket w \rrbracket$ для любой функции w понимается как скачок w при переходе через любое ребро из γ_z , вычисляемый в направлении против часовой стрелки по отношению к z . Таким образом, $\llbracket \nabla u_h \rrbracket \cdot \nu =: J_z$ — скачок, в направлении против часовой стрелки по отношению к z , нормальной производной u_h по направлению против часовой стрелки. Очевидно, что $|\llbracket \nabla u_h \rrbracket| = |J_z|$. При использовании обозначения $\llbracket \nabla u_h \rrbracket$ на ребрах, расположенных на границе области, будем считать, что $u_h = 0$ в $\mathbb{R}^2 \setminus \Omega$.

З а м е ч а н и е 3.4 (Квадратурная ошибка). *С точки зрения практических вычислений, используемая в оценке (3.13) норма $\|\cdot\|_{\varepsilon; \Omega}^*$, сопряженная норме $\|\cdot\|_{\varepsilon; \Omega}$, весьма неудобна. С другой стороны, из (3.13) следует, что при любом выборе $\mathcal{T}_0 \subset \mathcal{T}$ и любой сколь угодно малой постоянной $\beta > 0$, справедлива оценка*

$$|\mathcal{E}_{\text{quad}}| \lesssim \left\{ \varepsilon^{-2} \|f_h - f_h^I\|_{1+\beta; \mathcal{T}_0}^2 + \|f_h - f_h^I\|_{2; \Omega \setminus \mathcal{T}_0}^2 \right\}^{1/2}. \quad (3.15)$$

Здесь мы также воспользовались теоремой вложения Соболева $W_p^1(\Omega) \hookrightarrow L_{1+\beta^{-1}}(\Omega)$, где $p = p(\beta)$ удовлетворяет $\frac{1}{p} = \frac{1}{2} + (1 + \beta^{-1})^{-1} > \frac{1}{2}$, поэтому $p < 2$, и следовательно $\|v\|_{1,2; \Omega} \gtrsim \|v\|_{1,p; \Omega} \gtrsim \|v\|_{1+\beta^{-1}; \Omega} \forall v$.

3.5 Апостериорный анализ на частично структурированных анизотропных сетках

Чтобы проиллюстрировать основные новые идеи, начнем с изложения несколько упрощенной версии представляемого здесь апостериорного анализа для случая частично структурированных анизотропных сеток в квадратной области $\Omega = (0, 1)^2$. Будет предполагаться, что триангуляция удовлетворяет условиям A1–A3 из §2.5 (см. также Рис. 2.2, слева). Также будет использоваться определение (2.23) областей ω_z^* , для которых остается справедливым Замечание 2.5.

3.5.1 Выбор \bar{g}_z . Основные результаты

Следуя §2.5.1, выбор значений \bar{g}_z в (3.14) будет связан с локальной ориентацией анизотропных элементов. Пусть $\bar{g}_z = 0$ при $z \in \partial\Omega$, а в противном случае, при $x_z = x_i$ с некоторым $1 \leq i \leq n-1$, пусть \bar{g}_z определяется формулой (2.24). Заметим, что соотношение (2.25) также остается справедливым.

Т е о р е м а 3.5. Пусть $g = G - G_h$ с G из (3.11) и $\forall G_h \in S_h$, а

$$\Theta := \varepsilon^2 \|\nabla g\|_{2;\Omega}^2 + \sum_{z \in \mathcal{N}} (1 + \varepsilon^2 H_z^{-2}) \|g\|_{2;\omega_z}^2. \quad (3.16)$$

Тогда $\|u_h - u\|_{\varepsilon;\Omega} \lesssim I + II + \mathcal{E}_{\text{quad}}$, где для $\mathcal{E}_{\text{quad}}$ справедлива оценка (3.13), а при условиях A1–A3 из §2.5 справедливы соотношения

$$|I| \lesssim \left\{ \Theta \sum_{z \in \mathcal{N}} \lambda_z \|\varepsilon J_z\|_{\infty;\gamma_z}^2 \right\}^{1/2}, \quad \lambda_z := h_z H_z \min\{1, \varepsilon H_z h_z^{-2}\}, \quad (3.17)$$

$$|II| \lesssim \left\{ \Theta \sum_{z \in \mathcal{N}} \|\lambda'_z f_h^I\|_{2;\omega_z}^2 \right\}^{1/2}, \quad \lambda'_z := \min\{1, H_z \varepsilon^{-1}\}. \quad (3.18)$$

Также справедлива альтернативная оценка

$$\begin{aligned} |II| \lesssim & \left\{ \Theta \sum_{z \in \mathcal{N} \setminus \mathcal{N}_{\partial\Omega}^*} \|\min\{1, h_z \varepsilon^{-1}\} f_h^I\|_{2;\omega_z}^2 + \Theta \sum_{z \in \mathcal{N} \setminus \mathcal{N}_{\partial\Omega}^*} \|\lambda'_z \text{osc}(f_h^I; \omega_z)\|_{2;\omega_z}^2 \right. \\ & \left. + \Theta \sum_{z \in \mathcal{N}_{\partial\Omega}^*} \|\lambda'_z f_h^I\|_{2;\omega_z}^2 \right\}^{1/2}, \quad (3.19) \end{aligned}$$

где $\mathcal{N}_{\partial\Omega}^* = \{z \in \mathcal{N} : x_z = 0 \text{ or } x_z = 1\}$ (в соответствии с (3.6)).

Оценки (3.18) and (3.19) в Теореме 3.5 доказываются аналогично оценкам (2.28)–(2.30) для поэлементных невязок в §2.5.3 (подробные же доказательства даны в [169, §5.3]). Доказательство же оценки (3.17) для невязок, ассоциируемых с ребрами, несколько отличается от аналогичного доказательства в §2.5.2, и приводится ниже в §3.5.2.

З а м е ч а н и е 3.6. *Несколько улучшенная версия оценки (3.17) с весом $\lambda_z = \min\{h_z H_z, \varepsilon H_z\}$ установлена в Теореме 4.1 в работе [173]. Для этого используется вариант теоремы о следах для анизотропных элементов, описанный в Замечании 3.3.*

З а м е ч а н и е 3.7. *В §?? доказывается, что при более общих, чем A1–A3, условиях справедлива оценка $\Theta \lesssim \|G\|_{\varepsilon; \Omega} = 1$ (см. Теорему 3.18). Объединяя последнюю оценку с Теоремой 3.5 получаем две апостериорные оценки для $\|u_h - u\|_{\varepsilon; \Omega}$.*

3.5.2 Невязки, ассоциируемые с ребрами. Доказательство (3.17)

Д о к а з а т е л ь с т в о (3.17). Представим I из (3.14) в виде суммы $I = \sum_{z \in \mathcal{N}} I_z$, где

$$I_z := \varepsilon^2 \int_{\gamma_z} (g - \bar{g}_z) \phi_z [\nabla u_h] \cdot \nu. \quad (3.20)$$

Далее, воспользовавшись $[\nabla u_h] \cdot \nu = [\partial_x u_h] \nu_x + [\partial_y u_h] \nu_y$, можем разбить I_z как

$$\begin{aligned} I_z = I'_z + I''_z + I'''_z &:= \varepsilon^2 \int_{\gamma_z} (g - \bar{g}_z) \phi_z [\partial_x u_h] \nu_x \\ &+ \varepsilon^2 \int_{\gamma_z} [g - g(x, y_z)] \phi_z [\partial_y u_h] \nu_y \\ &+ \varepsilon^2 \int_{\gamma_z} [g(x, y_z) - \bar{g}_z] \phi_z [\partial_y u_h] \nu_y. \end{aligned} \quad (3.21)$$

Мы утверждаем, что для получения желаемой оценки (3.17), достаточно установить оценки

$$|I'_z| + |I''_z| \lesssim \varepsilon \|g\|_{1; \omega_z^*} \|\varepsilon J_z\|_{\infty; \gamma_z}, \quad I'''_z = 0, \quad (3.22)$$

и

$$|I_z| \lesssim \varepsilon \frac{H_z}{h_z} \left\{ h_z \|g\|_{2; \omega_z^*} \|g\|_{2; \omega_z^*} \right\}^{1/2} \|\varepsilon J_z\|_{\infty; \gamma_z}, \quad (3.23)$$

а затем воспользоваться неравенством Гельдера. Действительно, тогда $|I_z| \lesssim (\theta_z \lambda_z)^{1/2} \|\varepsilon J_z\|_{\infty; \gamma_z}$, где

$$\theta_z := \lambda_z^{-1} \varepsilon^2 \min \left\{ \|g\|_{1; \omega_z^*}^2, H_z^2 h_z^{-1} \|g\|_{2; \omega_z^*} \|g\|_{2; \omega_z^*} \right\},$$

что приводит к (3.17) в силу $\sum_{z \in \mathcal{N}} \theta_z \lesssim \Theta$. Для получения последней оценки заметим, что $\min(a, bc) / \min(1, c) \leq a + b$ (при любых $a, b, c > 0$). Поскольку $\lambda_z = h_z H_z \min\{1, \varepsilon H_z h_z^{-2}\}$, получаем $\theta_z \lesssim \varepsilon^2 \|g\|_{2; \omega_z^*}^2 + \varepsilon \|g\|_{2; \omega_z^*} \|g\|_{2; \omega_z^*}$, и остается лишь воспользоваться (3.7) и (2.23), а также Замечанием 2.5.

Таким образом, остается установить (3.22) и (3.23). При этом соотношения (3.22) идентичны с (2.32), переписанными в обозначениях настоящего параграфа, и были получены в §2.5.2. Соотношение же (3.23) можно получить из (3.20), воспользовавшись Следствием 3.2 (в духе Леммы 3.2 из [149]). Более подробно, воспользуемся $|\phi_z[\nabla u_h]| \leq |J_z|$, а также оценкой

$$\frac{h_z}{H_z} \int_{\gamma_z} (|g| + |\bar{g}_z|) \lesssim \left\{ h_z \|g\|_{2; \omega_z^*} \|g\|_{2; \omega_z^*} \right\}^{1/2}. \quad (3.24)$$

Последняя оценка доказывается с помощью (3.9), при этом используется $\|g\|_{2; \omega_z} \lesssim \|g\|_{2; \omega_z^*}$ (в силу $\omega_z \subseteq \omega_z^*$ и $\text{diam } \omega_z \simeq \text{diam } \omega_z^*$). При оценке \bar{g}_z в (3.24) используется также (2.25) (где $\bar{S}_z \subset \omega_z^*$, а $|\bar{S}_z| \simeq H_z$). Таким образом, (3.23) установлено. \square

3.6 Апостериорный анализ на неструктурированных анизотропных сетках

Аналогично §2.6, для упрощения изложения сделаем два дополнительных предположения о триангуляции.

A1*. Если $z \in \mathcal{N}$ является углом многоугольника Ω , то $z \in \mathcal{N}_{\text{reg}}$ (т. е. все углы области являются регулярными узлами, и ни один из них не принадлежит $\mathcal{N}_{\partial\Omega}^*$).

$\mathcal{A}2^*$. Максимальный угол при любом $z \in \mathcal{N}_{\partial\Omega}^*$ не превосходит $\frac{\pi}{2} + \alpha_1 \frac{h_z}{H_z}$ для некоторой положительной постоянной α_1 (см. Рис. 3.1).

З а м е ч а н и е 3.8 (Условия $\mathcal{A}1^$ и $\mathcal{A}2^*$). Условие $\mathcal{A}1^*$ является вариантом $\mathcal{A}1$ из §2.6. А вот условие $\mathcal{A}2^*$ значительно слабее, чем $\mathcal{A}2$ из §2.6, поскольку теперь мы накладываем ограничения лишь на геометрию граничных анизотропных узлов из $z \in \mathcal{N}_{\partial\Omega}^* \subset \partial\Omega$. Заметим, что $\mathcal{A}2$ используется в §2.6 лишь при получении оценки (2.44), в то время как аналогичная оценка (3.25) получена ниже лишь при условии $\mathcal{A}2^*$. Для получения последнего более тонкого результата ниже применяются соотношения (3.32) (вместо получения оценок для I_z'''). Похожий подход был ранее применен в [170, §6], и, более того, может быть применен для получения оценок в норме максимума модуля типа (2.44) при более слабом предположении $\mathcal{A}2^*$. (В §2 более сильное условие $\mathcal{A}2$ предполагалось с целью упрощения изложения.)*

С тем, чтобы получить оценители, аналогичные (3.17), (3.19), но для более общих триангуляций, будут сделаны некоторые дополнительные предположения. Как и в §2.6.4, весьма существенным оказывается, что поэлементные невязки будут оцениваться не для отдельных узлов, а для цепочек анизотропных узлов, объединенных короткими ребрами. Для этого мы вновь воспользуемся определением анизотропной цепочки из §2.6.4, такие цепочки будут опять обозначаться \mathcal{N}_i , а множество всех узлов в таких цепочках будет опять обозначаться $\mathcal{N}_{\text{paths}} := \cup_{i=1}^{n_{\text{ani}}} \mathcal{N}_i$. Будет также предполагаться, что выполняется условие $\mathcal{A}3$ из §2.6.4.

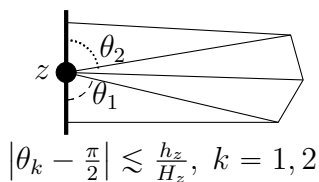


Рис. 3.1: Граничный анизотропный узел $z \in \mathcal{N}_{\partial\Omega}^*$, удовлетворяющий $\mathcal{A}2^*$.

3.6.1 Основные результаты

Основные результаты для случая неструктурированных анизотропных сеток представлены в виде Теорем 3.9 и 3.11.

Т е о р е м а 3.9. Пусть $g = G - G_h$, где G определено в (3.11), при любом $G_h \in S_h$, а Θ определено в (3.16). Тогда $\|u_h - u\|_{\varepsilon; \Omega} \lesssim I + II + \mathcal{E}_{\text{quad}}$, где для $\mathcal{E}_{\text{quad}}$ справедлива оценка (3.13), а при соблюдении условий $\mathcal{A}1^*$ и $\mathcal{A}2^*$, справедливы оценки

$$|I| \lesssim \left\{ \Theta \sum_{z \in \mathcal{N}} \lambda_z \|\varepsilon J_z\|_{\infty; \gamma_z}^2 \right\}^{1/2}, \quad \lambda_z := h_z H_z \min\{1, \varepsilon H_z h_z^{-2}\}, \quad (3.25)$$

$$|II| \lesssim \left\{ \Theta \sum_{z \in \mathcal{N}} \|\lambda'_z f_h^I\|_{2; \omega_z}^2 \right\}^{1/2}, \quad \lambda'_z := \min\{1, H_z \varepsilon^{-1}\}. \quad (3.26)$$

З а м е ч а н и е 3.10. Несколько улучшенная версия оценки (3.25) с весом $\lambda_z = \min\{h_z H_z, \varepsilon H_z\}$ установлена в Теореме 4.3 в работе [173]. Для этого используется вариант теоремы о следах для анизотропных элементов, описанный в Замечании 3.3.

Т е о р е м а 3.11. При выполнении условий Теоремы 3.9, а также условия $\mathcal{A}3$ из §2.6.4, справедлива оценка

$$\begin{aligned} |II| \lesssim \Theta^{1/2} \left\{ \sum_{z \in \mathcal{N}_{\text{paths}}} \|\min\{1, h_z \varepsilon^{-1}\} f_h^I\|_{2; \omega_z}^2 + \sum_{z \in \mathcal{N} \setminus \mathcal{N}_{\text{paths}}} \|\lambda'_z f_h^I\|_{2; \omega_z}^2 \right. \\ \left. + \sum_{z \in \mathcal{N}_{\text{paths}}} \|\lambda'_z \text{osc}(f_h^I; \omega_z)\|_{2; \omega_z}^2 \right\}^{1/2}. \end{aligned} \quad (3.27)$$

Теорема 3.9 будет доказана в §3.6.2, а Теорема 3.11 доказывается аналогично Теореме 2.15 в §2.6.4 (подробное доказательство дано в [169, §6]).

3.6.2 Доказательство Теоремы 3.9

Д о к а з а т е л ь с т в о (3.25). Для любого фиксированного узла $z \in \mathcal{N}_{\text{ani}}$, введем локальные обозначения, которые были использованы при доказательстве (2.45) в §2.6.2: локальные декартовы координаты (ξ, η) и соответствующую стандартную кусочно-линейную базисную функцию $\varphi_z(\xi)$ с

носителем $(\xi_z^-, \xi_z^+) \ni 0$. Дополнительно, в силу $\mathcal{A}2^*$, при $z \in \mathcal{N}_{\partial\Omega}^*$ выберем декартовы координаты таким образом, что ось ξ перпендикулярна участку границы $\partial\Omega$ в z .

Далее, положим $\bar{g}_z := 0$, если $z \in \mathcal{N}_{\text{reg}}$ является регулярным узлом или $z \in \partial\Omega$. В противном случае, вновь воспользуемся (2.48) с тем же определением функции $\bar{\eta}_z(\xi)$ (см. также Рис. 3.2, слева), т. е. пусть

$$\int_{\xi_z^-}^{\xi_z^+} [g(\xi, \bar{\eta}_z(\xi)) - \bar{g}_z] \varphi_z(\xi) d\xi = 0. \quad (3.28)$$

Вновь полагая $\bar{S}_z^- := \{(\xi, \bar{\eta}_z(\xi)) : \xi \in (\xi_z^-, 0)\}$ и $\bar{S}_z^+ := \{(\xi, \bar{\eta}_z(\xi)) : \xi \in (0, \xi_z^+)\}$, и затем воспользовавшись (3.28) и (3.8) приходим к оценке

$$h_z |\bar{g}_z| \lesssim \frac{h_z}{H_z} \|g\|_{1; \bar{S}_z^- \cup \bar{S}_z^+} \lesssim \|g\|_{1; \omega_z}. \quad (3.29)$$

Перейдем теперь к оценке I из (3.14). Для этого представим I в виде суммы $I = \sum_{z \in \mathcal{N}} I_z$, где I_z определяется в (3.20). Как и при доказательстве (3.17), вновь справедливы соотношения (3.24) и (3.23), с той лишь разницей, что ω_z^* заменяется на ω_z (поскольку $\bar{S}_z^- \cup \bar{S}_z^+ \subset \omega_z$ в (3.29)). Также, аналогично

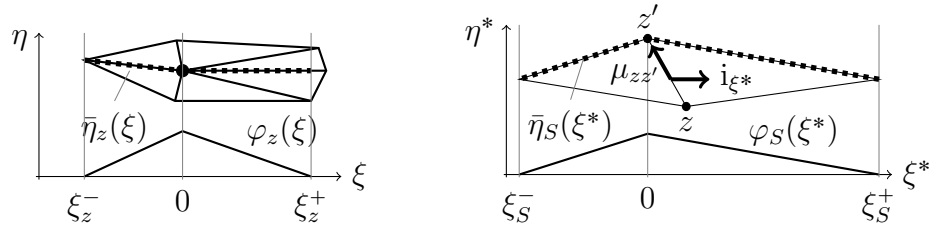


Рис. 3.2: Локальные обозначения, ассоциируемые с $z \in \mathcal{N}$ (слева), и коротким ребром $S \in \mathcal{S}^*$, соединяющим узлы z и z' (справа).

(3.21), I_z допускает представление

$$\begin{aligned}
I_z &= \varepsilon^2 \int_{\gamma_z} (g - \bar{g}_z) \phi_z [\nabla u_h] \cdot \nu = I'_z + I''_z + I'''_z + I''''_z \\
&:= \varepsilon^2 \int_{\hat{\gamma}_z} (g - \bar{g}_z) \phi_z [\nabla u_h] \cdot \nu + \varepsilon^2 \int_{\gamma_z \setminus \hat{\gamma}_z} (g - \bar{g}_z) \phi_z [\partial_\xi u_h] \nu_\xi \\
&\quad + \varepsilon^2 \int_{\gamma_z \setminus \hat{\gamma}_z} [g - g(\xi, \bar{\eta}_z(\xi))] \varphi_z [\partial_\eta u_h] \nu_\eta \\
&\quad + \varepsilon^2 \int_{\gamma_z \setminus \hat{\gamma}_z} [g(\xi, \bar{\eta}_z(\xi)) - \bar{g}_z] \varphi_z [\partial_\eta u_h] \nu_\eta \\
&\quad + \varepsilon^2 \int_{\gamma_z \setminus \hat{\gamma}_z} (g - \bar{g}_z) \{ \phi_z - \varphi_z \} [\partial_\eta u_h] \nu_\eta, \tag{3.30}
\end{aligned}$$

где, вновь несколько злоупотребляя обозначениями, $g = g(\xi, \eta)$.

Следующая оценка идентична (2.51), полученной при доказательстве (2.44) в §2.6.2; здесь она представлена в обозначениях настоящего параграфа:

$$|I'_z| + |I''_z| + |I'''_z| \lesssim \varepsilon \|g\|_{1; \omega_z} \|\varepsilon J_z\|_{\infty; \gamma_z} \quad \text{for } z \in \mathcal{N}. \tag{3.31}$$

Для завершения доказательства достаточно показать, что найдутся такое подмножество ребер $\mathcal{S}^* \subset \mathcal{S}$ и соответствующие величины $\mathcal{I}_{S;z}$, ассоциируемые с каждым $S \in \mathcal{S}_z \cap \mathcal{S}^*$ (последние будут определены ниже), что справедливы следующие соотношения:

$$\sum_{z \in \mathcal{N}} \sum_{S \in \mathcal{S}_z \cap \mathcal{S}^*} \mathcal{I}_{S;z} = 0, \tag{3.32a}$$

$$|I'''_z| + \sum_{S \in \mathcal{S}_z \cap \mathcal{S}^*} \mathcal{I}_{S;z} \lesssim \varepsilon \|g\|_{1; \omega_z} \|\varepsilon J_z\|_{\infty; \gamma_z}, \tag{3.32b}$$

$$|\mathcal{I}_{S;z}| \lesssim \varepsilon \frac{H_z}{h_z} \left\{ h_z \|g\|_{2; \omega_z} \|g\|_{2; \omega_z} \right\}^{1/2} \|\varepsilon J_z\|_{\infty; \gamma_z}. \tag{3.32c}$$

Действительно, объединяя оценку (3.23) для $|I_z|$ (где ω_z^* заменена на ω_z) с (3.31) и (3.32), получаем

$$|I_z| + \sum_{S \in \mathcal{S}_z \cap \mathcal{S}^*} \mathcal{I}_{S;z} \lesssim (\theta_z \lambda_z)^{1/2} \|\varepsilon J_z\|_{\infty; \gamma_z}, \tag{3.33}$$

где

$$\theta_z := \lambda_z^{-1} \varepsilon^2 \min \left\{ \|g\|_{1; \omega_z}^2, H_z^2 h_z^{-1} \|g\|_{2; \omega_z} \|g\|_{2; \omega_z} \right\}. \tag{3.34}$$

Отсюда получаем желаемую оценку (3.25), поскольку $\sum_{z \in \mathcal{N}} \theta_z \lesssim \Theta$ (последнее же соотношение выводится как и при доказательстве (3.17) в §3.5.2).

Перейдем к получению соотношений (3.32). Заметим, что в определении I_z''' в (3.30) используется функция $[g(\xi, \bar{\eta}_z(\xi)) - \bar{g}_z] \varphi_z$ одной переменной ξ , а $\nu_\eta |d\nu| = -\text{sgn}(\xi) d\xi$. Поэтому, в силу (3.28), приходим к представлению

$$I_z''' = -\varepsilon^2 \left(\sum_{S \in [\mathcal{S}_z^- \cup \mathcal{S}_z^+] \setminus \partial\Omega} \llbracket \partial_\eta u_h \rrbracket \right) \int_0^{\xi_z^+} [g(\xi, \bar{\eta}_z(\xi)) - \bar{g}_z] \varphi_z(\xi) d\xi. \quad (3.35)$$

Отсюда вытекает, что $I_z''' = 0$ при $z \in \mathcal{N}_{\text{reg}} \cup [(\mathcal{N}_{\text{ani}} \setminus \mathcal{N}_{\partial\Omega}^*) \cap \partial\Omega]$.

Для остальных узлов $z \in (\mathcal{N}_{\text{ani}} \setminus \partial\Omega) \cup \mathcal{N}_{\partial\Omega}^*$ получаем $\sum_{S \in \mathcal{S}_z^- \cup \mathcal{S}_z^+} \llbracket \partial_\eta u_h \rrbracket = -\sum_{S \in \mathcal{S}_z} \llbracket \partial_\eta u_h \rrbracket$ (где при $z \in \mathcal{N}_{\partial\Omega}^*$ при вычислении $\llbracket \partial_\eta u_h \rrbracket$ используется $u_h = 0$ в $\mathbb{R}^2 \setminus \Omega$). Заметим также, что $\llbracket \partial_\eta u_h \rrbracket = -\mu \cdot i_\xi J_z$ на любом ребре $S \in \mathcal{S}_z$, где μ — единичный вектор, направленный из z вдоль S . (Убедиться в последнем можно, объединяя $\partial_\eta = i_\eta \cdot \nabla$, $\llbracket \nabla u_h \rrbracket = \nu J_z$ и $i_\eta \cdot \nu = -i_\xi \cdot \mu$.) Объединяя вышеизложенные наблюдения с (3.35), при $z \in (\mathcal{N}_{\text{ani}} \setminus \partial\Omega) \cup \mathcal{N}_{\partial\Omega}^*$ получаем представление

$$I_z''' = -\varepsilon^2 \alpha_z \sum_{S \in \mathcal{S}_z} \mu \cdot i_\xi J_z, \quad \alpha_z := \int_0^{\xi_z^+} [g(\xi, \bar{\eta}_z(\xi)) - \bar{g}_z] \varphi_z(\xi) d\xi. \quad (3.36)$$

Заметим, что при $z \in \mathcal{N}_{\partial\Omega}^*$ вектор i_ξ был выбран перпендикулярным к $\partial\Omega$ в z , в то время как $\mathring{\mathcal{S}}_z \subset \partial\Omega$, поэтому $\mu \cdot i_\xi = 0$, и, таким образом, $I_z''' = 0$ при $z \in \mathcal{N}_{\partial\Omega}^*$.

В случае же $z \in \mathcal{N}_{\text{ani}} \setminus \partial\Omega$, воспользовавшись представлением (3.36), определим подмножество ребер $\mathcal{S}^* := \cup_{z \in \mathcal{N}_{\text{ani}} \setminus \partial\Omega} \mathring{\mathcal{S}}_z$. Далее, для любого ребра $S \in \mathcal{S}^*$, соединяющего узлы z и z' , положим

$$\mathcal{I}_{S;z} := \varepsilon^2 \alpha_S \mu_{zz'} \cdot i_{\xi^*} J_z|_S, \quad \alpha_S := \int_0^{\xi_S^+} [g(\xi^*, \bar{\eta}_S(\xi^*)) - \bar{g}_S] \varphi_S(\xi^*) d\xi^*. \quad (3.37)$$

Здесь $\mu_{zz'}$ — единичный вектор, направленный из z в z' , а i_{ξ^*} — единичный вектор вдоль оси ξ^* . Локальные декартовы координаты (ξ^*, η^*) , ассоциируемые с ребром S , пусть совпадают с локальными координатами (ξ, η) ,

ассоциируемыми либо с узлом $z \in \mathcal{N}_{\text{ani}}$, либо же с узлом $z' \in \mathcal{N}_{\text{ani}}$ (по крайней мере один из них принадлежит \mathcal{N}_{ani}). Величины α_S определяются аналогично α_z , при этом определением для \bar{g}_S пусть служит соотношение $\int_{\xi_S^-}^{\xi_S^+} [g(\xi^*, \bar{\eta}_S(\xi^*)) - \bar{g}_S] \varphi_S(\xi^*) d\xi^* = 0$ (являющееся вариантом (3.28)). Одномерная кусочно-линейная функция $\varphi_S(\xi^*)$ ассоциируется с интервалом (ξ_S^-, ξ_S^+) , который является проекцией области $\omega_z \cap \omega_{z'}$ (состоящей из двух треугольников) на ось ξ^* . Кусочно-линейная функция $\bar{\eta}_S(\xi^*)$ определяется аналогично $\bar{\eta}_z(\xi)$, при этом предполагается, что любая точка $(\xi^*, \bar{\eta}(\xi^*)) \in \omega_z \cap \omega_{z'}$ (см Рис. 3.2, справа).

Рассмотрим (3.37). Заметим, что $\mu_{zz'} + \mu_{z'z} = 0$, а $J_z|_S = J_{z'}|_S$, поэтому $\mathcal{I}_{S;z} + \mathcal{I}_{S;z'} = 0$, и следовательно получаем (3.32a). Далее, оценка (3.32c) может быть получена аналогично (3.23) для $|I_z|$. Для получения оценки (3.32b) при $z \in \mathcal{N}_{\text{ani}} \setminus \partial\Omega$ заметим, что μ в (3.36) соответствует $\mu_{zz'}$ в (3.37), в то время как либо $i_\xi = i_{\xi^*}$, либо, в силу условия локальной ориентации элементов, $|i_\xi - i_{\xi^*}| \simeq h_z H_z^{-1}$ (где, без ограничения общности, предполагается, что $i_\xi \cdot i_{\xi^*} > 0$). Также заметим, что $|\alpha_z - \alpha_S| \lesssim \|\nabla g\|_{1;\omega_z}$, в то время как $|\alpha_z| + |\alpha_S| \lesssim H_z h_z^{-1} \|g\|_{1;\omega_z}$ (последнее следует из (3.28), (3.29), объединенных с аналогичными оценками для g_S и \bar{g}_S). Объединяя вышеизложенные наблюдения, приходим к соотношению (3.32b) при любом $z \in \mathcal{N}_{\text{ani}} \setminus \partial\Omega$.

Наконец обсудим $z \notin \mathcal{N}_{\text{ani}} \setminus \partial\Omega$, для которых было показано, что $I_z''' = 0$, и поэтому в случае $\mathcal{S}_z \cap \mathcal{S}^* = \emptyset$ немедленно получаем (3.32b). Однако множество $\mathcal{S}_z \cap \mathcal{S}^*$ может оказаться непустым при некоторых $z \notin \mathcal{N}_{\text{ani}} \setminus \partial\Omega$ (поскольку $z \in \mathcal{N}_{\text{iso}}$ или $z \in (\mathcal{N}_{\text{ani}} \setminus \mathcal{N}_{\partial\Omega}^*) \cap \partial\Omega$ может соединяться коротким ребром с узлом в $\mathcal{N}_{\text{ani}} \setminus \partial\Omega$). Если такой узел $z \in \mathcal{N}_{\text{reg}}$, то соотношение (3.8) (с $h_z \simeq H_z$) приводит к $|\alpha_S| \lesssim \|g\|_{1;\omega_z}$, а если $z \in (\mathcal{N}_{\text{ani}} \setminus \mathcal{N}_{\partial\Omega}^*) \cap \partial\Omega$, то нужно воспользоваться $\alpha_z = 0$ и $|\alpha_S - \alpha_z| \lesssim \|\nabla u_h\|_{1;\omega_z}$. Таким образом, получаем (3.32b) также и при любом $z \notin \mathcal{N}_{\text{ani}} \setminus \partial\Omega$. Доказательство (3.32), а следовательно и (3.25), завершено. \square

З а м е ч а н и е 3.12. Замечание 2.13 остается верным, т. е. можно

заменить величины $\{\bar{g}_z\}_{z \in \mathcal{N}}$ на $\{\bar{g}_z^*\}_{z \in \mathcal{N}}$, воспользовавшись либо (2.56a), либо (2.56b), и при этом будет выполняться (2.37), которое в обозначениях данного параграфа принимает вид

$$H_z |\bar{g}_z - \bar{g}_z^*| \lesssim \|g\|_{1;\omega_z}, \quad h_z H_z |\bar{g}_z^*| \lesssim \|g\|_{1;\omega_z} \lesssim (h_z H_z)^{1/2} \|g\|_{2;\omega_z}. \quad (3.38)$$

Далее, внимательное изучение вышеизложенного доказательства показывает, что при такой замене значений $\{\bar{g}_z\}_{z \in \mathcal{N}}$ из (2.24) на $\{\bar{g}_z^*\}_{z \in \mathcal{N}}$ соотношение (3.25) остается справедливым. Действительно, тогда в I_z появится дополнительное слагаемое $I_z^* := \varepsilon^2 \int_{\gamma_z} (\bar{g}_z - \bar{g}_z^*) \phi_z [\nabla u_h] \cdot \nu$, которое должно быть учтено в оценке (3.33). Для этого определение (3.34) для θ_z должно быть несколько изменено и принимает вид

$$\theta_z := \lambda_z^{-1} \varepsilon^2 \min \left\{ \|g\|_{1;\omega_z}^2, H_z^2 h_z^{-1} \|g\|_{2;\omega_z} \|g\|_{2;\omega_z} + H_z h_z^{-1} \|g\|_{2;\omega_z}^2 \right\}. \quad (3.39)$$

Прокомментируем включение слагаемого $H_z h_z^{-1} \|g\|_{2;\omega_z}^2$, отсутствующего в (3.34). Вспомним, что при использовании оценки (3.23) для $|I_z|$ уже была учтена компонента $\varepsilon^2 \int_{\gamma_z} |\bar{g}_z| \phi_z [\nabla u_h] \cdot \nu$ из I_z^* , а для оставшейся компоненты $\varepsilon^2 \int_{\gamma_z} |\bar{g}_z^*| \phi_z [\nabla u_h] \cdot \nu$ используется второе соотношение из (3.38), которое и приводит к включению $H_z h_z^{-1} \|g\|_{2;\omega_z}^2$ в (3.39).

Доказательство (3.26). Воспользовавшись Замечанием 3.12, заменим $\{\bar{g}_z\}$ в (3.14) на $\{\bar{g}_z^*\}$. Далее (3.26) доказывается аналогично (3.18) с использованием соотношений (3.38). \square

3.7 Квазиинтерполяционный оператор Π_h на анизотропных сетках

Апостериорные оценки в вышеизложенных Теоремах 3.5, 3.9 и 3.11 содержат величину Θ , определенную в (3.16), которая связана с интерполяционной ошибкой для функции Грина, и которую нам еще предстоит оценить. Начнем с наиболее простого случая.

З а м е ч а н и е 3.13. Если $H_z \gtrsim \varepsilon$ при любом $z \in \mathcal{N}$, то $\Theta \lesssim 1$. Действительно, в силу (3.11), полагая $G_h := 0$ немедленно получаем желаемую оценку. Заметим, что этот результат остается справедливым и в случае сильно анизотропных сеток (поскольку здесь мы не делаем никаких предположений о h_z или $\frac{h_z}{H_z}$).

Для оценки $g = G - G_h$ было бы очень удобно воспользоваться стандартными квазиинтерполянтами, удовлетворяющими $|G - G_h|_{k,p;T} \lesssim H_T^{j-k} |G|_{j,p;\omega_T}$ при любых $0 \leq k \leq j \leq 1$, $1 \leq p \leq \infty$ (как это сделано в §1 для случая локально квазирваномерных сеток). Однако в случае неструктурированных анизотропных сеток построение таких интерполянтов является весьма сложной задачей. Этот вопрос обсуждается, например, в Главе 3 монографии [56] для случая прямоугольных анизотропных сеток. Для случая же более общих анизотропных сеток оценки ошибки для стандартных квазиинтерполянтов получены в работе [90, §3], но только в норме L_2 , в то время обобщение результатов этой работы для градиента квазиинтерполяционной ошибки приводит к оценкам, являющимся недостаточно точными с точки зрения целей настоящего параграфа.

Поэтому здесь будет построен новый квазиинтерполяционный оператор Π_h для случая анизотропных сеток, с помощью которого, выбирая $G_h := \Pi_h G \in S_h$, удастся получить оценку $\Theta \lesssim 1$. Перейдем к построению данного оператора.

Для любого узла $z \in \mathcal{N}$, пусть σ_z будет некоторым ребром, исходящим из z . При $z \in \partial\Omega$, будем дополнительно требовать, чтобы $\sigma_z \subset \partial\Omega$. Теперь положим $\Pi_h v(z)$ равным среднему значению v на σ_z (которое также будет обозначаться $\mathcal{A}_{\sigma_z} v$):

$$\Pi_h v(z) := \mathcal{A}_{\sigma_z} v := |\sigma_z|^{-1} \int_{\sigma_z} v, \quad (3.40)$$

а затем $\Pi_h v := \sum_{z \in \mathcal{N}} \phi_z |\sigma_z|^{-1} \int_{\sigma_z} v$.

Далее, наложим важные дополнительные ограничения на выбор σ_z (при

этом никакие дополнительные условия на триангуляцию наложены не будут).

- При любом $z \in \mathcal{N} \setminus \mathcal{N}_{\partial\Omega}^*$, пусть $|\sigma_z| \simeq H_z$.
- Для любой пары узлов z и z' из $\mathcal{N}_{\text{ani}} \setminus \mathcal{N}_{\partial\Omega}^*$ с общим коротким ребром (т. е. $|z - z'| \simeq h_z$), σ_z и $\sigma_{z'}$ выбираются одинаково ориентированными в том смысле, что если эти два ребра соответственно соединяют z с узлом \hat{z} и z' с узлом \hat{z}' , то справедливо соотношение $|\hat{z} - \hat{z}'| \lesssim h_z$; см. Рис. 3.3.

Нам потребуется следующий вспомогательный результат.

Л е м м а 3.14. Для любого треугольника T с вершинами z, z', z'' и соответствующими противоположными ребрами S, S', S'' справедливо соотношение

$$|\mathcal{A}_{S'} v - \mathcal{A}_{S''} v| \lesssim |S||T|^{-1} \|\nabla v\|_{1;T}.$$

Д о к а з а т е л ь с т в о. Желаемая оценка немедленно вытекает из наблюдения, что, если мы введем ось ζ по направлению внутренней нормали к S , и положим $\hbar := 2|T||S|^{-1}$, то $\mathcal{A}_{S'} v - \mathcal{A}_{S''} v = \hbar^{-1} \int_0^{\hbar} (v|_{S'} - v|_{S''}) d\zeta$. \square

Л е м м а 3.15. (i) Если $z \in \mathcal{N}_{\text{ani}}$ является вершиной треугольника T , то

$$|T|^{1/2} |\Pi_{\hbar} v(z)| \lesssim \|v\|_{2;T} + H_T \|\nabla v\|_{2;\omega_z}. \quad (3.41)$$

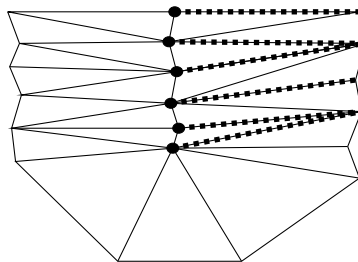


Рис. 3.3: Выбор σ_z в (3.40): допустимые отрезки σ_z выделены для каждого выделенного узла z .

(ii) Если же $z, z' \in \mathcal{N}_{\text{ani}}$ являются вершинами T , то, при $p = 1, 2$, справедливо соотношение

$$|T|^{1/p} \frac{|\Pi_h v(z) - \Pi_h v(z')|}{|z - z'|} \lesssim \|\nabla v\|_{p; \omega_z \cup \omega_{z'}}. \quad (3.42)$$

Д о к а з а т е л ь с т в о. (i) Достаточно рассмотреть лишь $z \in \mathcal{N}_{\text{ani}} \setminus \partial\Omega$, поскольку при $z \in \partial\Omega$ результат очевиден. Пусть S является любым длинным ребром в T , т. е. $|S| \simeq H_T$. Тогда (возможно, многократное) применение Леммы 3.14 приводит к $|\Pi_h v(z) - \mathcal{A}_S v| = |\mathcal{A}_{\sigma_z} v - \mathcal{A}_S v| \lesssim h_T^{-1} \|\nabla v\|_{1; \omega_z}$. Теперь применение варианта Леммы 3.1 для T (с областью ω_z , замененной на T , а $\gamma_z \setminus \dot{\gamma}_z$ на S , а также h_z на h_T и H_z на H_T) приводит к $|\mathcal{A}_S v| \lesssim H_T^{-1} \|v\|_{1; S} \lesssim h_T^{-1} \|v\|_{1; T}$. Поэтому получаем оценку

$$|\Pi_h v(z)| \lesssim (h_T H_T)^{-1} \|v\|_{1; T} + h_T^{-1} \|\nabla v\|_{1; \omega_z}, \quad (3.43)$$

эквивалентную (3.41) в силу $|T| \simeq h_T H_T \simeq |\omega_z|$; последнее справедливо поскольку $z \in \mathcal{N}_{\text{ani}}$.

(ii) Рассмотрим узлы $z, z' \in \mathcal{N}_{\text{ani}} \setminus \mathcal{N}_{\partial\Omega}^*$. Если $|z - z'| \simeq H_T$, то воспользовавшись вышеизложенными рассуждениями, получаем $|\Pi_h v(z) - \Pi_h v(z')| \lesssim h_T^{-1} \|\nabla v\|_{1; \omega_z \cup \omega_{z'}}$. Если же $|z - z'| \simeq h_T$, то применение Леммы 3.14 (при этом важно принять во внимание одинаковую ориентацию ребер σ_z и $\sigma_{z'}$) приводит к $|\Pi_h v(z) - \Pi_h v(z')| \lesssim H_T^{-1} \|\nabla v\|_{1; \omega_z \cup \omega_{z'}}$. Объединяя эти два случая получаем желаемую оценку (3.42) (опять воспользовавшись $|T| \simeq h_T H_T \simeq |\omega_z \cup \omega_{z'}|$).

Если же оба узла $z, z' \in \partial\Omega$, то результат очевиден. Остается рассмотреть случай, когда $z \in \mathcal{N}_{\partial\Omega}^*$, а $z' \notin \partial\Omega$. Тогда $|z - z'| \simeq H_T$, и мы опять получаем (3.42). \square

С л е д с т в и е 3.16. Если все вершины z, z', z'' треугольника $T \in \mathcal{T}$ принадлежат множеству \mathcal{N}_{ani} , то

$$H_T^{-1} \|v - \Pi_h v\|_{2; T} + \|\nabla(v - \Pi_h v)\|_{2; T} \lesssim \|\nabla v\|_{2; \omega_T},$$

где $\omega_T = \omega_z \cup \omega_{z'} \cup \omega_{z''}$ является множеством элементов, касающихся T .

Доказательство. Поскольку $\Pi_h v \in S_h$, то, объединяя (3.41) и (3.42), немедленно получаем $\|\Pi_h v\|_{2;T} \lesssim \|v\|_{2;T} + H_T \|\nabla v\|_{2;\omega_T}$ и $\|\nabla \Pi_h v\|_{2;T} \lesssim \|\nabla v\|_{2;\omega_T}$. Далее, для любой постоянной C справедливо $\Pi_h C = C$. Выберем C таким образом, чтобы $\|v - C\|_{2;T} \lesssim H_T \|\nabla v\|_{2;T}$. Следовательно

$$\|v - \Pi_h v\|_{2;T} = \|(v - C) - \Pi_h(v - C)\|_{2;T} \lesssim H_T \|\nabla v\|_{2;\omega_T}.$$

Это приводит к желаемой оценке. \square

Замечание 3.17 (Квазиинтерполяционный оператор Π_h^*). Заметим, что построенный оператор Π_h устойчив в нормализованной норме H^1 , но не в норме L_2 , в то время как последнее свойство нам понадобится при оценке Θ в случае $\varepsilon \lesssim H_T$. Поэтому мы несколько модифицируем определение Π_h для узлов $z \notin \partial\Omega$, что будет сделано в два шага.

Во-первых, введем квазиинтерполяционный оператор $\tilde{\Pi}_h$, заменив в определении Π_h каждое из ребер σ_z его половинкой $\tilde{\sigma}_z$, исходящей из z (см. Рис. 3.4). При этом и Лемма 3.15, и Следствие 3.16 остаются справедливыми. (В частности, в доказательстве Леммы 3.15(ii) при рассмотрении случая $|z - z'| \simeq H_T$ теперь понадобится оценка $|\mathcal{A}_{\tilde{\sigma}_z} v - \mathcal{A}_{\sigma_z \setminus \tilde{\sigma}_z} v| \lesssim h_T^{-1} \|\nabla v\|_{1;\omega_z}$. Последняя оценка может быть получена троекратным применением Леммы 3.14 для любого треугольника с ребром σ_z .)

Построение Π_h^* завершается заменой каждой половинки ребра $\tilde{\sigma}_z$ на параллелограмм σ_z^* , построенный на $\tilde{\sigma}_z$ и прилегающей к нему половине любого короткого ребра из $\hat{\gamma}_z$ как на сторонах (см. Рис 3.4; заметим, что всегда

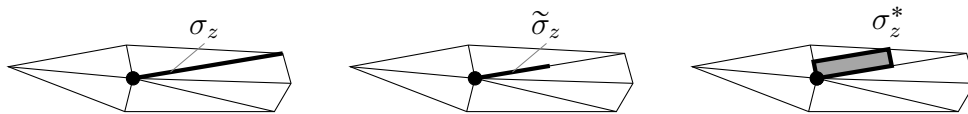


Рис. 3.4: Выбор половинки ребра $\tilde{\sigma}_z$ (центр) и параллелограмма σ_z^* (справа), описанных в Замечании 3.17.

можно выбрать ребро σ_z со всеми вышеперечисленными свойствами таким образом, что одно из прилегающих к нему ребер принадлежит $\hat{\gamma}_z$). Весьма важно, что выполняется соотношение $|\sigma_z^*| \simeq |\omega_z|$. Далее, положим $\Pi_h^* v(z) := \mathcal{A}_{\sigma_z^*} v$. Тогда несложно убедиться, что

$$|\omega_z|^{1/2} |\Pi_h^* v(z)| \lesssim \|v\|_{2;\omega_z}, \quad |\omega_z| |\Pi_h^* v(z) - \tilde{\Pi}_h v(z)| \lesssim h_z \|\nabla v\|_{1;\omega_z}. \quad (3.44)$$

Здесь первое соотношение приводит к $\|\Pi_h^* v\|_{2;T} \lesssim \|v\|_{2;\omega_T}$, из второго же следует справедливость оценок (3.41) и (3.42) для Π_h^* , т. е. и Лемма 3.15, и Следствие 3.16 остаются справедливыми для Π_h^* . Объединяя все эти свойства, получаем, что при $j = 0, 1$ справедливы соотношения

$$\|v - \Pi_h^* v\|_{2;T} \lesssim H_T^j |v|_{j,2;\omega_T}, \quad \|\nabla(v - \Pi_h^* v)\|_{2;T} \lesssim \|\nabla v\|_{2;\omega_T}. \quad (3.45)$$

Т е о р е м а 3.18. *При $\mathcal{N} = \mathcal{N}_{\text{ani}}$ (т. е. если все узлы являются анизотропными), существует такой интерполянт $G_h \in S_h$, что для Θ из (3.16) справедлива оценка $\Theta \lesssim 1$.*

Д о к а з а т е л ь с т в о. Положим $G_h := \Pi_h^* G$ и применим (3.45). Тогда $\Theta \lesssim \| \|G\|_{\varepsilon;\Omega} \lesssim 1$, где мы также воспользовались тем, что $H_T \simeq \text{diam } \omega_T \simeq H_z$ при любом $T \subset \omega_z$. (Если $H_T \lesssim \varepsilon$ при любом $T \in \mathcal{T}$, то, в силу Следствия 3.16, можно также воспользоваться более простым интерполянтом $G_h := \Pi_h G$.) \square

С л е д с т в и е 3.19. *При $\mathcal{N} = \mathcal{N}_{\text{ani}} \cup \mathcal{N}_{\text{reg}}$, существует такой интерполянт $G_h \in S_h$, что для Θ из (3.16) справедлива оценка $\Theta \lesssim 1$.*

Д о к а з а т е л ь с т в о. Достаточно заметить, что вариант Теоремы 3.18 справедлив и для локально квазирваномерных сеток. \square

З а м е ч а н и е 3.20. *Вариант Теоремы 3.18 для более общих типов узлов, определенных в §2.2, получен при некоторых дополнительных предположениях о триангуляции в [169, §7.2].*

Таблица 3.1: Значения ошибки $\{\varepsilon^2 \|\nabla u_h - (\nabla u)^I\|_{2;\Omega}^2 + \|u_h - u^I\|_{2;\Omega}^2\}^{1/2}$, оценителя $\{\mathcal{E}_{(3.25)}^2 + \mathcal{E}_{(3.27)}^2\}^{1/2}$ и отношения $\mathcal{E}_{(3.25)}$ к $\mathcal{E}_{(3.27)}$.

N	$\varepsilon = 1$	$\varepsilon = 2^{-5}$	$\varepsilon = 2^{-10}$	$\varepsilon = 2^{-15}$	$\varepsilon = 2^{-20}$	$\varepsilon = 2^{-25}$	$\varepsilon = 2^{-30}$
	Ошибки (нечетные ряды) & Порядки Сходимости (четные ряды)						
64	3.202e-2	5.081e-3	7.993e-4	1.408e-4	2.489e-5	4.399e-6	7.777e-7
	1.00	0.99	1.00	1.00	1.00	1.00	1.00
128	1.602e-2	2.564e-3	3.991e-4	7.028e-5	1.242e-5	2.196e-6	3.882e-7
	1.00	0.99	1.00	1.00	1.00	1.00	1.00
256	8.011e-3	1.289e-3	1.997e-4	3.511e-5	6.207e-6	1.097e-6	1.940e-7
	1.00	1.00	0.99	1.00	1.00	1.00	1.00
512	4.006e-3	6.461e-4	1.002e-4	1.755e-5	3.102e-6	5.484e-7	9.695e-8
	Оценители (нечетные ряды) & Индексы Эффективности (четные ряды)						
64	1.041e-1	2.102e-2	4.129e-3	7.393e-4	1.308e-4	2.311e-5	4.086e-6
	3.25	4.14	5.17	5.25	5.25	5.25	5.25
128	5.147e-2	1.051e-2	2.050e-3	3.711e-4	6.566e-5	1.161e-5	2.052e-6
	3.21	4.10	5.14	5.28	5.29	5.29	5.29
256	2.559e-2	5.269e-3	1.006e-3	1.858e-4	3.290e-5	5.817e-6	1.028e-6
	3.19	4.09	5.04	5.29	5.30	5.30	5.30
512	1.276e-2	2.645e-3	4.883e-4	9.287e-5	1.647e-5	2.912e-6	5.147e-7
	3.19	4.09	4.87	5.29	5.31	5.31	5.31
	Значения Отношения $\mathcal{E}_{(3.25)}/\mathcal{E}_{(3.27)}$						
64	1.66	1.31	0.82	0.81	0.81	0.81	0.81
128	1.67	1.48	0.83	0.80	0.80	0.80	0.80
256	1.67	1.57	0.86	0.80	0.79	0.79	0.79
512	1.66	1.60	0.93	0.79	0.79	0.79	0.79

3.8 Численные результаты

Протестируем оценители, вытекающие из Теорем 3.9 и 3.11, на простой тестовой задаче, которая получается из (3.1) при $\Omega = (0, 1)^2$ и $f = u - F(x, y)$, и где F соответствует точному решению $u = 4y(1 - y)[1 - x^2 - (e^{-x/\varepsilon} - e^{-1/\varepsilon})/(1 - e^{-x/\varepsilon})]$ (последнее имеет пограничный слой вдоль $x = 0$).

Таблица 3.2: Значения компоненты ошибки $\|u - u^I\|_{2;\Omega}$ из (3.47), оценивателя $\mathcal{E}_{(3.15)}$ и его индексов эффективности.

N	$\varepsilon = 1$	$\varepsilon = 2^{-5}$	$\varepsilon = 2^{-10}$	$\varepsilon = 2^{-15}$	$\varepsilon = 2^{-20}$	$\varepsilon = 2^{-25}$	$\varepsilon = 2^{-30}$
Ошибки $\ u - u^I\ _{2;\Omega}$ (нечетные ряды) & Порядки Сходимости (четные ряды)							
64	2.242e-4	6.120e-4	6.496e-4	6.567e-4	6.571e-4	6.571e-4	6.571e-4
	2.00	2.00	2.01	2.01	2.01	2.01	2.01
128	5.607e-5	1.525e-4	1.617e-4	1.635e-4	1.636e-4	1.636e-4	1.636e-4
	2.00	2.00	2.00	2.00	2.00	2.00	2.00
256	1.402e-5	3.807e-5	4.036e-5	4.077e-5	4.080e-5	4.080e-5	4.080e-5
	2.00	2.00	2.00	2.00	2.00	2.00	2.00
512	3.505e-6	9.510e-6	1.008e-5	1.018e-5	1.019e-5	1.019e-5	1.019e-5
$\mathcal{E}_{(3.15)}$ (нечетные ряды) & Индексы эффективности $\mathcal{E}_{(3.15)}/\ u - u^I\ _{2;\Omega}$ (четные ряды)							
64	2.661e-3	5.762e-4	6.484e-4	6.567e-4	6.571e-4	6.571e-4	6.571e-4
	11.87	0.94	1.00	1.00	1.00	1.00	1.00
128	6.671e-4	1.435e-4	1.614e-4	1.634e-4	1.636e-4	1.636e-4	1.636e-4
	11.90	0.94	1.00	1.00	1.00	1.00	1.00
256	1.670e-4	3.579e-5	4.029e-5	4.077e-5	4.079e-5	4.080e-5	4.080e-5
	11.91	0.94	1.00	1.00	1.00	1.00	1.00
512	4.178e-5	8.938e-6	1.006e-5	1.018e-5	1.019e-5	1.019e-5	1.019e-5
	11.92	0.94	1.00	1.00	1.00	1.00	1.00

Здесь будет рассмотрена одна априорно построенная адаптивная сетка типа Бахвалова, описанная в §2.5.4.

