СОДЕРЖАНИЕ

Том 47, номер 8, 2007 год

Contents

Vol. 47, No. 8, 2007