Объединяя оценки Теорем 3.9 и 3.11, приходим к апостериорной оценке

$$\|u_h - u\|_{\varepsilon;\Omega} \lesssim \mathcal{E} := \left\{ \mathcal{E}_{(3.25)}^2 + \mathcal{E}_{(3.27)}^2 + \mathcal{E}_{(3.15)}^2 \right\}^{1/2}, \quad (3.46)$$

где обозначение $\mathcal{E}_{(\cdot)}$ используется для правой части соответствующей оценки (\cdot) (например, $\mathcal{E}_{(3.15)} = \|f_h - f_h^I\|_{2;\Omega}$). При этом положим $\Theta := 1$ во всех используемых оценках, а в $\mathcal{E}_{(3.15)}$ будем использовать $\mathcal{T}_0 := \emptyset$. Также величины H_z в (3.4) были заменены на $\max_{T \subset \omega_z} H_T \simeq H_z$, а величины типа $\min\{1, a \varepsilon^{-1}\}$ были заменены соответствующими более гладкими значениями

Таблица 3.3: Значения оценителя $\mathcal{E}_{(3.26)}$ и его индексов эффективности (вычисляемых как отношение $\mathcal{E}_{(3.26)}$ к правой части оценки (3.47)).

N	$\varepsilon = 1$	$\varepsilon = 2^{-5}$	$\varepsilon = 2^{-10}$	$\varepsilon = 2^{-15}$	$\varepsilon = 2^{-20}$	$\varepsilon = 2^{-25}$	$\varepsilon = 2^{-30}$
	Оценители (нечетные ряды) & Индексы Эффективности (четные ряды)						
64	1.268e-1	4.661e-2	1.565e-2	2.850e-3	5.043e-4	8.916e-5	1.576e-5
	3.93	8.19	10.80	3.57	0.74	0.13	0.02
128	6.449e-2	3.111e-2	1.519e-2	2.847e-3	5.043e-4	8.915e-5	1.576e-5
	4.01	11.45	27.08	12.18	2.87	0.54	0.10
256	3.253e-2	1.868e-2	1.435e-2	2.842e-3	5.042e-4	8.915e-5	1.576e-5
	4.05	14.08	59.76	37.45	10.73	2.13	0.38
512	1.634e-2	1.039e-2	1.291e-2	2.831e-3	5.042e-4	8.915e-5	1.576e-5
	4.07	15.84	117.06	102.07	37.94	8.30	1.53

$\frac{a}{\varepsilon+a}$ (например, вместо λ'_z использовалось $\frac{H_z}{\varepsilon+H_z}$). Наконец, во всех вычислениях вместо функций f_h и u использовались соответствующие квадратичные интерполлянты Лагранжа.

Для ошибки численного решения в энергетической норме разумно предположить, что

$$\|u_h - u\|_{\varepsilon;\Omega} \simeq \{\varepsilon^2 \|\nabla u_h - (\nabla u)^I\|_{2;\Omega}^2 + \|u_h - u^I\|_{2;\Omega}^2\}^{1/2} + \|u - u^I\|_{2;\Omega}. \quad (3.47)$$

Данное представление ошибки оказывается полезным, поскольку две компоненты ошибки в правой части (3.47) вели себя несколько по-разному в наших экспериментах: соответственно как $\simeq \varepsilon^{1/2}N^{-1}$ и $\simeq N^{-2}$ (для этого достаточно сравнить верхние части Таблиц 3.1 и 3.2; см. также Таблицу 3.4 и Замечание 3.21). Более того, эксперименты показывают, что компоненты оценителя $\mathcal{E}_{(3.25)} \simeq \mathcal{E}_{(3.27)}$ хорошо оценивают первую компоненту ошибки в правой части (3.47) (см. Таблицу 3.1), в то время как $\mathcal{E}_{(3.15)}$ является хорошим оценителем для второй компоненты ошибки в (3.47) (см. Таблицу 3.2). Отметим также, что для оценителей, представленных в Таблицах 3.1 и 3.2, индексы эффективности (вычисляемые как отношение оценителя к ошибке) не превосходят 11.92. В Таблице 3.1 также приведены значения отношения $\mathcal{E}_{(3.25)}$ к

Таблица 3.4: Компоненты ошибки (см (3.47)) и соответствующие порядки сходимости.

N	$\varepsilon = 1$	$\varepsilon = 2^{-5}$	$\varepsilon = 2^{-10}$	$\varepsilon = 2^{-15}$	$\varepsilon = 2^{-20}$	$\varepsilon = 2^{-25}$	$\varepsilon = 2^{-30}$
	Ошибки $\varepsilon^{1/2}\ \nabla u_h - (\nabla u)^I\ _{2;\Omega}$ (нечетные ряды) & Порядки Сходимости (четные ряды)						
64	3.202e-2	2.872e-2	2.532e-2	2.523e-2	2.522e-2	2.522e-2	2.522e-2
	1.00	0.99	1.00	1.00	1.00	1.00	1.00
128	1.602e-2	1.450e-2	1.265e-2	1.260e-2	1.260e-2	1.260e-2	1.260e-2
	1.00	0.99	1.00	1.00	1.00	1.00	1.00
256	8.011e-3	7.290e-3	6.336e-3	6.297e-3	6.297e-3	6.297e-3	6.297e-3
	1.00	1.00	0.99	1.00	1.00	1.00	1.00
512	4.006e-3	3.655e-3	3.186e-3	3.148e-3	3.148e-3	3.148e-3	3.148e-3
	Ошибки $\varepsilon^{-1/2}\ u_h - u^I\ _{2;\Omega}$ (нечетные ряды) & Порядки Сходимости (четные ряды)						
64	1.714e-4	1.173e-3	3.612e-3	3.624e-3	3.624e-3	3.624e-3	3.624e-3
	2.00	2.32	1.06	1.04	1.04	1.04	1.04
128	4.290e-5	2.351e-4	1.736e-3	1.757e-3	1.757e-3	1.757e-3	1.757e-3
	2.00	2.47	1.07	1.02	1.02	1.02	1.02
256	1.073e-5	4.254e-5	8.267e-4	8.643e-4	8.643e-4	8.643e-4	8.643e-4
	2.00	2.44	1.17	1.01	1.01	1.01	1.01
512	2.682e-6	7.826e-6	3.667e-4	4.286e-4	4.287e-4	4.287e-4	4.287e-4

$\mathcal{E}_{(3.27)}$, которые, как мы видим, остаются между 0.79 и 1.67.

Далее, рассмотрим значения оценителя, получаемого из Теоремы 3.9, который является вариантом (3.46), но с $\mathcal{E}_{(3.27)}$, замененным на менее точный оценитель для поэлементных невязок $\mathcal{E}_{(3.26)}$. Таблица 3.3 показывает, что при использовании $\mathcal{E}_{(3.26)}$ при некоторых значениях ε и N эффективность глобального оценителя значительно ухудшается.

З а м е ч а н и е 3.21 (Сравнение энергетической нормы и нормы L_2). При $\varepsilon \ll 1$ Таблица 3.4 демонстрирует, что $\|u_h - u^I\|_{2;\Omega} \simeq \varepsilon \|\nabla u_h - (\nabla u)^I\|_{2;\Omega} \simeq \varepsilon^{1/2} N^{-1}$ (последнее наблюдение согласуется с численными результатами, представленными в Таблице 4 в работе [167]). С другой стороны, верхняя

часть Таблицы 3.2 показывает, что соответствующая ошибка интерполяции $\|u - u^I\|_{2;\Omega} \simeq N^{-2}$. Эти результаты предполагают, что при $\varepsilon \ll 1$ не стоит ожидать, что $\|u_h - u\|_{2;\Omega} \ll \|u_h - u\|_{\varepsilon;\Omega}$, поскольку и $\|u_h - u\|_{2;\Omega}$, и $\|u_h - u^I\|_{2;\Omega}$ могут быть того же порядка, что и $\|u_h - u\|_{\varepsilon;\Omega}$.

При рассмотренных значениях ε и N аспектные соотношения элементов сетки принимали значения от 1 до $3.6e+8$. При этом оценитель \mathcal{E} из (3.46) работает достаточно хорошо, а его индексы эффективности стабилизируются при $\varepsilon \rightarrow 0$. Использование же менее точной альтернативы $\mathcal{E}_{(3.26)}$ для компоненты $\mathcal{E}_{(3.27)}$ в глобальном оценителе (3.46) при некоторых значениях ε и N оказывается неудовлетворительным.

Глава 2

Апостериорные оценки в норме максимума модуля для полулинейных параболических уравнений с использованием эллиптических реконструкций

§4 Апостериорные оценки для полудискретных методов (без дискретизации по пространству)

Для временных полудискретизаций полулинейных параболических уравнений второго порядка получены апостериорные оценки в норме максимума модуля для ошибки соответствующих численных решений. При этом рассматриваются неявный метод Эйлера, метод Кранка-Николсон и разрывный метод Галеркина $dG(r)$ с квадратурой Радо. Полученные результаты являются новыми и для классических, и для сингулярно возмущенных уравнений. Простота предложенного теоретического подхода связана с использованием интерполянтов численного решения по времени, степень которых соответствует порядку метода (например, при анализе неявного метода Эйлера используется кусочно-постоянная интерполяция) и является более низкой по сравнению с типично используемыми в литературе интерполянтами. Также важную роль в представленном анализе играют оценки для функции Грина параболического оператора.

Основной материал параграфа опубликован в работе [175].

4.1 Введение

Рассмотрим полулинейное параболическое уравнение в виде

$$\mathcal{M}u := \partial_t u + \mathcal{L}u + f(x, t, u) = 0 \quad \text{при } (x, t) \in Q := \Omega \times (0, T], \quad (4.1a)$$

где $\mathcal{L} = \mathcal{L}(t)$ является линейным эллиптическим оператором второго порядка, а пространственная область $\Omega \subset \mathbb{R}^n$ имеет липшицеву границу. При этом накладываются следующие начальное и граничное условия:

$$u(x, 0) = \varphi(x) \quad \text{при } x \in \bar{\Omega}, \quad u(x, t) = 0 \quad \text{при } (x, t) \in \partial\Omega \times [0, T]. \quad (4.1b)$$

Будем предполагать, что функция f непрерывна в $\bar{\Omega} \times [0, T] \times \mathbb{R}$, дифференцируема по третьему аргументу, а также при некоторых неотрицательных постоянных γ и $\bar{\gamma}$ удовлетворяет соотношениям

$$0 \leq \gamma^2 \leq \partial_z f(x, t, z) \leq \bar{\gamma}^2 \quad \text{при } (x, t, z) \in \bar{\Omega} \times [0, T] \times \mathbb{R}. \quad (4.2)$$

Целью настоящего параграфа является получение явных апостериорных оценок ошибки некоторых полудискретных методов (без дискретизации по пространству) для задачи (4.1). Будут рассмотрены неявный метод Эйлера, метод Кранка-Николсон и разрывный метод Галеркина dG(r) с квадратурой Радо.

Полученные результаты будут обсуждаться в контексте модельного уравнения с $\mathcal{L} := -\varepsilon^2 \Delta = -\varepsilon^2 \sum_{i=1}^n \partial_{x_i}^2$:

$$\mathcal{M}u := \partial_t u - \varepsilon^2 \Delta u + f(x, t, u) = 0 \quad (4.3)$$

в ограниченной полигональной области $\Omega \subset \mathbb{R}^n$ при $n = 1, 2, 3$. Последнее уравнение будет рассмотрено в двух режимах:

$$(i) \quad \varepsilon = 1, \quad \gamma \geq 0; \quad (ii) \quad \varepsilon \ll 1, \quad \gamma > 0.$$

Заметим, что режим (ii) соответствует сингулярно возмущенному уравнению реакции-диффузии, для решений которого характерны резкие пограничные

и внутренние слои. Весьма важно, чтобы апостериорные оценки, полученные для данного случая, были робастными в том смысле, что любая зависимость от малого параметра ε должна быть представлена явным образом [112, 132], а также эффективность таких оценок не должна ухудшаться при уменьшении ε .

Ошибка решения в данной главе будет оцениваться в норме максимума модуля, поскольку эта норма является достаточно сильной и обеспечивает достаточную точность численного решения в окрестностях сингулярностей, а также в пограничных и внутренних слоях, что в особенности важно, если уравнение (4.1) является сингулярно возмущенным. Полученные нами оценки будут типа интерполяционных в том смысле, что они будут включать слагаемые, которые можно интерпретировать как аппроксимирующие $\tau_j^p |\partial_t^p u|$, где p равняется порядку метода, а τ_j обозначает локальный шаг по времени.

Важную роль в представленном анализе играют оценки для функции Грина параболического оператора, которые будут установлены аналогично работе [75], но с той существенной разницей, что полученные здесь оценки применимы для случая полулинейного и сингулярно возмущенного уравнения (4.3), в то время как в [75] рассматривался линейный параболический оператор $\partial_t - \Delta$.

Простота предлагаемого здесь теоретического подхода связана с использованием интерполянтов численного решения по времени степени $p - 1$ (для метода порядка p) и является более низкой по сравнению с типично используемыми в литературе интерполянтами [54, 75]. Например, при анализе неявного метода Эйлера используется кусочно-постоянная интерполяция. При анализе же методов Кранка-Николсон и dG(1) соответственно используются кусочно-линейные и кусочно-квадратичные интерполянты численного решения по времени. Это упрощает анализ, но при этом в невязках численного решения присутствуют производные разрывных функций, которые понима-

ются как обобщенные функции.

Заметим, что более ранние апостериорные оценки для параболических уравнений в норме максимума модуля были получены либо для регулярных линейных задач [83, 62, 75, 77], либо же не являются робастными, поскольку в них присутствуют нежелательные отрицательные степени малого параметра ε [62]. Более детальное сравнение полученных здесь оценок с более ранними результатами будет представлено в Замечаниях 4.12, 5.12, 5.14 и 5.25.

Параграф организован следующим образом. §4.2 будет посвящен функции Грина параболического оператора; в частности, будет установлена лемма, описывающая некоторую априорную оценку для рассматриваемой параболической задачи, которая будет играть ключевую роль в нашем анализе как здесь, так и в последующем §5. Содержание параграфов §§4.3–4.6, а также соответствующих §§5.2–5.5 для полностью дискретных методов описано в следующей таблице.

	Описание результатов	Неявный Эйлер	Кранк-Николсон	$dG(r)$
Полудискретизации	§4.3	§4.4	§4.5	§4.6
Полные дискретизации	§5.2	§5.3	§5.4	§5.5

Обозначения. В настоящей главе C , а также c будут обозначать некоторую постоянную, принимающую разные значение в разных формулах, которая может зависеть от Ω и f , но не зависит от шагов по времени, а также от ε . Будем писать $|x|$ для обозначения евклидовой нормы $x \in \mathbb{R}^n$. Будут использоваться стандартные пространства функций $C(\bar{\Omega})$ и $L_p(\Omega)$ при $1 \leq p \leq \infty$ с нормой $\|\cdot\|_{p,\Omega}$, а также стандартное пространство $H_0^1(\Omega)$ функций из $W_2^1(\Omega)$, обращающихся в нуль на $\partial\Omega$, в то время как $\langle \phi, \psi \rangle = \int_{\Omega} \phi(x)\psi(x) dx$ будет обозначать скалярное произведение в $L_2(\Omega)$.

Обобщенные функции, соглашение о непрерывности слева. Некоторые функции будут пониматься как обобщенные [93], на что в большинстве случаев будет указываться. Функции же, интегрируемые по Лебегу в области

$\Omega \times (0, T)$, будем называть регулярными. При рассмотрении любой кусочно-непрерывной регулярной функции, будем считать, что она является непрерывной слева при любом $t \in (0, T]$. В частности, данное соглашение будет удобно использовать при рассмотрении кусочно-непрерывных производных по времени.

4.2 Функция Грина для параболического оператора

Перейдем к рассмотрению функции Грина \mathcal{G} , ассоциируемой с параболическим оператором \mathcal{M} из (4.1). Наш интерес к функции Грина обусловлен тем, что она будет использоваться при представлении ошибки численного решения через невязки.

Определения и свойства фундаментальных решений и функций Грина для параболических операторов с переменными коэффициентами рассматриваются, например в Главе 1 и в §7 Главы 3 монографии [89]. Для любой пары ограниченных функций v и w , обращающихся в нуль на $\partial\Omega$, стандартная лиnearизация приводит к представлению $\mathcal{M}v - \mathcal{M}w = [\partial_t + \mathcal{L} + a(x, t)](v - w)$, в котором $a(x, t) := \int_0^1 \partial_z f(x, t, w + z[v - w]) dz$. Отсюда следует, что $v - w$ допускает представление

$$\begin{aligned} [v - w](x, t) &= \int_{\Omega} \mathcal{G}(x, t; \xi, 0) [v - w](\xi, 0) d\xi \\ &\quad + \int_0^t \int_{\Omega} \mathcal{G}(x, t; \xi, s) [\mathcal{M}v - \mathcal{M}w](\xi, s) d\xi ds \end{aligned} \quad (4.4)$$

через функцию Грина \mathcal{G} , которая определяется следующим образом. При фиксированных $(x, t) \in Q$, функция Грина $\mathcal{G}(x, t; \xi, s) =: \Gamma(\xi, s)$ является решением сопряженной задачи с конечным условием:

$$[-\partial_s - \mathcal{L}^* + a(\xi, s)] \Gamma(\xi, s) = 0 \quad \text{при } (\xi, s) \in \Omega \times [0, t), \quad (4.5a)$$

$$\Gamma(\xi, t) = \delta(\xi - x) \quad \text{при } \xi \in \Omega, \quad (4.5b)$$

$$\Gamma(\xi, s) = 0 \quad \text{при } (\xi, s) \in \partial\Omega \times [0, t]. \quad (4.5c)$$

Здесь используется дельта-функция Дирака $\delta(\cdot)$ в \mathbb{R}^n [93], а также оператор \mathcal{L}^* , сопряженный линейному оператору \mathcal{L} .

В представленном здесь апостериорном анализе будет предполагаться выполнение следующего условия.

У с л о в и е 4.1. Найдутся такие постоянные $\kappa_0, \kappa_1 > 0$ и $\kappa_2 \geq 0$, что функция Грина \mathcal{G} из (4.5), (4.2) удовлетворяет соотношениям

$$\|\mathcal{G}(x, t; \cdot, s)\|_{1, \Omega} \leq \kappa_0 e^{-\gamma^2(t-s)}, \quad \int_0^{t-\tau} \|\partial_s \mathcal{G}(x, t; \cdot, s)\|_{1, \Omega} ds \leq \kappa_1 \ell(\tau, t) + \kappa_2,$$

где $x \in \Omega$, $\tau \in (0, t]$, $t \in (0, T]$, а $\ell(\tau, t) := \int_\tau^t s^{-1} e^{-\frac{1}{2}\gamma^2 s} ds \leq \ln(t/\tau)$.

Заметим, что наша модельная задача удовлетворяет данному условию.

Л е м м а 4.2. Пусть $\varepsilon \in (0, 1]$ и $\gamma \geq 0$. При выполнении (4.2) модельная задача (4.3) удовлетворяет Условию 4.1 с постоянными $\kappa_0 := 1$, $\kappa_1 := \frac{3^n}{2^{n/2+1}}$ и некоторой независимой от ε постоянной $\kappa_2 \geq 0$. Если же $f(x, t, z) = a(x)z + b(x, t)$, то $\kappa_2 = 0$. В общем случае $\kappa_2 = (\bar{\gamma}^2 - \gamma^2) \hat{\kappa}_2$, где $\hat{\kappa}_2 = \hat{\kappa}_2(\gamma)$ при $\gamma > 0$ и $\hat{\kappa}_2 = \hat{\kappa}_2(T)$ при $\gamma = 0$.¹

Д о к а з а т е л ь с т в о. Доказательство приводится в §4.7. □

Вышеописанное Условие 4.1 будет использоваться в следующей лемме, играющей ключевую роль в нашем анализе. Данная лемма будет сформулирована в контексте произвольной неравномерной сетки по времени:

$$0 = t_0 < t_1 < t_2 < \dots < t_M = T, \quad \tau_j = t_j - t_{j-1} \quad \text{при } j = 1, \dots, M. \quad (4.6)$$

¹Постоянные κ_0 и κ_1 в Лемме 4.2 являются достаточно точными. Например, в частном случае задачи (4.3) с постоянными коэффициентами $\partial_t u - \varepsilon^2 \partial_x^2 u + \gamma^2 u = b(x, t)$ в пространственной области $\Omega := \mathbb{R}$ несложно убедиться [174], что $\|\mathcal{G}(x, t; \cdot, s)\|_{1, \Omega} = e^{-\gamma^2(t-s)}$ и $\|\partial_s \mathcal{G}(x, t; \cdot, s)\|_{1, \Omega} \leq (\sqrt{2(\pi e)^{-1}}(t-s)^{-1} + \gamma^2) e^{-\gamma^2(t-s)}$. Таким образом, Условие 4.1 удовлетворяется с $\kappa_0 = 1$ (т. е. с тем же значением, что и в Лемме 4.2) и $\kappa_1 = \sqrt{2(\pi e)^{-1}} \approx 0.48$, $\kappa_2 = 1$, в то время как Лемма 4.2 дает несколько большее значение $\kappa_1 = 3 \cdot 2^{-3/2} \approx 1.06$, при этом $\kappa_2 = 0$.

Л е м м а 4.3. *Предположим, что параболический оператор \mathcal{M} из (4.1) удовлетворяет (4.2), а также Условию 4.1, а функции v, w заданы и ограничены в $\bar{\Omega} \times [0, T]$. Пусть также $v(\cdot, t), w(\cdot, t) \in H_0^1(\Omega) \cap C(\bar{\Omega})$ при любом $t \in [0, T]$, а*

$$\mathcal{M}v - \mathcal{M}w = \partial_t \mu + \vartheta \quad \text{в } Q. \quad (4.7)$$

При этом пусть функция μ является непрерывной и ограниченной на $[t_0, t_1]$ и каждым $(t_{j-1}, t_j]$, а $\partial_t \mu$ является непрерывной и ограниченной на $(t_{m-1}, t_m]$ при некотором $1 \leq m \leq M$, в то время как $\|\vartheta(\cdot, s)\|_{\infty, \Omega}$ интегрируема по s на $(0, t_m)$. Тогда

$$\begin{aligned} & \| [v - w](\cdot, t_m) \|_{\infty, \Omega} \\ & \leq \kappa_0 e^{-\gamma^2 t_m} \| [v - w - \mu](\cdot, 0) \|_{\infty, \Omega} + (\kappa_1 \ell_m + \kappa_2) \sup_{s \in [0, t_{m-1}]} \| \mu(\cdot, s) \|_{\infty, \Omega} \\ & \quad + \kappa_0 \lim_{s \rightarrow t_{m-1}^+} \| \mu(\cdot, s) \|_{\infty, \Omega} + \kappa_0 \tau_m \sup_{s \in (t_{m-1}, t_m]} \| \partial_s \mu(\cdot, s) \|_{\infty, \Omega} \\ & \quad + \kappa_0 \int_0^{t_m} e^{-\gamma^2(t_m-s)} \| \vartheta(\cdot, s) \|_{\infty, \Omega} ds, \end{aligned} \quad (4.8)$$

где $\ell_m = \ell_m(\gamma) := \int_{\tau_m}^{t_m} s^{-1} e^{-\frac{1}{2}\gamma^2 s} ds \leq \ln(t_m/\tau_m)$.

З а м е ч а н и е 4.4. *Присутствующее в правой части (4.7) слагаемое $\partial_t \mu$ понимается как обобщенная функция. При этом типичная функция μ будет непрерывно дифференцируемой по времени на любом $(t_{j-1}, t_j]$ и разрывной при $t \in \{t_j\}_{j=1}^{m-1}$. Заметим, что соглашение о непрерывности слева позволяет избежать неопределенности при интегрировании по частям. С другой стороны, все вычисления данного типа допускают следующую эквивалентную интерпретацию. При некотором малом положительном значении λ можно заметить t_j^+ на $t_j + \lambda$, а функцию μ на ее непрерывное на $[0, T]$ приближение μ_λ , линейное на каждом $[t_j, t_j + \lambda]$ и определяемое как $\mu_\lambda = \mu$ при $t \in [t_{j-1} + \lambda, t_j]$. Тогда вместо обобщенной функции $\partial_t \mu$ в вычислениях будет использоваться регулярная функция $\partial_t \mu_\lambda$, а тот же конечный результат получается при $\lambda \rightarrow 0^+$.*

Аналогично, во всех вычислениях, в которых используется Γ , можно заменить последнюю ее регулярным аналогом Γ_λ получаемым заменой δ в (4.5b) на некоторую регулярную аппроксимацию δ_λ , а затем применить $\lambda \rightarrow 0^+$. Что касается регулярности Γ , то из Условия 4.1 следует, что $\partial_s \Gamma \in L_1(\Omega \times [0, t - \tau])$ при любом $\tau \in (0, t)$, в то время как доказательство Леммы 4.2 показывает, что $\partial_s \Gamma \in L_2(\Omega \times [0, t - \tau])$.

З а м е ч а н и е 4.5. Несложно убедиться, что при $\gamma = 0$ получаем $\ell_m = \ln(t_m/\tau_m)$. А при $\gamma > 0$ получаем $\ell_m(\gamma) = E_1(\frac{1}{2}\gamma^2\tau_m) - E_1(\frac{1}{2}\gamma^2t_m)$, где $E_1(t) = \int_t^\infty s^{-1}e^{-s} ds$, и поэтому $\ell_m(\gamma) \leq |\ln(\frac{1}{2}\gamma^2\tau_m)|$ при условии $\frac{1}{2}\gamma^2\tau_m \leq 0.67$. (Для проверки последнего наблюдения достаточно убедиться, что уравнение $E_1(s) = |\ln s|$ на интервале $(0, 1)$ имеет единственный корень ≈ 0.67 .) Заметим также, что $\ell_1 = 0$ при любом $\gamma \geq 0$.

Д о к а з а т е л ь с т в о Л е м м ы 4.3. Объединяя представление (4.4) с обозначением $\Gamma(\xi, s) := \mathcal{G}(x, t_m; \xi, s)$ для функции Грина из (4.5), получаем

$$[v - w](x, t_m) = \langle [v - w](\cdot, 0), \Gamma(\cdot, 0) \rangle + \int_0^{t_m} \langle [\mathcal{M}v - \mathcal{M}w](\cdot, s), \Gamma(\cdot, s) \rangle ds.$$

Здесь, в силу (4.7), в интеграле в правой части присутствуют функции μ и ϑ , поэтому данный интеграл можно представить как сумму соответствующих интегралов $J_\mu + J_\vartheta$, которые будут рассмотрены отдельно. Воспользуемся соглашением $\int^{b^+} := \lim_{\beta \rightarrow 0^+} \int^{b+\beta}$ и разобьем J_μ как

$$J_\mu = J_\mu^{(1)} + J_\mu^{(2)} := \int_0^{t_{m-1}^+} \langle \partial_s \mu, \Gamma(\cdot, s) \rangle ds + \int_{t_{m-1}^+}^{t_m} \langle \partial_s \mu, \Gamma(\cdot, s) \rangle ds.$$

Для $J_\mu^{(1)}$ воспользуемся интегрированием по частям:

$$J_\mu^{(1)} = \langle \mu(\cdot, t_{m-1}^+), \Gamma(\cdot, t_{m-1}) \rangle - \langle \mu(\cdot, 0), \Gamma(\cdot, 0) \rangle - \int_0^{t_{m-1}} \langle \mu(\cdot, s), \partial_s \Gamma(\cdot, s) \rangle ds.$$

Таким образом, приходим к представлению

$$\begin{aligned}
[v - w](x, t_m) &= \langle [v - w - \mu](\cdot, 0), \Gamma(\cdot, 0) \rangle - \int_0^{t_{m-1}} \langle \mu(\cdot, s), \partial_s \Gamma(\cdot, s) \rangle ds \\
&\quad + \langle \mu(\cdot, t_{m-1}^+), \Gamma(\cdot, t_{m-1}) \rangle + \int_{t_{m-1}^+}^{t_m} \langle \partial_s \mu, \Gamma(\cdot, s) \rangle ds \\
&\quad + \int_0^{t_m} \langle \vartheta(\cdot, s), \Gamma(\cdot, s) \rangle ds, \tag{4.9}
\end{aligned}$$

в котором последнее слагаемое представляет J_ϑ . Поскольку Условие 4.1 может быть представлено в виде

$$\|\Gamma(\cdot, s)\|_{1, \Omega} \leq \kappa_0 e^{-\gamma^2(t_m - s)} \leq \kappa_0, \quad \int_0^{t_{m-1}} \|\partial_s \Gamma(\cdot, s)\|_{1, \Omega} ds \leq \kappa_1 \ell_m + \kappa_2,$$

то из (4.9) получаем желаемую оценку. \square

Далее рассмотрим вариант Леммы 4.3, в котором будут использоваться некоторые аппроксимации Γ_h^j для $\Gamma(\cdot, t_j)$.

Л е м м а 4.3. При выполнении условий Леммы 4.3 предположим, что вместо (4.7) справедливо представление $\mathcal{M}v - \mathcal{M}w = \partial_t \mu + \vartheta + \vartheta_*$, в котором $\vartheta_*(\cdot, t) = \sum_{j=1}^{m-1} \vartheta^j \delta(t - t_j)$ при $t \in [0, t_m]$. Предположим также, что существуют некоторые функции $\{\Gamma_h^j\}_{j=1}^{m-1}$, удовлетворяющие $\langle \vartheta^j, \Gamma_h^j \rangle = 0$ при $j = 1, \dots, m-1$, а также $\sum_{j=1}^{m-1} \tau_j \|\mathcal{H}_j^{-2} \{\Gamma(\cdot, t_j) - \Gamma_h^j\}\|_{1, \Omega} \leq \kappa_3 \ell(\tau, t)$ с некоторыми весовыми функциями $\{\mathcal{H}_j\}$ при некоторой постоянной κ_3 . Тогда оценка Леммы 4.3 остается справедливой, с той лишь разницей, что в последней строке в (4.8) будет присутствовать дополнительное слагаемое $\kappa_3 \ell(\tau, t) \max_{j=1, \dots, m-1} \{\tau_j^{-1} \|\mathcal{H}_j^2 \vartheta^j\|_{\infty, \Omega}\}$.*

Д о к а з а т е л ь с т в о. Рассуждения, аналогичные используемым при доказательстве Леммы 4.3, вновь приводят к представлению для ошибки (4.9), только в нем теперь присутствует дополнительное слагаемое $\sum_{j=1}^{m-1} \langle \vartheta^j, \Gamma(\cdot, t_j) \rangle = \sum_{j=1}^{m-1} \langle \vartheta^j, \Gamma(\cdot, t_j) - \Gamma_h^j \rangle$. \square

4.3 Описание результатов для полудискретных методов (без дискретизации по пространству)

Перейдем к описанию основных результатов параграфа для абстрактной параболической задачи (4.1) и ее полудискретизаций по времени на произвольной неравномерной сетке (4.6) на основе неявного метода Эйлера, метода Кранка-Николсон и разрывных методов Галеркина.

Пусть u является точным решением задачи (4.1), в которой параболический оператор \mathcal{M} удовлетворяет (4.2) и Условию 4.1, а численное решение $U^j \in H_0^1(\Omega) \cap C(\bar{\Omega})$, ассоциируемое с временным уровнем t_j , является решением соответствующей полудискретной задачи, при этом $U^0 = \varphi$. Тогда при $m = 1, \dots, M$ будут получены апостериорные оценки в виде

$$\begin{aligned} \|U^m - u(\cdot, t_m)\|_{\infty, \Omega} \leq & C_1(\kappa_1 \ell_m + \kappa_2) \max_{j=1, \dots, m-1} \|\chi^j\|_{\infty, \Omega} + C_2 \kappa_0 \|\chi^m\|_{\infty, \Omega} \\ & + \kappa_0 \sum_{j=1}^m \int_{t_{j-1}}^{t_j} e^{-\gamma^2(t_m-s)} \|\vartheta(\cdot, s)\|_{\infty, \Omega} ds. \end{aligned} \quad (4.10)$$

Величины, присутствующие в данных оценках, описаны в Таблице 4.1. Соответствующие же оценки будут представлены в Теоремах 4.8 и 4.10, а также в Следствии 4.15. При этом при вычислении χ^{j+1} и ϑ используются

$$\psi^{j+\alpha} := \mathcal{L}(t_{j+\alpha}) U^{j+\alpha} + f(\cdot, t_{j+\alpha}, U^{j+\alpha}), \quad \tilde{\psi} := \mathcal{L}(t) \tilde{U} + f(\cdot, t, \tilde{U}), \quad (4.11)$$

где $\alpha \in (0, 1]$ — любое значение, соответствующее временному уровню $t_{j+\alpha} := t_j + \alpha \tau_{j+1}$, для которого определение полудискретного метода предполагает вычисление соответствующего численного решения $U^{j+\alpha}$. Также нам понадобится кусочно-полиномиальный интерполянт \tilde{U} степени $p-1$ для численного решения, в то время как $I_{p-1, t} \tilde{\psi}$ будет обозначать кусочно-полиномиальный интерполянт той же степени и с теми же интерполяционными точками для $\tilde{\psi}$.

З а м е ч а н и е 4.6 (Апостериорные оценки типа интерполяционных). Величина $|\chi^j|$ в (4.10) аппроксимирует $\tau_j^p |\partial_t^p u(\cdot, t_j)|$. Это достаточ-

	p	χ^{j+1}	ϑ	C_1	C_2
неявный Эйлер	1	$U^{j+1} - U^j$	$\tilde{\psi} - \psi^j$ on $(t_{j-1}, t_j]$	1	2
Кранк-Николсон	2	$\tau_{j+1}(\psi^{j+1} - \psi^j)$	$\tilde{\psi} - I_{1,t}\tilde{\psi}$	$\frac{1}{8}$	$\frac{1}{2}$
dG(1)-Радо	3	$3\tau_{j+1}(2\psi^j - 3\psi^{j+1/3} + \psi^{j+1})$	$\tilde{\psi} - I_{2,t}\tilde{\psi}$	$\frac{2}{81}$	$\frac{1}{6}$

Таблица 4.1: Величины, присутствующие в оценке (4.10).

но очевидно для неявного метода Эйлера при $\chi^j = U^j - U^{j-1}$. Для методов же Кранка-Николсон и dG(1) заметим, что $\psi^{j+\alpha}$ аппроксимирует $\mathcal{L}u + f(\cdot, t, u)$ при $t = t_{j+\alpha}$, поэтому χ^j аппроксимирует $\tau_j^p |\partial_t^{p-1}(\mathcal{L}u + f(\cdot, t, u))|$ (более подробно, $\chi^j = \tau_j^p \partial_t^{p-1}(I_{p-1,t}\tilde{\psi})$), в то время как из (4.1) следует, что $|\partial_t^{p-1}(\mathcal{L}u + f(\cdot, t, u))| = |\partial_t^p u|$.

З а м е ч а н и е 4.7 (Оценки порядка p). Замечание 4.6, а также определения для ϑ в Таблице 4.1 позволяют заключить, что (4.10) является апостериорной оценкой порядка p , где $p = 1, 2$ и 3 , соответственно, для неявного метода Эйлера, метода Кранка-Николсон и метода dG(1) с квадратурой Радо.

4.4 Полудискретный неявный метод Эйлера

На произвольной неравномерной сетке по времени (4.6) для абстрактной параболической задачи (4.1) рассмотрим полудискретизацию по времени, использующую неявный метод Эйлера. Пусть $U^j \in H_0^1(\Omega) \cap C(\bar{\Omega})$, численное решение, ассоциируемое с временным уровнем t_j , удовлетворяет соотношениям

$$\delta_t U^j + \mathcal{L}^j U^j + f^j = 0 \quad \text{в } \Omega, \quad j = 1, \dots, M; \quad U^0 = \varphi, \quad (4.12a)$$

$$\text{где } \delta_t U^j := \frac{U^j - U^{j-1}}{\tau_j}, \quad \mathcal{L}^j := \mathcal{L}(t_j) \quad \text{и} \quad f^j := f(\cdot, t_j, U^j). \quad (4.12b)$$

Для данной дискретизации справедлива следующая апостериорная оценка.

Т е о р е м а 4.8. Пусть u является решением задачи (4.1) с параболическим оператором \mathcal{M} , удовлетворяющим (4.2) и Условию 4.1, а U^j пусть является решением соответствующей полудискретной задачи (4.12). Тогда при $m = 1, \dots, M$ справедлива оценка (4.10), в которой $\chi^j = U^j - U^{j-1}$, $C_1 = 1$, $C_2 = 2$, а ϑ определяется соотношениями

$$\vartheta(\cdot, t) = \tilde{\psi}(\cdot, t) - \tilde{\psi}(\cdot, t_j), \quad \tilde{\psi}(\cdot, t) = \mathcal{L}(t)U^j + f(\cdot, t, U^j) \quad \text{при } t \in (t_{j-1}, t_j]. \quad (4.13)$$

Д о к а з а т е л ь с т в о. Пусть $I_{1,t}U$ обозначает стандартный кусочно-линейный интерполянт для U^j по времени:

$$I_{1,t}U(\cdot, t) := \frac{t_j - t}{\tau_j} U^{j-1} + \frac{t - t_{j-1}}{\tau_j} U^j \quad \text{при } t \in [t_{j-1}, t_j], \quad j = 1, \dots, M. \quad (4.14)$$

Дополнительно введем кусочно-постоянный интерполянт \tilde{U} для U^j в виде

$$\tilde{U}(\cdot, t) := U^j \quad \text{при } t \in (t_{j-1}, t_j], \quad j = 1 \dots, M; \quad \tilde{U}(\cdot, 0) := U^1, \quad (4.15)$$

(т. е. интерполянт \tilde{U} непрерывен на $[t_0, t_1]$). Заметим, что производная по времени $\partial_t \tilde{U}$ понимается как обобщенная функция, в то время как $\partial_t(I_{1,t}U)$ является регулярной функцией, равной $\delta_t U^j$ при $t \in (t_{j-1}, t_j]$ (в соответствии с соглашением о непрерывности слева для всех разрывных функций). Поэтому из (4.12а) получаем

$$\partial_t(I_{1,t}U) + \tilde{\psi} = \vartheta \quad \text{при } (x, t) \in Q. \quad (4.16)$$

Здесь мы также воспользовались наблюдением, что, в силу (4.15), регулярная функция ϑ из (4.13) может быть представлена в виде $\vartheta = \tilde{\psi} - [\mathcal{L}^j U^j + f^j]$ при $t \in (t_{j-1}, t_j]$.

Поскольку $\mathcal{M}\tilde{U} = \partial_t \tilde{U} + \tilde{\psi}$, а $\mathcal{M}u = 0$, то из (4.16) следует, что

$$\mathcal{M}\tilde{U} - \mathcal{M}u = \partial_t[\tilde{U} - I_{1,t}U] + \vartheta \quad \text{в } Q.$$

Отсюда получаем желаемую оценку для $U^m - u(\cdot, t_m) = [\tilde{U} - u](\cdot, t_m)$, воспользовавшись Леммой 4.3 с $\mu := \tilde{U} - I_{1,t}U$ и ϑ из (4.13). Также используются следующие два соображения. Во-первых, $[\tilde{U} - u - \mu](\cdot, 0) = U^1 - \varphi - (U^1 - \varphi) = 0$. А во-вторых, при $t \in (t_{j-1}, t_j]$ справедливы соотношения

$$\mu = \frac{t_j - t}{\tau_j} (U^j - U^{j-1}) = \frac{t_j - t}{\tau_j} \chi^j \quad \Longrightarrow \quad |\mu| \leq |\chi^j|, \quad \tau_j |\partial_t \mu| = |\chi^j|.$$

Теорема доказана □

С л е д с т в и е 4.9. *При условии (4.2) апостериорная оценка Теоремы 4.8 справедлива для модельной задачи (4.3) с $\vartheta = f(\cdot, t, U^j) - f(\cdot, t_j, U^j)$ и постоянными $\kappa_0, \kappa_1, \kappa_2$ из Леммы 4.2.*

4.5 Полудискретный метод Кранка-Николсон

На произвольной неравномерной сетке по времени (4.6) для абстрактной параболической задачи (4.1) рассмотрим полудискретизацию по времени, использующую метод Кранка-Николсон. Пусть $U^j \in H_0^1(\Omega) \cap C(\bar{\Omega})$, численное решение, ассоциируемое с временным уровнем t_j , удовлетворяет соотношениям

$$\delta_t U^j + \frac{1}{2}(\mathcal{L}^{j-1} U^{j-1} + \mathcal{L}^j U^j) + \frac{1}{2}(f^{j-1} + f^j) = 0 \quad \text{в } \Omega, \quad j = 1, \dots, M, \quad (4.17a)$$

где мы вновь полагаем

$$U^0 = \varphi, \quad \delta_t U^j := \frac{U^j - U^{j-1}}{\tau_j}, \quad \mathcal{L}^j := \mathcal{L}(t_j) \quad \text{и} \quad f^j := f(\cdot, t_j, U^j). \quad (4.17b)$$

Чтобы описать апостериорную ошибку для данного метода, воспользуемся стандартным кусочно-линейным интерполяционным оператором $I_{1,t}$, который для любой непрерывной функции $w = w(t)$ определяется соотношением

$$I_{1,t}w(t) := \frac{t_j - t}{\tau_j} w(t_{j-1}) + \frac{t - t_{j-1}}{\tau_j} w(t_j) \quad \text{при } t \in [t_{j-1}, t_j], \quad j = 1, \dots, M. \quad (4.18)$$

Данное определение идентично определению (4.14) кусочно-линейного интерполянта $I_{1,t}U$ для численного решения.

Т е о р е м а 4.10. Пусть u является решением задачи (4.1) с параболическим оператором \mathcal{M} , удовлетворяющим (4.2) и Условию 4.1, а U^j пусть является решением соответствующей полудискретной задачи (4.17). Тогда при $m = 1, \dots, M$ справедлива оценка (4.10), в которой $\chi^j = \tau_j (\psi^j - \psi^{j-1})$, где $\psi^j = \mathcal{L}^j U^j + f^j$, $C_1 = \frac{1}{8}$, $C_2 = \frac{1}{2}$, а ϑ определяется соотношениями

$$\vartheta = \tilde{\psi} - I_{1,t}\tilde{\psi}, \quad \tilde{\psi} = \mathcal{L}(t)\tilde{U} + f(\cdot, t, \tilde{U}), \quad \tilde{U}(\cdot, t) = I_{1,t}U(\cdot, t) \quad (4.19)$$

при $t \in [0, T]$, и где также используются $I_{1,t}U(\cdot, t)$ из (4.14) и $I_{1,t}$ из (4.18).

Д о к а з а т е л ь с т в о. Рассмотрим $t \in [t_{j-1}, t_j]$. Заметим, что $\tilde{\psi} = I_{1,t}\tilde{\psi} + \vartheta = \frac{1}{2}(\psi^{j-1} + \psi^j) + \partial_t \mu + \vartheta$, где $\mu := \int_{t_j}^t [I_{1,t}\tilde{\psi} - \frac{1}{2}(\psi^{j-1} + \psi^j)] dt$, и поэтому

$$\mu = \tau_j^{-1} \chi^j \int_{t_j}^t (t - t_{j-1/2}) dt = -\frac{1}{2}(t_j - t)(t - t_{j-1}) \cdot \tau_j^{-2} \chi^j \quad \text{при } t \in [t_{j-1}, t_j]. \quad (4.20)$$

Далее, из $\tilde{U}(\cdot, t) = I_{1,t}U(\cdot, t)$ следует, что $\partial_t \tilde{U} = \delta_t U^j = -\frac{1}{2}(\psi^{j-1} + \psi^j)$ при $t \in (t_{j-1}, t_j]$ (где также использовалось соотношение (4.17a)). Объединяя эти два соображения, заключаем, что $\partial_t \tilde{U} + \tilde{\psi} = \partial_t \mu + \vartheta$. Поскольку же $\mathcal{M}\tilde{U} = \partial_t \tilde{U} + \tilde{\psi}$, а $\mathcal{M}\mu = 0$, то

$$\mathcal{M}\tilde{U} - \mathcal{M}\mu = \partial_t \mu + \vartheta \quad \text{в } Q. \quad (4.21)$$

Обе стороны данного соотношения являются регулярными функциями, т. е. оно справедливо при $t \in (0, T]$, поскольку функция μ из (4.20) непрерывна при $t \in [0, T]$.

Далее, воспользовавшись Леммой 4.3, из (4.21) получаем желаемую оценку для $U^m - u(\cdot, t_m) = [\tilde{U} - u](\cdot, t_m)$ с μ из (4.20) и ϑ из (4.19). Также используются следующие два соображения. Во-первых, заметим, что $[\tilde{U} - u - \mu](\cdot, 0) = U^0 - \varphi - 0 = 0$. Во-вторых, при $t \in (t_{j-1}, t_j]$ справедливы соотношения

$$|\mu| \leq \frac{1}{8} |\chi^j| \quad \text{и} \quad \tau_j |\partial_t \mu| \leq \frac{1}{2} |\chi^j|, \quad (4.22)$$

в то время как $\mu(\cdot, t_{m-1}^+) = 0$. Теорема доказана. \square

С л е д с т в и е 4.11. При выполнении условия (4.2) для модельной задачи (4.3) справедлива апостериорная оценка Теоремы 4.10 с $\vartheta = f(\cdot, t, I_{1,t}U) - I_{1,t}[f(\cdot, t, I_tU)]$ и постоянными $\kappa_0, \kappa_1, \kappa_2$ из Леммы 4.2.

З а м е ч а н и е 4.12. Апостериорная оценка Теоремы 4.10 аналогична оценкам, полученным в [54] (хотя данные оценки не являются идентичными). Представленный здесь анализ для метода Кранка-Николсон представляется несколько более простым, поскольку мы используем стандартный кусочно-линейный интерполянт численного решения, в то время как в [54] конструируется более сложный кусочно-квадратичный интерполянт численного решения по времени. В дополнение, в §5.4 будут получены апостериорные оценки для полностью дискретных методов на основе временной дискретизации типа Кранка-Николсон, которые в [54] не рассматриваются.

4.6 Полудискретный разрывный метод Галеркина $dG(\mathbf{r})$ с квадратурой Радо

На произвольной неравномерной сетке по времени (4.6) для абстрактной параболической задачи (4.1) рассмотрим полудискретизацию по времени, использующую полудискретный разрывный метод Галеркина $dG(\mathbf{r})$ (описанный, например, в [84, 145]).

Введем множество квадратурных узлов Радо $\mathcal{A}_R := \{\alpha_k : 0 < \alpha_0 < \alpha_1 < \dots < \alpha_r = 1\}$ (где, например, $r = 1$ соответствует $\mathcal{A}_R = \{\frac{1}{3}, 1\}$). Полезным также будет дополненное нулем множество $\mathcal{A} := \{0\} \cup \mathcal{A}_R$, включающее $r + 2$ точек. Далее на отрезке $[0, 1]$ введем базисные функции $\{\phi_k(s)\}_{k=0}^r$ для многочленов степени r , удовлетворяющие соотношению $\varphi_k(\alpha_l) = \delta_{kl}$, а также многочлен ζ_{r+1} степени $r + 1$, для которого справедливы соотношения

$$\zeta_{r+1}(0) = 1, \quad \zeta_{r+1}(\alpha_k) = 0 \quad \text{при } k = 0, \dots, r, \quad C_\zeta := \frac{d^{r+1}}{ds^{r+1}} \zeta_{r+1}(s). \quad (4.23)$$

Также нам потребуются два интерполяционных оператора на $(t_j, t_{j+1}]$, соответственно, степеней r и $r + 1$, которые для любой гладкой функции ϕ определяются как $\hat{I}_{r,t}\phi \in \Pi_r$, при этом $(\hat{I}_{r,t}\phi)(t_{j+\alpha}) = \phi(t_{j+\alpha})$ при $\alpha \in \mathcal{A}_R$, а также $I_{r+1,t}\phi \in \Pi_{r+1}$, при этом $(I_{r+1,t}\phi)(t_{j+\alpha}) = \phi(t_{j+\alpha})$ при $\alpha \in \mathcal{A}$.

Пусть $U^0 := \varphi$. При заданном численном решении $U^j \in H_0^1(\Omega) \cap C(\bar{\Omega})$, ассоциируемым с временным уровнем t_j , пусть $U^{j+\alpha_k} \in H_0^1(\Omega) \cap C(\bar{\Omega})$ при $k = 0, \dots, r$, соответственно ассоциируемые с временными уровнями $t_{j+\alpha_k}$, удовлетворяют соотношениям

$$[U(\cdot, t_j^+) - U^j] \varphi_k(0) + \int_{t_j^+}^{t_{j+1}} [\partial_t U + \hat{I}_{r,t}\psi] \varphi_k\left(\frac{t-t_j}{\tau_{j+1}}\right) dt = 0 \quad \text{при } k = 0, \dots, r, \quad (4.24a)$$

$$\text{где } U := \sum_{k=0}^r U^{j+\alpha_k} \varphi_k\left(\frac{t-t_j}{\tau_{j+1}}\right), \quad \psi := \mathcal{L}(t)U + f(\cdot, t, U) \quad \text{при } t \in (t_j, t_{j+1}]. \quad (4.24b)$$

Заметим, что соотношения (4.24) описывают метод $dG(r)$ с квадратурой Радо, которая является точной для многочленов степени $2r$. Если же слагаемое $\hat{I}_{r,t}\psi$ заменить на ψ , то получим метод $dG(r)$ без квадратуры.

Далее, применив интерполяционный оператор $I_{r+1,t}$ к $\{U^{j+\alpha}, \alpha \in \mathcal{A}\}$, получаем интерполянт \tilde{U} и связанную с ним функцию $\tilde{\psi}$:

$$\tilde{U} := U - [U(\cdot, t_j^+) - U^j] \zeta_{r+1}\left(\frac{t-t_j}{\tau_{j+1}}\right), \quad \tilde{\psi} = \mathcal{L}(t)\tilde{U} + f(\cdot, t, \tilde{U}). \quad (4.25)$$

Заметим, что $I_{r+1,t}\tilde{\psi}$ допускает представление

$$I_{r+1,t}\tilde{\psi} = \hat{I}_{r,t}\psi + \chi^{j+1} \tau_{j+1}^{-1} C_\zeta^{-1} \zeta_{r+1}\left(\frac{t-t_j}{\tau_{j+1}}\right), \quad (4.26)$$

где $\chi^{j+1} := \tau_{j+1}^{r+2} \partial_t^{r+1}[I_{r+1,t}\tilde{\psi}]$, поэтому, обозначая $\psi^{j+\alpha} = \psi(\cdot, t_{j+\alpha})$, получаем

$$\chi^{j+1} = \tau_{j+1} C_\zeta \left[\psi^j - \hat{I}_{r,t}\psi(\cdot, t_j^+) \right] = \tau_{j+1} C_\zeta \left[\psi^j - \sum_{k=0}^r \psi^{j+\alpha_k} \varphi_k(0) \right]. \quad (4.27)$$

Т е о р е м а 4.13. Пусть u является решением задачи (4.1) с параболическим оператором \mathcal{M} , удовлетворяющим (4.2) и Условию 4.1, а U^j пусть

является решением соответствующей полудискретной задачи (4.24). Тогда при $m = 1, \dots, M$ справедлива оценка (4.10), в которой используются $\vartheta = \tilde{\psi} - I_{r+1,t}\tilde{\psi}$ и постоянные $C_1 = C_\zeta^{-1} \max_{s \in [0,1]} |\int_0^s \zeta_{r+1}(\sigma) d\sigma|$ и $C_2 = C_\zeta^{-1} \max_{s \in [0,1]} |\zeta_{r+1}(s)|$, а также обозначения (4.23), (4.25), (4.27).

Доказательство. Заметим, что соотношение (4.24а) эквивалентно соотношению

$$\int_{t_j}^{t_{j+1}} [\partial_t \tilde{U} + \hat{I}_{r,t}\psi] \varphi_k\left(\frac{t-t_j}{\tau_{j+1}}\right) dt = 0 \quad \text{при } k = 0, \dots, r. \quad (4.28)$$

В этом несложно убедиться, представив $\partial_t[\tilde{U} - U]$ с помощью первого соотношения из (4.25), а также воспользовавшись $\int_{t_j}^{t_{j+1}} \partial_t \left\{ -\zeta_{r+1}\left(\frac{t-t_j}{\tau_{j+1}}\right) \right\} \cdot \varphi_k\left(\frac{t-t_j}{\tau_{j+1}}\right) dt = \varphi_k(0)$ (последнее же соотношение получается интегрированием по частям, поскольку $\int_{t_j}^{t_{j+1}} p(t) dt = 0$ для любого многочлена p степени $2r$, равного нулю в точках Радо).

Далее заметим, что из (4.28) следует $\partial_t \tilde{U} + \hat{I}_{r,t}\psi = 0$ (поскольку данная функция является многочленом степени r на $[t_j, t_{j+1}]$). А поскольку $\mathcal{M}\tilde{U} = \partial_t \tilde{U} + \tilde{\psi}$ и $\mathcal{M}u = 0$, в то время как $\tilde{\psi} = I_{r+1,t}\tilde{\psi} + \vartheta$, то

$$\mathcal{M}\tilde{U} - \mathcal{M}u = \partial_t \mu + \vartheta, \quad \mu := \int_{t_j}^t [I_{r+1,t}\tilde{\psi} - \hat{I}_{r,t}\psi] dt \quad \text{при } t \in [t_j, t_{j+1}].$$

Следует заметить, что в силу (4.26) функция μ непрерывна по временной переменной (поскольку многочлен ζ_{r+1} равен нулю в точках Радо). Также μ удовлетворяет оценке (4.22), в которой постоянные $\frac{1}{8}$ и $\frac{1}{2}$ соответственно заменены на C_1 и C_2 , в то время как $\mu(\cdot, t_{m-1}^+) = 0$. Отсюда, воспользовавшись Леммой 4.3, получаем желаемую оценку для $U^m - u(\cdot, t_m) = [\tilde{U} - u](\cdot, t_m)$. \square

З а м е ч а н и е 4.14. Аналогично Замечанию 4.6, величина $|\chi^j|$ в оценке (4.10) аппроксимирует $\tau_j^{r+2} |\partial_t^{r+2} u(\cdot, t_j)|$, поэтому Теорема 4.13 описывает апостериорную оценку порядка $r + 2$.

4.6.1 Частный случай dG(1)

При $r = 1$ используются точки Радо $\mathcal{A}_R = \{\frac{1}{3}, 1\}$, поэтому (4.24) эквивалентно соотношениям²

$$U^{j+1} - U^j + \frac{1}{4} \tau_{j+1} (3\psi^{j+1/3} + \psi^{j+1}) = 0, \quad (4.29a)$$

$$U^{j+1/3} - U^j + \frac{1}{12} \tau_{j+1} (5\psi^{j+1/3} - \psi^{j+1}) = 0. \quad (4.29b)$$

Воспользовавшись же (4.23) и (4.27), приходим к $\zeta_2(s) = 3(s-1)(s-\frac{1}{3})$ и $C_\zeta = 6$, а также

$$I_{2,t}\tilde{\psi} = \psi^{j+1} - \{\psi^{j+1} - \psi^{j+1/3}\} \cdot \frac{3}{2} \left(\frac{t_{j+1}-t}{\tau_{j+1}} \right) + \chi^{j+1} \cdot \frac{1}{6} \tau_{j+1}^{-1} \zeta_2 \left(\frac{t-t_j}{\tau_{j+1}} \right), \quad (4.30)$$

где для $\chi^{j+1} = \tau_{j+1}^3 \partial_t^2 (I_{2,t}\tilde{\psi})$ получаем явное представление

$$\chi^{j+1} = 3\tau_{j+1} (2\psi^j - 3\psi^{j+1/3} + \psi^{j+1}). \quad (4.31)$$

Отметим, что \tilde{U} конструируется аналогично $I_{2,t}\tilde{\psi}$, применяя интерполяционный оператор $I_{2,t}$ к приближенным решениям $\{U^j, U^{j+1/3}, U^{j+1}\}$.

С л е д с т в и е 4.15 (dG(1)). Для полудискретного метода (4.29) Теорема 4.13 остается справедливой с обозначениями (4.30), (4.31) и постоянными $C_1 = \frac{2}{81}$, $C_2 = \frac{1}{6}$.

З а м е ч а н и е 4.16. Вычисление правой стороны оценки (4.10) требует вычисления величин χ^{j+1} из (4.31) при $j < t$. Заметим, что $\psi^{j+1/3}$ и ψ^{j+1} , которые входят в (4.31), допускают явное представление из соотношений (4.29).

²Отметим, что это неявный двухстадийных метод Рунге-Кутты порядка 3. Функции $U^{j+1/3}$ и U^{j+1} , получаемые из (4.29), являются аппроксимациями третьего порядка для точного решения и соответственно на временных уровнях $t_{j+1/3}$ и t_{j+1} .

4.6.2 Применение к независимому от t оператору \mathcal{L} и модельной задаче (4.3)

Предположим, что коэффициенты линейного эллиптического оператора $\mathcal{L}(t)$ не зависят от переменной t ; для этого случая будет использоваться обозначение $\mathring{\mathcal{L}} := \mathcal{L}$.

Вспомним, что в оценителе из Теоремы 4.13 (а также из Следствия 4.15) присутствует $\vartheta = \tilde{\psi} - I_{r+1,t}\tilde{\psi}$, где $\tilde{\psi} = \mathring{\mathcal{L}}\tilde{U} + f(\cdot, t, \tilde{U})$. Поскольку же $I_{r+1,t}[\mathring{\mathcal{L}}\tilde{U}] = \mathring{\mathcal{L}}[I_{r+1,t}\tilde{U}] = \mathring{\mathcal{L}}\tilde{U}$, то для данного случая получаем более простое представление

$$\vartheta = f(\cdot, t, \tilde{U}) - I_{r+1,t}[f(\cdot, t, \tilde{U})]. \quad (4.32)$$

Таким образом, в ϑ не используется оператор $\mathring{\mathcal{L}}$, и мы приходим к следующему результату.

С л е д с т в и е 4.17. Пусть эллиптический оператор $\mathcal{L}(t) = \mathring{\mathcal{L}}$ не зависит от переменной t . Тогда Теорема 4.13 и Следствие 4.15 остаются справедливыми при более простом соотношении (4.32) для ϑ .

Наконец, вспомним, что в нашей модельной задаче (4.3) не зависит от t , поэтому к данной задаче применимо Следствие 4.17.

С л е д с т в и е 4.18. При выполнении условия (4.2) апостериорные оценки Теоремы 4.13 и Следствия 4.15 применимы к модельной задаче (4.3) с постоянными $\kappa_0, \kappa_1, \kappa_2$ из Леммы 4.2 и упрощением (4.32).

4.7 Доказательство Леммы 4.2

Начнем с того, что заметим, что функция Грина \mathcal{G} , ассоциируемая с нашей задачей (4.3) в пространственной области Ω , и функция Грина $\hat{\mathcal{G}}$, ассоциируемая с аналогичной задачей $\mathcal{M}\hat{u} := \partial_t\hat{u} - \Delta\hat{u} + f(x/\varepsilon, t, \hat{u}) = 0$ в пространственной области $\hat{\Omega} := \Omega/\varepsilon$, связаны соотношением $\|\partial_s^k \mathcal{G}(x, t; \cdot, s)\|_{1,\Omega} =$

$\|\partial_s^k \hat{\mathcal{G}}(x/\varepsilon, t; \cdot, s)\|_{1, \hat{\Omega}}$ при $k = 0, 1$. Следовательно достаточно установить выполнение Условия 4.1 для случая $\varepsilon = 1$ с некоторыми κ_0, κ_1 и κ_2 , независящими от $|\Omega|$. Поэтому положим $\mathcal{L}^* = -\Delta$ в (4.5а).

(i) Начнем с проверки первого соотношения в Условии 4.1. Функция Грина на $\bar{\mathcal{G}}$, ассоциируемая с $\bar{\mathcal{M}} := \partial_t - \Delta + \gamma^2$ в пространственной области $\bar{\Omega} := \mathbb{R}^n$, допускает явное представление через фундаментальное решение для уравнения теплопроводности (см., например, [46, §III.3], [85, §2.3.1]). Таким образом, получаем

$$\bar{\mathcal{G}}(x, t; \xi, s) = g(x - \xi, t - s), \quad \text{где} \quad g(x, t) := \frac{e^{-\gamma^2 t}}{(4\pi t)^{n/2}} \exp\left(-\frac{|x|^2}{4t}\right). \quad (4.33)$$

Далее заметим, что в силу (4.2) коэффициент a в (4.5а) удовлетворяет $a \geq \gamma^2$, поэтому, воспользовавшись принципом максимума для задачи (4.5), получаем $0 \leq \mathcal{G} \leq \bar{\mathcal{G}}$. Наконец, заметим, что

$$\bar{\mathcal{G}}(x, t; \xi, s) d\xi = e^{-\gamma^2(t-s)} \psi(\zeta) d\zeta, \quad \text{где} \quad \psi(\zeta) := \frac{e^{-|\zeta|^2}}{\pi^{n/2}}, \quad \zeta := \frac{\xi - x}{2\sqrt{t-s}}. \quad (4.34)$$

Поскольку $\int_{\mathbb{R}^n} \psi(\zeta) d\zeta = 1$, то немедленно получаем $\|\bar{\mathcal{G}}(x, t; \cdot, s)\|_{1, \Omega} \leq 1$, а затем и первое соотношение в Условии 4.1 с $\kappa_0 = 1$.

(ii) Установим второе соотношение в Условии 4.1 с $\kappa_2 = 0$ в случае линейного уравнения с $f(x, t, z) = a(x)z + b(x, t)$. В этом случае в дифференциальном операторе в (4.5) переменная s не присутствует, поэтому можно воспользоваться Следствием 5 из работы [73] (применяя последний результат в этой части доказательства, мы имитируем доказательство Леммы 2.1 в работе [75]). В силу полученной выше оценки $0 \leq \mathcal{G} \leq \bar{\mathcal{G}}$, применяя Следствие 5 из [73] с параметрами $\beta = 2$, $\gamma = 1$, $c_1 = \frac{1}{4}$, $c_2 = \frac{4}{9}c_1$ и функцией $\alpha(t) = \frac{e^{-\gamma^2 t}}{(4\pi t)^{n/2}}$, приходим к оценке $|\partial_s \mathcal{G}(x, t; \xi, s)| \leq 18c_1 c_2 (t-s)^{-1} \alpha(\frac{1}{2}[t-s]) e^{-(c_2/c_1)|\zeta|^2}$, где используется та же вспомогательная переменная ζ , что и в части (i) данного доказательства. Далее представление, аналогичное (4.34), приводит к оценке $\|\partial_s \mathcal{G}(x, t; \cdot, s)\|_{1, \Omega} \leq \kappa_1 (t-s)^{-1} e^{-\frac{1}{2}\gamma^2(t-s)}$, из которой немедленно вытекает второе соотношение в Условии 4.1 с $\kappa_2 = 0$.

(iii) Остается установить второе соотношение в Условии 4.1 в общем случае, т. е. при $f(x, t, z)$, удовлетворяющей (4.2). Отсюда очевидно, что коэффициент a в (4.5а) удовлетворяет $\gamma^2 \leq a(\xi, s) \leq \bar{\gamma}^2$. При любых фиксированных $(x, t) \in \Omega \times (0, T]$ введем в рассмотрение функцию Грина $\hat{\mathcal{G}}(x, t; \xi, s) =: \hat{\Gamma}(\xi, s)$, ассоциируемую с оператором $\partial_t - \Delta + \gamma^2$ в пространственной области Ω . Т. е. $\hat{\Gamma}(\xi, s)$ удовлетворяет варианту (4.5) с коэффициентом a , замененным на γ^2 . Сравнивая последнюю задачу с задачей (4.5) для Γ , при этом в обеих $\mathcal{L} = \mathcal{L}^* = \Delta$, заключаем, что при любых фиксированных (x, t) функция $v(\xi, s) := \hat{\Gamma}(\xi, s) - \Gamma(\xi, s)$ является решением следующей задачи с конечным условием:

$$[-\partial_s - \Delta + \gamma^2]v(\xi, s) = F(\xi, s) \quad \text{при } (\xi, s) \in \Omega \times [0, t), \quad (4.35a)$$

$$v(\xi, t) = 0 \quad \text{при } \xi \in \Omega, \quad (4.35b)$$

$$v(\xi, s) = 0 \quad \text{при } (\xi, s) \in \partial\Omega \times [0, t]. \quad (4.35c)$$

Здесь $F(\xi, s) := [a(\xi, s) - \gamma^2]\Gamma(\xi, s)$, поэтому, воспользовавшись $\Gamma \leq \bar{G}$ и (4.33), приходим к оценке

$$0 \leq F(\xi, s) \leq (\bar{\gamma}^2 - \gamma^2)g(x - \xi, t - s). \quad (4.35d)$$

Заметим, что в части (ii) уже было показано, что для $\hat{\Gamma}$ справедливо второе соотношение в Условии 4.1 с $\kappa_2 = 0$. Поэтому достаточно показать, что для v также справедливо второе соотношение в Условии 4.1 с постоянными $\kappa_1 = 0$ и $\kappa_2 = (\bar{\gamma}^2 - \gamma^2)\hat{\kappa}_2$. Последняя же оценка немедленно следует из Леммы 4.19, приведенной ниже, которая применяется к задаче с конечным условием (4.35). \square

Целью следующей леммы является ее применение к задаче с конечным условием (4.35), однако будет удобнее сформулировать эту лемму в контексте задачи с начальным условием.

Л е м м а 4.19. Пусть v удовлетворяет $[\partial_t - \Delta + \gamma^2]v = F$ в Q , равняется нулю при $t = 0$, а также при $x \in \partial\Omega$, где $0 \leq F(x, t) \leq g(x - x_0, t)$ с g из

(4.33) при некотором $x_0 \in \Omega$. Тогда $\int_0^T \|\partial_t v(\cdot, t)\|_{1, \Omega} dt \leq \hat{\kappa}_2$, где постоянная $\hat{\kappa}_2$ не зависит от $|\Omega|$, а $\hat{\kappa}_2 = \hat{\kappa}_2(\gamma)$ при $\gamma > 0$, в то время как $\hat{\kappa}_2 = \hat{\kappa}_2(T)$ при $\gamma = 0$.

Д о к а з а т е л ь с т в о. Без ограничения общности будем предполагать, что $x_0 = 0 \in \Omega$, и следовательно $F(x, t) \leq g(x, t)$. Вспомним, что $\bar{\mathcal{M}}g = 0$, где $\bar{\mathcal{M}} = \partial_t - \Delta + \gamma^2$. Отсюда $\bar{\mathcal{M}}[tg] = g$, и, воспользовавшись принципом максимума, приходим к

$$0 \leq v(x, t) \leq tg(x, t). \quad (4.36)$$

(i) Сначала установим желаемую оценку, в которой $\hat{\kappa}_2$ зависит от $|\Omega|$. Положим $w(x, t) := \varrho(t)v$, где используется вес $\varrho := t^{\frac{1}{3}} e^{\frac{1}{2}\gamma^2 t}$, при этом $\varrho' = (\frac{1}{3}t^{-1} + \frac{1}{2}\gamma^2)\varrho$. Заметим, что

$$\begin{aligned} \|\partial_t v\|_{1, \Omega \times [0, T]} &\leq \|\varrho^{-1}\|_{2, \Omega \times [0, T]} \|\varrho \partial_t v\|_{2, \Omega \times [0, T]} \\ &\leq \hat{\kappa}_3 |\Omega|^{\frac{1}{2}} \left(\|\partial_t w\|_{2, \Omega \times [0, T]} + \|\varrho' v\|_{2, \Omega \times [0, T]} \right), \end{aligned} \quad (4.37)$$

поскольку $\varrho \partial_t v = \partial_t w - \varrho' v$, в то время как

$$\|\varrho^{-1}\|_{2, \Omega \times [0, T]}^2 = |\Omega| \int_0^T t^{-\frac{2}{3}} e^{-\gamma^2 t} dt =: |\Omega| \hat{\kappa}_3^2.$$

(Таким образом, $\hat{\kappa}_3^2 \leq 3T^{1/3}$ при $\gamma \geq 0$, а из $\int_0^\infty t^{-\frac{2}{3}} e^{-t} dt \approx 2.7$ получаем $\hat{\kappa}_3^2 \lesssim 2.7\gamma^{-2/3}$ при $\gamma > 0$). Чтобы оценить $\partial_t w$ в (4.37), заметим, что $\bar{\mathcal{M}}w = \varrho F + \varrho' v \leq \varrho g + \varrho' v$, и воспользуемся априорной оценкой [31, (6.6) из Главы III]:

$$\|\partial_t w\|_{2, \Omega \times [0, T]} \leq \|\bar{\mathcal{M}}w\|_{2, \Omega \times [0, T]}. \quad (4.38)$$

(Строго говоря, процитированная оценка дана для несколько отличного дифференциального оператора, но несложно убедиться, что доказательство применимо и к $\bar{\mathcal{M}}$). Объединяя (4.38) с (4.37), а также с $\varrho' v \leq (\frac{1}{3} + \frac{1}{2}\gamma^2 t)\varrho g$ (а последнее соотношение следует из (4.36)), приходим к оценке

$$\|\partial_t v\|_{1, \Omega \times [0, T]} \leq 2 \hat{\kappa}_3 |\Omega|^{\frac{1}{2}} \|\hat{\varrho} g\|_{2, \Omega \times [0, T]}, \quad \text{где } \hat{\varrho} := \left(\frac{4}{3} + \frac{1}{2}\gamma^2 t\right)\varrho. \quad (4.39)$$

Наконец, воспользовавшись $\zeta := \frac{x}{\sqrt{2t}}$ и $\psi(\zeta)$ из (4.34), можно убедиться, что

$$\|\hat{\rho} g\|_{2,\Omega \times [0,T]}^2 \leq \int_0^T \frac{\hat{\rho}^2(t) e^{-2\gamma^2 t}}{(8\pi t)^{n/2}} \int_{\mathbb{R}^n} \psi(\zeta) d\zeta dt = \int_0^T \frac{(\frac{4}{3} + \frac{1}{2}\gamma^2 t)^2 t^{2/3} e^{-\gamma^2 t}}{(8\pi t)^{n/2}} dt =: \hat{\kappa}_4^2.$$

(Присутствующий здесь несобственный интеграл сходится, поскольку $\frac{n}{2} - \frac{2}{3} < 1$ при $n \leq 3$.) Объединяя данный результат с (4.39), приходим к желаемой оценке с постоянной $\hat{\kappa}_2 := 2 \hat{\kappa}_3 \hat{\kappa}_4 |\Omega|^{\frac{1}{2}}$.

(ii) Покажем, что желаемая оценка справедлива и с постоянной $\hat{\kappa}_2$, независимой от $|\Omega|$ (для чего потребуются несколько более тонкие оценки). Разобьем \mathbb{R}^n на непересекающиеся подобласти $\Omega_0 := \{|x| < 2\}$ и $\Omega_j := \{2^j < |x| < 2^{j+1}\}$ при $j = 1, \dots$; положим также $\Omega'_j := \{2^{j-1} < |x| < 2^{j+2}\} \supset \Omega_j$ при $j = 1, \dots$. Очевидно, что

$$|\Omega_j|^{\frac{1}{2}} \leq c_n 2^{\frac{1}{2}nj}. \quad (4.40)$$

Далее мы частично имитируем доказательство из части (i). Во-первых, остается справедливой оценка (4.37) с Ω , замененной на Ω_j при $j = 0, 1, \dots$. Поэтому при $j = 0$, воспользовавшись промежуточными оценками из части (i), немедленно получаем

$$\|\partial_t v\|_{1,(\Omega \cap \Omega_0) \times [0,T]} \leq 2 \hat{\kappa}_3 \hat{\kappa}_4 |\Omega_0|^{\frac{1}{2}}. \quad (4.41)$$

При этом, в отличие от $\hat{\kappa}_2$ из части (i), присутствующий здесь множитель $\hat{\kappa}_3 \hat{\kappa}_4$ не зависит от $|\Omega|$.

При $j \geq 1$ объединим вышеописанный локальный вариант оценки (4.37) со следующим локальным вариантом оценки (4.38):

$$\|\partial_t w\|_{2,(\Omega \cap \Omega_j) \times [0,T]} \leq \bar{C} \left\{ \|\bar{\mathcal{M}}w\|_{2,(\Omega \cap \Omega'_j) \times [0,T]} + \|w\|_{2,(\Omega \cap \Omega'_j) \times [0,T]} \right\}.$$

В последней оценке постоянная \bar{C} не зависит от Ω или T (см. [31, (6.6), (6.11) из Главы III]). Далее $\bar{\mathcal{M}}w$ оценивается как в части (i), в то время как $w = \rho v \leq t\rho g$ в силу (4.36). Объединяя эти наблюдения, приходим к следующему локальному варианту оценки (4.39):

$$\|\partial_t v\|_{1,(\Omega \cap \Omega_j) \times [0,T]} \leq 2 \hat{\kappa}_3 |\Omega_j|^{\frac{1}{2}} \bar{C} \|(\hat{\rho} + t\rho) g\|_{2,\Omega'_j \times [0,T]} \quad \text{при } j \geq 1. \quad (4.42)$$

Остается вновь воспользоваться $\zeta := \frac{x}{2\sqrt{t}}$ и $\psi(\zeta)$ из (4.34), а также наблюдением, что при $j \geq 1$ справедлива оценка $(\exp(-\frac{|x|^2}{4t}))^2 \leq e^{-\frac{4j-2}{t}} e^{-|\zeta|^2} \leq c'_n (\frac{t}{4^j})^n e^{-|\zeta|^2}$. Таким образом, при $j \geq 1$ справедливо соотношение

$$\|(\hat{\rho} + t\rho)g\|_{2,\Omega'_j \times [0,T]}^2 \leq c'_n 4^{-jn} \int_0^T \frac{(\hat{\rho} + t\rho)^2 e^{-2\gamma^2 t} t^n}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} \psi(\zeta) d\zeta dt = c''_n 4^{-jn}.$$

Объединяя его с (4.42), а затем с (4.41) и (4.40), приходим к оценке

$$\|\partial_t v\|_{1,(\Omega \cap \Omega_j) \times [0,T]} \leq 2 \hat{\kappa}_3 c_n \begin{cases} \hat{\kappa}_4 & \text{при } j = 0, \\ \sqrt{c''_n} 2^{-\frac{1}{2}nj} & \text{при } j \geq 1. \end{cases}$$

Отсюда получаем желаемую оценку, в которой постоянная $\hat{\kappa}_2 := 2 \hat{\kappa}_3 c_n [\hat{\kappa}_4 + \sqrt{c''_n} (2^{\frac{1}{2}n} - 1)^{-1}]$ не зависит от $|\Omega|$. \square

§5 Апостериорные оценки для полностью дискретных методов

Настоящий параграф посвящен обобщению результатов, полученных в §4, на случай полностью дискретных методов. Для полулинейных параболических уравнений второго порядка будут получены апостериорные оценки в норме максимума модуля для ошибки конечноэлементных решений. При этом будут рассмотрены полностью дискретные методы на основе неявного метода Эйлера, метода Кранка-Николсон и разрывного метода Галеркина с квадратурой Радо.

При обобщении результатов предыдущего параграфа на случай полностью дискретных методов мы будем опираться на так называемые эллиптические реконструкции (elliptic reconstructions). Данный технический аппарат был предложен в работах [114, 108, 75] и является аналогом техники на основе проекции Ритца (Ritz projection) [153], часто используемой при получении априорных оценок ошибки для конечноэлементных решений параболических задач.

Основной материал параграфа опубликован в работе [175], а некоторые его результаты обсуждаются и обобщаются в [174, 176, 177].

Параграф организован следующим образом. В §5.1 будут описаны предположения о дискретизациях по пространству. Примеры таких дискретизаций, а также доступные для них эллиптические оценщики приведены в конце параграфа, в §5.6. Апостериорные оценки для полностью дискретных методов на основе неявного метода Эйлера, метода Кранка-Николсон и разрывного метода Галеркина будут сначала описаны в §5.2, а затем установлены, соответственно, в §5.3, §5.4 и §5.5.

5.1 Предположения о дискретизациях по пространству

Использование аппарата эллиптических реконструкций позволяет оценить ошибку, порожденную дискретизацией по пространству, через апостериорные оценщики, полученные для соответствующих эллиптических задач. Этот подход значительно упрощает анализ. В частности, нам не понадобятся оценки пространственных производных функций Грина для параболических операторов. (При этом важно понимать, что при получении эллиптических оценщиков в норме максимума модуля, как правило, используются оценки пространственных производных функции Грина для соответствующих эллиптических операторов; см., например, §§1–3.)

Рассмотрим стационарный вариант абстрактной параболической задачи (4.1):

$$\mathcal{L}v + g(\cdot, v) = 0 \quad \text{в } \Omega, \quad v = 0 \quad \text{на } \partial\Omega, \quad (5.1)$$

а также его дискретизацию в виде

$$\text{Найти } v_h \in \mathring{V}_h : \quad \mathcal{L}_h v_h + \mathcal{P}_h[g(\cdot, v_h)] = 0, \quad \text{где } \mathring{V}_h := V_h \cap H_0^1(\Omega). \quad (5.2a)$$

Здесь $V_h \subset C(\bar{\Omega})$ является некоторым конечноэлементным пространством, и для некоторого интерполяционного оператора $I_h : C(\bar{\Omega}) \rightarrow V_h$ операторы \mathcal{L}_h и \mathcal{P}_h таковы, что

$$\begin{aligned} \mathcal{L}_h : H_0^1(\Omega) &\rightarrow \mathring{V}_h - I_h[g(\cdot, 0)], \\ \mathcal{P}_h v \in \mathring{V}_h + I_h v \quad \forall v \in C(\bar{\Omega}), \quad \mathcal{P}_h v_h &= v_h \quad \forall v_h \in V_h. \end{aligned} \quad (5.2b)$$

Заметим, что поскольку любая функция $v_h \in \mathring{V}_h$ обращается в нуль на $\partial\Omega$, то $\mathring{V}_h - I_h[g(\cdot, 0)] = \mathring{V}_h - I_h[g(\cdot, v_h)]$, поэтому формулировка (5.2) корректна.

Предположения.

Будем предполагать, что при всем рассматриваемых здесь функциях g

- (i) существуют единственные решения v и v_h задач (5.1) и (5.2);
- (ii) для данных задач справедлива некоторая апостериорная оценка вида

$$\|v - v_h\|_{\infty, \Omega} \leq \eta(V_h, v_h, g(\cdot, v_h)). \quad (5.3)$$

5.1.1 Эллиптическая модельная задача

Стандартные конечноэлементные формулировки для эллиптических уравнений (включая формулировки с квадратурными формулами) довольно часто представимы в виде (5.2). Например, рассмотрим стационарный вариант параболической модельной задачи (4.3) в полигональной области $\Omega \subset \mathbb{R}^n$:

$$-\varepsilon^2 \Delta v + g(x, v) = 0 \quad \text{при } x \in \Omega, \quad v = 0 \quad \text{при } x \in \partial\Omega, \quad \partial_z g(x, z) \geq \gamma^2 \geq 0. \quad (5.4)$$

Пусть $V_h \subset C(\bar{\Omega})$ является некоторым конечноэлементным пространством, а $\dot{V}_h := V_h \cap H_0^1(\Omega)$. Тогда стандартный метод Галеркина для данной задачи принимает вид

$$\text{Найти } v_h \in \dot{V}_h : \quad \varepsilon^2 \langle \nabla v_h, \nabla w_h \rangle + \langle g(\cdot, v_h), w_h \rangle_h = 0 \quad \forall w_h \in \dot{V}_h, \quad (5.5)$$

где $\langle \cdot, \cdot \rangle_h$ — либо точное скалярное произведение $\langle \cdot, \cdot \rangle$ в $L_2(\Omega)$, либо же некоторая квадратурная аппроксимация для $\langle \cdot, \cdot \rangle$.

З а м е ч а н и е 5.1. Метод Галеркина (5.5) является дискретизацией типа (5.2) при условии, что матрица Грама с элементами $\langle \phi_i, \phi_j \rangle_h$, соответствующими базису $\{\phi_i\}$ в \dot{V}_h , обратима. Тогда \mathcal{L}_h и \mathcal{P}_h могут быть соответственно определены соотношениями $\langle \mathcal{L}_h \varphi, w_h \rangle_h = \varepsilon^2 \langle \nabla \varphi, \nabla w_h \rangle$ и $\langle \mathcal{P}_h q, w_h \rangle_h = \langle q, w_h \rangle_h$ для любых $\varphi \in H_0^1(\Omega)$, $q \in C(\bar{\Omega})$, $w_h \in \dot{V}_h$. При этом, разумеется, предполагается выполнение ограничений (5.2b).

Предположим, например, что при любых $q_h, w_h \in V_h$ справедливо соотношение $\langle q_h, w_h \rangle_h = \langle q_h, w_h \rangle$. Тогда вышеописанные определения для \mathcal{L}_h и \mathcal{P}_h могут быть несколько упрощены к виду $\langle \mathcal{L}_h \varphi, w_h \rangle = \varepsilon^2 \langle \nabla \varphi, \nabla w_h \rangle$ и $\langle \mathcal{P}_h q, w_h \rangle = \langle q, w_h \rangle_h$ для любых $\varphi \in H_0^1(\Omega)$, $q \in C(\bar{\Omega})$ и $w_h \in \dot{V}_h$ при выполнении ограничений (5.2b). В частности,

(i) если $\langle \cdot, \cdot \rangle_h := \langle \cdot, \cdot \rangle$ (т. е. в случае метода Галеркина без квадратур), то \mathcal{P}_h становится просто L_2 проекцией;

(ii) если же используется квадратурная формула вида $\langle q, w_h \rangle_h := \langle I_h q, w_h \rangle$ с некоторым интерполяционным оператором I_h на V_h , то $\mathcal{P}_h := I_h$.

З а м е ч а н и е 5.2. Допустим, что используется квадратурная формула с сосредоточенными массами вида $\langle q, \phi_i \rangle_h = \langle I_h(q\phi_i), 1 \rangle = q_i \langle \phi_i, 1 \rangle$ для любой базисной функции ϕ_i из V_h , где $q \in C(\bar{\Omega})$, при этом $\sum q_i \phi_i = I_h q$. Тогда вновь $\mathcal{P}_h := I_h$, при этом $\mathcal{L}_h v_h := \sum a_i \phi_i$, где для внутренних узлов сетки $a_i := \varepsilon^2 \frac{\langle \nabla v_h, \nabla \phi_i \rangle}{\langle \phi_i, 1 \rangle}$, а для граничных узлов $a_i := -[g(\cdot, 0)]_i$. Таким образом, $\mathcal{L}_h v_h$ легко вычисляется при любом $v_h \in \dot{V}_h$ умножением соответствующим образом нормализованной матрицы жесткости на вектор узловых значений $\{v_{h,i}\}$.

Примеры дискретизаций типа (5.5) для некоторых частных случаев (5.4), а также доступные для них эллиптические оцениватели вида (5.3) приведены в конце параграфа, в §5.6.

5.2 Описание результатов для полностью дискретных методов

Далее мы опишем наши апостериорные оценки для полностью дискретных задач на основе неявного метода Галеркина, метода Кранка-Николсон и разрывного метода Галеркина для абстрактной параболической задачи (4.1), удовлетворяющей (4.2) и Условию 4.1. При этом будут рассматриваться пространственные дискретизации типа (5.2).

Будем ассоциировать с временным уровнем t_{j+1} некоторое конечноэлементное пространство $V_h^{j+1} \subset C(\bar{\Omega})$ и соответствующее численное решение $u_h^{j+1} \in \dot{V}_h^{j+1} := V_h^{j+1} \cap H_0^1(\Omega)$. Аналогично, вспомогательное численное решение $\hat{u}_h^j \in H_0^1(\Omega)$ будет ассоциироваться с временным уровнем t_j^+ (заметим, что, как правило, либо $\hat{u}_h^j \in V_h^j$, либо же $\hat{u}_h^j \in V_h^{j+1}$). Полная дискретизация использует операторы $\mathcal{L}_h(t)$ и \mathcal{P}_h^{j+1} , для которых, в соответствии с (5.2b), предполагается, что с некоторым интерполяционным оператором I_h^{j+1}

на V_h^{j+1} справедливы соотношения

$$\begin{aligned} \mathcal{L}_h(t) : H_0^1(\Omega) &\rightarrow \mathring{V}_h^{j+1} - I_h^{j+1}[f(\cdot, t, 0)] \text{ при } t \in (t_j, t_{j+1}], \\ \mathcal{P}_h^{j+1}v &\in \mathring{V}_h^{j+1} + I_h^{j+1}v \quad \forall v \in C(\bar{\Omega}), \quad \mathcal{P}_h^{j+1}v_h = v_h \quad \forall v_h \in V_h^{j+1}. \end{aligned} \quad (5.6)$$

Отметим две возможные стратегии выбора \hat{u}_h^j :

$$\text{Случай А: } \hat{u}_h^j := I_*^{j+1}u_h^j, \quad I_*^{j+1} : \mathring{V}_h^j \rightarrow \mathring{V}_h^{j+1} \quad \Rightarrow \quad \mathcal{P}_h^{j+1}\hat{u}_h^j = \hat{u}_h^j, \quad (5.7a)$$

$$\text{Случай В: } \hat{u}_h^j := u_h^j \quad \Rightarrow \quad \hat{u}_h^j \in \mathring{V}_h^j, \quad u_h^{j+1} \in \mathring{V}_h^{j+1}. \quad (5.7b)$$

Здесь в случае А \hat{u}_h^j получается применением некоторого линейного интерполяционного оператора I_*^{j+1} к u_h^j , при этом предполагается, что $I_*^{j+1}w_h = w_h$ при любом $w_h \in \mathring{V}_h^{j+1}$. В качестве I_*^{j+1} может использоваться либо стандартная интерполяция Лагранжа, либо же L_2 проекция. Заметим, что при $V_h^j \subseteq V_h^{j+1}$ случаи А и В становятся идентичными.

Ниже будет показано, что для рассматриваемых полностью дискретных методов при $m = 1, \dots, M$ справедлива апостериорная оценка вида

$$\begin{aligned} \|u_h^m - u(\cdot, t_m)\|_{\infty, \Omega} &\leq \kappa_0 e^{-\gamma^2 t_m} \|u_h^0 - \varphi\|_{\infty, \Omega} \\ &+ (\kappa_1 \ell_m + \kappa_2) \max_{j=1, \dots, m-1} \left\{ C_1 \|\chi_h^j\|_{\infty, \Omega} + C_1^* \eta^j \right\} \\ &+ C_2 \kappa_0 \|\chi_h^m\|_{\infty, \Omega} + (C_2^* \kappa_0 + 1) \eta^m \\ &+ \kappa_0 \sum_{j=1}^m \int_{t_{j-1}}^{t_j} e^{-\gamma^2(t_m-s)} \|\vartheta_h(\cdot, s)\|_{\infty, \Omega} ds \\ &+ \kappa_0 \sum_{j=1}^{m-1} e^{-\gamma^2(t_m-t_j)} \|\hat{u}_h^j - u_h^j\|_{\infty, \Omega}. \end{aligned} \quad (5.8)$$

При этом величины, присутствующие в данной оценке, описаны в последующих Теоремах 5.9, 5.17 and 5.22, а также собраны в следующей таблице:

	p	χ_h^{j+1}	C_1	C_2	C_1^*	C_2^*	\mathcal{A}
неявный Эйлер	1	$u_h^{j+1} - \hat{u}_h^j$	1	2	1	1	$\{1\}$
Кранк-Николсон	2	$\tau_{j+1}(\psi_h^{j+1} - \hat{\psi}_h^j)$	$\frac{1}{8}$	$\frac{1}{2}$	2	3	$\{0^+, 1\}$
dG(1)-Радо	3	$3\tau_{j+1}(2\hat{\psi}_h^j - 3\psi_h^{j+1/3} + \psi_h^{j+1})$	$\frac{2}{81}$	$\frac{1}{6}$	$\frac{5}{3}$	10	$\{0^+, \frac{1}{3}, 1\}$

Отсюда видно, что при вычислении χ_h^{j+1} (аналогично (4.11)) используются функции $\hat{\psi}_h^j$ и $\psi_h^{j+\alpha}$, удовлетворяющие

$$\mathcal{P}_h^{j+1} \hat{\psi}_h^j = \mathcal{L}_h(t_j^+) \hat{u}_h^j + \mathcal{P}_h^{j+1}[f(\cdot, t_j, \hat{u}_h^j)], \quad (5.9a)$$

$$\mathcal{P}_h^{j+1} \psi_h^{j+\alpha} = \mathcal{L}_h(t_{j+\alpha}) u_h^{j+\alpha} + \mathcal{P}_h^{j+1}[f(\cdot, t_{j+\alpha}, u_h^{j+\alpha})], \quad (5.9b)$$

где $\alpha \in (0, 1]$ принимает значения из $\mathcal{A} \setminus \{0^+\}$ и соответствует временному уровню $t_{j+\alpha} := t_j + \alpha \tau_{j+1}$, для которого определение метода предполагает вычисление соответствующего численного решения $u_h^{j+\alpha}$. При $\alpha = 0^+$ используются \hat{u}_h^j и $\hat{\psi}_h^j$. Заметим, что в случае А из (5.7a) соотношения (5.9) несколько упрощаются в силу $\hat{\psi}_h^j = \mathcal{P}_h^{j+1} \hat{\psi}_h^j$ и $\psi_h^{j+\alpha} = \mathcal{P}_h^{j+1} \psi_h^{j+\alpha}$.

Также в (5.8) присутствуют величины η^j , соответствующие ошибке, порождаемой пространственной дискретизацией. Они определяются с помощью эллиптического оценителя η из (5.3):

$$\eta^{j+1} := \eta(V_h^{j+1}, u_h^{j+1}, g^{j+1}(\cdot, u_h^{j+1})) \quad \text{при } \mathcal{A} = \{1\}, \quad (5.10a)$$

т. е. в случае неявного метода Эйлера, а в противном случае

$$\eta^{j+1} := \max \left\{ \eta(V_h^{j+1}, \hat{u}_h^j, \hat{g}^j(\cdot, \hat{u}_h^j)), \max_{\alpha \in \mathcal{A} \setminus \{0^+\}} \eta(V_h^{j+1}, u_h^{j+\alpha}, g^{j+\alpha}(\cdot, u_h^{j+\alpha})) \right\}, \quad (5.10b)$$

где

$$\hat{g}^j(\cdot, v) := f(\cdot, t_j, v) - \hat{\psi}_h^j, \quad g^{j+\alpha}(\cdot, v) := f(\cdot, t_{j+\alpha}, v) - \psi_h^{j+\alpha}. \quad (5.10c)$$

Величина ϑ_h аналогична ϑ в (4.10), но в ее определении будут присутствовать так называемые эллиптические реконструкции численного решения. Определение и оценки для данной величины будут обсуждаться в §5.2.2 and §5.2.3. Постоянные C_1 и C_2 в (5.8) те же самые, что и в апостериорной оценке (4.10) для соответствующего полудискретного метода.

З а м е ч а н и е 5.3 (Апостериорные оценки типа интерполяционных). Аналогично Замечаниям 4.6 и 4.7 для полудискретных методов, величина

$|\chi_h^j|$ в (5.8) аппроксимирует $\tau_j^p |\partial_t^p u(\cdot, t_j)|$, поэтому апостериорная оценка (5.8) имеет порядок p по времени с $p = 1, 2$ и 3 для, соответственно, неявного метода Эйлера, метода Кранка-Николсон и метода $dG(1)$.

З а м е ч а н и е 5.4. Последнее слагаемое в оценке (5.8) обращается в нуль, если $\hat{u}_h^j = u_h^j$ при любом $j = 1, \dots, M$, т. е. в случае B из (5.7), а также в случае A при условии, что сетка по пространству может лишь сгущаться при возрастании j . Отметим также, что в некоторых случаях последнее слагаемое в оценке (5.8) может быть улучшено к виду (5.31); см. Замечание 5.10, которое применимо как к неявному методу Эйлера, так и к другим рассматриваемым здесь временным дискретизациям.

5.2.1 Вычисление величин χ_h^j и η^j в апостериорной оценке (5.8)

Начнем с наиболее простой ситуации и рассмотрим вычисление χ_h^j и η^j для неявного метода Эйлера. В этом случае (как будет показано в Замечании 5.8, посвященном ψ_h^{j+1})

$$\chi_h^{j+1} = u_h^{j+1} - \hat{u}_h^j, \quad \psi_h^{j+1} = -\frac{u_h^{j+1} - \hat{u}_h^j}{\tau_{j+1}}, \quad g^{j+1}(\cdot, u_h^{j+1}) = f(\cdot, t_{j+1}, u_h^{j+1}) - \psi_h^{j+1}, \quad (5.11)$$

где соотношение для g^{j+1} соответствует (5.10с). Поскольку u_h^{j+1} и \hat{u}_h^j доступны во время вычислений, то величины χ_h^{j+1} и η^{j+1} из (5.10а) легко вычисляемы в явном виде.

Для методов же Кранка-Николсон и $dG(1)$ вычисление величин χ_h^j и η^j из (5.10) является несколько менее очевидным, и сводится к вычислению $\hat{\psi}_h^j$. Действительно, для метода Кранка-Николсон можно явным образом получить ψ_h^{j+1} (воспользовавшись (5.35), в предположении, что величина $\hat{\psi}_h^j$ доступна). Для метода же $dG(1)$ величины $\psi_h^{j+1/3}$ и ψ_h^{j+1} могут вычислены явным образом (в силу (5.44а)). Таким образом, для вычисления χ_h^{j+1} и η^{j+1} достаточно вычислить $\hat{\psi}_h^j$.

Обсудим возможные подходы к вычислению $\hat{\psi}_h^j$ для модельной задачи (4.3) в случае А из (5.7). В этом случае $\hat{u}_h^j \in \mathring{V}_h^{j+1}$, поэтому (5.9а) упрощается к виду $\hat{\psi}_h^j = \mathcal{L}_h(t_j^+) \hat{u}_h^j + \mathcal{P}_h^{j+1}[f(\cdot, t_j, \hat{u}_h^j)]$. При описании этих подходов, мы будем иметь в виду Замечания 5.1 а 5.2; см. также Замечание 5.26. (А случай В будет рассмотрен в Замечании 5.5.)

(i) Предположим, что $V_h^j = V_h^{j+1}$. Тогда, объединяя (5.9) с $\psi_h^{j+1} = \mathcal{P}_h^{j+1}\psi_h^{j+1}$, заключаем, что $\hat{\psi}_h^j = \psi_h^j$, где ψ_h^j уже было вычислено.

(ii) Допустим, что \mathcal{P}_h^{j+1} соответствует квадратуре с *сосредоточенными массами* $\langle q, \phi_i \rangle_h$. Тогда, как описано в Замечании 5.2, $\mathcal{P}_h^{j+1} = I_h^{j+1}$ совпадает с интерполяционным оператором на V_h^{j+1} , а $\mathcal{L}_h(t_j^+) \hat{u}_h^j$ легко вычисляется при любом $\hat{u}_h^j \in \mathring{V}_h^{j+1}$ умножением нормализованной матрицы жесткости на вектор узловых значений $\{\hat{u}_{h,i}^j\}$. Таким образом, при получении величины $\hat{\psi}_h^j$, равной в данном случае правой части из (5.9а), используются лишь явные вычисления.

(iii) Вычисление $\hat{\psi}_h^j$ в общем случае также сводится к вычислению правой части в (5.9а), при этом требуется применение операторов $\mathcal{L}_h(t_j^+)$ и \mathcal{P}_h^{j+1} . Как обсуждалось в Замечании 5.1, грубо говоря, $\mathcal{L}_h(t_j^+)v_h$ при некотором $v_h \in \mathring{V}_h^{j+1}$ может быть получено применением $M_{j+1}^{-1}K_{j+1}$ к вектору узловых значений $\{v_{h,i}\}$, где M_{j+1} является матрицей массы а K_{j+1} — матрицей жесткости, ассоциируемыми с временным уровнем t_{j+1} . Такие вычисления несколько более дорогостоящи.

В то же время, в некоторых случаях можно попытаться избежать обращения матрицы массы. Допустим, в оценителе присутствует $\hat{\psi}_h^j - w_h$ с некоторой функцией w_h , при этом требуется обращение матрицы массы $M := M_{j+1}$ при вычислении $\hat{\psi}_h^j$. Тогда можно воспользоваться оценкой $\|\hat{\psi}_h^j - w_h\|_{\infty, \Omega} \leq \|M^{-1}\|_{\infty} \cdot \|M(\hat{\psi}_h^j - w_h)\|_{\infty, \Omega}$, где $\|M^{-1}\|_{\infty}$ может быть оценена априори. Поскольку $M\hat{\psi}_h^j$ вычисляется явно (применением нормализованной матрицы жесткости к вектору узловых значений, соответствующему \hat{u}_h^j), то все вычисления становятся явными.

З а м е ч а н и е 5.5 (Случай В). В случае (5.7b) при $V_h^j \not\subseteq V_h^{j+1}$ функция $\hat{\psi}_h^j$ для метода Кранка-Николсон уже не равна правой части из (5.9a), поэтому ψ_h^{j+1} и $\hat{\psi}_h^j$ вычисляются с помощью (5.34), при этом вышеизложенные подходы (ii) или (iii) применяются при вычислении ψ_h^{j+1} . Для метода $dG(1)$ в этом случае просто используется $\hat{\psi}_h^j = \psi_h^j$ (как будет описано в Замечании 5.21).

5.2.2 Эллиптические реконструкции. Определение \mathcal{V}_h

При выводе апостериорных оценок для полностью дискретных методов мы будем опираться на так называемые эллиптические реконструкции (elliptic reconstructions). Данный технический аппарат был предложен в работах [114, 108, 75] и является аналогом техники на основе проекции Ритца (Ritz projection) [153], часто используемой при получении априорных оценок ошибки для конечноэлементных решений параболических задач.

Будет ассоциировать эллиптическую реконструкцию \hat{R}^j с временным уровнем t_j^+ , а $R^{j+\alpha}$ при $\alpha \in \mathcal{A} \setminus \{0^+\}$ — с временным уровнем $t_{j+\alpha}$. Определим последние с помощью \hat{g}^j и $g^{j+\alpha}$ из (5.10c), как единственные решения, принадлежащие $H_0^1(\Omega) \cap C(\bar{\Omega})$, следующих эллиптических задач:

$$\mathcal{L}(t_j) \hat{R}^j + \hat{g}^j(x, \hat{R}^j) = 0, \quad \mathcal{L}(t_{j+\alpha}) R^{j+\alpha} + g^{j+\alpha}(x, R^{j+\alpha}) = 0. \quad (5.12)$$

Заметим, что (5.12) описывает два варианта эллиптической задачи (5.1), с $\mathcal{L} := \mathcal{L}(t_j)$, $g := \hat{g}^j$, а также с $\mathcal{L} := \mathcal{L}(t_{j+\alpha})$, $g := g^{j+\alpha}$, с соответствующими точными решениями \hat{R}^j и $R^{j+\alpha}$. Важно отметить, что при применении дискретизации по пространству (5.2), в которой используется V_h^{j+1} , к вышеописанным эллиптическим задачам, получаем

$$\mathcal{L}_h(t_j^+) \hat{R}_h^j + \mathcal{P}_h^{j+1} [\hat{g}^j(x, \hat{R}_h^j)] = 0, \quad \mathcal{L}_h(t_{j+\alpha}) R_h^{j+\alpha} + \mathcal{P}_h^{j+1} [g^{j+\alpha}(x, R_h^{j+\alpha})] = 0. \quad (5.13)$$

Поскольку, при наших предположениях решения данных дискретных задач единственны, заключаем, что $\hat{R}_h^j = \hat{u}_h^j$ и $R_h^{j+\alpha} = u_h^{j+\alpha}$. В этом легко убедиться,

объединив (5.13) с определениями для \hat{g}^j и $g^{j+\alpha}$ в (5.10с), а затем воспользовавшись (5.9). Следовательно, в силу эллиптической апостериорной оценки (5.3), для точных решений \hat{R}^j и $R^{j+\alpha}$ и соответствующих численных решений \hat{u}_h^j и $u_h^{j+\alpha}$ справедливы следующие оценки, в которых используется η^{j+1} из (5.10):

$$\|\hat{R}^j - \hat{u}_h^j\|_{\infty, \Omega} \leq \eta^{j+1} \text{ если } 0^+ \in \mathcal{A}, \quad \|R^{j+\alpha} - u_h^{j+\alpha}\|_{\infty, \Omega} \leq \eta^{j+1} \text{ при } \alpha \in \mathcal{A} \setminus \{0^+\}. \quad (5.14)$$

Далее, аналогично определениям для \tilde{U} , $\tilde{\psi}$ и ϑ из §4.3, введем кусочно-полиномиальную функцию \tilde{R} , а затем $\tilde{\psi}_R$ и ϑ_h :

$$\tilde{R} := I_{p-1,t}^* R, \quad \tilde{\psi}_R := \mathcal{L}(t) \tilde{R} + f(\cdot, t, \tilde{R}), \quad \vartheta_h := \tilde{\psi}_R - I_{p-1,t}^* \tilde{\psi}_R. \quad (5.15)$$

Здесь $I_{p-1,t}^*$ является кусочно-полиномиальным интерполяционным оператором степени $p-1$, построенным с помощью узлов $\{t_{j+\alpha}, \alpha \in \mathcal{A}\}$ на каждом из отрезков $(t_j, t_{j+1}]$ (при этом $I_{p-1,t}^*$ отличается от $I_{p-1,t}$ тем что используется узел t_j^+ вместо t_j , при этом $I_{0,t}^* = I_{0,t}$).

Заметим, что, объединяя (5.12), (5.10с) и определение для $\tilde{\psi}_R$ в (5.15), приходим к соотношениям

$$\tilde{\psi}_R(\cdot, t_j^+) = \hat{\psi}_h^j \text{ если } 0^+ \in \mathcal{A}, \quad \tilde{\psi}_R(\cdot, t_{j+\alpha}) = \psi_h^{j+\alpha} \text{ при } \alpha \in \mathcal{A} \setminus \{0^+\}. \quad (5.16)$$

5.2.3 Вычисление ϑ_h

Обсудим возможные подходы к вычислению ϑ_h в случае независимого от t оператора \mathcal{L} , который включает и модельную задачу (4.3). Тогда ϑ_h из (5.15) упрощается к виду

$$\vartheta_h = \vartheta_{f, \tilde{R}} := f(\cdot, t, \tilde{R}) - I_{p-1,t}^*[f(\cdot, t, \tilde{R})]. \quad (5.17)$$

З а м е ч а н и е 5.6 (Неявный метод Эйлера). Для метода Эйлера $\tilde{R} = R^j$, поэтому (5.17) упрощается к виду $\vartheta_h = \vartheta_{f, \tilde{R}} = f(\cdot, t, R^j) - f(\cdot, t_j, R^j)$ при $t \in$

$(t_{j-1}, t_j]$. Поскольку в данном представлении $\vartheta_{f, \tilde{R}}$ присутствует эллиптическая реконструкция R^j , явное вычисление которой численным методом не предусмотрено, то можно воспользоваться аппроксимацией $\vartheta_{f, \tilde{u}_h}$ (где $\tilde{u}_h = u_h^j$), которая хорошо оценивается вычислением ее значений для некоторого набора значений t на каждом из интервалов $(t_{j-1}, t_j]$. При этом разница между $\vartheta_{f, \tilde{R}}$ и $\vartheta_{f, \tilde{u}_h}$ может, например, быть оценена следующим образом. При $t \in (t_{j-1}, t_j]$

$$\begin{aligned} \|\vartheta_{f, \tilde{R}} - \vartheta_{f, \tilde{u}_h}(\cdot, t)\|_{\infty, \Omega} &\leq \eta^j \sup_{(t_{j-1}, t_j] \times \mathbb{R}} \|\partial_z f(\cdot, t, z) - \partial_z f(\cdot, t_j, z)\|_{\infty, \Omega} \\ &\leq \tau_j \eta^j \sup_{(t_{j-1}, t_j] \times \mathbb{R}} \|\partial_t \partial_z f(\cdot, t, z)\|_{\infty, \Omega}, \end{aligned}$$

где мы воспользовались (5.14), а η^j вычисляется из (5.10а). Если же выполняется $|\partial_t \partial_z f| \leq C$, то $\|\vartheta_{f, \tilde{R}} - \vartheta_{f, \tilde{u}_h}(\cdot, t)\|_{\infty, \Omega}$ становится $\mathcal{O}(\tau_j \eta^j)$, т. е. незначительной по сравнению с величиной η^j , которая явно присутствует в (5.8).

З а м е ч а н и е 5.7 (Методы Кранка-Николсон и dG(1)). В более общем случае при оценке $\vartheta_{f, \tilde{R}}$ в (5.17) можно использовать $\vartheta_{f, \tilde{u}_h}$, где $\tilde{u}_h := I_{p-1, t}^* u_h$, поскольку можно ожидать, что $\vartheta_{f, \tilde{R}} \approx \vartheta_{f, \tilde{u}_h}$. Например, при $|\partial_z f| \leq C_f$ с некоторой постоянной C_f , воспользовавшись

$$|\vartheta_{f, \tilde{R}} - \vartheta_{f, \tilde{u}_h}| \leq |f(\cdot, t, \tilde{u}_h) - f(\cdot, t, \tilde{R})| + |I_{p, t}^*[f(\cdot, t, \tilde{u}_h) - f(\cdot, t, \tilde{R})|,$$

легко получить довольно грубую оценку $\|\vartheta_{f, \tilde{R}} - \vartheta_{f, \tilde{u}_h}(\cdot, t)\|_{\infty, \Omega} \leq C_* C_f \eta^{j+1}$ при $t \in (t_j, t_{j+1}]$, где $C_* = 2$ для метода Кранка-Николсон и $C_* = \frac{10}{3}$ для метода dG(1). В некоторых же специальных случаях (например, если f линейна по третьему аргументу) при $t \in (t_j, t_{j+1}]$ справедлива более точная оценка $\|\vartheta_{f, \tilde{R}} - \vartheta_{f, I_{\tilde{u}_h}}(\cdot, t)\|_{\infty, \Omega} \leq C \tau_{j+1} \eta^{j+1}$ с некоторой постоянной C . В этом случае разница между $\vartheta_{f, \tilde{R}}$ и $\vartheta_{f, \tilde{u}_h}$ вновь является незначительной по сравнению с величиной η^{j+1} , которая явно присутствует в (5.8).

5.3 Полностью дискретный неявный метод Эйлера

Рассмотрим полностью дискретный неявный метод Эйлера для задачи (4.3), получаемый применением пространственной дискретизации (5.5) к полудискретному неявному методу Эйлера (4.12):

$$\text{Найти } u_h^j \in \mathring{V}_h^j : \quad \varepsilon^2 \langle \nabla u_h^j, \nabla w_h \rangle + \langle f(\cdot, t_j, u_h^j) + \delta_t^* u_h^j, w_h \rangle_h = 0 \quad \forall w_h \in \mathring{V}_h^j, \quad (5.18)$$

при $j = 1, \dots, M$, где $\langle \cdot, \cdot \rangle_h$ — либо точное скалярное произведение $\langle \cdot, \cdot \rangle$ в $L_2(\Omega)$, либо же некоторая квадратурная аппроксимация для $\langle \cdot, \cdot \rangle$.

Заметим, что $\delta_t^* u_h^j$ аппроксимирует $\partial_t u$ и определяется аналогично δ_t из (4.12b):

$$\delta_t^* u_h^j := \frac{u_h^j - \hat{u}_h^{j-1}}{\tau_j}, \quad \text{где } \hat{u}_h^0 := u_h^0. \quad (5.19a)$$

Точнее говоря, δ_t^* идентично δ_t при $j = 1$, а при $j > 1$ несколько отличается, поскольку используется промежуточное численное решение $\hat{u}_h^{j-1} \in H_0^1(\Omega)$, ассоциируемое с временным уровнем t_{j-1}^+ , возможные определения которого описаны в (5.7).

Таким образом полностью дискретный метод (5.18) допускает представление

$$\mathcal{L}_h^j u_h^j + \mathcal{P}_h^j [f(\cdot, t_j, u_h^j) + \delta_t^* u_h^j] = 0, \quad (5.19b)$$

где $\mathcal{L}_h^j := \mathcal{L}_h(t_j)$, а \mathcal{P}_h^j удовлетворяет (5.6). Для некоторых частных случаев $\langle \cdot, \cdot \rangle_h$ операторы \mathcal{L}_h^j и \mathcal{P}_h^j описаны в Замечаниях 5.1 и 5.2, только сейчас необходимо использовать V_h^j вместо V_h . Также (5.19b) может интерпретироваться как полностью дискретный метод для абстрактной параболической задачи (4.1), полученный применением пространственной дискретизации (5.2) к полудискретной задаче (4.12).

Заметим, что, в силу (5.6), $\mathcal{L}_h^j u_h^j \in \mathring{V}_h^j - I_h^j[f(\cdot, t_j, 0)]$, в то время как и u_h^j , и $\delta_t^* u_h^j$ обращаются в нуль на $\partial\Omega$, и поэтому $\mathring{V}_h^j - I_h^j[f(\cdot, t_j, 0)]$ совпадает с $\mathring{V}_h^j - I_h^j[f(\cdot, t_j, u_h^j) + \delta_t^* u_h^j]$, т. е. формулировка (5.19) корректна.

5.3.1 Апостериорный анализ на основе кусочно-постоянной эллиптической реконструкции

Для оценки ошибки численного решения полностью дискретного неявного метода Эйлера (5.19) положим $\mathcal{A} := \{1\}$ (т. е. всегда справедливо $j + \alpha = j + 1$) и воспользуемся эллиптическими реконструкциями R^j , определенными в (5.12) при $j = 1, \dots, M$. Заметим, что в последнем определении используется g^j , которое, в свою очередь, определяется через $\psi_h^j = -\delta_t^* u_h^j$, оба описаны в (5.11).

З а м е ч а н и е 5.8. В силу (5.19b), из $\psi_h^j = -\delta_t^* u_h^j$ следует, что $\mathcal{P}_h^j \psi_h^j = \mathcal{L}_h^j u_h^j + \mathcal{P}_h^j [f(\cdot, t_j, u_h^j)]$, т. е. ψ_h^j удовлетворяет (5.9b). (В случае \mathcal{A} из (5.7) ситуация упрощается, поскольку $\mathcal{P}_h^j \psi_h^j = \psi_h^j$.) Следовательно R^j удовлетворяет (5.14) с $\mathcal{A} := \{1\}$.

Мы готовы получить апостериорную оценку для полностью дискретного метода (5.19).

Т е о р е м а 5.9. Пусть u является точным решением задачи (4.1), (4.2) с параболическим оператором \mathcal{M} , удовлетворяющим Условию 4.1, а u_h^j является решением (5.19). Тогда при $m = 1, \dots, M$ справедлива оценка (5.8), в которой η^j и χ_h^j определены в (5.10a), (5.11), $C_1 = 1$, $C_2 = 2$, $C_1^* = C_2^* = 1$, а регулярная функция ϑ_h при $t \in (t_{j-1}, t_j]$, $j = 1, \dots, M$ определяется соотношением

$$\vartheta_h(\cdot, t) = \tilde{\psi}_R(\cdot, t) - \tilde{\psi}_R(\cdot, t_j), \quad \tilde{\psi}_R(\cdot, t) = \mathcal{L}(t)R^j + f(\cdot, t, R^j) \quad \text{при } t \in (t_{j-1}, t_j]. \quad (5.20)$$

Здесь также используется R^j , эллиптическая реконструкция из (5.12), (5.11), при этом $\mathcal{A} := \{1\}$.

Т е о р е м а 5.9*. Теорема 5.9 остается справедливой, если заменить величины $\|\chi_h^j\|_{\infty, \Omega}$ и $\|\chi_h^m\|_{\infty, \Omega}$ в оценке (5.8) на, соответственно, $\|u_h^j - u_h^{j-1}\|_{\infty, \Omega}$ и $\|u_h^m - u_h^{m-1}\|_{\infty, \Omega}$, а также заменить $e^{-\gamma^2(t_m - t_j)}$ на $e^{-\gamma^2(t_m - t_{j+1})}$.

Удобно сначала установить Теорему 5.9*, а затем обобщить доказательство последней для Теоремы 5.9.

Доказательство Теоремы 5.9*. В силу Замечания 5.8, $\|R^j - u_h^j\|_{\infty, \Omega} \leq \eta^j$, поэтому для доказательства желаемой оценки типа (5.8) для $u_h^m - u(\cdot, t_m)$ достаточно установить соответствующую оценку типа (5.8) для $R^m - u(\cdot, t_m)$, с той лишь разницей, что $(C_2^* \kappa_0 + 1)$ должна быть заменена на $C_2^* \kappa_0 = \kappa_0$. После этого останется лишь применить неравенство треугольника.

Перейдем к оценке $R^m - u(\cdot, t_m)$. Для этого, в некотором смысле, будет обобщено доказательство Теоремы 4.8. Введем $I_{1,t}u_h$, стандартный кусочно-линейный интерполянт для u_h^j по времени:

$$I_{1,t}u_h(\cdot, t) := \frac{t_j - t}{\tau_j} u_h^{j-1} + \frac{t - t_{j-1}}{\tau_j} u_h^j \quad \text{при } t \in [t_{j-1}, t_j], \quad j = 1, \dots, M. \quad (5.21)$$

Далее введем кусочно-постоянный интерполянт \tilde{R} по времени для R^j :

$$\tilde{R}(\cdot, t) := R^j \quad \text{при } t \in (t_{j-1}, t_j], \quad j = 1 \dots, M; \quad \tilde{R}(\cdot, 0) := R^1, \quad (5.22)$$

(т. е. \tilde{R} непрерывен на $[t_0, t_1]$, и определяется аналогично \tilde{U} в (4.15)). Производная по времени $\partial_t \tilde{R}$ понимается как обобщенная функция, в то время как $\partial_t(I_{1,t}u_h)$ является регулярной функцией.

Заметим, что данное определение \tilde{R} , а также определения функций в (5.20) идентичны с соответствующими определениями в (5.15) (при $p = 1$), поэтому справедливо и наблюдение (5.16), которое можно переписать в виде $\tilde{\psi}_R(\cdot, t_j) = \psi_h^j = -\delta_t^* u_h^j$. Объединяя это с (5.20), получаем $\tilde{\psi}_R = \vartheta_h - \delta_t^* u_h^j$, и следовательно

$$\partial_t(I_{1,t}u_h) + \tilde{\psi}_R = \vartheta_h + \vartheta_* \quad \text{в } Q, \quad (5.23)$$

где ϑ_* является следующей регулярной функцией:

$$\vartheta_*(\cdot, t) := \partial_t(I_{1,t}u_h) - \delta_t^* u_h^j \quad \text{при } t \in (t_{j-1}, t_j]. \quad (5.24)$$

Поскольку $\mathcal{M}\tilde{R} = \partial_t \tilde{R} + \tilde{\psi}_R$, и при этом $\mathcal{M}u = 0$, то в силу (5.23) получаем

$$\mathcal{M}\tilde{R} - \mathcal{M}u = \partial_t[\tilde{R} - I_{1,t}u_h] + [\vartheta_h + \vartheta_*] \quad \text{в } Q.$$

Отсюда получаем желаемую оценку для $R^m - u(\cdot, t_m) = [\tilde{R} - u](\cdot, t_m)$ типа (5.8), но с $(C_2^* \kappa_0 + 1)$ замененной на $C_2^* \kappa_0 = \kappa_0$. Для этого достаточно воспользоваться Леммой 4.3 с $\mu := \tilde{R} - I_{1,t} u_h$ и $\vartheta := \vartheta_h + \vartheta_*$, а также следующими тремя наблюдениями. Во-первых,

$$[\tilde{R} - u - \mu](\cdot, 0) = R^1 - \varphi - (R^1 - u_h^0) = u_h^0 - \varphi. \quad (5.25)$$

Во-вторых, при $t \in (t_{j-1}, t_j]$ справедливо $\mu = R^j - u_h^j + \frac{t_j - t}{\tau_j} (u_h^j - u_h^{j-1})$. Поэтому

$$|\mu| \leq |R^j - u_h^j| + |u_h^j - u_h^{j-1}| \quad \text{и} \quad \tau_j |\partial_t \mu| = |u_h^j - u_h^{j-1}|, \quad (5.26)$$

где $\|R^j - u_h^j\|_{\infty, \Omega} \leq \eta^j$. Наконец, объединяя (5.24) с (5.19а), (5.21), заключаем, что $\vartheta_*(\cdot, t) = \frac{1}{\tau_j} (\hat{u}_h^{j-1} - u_h^{j-1})$ при $t \in (t_{j-1}, t_j]$. Следовательно

$$\int_{t_j}^{t_{j+1}} e^{-\gamma^2(t_m - s)} \|\vartheta_*(\cdot, s)\|_{\infty, \Omega} ds \leq e^{-\gamma^2(t_m - t_{j+1})} \|\hat{u}_h^j - u_h^j\|_{\infty, \Omega}, \quad (5.27)$$

где $\hat{u}_h^0 - u_h^0 = 0$. Данные три наблюдения (5.25), (5.26) и (5.27) приводят к желаемой оценке для $\|R^m - u(\cdot, t_m)\|_{\infty, \Omega}$. \square

Доказательство Теоремы 5.9. Будем следовать доказательству Теоремы 5.9*, но теперь интерполянт $I_{1,t} u_h$ из (5.21) будет заменен во всех рассуждениях на кусочно-непрерывный интерполянт

$$I_{1,t}^* u_h(\cdot, t) := \frac{t_j - t}{\tau_j} \hat{u}_h^{j-1} + \frac{t - t_{j-1}}{\tau_j} u_h^j \quad \text{при } t \in (t_{j-1}, t_j], \quad j = 1, \dots, M, \quad (5.28)$$

где $I_{1,t}^* u_h(\cdot, 0) := \hat{u}_h^0 = u_h^0$. Также, в качестве определения для ϑ_* будет использоваться не (5.24), а следующее:

$$\vartheta_*(\cdot, t) := \partial_t (I_{1,t}^* u_h) - \delta_t^* u_h^j = [\hat{u}_h^{j-1} - u_h^{j-1}] \delta(t - t_{j-1}^+) \quad \text{при } t \in (t_{j-1}, t_j], \quad (5.29)$$

где $\delta(\cdot)$ — одномерная дельта-функция Дирака. (Заметим, что в силу $\hat{u}_h^0 = u_h^0$ и соглашения о непрерывности справа в $t = 0$ заключаем, что $\vartheta_* = 0$ на $[0, t_1]$.) Таким образом, вместо (5.27) будет использоваться оценка

$$\int_0^{t_m} e^{-\gamma^2(t_m - s)} \|\vartheta_*(\cdot, s)\|_{\infty, \Omega} ds \leq \sum_{j=1}^{m-1} e^{-\gamma^2(t_m - t_j)} \|\hat{u}_h^j - u_h^j\|_{\infty, \Omega}. \quad (5.30)$$

Необходимая оценка для $R^m - u(\cdot, t_m) = [\tilde{R} - u](\cdot, t_m)$ вновь устанавливается с помощью Леммы 4.3, только теперь $\mu := \tilde{R} - I_{1,t}^* u_h$, поэтому справедлив вариант соотношений (5.26), в котором u_h^{j-1} заменено на \hat{u}_h^{j-1} .

З а м е ч а н и е 5.10 (Эффект изменения сетки по пространству). В некоторых случаях слагаемое в последней строке (5.8), отвечающее за ошибку при изменении (в частности, при огрублении) пространственной сетки, может быть несколько улучшено и принимает вид

$$\kappa_3 \ell(\tau, t) \max_{j=1, \dots, m-1} \{ \tau_j^{-1} \|\mathcal{H}_j^2 \vartheta^j\|_{\infty, \Omega} \}, \quad \text{где } \vartheta^j = u_h^{j-1} - \hat{u}_h^{j-1}. \quad (5.31)$$

При этом вес \mathcal{H}_j соответствует локальному диаметру сетки по пространству, ассоциируемой с конечноэлементным пространством \mathring{V}_h^j . Описанный вариант оценки (5.8) легко установить, воспользовавшись Леммой 4.3* с ϑ_* из (5.29), при условии, что выполняется вариант Леммы 4.2 для пространственных производных функции Грина. Действительно, пусть \hat{u}_h^{j-1} будет L_2 проекцией для u_h^{j-1} на \mathring{V}_h^j . Тогда $\langle u_h^{j-1} - \hat{u}_h^{j-1}, \Gamma_h^j \rangle = 0$ при любом $\Gamma_h^j \in \mathring{V}_h^j$. Выберем $\Gamma_h^j = I_h^j \Gamma(\cdot, t_j)$, тогда достаточно убедиться, что $\|\mathcal{H}_j^{-2} \{ \Gamma(\cdot, t_j) - \Gamma_h^j \}\|_{1, \Omega} \leq \kappa_3 \frac{C}{t_m - t_j}$. Для этого, в свою очередь, достаточно показать, что $\|\Gamma(\cdot, t_j)\|_{W_1^2(\Omega)} \leq \kappa_3 \frac{C}{t_m - t_j}$. Последняя оценка является здесь ключевой, и требует оценок для пространственных производных Γ . Такая оценка может быть установлена, например, в силу [77, (2.2)] для случая гладкой области и $\mathcal{L} = -\Delta + 1$, $f = f(x, t)$ с некоторой постоянной $\kappa_3 = \mathcal{O}(1)$. Также данная оценка может быть получена при $\mathcal{L} = -\varepsilon^2 \partial_{x_1}^2 u + a(x_1)$ с $\kappa_3 = \mathcal{O}(\varepsilon^{-2})$ в силу [174, (2.18b)].

Важно заметить, что $\kappa_3 = \mathcal{O}(\varepsilon^{-2})$ (поскольку, в силу (4.5a), $\Delta \Gamma = \varepsilon^{-2} [-\partial_s + a] \Gamma$; см. также [174]). Поэтому в сингулярно возмущенном режиме $\varepsilon \ll 1$, компонента оценителя вида (5.31) может оказаться существенно больше, чем последнее слагаемое в (5.8). Некоторые численные результаты для этого случая представлены в [176].

Заметим также, что если сетка по пространству изменяется в сторо-

ну огрубления достаточно часто, то выбор стратегий интерполяции численных решений со старой сетки на новую является достаточно деликатным даже в регулярном режиме; см., например, контрпример в [80, §4].

5.3.2 Модельная задача (4.3)

Пусть u является точным решением задачи (4.3) в ограниченной полигональной области $\Omega \subset \mathbb{R}^n$, $n = 1, 2, 3$, а u_h^j — решением дискретной задачи (5.18), при этом V_h^j и $\langle \cdot, \cdot \rangle_h$ на каждом временном уровне t_j таковы, что справедлива апостериорная ошибка типа (5.3) для соответствующей эллиптической задачи.

С л е д с т в и е 5.11. Апостериорные оценки Теорем 5.9 and 5.9* справедливы для задачи (4.3), при этом ϑ_h упрощается к виду (5.17) и может быть оценено, как описано в Замечании 5.6, а

$$\eta^j := \eta(V_h^j, u_h^j, f(\cdot, t_j, u_h^j) + \delta_t^* u_h^j) \quad \text{при } j = 1, \dots, M,$$

где используется эллиптический оценщик η из (5.3).

З а м е ч а н и е 5.12 (Регулярный режим $\varepsilon = 1$). Пусть $\varepsilon = 1$, $\gamma \geq 0$, а V_h^j и $\langle \cdot, \cdot \rangle_h$ определены как описано в §5.6.1, т. е. \mathcal{T}_h^j является локально квазиравномерной триангуляцией области $\bar{\Omega}$, состоящей из элементов T , V_h^j — соответствующее конечноэлементное пространство непрерывных кусочно-полиномиальных функций степени $l \geq 1$, а $\mathring{V}_h^j := V_h^j \cap H_0^1(\Omega)$. При этом используется квадратурная формула вида $\langle \varphi, w \rangle_h := \sum_{T \in \mathcal{T}_h^j} Q_T(\varphi w)$, описанная в §5.6.1. Тогда Следствие 5.11 применимо с $\eta = \eta_0$ из (5.51).

Заметим, что неявный метод Эйлера для линейного уравнения типа (4.3) при $\varepsilon = 1$ рассматривался в работах [83, 62, 75, 77] в случае B из (5.7). (Заметим, что случай B эквивалентен случаю A , если \hat{u}_h^{j-1} выбирается равной L_2 проекции для u_h^{j-1} на \mathring{V}_h^j .) В этом случае вышеописанная апостериорная оценка аналогична полученным в [83, (1.13)] и [62] (хотя они

и неидентичны), поскольку во всех этих оценках присутствуют величины $|u_h^j - u_h^{j-1}|$, которые можно интерпретировать как аппроксимирующие $\tau_j |\partial_t u|$. Однако оценка [83, (1.13)] представлена без доказательства. Доказательства же в [62] основаны на оценках временных и пространственных производных обобщенной функции Грина в норме $L_1(\Omega)$, и представляются весьма сложными по сравнению с представленным здесь анализом. (Работа [62] будет также обсуждаться в Замечаниях 5.14 и 5.25.) В апостериорных же оценках [75, 77] присутствуют величины (обозначенные как $\tau_j |g^j - g^{j-1}|$ в [75]), которые можно интерпретировать как аппроксимирующие $\tau_j |\partial_t^2 u + \dots|$, что представляется несколько неестественным для метода первого порядка по времени. Слагаемые же, отвечающие за ошибку при изменении (в частности, при огрублении) пространственной сетки, в [83, 77] аналогичны (5.31).

З а м е ч а н и е 5.13 (Сингулярно возмущенный режим $\varepsilon \ll 1$). Пусть $\varepsilon \in (0, 1]$, $\gamma > 0$, $\Omega := (0, 1)$, а V_h^j и $\langle \cdot, \cdot \rangle_h$ определены как описано в §5.6.2. Тогда Следствие 5.11 применимо с $\eta = \eta_\varepsilon$ из (5.54). (В случае же полигональной области, в качестве η можно использовать эллиптические оценщики из §§1–3.)

Заметим, что получаемые таким образом параболические оценщики являются робастными. Действительно, малый параметр ε присутствует в оценке (5.8) лишь в оценщиках ошибки дискретизации по пространству η^j . Робастность же последних следует из их робастности для стационарного случая; см. Замечание 5.27. Последнее, действительно, применимо к η^j при $g_* = f(\cdot, t_j, u_h^j) + \delta_t^* u_h^j$ и v , замененном на $u(\cdot, t_j)$. В этом случае в η^j присутствуют величины $\varepsilon^{-2} h_i^2 |I_h g_*|$ и $\varepsilon^{-1} h_i^2 |\partial_x (I_h g_*)|$, соответственно аппроксимирующие $h_i^2 |\partial_x^2 u(\cdot, t_j)|$ и $\varepsilon h_i^2 |\partial_x^3 u(\cdot, t_j)| \simeq h_i^2 |\partial_x^2 u(\cdot, t_j)|$. Численные результаты, приведенные в [174, §4], также подтверждают эффективность и робастность полученного здесь оценщика.

З а м е ч а н и е 5.14. *Неявный метод Эйлера для более общего сингулярно возмущенного уравнения с малым параметром $\varepsilon \ll 1$ рассматривался в работе [62]. Заметим, что полученные в [62] апостериорные оценки не являются робастными, поскольку в них присутствуют слагаемые вида $\varepsilon^{-1} \max_j \|u_h^j - u_h^{j-1}\|_{\infty, \Omega}$ (вместо $\max_j \|u_h^j - \hat{u}_h^{j-1}\|_{\infty, \Omega}$, присутствующих в полученных здесь оценителях).*

Аналогично, в апостериорных оценках [60] для более сложного сингулярно возмущенного уравнения Аллена-Кана (полученных в более слабой норме $L_\infty(L_2)$) присутствуют отрицательные степени ε . Заметим также, что, строго говоря, наши результаты неприменимы к уравнению Аллена-Кана, в силу условия (4.2). С другой стороны, некоторые наши результаты могут быть обобщены на случай $|\partial_z f(x, t, z)| \leq \bar{\gamma}^2$ (см. Замечание 2.2 в [174]), при этом в апостериорных оценках будут присутствовать дополнительные множители типа $e^{\bar{\gamma}^2 t_m}$.

5.4 Полностью дискретный метод Кранка-Николсон

Рассмотрим полностью дискретный метод Кранка-Николсон для задачи (4.3), получаемый применением пространственной дискретизации (5.5) к полудискретному методу Кранка-Николсон (4.17): Найти численное решение $u_h^j \in \mathring{V}_h^j$, удовлетворяющее соотношению

$$\varepsilon^2 \left\langle \frac{1}{2} \nabla (\hat{u}_h^{j-1} + u_h^j), \nabla w_h \right\rangle + \left\langle \frac{1}{2} [f(\cdot, t_{j-1}, \hat{u}_h^{j-1}) + f(\cdot, t_j, u_h^j)] + \delta_t^* u_h^j, w_h \right\rangle_h = 0, \quad (5.32)$$

$\forall w_h \in \mathring{V}_h^j$ при $j = 1, \dots, M$, где $\langle \cdot, \cdot \rangle_h$ — либо точное скалярное произведение $\langle \cdot, \cdot \rangle$ в $L_2(\Omega)$, либо же некоторая квадратурная аппроксимация для $\langle \cdot, \cdot \rangle$. Здесь численное решение $u_h^j \in \mathring{V}_h^j := V_h^j \cap H_0^1(\Omega)$ и вспомогательное численное решение $\hat{u}_h^{j-1} \in H_0^1(\Omega)$ соответственно ассоциируются с временными слоями t_j и t_{j-1}^+ .

Также здесь используется $\delta_t^* u_h^j$, которое аппроксимирует $\partial_t u$ и определяется идентично (5.19а) (и аналогично δ_t из (4.17b)):

$$\delta_t^* u_h^j := \frac{u_h^j - \hat{u}_h^{j-1}}{\tau_j}, \quad \text{где } \hat{u}_h^0 := u_h^0. \quad (5.33a)$$

Заметим, что полностью дискретный метод (5.32) может быть представлен в виде

$$\mathcal{P}_h^j[\delta_t^* u_h^j] + \frac{1}{2}(\hat{\mathcal{L}}_h^{j-1} \hat{u}_h^{j-1} + \mathcal{L}_h^j u_h^j) + \frac{1}{2} \mathcal{P}_h^j[f(\cdot, t_{j-1}, \hat{u}_h^{j-1}) + f(\cdot, t_j, u_h^j)] = 0, \quad (5.33b)$$

где $\hat{\mathcal{L}}_h^{j-1} := \mathcal{L}_h(t_{j-1}^+)$, $\mathcal{L}_h^j := \mathcal{L}_h(t_j)$, а \mathcal{P}_h^j удовлетворяет (5.6). Для некоторых частных случаев $\langle \cdot, \cdot \rangle_h$ операторы \mathcal{L}_h^j и \mathcal{P}_h^j описаны в Замечаниях 5.1 и 5.2. С другой стороны, (5.33) можно интерпретировать как полную дискретизацию для абстрактной задачи (4.1), полученную применением пространственной дискретизации типа (5.2) к полудискретной задаче (4.17).

Заметим также, что, в силу (5.6), $\hat{\mathcal{L}}_h^{j-1} \hat{u}_h^{j-1} \in \mathring{V}_h^j - I_h[f(\cdot, t_{j-1}, 0)]$ и $\mathcal{L}_h^j u_h^j \in \mathring{V}_h^j - I_h[f(\cdot, t_j, 0)]$, в то время как при $w \in H_0^1(\Omega)$ справедливо $\mathring{V}_h^j - I_h[f(\cdot, t_k, 0)] = \mathring{V}_h^j - I_h[f(\cdot, t_k, w)]$. Поскольку $\delta_t^* u_h^j \in H_0^1(\Omega)$, заключаем, что формулировка (5.33b) корректна.

5.4.1 Апостериорный анализ на основе кусочно-постоянных эллиптических реконструкций

Для получения оценки ошибки для полностью дискретного метода Кранка-Николсон (5.33) положим $\mathcal{A} := \{0^+, 1\}$ и воспользуемся эллиптическими реконструкциями \hat{R}^{j-1} и R^j (где $j = 1, \dots, M$), описанными в (5.12). При этом используются величины \hat{g}^{j-1} и g^j , определяемые соотношениями (5.10с), в которых, в свою очередь, будут использоваться следующие $\hat{\psi}_h^{j-1}$ и ψ_h^j :

$$\hat{\psi}_h^{j-1} := -\psi_h^j - 2\delta_t^* u_h^j, \quad \psi_h^j := \mathcal{L}_h(t_j) u_h^j + \mathcal{P}_h^j[f(\cdot, t_j, u_h^j)]. \quad (5.34)$$

Отсюда (из первого соотношения) следует, что

$$\delta_t^* u_h^j + \frac{1}{2}(\hat{\psi}_h^{j-1} + \psi_h^j) = 0. \quad (5.35)$$

З а м е ч а н и е 5.15. Определение ψ_h^j в (5.34) приводит к $\psi_h^j \in \mathring{V}_h^j$, следовательно $\mathcal{P}_h^j \psi_h^j = \psi_h^j$, и поэтому ψ_h^j удовлетворяет (5.9b). Далее, $\hat{\psi}_h^{j-1}$ из (5.34) удовлетворяет $\mathcal{P}_h^j \hat{\psi}_h^{j-1} = -\psi_h^j - 2\mathcal{P}_h^j[\delta_t^* u_h^j]$ при любом $\hat{u}_h^{j-1} \in H_0^1(\Omega)$, что, в силу (5.33b), приводит к $\mathcal{P}_h^j \hat{\psi}_h^{j-1} = \mathcal{L}_h(t_{j-1}^+) \hat{u}_h^{j-1} + \mathcal{P}_h^j[f(\cdot, t_{j-1}, \hat{u}_h^{j-1})]$, т. е. $\hat{\psi}_h^{j-1}$ удовлетворяет (5.9a). Отсюда заключаем, что \hat{R}^{j-1} и R^j удовлетворяют (5.14).

З а м е ч а н и е 5.16. Теорема 5.17 остается справедливой для любой пары $\hat{\psi}_h^{j-1}$, ψ_h^j , удовлетворяющей (5.9) и (5.35). Например, в качестве определения ψ_h^j вместо (5.34), можно использовать $\psi_h^j := \mathcal{L}_h(t_j) u_h^j + f(\cdot, t_j, u_h^j)$, однако данная модификация несколько усложняет вычисление ψ_h^j .

Чтобы сформулировать апостериорную оценку для ошибки $u_h - u$, полезным оказывается обобщить кусочно-линейный интерполирующий оператор $I_{1,t}$ из (4.18) для непрерывных слева функций $w = w(t)$ следующим образом:

$$I_{1,t}^* w(t) = \frac{t_j - t}{\tau_j} w(t_{j-1}^+) + \frac{t - t_{j-1}}{\tau_j} w(t_j) \quad \text{при } t \in (t_{j-1}, t_j], \quad j = 1, \dots, M. \quad (5.36)$$

Аналогично, применив кусочно-линейный интерполирующий оператор $I_{1,t}^*$ к эллиптическим реконструкциям \hat{R}^{j-1} и R^j , определим

$$\tilde{R}(\cdot, t) = \frac{t_j - t}{\tau_j} \hat{R}^{j-1} + \frac{t - t_{j-1}}{\tau_j} R^j \quad \text{при } t \in (t_{j-1}, t_j], \quad j = 1, \dots, M, \quad \tilde{R}(\cdot, 0) = \hat{R}^0. \quad (5.37)$$

Заметим, что и \tilde{R} , и $I_{1,t}^* w$ понимаются как непрерывные справа при $t = 0$.

Т е о р е м а 5.17. Пусть u является решением задачи (4.1), (4.2) с параболическим оператором \mathcal{M} , удовлетворяющим Условию 4.1, а u_h^j — решением дискретной задачи (5.33). Тогда при $m = 1, \dots, M$ справедлива оценка (5.8), в которой $\chi_h^j = \tau_j (\psi_h^j - \hat{\psi}_h^{j-1})$, где используются ψ_h^j и $\hat{\psi}_h^{j-1}$ из (5.34), η^j определено в (5.10) с $\mathcal{A} = \{0^+, 1\}$, $C_1 = \frac{1}{8}$, $C_2 = \frac{1}{2}$, $C_1^* = 2$, $C_2^* = 3$, а

$$\vartheta_h := \tilde{\psi}_R - I_{1,t}^* \tilde{\psi}_R, \quad \tilde{\psi}_R := \mathcal{L}(t) \tilde{R} + f(\cdot, t, \tilde{R}) \quad (5.38)$$

при $t \in [0, T]$, где $I_{1,t}^*$ и \tilde{R} описаны в (5.36) и (5.37).

Доказательство. В силу Замечания 5.15, $\|R^j - u_h^j\|_{\infty, \Omega} \leq \eta^j$, поэтому для доказательства желаемой оценки типа (5.8) для $u_h^m - u(\cdot, t_m)$ достаточно установить соответствующую оценку типа (5.8) для $R^m - u(\cdot, t_m) = [\tilde{R} - u](\cdot, t_m)$, с той лишь разницей, что $(C_2^* \kappa_0 + 1)$ должна быть заменена на $C_2^* \kappa_0 = 3\kappa_0$. После этого останется лишь применить неравенство треугольника.

Перейдем к оценке $R^m - u(\cdot, t_m)$. Наши рассуждения будут аналогичны использованным при доказательстве Теоремы 4.10. При $t \in (t_{j-1}, t_j]$ справедливо представление $\tilde{\psi}_R = I_{1,t}^* \tilde{\psi}_R + \vartheta_h = \frac{1}{2}(\hat{\psi}_h^{j-1} + \psi_h^j) + \partial_t \mu_h + \vartheta_h$, где $\mu_h := \int_{t_j}^t [I_{1,t}^* \tilde{\psi}_R - \frac{1}{2}(\hat{\psi}_h^{j-1} + \psi_h^j)] dt$. Заметим, что интерполянт \tilde{R} определен таким образом, что определения в (5.38) идентичны с (5.15) (при $p = 2$), поэтому мы можем воспользоваться наблюдением (5.16), которое приводит к $\tilde{\psi}_R(\cdot, t_{j-1}^+) = \hat{\psi}_h^{j-1}$ и $\tilde{\psi}_R(\cdot, t_j) = \psi_h^j$. Следовательно мы получаем вариант представления (4.20):

$$\mu_h = \tau_j^{-1} \chi_h^j \int_{t_j}^t (t - t_{j-1/2}) dt = -\frac{1}{2}(t_j - t)(t - t_{j-1}) \cdot \tau_j^{-2} \chi_h^j \quad \text{при } t \in [t_{j-1}, t_j]. \quad (5.39)$$

Объединяя вышеизложенные соотношения с (5.35), приходим к

$$\delta_t^* u_h^j + \tilde{\psi}_R = \partial_t \mu_h + \vartheta_h \quad \text{при } t \in (t_{j-1}, t_j]. \quad (5.40)$$

Далее воспользуемся $I_{1,t}^* u_h$ из (5.28), для которого справедливы (5.29) и (5.30). Поскольку $\mathcal{M}\tilde{R} = \partial_t \tilde{R} + \tilde{\psi}_R$, а $\mathcal{M}u = 0$, то из (5.40) следует представление

$$\mathcal{M}\tilde{R} - \mathcal{M}u = \partial_t(\tilde{R} - I_{1,t}^* u_h) + \partial_t \mu_h + [\vartheta_h + \vartheta_*] \quad \text{в } Q. \quad (5.41)$$

Заметим, что $\partial_t \mu_h$ является регулярной функцией, поскольку μ_h , определенная в (5.39), непрерывна при $t \in [0, T]$.

Далее желаемая оценка типа (5.8) для $R^m - u(\cdot, t_m) = [\tilde{R} - u](\cdot, t_m)$ с $(C_2^* \kappa_0 + 1)$, замененной на $C_2^* \kappa_0 = 3\kappa_0$, может быть получена применением Леммы 4.3 к уравнению (5.41). При этом положим $\mu := \mu_R + \mu_h := (\tilde{R} - I_{1,t}^* u_h) + \mu_h$ и $\vartheta := \vartheta_h + \vartheta_*$, а также воспользуемся (5.30) и следующими двумя

наблюдениями. Во-первых, $[\tilde{R} - u - \mu](\cdot, 0) = \hat{R}^0 - \varphi - [(\hat{R}^0 - u_h^0) + 0] = u_h^0 - \varphi$. Во-вторых, при $t \in (t_{j-1}, t_j]$ справедливы оценки

$$\begin{aligned} |\mu_R| &\leq |\hat{R}^{j-1} - \hat{u}_h^{j-1}| + |R^j - u_h^j| \leq 2\eta^j, & |\mu_h| &\leq \frac{1}{8} |\chi_h^j|, \\ \tau_j |\partial_t \mu_R| &\leq |\hat{R}^{j-1} - \hat{u}_h^{j-1}| + |R^j - u_h^j| \leq 2\eta^j, & \tau_j |\partial_t \mu_h| &\leq \frac{1}{2} |\chi_h^j|, \end{aligned}$$

которые следуют из представления $\mu_R = \tilde{R} - I_{1,t}^* u_h = I_{1,t}^*(R - u_h)$, объединенного с Замечанием 5.15, а также здесь используется (5.39). Наконец, заметим, что $|\mu(\cdot, t_{m-1}^+)| = |\mu_R(\cdot, t_{m-1}^+)| \leq \eta^m$, поэтому множитель при η^m равен $1 + 2 = 3 = C_2^*$. Теорема доказана. \square

5.4.2 Модельная задача (4.3)

Пусть u является точным решением задачи (4.3) в ограниченной полигональной области $\Omega \subset \mathbb{R}^n$, $n = 1, 2, 3$, а u_h^j — решением дискретной задачи (5.18), при этом V_h^j и $\langle \cdot, \cdot \rangle_h$ на каждом временном уровне t_j таковы, что справедлива апостериорная ошибка типа (5.3) для соответствующей эллиптической задачи.

С л е д с т в и е 5.18. Апостериорная оценка Теоремы 5.17 справедлива для задачи (4.3), при этом ϑ_h упрощается к виду (5.17) при $p = 2$, а η^{j+1} описано в (5.10b) и использует эллиптический оценщик η из (5.3).

5.5 Полностью дискретный разрывный метод Галеркина на dG(1)

Для упрощения изложения, наше внимание будет в основном посвящено методу dG(r) при $r = 1$; при $r > 1$ см. Замечание 5.23. Рассмотрим полностью дискретный разрывный метод Галеркина dG(1) для задачи (4.3), получаемый применением пространственной дискретизации (5.5) к полудискретному методу dG(1) (4.29): При $j = 0, \dots, M - 1$ найти $u_h^{j+1/3}, u_h^{j+1} \in \mathring{V}_h^{j+1}$,

удовлетворяющие соотношениям

$$\begin{aligned} \left\langle \frac{u_h^{j+1} - \hat{u}_h^j}{\tau_{j+1}}, w_h \right\rangle_h + \varepsilon^2 \left\langle \frac{1}{4} \nabla (3u_h^{j+1/3} + u_h^{j+1}), \nabla w_h \right\rangle + \frac{1}{4} \left\langle 3f_h^{j+1/3} + f_h^{j+1}, w_h \right\rangle_h &= 0, \\ \left\langle \frac{u_h^{j+1/3} - \hat{u}_h^j}{\tau_{j+1}}, w_h \right\rangle_h + \varepsilon^2 \left\langle \frac{1}{12} \nabla (u_h^{j+1/3} - u_h^{j+1}), \nabla w_h \right\rangle + \frac{1}{12} \left\langle 5f_h^{j+1/3} - f_h^{j+1}, w_h \right\rangle_h &= 0, \end{aligned} \quad (5.42)$$

$\forall w_h \in \mathring{V}_h^{j+1}$, где используются $f_h^{j+\alpha} := f(\cdot, t_{j+\alpha}, u_h^{j+\alpha})$. Как и раньше, $\langle \cdot, \cdot \rangle_h$ — либо точное скалярное произведение $\langle \cdot, \cdot \rangle$ в $L_2(\Omega)$, либо же некоторая квадратичная аппроксимация для $\langle \cdot, \cdot \rangle$.

Заметим, что полностью дискретный метод (5.42) может быть представлен в виде

$$\mathcal{P}_h^{j+1}[u_h^{j+1} - \hat{u}_h^j] + \frac{1}{4} \tau_{j+1} (3\Psi^{j+1/3} + \Psi^{j+1}) = 0, \quad (5.43a)$$

$$\mathcal{P}_h^{j+1}[u_h^{j+1/3} - \hat{u}_h^j] + \frac{1}{12} \tau_{j+1} (5\Psi^{j+1/3} - \Psi^{j+1}) = 0. \quad (5.43b)$$

Здесь

$$\Psi^{j+\alpha} := \mathcal{L}_h(t_{j+\alpha}) u_h^{j+\alpha} + \mathcal{P}_h^{j+1}[f(\cdot, t_{j+\alpha}, u_h^{j+\alpha})] \quad \text{при } \alpha \in \{\frac{1}{3}, 1\}, \quad (5.43c)$$

где $\mathcal{L}_h(t)$ и \mathcal{P}_h^{j+1} удовлетворяют (5.6). Для некоторых частных случаев $\langle \cdot, \cdot \rangle_h$ операторы $\mathcal{L}_h(t)$ и \mathcal{P}_h^{j+1} описаны в Замечаниях 5.1 и 5.2. Таким образом, (5.43) можно интерпретировать как полную дискретизацию типа dG(1) для абстрактной параболической задачи (4.1), получаемую применением пространственной дискретизации (5.2) к полудискретной задаче (5.5).

Заметим, что $u_h^{j+\alpha}$ обращается в нуль на $\partial\Omega$, поэтому $I_h[f(\cdot, t_{j+\alpha}, 0)] = I_h[f(\cdot, t_{j+\alpha}, u_h^{j+\alpha})]$ на $\partial\Omega$, поэтому $\Psi^{j+\alpha} \in \mathring{V}_h^{j+1}$ при $\alpha \in \{\frac{1}{3}, 1\}$, и следовательно формулировка (5.43) корректна.

5.5.1 Апостериорный анализ на основе кусочно-квадратичных эллиптических реконструкций

Для оценки ошибки полностью дискретного метода dG(1) из (5.43) воспользуемся рассуждениями, аналогичными использованным в §4.6 при анализе

полудискретного метода. Положим $\mathcal{A} := \{0^+, \frac{1}{3}, 1\}$, при этом эллиптические реконструкции \hat{R}^j , $R^{j+1/3}$ и R^{j+1} описаны в (5.12). Их определения используют \hat{g}^j , $g^{j+1/3}$ и g^{j+1} , описанные в (5.10с), которые, в свою очередь, используют $\hat{\psi}_h^j$, $\psi_h^{j+1/3}$ и ψ_h^{j+1} . Определим последние следующим образом:

$$\psi_h^{j+1/3} := -\frac{u_h^{j+1} - \hat{u}_h^j}{2\tau_{j+1}} - 3\frac{u_h^{j+1/3} - \hat{u}_h^j}{2\tau_{j+1}}, \quad \psi_h^{j+1} := -5\frac{u_h^{j+1} - \hat{u}_h^j}{2\tau_{j+1}} + 9\frac{u_h^{j+1/3} - \hat{u}_h^j}{2\tau_{j+1}}, \quad (5.44a)$$

$$\hat{\psi}_h^j := \mathcal{L}_h(t_j^+) \hat{u}_h^j + \mathcal{P}_h^{j+1}[f(\cdot, t_j, \hat{u}_h^j)]. \quad (5.44b)$$

Заметим, что из (5.44a) следует вариант (5.43):

$$u_h^{j+1} - \hat{u}_h^j + \frac{1}{4} \tau_{j+1} (3\psi_h^{j+1/3} + \psi_h^{j+1}) = 0, \quad (5.45a)$$

$$u_h^{j+1/3} - \hat{u}_h^j + \frac{1}{12} \tau_{j+1} (5\psi_h^{j+1/3} - \psi_h^{j+1}) = 0. \quad (5.45b)$$

Если же $\hat{u}_h^j \in \mathring{V}_h^{j+1}$ (что соответствует случаю А из (5.7a)), то (5.45) и (5.43) становятся эквивалентными (при этом $\Psi^{j+\alpha} = \psi_h^{j+\alpha}$ при $\alpha \in \{\frac{1}{3}, 1\}$).

З а м е ч а н и е 5.19. Сравнение (5.45) и (5.43) приводит к $\Psi^{j+\alpha} = \mathcal{P}_h^{j+1} \psi_h^{j+\alpha}$ при $\alpha \in \{\frac{1}{3}, 1\}$. Поэтому, в силу (5.43с), заключаем, что $\psi_h^{j+1/3}$ и ψ_h^{j+1} удовлетворяют (5.9b). Также, $\hat{\psi}_h^j$ из (5.44b) принадлежит \mathring{V}_h^{j+1} , поэтому $\hat{\psi}_h^j$ удовлетворяет (5.9a). Следовательно \hat{R}^j , $R^{j+1/3}$ и R^{j+1} удовлетворяют (5.14) при $\mathcal{A} := \{0^+, \frac{1}{3}, 1\}$ и η^{j+1} из (5.10b).

З а м е ч а н и е 5.20. Теорема 5.22 остается справедливой для любой тройки функций $\hat{\psi}_h^j, \psi_h^{j+1/3}, \psi_h^{j+1}$, которые удовлетворяют (5.9) и (5.45). Например, можно заменить $\mathcal{P}_h^{j+1}[f(\cdot, t_j, \hat{u}_h^j)]$ в (5.44b) на $f(\cdot, t_j, \hat{u}_h^j)$, однако данная модификация несколько усложнит вычисление $\hat{\psi}_h^j$.

З а м е ч а н и е 5.21 (Случай В). В случае (5.7b) при $\hat{u}_h^j := u_h^j$ более естественно заменить (5.44b) на $\hat{\psi}_h^j := \psi_h^j$ (поскольку тогда $\hat{\psi}_h^j$ легко вычисляется). Заметим, что при данной замене (5.9a) перестает выполняться, а вот оценки (5.14) остаются справедливыми, если мы заменим $\eta(V_h^{j+1}, \hat{u}_h^j, \hat{g}^j(\cdot, \hat{u}_h^j))$ в (5.10b) на $\eta(V_h^j, u_h^j, g^j(\cdot, u_h^j))$. Следовательно Теорема 5.22 остается справедливой при описанных модификациях.

Чтобы сформулировать нашу оценку для ошибки $u_h - u$, обобщим кусочно-квадратичный интерполяционный оператор $I_{2,t}$ из (4.30) для любой непрерывной слева функции $w = w(t)$. Для этого будут использоваться интерполяционные узлы t_j^+ , $t_{j+1/3}$ и t_{j+1} :

$$I_{2,t}^* w(0) := w^0, \quad I_{2,t}^* w(t) := \hat{I}_{1,t} w(t) + 3 [2\hat{w}^j - 3w^{j+1/3} + w^{j+1}] \cdot \frac{1}{6} \tau_{j+1}^{-1} \zeta_2\left(\frac{t-t_j}{\tau_{j+1}}\right),$$

где $\hat{I}_{1,t} w(t) := w^{j+1} - \{w^{j+1} - w^{j+1/3}\} \cdot \frac{3}{2} \left(\frac{t_{j+1}-t}{\tau_{j+1}}\right)$ (5.46)

при $t \in (t_j, t_{j+1}]$, $j = 0, \dots, M-1$, при этом используются обозначения $\hat{w}^j := w(t_j^+)$ и $w^{j+\alpha} := w(t_{j+\alpha})$. Применяя $I_{2,t}^*$ к эллиптическим реконструкциям \hat{R}^j , $R^{j+1/3}$ и R^{j+1} , ассоциируемым с временными уровнями t_j^+ , $t_{j+1/3}$ и t_{j+1} , определим

$$\tilde{R}(\cdot, 0) := \hat{R}^0 = R^0, \quad \tilde{R}(\cdot, t) := I_{2,t}^* \{\hat{R}^j, R^{j+1/3}, R^{j+1}\} \quad \text{при } t \in (0, T]. \quad (5.47)$$

Аналогично определяется и кусочно-квадратичное по времени численное решение:

$$\tilde{u}_h(\cdot, 0) := \hat{u}_h^0 = u_h^0, \quad \tilde{u}_h(\cdot, t) := I_{2,t}^* \{\hat{u}_h^j, u_h^{j+1/3}, u_h^{j+1}\} \quad \text{при } t \in (0, T]. \quad (5.48)$$

Мы готовы к тому, чтобы сформулировать наш основной результат для полностью дискретного метода dG(1).

Т е о р е м а 5.22. Пусть u является решением задачи (4.1), (4.2), удовлетворяющей Условию 4.1, а u_h^j является решением (5.43) при любых $\hat{u}_h^j \in H_0^1(\Omega)$. Тогда справедлива апостериорная оценка (5.8), в которой

$$\chi_h^{j+1} := 3\tau_{j+1} [2\hat{\psi}_h^j - 3\psi_h^{j+1/3} + \psi_h^{j+1}], \quad (5.49)$$

где $\hat{\psi}^j$, $\psi_h^{j+1/3}$ и ψ_h^{j+1} описаны в (5.44), η^j — в (5.10). При этом $\mathcal{A} = \{0^+, \frac{1}{3}, 1\}$, постоянные $C_1 = \frac{2}{81}$, $C_2 = \frac{1}{6}$, $C_1^* = \frac{5}{3}$, $C_2^* = 10$, а

$$\vartheta_h := \tilde{\psi}_R - I_{2,t}^* \tilde{\psi}_R, \quad \tilde{\psi}_R := \mathcal{L}(t) \tilde{R} + f(\cdot, t, \tilde{R}) \quad (5.50)$$

при $t \in [0, T]$, где $I_{2,t}^*$ и \tilde{R} описаны в (5.46) и (5.47).

Д о к а з а т е л ь с т в о. Достаточно воспользоваться рассуждениями, аналогичными использованным при доказательстве Теоремы 4.13. (Подробное доказательство приведено в [175, §11.1].) \square

З а м е ч а н и е 5.23 ($dG(r)$ при $r > 1$). Вышеописанные результаты, включая апостериорную оценку из Теоремы 5.22, могут быть обобщены на случай полностью дискретных методов типа $dG(r)$ с квадратурой Радо при $r > 1$ (аналогично анализу, приведенному в §4.6 для полудискретных методов типа $dG(r)$). В этом случае в апостериорной оценке (5.8) будут присутствовать величины χ_h^{j+1} из (4.27), при этом $\psi^{j+\alpha}$ будут заменены на $\psi_h^{j+\alpha}$, а значения постоянных C_1 и C_2 останутся теми же, что и в Теореме 4.13.

5.5.2 Модельная задача (4.3)

Пусть u является точным решением задачи (4.3) в ограниченной полигональной области $\Omega \subset \mathbb{R}^n$, $n = 1, 2, 3$, а u_h^j — решением дискретной задачи (5.42), при этом V_h^j и $\langle \cdot, \cdot \rangle_h$ на каждом временном уровне t_j таковы, что справедлива апостериорная ошибка типа (5.3) для соответствующей эллиптической задачи.

С л е д с т в и е 5.24. Апостериорная оценка Теоремы 5.22 справедлива для задачи (4.3), при этом ϑ_h упрощается к виду (5.17) при $p = 3$, а в (5.10b) используется эллиптический оценщик η из (5.3).

З а м е ч а н и е 5.25 (Регулярный режим $\varepsilon = 1$). Пусть $\varepsilon = 1$, $\gamma \geq 0$, а V_h^j и $\langle \cdot, \cdot \rangle_h$ определены как описано в §5.6.1, т. е. \mathcal{T}_h^j является локально квазиравномерной триангуляцией области $\bar{\Omega}$, состоящей из элементов T , V_h^j — соответствующее конечноэлементное пространство непрерывных кусочно-полиномиальных функций степени $l \geq 1$, а $\dot{V}_h^j := V_h^j \cap H_0^1(\Omega)$. При этом используется квадратурная формула вида $\langle \varphi, w \rangle_h := \sum_{T \in \mathcal{T}_h^j} Q_T(\varphi w)$, описанная в §5.6.1. Тогда Следствие 5.24 применимо с $\eta = \eta_0$ из (5.51).

Разрывный метод Галеркина $dG(1)$ для линейного уравнения типа (4.3) при $\varepsilon = 1$ рассматривался в работах [83, 62]. В этом случае $f = f(x, t)$, поэтому (5.49) допускает представление $\chi_h^{j+1} = 6\tau_{j+1} \mathring{\mathcal{L}}_h^{j+1}[\hat{u}_h^j - \hat{I}_{1,t}u_h(t_j^+)] + \tau_{j+1}^3 \partial_t^2 f(x, t')$ при некотором промежуточном $t' \in [t_j, t_{j+1}]$. Здесь используется кусочно-линейный интерполяционный оператор $\hat{I}_{1,t}$ из (5.46), а также дискретный оператор $\mathring{\mathcal{L}}_h^{j+1}$, аналогичный \mathcal{L}_h^{j+1} , только теперь $\mathring{\mathcal{L}}_h^{j+1} : H_0^1(\Omega) \rightarrow \mathring{V}_h^{j+1}$. С описанными упрощениями, апостериорная оценка из Следствия 5.24 аналогична соответствующей оценке [83, (1.14)] (хотя они и неидентичны). Более подробно, в последней оценке также присутствуют $\tau_{j+1} \|\mathring{\mathcal{L}}_h^{j+1}[\hat{u}_h^j - \hat{I}_{1,t}u_h(t_j^+)]\|_{\infty, \Omega}$ и $\tau_{j+1}^3 \|\partial_t^2 f\|_{\infty, \Omega}$. (Важно отметить, что оценка [83, (1.14)] приводится без доказательства.) Апостериорная же оценка из [62] имеет порядок 2 по времени, поскольку в ней присутствуют слагаемые типа $\|\hat{u}_h^j - \hat{I}_{1,t}u_h(t_j^+)\|_{\infty, \Omega} = \mathcal{O}(\tau_{j+1}^2)$. Отметим также работу [115], в которой приводятся оценки для разрывного метода Галеркина в других нормах.

5.6 Эллиптические оценители

Здесь будут приведены некоторые известные эллиптические оценители типа (5.3) для частных случаев эллиптической модельной задачи (5.4) и ее дискретизаций в виде (5.5). Заметим, что в соответствующих ситуациях полученные в данном параграфе параболические оценители могут быть объединены с эллиптическими оценителями из §§1–3. Однако, главным образом с целью упрощения изложения, здесь будут рассмотрены более простые, по сравнению с §§1–3, случаи, и соответственно более простые эллиптические оценители.

5.6.1 Эллиптическая модельная задача: регулярный режим

Рассмотрим стационарный вариант (5.4) нашей модельной задачи (4.3) при $\varepsilon := 1$. Пусть v является точным решением задачи (5.4) при $\varepsilon = 1$, $\gamma \geq 0$

в ограниченной полигональной области $\Omega \subset \mathbb{R}^n$, $n = 2, 3$, а v_h — решением дискретной задачи (5.5) где V_h и $\langle \cdot, \cdot \rangle_h$ определяются следующим образом. Пусть \mathcal{T}_h^j является локально квазиравномерной триангуляцией области $\bar{\Omega}$, состоящей из элементов T , V_h^j — соответствующее конечноэлементное пространство непрерывных кусочно-полиномиальных функций степени $l \geq 1$, а $\mathring{V}_h^j := V_h^j \cap H_0^1(\Omega)$. При этом используется $\langle \varphi, w \rangle_h := \sum_{T \in \mathcal{T}_h^j} Q_T(\varphi w)$, где Q_T — некоторая квадратурная формула для интегралов по T с положительными весами и квадратурными точками в T , при этом формула Q_T является точной для многочленов степени q при $q \geq \max\{2l - 2, 1\}$.

При данных условиях Теорема 4.2 в работе [125] описывает следующую апостериорную оценку типа (5.3):

$$\begin{aligned} \eta_0(V_h, v_h, g(\cdot, v_h)) := & \left[c_0 \max_{T \in \mathcal{T}_h} \left\{ h_T^2 \|(\Delta v_h - g(\cdot, v_h))\|_{\infty, T} + h_T \|[[\partial_n v_h]]\|_{\infty, \partial T \setminus \partial \Omega} \right\} \right. \\ & \left. + c_1 \|\nu_{n/2, T}^q\|_{l_{n/2}} + c_2 \|h_T \nu_{n, T}^{q-1}\|_{l_n} \right] \times |\ln h_{\min}|^2. \end{aligned} \quad (5.51)$$

Здесь h_{\min} обозначает минимальный диаметр элементов сетки, h_T — диаметр элемента T , $[[\partial_n v_h]]$ — скачок нормальной производной при пересечении границы элемента. Также, $\|\cdot\|_{l_p}$ обозначает норму l_p , а величины

$$\nu_{n', T}^{q'} := |T|^{1/n'} \|g(\cdot, v_h) - I_{h, q'}[g(\cdot, v_h)]\|_{\infty, T}$$

определяются с использованием стандартного интерполяционного оператора Лагранжа $I_{h, q'}$ степени $\leq q'$.

5.6.2 Эллиптическая модельная задача: сингулярно-возмущенный режим

Далее рассмотрим стационарный вариант (5.4) нашей модельной задачи (4.3) при $\varepsilon \ll 1$ и $n = 1$. Пусть v является точным решением (5.4) при $\varepsilon \in (0, 1]$ и $\gamma > 0$ в области $\Omega := (0, 1)$, а v_h — решением дискретной задачи (5.5), где конечноэлементное пространство V_h состоит из непрерывных кусочно-линейных функций на произвольной неравномерной сетке $\{x_i\}_{i=1}^N$, при этом $0 = x_0 < x_1 < \dots < x_N = 1$ и $h_i := x_i - x_{i-1}$.

Рассмотрим два простых варианта для $\langle \cdot, \cdot \rangle_h$, которые описываются с помощью кусочно-линейного интерполяционного оператора I_h :

$$\langle \varphi, w_h \rangle_h := \langle I_h \varphi, w_h \rangle, \quad (\text{стандартная квадратура}) \quad (5.52a)$$

$$\langle \varphi, w_h \rangle_h := \langle I_h[\varphi w_h], 1 \rangle. \quad (\text{сосредоточенные массы}) \quad (5.52b)$$

З а м е ч а н и е 5.26. Чтобы проиллюстрировать Замечания 5.1 and 5.2, покажем что и (5.52a), и (5.52b) приводят к дискретизациям типа (5.2). Действительно, (5.52a) соответствует $\mathcal{L}_h := -\varepsilon^2[\partial_x^2]_h$ и $\mathcal{P}_h := I_h$. При этом оператор $[\partial_x^2]_h : H_0^1(\Omega) \rightarrow \dot{V}_h + \varepsilon^{-2}I_h[g(\cdot, 0)]$ определяется соотношением $\langle -[\partial_x^2]_h \varphi, w_h \rangle = \langle \varphi', w_h' \rangle$ при $\forall \varphi \in H_0^1(\Omega)$, $w_h \in \dot{V}_h$. Следовательно дискретная задача на основе (5.52a) может быть представлена в виде

$$-\varepsilon^2 [\partial_x^2]_h v_h + I_h[g(\cdot, v_h)] = 0. \quad (5.53a)$$

С другой стороны, (5.52b) приводит к стандартной трехточечной разностной схеме $-\varepsilon^2 \delta_x^2 v_{h,i} + g(x_i, v_{h,i}) = 0$ при $i = 1, \dots, N-1$, где используется $\delta_x^2 v_{h,i} := \frac{2}{h_i + h_{i+1}} \left[\frac{1}{h_{i+1}}(v_{h,i+1} - v_{h,i}) - \frac{1}{h_i}(v_{h,i} - v_{h,i-1}) \right]$. Положим $\delta_x^2 v_{h,i} := \varepsilon^{-2} g(x_i, v_{h,i})$ при $i = 0, N$. Тогда, применив линейную интерполяцию I_h к $\{\delta_x^2 v_{h,i}\}_{i=0}^N$, можем представить дискретную задачу на основе (5.52b) в виде

$$-\varepsilon^2 I_h[\delta_x^2 v_h] + I_h[g(\cdot, v_h)] = 0, \quad (5.53b)$$

где значения $\delta_x^2 v_{h,i}$ легко вычисляемы.

Опишем апостериорные оценки типа (5.3) из [164, 111, 112] с $\eta := \eta_\varepsilon$, где $\eta_\varepsilon = \eta_{\varepsilon, (5.52a)}$ при (5.52a) и $\eta_\varepsilon = \eta_{\varepsilon, (5.52b)}$ при (5.52b):

$$\eta_{\varepsilon, (5.52a)}(V_h, g_*) := \max_{i=1, \dots, N} \left\{ \frac{h_i^2}{4\varepsilon^2} \|I_h g_*\|_{\infty, (x_{i-1}, x_i)} \right\} + \gamma^{-2} \|g_* - I_h g_*\|_{\infty, (0,1)}, \quad (5.54a)$$

$$\eta_{\varepsilon, (5.52b)}(V_h, g_*) := \eta_{\varepsilon, (5.52a)} + \max_{i=1, \dots, N} \left\{ \frac{h_i^2}{6\gamma\varepsilon} \|\partial_x(I_h g_*)\|_{\infty, (x_{i-1}, x_i)} \right\}, \quad (5.54b)$$

где $g_* := g(\cdot, v_h)$.

З а м е ч а н и е 5.27. Эллиптические апостериорные оценки (5.54a) и (5.54b) являются робастными несмотря на то, что в них присутствуют отрицательные степени ε . Действительно, воспользовавшись представлениями (5.53a) и (5.53b), заключаем, что для данных двух дискретизаций величина $\varepsilon^{-2}h_i^2|I_h g_*| = \varepsilon^{-2}h_i^2|I_h[g(\cdot, v_h)]|$ становится, соответственно, $h_i^2|[\partial_x^2]_h v_h|$ и $h_i^2|\delta_x^2 v_h|$, т. е. может рассматриваться, как аппроксимирующая $h_i^2|\partial_x^2 v|$, где v является точным решением $-\varepsilon^2 \partial_x^2 v + g(\cdot, v) = 0$. Аналогично, $\varepsilon^{-1}h_i^2|\partial_x(I_h g_*)|$ аппроксимирует $\varepsilon|\partial_x^3 v| \simeq h_i^2|\partial_x^2 v|$.

Глава 3

Немонотонные эллиптические сингулярно возмущенные уравнения реакции-диффузии: асимптотический анализ и априорные оценки ошибки численного решения

§6 Априорные оценки ошибки численного ре- шения в норме максимума модуля в гладкой области

Под немонотонными полулинейными эллиптическими уравнениями будут пониматься уравнения, в которых нелинейность по неизвестной функции не является монотонной, поэтому такие уравнения могут иметь несколько решений. В гладкой двумерной области рассматривается полулинейное эллиптическое уравнение реакции-диффузии, коэффициент диффузии ε^2 в котором может принимать сколь угодно малые значения, что может приводить к образованию пограничных слоев. Целью настоящего параграфа является доказательство существования численных решений, а также исследование их точности на сгущающихся в пограничных слоях сетках типа Бахвалова и Шишкина. Установлен второй порядок сходимости (с логарифмическим множителем в случае сетки Шишкина) в сеточной норме максимума модуля равномерно по малому параметру ε при условии $\varepsilon \leq Ch$. Здесь $h > 0$ являет-

ся максимальным диаметром элементов сетки, в то время как число степеней свободы не превосходит Ch^{-2} . Полученные теоретические оценки сходимости будут проиллюстрированы результатами численных экспериментов.

Отметим, что, в силу немонотонности исходного уравнения в сочетании с наличием малого параметра при старших производных, стандартный технический аппарат, используемый при получении априорных оценок ошибки как для конечно-разностных, так и для конечноэлементных численных решений, становится неприменимым. В настоящем параграфе, при доказательстве существования и точности численных решений используется принципиально иной подход, основанный на построении дискретных верхних и нижних решений, при этом последние являются нелинейными модификациями асимптотического разложения для исходной задачи. Данный подход ранее использовался в работах [70, 143, 182] для обыкновенных дифференциальных уравнений, и впервые был обобщен автором диссертации на случаи эллиптических и параболических уравнений в [163, 181]. Основной материал параграфа опубликован в работе [163], а обобщения данного подхода кратко описаны в §6.5 и опубликованы в работах [179, 180, 181, 183].

6.1 Введение

Рассмотрим полулинейное сингулярно возмущенное уравнение реакции-диффузии

$$Fu \equiv -\varepsilon^2 \Delta u + b(x, u) = 0, \quad x = (x_1, x_2) \in \Omega \subset \mathbb{R}^2, \quad (6.1a)$$

$$u(x) = g(x), \quad x \in \partial\Omega, \quad (6.1b)$$

где ε — положительный параметр, принимающий сколь угодно малые значения, $\Delta = \partial_{x_1}^2 + \partial_{x_2}^2$ обозначает оператор Лапласа, а Ω является ограниченной двумерной областью с достаточно гладкой границей $\partial\Omega$. Также предполагается, что функции b и g являются достаточно гладкими. Нас интересуют погранслойные решения (6.1). Вообще говоря, решения (6.1) могут также

иметь внутренние переходные слои, которые в данной работе не рассматриваются (см., например, [19, 38, 147, 183]).

Полагая $\varepsilon = 0$ в (6.1a), получаем вырожденную версию (6.1):

$$b(x, u_0(x)) = 0 \quad \text{для } x \in \Omega. \quad (6.2)$$

Заметим, что решение u_0 уравнения (6.2), вообще говоря, не удовлетворяет граничному условию в (6.1b).

В литературе по вычислительной математике нередко предполагается (см., например, [14, 133, 10]), что уравнение (6.1a) монотонно по u и, более того, $b_u(x, u) > \gamma^2 > 0$ при любых $(x, u) \in \Omega \times \mathbb{R}^1$ для некоторой положительной постоянной γ . При данном условии вырожденная задача имеет единственное решение u_0 , которое является достаточно гладким в $\bar{\Omega}$ (что, в силу компактности $\bar{\Omega}$, следует из применения теоремы о неявной функции). Тем не менее, данное глобальное условие является довольно серьезным ограничением. Например, математические модели возникающие при моделировании биологических и химических процессов нередко используют уравнения типа (6.1), в которых $b(x, u)$ *немонотонна* по u [120, §14.7], [94, §2.3]. В связи с этим, в настоящем параграфе (6.1) будет рассмотрено при следующих более слабых *предположениях* [88, 38]:

A1 Существует *устойчивое вырожденное решение*, т. е. вырожденная задача (6.2) имеет достаточно гладкое решение u_0 , удовлетворяющее при некоторой постоянной $\gamma > 0$ условию

$$b_u(x, u_0) > \gamma^2 > 0 \quad \text{для любого } x \in \Omega.$$

A2 Граничное условие $g(x)$ удовлетворяет

$$\int_{u_0(x)}^v b(x, s) ds > 0 \quad \text{для любого } v \in (u_0(x), g(x)]', \quad x \in \partial\Omega.$$

Используемое здесь обозначение $(a, b]'$ совпадает с $(a, b]$ при $a < b$ и $[b, a)$ при $a > b$, в то время как $(a, b]' = \emptyset$ при $a = b$.

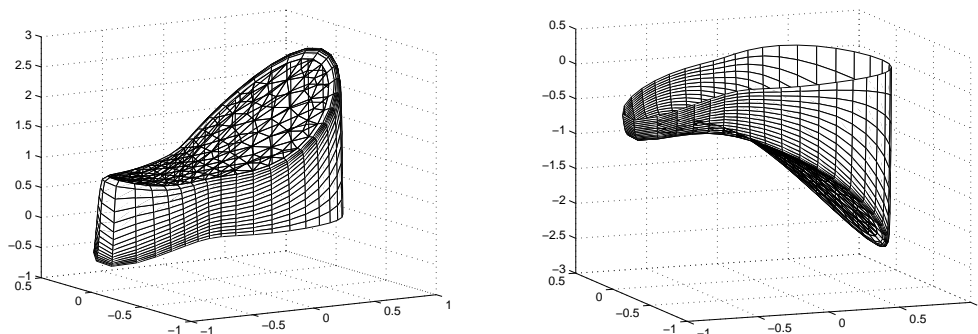


Рис. 6.1: Погранслойные решения модельной задачи (6.37): во внутренней части области $u(x) \approx \bar{u}_0(x)$ (слева) или $u(x) \approx -\bar{u}_0(x)$ (справа), где $\pm\bar{u}_0(x)$ — устойчивые решения вырожденной задачи (6.2).

Заметим, что если $g(x) \approx u_0(x)$, то A2 следует из A1 и (6.2), в то время как если $g(x) = u_0(x)$ в некоторой точке $x \in \partial\Omega$, то A2 не является ограничением на g в рассматриваемой точке.

Условия A1 и A2 естественным образом возникают при построении асимптотического разложения для задачи (6.1) и гарантируют, что существует погранслойное решение u , для которого справедливо соотношение $u \approx u_0$ во внутренней части области, в то время как вдоль границы области возникает пограничный слой ширины $O(\varepsilon |\ln \varepsilon|)$ [88, 38]; см. Теорему 6.2, а также [182], где условия A1 и A2 обсуждаются в одномерном случае. Заметим, что условие A1 локально, т. е. вырожденная задача (6.2) может обладать более, чем одним устойчивым решением. Если же несколько устойчивых решений вырожденной задачи удовлетворяют условиям A2, то исходная задача (6.1) обладает несколькими погранслойными решениями; см. Рис. 6.1.

Рассматриваемый нами численный метод использует дискретизацию области, которую иллюстрирует Рис. 6.2, (более подробное описание дано в §6.3.1) с использованием сгущающихся в окрестности погранслоя сеток типа Бахвалова и Шишкина, число узлов которых не превосходит Ch^{-2} . Здесь

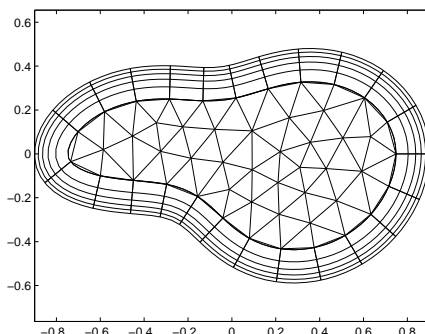


Рис. 6.2: Пример сгущающейся в погранслое сетки.

$h > 0$ — максимальный диаметр элементов сетки.

Что касается дискретизации уравнения (6.1a), то оно аппроксимируется с использованием метода конечных разностей на криволинейной прямоугольной части сетки и линейного метода конечных элементов с сосредоточенными массами на квазиравномерной триангуляции Делоне (Delaunay) во внутренней части области. С точки зрения используемого здесь технического аппарата, ключевым свойством рассматриваемого численного метода является то, что дискретизация оператора $-\Delta$ соответствует M -матрице. Последнее позволяет использовать дискретные нижние и верхние решения при доказательстве существования численного решения и исследовании его точности. Основным результатом настоящего параграфа, Теорема 6.28, устанавливает второй порядок сходимости (с логарифмическим множителем в случае сетки Шишкина) в сеточной норме максимума модуля равномерно по малому параметру ε .

С целью упрощения изложения будут сделаны еще два предположения. Чтобы не рассматривать несколько случаев, предполагается, что

$$u_0(x) < g(x) \quad \text{для всех } x \in \partial\Omega. \quad (6.3)$$

Также в данном параграфе предполагается, что

$$\varepsilon \leq Ch. \quad (6.4)$$

Последнее условие не является ограничением на практике при малых значениях ε , в то время как $\varepsilon = O(1)$ приводит к задаче (6.1), которая принципиально отлична от рассматриваемой здесь и с теоретической, и с вычислительной точек зрения.

Обыкновенное дифференциальное уравнение типа (6.1) было изучено в работах [70, 143, 182]. Данный параграф в большой степени обобщает подход [182] на случай эллиптического уравнения. Линейные монотонные версии задачи (6.1) рассматривались в [117, 66]. В [117] Меленк устанавливает оценки ошибки численного решения в энергетической норме для hp методов конечных элементов на сгущающихся в погранслоях сетках в случае криволинейной многоугольной области. В работе [66] получены оценки численного решения в норме максимума модуля для метода конечных разностей на прямоугольной сетке Шишкина в квадратной области. Полулинейная задача (6.1) при условии $b_u > \gamma^2 > 0$ для всех (x, u) рассмотрена в [133, 10, 11]. В работе [133], авторы устанавливают поточечные оценки ошибки для метода конечных элементов на квазиравномерных сетках в многоугольных областях. Блатов [10, 11] показывает сходимость второго порядка в норме максимума модуля (при условии $\varepsilon |\ln \varepsilon| \leq Ch$ аналогичном (6.4)). Заметим, что рассматриваемые здесь сетки аналогичны [10, 11], в то время как рассматриваемый численный метод отличен, поскольку в [10, 11] рассматриваются билинейные конечные элементы на прямоугольной в криволинейных координатах части сетки (используемой вдоль границы). Отметим также, что техника анализа, используемая в данном параграфе, весьма сильно отличается от техники Блатова. В частности, Блатов оценивает ошибку численного решения через ошибку интерполяции, что приводит к весьма технически трудоемким доказательствам, в то время как в данном параграфе при получении оценок численного решения будет использован принципиально иной подход, основан-

ный на нелинейной модификации асимптотических разложений для исходной задачи (6.1) при построении дискретных нижних и верхних решений.

Настоящий параграф организован следующим образом. Мы начнем с обсуждения асимптотических свойств решений задачи (6.1) и построения верхних и нижних решений в §6.2. Далее, в §6.3 будут описаны сгущающиеся в окрестности погранслоя сетки, а также рассматриваемый нами численный метод для задачи (6.1). Также будут установлены существование и оценки сходимости для численных решений, полученных на сетках типа Бахвалова и Шишкина. Полученные теоретические оценки будут проиллюстрированы результатами численных экспериментов в §6.4. В заключение, обобщения используемого подхода будут кратко описаны в §6.5

Обозначения. В данном параграфе C используется для обозначения положительных постоянных, возможно принимающих различные значения, но всегда независимых от h и ε . Пронумерованные постоянные C (например, C_1) принимают фиксированные значения, оставаясь независимыми от h и ε . Для любых двух величин w_1 и w_2 , обозначение $w_1 = O(w_2)$ подразумевает, что $|w_1| \leq Cw_2$. Для любой функции $g \in C(\bar{\Omega})$, обозначения g_i или g_{ij} соответственно используются для $g(X_i)$ или $g(x_{ij})$, где $X_i \in \bar{\Omega}$ и $x_{ij} \in \bar{\Omega}$ являются узлами сетки.

6.2 Асимптотический анализ, верхние и нижние решения

Кратко опишем асимптотическое разложение из [38]; см. также [21, §4], [147, §3.1.1]) для линейного случая. Далее, оно будет модифицировано, опять частично следуя [38], с целью построения верхних и нижних решений для задачи (6.1).

6.2.1 Локальные криволинейные координаты

Параметризуем границу области $\partial\Omega$:

$$x_1 = \varphi(l), \quad x_2 = \psi(l), \quad 0 \leq l \leq L,$$

где $(\varphi(0), \psi(0)) = (\varphi(L), \psi(L))$, и при возрастании l область остается слева. Все функции переменной l определенные за пределами интервала $[0, L]$ будут пониматься как L -периодические. Будет использоваться обозначение $\tau > 0$ для длины тангенциального вектора (φ', ψ') , а также кривизна κ границы в точке $(\varphi(l), \psi(l))$, которые определяются как

$$\tau = \sqrt{\varphi'^2 + \psi'^2}, \quad \kappa = \kappa(l) = \frac{\varphi'\psi'' - \psi'\varphi''}{\tau^3}.$$

В узкой окрестности $\partial\Omega$, которая будет определена позже, введем криволинейные локальные координаты (r, l) :

$$x_1 = \varphi(l) + rn_1(l), \quad x_2 = \psi(l) + rn_2(l), \quad (6.5)$$

где (n_1, n_2) единичный внутренний нормальный вектор к границе $\partial\Omega$ в точке $(\varphi(l), \psi(l))$. Таким образом этот вектор ортогонален к тангенциальному вектору (φ', ψ') и определяется соотношениями

$$n_1 = \frac{-\psi'}{\tau}, \quad n_2 = \frac{\varphi'}{\tau}.$$

Поскольку граница $\partial\Omega$ является гладкой, найдется такая достаточно малая постоянная C_1 , что в подобласти $\bar{\Omega}_{C_1} = \{0 \leq r \leq C_1\}$ новые координаты однозначны и корректно определены, а отображение $(r, l) \mapsto (x_1, x_2)$ является взаимно однозначным. При использовании новых координат, удобно использовать неотрицательную гладкую функцию-срезку $\omega(x)$, которая равна 1 при $r \leq C_1/2$ и исчезает в $\bar{\Omega} \setminus \bar{\Omega}_{C_1}$.

Л е м м а 6.1 ([163, Appendix A]). *Криволинейные координаты (6.5) являются ортогональными, а оператор Лапласа допускает представление*

$$\Delta u = \eta^{-1} \frac{\partial}{\partial r} \left(\eta \frac{\partial u}{\partial r} \right) + \zeta \frac{\partial}{\partial l} \left(\zeta \frac{\partial u}{\partial l} \right), \quad \text{где } \eta := 1 - \kappa r, \quad \zeta := (\tau\eta)^{-1}. \quad (6.6)$$

6.2.2 Асимптотическое разложение первого порядка

Введем растянутую переменную

$$\xi := r/\varepsilon \quad (6.7)$$

и функции $v_0(\xi, l)$ и $v_1(\xi, l)$, удовлетворяющие

$$-\frac{\partial^2 v_0}{\partial \xi^2} + b(\bar{x}, u_0(\bar{x}) + v_0) = 0, \quad (6.8)$$

$$-\frac{\partial^2 v_1}{\partial \xi^2} + v_1 b_u(\bar{x}, u_0(\bar{x}) + v_0) = -\xi \frac{d}{dr} b(x, u_0(x) + t) \Big|_{x=\bar{x}, t=v_0} - \kappa \frac{\partial v_0}{\partial \xi} \quad (6.9)$$

при $\xi > 0$, а также граничным условиям

$$v_0(0, l) = g(\bar{x}) - u_0(\bar{x}), \quad v_1(0, l) = 0, \quad v_0(\infty, l) = v_1(\infty, l) = 0. \quad (6.10)$$

Здесь $x = x(r, l)$ определяется (6.5), а

$$\bar{x} = \bar{x}(l) := (\varphi(l), \psi(l)), \quad \eta^{-1} \frac{\partial \eta}{\partial r} \Big|_{r=0} = \frac{-\kappa}{1 - \kappa r} \Big|_{r=0} = -\kappa.$$

Заметим, что для любого l уравнение, определяющее v_0 , является автономным нелинейным обыкновенным дифференциальным уравнением, а условия A1 и A2 обеспечивают существование и асимптотические свойства решений v_0 и v_1 [182, 38]; см. также Лемму 6.4.

Следующий результат описывает значительно упрощенную версию Теоремы 3 из [38].

Т е о р е м а 6.2. [38] *При выполнении условий A1 и A2, для достаточно малых значений ε существует решение $u(x)$ задачи (6.1) в окрестности асимптотического разложения нулевого порядка $u_0(x) + v_0(\xi, l)\omega(x)$. В то же время для асимптотического разложения первого порядка*

$$u_{\text{as}}(x) := u_0(x) + [v_0(\xi, l) + \varepsilon v_1(\xi, l)]\omega(x)$$

справедливо

$$|Fu_{\text{as}}(x)| \leq C\varepsilon^2, \quad |u(x) - u_{\text{as}}(x)| \leq C\varepsilon^2 \quad \text{для любого } x \in \bar{\Omega}. \quad (6.11)$$

З а м е ч а н и е 6.3. *Асимптотические разложения из Теоремы 6.2 описывает некоторые аппроксимации для u . Заметим, что, вообще говоря, вычисление приближенных решений по асимптотическим формулам является весьма дорогостоящим, поскольку при этом необходимо решить, возможно численно, несколько вспомогательных, возможно нелинейных, задач. Поэтому при наличии эффективных численных методов, применимых к рассматриваемому классу задач, применение последних к решению исходного уравнения представляется более простым и эффективным.*

6.2.3 Модифицированное асимптотическое разложение, верхние и нижние решения

При построении верхних и нижних решений будет использоваться функция $v(\xi, l; p)$, удовлетворяющая

$$-\frac{\partial^2 v}{\partial \xi^2} + b(\bar{x}, u_0(\bar{x}) + v) = pv, \quad (6.12)$$

с граничными условиями

$$v(0, l; p) = g(\bar{x}) - u_0(\bar{x}), \quad v(\infty, l; p) = 0. \quad (6.13)$$

Очевидно, что $v(\xi, l; p)$ является обобщением $v_0(\xi, l)$ из (6.8), поскольку $v(\xi, l; 0) = v_0(\xi, l)$. Далее введем

$$\beta(x; p) := u_0(x) + [v(\xi, l; p) + \varepsilon v_1(\xi, l)] \omega(x) + C_0 p, \quad (6.14)$$

где C_0 является достаточно малой положительной постоянной, которая будет определена ниже. Отметим, что (6.14) можно переписать в виде

$$\beta(x; p) = u_{\text{as}}(x) + w(\xi, l; p) + C_0 p, \quad w(\xi, l; p) := [v(\xi, l; p) - v_0(\xi, l)] \omega(x). \quad (6.15)$$

Для некоторых малых значений $p > 0$ функции $\beta(x; -p)$ и $\beta(x; p)$ будут соответственно выступать в роли нижнего и верхнего решений. Параметр p ,

используемый в определениях v и β , всегда предполагается достаточно малым и обычно выбирается $o(\varepsilon)$ или, при рассмотрении численных решений, $o(h)$. Таким образом, в силу следующей леммы, определение функции β имеет смысл.

Л е м м а 6.4. Положим $\gamma_0^2 := \min_{x \in \partial\Omega} b_u(x, u_0(x)) > 0$. Найдется такое значение $p_0 \in (0, \gamma_0^2)$, что для любого $|p| \leq p_0$ существуют $v_0(\xi, l)$, $v_1(\xi, l)$ и $v(\xi, l; p)$, которые удовлетворяют (6.8), (6.9), (6.10), (6.12) и (6.13). Также $v_0 \geq 0$ и $v \geq 0$. Более того, при $\tilde{\gamma}_0^2 := \gamma_0^2 - p_0$ и произвольном фиксированном $\delta \in (0, \tilde{\gamma}_0)$ существует такая положительная постоянная C_δ , что

$$\frac{\partial v(\xi, l; p)}{\partial p} \geq 0 \quad (6.16)$$

и

$$\left| \frac{\partial^k v(\xi, l; p)}{\partial \xi^k} \right| + \left| \frac{\partial^k v_1(\xi, l)}{\partial \xi^k} \right| + \left| \frac{\partial v(\xi, l; p)}{\partial p} \right| + \left| \frac{\partial^2 v(\xi, l; p)}{\partial p \partial \xi} \right| \leq C_\delta e^{-(\tilde{\gamma}_0 - \delta)\xi} \quad (6.17)$$

для $0 \leq \xi < \infty$, и любого l и $k = 0, 1, \dots, 4$.

Д о к а з а т е л ь с т в о. Настоящая лемма обобщает [182, Лемму 2.3 и (2.15)]. Оценка для $\partial_p \partial_\xi v$ доказывается аналогично похожей оценке для $\partial_p v$. \square

З а м е ч а н и е 6.5. В силу A1, $\gamma_0 > \gamma$. Выбирая p_0 и δ достаточно малыми, можно обеспечить, что величина $\tilde{\gamma}_0 - \delta$ в (6.17) достаточно близка к γ_0 , а $\tilde{\gamma}_0 - \delta > \gamma$.

С л е д с т в и е 6.6. Справедливо соотношение $(1 + \xi)|w| + |w_\xi| \leq Cp$, где w определяются в (6.15).

Д о к а з а т е л ь с т в о. Желаемое утверждение следует из (6.17) для $\partial_p v$ и $\partial_p \partial_\xi v$. \square

С л е д с т в и е 6.7. Справедливо соотношение $\beta(x; p) = u_{\text{as}}(x) + O(p) = u(x) + O(p + \varepsilon^2)$. Также, если $p \geq 0$, то $\beta(x; -p) \leq \beta(x; p)$.

Доказательство. Эти два утверждения вытекают, соответственно, из Следствия 6.6 и (6.11), а также (6.14) и (6.16). \square

Лемма 6.8. *Справедливо соотношение*

$$F\beta = C_0 p b_u(x, u_0) + [1 + C_0 \lambda(x)] p v_0(\xi, l) + O(\varepsilon^2 + p^2),$$

где $\lambda(x) := b_{uu}(x, u_0 + \vartheta v_0)$ с $\vartheta = \vartheta(x) \in (0, 1)$.

Доказательство. Для упрощения изложения предположим, что функция-срезка ω равна 1 во всей области $\bar{\Omega}$. Поскольку это неверно при $\varepsilon\xi > C_1/2$, где функции v и v_1 вместе со своими производными являются пренебрежимо малыми, данное предположение не влияет на наши теоретические выводы.

На протяжении доказательства будут использованы некоторые дополнительные обозначения: $\bar{x} := (\varphi(l), \psi(l))$ и $\bar{u}_0 := u_0(\bar{x})$, а u_0 , v_0 , v и w соответственно понимаются как $u_0(x)$, $v_0(\xi, l)$, $v(\xi, l; p)$ and $w(\xi, l; p)$. Отметим также, что $v_0 = v(\xi, l; 0)$.

Из (6.11) следует, что $\varepsilon^2 \Delta u_{\text{as}} = b(x, u_{\text{as}}) + O(\varepsilon^2)$. Далее, в силу (6.6), (6.7), (6.15) справедливо, что $\varepsilon^2 \Delta w = w_{\xi\xi} + O(\varepsilon|w_\xi| + \varepsilon^2)$. Поскольку в силу Следствия 6.6 $|w_\xi| \leq Cp$, получаем

$$F\beta = -b(x, u_{\text{as}}) - w_{\xi\xi} + b(x, u_{\text{as}} + w + C_0 p) + O(p^2 + \varepsilon^2).$$

Далее, воспользовавшись (6.12), а также $v = v_0 + w$ и $w = O(p)$,

$$w_{\xi\xi} = [b(\bar{x}, \bar{u}_0 + v_0 + w) - b(\bar{x}, \bar{u}_0 + v_0)] - p v_0 + O(p^2).$$

Следовательно

$$\begin{aligned} F\beta = & -b(x, u_{\text{as}}) - [b(\bar{x}, \bar{u}_0 + v_0 + w) - b(\bar{x}, \bar{u}_0 + v_0)] + p v_0 \\ & + b(x, u_{\text{as}} + w + C_0 p) + O(p^2 + \varepsilon^2). \end{aligned}$$

Воспользовавшись функцией

$$\mu(t) := b(x, u_{\text{as}} + t) - b(\bar{x}, \bar{u}_0 + v_0 + t),$$

перепишем соотношение для $F\beta$ в виде

$$F\beta = [b(x, u_{\text{as}} + w + C_0p) - b(x, u_{\text{as}} + w)] + [\mu(w) - \mu(0)] + pv_0 + O(p^2 + \varepsilon^2). \quad (6.18)$$

Поскольку $u_{\text{as}} = u_0 + v_0 + O(\varepsilon)$ и $w = O(p)$, получаем

$$b(x, u_{\text{as}} + w + C_0p) - b(x, u_{\text{as}} + w) = [b_u(x, u_0 + v_0) + O(\varepsilon + p)]C_0p, \quad (6.19)$$

где

$$b_u(x, u_0 + v_0)C_0p = [b_u(x, u_0) + \lambda(x)v_0]C_0p. \quad (6.20)$$

Далее,

$$\mu(w) - \mu(0) = w\mu'(\bar{\vartheta}w) = [b_u(x, u_{\text{as}} + \bar{\vartheta}w) - b_u(\bar{x}, \bar{u}_0 + v_0 + \bar{\vartheta}w)]w,$$

где $\bar{\vartheta} = \bar{\vartheta}(x) \in (0, 1)$. Поэтому, в силу Следствия 6.6, заключаем, что

$$\mu(w) - \mu(0) = O(\varepsilon)[1 + \xi]|w| = O(\varepsilon p). \quad (6.21)$$

Объединяя (6.18)–(6.21), завершаем доказательство. \square

С л е д с т в и е 6.9. *Существует $C_0 > 0$ такая, что для любого $|p| \leq p_0$ выполняется*

$$\begin{aligned} F\beta &\geq C_0p\gamma^2 + O(\varepsilon^2 + p^2) \quad \text{при } p > 0, \\ F\beta &\leq -C_0|p|\gamma^2 + O(\varepsilon^2 + p^2) \quad \text{при } p < 0. \end{aligned}$$

Д о к а з а т е л ь с т в о. Вспомним условие A1 и соотношение $v_0 \geq 0$, вытекающее из Леммы 6.4. Выбирая $0 < C_0 \leq |\lambda(x)|^{-1}$ для любого x , получаем $1 + C_0\lambda(x) \geq 0$. \square

Таким образом, если $p > 0$ и слагаемое $C_0p\gamma^2$ доминирует в соотношениях Следствия 6.9 по сравнению с $O(\varepsilon^2 + p^2)$, то $F\beta(x; -p) \leq 0 \leq F\beta(x; p)$, в то время как в силу Следствия 6.7 $\beta(x; -p) \leq \beta(x; p)$. Такие функции $\beta(x; -p)$ и $\beta(x; p)$ называются нижним и верхним решениями задачи (6.1).

6.3 Анализ численного метода

6.3.1 Сгущающиеся в окрестности погранслоя сетки

Введем достаточно малый положительный параметр σ , который будет определен ниже. Пусть $\sigma \leq C_1$, с тем, чтобы замкнутая кривая $\partial\Omega_\sigma$, которая определяется уравнением $r = \sigma$, не имела самопересечений. Далее, пусть Ω_σ является областью внутри $\partial\Omega_\sigma$. Наше уравнение (6.1a) будет аппроксимироваться отдельно в подобластях Ω_σ и $\Omega \setminus \Omega_\sigma$, которые мы, соответственно, будем называть внутренней подобластью и погранслошной подобластью (см. Рис. 6.2).

Погранслошная подобласть $\Omega \setminus \Omega_\sigma$ является прямоугольником $(0, \sigma) \times [0, L]$ в координатах (r, l) . В данных координатах, в этой подобласти введем прямоугольную сетку $\{(r_i, l_j), i = 0, \dots, N, j = -1, \dots, N_l\}$, где, как обычно, $r_0 = 0$, $r_N = \sigma$, $l_0 = 0$, $l_{N_l} = L$, а $l_{-1} = l_{N_l-1} - L$. Будем предполагать, что $\{l_j\}$ является квазиравномерной сеткой на отрезке $[0, L]$, т. е. $C^{-1}h \leq l_j - l_{j-1} \leq Ch$. Выбор сгущения $\{r_i\}$ в погранслое на отрезке $[0, \sigma]$ является ключевым для сходимости метода и будет описан ниже. Пока предполагается лишь, что $r_i - r_{i-1} \leq h$ и

$$C^{-1}h^{-1} \leq N \leq Ch^{-1}. \quad (6.22)$$

Во внутренней подобласти Ω_σ вводится квазиравномерная триангуляция Делоне (Delaunay), т. е. максимальный диаметр элементов сетки не превосходит h , площадь любого треугольника сетки ограничена снизу величиной Ch^2 , а сумма углов, противоположных любому ребру не превосходит π (в то время, как любой угол напротив $\partial\Omega_\sigma$ не превосходит $\pi/2$). Для таких сеток известно, что аппроксимация оператора $-\Delta$ линейными конечными элементами приводит к M -матрице.

Далее, пусть объединение всех треугольных элементов сетки в Ω_σ определяет полигональную область Ω_σ^h , все граничные узлы которой расположены на $\partial\Omega_\sigma$. Также будет предполагаться, что узлы сеток во внутренней

и погранслошной подобластях совпадают на $\partial\Omega_\sigma$. Отметим, при численном решении исходная область Ω не заменяется некоторой полигональной областью Ω^h , поскольку в этом случае значительная часть пограничного слоя не войдет в $\Omega \setminus \Omega^h$.

Будут рассмотрены два частных случая погранслошного сгущения $\{r_i\}$:

6.3.1(a) Сетка Бахвалова. [8] Положим $\sigma := 2\gamma^{-1}\varepsilon |\ln \varepsilon|$ и определим сетку соотношениями

$$r_i := r([1 - \varepsilon]i/N), \quad i = 0 \dots, N, \quad r(t) := -2\gamma^{-1}\varepsilon \ln(1 - t) \quad \text{при } t \in [0, 1 - \varepsilon].$$

6.3.1(b) Сетка Шишкина. [49] Положим $\sigma = 2\gamma^{-1}\varepsilon \ln N$ и введем равномерную сетку $\{r_i\}_{i=0}^N$ на $[0, \sigma]$, т. е. $r_i - r_{i-1} = \sigma/N = 2\gamma^{-1}\varepsilon N^{-1} \ln N$.

Отметим, что если ε является достаточно малым (в силу (6.4)), то выполняется условие $\sigma \leq C_1$ и обе сетки (a) и (b) определены корректно.

З а м е ч а н и е 6.10. Если (6.4) не удовлетворяется, но (a) $\sigma \leq C_1$ и $\varepsilon \leq 1/2$, или (b) $\sigma \leq C_1$, то сетки 6.3.1(a) и 6.3.1(b) остаются корректно определенными. В противном случае, т. е. при $\varepsilon > C$, рассматриваемая задача не является сингулярно возмущенной. Тогда, имитируя [8, 49], можно обобщить сетки 6.3.1(a) и 6.3.1(b), например, используя (b), в котором полагаем $\sigma := C_1$. В качестве альтернативы, можно просто использовать линейные конечные элементы на квазиравномерной триангуляции Делоне во всей области $\bar{\Omega}$.

З а м е ч а н и е 6.11. В определениях сеток 6.3.1(a) и 6.3.1(b) постоянная γ из A1 может быть заменена любой постоянной $\tilde{\gamma} \in (0, \gamma_0)$, где используется γ_0 из Леммы 6.4 (см. Замечание 6.5).

6.3.2 Z-поля

О п р е д е л е н и е 6.12. Оператор $H : \mathbb{R}^n \rightarrow \mathbb{R}^n$ называется Z-полем, если для любых $i \neq j$ отображение $x_j \mapsto (H(x_1, x_2, \dots, x_n))_i$ является монотонно убывающей функцией из \mathbb{R} в \mathbb{R} при фиксированных $x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n$.

З а м е ч а н и е 6.13. Если H дифференцируемо, то H является Z -полем тогда и только тогда, когда внедиагональные элементы матрицы Якоби этого преобразования являются неположительными.

Мы воспользуемся следующим результатом Лоренца [113], доказательство которого также можно найти в [182].

Л е м м а 6.14. [113] Пусть отображение $H : \mathbb{R}^n \rightarrow \mathbb{R}^n$ непрерывно и является Z -полем, и пусть дано некоторое $r \in \mathbb{R}^n$. Предположим, что функции $\alpha, \beta \in \mathbb{R}^n$ таковы, что $\alpha \leq \beta$ и $H\alpha \leq r \leq H\beta$. (Здесь векторные неравенства понимаются поточечно.) Тогда уравнение $Hu = r$ имеет решение $u \in \mathbb{R}^n$, которое удовлетворяет $\alpha \leq u \leq \beta$.

З а м е ч а н и е 6.15. Функции α и β в Лемме 6.14 называются соответственно нижним и верхним решениями дискретной задачи $Hu = r$.

З а м е ч а н и е 6.16. Пусть X_1, X_2, \dots, X_n являются внутренними точками Ω , точки $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_m$ расположены на $\partial\Omega$, а $U \in \mathbb{R}^{n+m}$ является дискретной функцией, определенной в этих точках. Предположим, что оператор $F^h : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^n$ может быть представлен в виде

$$F^h U = \varepsilon^2 \Lambda U + [b(X_i, U_i)]_{i=1}^n, \quad (6.23)$$

где Λ соответствует M -матрице и является аппроксимацией оператора $-\Delta$. Тогда отображение

$$(X_1, \dots, X_n, \bar{X}_1, \dots, \bar{X}_m) \mapsto (F^h U_1, \dots, F^h U_n, g(\bar{X}_1), \dots, g(\bar{X}_m))$$

является Z -полем.

З а м е ч а н и е 6.17. Таким образом, чтобы воспользоваться теорией Z -полей в нашем анализе, будут наложены следующие требования на аппроксимации нашей задачи. Пусть (i) аппроксимация оператора $-\Delta$ соответствует M -матрице; (ii) аппроксимация $b(x, u)$ в любой внутренней

точке сетки X_i использует только U_k с $k = i$. Чтобы удовлетворить этим условиям, удобно использовать метод конечных разностей в погранслойной подобласти и линейные конечные элементы с сосредоточенными массами на триангляциях Делоне во внутренней подобласти.

6.3.3 Аппроксимация во внутренних точках криволинейной прямоугольной сетки в погранслойной подобласти

Вспомним, что $\Omega \setminus \Omega_\sigma$ является прямоугольником $(0, \sigma) \times [0, L]$ в криволинейных координатах (r, l) . Перепишем (6.1a) в координатах (r, l) , используя (6.6), а затем аппроксимируем получившееся уравнение на прямоугольной сетке $\{(r_i, l_j)\}$ с помощью стандартного метода конечных разностей [41, 42]. Во внутренних точках сетки в $\Omega \setminus \Omega_\sigma$, т. е. при $i = 1, \dots, N-1, j = 0, \dots, N_l-1$, будет использована следующая аппроксимация:

$$\begin{aligned} F^h U_{ij} &:= -\varepsilon^2 \eta_{ij}^{-1} D_r [\tilde{\eta}_{ij} D_r^- U_{ij}] - \varepsilon^2 \zeta_{ij} D_l [\tilde{\zeta}_{ij} D_l^- U_{ij}] + b(x_{ij}, U_{ij}) = 0, \\ U_{i, N_l} &= U_{i, 0}, \quad U_{i, -1} = U_{i, N_l - 1}, \quad U_{0, j} = g(x_{0, j}). \end{aligned} \quad (6.24)$$

Здесь U_{ij} обозначает значение численного решения в узле сетки x_{ij} ,

$$\begin{aligned} D_r^- v_{ij} &:= \frac{v_{ij} - v_{i-1, j}}{r_i - r_{i-1}}, & D_r v_{ij} &:= \frac{v_{i+1, j} - v_{ij}}{(r_{i+1} - r_{i-1})/2}, \\ D_l^- v_{ij} &:= \frac{v_{ij} - v_{i, j-1}}{l_j - l_{j-1}}, & D_l v_{ij} &:= \frac{v_{i, j+1} - v_{ij}}{(l_{j+1} - l_{j-1})/2}, \end{aligned}$$

и

$$\begin{aligned} \eta_{ij} &:= \eta(r_i, l_j), & \zeta_{ij} &:= \zeta(r_i, l_j), & x_{ij} &:= x(r_i, l_j), \\ \tilde{\eta}_{ij} &:= \eta(r_{i-1/2}, l_j), & \tilde{\zeta}_{ij} &:= \zeta(r_i, l_{j-1/2}). \end{aligned}$$

З а м е ч а н и е 6.18. Аппроксимация (6.24) принадлежит типу (6.23).

Л е м м а 6.19. Пусть $\beta(x; p)$ определяется из (6.14), а сетка $\{r_i\}_{i=0}^N$ является либо сеткой Бахвалова из §6.3.1(a), либо сеткой Шишкина из

§6.3.1(b). Тогда при любом $|p| \leq p_0$ для любого внутреннего узла сетки $x_{ij}, i = 1, \dots, N-1, j = 0, \dots, N_l-1$, справедливо соотношение

$$|F^h \beta_{ij} - F\beta(x_{ij})| \leq Ch^2 |\ln h|^m,$$

где $m = 0$ для сетки Бахвалова (а) и $m = 2$ для сетки Шлишкина (б).

Д о к а з а т е л ь с т в о. Заметим, что компонента $\varepsilon^2 \zeta \partial_l (\zeta \partial_l \beta)$ в $F\beta$ и ее дискретный аналог в $F^h \beta$ имеют порядок $O(\varepsilon^2) \leq O(h^2)$ в силу (6.4). Далее, желаемое утверждение устанавливается имитацией доказательств в [182, Лемме 3.3 и §3.4.2] с использованием (6.22). \square

6.3.4 Аппроксимация в погранслойной подобласти на интерфейсной границе $\partial\Omega_\sigma$

Для аппроксимации на интерфейсной границе $\partial\Omega_\sigma$ удобно ввести фиктивное условие Неймана

$$\frac{\partial u}{\partial r} = \phi(x) \quad \text{для } x \in \partial\Omega_\sigma, \quad (6.25)$$

которое будет исключено при объединении аппроксимаций во внутренней и погранслойной подобластях.

При $i = N, j = 0, \dots, N_l-1$, используя подход из [41, 42], уравнения (6.1а), (6.6) с граничным условием (6.25) аппроксимируются следующим образом:

$$\begin{aligned} F_-^h U_{Nj} &:= -\varepsilon^2 \delta_r^2 U_{Nj} - \varepsilon^2 \zeta_{Nj} D_l [\tilde{\zeta}_{Nj} D_l^- U_{Nj}] + b(x_{Nj}, U_{Nj}) = 0, \\ U_{N, N_l} &= U_{N, 0}, \quad U_{N, -1} = U_{N, N_l-1}, \end{aligned} \quad (6.26)$$

где

$$\begin{aligned} \delta_r^2 U_{Nj} &:= \eta_{Nj}^{-1} \frac{\eta_{Nj} \phi_j - \tilde{\eta}_{Nj} D_r^- U_{Nj}}{h_N/2} = \frac{2}{h_N} \phi_j - \eta_{Nj}^{-1} \frac{2}{h_N} \tilde{\eta}_{Nj} D_r^- U_{Nj}, \\ h_N &:= r_N - r_{N-1}, \quad \phi_j := \phi(x_{Nj}). \end{aligned} \quad (6.27)$$

Обозначение F_-^h отражает тот факт, что здесь мы описываем вклад в аппроксимацию на $\partial\Omega_\sigma$ лишь из окрестности этой кривой, соответствующей

$r = \sigma^-$. Отметим также, что F_-^h включает неизвестную вспомогательную функцию ϕ . Окончательная аппроксимация будет получена после того, как мы также аппроксимируем уравнение в окрестности $\partial\Omega_\sigma$, соответствующей $r = \sigma^-$, и объединим две полученные аппроксимации, исключив ϕ (см. §6.3.6).

Следующая лемма представляет вспомогательный результат, который будет использован ниже при оценке погрешности аппроксимации в Ω_σ в окрестности интерфейсной границы $\partial\Omega_\sigma$.

Л е м м а 6.20. Пусть ограниченная область $\tilde{\Omega}_\sigma \supset \Omega_\sigma$ ограничена замкнутой кривой $r = \sigma - h_N$, где σ и $\{r_i\}_{i=0}^N$ выбраны в соответствии с §6.3.1(a) или §6.3.1(b). Тогда для $\beta(x; p)$ из (6.14) справедливо $\|\beta\|_{C^2(\tilde{\Omega}_\sigma)} \leq C(1 + h^2/\varepsilon^2)$.

Д о к а з а т е л ь с т в о. Для u_0 из (6.2) справедливо $\|u_0\|_{C^2(\Omega)} \leq C$. Производные v и v_1 (см. (6.12), (6.9)) по переменной l оцениваются аналогично. Желаемые оценки для производных v и v_1 по переменной r доказываются с использованием оценок (6.17) при $r = \varepsilon\xi$, а также с учетом Замечания 6.5 и значений, принимаемых параметрами σ и h_N в случае сеток (a) и (b). В частности, используется (a) $\sigma - h_N = 2\gamma^{-1}\varepsilon \ln[1/O(h)]$ и (b) $h_N = O(\varepsilon h |\ln h|)$. Остается проверить, что $\varepsilon^{-2}e^{-\gamma(\sigma-h_N)/\varepsilon}$ ограничено величиной Ch^2/ε^2 . \square

Л е м м а 6.21. Пусть $\beta(x; p)$ определяется (6.14), а сетка $\{r_i\}_{i=0}^N$ является либо сеткой Бахвалова §6.3.1(a), либо сеткой Шихкина §6.3.1(b). Тогда при любом $|p| \leq p_0$ во всех узлах сетки на интерфейсной границе $x_{Nj} \in \partial\Omega_\sigma$ справедливо

$$(F_-^h\beta - F\beta)\Big|_{x_{Nj} \in \partial\Omega_\sigma} = \frac{2\varepsilon^2}{h_N} \left(\frac{\partial\beta}{\partial r} - \phi \right)\Big|_{x_{Nj}} + O(h^2). \quad (6.28)$$

Д о к а з а т е л ь с т в о. Прежде всего, заметим, что

$$F_-^h\beta_{Nj} - F\beta(x_{Nj}) = -\varepsilon^2 \delta_r^2 \beta_{Nj} + O(\varepsilon^2) \|\beta\|_{C^2(\tilde{\Omega}_\sigma)}.$$

Далее,

$$\delta_r^2 \beta_{Nj} = \frac{\phi_j - \frac{\partial\beta}{\partial r}\Big|_{x_{Nj}}}{h_N/2} + \eta_{Nj}^{-1} \frac{\eta_{Nj} \frac{\partial\beta}{\partial r}\Big|_{x_{Nj}} - \tilde{\eta}_{Nj} D_r^- \beta_{Nj}}{h_N/2},$$

где второе слагаемое в правой части ограничено величиной $C\|\beta\|_{C^2(\bar{\Omega}_\sigma)}$. Объединяя эти соображения с Леммой 6.20 и (6.4), завершаем доказательство.

□

6.3.5 Аппроксимация во внутренней подобласти

Уравнение (6.1a) в Ω_σ , объединенное с фиктивным граничным условием (6.25), может быть описано с помощью следующей стандартной вариационной формулировки: найти функцию $u \in W_2^1(\Omega_\sigma)$ такую, что $\forall v \in W_2^1(\Omega_\sigma)$ справедливо

$$\varepsilon^2(\nabla u, \nabla v) + \varepsilon^2 \oint_{\partial\Omega_\sigma} \phi v ds + (b(x, u), v) = 0,$$

где обозначение (\cdot, \cdot) используется для скалярного произведения в пространстве $L_2(\Omega_\sigma)$. Здесь мы также воспользовались (6.25), которое удобно представить в виде $\frac{\partial u}{\partial n} = -\phi$, поскольку внешняя нормаль на границе Ω_σ имеет направление, противоположное положительному r -направлению.

Введем стандартное конечноэлементное пространство $S^h \subset W_2^1(\Omega_\sigma^h)$, состоящее из непрерывных функций, являющихся линейными на каждом треугольном элементе сетки в Ω_σ^h . Далее, пусть численное решение $U \in S^h$ во внутренней подобласти удовлетворяет соотношению

$$\varepsilon^2(\nabla U, \nabla \chi_i) + \varepsilon^2 \phi_i \oint_{\partial\Omega_\sigma^h} \chi_i ds + b(X_i, U_i)(1, \chi_i) = 0 \quad \forall \chi_i \in S^h, \quad (6.29)$$

где X_i — узел сетки в $\bar{\Omega}_\sigma^h$, $U_i = U(X_i)$, $\phi_i = \phi(X_i)$, а $\chi_i \in S^h$ являются соответствующей X_i стандартной базисной функцией, т. е. $\chi_i(X_j)$ равно 1 при $i = j$ и 0 в противном случае. Отметим, что здесь используется конечноэлементная аппроксимация с сосредоточенными массами для интегралов по границе и для интегралов с b (см. Замечания 6.16 и 6.17).

Начнем с рассмотрения (6.29) во внутренних узлах сетки X_i в Ω_σ . Получающиеся разностные уравнения будут нормализованы аналогично опреде-

ленному ранее конечно-разностному оператору F^h следующим образом:

$$F^h U_i := \frac{\varepsilon^2}{(1, \chi_i)} (\nabla U, \nabla \chi_i) + b(X_i, U_i) = 0 \quad \forall X_i \in \Omega_\sigma. \quad (6.30)$$

З а м е ч а н и е 6.22. Поскольку используется триангуляция Делоне, первое слагаемое в (6.30) соответствует M -матрице. Поэтому аппроксимация (6.30) во внутренних сеточных узлах в Ω_σ принадлежит классу (6.23).

Л е м м а 6.23. Пусть функция $\beta = \beta(x; p)$ определяется (6.14), а $\beta^I \in S^h$ является ее кусочно-линейным интерполянтom, удовлетворяющим $\beta^I(X_i) = \beta(X_i)$ во всех узлах сетки $X_i \in \bar{\Omega}_\sigma$. Также, пусть параметр σ выбран в соответствии с §6.3.1(a) или §6.3.1(b). Тогда для всех $|p| \leq p_0$ справедливо

$$|F^h \beta_i^I - F \beta_i| \leq Ch^2 \quad \forall X_i \in \Omega_\sigma.$$

Д о к а з а т е л ь с т в о. Нашей целью является получение оценки для

$$F^h \beta_i^I - F \beta|_{X_i} = \frac{\varepsilon^2}{(1, \chi_i)} (\nabla \beta^I, \nabla \chi_i) + \varepsilon^2 \Delta \beta|_{X_i}.$$

Поскольку X_i является внутренним узлом в Ω_σ , то $\chi_i = 0$ на $\partial\Omega_\sigma^h$, и следовательно

$$(\nabla \beta^I, \nabla \chi_i) = (\nabla[\beta^I - \beta], \nabla \chi_i) - (\Delta \beta, \chi_i). \quad (6.31)$$

Объединяя это с оценкой ошибки интерполяции $|\nabla[\beta^I - \beta]| \leq Ch \|\beta\|_{C^2(\bar{\Omega}_\sigma)}$ и стандартными свойствами квазиравномерных сеток $(1, |\nabla \chi_i|) \leq Ch$ и $(1, \chi_i) \geq Ch^2$, получаем

$$|F^h \beta_i^I - F \beta_i| \leq C \varepsilon^2 \|\beta\|_{C^2(\bar{\Omega}_\sigma)}.$$

Чтобы завершить доказательство, остается воспользоваться Леммой 6.20 и (6.4).

□

6.3.6 Аппроксимация на интерфейсной границе $\partial\Omega_\sigma$

Далее рассмотрим (6.29) в узлах сетки X_j на интерфейсной границе $\partial\Omega_\sigma$ и определим

$$F_+^h U_j := \frac{\varepsilon^2}{(1, \chi_j)} (\nabla U, \nabla \chi_j) + \frac{\varepsilon^2 a_j}{h} \phi_j + b(X_j, U_j) = 0 \quad \forall X_j \in \partial\Omega_\sigma, \quad (6.32)$$

где

$$a_j := \frac{h}{(1, \chi_j)} \oint_{\partial\Omega_\sigma^h} \chi_j ds, \quad 0 < C^{-1} < a_j < C. \quad (6.33)$$

Второе соотношение в (6.33) справедливо при любом j поскольку триангуляция в Ω_σ является квазиравномерной.

Лемма 6.24. При выполнении условий Леммы 6.23, для любого $|p| \leq p_0$ справедливо соотношение

$$F_+^h \beta_j^I - F \beta_j = -a_j \frac{\varepsilon^2}{h} \left(\frac{\partial \beta}{\partial r} - \phi \right) \Big|_{X_j} + O(h^2) \quad \forall X_j \in \partial\Omega_\sigma. \quad (6.34)$$

Доказательство. Нашей целью является получение оценки для

$$F_+^h \beta_j^I - F \beta \Big|_{X_j} = \frac{\varepsilon^2}{(1, \chi_j)} (\nabla \beta^I, \nabla \chi_j) + \varepsilon^2 \frac{a_j}{h} \phi_j + \varepsilon^2 \Delta \beta \Big|_{X_j}.$$

Для этого мы имитируем доказательство Леммы 6.23, с той лишь разницей, что теперь χ_j не исчезает на $\partial\Omega_\sigma^h$. В результате появляются дополнительные слагаемые, содержащие интегралы по $\partial\Omega_\sigma^h$. Поэтому вместо (6.31) получаем

$$(\nabla \beta^I, \nabla \chi_j) = (\nabla [\beta^I - \beta], \nabla \chi_j) - (\Delta \beta, \chi_j) + \oint_{\partial\Omega_\sigma^h} \chi_j \frac{\partial \beta}{\partial n} ds.$$

Здесь $\frac{\partial \beta}{\partial n}$ является производной по направлению внешней нормали, которая вычисляется в $O(h)$ -окрестности точки X_j , и для которой поэтому справедливо

$$\frac{\partial \beta}{\partial n} = - \frac{\partial \beta}{\partial r} \Big|_{X_j} + O(h) \|\beta\|_{C^2(\bar{\Omega}_\sigma)}.$$

При получении последнего соотношения мы также воспользовались тем, что, в силу гладкости Ω_σ , угол между направлением $-r$ в X_j и нормалью n к $\partial\Omega_\sigma^h$ в $O(h)$ -окрестности точки X_j остается в пределах $O(h)$. Для завершения доказательства, остается воспользоваться соотношением $\oint_{\partial\Omega_\sigma^h} \chi_j ds \leq Ch$, Леммой 6.20, и условием (6.4). \square

Наконец, объединяем F_-^h (6.26), (6.27) и F_+^h (6.32), (6.33), исключая вспомогательную неизвестную функцию ϕ , и получаем окончательную аппроксимацию на интерфейсной границе $\partial\Omega_\sigma$:

$$F^h U_j := \frac{(h_N/2) F_-^h U_j + (h/a_j) F_+^h U_j}{h_N/2 + h/a_j} \quad \forall X_j \in \partial\Omega_\sigma. \quad (6.35)$$

З а м е ч а н и е 6.25. В силу (6.26), (6.27), (6.32), (6.33), аппроксимация (6.35) принадлежит классу (6.23).

Л е м м а 6.26. При выполнении условий Леммы 6.21, для F^h из (6.35) справедливо соотношение

$$|F^h \beta_j - F \beta_j| \leq Ch^2 \quad \forall X_j \in \partial\Omega_\sigma.$$

Д о к а з а т е л ь с т в о. Сложим (6.28), умноженное на $(h_N/2)$, и (6.34), умноженное на (h/a_j) , и разделим сумму на $(h_N/2 + h/a_j)$. \square

6.3.7 Существование и точность численных решений, дискретные верхние и нижние решения

Л е м м а 6.27. Положим $\bar{p} := C_2 h^2 |\ln h|^m$, где $m = 0$ для сетки Бахвалова из §6.3.1(a) и $m = 2$ для сетки Шишкина из §6.3.1(b), а $C_2 > 0$ является достаточно большой постоянной. Пусть $\beta(x; p)$ задается соотношением (6.14), а $h \leq C_3$ при некоторой достаточно малой $C_3 > 0$. Тогда $\beta(X_i; -\bar{p})$ и $\beta(X_i; \bar{p})$ соответственно являются дискретными нижним и верхним решениями (см. Замечание 6.15).

Д о к а з а т е л ь с т в о. Объединяя результаты Лемм 6.19, 6.23 и 6.26, и выбирая постоянную C_2 достаточно большой, получаем $|F^h\beta - F\beta| \leq Ch^2 |\ln h|^m \leq \bar{p}/2$ для любого $|p| \leq p_0$. Далее, в силу Следствия 6.9 и (6.4), при достаточно большой постоянной C_2 и достаточно малом h справедливы соотношения $F\beta(x; -\bar{p}) \leq -\bar{p}/2$ и $F\beta(x; \bar{p}) \geq \bar{p}/2$. Следовательно $F^h\beta(X_i; -\bar{p}) \leq 0$ и $F^h\beta(X_i; \bar{p}) \geq 0$. В силу Следствия 6.7, лемма доказана. \square

Следующая теорема устанавливает существование и ε -равномерную точность численных решений нашей дискретной задачи.

Т е о р е м а 6.28. Пусть $\{r_i\}_{i=0}^N$ является либо сеткой Бахвалова из §6.3.1(a), либо сеткой Шишкина из §6.3.1(b). Существует дискретное решение U задачи (6.24), (6.30), (6.35) такое, что при достаточно малых значениях h справедлива оценка

$$|U(X_i) - u(X_i)| \leq Ch^2 |\ln h|^m \quad \text{для любого узла сетки } X_i \in \bar{\Omega}, \quad (6.36)$$

где $m = 0$ для сетки Бахвалова (a) и $m = 2$ для сетки Шишкина (b).

Д о к а з а т е л ь с т в о. Дискретный оператор F^h является Z -полем в силу Замечаний 6.16, 6.18, 6.22 и 6.25. В силу Леммы 6.27, заключаем, что $\beta(X_i; -\bar{p})$ и $\beta(X_i; \bar{p})$ являются соответственно дискретными нижним и верхним решениями. Поэтому, в силу Леммы 6.14, существует дискретное решение U такое, что $\beta(X_i; -\bar{p}) \leq U(X_i) \leq \beta(X_i; \bar{p})$. Следовательно, в силу Следствия 6.7, получаем $U(X_i) = u(X_i) + O(\bar{p} + \varepsilon^2)$. Наконец, применение (6.4) завершает доказательство. \square

З а м е ч а н и е 6.29. Теорема 6.28 получена при выполнении условия (6.4), т. е. при $\varepsilon \leq Ch$. Если же $\varepsilon > Ch$, из оценок для метода конечных элементов [133], можно ожидать несколько более общую версию оценки (6.36):

$$|u - U| \leq C[h^2 \ln(1 + \varepsilon/h) + h^2 |\ln h|^m].$$

Таблица 6.1: Сетка Бахвалова: вычислительные порядки сходимости r в $(N^{-1})^r$ и ошибки решения в сеточной норме максимума модуля

N	$\varepsilon = 10^{-2}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-8}$
32	2.010	2.011	2.011
64	1.995	1.997	1.997
128	1.995	2.001	2.001
32	3.745e-3	3.842e-3	3.843e-3
64	9.296e-4	9.534e-4	9.536e-4
128	2.333e-4	2.388e-4	2.388e-4
256	5.854e-5	5.967e-5	5.968e-5

Последняя оценка отличается от (6.36) дополнительным слагаемым $h^2 \ln(1 + \varepsilon/h)$, которое присутствует в локальной оценке в норме максимума модуля [133, (1.8)]. Можно ожидать, что данная локальная оценка будет справедлива во внутренней подобласти, где используется квазиравномерная триангуляция Делоне, с возможной дополнительной ошибкой $O(h^2 |\ln h|^m)$ при $m = 2$, появляющейся в силу недостаточно оптимального сгущения сетки в окрестности погранслоя. Отметим, что $h^2 \ln(1 + \varepsilon/h)$ становится $O(h^2)$ при $\varepsilon \leq Ch$, а при ε , постепенно возрастающем к $O(1)$, приближается к классической конечноэлементной оценке ошибки в норме максимума модуля $O(h^2 |\ln h|)$.

6.4 Численные результаты

В качестве теста, рассмотрим задачу типа (6.1) в области Ω (см. также Рис. 6.2), граница которой $\partial\Omega$ задается соотношениями $x_1 = \varphi(l) := R \cos \theta$ и $x_2 = \psi(l) := R \sin \theta$, где $l \in [0, 2\pi]$,

$$R = R(l) = 0.4 + 0.5 \cos^2 l, \quad \theta = \theta(l) = l + e^{(l-5)/2} \sin(l/2) \sin l.$$

Таблица 6.2: Сетка Шишкина: вычислительные порядки сходимости r в $(N^{-1} \ln N)^r$ и ошибки решения в сеточной норме максимума модуля

N	$\varepsilon = 10^{-2}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-8}$
32	2.131	2.137	2.137
64	2.210	2.045	2.045
128	2.465	2.009	2.009
32	3.915e-2	3.947e-2	3.948e-2
64	1.318e-2	1.325e-2	1.325e-2
128	4.004e-3	4.400e-3	4.401e-3
256	1.008e-3	1.430e-3	1.430e-3

В уравнении (6.1) будем использовать

$$b(x, u) = (u - \bar{u}_0(x))u(u + \bar{u}_0(x)), \quad \bar{u}_0(x) = x_1^2 + x_1 + 1. \quad (6.37)$$

При этом $\pm \bar{u}_0(x)$ являются двумя устойчивыми (а 0 — неустойчивым) решениями соответствующей вырожденной задачи. Граничное условие $g(x) = (x_1 - x_1^2)/3$ удовлетворяет А2 для обоих решений $\pm \bar{u}_0$ (см. Рис. 6.1). Здесь будут приведены численные результаты лишь для точного решения u , которое близко к \bar{u}_0 во внутренней части области (см. Рис. 6.1, слева). Отметим, что численные результаты для решения близкого к $-\bar{u}_0$ во внутренней части области (см. Рис. 6.1, справа) принципиально не отличаются.

При решении нелинейной дискретной задачи использовался метод Ньютона. В таблицах 6.1 и 6.2 представлены численные результаты соответственно для сеток Бахвалова и Шишкина при $\gamma := 0.8\gamma_0$, где $\gamma_0 = 3\sqrt{2}/4$. В таблицах представлены вычислительные порядки сходимости, а также ошибки численного решения в сеточной норме максимума, полученные аналогично [182, §4].

Заметим, что представленные численные результаты согласуются с тео-

ретическими оценками Теоремы 6.28.

6.5 Обобщения результатов параграфа

Как уже отмечалось, в силу немонотонности исходного уравнения в сочетании с наличием малого параметра при старших производных, стандартный технический аппарат, используемый при получении априорных оценок ошибки как для конечно-разностных, так и для конечноэлементных численных решений, становится неприменимым. В настоящем параграфе, при доказательстве существования и точности численных решений был использован принципиально иной подход, основанный на построении дискретных верхних и нижних решений, при этом последние являются нелинейными модификациями асимптотического разложения для исходной задачи. Данный подход ранее использовался в работах [70, 143, 182] для обыкновенных дифференциальных уравнений, в то время как основной материал параграфа опубликован в [163], где описанный подход был впервые обобщен на случай эллиптических уравнений.

Отметим, что в более поздних работах автора аналогичный подход был обобщен на случай параболических уравнений [181], а также применялся при исследовании численных решений в случае внутренних переходных слоев [183] и в случае применения метода декомпозиции Шварца [179, 180]. Также, результаты параграфа обобщены автором в работе [166] для случая, когда явное параметрическое описание границы области неизвестно.

§7 Немонотонное полулинейное сингулярно возмущенное уравнение реакции-диффузии в многоугольной области

В выпуклой многоугольной области рассматривается полулинейное уравнение реакции-диффузии $-\varepsilon^2 \Delta u + b(x, u) = 0$ с граничными условиями Дирихле. Параметр ε предполагается сколь угодно малым, а вырожденное уравнение $b(x, u_0(x)) = 0$ может иметь несколько решений. Построено асимптотическое разложение для решения u , в котором будут использованы погранслойные и угловые функции. С помощью нелинейной модификации этого асимптотического разложения, построены также верхние и нижние решения для исходной задачи, и, таким образом, установлено существование решения исходной задачи u в окрестности асимптотического разложения. Наличие угловых точек границы серьезно усложняет анализ, поскольку приводит к исследованию нелинейного автономного эллиптического уравнения $-\Delta z + f(z) = 0$ в бесконечном секторе с немонотонной функцией f . Для данного уравнения также будет установлена корректность соответствующей линеаризованной задачи.

Наш интерес к асимптотическому анализу задачи (7.1) обусловлен тем, что построенное асимптотическое разложение и соответствующие верхние и нижние решения могут быть использованы при построении и теоретическом исследовании ε -равномерных численных методов для рассматриваемой задачи (аналогично §6).

Основной материал параграфа опубликован в работах [160, 161].

7.1 Введение

Рассмотрим краевую задачу для полулинейного сингулярно возмущенного уравнения реакции-диффузии

$$Fu \equiv -\varepsilon^2 \Delta u + b(x, u) = 0, \quad x = (x_1, x_2) \in \Omega \subset \mathbb{R}^2, \quad (7.1a)$$

$$u(x) = g(x), \quad x \in \partial\Omega. \quad (7.1b)$$

Здесь Ω является выпуклой многоугольной областью, положительный параметр ε принимает сколь угодно малые значения, а $\Delta = \partial^2/\partial x_1^2 + \partial^2/\partial x_2^2$ — обычный оператор Лапласа.

Формально полагая $\varepsilon = 0$ в (7.1a), получаем так называемую вырожденную задачу для (7.1):

$$b(x, u_0(x)) = 0 \quad \text{при } x \in \bar{\Omega}. \quad (7.2)$$

Предполагается, что (7.2) имеет достаточно гладкое решение u_0 , которое является устойчивым (точное определение будет дано ниже). Отметим, что при наших предположениях о b уравнение (7.2), а следовательно и задача (7.1), может иметь несколько решений. Вообще говоря, $u_0 \neq g$ на $\partial\Omega$, поэтому решения исходной задачи могут иметь пограничные слои в окрестности $\partial\Omega$. Задачи типа (7.1) в одномерном случае были рассмотрены в [97], а в случае гладкой двумерной области Ω — в [88, 163, 38]. В указанных работах для достаточно малых значений ε установлено, что существует решение (7.1), которое близко к u_0 во внутренней части Ω . Что касается ε -равномерных численных методов для (7.1), одномерному случаю посвящены работы [70, 143, 182], а случай гладкой двумерной области рассмотрен в [163] (более подробное обсуждение приводится в §6).

В настоящем параграфе задача (7.1) рассматривается в выпуклой многоугольной области. Наличие угловых точек на границе $\partial\Omega$ приводит к дополнительным сложностям в асимптотическом анализе. В дополнение к достаточно стандартным погранслойным функциям, в асимптотическом разложении появляются угловые пограничные функции. Эти угловые функции

определяются как решения некоторых нелинейных краевых задач в бесконечном выпуклом секторе, что серьезно усложняет анализ, поскольку даже существование решений таких задач далеко неочевидно.

Отметим, что наш интерес к асимптотическому анализу задачи (7.1) обусловлен тем, что построенное асимптотическое разложение и соответствующие верхние и нижние решения могут быть использованы при построении и теоретическом исследовании ε -равномерных численных методов для рассматриваемой задачи (аналогично §6; см. также §7.5, где кратко обсуждаются некоторые аспекты численного решения (7.1)).

Для угловых точек области Ω будут использоваться обозначения $\{P_j\}_{j=1}^M$, а для сторон $\{\Gamma_j\}_{j=1}^M$. При этом угловые точки пронумерованы против часовой стрелки, и вершина P_{j-1} находится на пересечении сторон Γ_{j-1} и Γ_j (при это считаем, что $\Gamma_{M+1} = \Gamma_1$). Предполагается, что функция b достаточно гладкая, а g является гладкой на каждой из сторон Γ_j и непрерывной в каждой вершине P_j . Будут также сделаны следующие предположения.

A1 (*Устойчивое вырожденное решение*) Существует устойчивое вырожденное решение, т. е. вырожденная задача (7.2) имеет достаточно гладкое решение u_0 , удовлетворяющее при некоторой постоянной $\gamma > 0$ условию

$$b_u(x, u_0(x)) > \gamma^2 > 0 \quad \text{при любом } x \in \Omega.$$

A2 (*Граничные условия*) Функция $g(x)$, используемая в граничном условии (7.1b), удовлетворяет соотношению

$$\int_{u_0(x)}^v b(x, s) ds > 0 \quad \text{для любого } v \in (u_0(x), g(x)]', \quad x \in \partial\Omega.$$

Используемое здесь обозначение $(a, b]'$ совпадает с $(a, b]$ при $a < b$ и $[b, a)$ при $a > b$, в то время как $(a, b]' = \emptyset$ при $a = b$.

A3 (*Угловые условия*) Для каждой угловой точки P_j , при $g(P_j) \neq u_0(P_j)$ предполагается, что

$$\frac{b(P_j, g(P_j))}{g(P_j) - u_0(P_j)} > 0.$$

A4 Лишь с целью упрощения изложения, также предполагается, что

$$u_0(x) < g(x) \quad \text{при любом } x \in \partial\Omega.$$

Заметим, что с использованием A4, условие A3 может быть переписано в упрощенной форме $b(P_j, g(P_j)) > 0$.

Отметим, что если $g(x) \approx u_0(x)$, то A2 следует из A1 и (7.2), в то время как если $g(x) = u_0(x)$ в некоторой точке $x \in \partial\Omega$, то A2 не приводит к дополнительным ограничениям на g в этой точке. Аналогично, если $g(P_j) \approx u_0(P_j)$, то A3 следует из A1 и (7.2), в то время как если $g(P_j) = u_0(P_j)$ в некоторой угловой точке P_j , то A3 не приводит к дополнительным ограничениям на g в этой точке.

Предположение A1 является локальным, и позволяет использовать другие устойчивые решения вырожденной задачи (7.2) при доказательстве существования других решений исходной задачи (7.1). Предположение A2 является стандартным при рассмотрении гладких областей [88, 163, 38]; оно обеспечивает существование погранслойных ингредиентов в асимптотическом разложении.

Обсудим предположение A3, которое обусловлено наличием угловых точек границы $\partial\Omega$. Ключевым элементом анализа, изложению которого посвящен данный параграф, будет исследование некоторых решений полулинейного эллиптического уравнения

$$-\Delta z + f(z) = 0 \tag{7.3}$$

в бесконечном секторе. Уравнение типа (7.3) представляет интерес, поскольку угловые погранфункции, ассоциируемые с вершиной P_j могут быть описаны с использованием некоторого решения уравнения (7.3) при $f(z) = b(P_j, z)$ с условием $z = g(P_j)$ на границе сектора (ниже см. описание задачи (7.12)). Как будет показано, предположение A3 не только является достаточным для

существования решения z . Результат из [144] показывает, что АЗ также является необходимым для существования решения z , если мы хотим исключить появление острых пиков в решении u задачи (7.1) в весьма малых окрестностях углов Ω (см. Замечание 7.7). Более того, с использованием АЗ будет установлена устойчивость решений (7.3), понимаемая в том смысле, что главное собственное значение оператора, возникающего при линеаризации (7.3) около решения z и рассматриваемого в соответствующем секторе радиуса R , является положительным и отделено от нуля положительной постоянной равномерно по $R \rightarrow \infty$ (см. §7.3.4). Раздел, посвященный задаче (7.3) и соответствующему линеаризованному оператору, является ключевым в настоящем параграфе, и может также представлять независимый интерес.

Основным результатом параграфа является построение асимптотического разложения первого порядка u_{as} для задачи (7.1) и доказательство существования решения этой задачи $u(x)$, для которого справедливо $|u - u_{\text{as}}| \leq C\varepsilon^2$. Также будут установлены поточечные оценки производных для компонент построенного асимптотического разложения. Последние могут быть полезными при построении и теоретическом исследовании сходимости численных методов для рассматриваемой задачи. Аналогичные результаты для гладких областей были получены Файфом [88], а также Нефедовым [38].

Результаты данного параграфа являются принципиально новыми для многоугольных областей. Отметим, что в случае прямоугольной области результаты, аналогичные нашим, были получены в работе [15] для линейного варианта задачи (10). Нелинейная же задача (10) в прямоугольной области рассматривалась в работах [23, 24, 25, 26, 27], но при весьма более сильных, по сравнению с АЗ, ограничениях в угловых точках области. (Например, в [25] предполагается, что $b(P_j, \cdot)$ монотонна между $g(P_j)$ и $u_0(P_j)$, а в [23, 26] $b(P_j, \cdot)$ предполагается квадратичной.) Отметим также, что результаты параграфа могут быть обобщены на случай кусочно-гладких выпуклых областей, а также при построении асимптотических разложений более высокого порядка.

Следуя работам [97, 38], для доказательства существования решения будет использована теория верхних и нижних решений. При этом последние будут построены с помощью нелинейной модификации асимптотического разложения.

Параграф организован следующим образом. В §7.2 будут определены погранслойные функции, ассоциируемые со сторонами многоугольной области Ω , а также угловые функции, ассоциируемые с угловыми точками Ω . Погранслойные функции определяются как решения некоторых обыкновенных дифференциальных уравнений в растянутой переменной, направленной по нормали к сторонам области. Угловые функции будут определены как решения некоторых эллиптических уравнений в растянутых переменных (x, y) . Существование и свойства угловых функций будут установлены в §7.3. Данный раздел представляет ключевой материал параграфа. Далее в §7.4 построенные погранслойные и угловые функции будут использованы при построении верхних и нижних решений исходной задачи. Последние же потребуются при доказательстве существования и некоторых свойств соответствующего решения (7.1). Некоторые аспекты построения ε -равномерных численных методов для (7.1) кратко обсуждаются в §7.5. Отметим также, что некоторые технические вычисления и доказательства помещены в [160].

Обозначения. В настоящем параграфе C , \bar{C} , c , c' обозначают положительные постоянные, которые могут принимать разные значения в разных формулах, но являются независимыми от ε (при этом \bar{C} обычно используется для достаточно больших постоянных). Обозначение C с индексом (например, C_1) будет использоваться для положительной постоянной, которая не зависит от ε и принимает некоторое фиксированное значение. Для любых двух величин w_1 и w_2 , будем писать $w_1 = O(w_2)$ при выполнении $|w_1| \leq C|w_2|$.

7.2 Пограничные и угловые функции

Приступим к описанию пограничных функций, ассоциируемых со сторонами рассматриваемой многоугольной области Ω , а также угловых функций, ассоциируемых с вершинами Ω . Существование и свойства угловых функции будут установлены в §7.3.

Удобно ввести вспомогательные функции

$$B(x, t) = b(x, u_0(x) + t), \quad \tilde{B}(x, t; p) = b(x, u_0(x) + t) - pt. \quad (7.4)$$

Возмущенная версия \tilde{B} функции B будет использоваться, при достаточно малых значениях $|p|$, при построении верхних и нижних решений для исходной задачи. При этом волнистая черта показывает, что используется возмущенная функция, которая зависит от параметра p , даже если эта зависимость не всегда будет показана явно (например, иногда мы будем использовать $\tilde{B}(x, t)$ для $\tilde{B}(x, t; p)$). Также нам потребуются производные \tilde{B} . Для производных по x будем писать $\nabla_x \tilde{B}$, $\nabla_x^2 \tilde{B}$ и т. д. соответственно для обозначения вектора первых производных, матрицы вторых производных и т. д. Для производных же по t будем писать \tilde{B}_t , \tilde{B}_{tt} и т. д. Заметим, что $\tilde{B}(x, 0) = 0$, поэтому $\nabla_x^k \tilde{B}(x, 0) = 0$ при $k = 1, 2, \dots$, и следовательно

$$|\nabla_x^k \tilde{B}(x, t)| \leq C|t| \quad \text{при } k = 0, 1, 2, \dots \quad (7.5)$$

Для произвольной функции f будут использоваться обозначения

$$f|_a^b = f(b) - f(a), \quad f|_{a;b}^c = f(c) - f(b) - f(a). \quad (7.6)$$

Поскольку $f|_{a;b}^{a+b} + f(0) = abf''(t)$, в случае $f(0) = 0$ получаем $f|_{a;b}^{a+b} = O(|ab|)$, и следовательно $f|_{a;b}^{a+b+c} = O(|c| + |ab|)$. Таким образом, в силу (7.5), получаем

$$\nabla_x^k \tilde{B}(x, \cdot) \Big|_{a;b}^{c+a+b} = O(|c| + |ab|). \quad (7.7)$$

7.2.1 Решение в окрестности стороны многоугольной области

Приступим к построению погранслойных функций, ассоциируемых со стороной Γ_j многоугольной области Ω . Пусть бесконечная линия Γ является продолжением Γ_j . Продолжим u_0 и b гладким образом, с сохранением обозначений u_0 и b , соответственно на \mathbb{R}^2 и $\mathbb{R}^2 \times \mathbb{R}$ таким образом, что (7.2) и A1 выполняются при любом $x \in \mathbb{R}^2$. Также продолжим $g|_{\Gamma_j}$ на Γ гладким образом и с выполнением A2 и A4 при любом $x \in \Gamma$.

Пусть \mathbf{e}_s обозначает единичный вектор вдоль Γ , ориентированный в направлении от P_{j-1} к P_j . А \mathbf{e}_r пусть обозначает единичный вектор перпендикулярный \mathbf{e}_s и ориентированный внутрь области Ω . При этом s будет обозначать (возможно отрицательную) длину вдоль Γ с $s = 0$ в P_{j-1} . Таким образом, при $x \in \mathbb{R}^2$ справедливо $x = P_{j-1} + s\mathbf{e}_s + r\mathbf{e}_r$. При этом точка $\bar{x} = P_{j-1} + s\mathbf{e}_s$ расположена на Γ и является ближайшей к x , а r является (возможно отрицательным) расстоянием от \bar{x} к x , с $r > 0$ при $x \in \Omega$ (\mathbf{e}_s , \mathbf{e}_r , x и \bar{x} показаны на Рис. 7.1).

Определим $\tilde{v}_0(\xi, s; p)$ как решение автономной нелинейной двухточечной краевой задачи

$$\begin{aligned} -\frac{\partial^2 \tilde{v}_0}{\partial \xi^2} + \tilde{B}(\bar{x}, \tilde{v}_0; p) &= 0, \\ \tilde{v}_0(0, s; p) &= g(\bar{x}) - u_0(\bar{x}), \quad \tilde{v}_0(\infty, s; p) = 0. \end{aligned} \tag{7.8}$$

Здесь используется растянутая переменная $\xi = r/\varepsilon$. Переменные же p и s в задаче (7.8) являются параметрами. Параметр p удовлетворяет $|p| < \gamma^2$ и предполагается близким к нулю. Иногда явная зависимость \tilde{v}_0 от p не будет показана, и мы будем писать $\tilde{v}_0(\xi, s)$, подразумевая $\tilde{v}_0(\xi, s; p)$.

Положим $v_0(\xi, s) = \tilde{v}_0(\xi, s; 0)$. Функция v_0 используется в асимптотическом разложении в окрестности стороны Γ_j , а также при определении $v_1(\xi, s)$. Последняя является решением следующей линейной двухточечной краевой

задачи

$$-\frac{\partial^2 v_1}{\partial \xi^2} + v_1 B_t(\bar{x}, v_0) = -\xi \mathbf{e}_r \cdot \nabla_x B(\bar{x}, v_0), \quad (7.9)$$

$$v_1(0, s) = v_1(\infty, s) = 0.$$

Отметим, что v_1 не зависит от p . Также введем

$$\begin{aligned} \overset{\circ}{v}_0(\xi; p) &= \tilde{v}_0(\xi, 0; p), & \overset{\circ}{v}_0(\xi) &= v_0(\xi, 0), & \overset{\circ}{v}_1(\xi) &= v_1(\xi, 0), \\ \tilde{v} &= \tilde{v}_0 + \varepsilon v_1, & v &= v_0 + \varepsilon v_1, & \overset{\circ}{v} &= \overset{\circ}{v}_0 + \varepsilon \overset{\circ}{v}_1, & \overset{\circ}{v} &= \overset{\circ}{v}_0 + \varepsilon \overset{\circ}{v}_1. \end{aligned} \quad (7.10)$$

В используемых здесь обозначениях маленький кружок над именем функции показывает, что при ее определении используется $s = 0$.

Что касается разрешимости и свойств задач (7.8) и (7.9), справедливо следующее утверждение.

Л е м м а 7.1. *Найдется такое значение $p_0 \in (0, \gamma^2)$, что при любом $|p| \leq p_0$ существуют функции \tilde{v}_0 и v_1 , удовлетворяющие (7.8), (7.9). При этом для функции $\tilde{v}_0 = \tilde{v}_0(\xi, s; p)$ справедливы соотношения*

$$\tilde{v}_0 \geq 0, \quad \frac{\partial \tilde{v}_0}{\partial p} \geq 0. \quad (7.11)$$

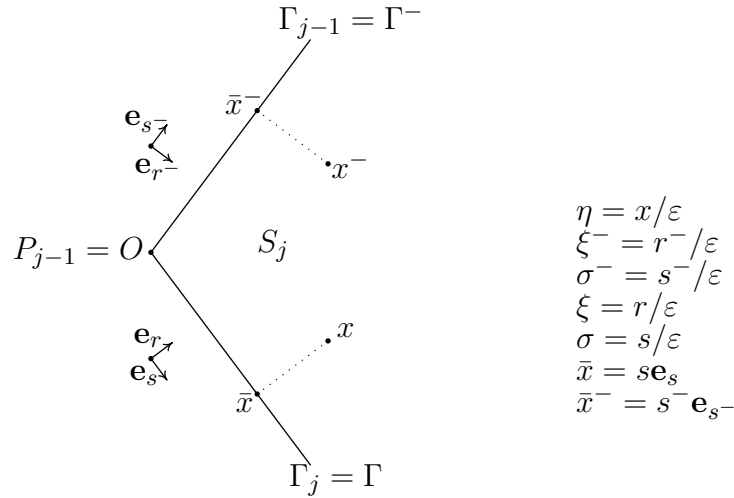
Также, при любом $k \geq 0$ и сколь угодно малом фиксированном δ , найдется такая постоянная $C > 0$, что при $0 \leq \xi < \infty$, $s \in \mathbb{R}$ и $k = 0, 1, \dots$, справедлива оценка

$$\left| \frac{\partial^k \tilde{v}_0}{\partial \xi^k} \right| + \left| \frac{\partial^k \tilde{v}_0}{\partial s^k} \right| + \left| \frac{\partial^k v_1}{\partial \xi^k} \right| + \left| \frac{\partial^k v_1}{\partial s^k} \right| + \left| \frac{\partial \tilde{v}_0}{\partial p} \right| + \left| \frac{\partial^2 \tilde{v}_0}{\partial p \partial s} \right| \leq C e^{-(\gamma - \sqrt{|p| - \delta})\xi}.$$

Д о к а з а т е л ь с т в о. Существование и свойства \tilde{v}_0 следуют из Лемм 2.2 и 2.3 из работы [182]. Для v_1 же можно воспользоваться Леммой 2.2 [88] (см. также [147, §2.3.1]). \square

7.2.2 Решение в окрестности вершины

Перейдем к построению угловых функций, которые будут ассоциироваться с некоторой вершиной P_{j-1} рассматриваемой многоугольной области. Эти

Рис. 7.1: Геометрия сектора $S_j = S$

угловые функции будут использоваться в асимптотическом разложении решения, а также при построении верхних и нижних решений.

Нам потребуются некоторые новые обозначения. Поместим вершину P_{j-1} в начало координат O . Пусть S_j , или просто S , обозначает бесконечный сектор, получаемый продолжением двух сторон Γ_j и Γ_{j-1} в направлении от O . При этом луч, продолжающий Γ_j , будет обозначаться Γ , а луч, продолжающий сторону Γ_{j-1} — соответственно Γ^- . Продолжим функцию $g|_{\Gamma_j}$ гладким образом на Γ с сохранением обозначения g . Аналогично, продолжим $g|_{\Gamma_{j-1}}$ на Γ^- и обозначим продолженную функцию g^- . Отметим, что данные продолжения функций выполняются с сохранением свойств А2 и А4. Пусть s обозначает дистанцию вдоль Γ от O , а r — расстояние от точки $x \in S$ до стороны Γ . Таким образом, $x \rightarrow (s, r)$ является линейным ортогональным отображением. При этом \mathbf{e}_r и \mathbf{e}_s будут обозначать единичные векторы, соответственно направленные вдоль и ортогонально к Γ таким образом, что $x = r\mathbf{e}_r + s\mathbf{e}_s$. Обозначение же $\bar{x} = s\mathbf{e}_s$ будет использоваться для ближайшей к x точке на Γ . Аналогичным образом определяются переменные (s^-, r^-) и \bar{x}^- , ассоциируемые со стороной Γ^- . Переменная s^- обозначает дистанцию

вдоль Γ^- , отмеряемую от O . Также будут использованы растянутые переменные $\eta = x/\varepsilon$, $\xi = r/\varepsilon$, $\sigma = s/\varepsilon$, $\xi^- = r^-/\varepsilon$, $\sigma^- = s^-/\varepsilon$. Описанные переменные также показаны на Рис. 7.1.

С использованием вновь введенных обозначений, заметим, что в § 7.2.1 были определены погранслойные функции $\tilde{v}_0(\xi, s; p)$ и $v_1(\xi, s)$, ассоциируемые со стороной Γ , а также $\tilde{v}_0^-(\xi^-, s^-; p)$ и $v_1^-(\xi^-, s^-)$, ассоциируемые со стороной Γ^- . Вспомним также обозначения из (7.10) и воспользуемся аналогичными обозначениями для Γ^- . Функция \tilde{v} компенсирует разницу между граничными условиями (7.1b) и значениями u_0 на Γ , но оставляет быстроубывающий след на части границы Γ^- . Аналогично, \tilde{v}^- компенсирует разницу между граничными условиями (7.1b) и значениями u_0 на Γ^- , но оставляет быстроубывающий след на части границы Γ . Для компенсации этих быстроубывающих следов на сторонах сектора S , будут построены угловые функции $\tilde{z}_0(\eta; p)$ и $z_1(\eta)$, определяемые с использованием растянутой переменной η .

Функция \tilde{z}_0 определяется как ограниченное решение автономной нелинейной эллиптической краевой задачи

$$\begin{aligned} -\Delta_\eta \tilde{z}_0 + \tilde{B}(O, \tilde{z}_0; p) &= 0 & \text{в } S, \\ \tilde{z}_0 &= A := g(O) - u_0(O) & \text{на } \partial S. \end{aligned} \quad (7.12)$$

Заметим, что используемое здесь значение $A > 0$, в силу предположения A4 в точке $P_{j-1} = O$. Также введем $z_0(\eta) = \tilde{z}_0(\eta; 0)$. Существование и свойства \tilde{z}_0 устанавливаются в следующей теореме, доказательству которой будет посвящен §7.3.

Т е о р е м а 7.2. *Найдется такая положительная постоянная p^* , что при любом $|p| \leq p^*$ задача (7.12) имеет решение \tilde{z}_0 , удовлетворяющее соотношениям $\tilde{z}_0 \leq A$ и*

$$0 < \max\{\overset{\circ}{v}_0, \overset{\circ}{v}_0^-\} \leq \tilde{z}_0(\eta; p) \leq \max\{\overset{\circ}{v}_0, \overset{\circ}{v}_0^-\} + C|\eta|^{-1}, \quad (7.13)$$

а также являющееся возрастающей функцией переменной p . При этом, $|\nabla \tilde{z}_0|$ ограничено в S . Наконец, найдется такая постоянная $C > 0$, что

справедлива оценка

$$\tilde{z}_0(\eta) \leq C \left(e^{-\gamma\xi} + e^{-\gamma\xi^-} \right). \quad (7.14)$$

Далее введем функцию $z_1(\eta)$, удовлетворяющую линейной эллиптической краевой задаче

$$\begin{aligned} -\Delta_\eta z_1 + z_1 B_t(O, z_0) &= -\eta \cdot \nabla_x B(O, z_0) \quad \text{в } S, \\ z_1 = \sigma \frac{\partial}{\partial s} (g - u_0) \Big|_{x=O} \quad \text{на } \Gamma, \quad z_1 &= \sigma^- \frac{\partial}{\partial s^-} (g^- - u_0) \Big|_{x=O} \quad \text{на } \Gamma^-. \end{aligned} \quad (7.15)$$

Функции \tilde{z}_0 и z_1 позволяют построить коррекцию $\tilde{z}_0 + \varepsilon z_1$ к решению вырожденной задачи u_0 в окрестности вершины O . За пределами же окрестности O используются корректирующие конструкции $\tilde{v}_0 + \varepsilon v_1$ и $\tilde{v}_0^- + \varepsilon v_1^-$ для u_0 соответственно вдоль сторон Γ and Γ^- . Заметим, что далее мы воспользуемся угловыми функциями \tilde{z}_0 и z_1 и погранслойными функциями $\tilde{v}_0, v_1, \tilde{v}_0^-, v_1^-$ при построении еще одной пары угловых функций \tilde{q}_0 и q_1 . Именно последние (а не \tilde{z}_0 и z_1) будут фигурировать в формальном асимптотическом разложении решения (7.1) (см. §7.4).

Чтобы описать \tilde{q}_0 и q_1 , выберем произвольную точку $\eta \in S$. Данной точке соответствуют $\xi, \sigma, \xi^-, \sigma^-$, определяемые по формулам

$$\eta = \xi \mathbf{e}_r + \sigma \mathbf{e}_s = \xi^- \mathbf{e}_{r^-} + \sigma^- \mathbf{e}_{s^-} \quad (7.16)$$

(см. Рис. 7.1). Теперь, с использованием \tilde{z}_0, z_1 и $\mathring{v}_0, \mathring{v}_0^-, \mathring{v}_1, \mathring{v}_1^-$ из (7.8), (7.9), (7.10), положим

$$\tilde{q}_0(\eta; p) = \tilde{z}_0(\eta; p) - \mathring{v}_0(\xi; p) - \mathring{v}_0^-(\xi^-; p), \quad (7.17a)$$

$$q_1(\eta) = z_1(\eta) - [\mathring{v}_1(\xi) + \sigma \mathring{v}_{0,s}(\xi)] - [\mathring{v}_1^-(\xi^-) + \sigma^- \mathring{v}_{0,s^-}(\xi^-)], \quad (7.17b)$$

а также

$$\tilde{q}(\eta; p) = \tilde{q}_0(\eta; p) + \varepsilon q_1(\eta), \quad q_0(\eta) = \tilde{q}_0(\eta; 0), \quad q(\eta) = q_0(\eta) + \varepsilon q_1(\eta). \quad (7.17c)$$

В этих формулах, с использованием обозначений (7.10), подразумевается, что

$$\mathring{v}_{0,s}(\xi) = \frac{\partial}{\partial s} v_0(\xi, s) \Big|_{s=0}, \quad \mathring{v}_{0,s^-}^- = \frac{\partial}{\partial s^-} v_0^-(\xi^-, s^-) \Big|_{s^-=0}. \quad (7.17d)$$

При этом краевые условия из (7.15) могут быть представлены в виде $z_1 = \sigma \dot{v}_{0,s}$ на Γ и $z_1 = \sigma^- \dot{v}_{0,s}^-$ на Γ^- .

Объединяя определение \tilde{q}_0 с (7.8), (7.12), получаем нелинейную краевую задачу, которой удовлетворяет \tilde{q}_0 :

$$\Delta_\eta \tilde{q}_0 = \tilde{B}(O, \tilde{q}_0 + \dot{v}_0 + \dot{v}_0^-) - \tilde{B}(O, \dot{v}_0) - \tilde{B}(O, \dot{v}_0^-), \quad (7.18a)$$

$$\tilde{q}_0 = -\dot{v}_0^- \quad \text{на } \Gamma, \quad \tilde{q}_0 = -\dot{v}_0 \quad \text{на } \Gamma^-. \quad (7.18b)$$

Аналогично, с использованием (7.8), (7.9) и (7.15) (см Лемму 2.4 в [160]) мы можем формально вывести линейную краевую задачу, которой удовлетворяет q_1 :

$$\begin{aligned} -\Delta_\eta q_1 + q_1 B_t(O, z_0) &= -\eta \cdot \nabla_x B(O, \cdot) \Big|_{\dot{v}_0; \dot{v}_0^-}^{z_0} \\ &\quad - (\dot{v}_1 + \sigma \dot{v}_{0,s}) B_t(O, \cdot) \Big|_{\dot{v}_0}^{z_0} - (\dot{v}_1^- + \sigma^- \dot{v}_{0,s}^-) B_t(O, \cdot) \Big|_{\dot{v}_0^-}^{z_0}, \end{aligned} \quad (7.19)$$

$$q_1 = -(\dot{v}_1^- + \sigma^- \dot{v}_{0,s}^-) \quad \text{на } \Gamma, \quad q_1 = -(\dot{v}_1 + \sigma \dot{v}_{0,s}) \quad \text{на } \Gamma^-,$$

где используются обозначения (7.6). Наконец, формально дифференцируя по p соотношение (7.17a) и задачу (7.12) (или эквивалентную ей задачу (7.18)) и затем воспользовавшись (7.4), мы получаем краевую задачу, которой формально удовлетворяет $\tilde{q}_{0,p}$:

$$\begin{aligned} -\Delta_\eta \tilde{q}_{0,p} + \tilde{q}_{0,p} \tilde{B}_t(O, \tilde{z}_0) &= \tilde{q}_0 - \dot{v}_{0,p} \tilde{B}_t(O, \cdot) \Big|_{\dot{v}_0}^{\tilde{z}_0} - \dot{v}_{0,p}^- \tilde{B}_t(O, \cdot) \Big|_{\dot{v}_0^-}^{\tilde{z}_0}, \\ \tilde{q}_{0,p} &= -\dot{v}_{0,p}^- \quad \text{на } \Gamma, \quad \tilde{q}_{0,p} = -\dot{v}_{0,p} \quad \text{на } \Gamma^-. \end{aligned} \quad (7.20)$$

Задача (7.18) будет использована в §7.3.3 при доказательстве экспоненциального убывания \tilde{q}_0 при $|\eta| \rightarrow \infty$. Также мы увидим, что все данные в линейных задачах (7.19) и (7.20) экспоненциально убывают. Эти наблюдения будут использованы в §7.3.5 при доказательстве корректной постановки этих линейных задач, а также того, что они имеют экспоненциально убывающие решения. В силу (7.17b) существование q_1 немедленно приводит к существованию z_1 . Аналогично, доказав существование решения (7.20), с помощью интегрирования можно убедиться в том, что данное решение является производной \tilde{q}_0 по p .

7.3 Существование и свойства угловых погранфункций

Перейдем к доказательству существования и некоторых свойств функций \tilde{z}_0 and z_1 . Существование решения (7.12) вытекает из теории верхних и нижних решений (см. §7.3.1). Последняя теория также применяется в §7.4 при доказательстве существования решения (7.1). Существование решения \tilde{z}_0 для (7.12) и его свойства, описанные в Теореме 7.2, будут доказаны в §7.3.2. Поскольку (7.12) может иметь несколько решений, сначала мы построим определенные нижнее и верхнее решения для (7.12), после чего функция \tilde{z}_0 будет выбрана как *единственное* минимальное решение между выбранными нижним и верхним решениями. В §7.3.3 будет рассмотрено экспоненциальное убывание компоненты \tilde{q}_0 в \tilde{z}_0 .

Прежде чем мы перейдем к доказательству свойств z_1 и $\partial\tilde{z}_0/\partial p$, потребуется рассмотреть линеаризацию (7.12) около функции \tilde{z}_0 (см. §7.3.4). Будет установлено, что главное собственное значение линеаризованной задачи строго положительно, и, следовательно, линеаризованный оператор обратим. Эта часть анализа является важнейшей в данном параграфе и может представлять независимый интерес. Существование и необходимые нам свойства z_1 и $\partial\tilde{z}_0/\partial p$ будут затем установлены в §7.3.5.

7.3.1 Верхние и нижние решения

Теория верхних и нижних решений описана, например, в [68, 85, 129]. Мы опишем определения и некоторые важные факты этой теории в контексте задачи (7.1), поставленной в многоугольной области Ω . Отметим, что эта теория применима и в более общих ситуациях, например, в случае задачи (7.1), поставленной в секторе S . Последнее позволит нам воспользоваться результатами, которые будут представлены ниже, при изучении задач (7.1) и (7.12) (последняя, очевидно, принадлежит классу (7.1) при $\varepsilon = 1$).

Функция β называется верхним решением для задачи (7.1), если β является непрерывной и ограниченной в $\bar{\Omega}$, при этом $\beta \geq g$ на $\partial\Omega$, а для каждого

$\chi \in C_0^\infty(\Omega)$, удовлетворяющего $\chi \geq 0$, справедливо соотношение

$$\iint_{\Omega} [-\beta \varepsilon^2 \Delta \chi + b(x, \beta) \chi] dx \geq 0. \quad (7.21)$$

Здесь $C_0^\infty(\Omega)$ обозначает множество бесконечно дифференцируемых функций с компактным носителем в Ω . Аналогично, α называется нижним решением для (7.1), если α является непрерывной и ограниченной в $\bar{\Omega}$, при этом $\alpha \leq g$ на $\partial\Omega$, а также справедливо обратное к (7.21) неравенство, с β замененной на α . Следующая лемма установлена в [68, 85, 129].

Л е м м а 7.3. Если β_1 и β_2 являются двумя верхними решениями для (7.1), которые принадлежат пространству $C^2(\Omega)$, то $\min\{\beta_1, \beta_2\}$ тоже является верхним решением (7.1). Если α_1 и α_2 являются двумя нижними решениями (7.1), которые принадлежат пространству $C^2(\Omega)$, то $\max\{\alpha_1, \alpha_2\}$ тоже является нижним решением (7.1).

Следующая лемма демонстрирует полезность построения верхних и нижних решений, заключающуюся в том, что они позволяют обосновывать существование решения соответствующей задачи. Этот результат приводится в [68] и [85] для гладких ограниченных областей, а также в [129] для более общих областей.

Л е м м а 7.4. Пусть α и β являются, соответственно, нижним и верхним решениями задачи (7.1), при этом $\alpha \leq \beta$ в Ω . Тогда (7.1) имеет решение u , для которого справедливо $\alpha \leq u \leq \beta$ в Ω . Также для (7.1) существует минимальное решение u_m в том смысле, что $\alpha \leq u_m \leq \beta$ в Ω , а при существовании другого решения u , также удовлетворяющего $\alpha \leq u \leq \beta$ в Ω , справедливо соотношение $u_m \leq u$ в Ω .

Также нам потребуется следующая лемма

Л е м м а 7.5. Пусть u_m является минимальным решением (7.1) в области Ω , соответствующим нижнему решению α , а $\hat{\Omega} \subset \Omega$. Пусть \hat{u}_m является минимальным решением задачи, состоящей из уравнения из (7.1),

но рассматриваемого в области $\hat{\Omega}$, с граничным условием $\hat{u}_m|_{\partial\hat{\Omega}} = u_m|_{\partial\hat{\Omega}}$, и соответствующим тому же нижнему решению α , теперь рассматриваемому в $\hat{\Omega}$. Тогда $\hat{u}_m = u_m$ в $\hat{\Omega}$.

Д о к а з а т е л ь с т в о. Поскольку $u_m \geq \alpha$ является решением задачи, которой удовлетворяет $\hat{u}_m \geq \alpha$, причем \hat{u}_m является минимальным решением этой задачи, то $\hat{u}_m \leq u_m$ в $\hat{\Omega}$. Введем функцию β в Ω , удовлетворяющую $\beta = \hat{u}_m$ в $\hat{\Omega}$ и $\beta = u_m$ в $\Omega \setminus \hat{\Omega}$. Мы утверждаем, что β является верхним решением (7.1). Действительно, пусть Γ обозначает часть $\partial\hat{\Omega}$, которая лежит внутри Ω , а n — единичную нормаль к $\partial\hat{\Omega}$, направленную из $\hat{\Omega}$. Если $\chi \in C_0^\infty(\Omega)$ и при этом $\chi \geq 0$, то, в силу соотношений $\varepsilon^2 \Delta \hat{u}_m = b(x, \hat{u}_m)$ в $\hat{\Omega}$ и $\varepsilon^2 \Delta u_m = b(x, u_m)$ в $\Omega \setminus \hat{\Omega}$, получаем

$$\begin{aligned} \iint_{\Omega} [-\beta \varepsilon^2 \Delta \chi + b(x, \beta) \chi] dx &= \iint_{\hat{\Omega}} [-\hat{u}_m \varepsilon^2 \Delta \chi + b(x, \hat{u}_m) \chi] dx \\ &+ \iint_{\Omega \setminus \hat{\Omega}} [-u_m \varepsilon^2 \Delta \chi + b(x, u_m) \chi] dx \\ &= \int_{\Gamma} \varepsilon^2 \chi \left(\frac{\partial \hat{u}_m}{\partial n_{\Gamma}} - \frac{\partial u_m}{\partial n_{\Gamma}} \right) \geq 0. \end{aligned}$$

Здесь последнее соотношение следует из $\hat{u}_m \leq u_m$ в $\hat{\Omega}$. Убедившись, что β является верхним решением, немедленно получаем $u_m \leq \beta$ в Ω . Следовательно $u_m \leq \hat{u}_m$ в $\hat{\Omega}$, поэтому $u_m = \hat{u}_m$ в $\hat{\Omega}$. \square

Л е м м а 7.6. Пусть u_m является минимальным решением (7.1), соответствующим нижнему решению α , а $\hat{\Omega} \subset \Omega$. Пусть \bar{u}_m является минимальным решением задачи, состоящей из уравнения из (7.1), но рассматриваемого в области $\hat{\Omega}$, с граничным условием $\bar{u}_m|_{\partial\hat{\Omega}} \geq u_m|_{\partial\hat{\Omega}}$, и соответствующим тому же нижнему решению α , теперь рассматриваемому в $\hat{\Omega}$. Тогда $\bar{u}_m \geq u_m$ в $\hat{\Omega}$.

Д о к а з а т е л ь с т в о. Очевидно, что для \hat{u}_m из Леммы 7.5 имеем $\hat{u}_m \leq \bar{u}_m$. Объединяя это с $\hat{u}_m = u_m$ в $\hat{\Omega}$, получаем утверждение Леммы. \square

7.3.2 Существование \tilde{z}_0

Перейдем к доказательству Теоремы 7.2. Пусть S является выпуклым сектором с вершиной O и границей $\Gamma \cup \Gamma^-$. Пусть также $z = \tilde{z}_0$ и $f(z) = \tilde{B}(O, z)$. Тогда интересующая нас краевая задача (7.12) может быть представлена, в терминах z и f , в виде

$$-\Delta z + f(z) = 0 \quad \text{в } S, \quad z = A \quad \text{на } \partial S. \quad (7.22)$$

Здесь $A > 0$, в силу А4, и при достаточно малых значениях $|p|$, в силу (7.4), (7.2) и А1–А3, можно показать, что для функции f справедливо

$$f(0) = 0, \quad f(A) > 0, \quad f'(0) > \gamma^2, \quad \int_0^s f(t)dt > 0 \quad \text{при } s \in (0, A]. \quad (7.23)$$

Заметим, что из (7.23) следует, что 0 и A соответственно являются нижним и верхним решениями для (7.22). Следовательно существует по крайней мере одно решение $z \in [0, A]$ задачи (7.22). Однако, чтобы установить желаемые оценки решения, нам потребуются более точные нижнее и верхнее решения. Для более точного нижнего решения заметим, что, в силу Леммы 7.3, функция $\alpha = \max\{\tilde{v}_0^{\circ}, \tilde{v}_0^{\circ-}\}$ является нижним решением (7.22) и удовлетворяет $0 \leq \alpha \leq A$. Теперь определим \tilde{z}_0 как *единственное минимальное решение*, соответствующее нижнему решению α и верхнему решению A . Тогда $\tilde{z}_0 \leq A$, а также $\tilde{z}_0 \leq \beta$ для любого верхнего решения β , удовлетворяющего $\beta \geq \alpha$. Более точная верхняя оценка для \tilde{z}_0 будет построена при доказательстве Теоремы 7.2.

Доказательство Теоремы 7.2. Без ограничения общности, будем считать, что произвольная точка $\eta_0 = \xi_0 \mathbf{e}_r + \sigma_0 \mathbf{e}_s$ в S расположена ближе к стороне Γ , чем к Γ^- . Пусть $\bar{\eta} = \sigma_0 \mathbf{e}_s$, $\bar{\eta}^- = \sigma_0 \mathbf{e}_{s^-}$, а \mathcal{O} обозначает круг внутри S , касающийся Γ в точке $\bar{\eta}$ и Γ^- в точке $\bar{\eta}^-$. Заметим, что радиус \mathcal{O} , равный $\rho = \sigma_0 \tan(\omega/2)$, растёт при удалении η_0 от O . Рассмотрим задачу (7.22) в области \mathcal{O} (вместо S) и обозначим соответствующее решение $z_{\mathcal{O}}$:

$$-\Delta z_{\mathcal{O}} + f(z_{\mathcal{O}}) = 0 \quad \text{в } \mathcal{O}, \quad z_{\mathcal{O}} = A \quad \text{на } \partial \mathcal{O}. \quad (7.24)$$

Замена переменных $\hat{\eta} = \eta/\rho$ преобразует \mathcal{O} в единичный круг $\hat{\mathcal{O}}$. При этом задача (7.24) преобразуется в следующую задачу для $\hat{z}_{\mathcal{O}}(\hat{\eta}) := z_{\mathcal{O}}(\eta)$:

$$-\rho^{-2}\Delta_{\hat{\eta}}\hat{z}_{\mathcal{O}} + f(\hat{z}_{\mathcal{O}}) = 0 \quad \text{в } \hat{\mathcal{O}}, \quad \hat{z}_{\mathcal{O}} = A \quad \text{на } \partial\hat{\mathcal{O}}, \quad (7.25)$$

где ρ^{-1} становится малым параметром при достаточно больших значениях σ_0 , т. е. мы имеем сингулярно возмущенную задачу типа (7.1), но в гладкой области (являющейся единичным кругом). Такие задачи рассматривались, например, в [163]. В частности, воспользовавшись Следствием 2.9 из [163], заключаем, что найдется такая постоянная C , что $\omega(\hat{d})\overset{\circ}{v}_0(\rho\hat{d}; p + \hat{p}) + b_2^{-1}\hat{p} \leq \hat{z}_{\mathcal{O}} \leq C\rho^{-1}$ является верхним решением задачи (7.25). Здесь постоянная $b_2 \geq |b_{uu}(O, t)|$ при $t \in [u_0(O), g(O)]$, $\hat{d} = \hat{d}(\hat{\eta})$ обозначает расстояние от точки $\hat{\eta} \in \hat{\mathcal{O}}$ до $\partial\hat{\mathcal{O}}$, а $\omega(\hat{d})$ является гладкой функцией-срезкой, равной 1 при $\hat{d} > 1/2$ и 0 при $\hat{d} < 1/4$. (Отметим, что $\overset{\circ}{v}_0(\rho\hat{d}; p + \hat{p})$ заменяет $v(\rho\hat{d}, l; p + \hat{p})$ в обозначениях [163], и в настоящей ситуации не зависит от l .)

Интерпретируя наши выводы в терминах переменной η , для чего нам также потребуется расстояние $d = d(\eta) = \rho\hat{d}$ от точки $\eta \in \mathcal{O}$ до $\partial\mathcal{O}$, убеждаемся, что $\omega(d/\rho)\overset{\circ}{v}_0(d; p + \hat{p}) + b_2^{-1}\hat{p}$, где $\hat{p} = C\rho^{-1}$, является верхним решением задачи (7.24). У нас уже есть пара функций $\alpha = \max\{\overset{\circ}{v}_0, \overset{\circ}{v}_0^-\}$ и A , которые соответственно являются нижним и верхним решением для (7.24). Объединяя два верхних решения, как показано в Лемме 7.3, получаем более точное верхнее решение $\beta_{\mathcal{O}} := \min\{\omega(d/\rho)\overset{\circ}{v}_0(d; p + \hat{p}) + b_2^{-1}\hat{p}, A\}$, где $\hat{p} = C\rho^{-1}$.

Для существования решения $z_{\mathcal{O}}(\eta)$ задачи (7.24) между α и $\beta_{\mathcal{O}}$, важно проверить, что $\alpha \leq \beta_{\mathcal{O}}$, что следует из $\alpha \leq \omega(d/\rho)\overset{\circ}{v}_0(d; p + \hat{p}) + b_2^{-1}\hat{p}$. Поскольку точка η расположена ближе к Γ , последнее соотношение эквивалентно $\overset{\circ}{v}_0(\xi; p) \leq \omega(d/\rho)\overset{\circ}{v}_0(d; p + \hat{p}) + b_2^{-1}\hat{p}$. В той части области, где $\omega = 1$, это вытекает из неравенств $\overset{\circ}{v}_0(\xi; p) \leq \overset{\circ}{v}_0(\xi; p + \hat{p}) \leq \overset{\circ}{v}_0(d; p + \hat{p})$, где мы воспользовались монотонностью $\overset{\circ}{v}$ по обоим аргументам, а также тем, что $\xi \geq d(\eta)$ при любом $\eta = (\xi, \sigma)$. В противном случае, т. е. при $\omega < 1$, имеем $\xi \geq d(\eta) \geq \rho/2$ и $\hat{p} = C\rho^{-1}$, следовательно $\overset{\circ}{v}_0(\xi; p) \leq Ce^{-C\xi} \leq b_2^{-1}\hat{p}$ при достаточно больших

значениях ρ . Таким образом мы показали, что действительно $\alpha \leq \beta_{\mathcal{O}}$ при достаточно больших ρ .

Пусть $z_{\mathcal{O}}(\eta)$ является минимальным решением (7.24) между α и $\beta_{\mathcal{O}}$. В силу Леммы 7.6, примененной к задачам (7.22) и (7.24), заключаем, что $\tilde{z}_0(\eta) \leq z_{\mathcal{O}}(\eta) \leq \beta_{\mathcal{O}}$ в \mathcal{O} и, в частности, $\tilde{z}_0(\eta_0) \leq \beta_{\mathcal{O}}(\eta_0)$. Заметим что при $\eta = \eta_0$ справедливо $d = \xi_0$, и следовательно $\beta_{\mathcal{O}}(\eta_0) \leq \check{v}_0(\xi_0; p + \hat{p}) + b_2^{-1}\hat{p} \leq \check{v}_0(\xi_0; p) + C\hat{p}$. Наконец, вспомним, что $\hat{p} = C\rho^{-1} = C\sigma_0^{-1} \leq C|\eta_0|^{-1}$. Это доказывает верхнюю оценку для \tilde{z}_0 в (7.13) при любом $\eta = \eta_0$, расположенном ближе к Γ чем к Γ^- . Таким образом, (7.13) доказано.

В силу (7.13) и (7.23), найдется такое достаточно большое число $\Xi > 0$, что при

$$S_{\Xi} := \{\eta \in S : \min\{\xi, \xi^-\} > \Xi\}, \quad (7.26)$$

для любого $\eta \in S_{\Xi}$ величина $\tilde{z}_0(\eta; p)$ настолько мала, что $f'(\tilde{z}_0) > \gamma^2$. По теореме о среднем значении, \tilde{z}_0 удовлетворяет линейному уравнению $-\Delta_{\eta}\tilde{z}_0 + a(\eta)\tilde{z}_0 = 0$ на S_{Ξ} с коэффициентом $a(\eta) > \gamma^2$. Положим $W(\eta) = e^{-\gamma\xi} + e^{-\gamma\xi^-}$. Тогда $\Delta_{\eta}W = \gamma^2W$ и следовательно $[-\Delta_{\eta} + a(\eta)]W = [a(\eta) - \gamma^2]W > 0$. При этом $W \geq e^{-\gamma\Xi}$ на границе ∂S_{Ξ} , поскольку последняя состоит из двух лучей, один на расстоянии Ξ от Γ , а другой на расстоянии Ξ от Γ^- . Поэтому найдется такая постоянная $C > 0$, что $CW \geq A \geq \tilde{z}_0$ на ∂S_{Ξ} . Теперь, в силу принципа максимума, $\tilde{z}_0(\eta) \leq CW(\eta)$ в S_{Ξ} . Это доказывает (7.14).

Если $p < p'$, то $\tilde{z}_0(\cdot, p')$ является верхним решением для задачи, удовлетворяемой $\tilde{z}_0(\cdot, p)$, поэтому $\tilde{z}_0(\eta, p) \leq \tilde{z}_0(\eta, p')$, что доказывает свойство монотонности, заявленное в утверждении теоремы. Наконец, поскольку \tilde{z}_0 удовлетворяет линейному уравнению $\Delta_{\eta}\tilde{z}_0 = F$ с ограниченной функцией $F(\eta) = \tilde{B}(O, \tilde{z}_0)$, стандартная оценка для первых производных приводит к $|\nabla_{\eta}\tilde{z}_0| \leq C$ в S . Это завершает доказательство Теоремы 7.2. ■

З а м е ч а н и е 7.7. В контексте задачи (7.22), предположение A3 эквивалентно условию $f(A) > 0$, которое присутствует в (7.23) и используется при доказательстве существования решения z , удовлетворяющего

$0 \leq z \leq A$. Более того, $A3$ необходимо для существования такого решения в следующем смысле. Результат из [144] показывает, что если в (7.23) мы заменим $f(A) > 0$ на $f(A) < 0$, то найдется такое значение $\theta_f \in (\pi/2, \pi)$, что в любом секторе S , угол при вершине которого меньше θ_f , не существуют решений z задачи (7.22), удовлетворяющих $0 \leq z \leq A$. (Отметим, что в [144] рассматриваются ограниченные области, но изучение доказательств показывает, что результаты этой статьи также применимы к неограниченному сектору S .) Таким образом, если условие $A3$ нарушено, и вместо этого выполняется $f(A) < 0$, то задача (7.22) не имеет решений, удовлетворяющих $0 \leq z \leq A$, даже если она поставлена в четверти плоскости.

Отметим также, что условие $0 \leq z \leq A$ исключает появление резких пиков в решениях u задачи (7.1) в окрестностях углов Ω . Действительно, вспомним, что $z = \tilde{z}_0$ является угловой погранфункцией, которая будет присутствовать в асимптотическом разложении, т. е. z должна быть пренебрежимо малой вдали от $\Gamma \cup \Gamma^-$, и также должна удовлетворять $z \approx \mathring{v}_0(\xi)$ и $z \approx \mathring{v}_0^-(\xi^-)$, соответственно, около границ Γ и Γ^- вдали от соответствующей вершины. Поскольку $\mathring{v}_0, \mathring{v}_0^- \in [0, A]$, заключаем, что соотношение $0 \leq z \leq A$ может быть нарушено только при $|x - P_{j-1}| \leq C\varepsilon$, и следовательно такое нарушение привело бы к появлению острых пиков в решении $u \approx z(\eta) + u_0(P_{j-1})$ в $O(\varepsilon)$ -окрестности вершины P_{j-1} .

7.3.3 Экспоненциальное убывание функции \tilde{q}_0

Функция \tilde{q}_0 определяется (7.17a), а также, как было показано выше, удовлетворяет нелинейной краевой задаче (7.18). В данном разделе, чтобы установить экспоненциальное убывание \tilde{q}_0 , мы воспользуемся некоторым эквивалентным вариантом краевой задачи (7.18).

Л е м м а 7.8. *Существуют постоянные C_1 и c_1 , при которых справед-*

ливо соотношение

$$|\tilde{q}_0| + |\nabla \tilde{q}_0| \leq C_1 e^{-c_1 |\eta|} \quad \text{в } S.$$

Д о к а з а т е л ь с т в о. Граничные условия (7.18b) являются экспоненциально убывающими, но проанализировать поведение правой части в (7.18a) представляется не слишком простой задачей. Вместо этого удобно исследовать функцию $\tilde{q}_0 + \overset{\circ}{v}_0^-$. Введем операторы \mathcal{N} и \mathcal{L} :

$$\mathcal{N}[W] = -\Delta W + \tilde{B}(O, \cdot) \Big|_{\overset{\circ}{v}_0}^{W + \overset{\circ}{v}_0}, \quad \mathcal{L}W = -\Delta W + W \tilde{B}_t(O, \overset{\circ}{v}_0).$$

Тогда имеем

$$\tilde{B}(O, \cdot) \Big|_{\overset{\circ}{v}_0}^{W + \overset{\circ}{v}_0} = W \tilde{B}_t(O, \overset{\circ}{v}_0) + W^2 R,$$

где $|R| \leq C_R = \frac{1}{2} \sup |\tilde{B}_{tt}|$. Следовательно

$$\mathcal{N}[W] = \mathcal{L}W + W^2 R. \quad (7.27)$$

Оператор \mathcal{N} будет присутствовать в краевой задаче для $\tilde{q}_0 + \overset{\circ}{v}_0^-$. Действительно, из (7.18a) и (7.8) получаем

$$\mathcal{N}[\tilde{q}_0 + \overset{\circ}{v}_0^-] = 0 \quad \text{в } S. \quad (7.28a)$$

Далее, из (7.18b) немедленно следует

$$\tilde{q}_0 + \overset{\circ}{v}_0^- = 0 \quad \text{на } \Gamma. \quad (7.28b)$$

Поскольку $\tilde{q}_0 + \overset{\circ}{v}_0^-$ экспоненциально убывает не во всем секторе S , сформулируем нашу краевую задачу в соответствующей подобласти сектора S . Пусть $\omega_1 \leq \omega/2$, и пусть Γ_{ω_1} является лучом в S , который имеет угол ω_1 с Γ , а $S_{\omega_1} \subset S$ пусть будет сектором со сторонами Γ и Γ_{ω_1} . Легко увидеть, что $\overset{\circ}{v}_0^- \leq \overset{\circ}{v}_0$ в S_{ω_1} . Следовательно, из (7.13), $0 < \overset{\circ}{v}_0 \leq \tilde{z}_0$ в S_{ω_1} , и далее $0 \leq \tilde{z}_0 - \overset{\circ}{v}_0 = q_0 + \overset{\circ}{v}_0^- < \tilde{z}_0$ в S_{ω_1} . Также воспользовавшись (7.14), заключаем, что при любом сколь угодно малом $c_2 > 0$, существует такое достаточно большое $\bar{\rho} = \bar{\rho}(c_2) > 0$, что

$$|\tilde{q}_0 + \overset{\circ}{v}_0^-| \leq C e^{-\gamma \xi} = C e^{-\gamma(\tan \omega_1) \sigma} \leq c_2 e^{-\gamma(\tan \omega_1) \sigma / 2} \quad \text{на } \Gamma_{\omega_1}, \quad \sigma > \bar{\rho}. \quad (7.28c)$$

Пусть $S'_{\omega_1} = \{\eta \in S_{\omega_1} : \sigma > \bar{\rho}\}$, при этом Γ'_{ω_1} обозначает часть луча Γ_{ω_1} , лежащую вдоль S'_{ω_1} , а $\Gamma_{\bar{\rho}}$ пусть обозначает порцию $\partial S'_{\omega_1}$ с $\sigma = \bar{\rho}$. Выбирая $\bar{\rho}$ достаточно большим, в силу (7.13), получаем

$$|\tilde{q}_0 + \overset{\circ}{v}_0^-| = |\tilde{z}_0 - \overset{\circ}{v}_0| \leq C|\eta|^{-1} \leq C/\bar{\rho} \leq c_2 \quad \text{на } \Gamma_{\bar{\rho}}. \quad (7.28d)$$

Рассмотрим нелинейную задачу

$$\mathcal{N}[W] = 0 \quad \text{в } S'_{\omega_1}, \quad W = \tilde{q}_0 + \overset{\circ}{v}_0^- \quad \text{на } \partial S'_{\omega_1}. \quad (7.29)$$

Построим экспоненциально-убывающее верхнее решение \bar{W} для (7.29). Это верхнее решение определим как $\bar{W} = w(\xi)\phi(\sigma)$, где $\phi(\sigma) = c_\phi e^{-a\sigma}$, при этом a и c_ϕ выбираются достаточно малыми, а w является частным решением уравнения

$$-w'' + w\tilde{B}_t(O, \overset{\circ}{v}_0) = 1.$$

Чтобы определить w , вспомним, что функция $\overset{\circ}{v}_0(\xi) \geq 0$ монотонно экспоненциально убывает, а также удовлетворяет $-\overset{\circ}{v}_0'' + \tilde{B}(O, \overset{\circ}{v}_0) = 0$. Полагая $\chi(\xi) = -\overset{\circ}{v}_{0,\xi} \geq 0$, получаем (см. доказательство Леммы 2.2 в [88])

$$w(\xi) = \chi(\xi) \int_0^\xi \chi(\eta)^{-2} \overset{\circ}{v}_0(\eta) d\eta + \chi(\xi). \quad (7.30)$$

Поскольку $\underline{\gamma}\overset{\circ}{v}_0 \leq \chi \leq \bar{\gamma}\overset{\circ}{v}_0$ (см., например, [182, оценку (A.2)]), несложно убедиться, что при некоторых c_w и C_w справедливо соотношение

$$0 \leq c_w \leq w(\xi) \leq C_w.$$

Теперь покажем, что при соответствующем выборе a и c_ϕ , функция \bar{W} действительно становится верхним решением. Заметим, что

$$\begin{aligned} \mathcal{L}[\bar{W}] &= -w''\phi - w\phi'' + [w\phi]\tilde{B}_t(O, \overset{\circ}{v}_0) = [-w'' + w\tilde{B}_t(O, \overset{\circ}{v}_0)]\phi - w\phi'' \\ &= \phi - w\phi'' = \phi(1 - a^2w). \end{aligned}$$

Таким образом $\mathcal{L}[\bar{W}] \geq \frac{1}{2}\phi$ при $a^2 \leq \frac{1}{2}C_w^{-1}$. Также, воспользовавшись (7.27), получаем

$$\begin{aligned} \mathcal{N}[\bar{W}] &= \mathcal{L}\bar{W} + w^2\phi^2R \\ &\geq \frac{1}{2}\phi - C_w^2\phi^2C_R = \phi\left[\frac{1}{2} - C_w^2C_R\phi\right] = \phi\left[\frac{1}{2} - C_w^2C_Rc_\phi e^{-a\sigma}\right] \\ &\geq 0 = \mathcal{N}[\tilde{q}_0 + \overset{\circ}{v}_0^-] \quad \text{при } c_\phi = \frac{1}{2}C_w^{-2}C_R^{-1}. \end{aligned}$$

Поскольку $\bar{W} > 0$ на Γ , из (7.28b) получаем $\bar{W} \geq \tilde{q}_0 + \overset{\circ}{v}_0^-$ на Γ . Теперь покажем, что $\bar{W} \geq \tilde{q}_0 + \overset{\circ}{v}_0^-$ на $\partial S'_{\omega_1}$. Воспользовавшись (7.28c), получаем

$$|\tilde{q}_0 + \overset{\circ}{v}_0^-| \leq c_2 e^{-\gamma(\tan\omega_1)\sigma/2} \leq c_w \cdot c_\phi e^{-a\sigma} \leq w\phi \quad \text{на } \Gamma'_{\omega_1}$$

при выполнении $c_2 < c_w c_\phi$ и $a \leq \gamma(\tan\omega_1)/2$. Здесь мы воспользовались (7.28c), которое позволяет выбрать сколь угодно малую постоянную c_2 за счет выбора достаточно большого $\bar{\rho}$. Наконец, на $\Gamma_{\bar{\rho}}$ в силу (7.28d) получаем

$$|\tilde{q}_0 + \overset{\circ}{v}_0^-| \leq c_2 \leq c_w \cdot c_\phi e^{-a\bar{\rho}} \leq w\phi \quad \text{на } \Gamma_{\bar{\rho}}$$

при $c_2 < c_w c_\phi$ и достаточно малом a . Для выполнения данного соотношения $\bar{\rho} = \bar{\rho}(c_2)$ вновь выбирается достаточно большим независимо от a , а $a = a(\bar{\rho})$ далее выбирается достаточно малым. В целом, если $\bar{\rho}$ выбирается достаточно большим, а a достаточно близким к 0, то функция \bar{W} становится верхним решением задачи (7.29).

Поскольку \tilde{z}_0 является единственным минимальным решением, соответствующим нижнему решению $\max\{\overset{\circ}{v}_0, \overset{\circ}{v}_0^-\}$, то, в силу Леммы 7.5, функция $W = \tilde{q}_0 + \overset{\circ}{v}_0^- = \tilde{z}_0 - \overset{\circ}{v}_0$, рассматриваемая в области S'_{ω_1} , является единственным минимальным решением (7.29), соответствующим нижнему решению $\underline{W} = \max\{\overset{\circ}{v}_0, \overset{\circ}{v}_0^-\} - \overset{\circ}{v}_0 = 0$. Очевидно, что $\underline{W} \leq \bar{W}$, поэтому заключаем, что $0 = \underline{W} \leq W = \tilde{q}_0 + \overset{\circ}{v}_0^- \leq \bar{W} \leq Ce^{-a|\eta|}$ в S_{ω_1} .

Объединяя это с $|\overset{\circ}{v}_0^-| \leq Ce^{-a|\eta|}$ в S_{ω_1} , получаем $|\tilde{q}_0| \leq Ce^{-a|\eta|}$ в S_{ω_1} . Аналогичные рассуждения показывают, что $|\tilde{q}_0| \leq Ce^{-a|\eta|}$ в секторе $S_{\omega_1}^-$ с углом ω_1 , расположенным вдоль стороны Γ^- сектора S . Из неравенства (7.14) следует, что $|\tilde{q}_0| \leq C_1 e^{-c_1|\eta|}$ в $S \setminus (S_{\omega_1} \cup S_{\omega_1}^-)$, поэтому $|\tilde{q}_0| \leq C_1 e^{-c_1|\eta|}$ во всем секторе S .

Чтобы оценить первые производные \tilde{q}_0 , введем $\omega_2 \in (\frac{1}{2}\omega, \omega)$, и пусть Γ_{ω_2} является лучом в S , имеющим угол ω_2 с Γ , а $S_{\omega_2} \subset S$ будет сектором со сторонами Γ и Γ_{ω_2} . Воспользовавшись (7.27), (7.28a) и экспоненциальным убыванием \tilde{q}_0 и \tilde{v}_0^- в S_{ω_2} , убеждаемся, что $|\Delta(\tilde{q}_0 + \tilde{v}_0^-)| \leq C|\tilde{q}_0 + \tilde{v}_0^-| \leq Ce^{-c|\eta|}$ в S_{ω_2} . Далее, применяя локальные оценки типа Шаудера для оценки первых производных к парам концентрических кругов радиусов 1 и 2, которые могут пересекать Γ , но не пересекают Γ_{ω_2} , убеждаемся, что $|\nabla(\tilde{q}_0 + \tilde{v}_0^-)| \leq Ce^{-c|\eta|}$, и следовательно $|\nabla\tilde{q}_0| \leq Ce^{-c|\eta|}$ внутри любого допустимого круга радиуса 1. Поскольку такие единичные круги покрывают $\bar{S}_{\omega/2} \cap \{|\eta| \geq \bar{C}\}$, в силу Теоремы 7.2 получаем желаемое экспоненциальное убывание во всем секторе $\bar{S}_{\omega/2}$. \square

7.3.4 Корректность линеаризованной задачи в секторе

Пусть \tilde{z}_0 является решением задачи (7.12), описанным в Теореме 7.2, а $\tilde{a}(\eta) = \tilde{B}_t(O, \tilde{z}_0(\eta))$. В данном разделе мы установим корректность задачи

$$\mathcal{M}W := -\Delta W + \tilde{a}W = F \quad \text{в } S, \quad W = 0 \quad \text{на } \partial S. \quad (7.31)$$

Как следствие, будет установлено существование функций z_1 и $\tilde{z}_{0,p}$ и экспоненциальное убывание их компонент q_1 и $\tilde{q}_{0,p}$.

Пусть S_R обозначает ограниченный сектор радиуса R . Стороны S_R будут обозначены Γ_R и Γ_R^- . Рассмотрим задачу на собственные значения

$$\mathcal{M}\Phi_R = \lambda_R\Phi_R \quad \text{в } S_R, \quad \Phi_R = 0 \quad \text{на } \partial S_R. \quad (7.32)$$

Воспользовавшись стандартными результатами для задач на собственные значения [85, §6.5.1] для оператора $\mathcal{M} + C_a$, где $C_a > \max|\tilde{a}|$, заключаем, что задача (7.32) имеет счетное множество действительных собственных значений $\lambda_{R,1} < \lambda_{R,2} \leq \dots$ и соответствующих собственных функций. При этом главному собственному значению $\lambda_{R,1}$ соответствует единственная собственная функция, которую обозначим ϕ , и для которой справедливо $\phi > 0$ в S_R .

Вообще говоря, коэффициент \tilde{a} не является неотрицательным. Тем не менее будет установлено (см. Теорему 7.15), что $\lambda_{R,1}$ положительно и, более того, отделено от 0 положительной постоянной равномерно по R . Из этого следует (см. Лемму 7.16), что задача (7.31), рассматриваемая в ограниченном секторе, является корректно поставленной и имеет решение, ограниченное равномерно по R . Далее, полагая $R \rightarrow \infty$, будет установлена корректность постановки задачи для бесконечного сектора (см. Теорему 7.17).

Пусть направление оси $\bar{\xi}$ совпадает с лучом $\omega/2$, и положим

$$Z(\eta) := -\frac{\partial \tilde{z}_0}{\partial \bar{\xi}}. \quad (7.33)$$

Очевидно, что $\mathcal{M}Z = 0$ и $Z \geq 0$ на ∂S , поскольку $\tilde{z}_0 = A$ на ∂S и $\tilde{z}_0 \leq A$ на S . Отметим также, что поскольку функция \tilde{z}_0 равна постоянной на Γ , то $Z = -(\sin \frac{1}{2}\omega)\tilde{z}_{0,\xi}$ на Γ .

Л е м м а 7.9. *Справедливо, что $0 \leq Z \leq C$ на S . Также, для любого $R' > 0$ найдется такая постоянная $C_1 = C_1(R') > 0$, что, если $\eta \in \Gamma \cup \Gamma^-$ и $|\eta| > R'$, то $Z(\eta) \geq C_1$.*

Д о к а з а т е л ь с т в о. Ограниченность Z вытекает из Теоремы 7.2. Чтобы показать, что $Z \geq 0$, вспомним, что \tilde{z}_0 было определено в §7.3.2 как минимальное решение (7.12), соответствующее нижнему решению $\alpha = \max\{\overset{\circ}{v}_0, \overset{\circ}{v}_0^-\}$. Проследивая доказательство Теоремы 7.1 из [129], убеждаемся, что \tilde{z}_0 может быть получено как предел возрастающей последовательности нижних решений $\{\alpha^{(k)}\}$, т. е. $\tilde{z}_0 = \lim_{k \rightarrow \infty} \alpha^{(k)}$, где $\alpha^{(0)} := \alpha$, а последующие $\alpha^{(k)}$ определяются индуктивно из

$$-\Delta \alpha^{(k)} + \bar{C} \alpha^{(k)} = \bar{C} \alpha^{(k-1)} - \tilde{B}(O, \alpha^{(k-1)}), \quad \alpha^{(k)} \Big|_{\partial S} = A.$$

Здесь $\bar{C} \geq \tilde{B}_t(O, t)$ при любом $t \in [0, A]$. Далее, несложно проверить, что

$$[-\Delta + \bar{C}] \alpha_{\bar{\xi}}^{(k)} = [\bar{C} - \tilde{B}_t(O, \alpha^{(k-1)})] \alpha_{\bar{\xi}}^{(k-1)}, \quad \alpha_{\bar{\xi}}^{(k)} \Big|_{\partial S} \leq 0,$$

где соотношение на границе при каждом k следует из $\alpha^{(k)} \leq A$ в S . Отметим также, что для $\alpha^{(0)} = \max\{\overset{\circ}{v}_0, \overset{\circ}{v}_0^-\}$ также справедливо соотношение

$$\alpha_{\xi}^{(0)} \leq 0 \quad \text{в } S.$$

С помощью принципа максимума, по индукции получаем, что

$$-\alpha_{\xi}^{(k)} \geq 0 \quad \text{в } S,$$

и следовательно, полагая $k \rightarrow \infty$, получаем неотрицательность Z .

Далее покажем, что для любого $R' > 0$, найдется такая постоянная $C_1 > 0$, что справедливо $Z \geq C_1 > 0$ при выполнении $\eta \in \Gamma \cup \Gamma^-$ и $|\eta| \geq R'$. Вспомним, что $Z = -\partial\tilde{z}_0/\partial\bar{\xi} = -[\overset{\circ}{v}_0(\xi) + \overset{\circ}{v}_0^-(\xi^-)]_{\bar{\xi}} - \tilde{q}_{0,\bar{\xi}}$. Поскольку первое слагаемое удовлетворяет желаемому соотношению, а в силу Леммы 7.8 слагаемое $\tilde{q}_{0,\bar{\xi}}$ может быть сделано сколь угодно малым за счет выбора $|\eta|$ достаточно большим, найдутся такие положительные постоянные R'' и C'' , что если $|\eta| > R''$ и $\eta \in \Gamma \cup \Gamma^-$, то $Z(\eta) > C''$.

Остается убедиться, что при любом $R' > 0$, найдется такая постоянная $C' > 0$, что $Z \geq C'$ при выполнении $\eta \in \Gamma \cup \Gamma^-$ и $R' \leq |\eta| \leq R''$. Это немедленно следует из соотношения $Z > 0$ на $\partial S \setminus O$, которое нам еще нужно доказать. Для этого, вспомним, что $\tilde{z}_0 \leq A$ является решением задачи (7.12), где $A > 0$ и $\tilde{B}(O, A) > 0$. Отсюда несложно проверить, что

$$[-\Delta + \bar{C}](\tilde{z}_0 - A) < (\bar{C}t - \tilde{B}(O, t))\Big|_A^{\tilde{z}_0} \leq 0 \quad \text{в } S.$$

Объединяя это с $\tilde{z}_0 - A = 0$ на ∂S и применяя принцип максимума, получаем $\tilde{z}_0 - A < 0$ в S . Также, в силу леммы Хопфа [85, §6.4.2], получаем $\partial(\tilde{z}_0 - A)/\partial n > 0$ на $\partial S \setminus O$. Поэтому $Z = (\sin \frac{1}{2}\omega)\partial\tilde{z}_0/\partial n > 0$ на $\partial S \setminus O$. \square

Поскольку пара $(\lambda_{R,1}, \phi)$ является решением (7.32), приходим к уравнению

$$\lambda_{R,1}\phi Z = -(\Delta\phi)Z + \tilde{a}\phi Z = -(\Delta\phi)Z + \phi(\Delta Z),$$

интегрируя которое по S_R , а также воспользовавшись тем, что $\phi = 0$ на ∂S_R , получаем

$$\lambda_{R,1} \iint_{S_R} \phi Z = - \int_{\partial S_R} Z \frac{\partial \phi}{\partial n} > 0. \quad (7.34)$$

Поскольку здесь ϕ и Z положительны в S_R , следовательно $\lambda_{R,1} > 0$. Далее нам еще предстоит получить равномерную по R нижнюю оценку для $\lambda_{R,1}$.

Положим $\rho(\eta) = \tilde{a}(\eta) - \lambda_{R,1}$, тогда $-\Delta \phi + \rho \phi = 0$. Вспомним, что $\tilde{a}(\eta) = \tilde{B}_t(O, \tilde{z}_0(\eta))$, а, в силу (7.4) и A1, справедливо $\tilde{B}_t(O, 0) > 0$. Воспользовавшись (7.14), выберем такое число $\Xi > 0$, что справедливо $\tilde{a}(\eta) \geq (3/4)\tilde{B}_t(O, 0)$ при $\min\{\xi, \xi^-\} \geq \Xi$. Положим $S(\Xi) = \{\eta \in S_R : \min\{\xi, \xi^-\} \geq \Xi\}$. Таким образом,

$$\rho(\eta) \geq \frac{1}{4}\tilde{B}_t(O, 0) \quad \text{при } \lambda_{R,1} \leq \frac{1}{2}\tilde{B}_t(O, 0) \text{ и } \eta \in S(\Xi). \quad (7.35)$$

Наконец, при $R' < R$ пусть $\Gamma_{R',R}$ обозначает множество точек $\eta \in \Gamma_R$, удовлетворяющих $|\eta| > R'$.

Л е м м а 7.10. Если $\lambda_{R,1} < \frac{1}{2}\tilde{B}_t(O, 0)$, то для любого $R' \in (0, R)$ существует такая не зависящая от R положительная постоянная $C_2(R')$, что

$$\lambda_{R,1} \geq C_2 \frac{\int_{\Gamma_{R',R}} \partial \phi / \partial \xi}{\|\phi\|_{L_1(S_R \cap \{\xi \leq \Xi\})}}. \quad (7.36)$$

Д о к а з а т е л ь с т в о. Поскольку $\partial \phi / \partial n \leq 0$ на ∂S_R и $-\partial \phi / \partial n = \partial \phi / \partial \xi$ на Γ_R , соотношение (7.34) и Лемма 7.9 приводят к

$$\lambda_{R,1} = \frac{\int_{\partial S_R} |\partial \phi / \partial n| Z}{\int_{S_R} \phi Z} \geq 2C_1 \frac{\int_{\Gamma_{R',R}} \phi \xi}{\|Z\|_{L_\infty(S_R)} \|\phi\|_{L_1(S_R)}},$$

где $C_1 = C_1(R')$ из Леммы 7.9. Воспользовавшись оценками Z , описанными в Лемме 7.9, остается показать, что

$$\|\phi\|_{L_1(S_R)} \leq C \|\phi\|_{L_1(S_R \cap \{\xi \leq \Xi\})}. \quad (7.37)$$

Из (7.35) получаем

$$\begin{aligned} 0 &\geq \int_{\partial S_R} \frac{\partial \phi}{\partial n} = \int_{S_R} \Delta \phi = \int_{S_R} \rho \phi = \int_{S(\Xi)} \rho \phi + \int_{S_R \setminus S(\Xi)} \rho \phi \\ &\geq \frac{1}{4}\tilde{B}_t(O, 0) \|\phi\|_{L_1(S(\Xi))} - (\max_S |\rho|) \|\phi\|_{L_1(S_R \setminus S(\Xi))}. \end{aligned}$$

А поскольку $\tilde{B}_t(O, 0) > 0$, справедлива оценка

$$\|\phi\|_{L_1(S(\Xi))} \leq C\|\phi\|_{L_1(S_R \setminus S(\Xi))} \leq 2C\|\phi\|_{L_1(S: \xi \leq \Xi)},$$

где в последнем неравенстве мы воспользовались симметрией ϕ по ξ и ξ^- . Добавляя $\|\phi\|_{L_1(S_R \setminus S(\Xi))}$ к обеим сторонам, получаем (7.37). \square

Далее сформулируем вспомогательный результат, который будет полезен при получении нижней оценки для $\int_{\Gamma_{R', R}} \partial\phi/\partial\xi$ в последующей Лемме 7.13.

Л е м м а 7.11. Пусть даны положительные числа ρ_0 , $\bar{\delta}$ и a , при этом $0 < \delta \leq \bar{\delta}$. Пусть Ψ является ограниченной положительной функцией на \mathbb{R} с $\Psi(x) = 0$ при $|x| > a$, а ограниченная функция $\psi(x, y)$ определяется при $(x, y) \in (-\infty, \infty) \times [0, \delta]$ соотношениями

$$-\Delta\psi + \rho_0\psi = 0, \quad \psi(x, 0) = 0, \quad \psi(x, \delta) = \Psi(x).$$

Тогда $\psi \geq 0$, и найдутся такие положительные числа $\bar{a} = \bar{a}(\bar{\delta})$ и $C_0(\bar{\delta})$, что для $I = \int_{-a}^a \Psi(x)dx$ справедливы соотношения

$$\int_{-(a+\bar{a})}^{(a+\bar{a})} \psi_y(x, 0)dx \geq \frac{C_0(\bar{\delta})}{\delta} I, \quad \int_{|x| > |a+\bar{a}|} \psi_y(x, 0)dx \leq \frac{C_0(\bar{\delta})}{2\delta} I.$$

Д о к а з а т е л ь с т в о. Применяя к ψ преобразование Фурье по переменной x , получаем представление решения $\psi(x, y) = \int_{-a}^a G(x-t, y)\Psi(t)dt$, где

$$G(x, y) = \frac{1}{\pi} \int_0^\infty \cos(xs) \frac{\sinh(y\sqrt{s^2 + \rho_0})}{\sinh(\delta\sqrt{s^2 + \rho_0})} ds.$$

В силу принципа максимума, $G(x, y) \geq 0$, при этом $G(x, 0) = 0$, и следовательно

$$G_y(x, 0) = \frac{1}{\pi} \int_0^\infty \cos(xs) \frac{\sqrt{s^2 + \rho_0}}{\sinh(\delta\sqrt{s^2 + \rho_0})} ds \geq 0.$$

Поскольку $\psi_y(x, 0) = \int_{-a}^a G_y(x-t, 0)\Psi(t)dt$, получаем

$$\int_{-a-\bar{a}}^{a+\bar{a}} \psi_y(x, 0) dx = \int_{-a}^a I_1(t)\Psi(t)dt, \quad I_1(t) := \int_{-a-\bar{a}-t}^{a+\bar{a}-t} G_y(x, 0) dx.$$

Если $t \in [-a, a]$, то $[-a - \bar{a} - t, a + \bar{a} - t] \supset [-\bar{a}, \bar{a}] \supset [-\delta\tilde{a}, \delta\tilde{a}]$, где \tilde{a} — произвольное положительное число, удовлетворяющее $\tilde{a} \leq \bar{a}/\bar{\delta}$, а $\delta \leq \bar{\delta}$. Воспользовавшись также $G_y(x, 0) \geq 0$, получаем

$$I_1(t) \geq \frac{1}{\pi} \int_{-\delta\tilde{a}}^{\delta\tilde{a}} dx \int_0^\infty \cos(xs) \frac{\sqrt{s^2 + \rho_0}}{\sinh(\delta\sqrt{s^2 + \rho_0})} ds.$$

Воспользовавшись заменой переменных $\hat{x} = x/\delta$ и $\hat{s} = s\delta$ (при этом $xs = \hat{x}\hat{s}$ и $dx ds = d\hat{x} d\hat{s}$), получаем

$$I_1(t) \geq \frac{1}{\delta\pi} \int_{-\tilde{a}}^{\tilde{a}} d\hat{x} \int_0^\infty \cos(\hat{x}\hat{s}) \frac{\sqrt{\hat{s}^2 + \delta^2\rho_0}}{\sinh(\sqrt{\hat{s}^2 + \delta^2\rho_0})} d\hat{s}.$$

Заметим, что здесь $\delta^2\rho_0 \in (0, \bar{\delta}^2\rho_0)$, в то время как $e^{-s} \leq s/\sinh s \leq 2e^{-s/2}$ для любого положительного s . Следовательно

$$e^{-\bar{\delta}\sqrt{\rho_0}} e^{-\hat{s}} \leq \frac{\sqrt{\hat{s}^2 + \delta^2\rho_0}}{\sinh(\sqrt{\hat{s}^2 + \delta^2\rho_0})} \leq 2e^{-\hat{s}/2}.$$

Объединяя это с $\cos(\hat{x}\hat{s}) > \frac{1}{2}$ при $\hat{s} \leq 1/|\hat{x}|$ и $|\cos(\hat{x}\hat{s})| \leq 1$ в противном случае, получаем

$$\begin{aligned} I_1 &\geq \frac{1}{\delta\pi} \int_{-\tilde{a}}^{\tilde{a}} d\hat{x} \left[\frac{1}{2} e^{-\bar{\delta}\sqrt{\rho_0}} \int_0^{1/|\hat{x}|} e^{-\hat{s}} d\hat{s} - 2 \int_{1/|\hat{x}|}^\infty e^{-\hat{s}/2} d\hat{s} \right] \\ &= \frac{C}{\delta} \int_{-\tilde{a}}^{\tilde{a}} d\hat{x} \left[1 - e^{-1/|\hat{x}|} - C' e^{-1/(2|\hat{x}|)} \right], \end{aligned}$$

где $C = C(\bar{\delta})$ и $C' = C'(\bar{\delta})$. Выберем число $\tilde{a} = \tilde{a}(\bar{\delta}) > 0$ достаточно малым, с тем чтобы подинтегральная функция была $\geq C > 0$, что приводит к $I_1 \geq C_0(\bar{\delta})/\delta$.

Далее рассмотрим

$$\int_{a+\bar{a}}^\infty \psi_y(x, 0) dx = \int_{-a}^a I_2(t) \Psi(t) dt, \quad I_2(t) = \int_{a+\bar{a}-t}^\infty G_y(x, 0) dx.$$

Поскольку $t \in [-a, a]$, убеждаемся, что $a + \bar{a} - t \geq \bar{a} \geq \delta\bar{a}/\bar{\delta}$. Объединяя это с $G_y(x, 0) \geq 0$, получаем

$$0 \leq I_2(t) \leq \int_{\delta\bar{a}/\bar{\delta}}^\infty G_y(x, 0) dx = \frac{1}{2\pi} \int_{\delta\bar{a}/\bar{\delta}}^\infty dx \int_{-\infty}^\infty e^{ixs} \frac{\sqrt{s^2 + \rho_0}}{\sinh(\delta\sqrt{s^2 + \rho_0})} ds.$$

Вновь воспользовавшись $\hat{x} = x/\delta$ и $\hat{s} = s\delta$, получаем

$$I_2(t) \leq \frac{\delta^{-1}}{2\pi} \int_{\bar{a}/\delta}^{\infty} d\hat{x} \int_{-\infty}^{\infty} e^{i\hat{x}\hat{s}} g(\hat{s}) d\hat{s} = \frac{\delta^{-1}}{2\pi} \int_{\bar{a}/\delta}^{\infty} \frac{d\hat{x}}{(i\hat{x})^2} \int_{-\infty}^{\infty} e^{i\hat{x}\hat{s}} g''(\hat{s}) d\hat{s},$$

где

$$g(\hat{s}) = \frac{\sqrt{\hat{s}^2 + \delta^2 \rho_0}}{\sinh(\sqrt{\hat{s}^2 + \delta^2 \rho_0})},$$

и дважды было использовано интегрирование по частям. Несложно убедиться, что функция $g(\hat{s})$ и ее производные определены корректно (например, они ограничены в 0), а также

$$|g(\hat{s})| + |g'(\hat{s})| + |g''(\hat{s})| \leq C e^{-C|\hat{s}|}.$$

Поэтому получаем

$$I_2(t) \leq \frac{C}{\delta} \int_{\bar{a}/\delta}^{\infty} \frac{d\hat{x}}{\hat{x}^2} = \frac{C}{\delta \bar{a}/\delta} = \frac{C_0(\bar{\delta})}{4\delta}$$

при условии, что $\bar{a} = \bar{a}(\bar{\delta})$ выбрано достаточно большим. Оставшийся интеграл по $(-\infty, -(a + \bar{a}))$ оценивается аналогичным образом. \square

С л е д с т в и е 7.12. *При выполнении условий Леммы 7.11, существует функция $\tilde{\psi}$, которая удовлетворяет соотношениям*

$$-\Delta \tilde{\psi} + \rho_0 \tilde{\psi} = 0, \quad \tilde{\psi}(x, 0) = 0, \quad \tilde{\psi}(x, \delta) = \Psi(x), \quad \tilde{\psi}(x, \pm(a + \bar{a})) \leq 0$$

в области $[-(a + \bar{a}), a + \bar{a}] \times [0, \delta]$, при этом

$$\int_{-(a+\bar{a})}^{(a+\bar{a})} \tilde{\psi}_y(x, 0) dx \geq C'_0(\bar{\delta}) \int_{-a}^a \Psi(x) dx.$$

Д о к а з а т е л ь с т в о. Пусть ψ является функцией, описанной в Лемме 7.11, и пусть

$$\tilde{\psi}(x, y) := \psi(x, y) - \psi(2(a + \bar{a}) - x, y) - \psi(-2(a + \bar{a}) - x, y).$$

Тогда $\tilde{\psi}(\pm(a + \bar{a})) = -\psi(\mp 3(a + \bar{a})) \leq 0$, а $\tilde{\psi}$ удовлетворяет тому же уравнению, что и ψ . Отсюда вытекает желаемое неравенство с $C'_0(\bar{\delta}) = \frac{1}{2}\bar{\delta}^{-1}C_0(\bar{\delta})$. \square

Л е м м а 7.13. *Найдутся такие независящие от R положительные числа \bar{a} и \bar{C} , что при любых σ_1, σ_2 , удовлетворяющих $0 < \sigma_1 < \sigma_2$ и $[0, \Xi] \times [\sigma_1 - \bar{a}, \sigma_2 + \bar{a}] \subset S_R$, где η интерпретируется как $\eta = (\xi, \sigma)$, справедлива оценка*

$$\int_{\sigma_1 - \bar{a}}^{\sigma_2 + \bar{a}} \phi_\xi(0, \sigma) d\sigma \geq \bar{C} \|\phi\|_{L_1([0, \Xi] \times [\sigma_1, \sigma_2])}.$$

Д о к а з а т е л ь с т в о. Поскольку $\phi \geq 0$, применяя теорему о среднем значении к положительной функции $\Phi(\xi) := \int_{\sigma_1}^{\sigma_2} \phi(\xi, \sigma) d\sigma$, получаем

$$\|\phi\|_{L_1([0, \Xi] \times [\sigma_1, \sigma_2])} = \Xi \int_{\sigma_1}^{\sigma_2} \phi(\delta, \sigma) d\sigma$$

при некотором $\delta \in (0, \Xi)$.

Применим Следствие 7.12 с $\rho_0 > 0$ таким, что $\rho_0 > \max_S \rho(\eta)$, $\bar{\delta} = \Xi$, а также с функцией Ψ , определенной как $\Psi(\sigma) = \phi(\delta, \sigma)$ при $\sigma \in [\sigma_1, \sigma_2]$, и $\Psi(\sigma) = 0$ в противном случае. Таким образом, существует функция $\tilde{\psi}$, для которой справедливы соотношения $-\Delta \tilde{\psi} + \rho_0 \tilde{\psi} = 0$ и

$$\int_{\sigma_1 - \bar{a}}^{\sigma_2 + \bar{a}} \tilde{\psi}_\xi(0, \sigma) d\sigma \geq C'_0(\Xi) \int_{\sigma_1}^{\sigma_2} \phi(\delta, \sigma) d\sigma.$$

Отметим, что вышеописанный выбор ρ_0 обеспечивает $-\Delta \phi + \rho_0 \phi \geq 0$. Тогда, в силу принципа максимума, $\phi \geq \tilde{\psi}(\xi, \sigma)$, и следовательно $\phi_\xi(0, \sigma) \geq \tilde{\psi}_\xi(0, \sigma)$, что приводит к

$$\int_{\sigma_1 - \bar{a}}^{\sigma_2 + \bar{a}} \phi_\xi(0, \sigma) d\sigma \geq C'_0(\Xi) \int_{\sigma_1}^{\sigma_2} \phi(\delta, \sigma) d\sigma = \Xi^{-1} C'_0(\Xi) \|\phi\|_{L_1([0, \Xi] \times [\sigma_1, \sigma_2])}.$$

Таким образом получаем желаемый результат с $\bar{C} = \Xi^{-1} C'_0(\Xi)$. \square

Следующая лемма описывает еще одну нижнюю оценку для интеграла $\int_{\Gamma_{R', R}} \phi_\xi$ из (7.36).

Л е м м а 7.14. *Найдутся такие независящие от R положительные числа R' и C^* , что, если $\lambda_{R,1} < \frac{1}{2} \tilde{B}_t(O, 0)$, то $R > 2R'$ и*

$$\int_{\Gamma_{R', R}} \phi_\xi \geq C^* \max_{S_R} \phi.$$

Д о к а з а т е л ь с т в о. Достаточно рассмотреть случай такой нормализации ϕ , что $\max_{S_R} \phi = 1$. Пусть этот максимум достигается в (ξ^*, σ^*) , при этом, конечно, $\xi^* = \xi^*(R)$ и $\sigma^* = \sigma^*(R)$. Поскольку ϕ симметрична по отношению к Γ_R и Γ_R^- , найдется такая точка (ξ^*, σ^*) , расположенная ближе к Γ . Отметим, что $\xi^* \leq \Xi$, поскольку, в силу принципа максимума, ϕ не может достигать своего положительного максимума в $S(\Xi)$, где $-\Delta\phi + \rho\phi = 0$ при $\rho > 0$. Далее, объединяя $\Delta\phi = \rho\phi$ с $\phi \leq 1$ в S_R , получаем $|\nabla\phi| \leq C$ в S_R , где C не зависит от R . Поэтому, выбирая $\delta = 1/(4C)$, получаем $\phi(\xi^*, \sigma) \geq \frac{1}{2}$ при $\sigma \in [\sigma^*, \sigma^* + 2\delta]$. Следовательно $R > \sigma^* + 2\delta \geq 2\delta$, прямоугольник $(0, \xi^*) \times (\sigma^*, \sigma^* + 2\delta)$ находится в S_R , а на его границе $\phi(\xi, \sigma)$ удовлетворяет $\phi(\xi^*, \sigma) \geq \frac{1}{2} \geq \frac{1}{4} [\cos([\sigma - (\sigma^* + \delta)]\pi/\delta) + 1]$, $\phi(\xi, (\sigma^* + \delta) \pm \delta) \geq 0$, $\phi(0, \sigma) = 0$.

Мы утверждаем, что справедлива оценка

$$\int_{\delta}^R \phi_{\xi}(0, \sigma) d\sigma \geq \int_{\sigma^* + \delta}^{\sigma^* + 2\delta} \phi_{\xi}(0, \sigma) d\sigma \geq C^*, \quad (7.38)$$

из которой немедленно вытекает утверждение леммы с $R' = \delta$.

Чтобы доказать (7.38), положим $\sigma' = \sigma - (\sigma^* + \delta)$, и пусть $\rho_0 > 0$ удовлетворяет $\rho_0 > \max_S \rho(\eta)$, а $\kappa = \sqrt{\rho_0 + (\pi/\delta)^2}$. Рассмотрим барьерную функцию

$$\psi(\xi, \sigma') := [\cos(\sigma'\pi/\delta) + 1] \frac{\sinh(\kappa\xi)}{\sinh(\kappa\xi^*)} \quad (\xi, \sigma') \in [0, \xi^*] \times [-\delta, \delta].$$

Очевидно, что

$$\psi \geq 0, \quad \psi(\xi, \pm\delta) = \psi(0, \sigma') = 0, \quad \psi(\xi^*, \sigma') = \cos(\sigma'\pi/\delta) + 1,$$

а также

$$-\Delta\psi + \rho_0\psi = -(\pi/\delta)^2 \frac{\sinh(\kappa\xi)}{\sinh(\kappa\xi^*)} \leq 0.$$

Наконец, заметим, что $\xi^* \leq \Xi$,

$$\left. \frac{\partial\psi}{\partial\xi} \right|_{\xi=0} = [\cos(\sigma'\pi/\delta) + 1] \frac{\kappa}{\sinh(\kappa\xi^*)} \geq [\cos(\sigma'\pi/\delta) + 1] \frac{\kappa}{\sinh(\kappa\Xi)},$$

Поэтому

$$\left. \frac{\partial \psi}{\partial \xi} \right|_{\xi=0, |\sigma'| \leq \delta/2} \geq \frac{\kappa}{\sinh(\kappa \Xi)}.$$

Далее, при $(\xi, \sigma') \in [0, \xi^*] \times [-\delta, \delta]$ можно показать, что $\phi \geq \frac{1}{4}\psi$. Это следует из $-\Delta \phi + \rho_0 \phi \geq -\Delta \phi + \rho \phi = 0$ и граничных условий на ϕ в силу принципа максимума, применяемого в области $(\xi, \sigma') \in [0, \xi^*] \times [-\delta, \delta]$. Поэтому $\phi_\xi|_{\xi=0, |\sigma'| \leq \delta/2} \geq \frac{1}{4}\kappa/\sinh(\kappa \Xi)$, и следовательно

$$\int_{\sigma^* + \delta}^{\sigma^* + 2\delta} \phi_\xi(0, \sigma) d\sigma \geq \int_0^{\delta/2} \phi_\xi(0, \sigma') d\sigma' \geq \frac{\kappa \delta}{8 \sinh(\kappa \Xi)} = C^* > 0.$$

Таким образом оценка (7.38) установлена. \square

Т е о р е м а 7.15. *Для главного собственного значения в области S_R справедлива оценка $\lambda_{R,1} \geq C > 0$, где постоянная C не зависит от R .*

Д о к а з а т е л ь с т в о. При $\lambda_{1,R} \geq \frac{1}{2}\tilde{B}_t(O, 0)$ желаемое утверждение очевидно. Остается рассмотреть случай $\lambda_{1,R} < \frac{1}{2}\tilde{B}_t(O, 0)$. При этом пусть R' выбрано как в Лемме 7.14. Рассмотрим максимально большую область типа $\Omega_0 = \{(\xi, \sigma) \in [0, \Xi] \times [\sigma_1, \sigma_2]\}$, для которой справедливо $\Omega'_0 = \{(\xi, \sigma) \in [0, \Xi] \times [\sigma_1 - \bar{a}, \sigma_2 + \bar{a}]\} \subset S_R \setminus S_{R'}$. Тогда, в силу Леммы 7.13, получаем

$$\int_{\Gamma_{R',R}} \phi_\xi \geq \int_{\sigma_1 - \bar{a}}^{\sigma_2 + \bar{a}} \phi_\xi(0, \sigma) d\sigma \geq \bar{C} \|\phi\|_{L_1(\Omega_0)}.$$

Воспользовавшись также Леммой 7.14, заключаем, что

$$\int_{\Gamma_{R',R}} \phi_\xi \geq C [\|\phi\|_{L_1(\Omega_0)} + \max_{S_R} \phi]. \quad (7.39)$$

Заметим, что подобласть $S_R \cap \{\xi \leq \Xi\} \setminus \Omega_0$ имеет площадь $O(1)$, следовательно

$$\|\phi\|_{L_1, S_R \cap \{\xi \leq \Xi\}} \leq \|\phi\|_{L_1(\Omega_0)} + C \max_{S_R} \phi. \quad (7.40)$$

Объединяя (7.36), (7.39) и (7.40), получаем утверждение теоремы. \square

Л е м м а 7.16. *Найдется такая независящая от R постоянная $C > 0$, что, если $R > 0$ и $F \in L_2(S_R)$, то задача*

$$\mathcal{M}W = F \quad \text{в } S_R, \quad W = 0 \quad \text{на } \partial S_R \quad (7.41)$$

имеет решение W , для которого справедлива оценка

$$\|W\|_{L_\infty(S_R)} + \|W\|_{H^2(S_R)} \leq C\|F\|_{L_2(S_R)}. \quad (7.42)$$

При этом, если $|F(\eta)| \leq Ce^{-c|\eta|}$, то $|W(\eta)| \leq C'e^{-c'|\eta|}$.

Д о к а з а т е л ь с т в о. Поскольку \mathcal{M} является самосопряженным оператором в $L_2(S_R)$, разрешимость нашей краевой задачи и соотношение $\|W\|_{L_2(S_R)} \leq C\|F\|_{L_2(S_R)}$ следуют из Теоремы 7.15 и разложения по собственным функциям. Перепишем наше дифференциальное уравнение как $-\Delta W = F_1 := -\tilde{a}W + F$. Тогда функция $F_1 \in L_2(R)$ ограничена в норме L_2 равномерно по R . Из "второго фундаментального неравенства"[32, Лемма 8.1] и выпуклости сектора S_R получаем соотношение $\|W\|_{H^2(S_R)} \leq C\|F_1\|_{L_2(S_R)} \leq C\|F\|_{L_2(S_R)}$, которой является одним из неравенств в (7.42). Воспользовавшись неравенством Соболева, также получаем оставшееся неравенство в (7.42).

Чтобы показать экспоненциальное убывание, заметим, что разложение W по собственным функциям \mathcal{M} приводит к

$$\lambda_{R,1}\|W\|_{L_2(S_R)}^2 \leq (\mathcal{M}W, W),$$

$$\|\nabla W\|_{L_2(S_R)}^2 = (\mathcal{M}W, W) - (\tilde{a}W, W) \leq C(\mathcal{M}W, W).$$

Поэтому получаем

$$\|W\|_{H^1(S_R)}^2 \leq C(\mathcal{M}W, W) \quad \text{при } W \in H_0^1(S_R) \quad (7.43)$$

с независящей от R постоянной C .

Заметим, что $|\eta| \cos(\omega/2) \leq \bar{\xi} \leq |\eta|$, где переменная $\bar{\xi}$ имеет то же значение, что и в (7.33). Поэтому достаточно установить свойство экспоненциального убывания, описанное в лемме, при $|\eta|$ замененном на $\bar{\xi}$. Предположим,

что F удовлетворяет $|F(\eta)| \leq Ce^{-c\bar{\xi}}$. Пусть $\kappa \in (0, c)$, и положим $\tilde{W} = e^{\kappa\bar{\xi}}W$, $\tilde{F} = e^{\kappa\bar{\xi}}F$. Таким образом, $|\tilde{F}(\eta)| \leq Ce^{-c'\bar{\xi}}$, где $c' = c - \kappa > 0$. Функция \tilde{W} удовлетворяет соотношению

$$\mathcal{M}\tilde{W} + 2\kappa\tilde{W}_{\bar{\xi}} - \kappa^2\tilde{W} = \tilde{F}. \quad (7.44)$$

Применяя (7.43), получаем

$$\|\tilde{W}\|_{H^1(S_R)}^2 \leq (\tilde{F} - 2\kappa\tilde{W}_{\bar{\xi}} + \kappa^2\tilde{W}, \tilde{W}) \leq \|\tilde{W}\|_{L_2(S_R)}\|\tilde{F}\|_{L_2(S_R)} + \kappa^2\|\tilde{W}\|_{L_2(S_R)}^2,$$

где мы также воспользовались $(\tilde{W}, \tilde{W}_{\bar{\xi}}) = 0$. Выбирая κ достаточно малым и воспользовавшись неравенством для среднего геометрического, получаем

$$\|\tilde{W}\|_{H^1(S_R)} \leq C\|\tilde{F}\|_{L_2(S_R)}.$$

Положим $\tilde{F}_1 = \tilde{F} - 2\kappa\tilde{W}_{\bar{\xi}} + \kappa^2\tilde{W}$, тогда $\|\tilde{F}_1\|_{L_2(S_R)} \leq C\|\tilde{F}\|_{L_2(S_R)}$. Уравнение (7.44) становится $\mathcal{M}\tilde{W} = \tilde{F}_1$, а неравенство (7.42), примененное к этому уравнению, приводит к $|\tilde{W}(\eta)| \leq C\|\tilde{F}\|_{L_2(S_R)}$ при $\eta \in S_R$. Следовательно $|W(\eta)| \leq Ce^{-\kappa\bar{\xi}}\|\tilde{F}\|_{L_2(S_R)} \leq Ce^{-\kappa\bar{\xi}}$ при $\eta \in S_r$. Отметим, что все постоянные в этих неравенствах не зависят от η . \square

Т е о р е м а 7.17. *Найдется такая постоянная $C > 0$, что, если $F \in L_2(S)$, то задача (7.31) имеет решение W , удовлетворяющее соотношению*

$$\|W\|_{L_\infty(S)} + \|W\|_{H^2(S)} \leq C\|F\|_{L_2(S)}. \quad (7.45)$$

При этом, если $|F(\eta)| \leq Ce^{-c|\eta|}$, тогда $|W(\eta)| \leq C'e^{-c'|\eta|}$ при $c' < c$.

Д о к а з а т е л ь с т в о. Выберем последовательность $R_j \rightarrow \infty$ и пусть W_j обозначает последовательность соответствующих решений (7.41). В силу компактности, можно выбрать подпоследовательность последовательности W_j , которую продолжим обозначать W_j , а также функцию $W \in H^1(S)$ такую, что W_j слабо сходится к W в $H^1(S)$. Пусть $\chi \in C_0^\infty(S)$, и возьмем предел соотношения

$$\iint_{S_{R_j}} \nabla W_j \cdot \nabla \chi + \tilde{a}W_j\chi = \iint_{S_{R_j}} F\chi,$$

из чего мы видим, что W является решением (7.31). При каждом $R > 0$ получаем

$$\|W\|_{H^2(S_R)} \leq \liminf \|W_j\|_{H^2(S_R)} \leq C\|F\|_{L_2(S)}.$$

Полагая $R \rightarrow \infty$, получаем второе неравенство в (7.45). Первое неравенство следует из неравенства Соболева. Утверждение, касающееся экспоненциального убывания, доказывается аналогично Лемме 7.16. \square

7.3.5 Существование и экспоненциальное убывание q_1 и $\tilde{q}_{0,p}$

Формула (7.17b) дает определение функции q_1 с помощью z_1 в предположении, что краевая задача (7.15) для z_1 имеет решение. Рассмотрим этот вопрос несколько с другой стороны. Из (7.17b) и (7.15) мы уже получили краевую задачу (7.19) для q_1 . Данные в этой задаче убывают экспоненциально. Это позволит нам установить существование q_1 , а далее, воспользовавшись (7.17b), определить z_1 .

Л е м м а 7.18. *Решение q_1 задачи (7.19) существует и экспоненциально убывает в S , т. е. $|q_1| \leq Ce^{-c|\eta|}$ при некоторых $C, c > 0$.*

Д о к а з а т е л ь с т в о. Применим Теорему 7.17 к версии задачи (7.19) для некоторой вспомогательной функции \hat{q}_1 , которая конструируется вычитанием из q_1 гладкой экспоненциально-убывающей функции, удовлетворяющей краевым условиям в (7.19), с тем чтобы $\hat{q}_1|_{\partial S} = 0$. Для этого достаточно показать, что q_1 экспоненциально убывает на Γ и Γ^- , а $\mathcal{M}q_1$ экспоненциально убывает в S . Экспоненциальное убывание граничных данных (7.19) и их производных следует из неравенств в Лемме 7.1, а также из того, что $\xi = |\eta| \sin \omega$ на Γ^- и $\xi^- = |\eta| \sin \omega$ на Γ . Далее мы покажем, что найдутся такие положительные постоянные c_1 и C_1 , что справедливо

$$|\mathcal{M}q_1| \leq C_1 e^{-c_1|\eta|}. \quad (7.46)$$

Воспользуемся (7.19) и обозначим 3 слагаемых в правой части $\mathcal{M}q_1$ как I , II , и II^- . Вспомним, что $z_0 = q_0 + \dot{v}_0 + \dot{v}_0^-$, и применим (7.7) к I , что приводит к $|I| \leq C|\eta|(|q_0| + \dot{v}_0\dot{v}_0^-)$. Для II и II^- справедливы $|II| \leq |\dot{v}_1 + \sigma\dot{v}_{0,s}|(|q_0| + \dot{v}_0^-)$ и $|II^-| \leq |\dot{v}_1^- + \sigma^-\dot{v}_{0,s}^-|(|q_0| + \dot{v}_0)$. Объединяя эти оценки с Леммой 7.8 и Леммой 7.1, получаем экспоненциальное убывание всех слагаемых I , II , и II^- , и следовательно желаемое соотношение (7.46). \square

Формально дифференцируя обе стороны (7.12) по p и воспользовавшись (7.4), получаем линейную краевую задачу для $\tilde{z}_{0,p}$. Однако данные в этом уравнении не являются квадратично интегрируемыми, поэтому Теорема 7.17 здесь неприменима. Вместо этого удобнее работать с \tilde{q}_0 . Краевая задача для $\tilde{q}_{0,p}$ дана в (7.20).

Л е м м а 7.19. Решение $\tilde{q}_{0,p}$ задачи (7.20) существует и экспоненциально убывает в S .

Д о к а з а т е л ь с т в о. Применим Теорему 7.17 к задаче (7.20). Для этого нужно показать, что $\tilde{q}_{0,p}$ экспоненциально убывает на Γ и Γ^- , а также что правая часть экспоненциально убывает в S . Это доказывается аналогично предыдущей лемме. \square

7.4 Возмущенное асимптотическое разложение, верхние и нижние решения, доказательство существования

В §7.2.1 мы определили погранслойные функции $\tilde{v} = \tilde{v}_0 + \varepsilon v_1$, ассоциируемые со стороной Γ области Ω , а в §7.2.2 были определены угловые погранфункции $\tilde{q} = \tilde{q}_0 + \varepsilon \tilde{q}_1$, ассоциируемые с вершиной P_{j-1} области Ω . Теперь мы готовы определить возмущенное асимптотическое разложение β_{S_j} , ассоциируемое с вершиной P_{j-1} многоугольника Ω . Эти функции далее будут собраны в асимптотическое разложение β_Ω для исходной задачи (7.1). При значении параметра $p = 0$ получаем стандартное асимптотическое разложе-

ние, для которого ниже будет установлена точность второго порядка. Возмущенные же асимптотические разложения (при $p \neq 0$) используются при построении нижних и верхних решений. Последние затем используются при доказательстве существования решения задачи (7.1).

Воспользуемся (7.10) и (7.17) при определении возмущенного асимптотического разложения β_{S_j} и более стандартного асимптотического разложения u_{as,S_j} , ассоциируемых с вершиной P_{j-1} :

$$\begin{aligned}\beta_{S_j}(x; p) &= u_0(x) + \tilde{v}(\xi, s; p) + \tilde{v}^-(\xi^-, s^-; p) + \tilde{q}(\eta; p) + \theta p, \\ u_{\text{as},S_j}(x) &= \beta_{S_j}(x; 0) = u_0(x) + v(\xi, s) + v^-(\xi^-, s^-) + q(\eta),\end{aligned}\tag{7.47}$$

где $x \in S_j$, а переменные ξ, ξ^-, s, s^-, η ассоциируются с сектором S_j с вершиной P_{j-1} (см. Рис. 7.1). Значение положительного параметра θ и область значений p будут определены ниже. Функции u_{as,S_j} будут использованы при построении глобального асимптотического разложения. Далее мы приводим Лемму 3.1 из [160].

Л е м м а 7.20. *Справедливо, что $Fu_{\text{as},S_j} = O(\varepsilon^2)$ при любом $x \in S_j$. При этом $u_{\text{as},S_j}(x) = g(x) + O(\varepsilon^2)$ на ∂S_j , где $g(x)$ продолжена с $\Gamma_{j-1} \cup \Gamma_j$ на ∂S_j , как описано в §7.2.2.*

Л е м м а 7.21. *Существуют такие положительные числа $\theta, \varepsilon^*, p^*, c_1$ и c_2 , что при $\varepsilon \leq \varepsilon^*$ и $|p| \leq p^*$ справедливо $|\beta_{S_j}(x; p) - u_{\text{as},S_j}| \leq Cp$, а также*

$$\beta_{S_j}(x; -p) \leq \beta_{S_j}(x; p) \quad \text{при } p > 0, \tag{7.48a}$$

$$F\beta_{S_j} \geq \frac{1}{2}\theta\gamma^2 p - c_1\varepsilon^2 \quad \text{при } p > 0, \tag{7.48b}$$

$$F\beta_{S_j} \leq -\frac{1}{2}\theta\gamma^2 |p| + c_1\varepsilon^2 \quad \text{при } p < 0, \tag{7.48c}$$

$$(-1)^{\text{sgn } p} \beta_{S_j} \geq (-1)^{\text{sgn } p} g + \frac{1}{2}\theta |p| - c_2\varepsilon^2 \quad \text{на } \partial S_j \quad \text{при } p \neq 0. \tag{7.48d}$$

Д о к а з а т е л ь с т в о. Неравенства (7.48b), (7.48c) получены в Лемме 4.4 из [160]. Отметим, что при их доказательстве существенно, что положительный параметр θ в определении (7.47) для β_{S_j} выбран достаточно

малым и удовлетворяет $0 < \theta \leq |\lambda(x)|^{-1}$, где $\lambda(x) = b_{uu}(x, u_0(x) + \vartheta[v_0 + v_0^- + q_0])$ с некоторым $\vartheta = \vartheta(x) \in (0, 1)$.

Далее воспользуемся Леммой 4.1 из [160], из которой вытекает, что $\beta_{S_j} = u_{as, S_j} + O(p)$, а также что при достаточно малом $\varepsilon^* > 0$, если $\varepsilon \leq \varepsilon^*$ и $p \geq 0$, то

$$\beta_{S_j}(x; -p) \leq u_{as, S_j}(x) - \frac{1}{2}\theta p \leq u_{as, S_j}(x) + \frac{1}{2}\theta p \leq \beta_{S_j}(x; p). \quad (7.49)$$

Это немедленно приводит к (7.48a). При этом при $x \in \partial S_j$ и $p > 0$, объединяя (7.49) с оценкой $u_{as, S_j}(x) = g(x) + O(\varepsilon^2)$ из Леммы 7.20, получаем $\beta_{S_j}(x; p) \geq g(x) - c_2\varepsilon^2 + \frac{1}{2}\theta p$, а также $\beta_{S_j}(x; -p) \leq g(x) + c_2\varepsilon^2 - \frac{1}{2}\theta p$ при $p < 0$. Таким образом последнее утверждение (7.48d) доказано. \square

Чтобы определить соответствующее возмущенное асимптотическое разложение для всей области Ω , нам потребуются неотрицательные гладкие функции $\{\chi_j\}_{j=1}^M: \bar{\Omega} \rightarrow [0, 1]$, которые удовлетворяют соотношениям

$$\chi_j(P_{j-1}) = 1, \quad \chi_j(x) + \chi_{j+1}(x) = 1 \quad \text{на } \bar{\Gamma}_j, \quad \sum_{j=1}^M \chi_j(x) = 1 \quad \text{на } \bar{\Omega}.$$

Определим возмущенное асимптотическое разложение β_Ω для задачи (7.1) следующим образом:

$$\beta_\Omega(x; p) = \sum_{j=1}^M \chi_j(x) \beta_{S_j}(x; p),$$

$$u_{as, \Omega}(x) = \beta_\Omega(x; 0).$$

Л е м м а 7.22. *Найдутся такие положительные числа $\theta, \varepsilon^*, p^*, c_1$ и c_2 , что при $\varepsilon \leq \varepsilon^*$ и $|p| \leq p^*$ справедливо $|\beta_\Omega(x; p) - u_{as, \Omega}| \leq Cp$, а также*

$$\beta_\Omega(x; -p) \leq \beta_\Omega(x; p) \quad \text{при } p > 0, \quad (7.50a)$$

$$F\beta_\Omega \geq \frac{1}{2}\theta\gamma^2 p - c_1\varepsilon^2 \quad \text{при } p > 0, \quad (7.50b)$$

$$F\beta_\Omega \leq -\frac{1}{2}\theta\gamma^2 |p| + c_1\varepsilon^2 \quad \text{при } p < 0, \quad (7.50c)$$

$$(-1)^{\text{sgn} p} \beta_\Omega \geq (-1)^{\text{sgn} p} g + \frac{1}{2}\theta |p| - c_2\varepsilon^2 \quad \text{на } \partial\Omega \quad \text{при } p \neq 0. \quad (7.50d)$$

Д о к а з а т е л ь с т в о. Доказательство каждого из неравенств леммы опирается на неотрицательность функций $\{\chi_j\}_{j=1}^M$ и соответствующее неравенство в Лемме 7.21.

Остановимся поподробнее на получении (7.50b) и (7.50c), получение которых несколько осложняется тем, что функция F нелинейна. В случае $\chi_j(x) = 1$ при некотором j , имеем $\beta_\Omega(x, p) = \beta_{S_j}(x, p)$, поэтому (7.50b), (7.50c) очевидны. В противном случае x находится на конечном расстоянии от любой вершины Ω . Теперь, если $\chi_j(x) + \chi_{j+1}(x) = 1$ при некотором j , то экспоненциальное убывание \tilde{q} приводит к $\beta_{S_{j+1}}(x, p) = \beta_{S_j}(x, p) + O(e^{-c/\varepsilon})$, поэтому $\beta_\Omega(x, p) = \beta_{S_j}(x, p) + O(e^{-c/\varepsilon})$, и следовательно $F\beta_\Omega(x, p) = F\beta_{S_j}(x, p) + O(e^{-c/\varepsilon})$, т. е. в этом случае справедливы (7.50b), (7.50c) с возможно несколько большей постоянной c_1 , чем в (7.48b), (7.48c). Если же $\chi_j(x) + \chi_{j+1}(x) \neq 1$ для любого j , то x находится на конечном расстоянии от $\partial\Omega$, поэтому экспоненциальное убывание \tilde{q} и \tilde{v} приводит к $\beta_{S_j}(x, p) = u_0(x) + \theta p + O(e^{-c/\varepsilon})$ для любого j , следовательно справедливо $\beta_\Omega(x, p) = \beta_{S_1}(x, p) + O(e^{-c/\varepsilon})$, и мы снова получаем (7.50b), (7.50c). \square

Следующий результат является основным в настоящем параграфе.

Т е о р е м а 7.23. Пусть b удовлетворяет А1–А4. Тогда найдется такая положительная постоянная ε^* , что при любом $\varepsilon \in (0, \varepsilon^*]$ задача (7.1) имеет решение $u(x)$, удовлетворяющее $|u(x) - u_{\text{as},\Omega}(x)| \leq C\varepsilon^2$.

Д о к а з а т е л ь с т в о. Сначала, выберем ε^* в Лемме 7.22 достаточно малым, чтобы были справедливы $c_1(\varepsilon^*)^2 \leq \frac{1}{2}\theta\gamma^2 p^*$ и $c_2(\varepsilon^*)^2 \leq \frac{1}{2}\theta p^*$. При любом $\varepsilon \leq \varepsilon^*$, выберем $\bar{p} = \max\{c_1\varepsilon^2/(\frac{1}{2}\theta\gamma^2), c_2\varepsilon^2/(\frac{1}{2}\theta)\}$. Далее, в силу Леммы 7.22, функции $\beta_\Omega(x; \bar{p})$ и $\beta_\Omega(x; -\bar{p})$ соответственно являются верхним и нижним решениями для задачи (7.1), при этом $\beta_\Omega(x; -\bar{p}) \leq \beta_\Omega(x; \bar{p})$. Следовательно, в силу Леммы 7.4, существует решение u задачи (7.1) расположенное между $\beta_\Omega(x; \bar{p})$ и $\beta_\Omega(x; -\bar{p})$. Далее, в силу Леммы 7.22, получаем оценку $|\beta_\Omega(x; \pm\bar{p}) - u_{\text{as},\Omega}(x)| \leq C\bar{p} \leq C\varepsilon^2$, которая ведет к желаемой оценке $|u(x) - u_{\text{as},\Omega}(x)| \leq C\varepsilon^2$. \square

7.5 Некоторые аспекты построения ε -равномерных численных методов

Как уже отмечалось, наш интерес к асимптотическому анализу задачи (7.1) обусловлен тем, что построенное асимптотическое разложение и соответствующие верхние и нижние решения могут быть использованы при построении и теоретическом исследовании ε -равномерных численных методов для рассматриваемой задачи (аналогично §6).

Заметим, что значения коэффициентов экспоненциального убывания (как, например, γ в последней оценке Леммы 7.1 для \tilde{v}_0) нередко являются весьма существенными при априорном построении сгущающихся в пограничных слоях сеток (см., например использование коэффициента экспоненциального убывания γ при построении сеток Бахвалова и Шишкина в §6.3.1(a),(b)). Это связано с тем, что данные коэффициенты явным образом присутствуют в определениях некоторых параметров таких сеток, в то время как правильный выбор этих параметров играет ключевую роль в обеспечении определенного порядка ε -равномерной сходимости численного метода (в контексте сеток Шишкина данный аспект обсуждается, например, в [49, 178]).

Отметим также, что в полученных в настоящем параграфе оценках для угловых погранфункций точные значения коэффициентов экспоненциального убывания (как, например, c_1 в оценке Леммы 7.8 для \tilde{q}_0) не описаны. Тем не менее, даже в этом случае можно построить сгущающиеся в пограничных слоях сетки, которые приведут к гарантированному порядку равномерной по ε сходимости в норме максимума модуля. Например, в случае сетки типа Шишкина, достаточно модифицировать параметр сетки $\sigma = 2\gamma^{-1}\varepsilon \ln N$ из §6.3.1(b) (где он выбирается с тем, чтобы $e^{-\gamma\sigma/\varepsilon} = N^{-2}$) следующим образом. Полагая $\sigma = \varepsilon \ln^{1+\kappa} N$, с любым фиксированным $\kappa > 0$, получаем $e^{-c_1\sigma/\varepsilon} = N^{-c_1 \ln^\kappa N} \leq CN^{-r}$ при любом ожидаемом порядке сходимости r (последний же определяется свойствами используемой дискретизации дифференциального уравнения, и не должен ухудшаться при адекватном выборе

сетки). Другими словами, при любом фиксированном значении $\kappa > 0$, можно ожидать, что вышеописанная модификация параметра сетки Шишкина σ приведет равномерной по ε сходимости порядка r (с соответствующим логарифмическим множителем) в норме максимума модуля. Разумеется, при этом мы предполагаем, что используется некоторая разумная дискретизация дифференциального уравнения (например, подобная дискретизации, описанной в §6, которая соответствует $r = 2$).

Заключение

Работа посвящена теоретическому исследованию конечноэлементных аппроксимаций некоторых полулинейных сингулярно возмущенных уравнений реакции-диффузии. Поскольку для таких задач характерны решения с пограничными и внутренними слоями, особое внимание уделяется случаю анизотропных сеток и оценкам ошибки в норме максимума модуля. Для рассмотренных конечноэлементных аппроксимаций установлен ряд новых апостериорных и априорных оценок.

Перечислим основные результаты работы.

1. Для полулинейных сингулярно возмущенных эллиптических уравнений реакции-диффузии в полигональных областях установлены явные апостериорные оценки ошибки численного решения на основе невязок в норме максимума модуля на локально квазиравномерных сетках. Постоянные в полученных оценках не зависят от диаметров элементов сетки и малого параметра.

2. Для полулинейных сингулярно возмущенных эллиптических уравнений реакции-диффузии в многоугольных областях получены явные апостериорные оценки ошибки численного решения в норме максимума модуля на неструктурированных анизотропных сетках. Постоянные в полученных оценках не зависят от диаметров элементов сетки, их аспектного соотношения (т. е. степени их сплюснутости) и малого параметра. Полученные апостериорные оценки являются новыми даже для уравнения Лапласа. Представленный подход обобщен для получения апостериорных оценок ошибки в энергетической норме.

3. Для полулинейных параболических уравнений второго порядка получены апостериорные оценки в норме максимума модуля для ошибки соответствующих численных решений. Рассмотрены временные полудискретизации

и полностью дискретные методы на основе неявного метода Эйлера, метода Кранка-Николсон и разрывного метода Галеркина $dG(r)$ с квадратурой Радо.

4. Для немонотонных полулинейных сингулярно возмущенных эллиптических уравнений реакции-диффузии в гладких областях исследованы численные решения на сгущающихся в пограничных слоях сетках типа Бахвалова и Шишкина. Доказано существование решений соответствующих нелинейных дискретных задач и установлен второй порядок сходимости (с логарифмическим множителем в случае сетки Шишкина) в сеточной норме максимума модуля равномерно по малому параметру ε при условии $\varepsilon \leq Ch$, при этом как число степеней свободы не превосходит Ch^{-2} .

5. Для немонотонного полулинейного сингулярно возмущенного эллиптического уравнения реакции-диффузии в выпуклой многоугольной области построено асимптотическое разложение и установлено существование решения исходной задачи в окрестности построенного асимптотического разложения.

Список литературы

1. Андреев В. Б. Логарифм в оценке L_∞ -сходимости метода конечных элементов на линейных треугольниках необходим // *Доклады семинара Института прикладной математики им. И. Н. Векуа*. 1989. Т. 4. N 3. С. 17–20.
2. Андреев В. Б. О точности сеточных аппроксимаций негладких решений сингулярно возмущенного уравнения реакции-диффузии в квадрате // *Дифференц. уравнения*. 2006. Т. 42. С. 895–906.
3. Андреев В. Б. Равномерная сеточная аппроксимация негладких решений смешанной краевой задачи для сингулярно возмущенного уравнения реакции-диффузии в прямоугольнике // *Ж. вычисл. матем. и матем. физ.* 2008. Т. 48. С. 90–114.
4. Андреев В. Б. Равномерная сеточная аппроксимация негладких решений сингулярно возмущенного уравнения конвекции-диффузии в прямоугольнике // *Дифференц. уравнения*. 2009. Т. 4. С. 954–964.
5. Андреев В. Б., Савин И. А. О равномерной по малому параметру сходимости монотонной схемы Самарского и ее модификации // *Ж. вычисл. матем. и матем. физ.* 1995. Т. 35. С. 739–752.
6. Багаев Б. М., Карпова Е. Д., Шайдуров В. В. *Сеточные методы решения задач с пограничным слоем*. Новосибирск: Наука, 1998. Ч. 1.
7. Багаев Б. М., Карпова Е. Д., Шайдуров В. В. *Сеточные методы решения задач с пограничным слоем*. Новосибирск: Наука, 2001. Ч. 2.
8. Бахвалов Н. С. К оптимизации методов решения краевых задач при наличии пограничного слоя // *Ж. вычисл. матем. и матем. физ.* 1969. Т. 9. N 4. С. 841–859.

9. Блатов И. А. О проекционном методе для сингулярно возмущенных краевых задач// *Ж. вычисл. матем. и матем. физ.* 1990. Т. 30. С. 1031–1044.
10. Блатов И. А. О методе конечных элементов Галеркина для эллиптических квазилинейных сингулярно возмущенных краевых задач. I// *Дифференц. уравнения.* 1992. Т. 28. С. 1168–1177.
11. Блатов И. А. О методе конечных элементов Галеркина для эллиптических квазилинейных сингулярно возмущенных краевых задач. II// *Дифференц. уравнения.* 1992. Т. 28. С. 1799–1810.
12. Блатов И. А. О методе конечных элементов Галеркина для эллиптических квазилинейных сингулярно возмущенных краевых задач. III. Задачи с угловыми погранслоями// *Дифференц. уравнения.* 1994. Т. 30. С. 467–479.
13. Блатов И. А., Стрыгин, В. В. *Элементы теории сплайнов и метод конечных элементов для задач с погранслоем.* Воронеж: ВГУ, 1997.
14. Боглаев И. П. Численный метод решения квазилинейного эллиптического уравнения с малым параметром при старших производных// *Ж. вычисл. матем. и матем. физ.* 1988. Т. 28. С. 492–502.
15. Бутузов В. Ф. Асимптотика решения уравнения $\mu^2 \Delta u - k^2(x, y)u = f(x, y)$ в прямоугольной области// *Дифференц. уравнения.* 1973. Т. 9. С. 1654–1660.
16. Вазов В. *Асимптотические разложения решений обыкновенных дифференциальных уравнений.* М.: Мир, 1968.
17. Василевский Ю. В., Данилов А. А., Липников К. Н., Чугунов В. Н. *Автоматизированные технологии построения неструктурированных расчетных сеток.* М.: Физматлит, 2016.

18. Васильева А. Б., Бутузов В. Ф. *Асимптотические разложения решений сингулярно возмущенных уравнений*. М.: Наука, 1973.
19. Васильева А. Б., Бутузов В. Ф. *Асимптотические методы в теории сингулярных возмущений*. М.: Высш. шк., 1990.
20. Васильева А. Б., Бутузов В. Ф., Нефедов, Н. Н. Сингулярно возмущенные задачи с пограничными и внутренними слоями// *Тр. МИАН*. 2010. Т. 268. С. 258–273.
21. Вишик М. И., Люстерник Л. А. Регулярное вырождение и пограничный слой для линейных дифференциальных уравнений с малым параметром// *Успехи матем. наук*. 1957. Т. 12. Вып. 5(77). С. 3–122.
22. Гринберг, Г. А. *Избранные вопросы математической теории электрических и магнитных явлений*. М.-Л.: Издательство Академии наук СССР, 1948.
23. Денисов И. В. Квазилинейные сингулярно возмущенные эллиптические уравнения в прямоугольнике// *Ж. вычисл. матем. и матем. физ.* 1995. Т. 35. С. 1666–1678.
24. Денисов И. В. Задача нахождения главного члена угловой части асимптотики решения сингулярно возмущенного эллиптического уравнения с нелинейностью// *Ж. вычисл. матем. и матем. физ.* 1999. Т. 39. С. 779–791.
25. Денисов И. В. Угловой погранслоем в нелинейных сингулярно возмущенных эллиптических уравнениях// *Ж. вычисл. матем. и матем. физ.* 2001. Т. 41. С. 390–406.
26. Денисов И. В. Угловой погранслоем в немонотонных сингулярно возмущенных краевых задачах с нелинейностями// *Ж. вычисл. матем. и матем. физ.* 2004. Т. 9. С. 1674–1692.

27. Денисов И. В. Угловой погранслоей в нелинейных сингулярно возмущенных эллиптических задачах// *Ж. вычисл. матем. и матем. физ.* 2008. Т. 48. С. 62–79.
28. Дулан Э., Миллер Дж., Шилдерс У. *Равномерные численные методы решения задач с пограничным слоем*. М.: Мир, 1983.
29. Ильин А. М. Разностная схема для дифференциального уравнения с малым параметром при старшей производной// *Матем. заметки*. 1969. Т. 6. С. 237–248.
30. Ильин А. М. *Согласование асимптотических разложений решений краевых задач*. М.: Наука, 1989.
31. Ладыженская О. А., Солонников В. А., Уральцева Н. Н. *Линейные и квазилинейные уравнения параболического типа*. М.: Наука, 1967.
32. Ладыженская О. А., Уральцева Н. Н. *Линейные и квазилинейные уравнения эллиптического типа*. М.: Наука, 1973.
33. Левашова Н. Т., Нефедов Н. Н., Орлов А. О. Стационарное уравнение реакции-диффузии с разрывным реактивным членом// *Ж. вычисл. матем. и матем. физ.* 2017. Т. 57. С. 854–866.
34. Лисейкин В. Д., Яненко Н. Н. О равномерно сходящемся алгоритме численного решения обыкновенного дифференциального уравнения второго порядка с малым параметром при старшей производной// *Численные методы механики сплошной среды*, Новосибирск. 1981. Т. 12. № 2. С. 45–56.
35. Ломов С. А. *Введение в общую теорию сингулярных возмущений*. М.: Наука, 1981.

36. Ломов С. А., Ломов И. С. Основы математической теории пограничного слоя. М.: Издательство Московского университета, 2011.
37. Нефедов Н. Н. Метод дифференциальных неравенств для некоторых сингулярно возмущенных задач в частных производных// *Дифференц. уравнения*. 1995. Т. 31. С. 719–722.
38. Нефедов Н. Н. Метод дифференциальных неравенств для некоторых классов нелинейных сингулярно возмущенных задач с внутренними слоями// *Дифференц. уравнения*. 1995. Т. 31. С. 1142–1149.
39. Нефедов Н. Н. Асимптотический метод дифференциальных неравенств в исследовании периодических контрастных структур: существование, асимптотика, устойчивость// *Дифференц. уравнения*. 2000. Т. 36. С. 298–305.
40. Оганесян Л. А. Вариационно-разностный метод для двумерных эллиптических уравнений с малым параметром// *Методы аппроксимации и интерполяции*. Новосибирск, 1981. С. 108–123.
41. Самарский А. А. *Теория разностных схем*. М.: Наука, 1989.
42. Самарский А. А., Андреев В. Б. *Разностные методы для эллиптических уравнений*. М.: Наука, 1976.
43. Сьярле Ф. *Метод конечных элементов для эллиптических задач*. М.: Мир, 1980.
44. Тихонов А. Н. О зависимости решений дифференциальных уравнений от малого параметра// *Матем. сб.* 1948. Т. 22(64). С. 193–204.
45. Тихонов А. Н. Системы дифференциальных уравнений, содержащие малые параметры при производных// *Матем. сб.* 1952. Т. 31(73). С. 575–586.

46. Тихонов А. Н., Самарский А. А. *Уравнения математической физики*. М.: Наука, 1977.
47. Флетчер К. *Численные методы на основе метода Галеркина*. М.: Мир, 1988.
48. Шишкин Г. И. Аппроксимация решений сингулярно возмущенных краевых задач с параболическим пограничным слоем // *Ж. вычисл. матем. и матем. физ.* 1989. Т. 29. С. 963–977.
49. Шишкин Г. И. *Сеточные аппроксимации сингулярно возмущенных эллиптических и параболических уравнений*. Екатеринбург: РАН, УрО, 1992.
50. Abramowitz, M., Stegun, I. A., eds. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. New York: Dover Publications, Inc., 1992.
51. Ainsworth, M., Babuška, I. Reliable and robust a posteriori error estimating for singularly perturbed reaction-diffusion problems // *SIAM J. Numer. Anal.* 1999. V. 36. Pp. 331–353.
52. Ainsworth, M., Oden, J. T. *A posteriori error estimation in finite element analysis*. New York : Wiley-Interscience, 2000.
53. Ainsworth, M., Vejchodský, T. Fully computable robust a posteriori error bounds for singularly perturbed reaction-diffusion problems. // *Numer. Math.* 2011. V. 119. Pp. 219–243.
54. Akrivis, G., Makridakis, C., Nochetto, R. H. A posteriori error estimates for the Crank-Nicolson method for parabolic equations // *Math. Comp.* 2006. V. 75. Pp. 511–531.

55. Andreev, V. B. Pointwise approximation of corner singularities for singularly perturbed elliptic problems with characteristic layers// *Int. J. Numer. Anal. Model.* 2010. V. 7. Pp. 416–427.
56. Apel, T. *Anisotropic finite elements: local estimates and applications.* Stuttgart: Teubner, 1999.
57. Babuška, I., Rheinboldt, W. C. Error estimates for adaptive finite element computations// *SIAM J. Numer. Anal.* 1978. V. 15. Pp. 736–754.
58. Babuška, I., Rheinboldt, W. C. A-posteriori error estimates for the finite element method// *International Journal for Numerical Methods in Engineering.* 1978. V. 12. Pp. 1597–1615.
59. Babuška, I., Strouboulis, T. *The finite element method and its reliability.* New York: Clarendon Press, Oxford University Press, 2001.
60. Bartels, S., Müller, R. Quasi-optimal and robust a posteriori error estimates in L^∞ (L^2) for the approximation of Allen-Cahn equations past singularities// *Math. Comp.* 2011. V. 80. Pp. 761–780.
61. Blum, H., Lin, Q., Rannacher, R. Asymptotic error expansion and Richardson extrapolation for linear finite elements// *Numer. Math.* 1986. V. 49. Pp. 11–37.
62. Boman, M. *On a posteriori error analysis in the maximum norm.* Ph.D. thesis. Göthenburg: Chalmers University of Technology and Göthenburg University, 2000.
63. Brenner, S. C., Scott, L. R. *The mathematical theory of finite element methods.* New York: Springer-Verlag, 2008.
64. Brézis, H., Strauss, W. A. Semi-linear second-order elliptic equations in L^1 .// *J. Math. Soc. Japan.* 1973. V. 25. Pp. 565–590.

65. Chen, L. *iFEM: An innovative finite element method package in Matlab*. Tech. rep. University of California-Irvine, 2009.
66. Clavero, C., Gracia, J. L., O’Riordan, E. A parameter robust numerical method for a two dimensional reaction-diffusion problem// *Math. Comp.* 2005. V. 74. Pp. 1743–1758.
67. Clément, Ph. Approximation by finite element functions using local regularization// *Rev. Française Automat. Informat. Recherche Opérationnelle Sér.* 1975. V. 9. Pp. 77–84.
68. Clément, Ph., Sweers, G. Getting a solution between sub- and supersolutions without monotone iteration// *Rendiconti dell’Istituto di Matematica dell’Università Trieste*. 1987. V. 19. Pp. 189–194.
69. Grisvard, P. *Elliptic problems in nonsmooth domains*. Boston, MA: Pitman, 1985.
70. D’Annunzio, C. M. *Numerical analysis of a singular perturbation problem with multiple solutions*. Ph.D. Dissertation. University of Maryland at College Park, 1986.
71. Dari, E., Durán, R. G., Padra, C. Maximum norm error estimators for three-dimensional elliptic problems// *SIAM J. Numer. Anal.* 2000. V. 37. Pp. 683–700.
72. Dauge, M. *Elliptic boundary value problems on corner domains*. Berlin: Springer, 1988.
73. Davies, E. B. Non-Gaussian aspects of heat kernel behaviour// *J. London Math. Soc.* 1997. V. 55. Pp. 105–125.
74. Demlow, A., Georgoulis, E. Pointwise a posteriori error control for discontinuous Galerkin methods for elliptic problems// *SIAM J. Numer. Anal.* 2012. V. 50. Pp. 2159–2181.

75. Demlow, A., Lakkis, O., Makridakis, C. A posteriori error estimates in the maximum norm for parabolic problems// *SIAM J. Numer. Anal.* 2009. V. 47. Pp. 2157–2176.
76. Demlow, A., Leykekhman, D., Schatz, A. H., Wahlbin, L. B. Best approximation property in the W_∞^1 norm for finite element methods on graded meshes// *Math. Comp.* 2012. V. 81. Pp. 743–764.
77. Demlow, A., Makridakis, C. Sharply local pointwise a posteriori error estimates for parabolic problems// *Math. Comp.* 2010. V. 79. Pp. 1233–1262.
78. Dörfler, W. Uniform error estimates for an exponentially fitted finite element method for singularly perturbed elliptic equations// *SIAM J. Numer. Anal.* 1999. V. 36. Pp. 1709–1738.
79. Duong, X. T., Hofmann, S., Mitrea, D., Mitrea, M., Yan, L. Hardy spaces and regularity for the inhomogeneous Dirichlet and Neumann problems.// *Rev. Mat. Iberoam.* 2013. V. 29. Pp. 183–236.
80. Dupont, T. Mesh modification for evolution equations// *Math. Comp.* 1982. V. 39. Pp. 85–107.
81. Durán, R. G. A note on the convergence of linear finite elements// *SIAM J. Numer. Anal.* 1988. V. 25. Pp. 1032–1036.
82. Eriksson, K. An adaptive finite element method with efficient maximum norm error control for elliptic problems// *Math. Models Methods Appl. Sci.* 1994. V. 4. Pp. 313–329.
83. Eriksson, K., Johnson, C. Adaptive finite element methods for parabolic problems II: Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$ // *SIAM J. Numer. Anal.* 1995. V. 32. Pp. 706–740.

84. Eriksson, K., Johnson, C., Thomée, V. Time discretization of parabolic problems by the discontinuous Galerkin method// *RAIRO Modél. Math. Anal. Numér.* 1985. V. 19. Pp. 611–643.
85. Evans, L. C. *Partial differential equations*. American Mathematical Society, 1998.
86. Farrell, P. A., Hegarty, A. F., Miller, J. J. H., O’Riordan, E., Shishkin, G. I. *Robust computational techniques for boundary layers*. Boca Raton, FL: Chapman & Hall/CRC, 2000.
87. Fellner, K., Kovtunenکو, V. *A singularly perturbed nonlinear Poisson-Boltzmann equation: uniform and super-asymptotic expansions*// Tech. rep., SFB F 32/Universität Graz, 2014.
88. Fife, P. C. Semilinear elliptic boundary value problems with small parameters// *Arch. Rational Mech. Anal.* 1973. V. 52. Pp. 205–232.
89. Friedman, A. *Partial differential equations of parabolic type*// Englewood Cliffs, N.J.: Prentice-Hall, 1964.
90. Formaggia, L., Perotto, S. New anisotropic a priori error estimates// *Numer. Math.* 2001. V. 89. Pp. 641–667.
91. Gartland, E.C. Graded-mesh difference schemes for singularly perturbed two-point boundary value problems// *Math. Comp.* 1988. V. 51. Pp. 631–657.
92. Gilbarg, D., Trudinger, N. S. *Elliptic partial differential equations of second order*. Berlin: Springer-Verlag, 1998.
93. Griffel, D. H. *Applied functional analysis*. Mineola, NY: Dover Publications, 2002.

94. Grindrod P. *Patterns and waves: the theory and applications of reaction-diffusion equations*. Oxford: Clarendon Press, 1991.
95. Guzmán, J., Leykekhman, D., Rossmann, J., Schatz, A. H. Hölder estimates for Green's functions on convex polyhedral domains and their applications to finite element methods// *Numer. Math.* 2009. V. 112. Pp. 221–243.
96. Haverkamp, R. Eine Aussage zur L_∞ -Stabilität und zur genauen Konvergenzordnung der H_0^1 -Projektionen// *Numer. Math.* 1984. V. 44. Pp. 393–405.
97. Howes, F. A. Boundary-interior layer interactions in nonlinear singular perturbation theory// *Mem. Amer. Math. Soc.* 1978. V. 15.
98. Huang, W, Russell, R. D. *Adaptive Moving Mesh Methods*. New York: Springer, 2011.
99. Jerison, D., Kenig, C. E. The inhomogeneous Dirichlet problem in Lipschitz domains// *J. Funct. Anal.* 1995, V. 130. Pp. 161–219.
100. Juntunen, M., Stenberg, R. A residual based a posteriori estimator for the reaction-diffusion problem// *C. R. Math. Acad. Sci. Paris.* 2009. V. 347. Pp. 555–558.
101. Juntunen, M., Stenberg, R. Analysis of finite element methods for the Brinkman problem// *Calcolo.* 2010. V. 47. Pp. 129–147.
102. Kellog, R. B., Tsan, A. Analysis of some difference approximations for a singular perturbation problem without turning points// *Math. Comp.* 1978. V. 32. Pp. 1025–1039.
103. Kreuzer, C., Siebert, K. G. Decay rates of adaptive finite elements with Dörfler marking// *Numer. Math.* 2011. V. 117. Pp. 679–716.

104. Kunert, G. An a posteriori residual error estimator for the finite element method on anisotropic tetrahedral meshes// *Numer. Math.* 2000. V. 86. Pp. 471–490.
105. Kunert, G. Robust a posteriori error estimation for a singularly perturbed reaction-diffusion equation on anisotropic tetrahedral meshes// *Adv. Comput. Math.* 2001. V. 15. Pp. 237–259.
106. Kunert, G. A posteriori H^1 error estimation for a singularly perturbed reaction diffusion problem on anisotropic meshes// *IMA J. Numer. Anal.* 2005. V. 25. Pp. 408–428.
107. Kunert, G., Verfürth, R. Edge residuals dominate a posteriori error estimates for linear finite element methods on anisotropic triangular and tetrahedral meshes// *Numer. Math.* 2000. V. 86. Pp. 283–303.
108. Lakkis, O., Makridakis, C. Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems// *Math. Comp.* 2006. V. 75. Pp. 1627–1658.
109. Leykekhman, D. Uniform error estimates in the finite element method for a singularly perturbed reaction-diffusion problem// *Math. Comp.* 2008. V. 77. Pp. 21–39.
110. Lin, R., Stynes, M. A balanced finite element method for singularly perturbed reaction-diffusion problems// *SIAM J. Numer. Anal.* 2012. V. 50. Pp. 2729–2743.
111. Linß, T. Maximum-norm error analysis of a non-monotone fem for a singularly perturbed reaction-diffusion problem// *BIT Numer. Math.* 2007. V. 47. Pp. 379–391.
112. Linß, T. *Layer-adapted meshes for reaction-convection-diffusion problems*// Berlin: Springer, 2010.

113. Lorenz, J. *Nonlinear singular perturbation problems and the Enquist-Osher scheme* Report 8115, Mathematical Institute, Catholic University of Nijmegen, 1981.
114. Makridakis, C., Nochetto, R. H. Elliptic reconstruction and a posteriori error estimates for parabolic problems// *SIAM J. Numer. Anal.* 2003. V. 41. Pp. 1585–1594.
115. Makridakis, C., Nochetto, R. H. A posteriori error analysis for higher order dissipative methods for evolution problems// *Numer. Math.* 2006. V. 104. Pp. 489–514.
116. Maz'ya, V., Rossmann, J. *Elliptic equations in polyhedral domains*. Providence, RI: American Mathematical Society, 2010.
117. Melenk, J. M. *hp-finite element methods for singular perturbations*. Springer, 2002.
118. Miller, J. J. H., O'Riordan, E., Shishkin, G. I. *Solution of singularly perturbed problems with ε -uniform numerical methods*. Singapore: World Scientific, 1996.
119. Morton, K. W. *Numerical solution of convection-diffusion problems*. London: Chapman & Hall/CRC, 1996.
120. Murray, J. D. *Mathematical Biology*. Berlin: Springer-Verlag, 1993.
121. Nefedov, N. N. Comparison principle for reaction-diffusion-advection problems with boundary and internal layers// *Lecture Notes in Comput. Sci.* 2013. V. 8236. Pp. 62–72.
122. Nefedov, N. N., Recke, L., Schnieder, K. R. Existence and asymptotic stability of periodic solutions with an interior layer of reaction-advection-diffusion equations// *J. Math. Anal. Appl.* 2013. V. 405. Pp. 90–103.

123. Nochetto, R. H. Pointwise a posteriori error estimates for elliptic problems on highly graded meshes// *Math. Comp.* 1995. V. 64. Pp. 1–22.
124. Nochetto, R. H. *Pointwise a posteriori error estimates for monotone semi-linear equations*. Lecture Notes at 2006 CNA Summer School Probabilistic and Analytical Perspectives on Contemporary PDEs, 2006. <http://www.math.cmu.edu/cna/Summer06/lecturenotes/nochetto/>
125. Nochetto, R. H., Schmidt, A., Siebert, K. G., Veerer, A. Pointwise a posteriori error estimates for monotone semilinear problems// *Numer. Math.* 2006. V. 104. Pp. 515–538.
126. Nochetto, R. H., Siebert, K. G., Veerer, A. Pointwise a posteriori error control for elliptic obstacle problems// *Numer. Math.* 2003. V. 95. Pp. 163–195.
127. Nochetto, R. H., Siebert, K. G., Veerer, A. Fully localized a posteriori error estimators and barrier sets for contact problems// *SIAM J. Numer. Anal.* 2005. V. 42. Pp. 2118–2135.
128. O'Malley, R. E. *Singular perturbation methods for ordinary differential equations*. New York: Springer, 1991.
129. Pao, C. V. *Nonlinear parabolic and elliptic equations*. New York: Plenum Press, 1992.
130. Rannacher, R., Scott, R. Some optimal error estimates for piecewise linear finite element approximations// *Math. Comp.* 1982. V. 38. Pp. 437–445.
131. Roos, H.-G., Stynes, M., Tobiska, L. *Numerical methods for singularly perturbed differential equations*. Berlin: Springer, 1996.
132. Roos, H.-G., Stynes, M., Tobiska, L. *Robust numerical methods for singularly perturbed differential equations*. Berlin: Springer, 2008.

133. Schatz, A. H., Wahlbin, L. B. On the finite element method for singularly perturbed reaction-diffusion problems in two and one dimensions// *Math. Comp.* 1983. V. 40. Pp. 47–89.
134. Schatz, A. H., Wahlbin, L. B. Interior maximum-norm estimates for finite element methods, Part II// *Math. Comp.* 1995. V. 64. Pp. 907–928.
135. Scott, L. R., Zhang, S. Finite element interpolation of nonsmooth functions satisfying boundary conditions// *Math. Comp.* 1990. V. 54. Pp. 483–493.
136. Shishkin, G. I., Shishkina, L. P. *Difference methods for singular perturbation problems*. Boca Raton, FL: CRC Press, 2009.
137. Siebert, K. G. An a posteriori error estimator for anisotropic refinement// *Numer. Math.* 1996. V. 73. Pp. 373–398.
138. Stevenson, R. P. The uniform saturation property for a singularly perturbed reaction-diffusion equation// *Numer. Math.* 2005. V. 101. Pp. 355–379.
139. Stynes, M. Steady-state convection-diffusion problems// *Acta Numerica*. 2005. V. 14. Pp. 445–508.
140. Stynes M., O’Riordan E. A finite element method for a singularly perturbed boundary value problem// *Numer. Math.* 1986. V. 50. Pp. 1–5.
141. Stynes M., O’Riordan E. A uniformly accurate finite element method for a singular perturbation problem in conservative form// *SIAM J. Numer. Anal.* 1986. V. 23. Pp. 369–375.
142. Stynes M., O’Riordan E. A super convergence result for a singularly perturbed boundary value problem// *BAIL III: Proc. 3-rd intern. conf. on Boundary and Interior Layers*, Dublin: Boole, 1984. Pp. 309–313.

143. Sun, G., Stynes, M. A uniformly convergent method for a singularly perturbed semilinear reaction-diffusion problem with multiple solutions// *Math. Comp.* 1996. V. 65. Pp. 1085–1109.
144. Sweers, G. Semilinear elliptic problems on domains with corners// *Comm. Partial Differential Equations.* 1989. V. 14. Pp. 1229–1247.
145. Thomée, V. *Galerkin finite element methods for parabolic problems.* Berlin: Springer, 2006.
146. Tobiska, L., Verfürth, R. Robust a posteriori error estimates for stabilized finite element methods// *IMA J. Numer. Anal.* 2015. V. 35. Pp. 1652–1671.
147. Vasil'eva, A. B., Butuzov, V. F., Kalachev, L. V. *The boundary function method for singular perturbation problems.* Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 1995.
148. Verfürth, R. *A review of a posteriori error estimation and adaptive mesh-refinement techniques.* Wiley-Teubner, 1996.
149. Verfürth, R. Robust a posteriori error estimators for a singularly perturbed reaction-diffusion equation// *Numer. Math.* 1998. V. 78. Pp. 479–493.
150. Verfürth, R. Robust a posteriori error estimates for stationary convection-diffusion equations// *SIAM J. Numer. Anal.* 2006. V. 43. Pp. 1766–1782.
151. Verfürth, R. *A posteriori error estimation techniques for finite element methods.* Oxford University Press, 2013.
152. Vulcanovic, R. On a numerical solution of a type of singularly perturbed boundary value problem by using a special discretization mesh// *Zb. Rad., Prir.-Mat. Fak., Univ. Novom Sadu, Ser. Mat.* 1983. V. 13. Pp. 187–201.

153. Wheeler, M. F. L_∞ estimates of optimal orders for Galerkin methods for one-dimensional second order parabolic and hyperbolic equations// *SIAM J. Numer. Anal.* 1973. V. 10. Pp. 908–913.

Публикации автора по теме диссертации

154. Андреев В. Б., Коптева Н. В. Об исследовании разностных схем с аппроксимацией первой производной центральным разностным отношением// *Ж. вычисл. матем. и матем. физ.* 1996. N 8. С. 101–117.
155. Andreev, V. B., Kopteva, N. Pointwise approximation of corner singularities for a singularly perturbed reaction-diffusion equation in an L -shaped domain// *Math. Comp.* 2008. V. 77. Pp. 2125–2139.
156. Chadha, N. M., Kopteva, N. Maximum norm a posteriori error estimate for a 3d singularly perturbed semilinear reaction-diffusion problem// *Adv. Comput. Math.* 2011. V. 35. Pp. 33–55.
157. Demlow, A., Kopteva, N. Maximum-norm a posteriori error estimates for singularly perturbed elliptic reaction-diffusion problems// *Numer. Math.* 2016. V. 133. Pp. 707–742.
158. Franz, S., Kopteva, N. Green's function estimates for a singularly perturbed convection-diffusion problem in three dimensions// *Int. J. Numer. Anal. Model. Ser. B.* 2011. V. 2. Pp. 124–141.
159. Franz, S., Kopteva, N. Green's function estimates for a singularly perturbed convection-diffusion problem// *J. Differential Equations.* 2012. V. 252. Pp. 1521–1545.
160. Kellogg, R. B., Kopteva, N. *Some asymptotic expansions for a semilinear reaction-diffusion problem in a sector*// arXiv:0902.0987, 2009.

161. Kellogg, R. B., Kopteva, N. A singularly perturbed semilinear reaction-diffusion problem in a polygonal domain// *J. Differential Equations*. 2010. V. 248. Pp. 184–208.
162. Kopteva, N. Error expansion for an upwind scheme applied to a two-dimensional convection-diffusion problem// *SIAM J. Numer. Anal.* 2003. V. 41. Pp. 1851–1869.
163. Kopteva, N. Maximum norm error analysis of a 2d singularly perturbed semilinear reaction-diffusion problem// *Math. Comp.* 2007. V. 76. Pp. 631–646.
164. Kopteva, N. Maximum norm a posteriori error estimates for a 1D singularly perturbed semilinear reaction-diffusion problem// *IMA J. Numer. Anal.* 2007. V. 27. Pp. 576–592.
165. Kopteva, N. Maximum norm a posteriori error estimate for a 2d singularly perturbed reaction-diffusion problem// *SIAM J. Numer. Anal.* 2008. V. 46. Pp. 1602–1618.
166. Kopteva, N. Numerical analysis of a 2d singularly perturbed semilinear reaction-diffusion problem// *Lecture Notes in Comput. Sci.* 2009. V. 5434. Pp. 80-91.
167. Kopteva, N. Linear finite elements may be only first-order pointwise accurate on anisotropic triangulations// *Math. Comp.* 2014. V. 83. Pp. 2061–2070.
168. Kopteva, N. Maximum-norm a posteriori error estimates for singularly perturbed reaction-diffusion problems on anisotropic meshes// *SIAM J. Numer. Anal.* 2015. V. 53. Pp. 2519–2544.
169. Kopteva, N. Energy-norm a posteriori error estimates for singularly perturbed reaction-diffusion problems on anisotropic meshes// *Numer. Math.* 2017. V. 137. Pp. 607–642.

170. Kopteva, N. *Fully computable a posteriori error estimator using anisotropic flux equilibration on anisotropic meshes*. arXiv:1704.04404, 2017.
171. Kopteva, N. Energy-norm a posteriori error estimates for singularly perturbed reaction-diffusion problems on anisotropic meshes. Neumann boundary conditions// *Boundary and Interior Layers, Computational and Asymptotic Methods BAIL 2016*. Lect. Notes Comput. Sci. Eng., Springer, 2017. Pp. 141–154.
172. Kopteva, N. Logarithm cannot be removed in maximum norm error estimates for linear finite elements in 3D// *Math. Comp.* 2018. Published electronically 28-Sep-2018.
173. Kopteva, N. *Improved energy-norm a posteriori error estimates for singularly perturbed reaction-diffusion problems on anisotropic meshes*. arXiv:1810.09211, 2018.
174. Kopteva, N., Linß, T. Maximum norm a posteriori error estimation for a time-dependent reaction-diffusion problem// *Comput. Methods Appl. Math.* 2012. V. 12. Pp. 189–205.
175. Kopteva, N., Linß, T. Maximum norm a posteriori error estimation for parabolic problems using elliptic reconstructions// *SIAM J. Numer. Anal.* 2013. V. 51. Pp. 1494–1524.
176. Kopteva, N., Linß, T. Numerical study of maximum norm a posteriori error estimates for singularly perturbed parabolic problems// *Lecture Notes in Comput. Sci.* 2013. V. 8236. Pp. 50–61.
177. Kopteva, N., Linß, T. Improved maximum-norm a posteriori error estimates for linear and semilinear parabolic equations// *Adv. Comput. Math.* 2017. V. 43. Pp. 999–1022.

178. Kopteva, N., O’Riordan, E. Shishkin meshes in the numerical solution of singularly perturbed differential equations// *Int. J. Numer. Anal. Model.* 2010. V. 7. Pp. 393–415.
179. Kopteva, N., Pickett, M. A second-order overlapping Schwarz method for a 2d singularly perturbed semilinear reaction-diffusion problem// *Math. Comp.* 2012. V. 81. Pp. 81–105.
180. Kopteva, N., Pickett, M., Purtil, H. A robust overlapping Schwarz method for a singularly perturbed semilinear reaction-diffusion problem with multiple solutions// *Int. J. Numer. Anal. Model.* 2009. V. 6. Pp. 680–695.
181. Kopteva, N., Savescu, S. B. Pointwise error estimates for a singularly perturbed time-dependent semilinear reaction-diffusion problem// *IMA J. Numer. Anal.* 2011. V. 31. Pp. 616-639.
182. Kopteva, N., Stynes, M. Numerical analysis of a singularly perturbed nonlinear reaction-diffusion problem with multiple solutions// *Appl. Numer. Math.* 2004. V. 51. Pp. 273–288.
183. Kopteva, N., Stynes, M. Stabilised approximation of interior-layer solutions of a singularly perturbed semilinear reaction-diffusion problem// *Numer. Math.* 2011. V. 119. Pp. 787–810.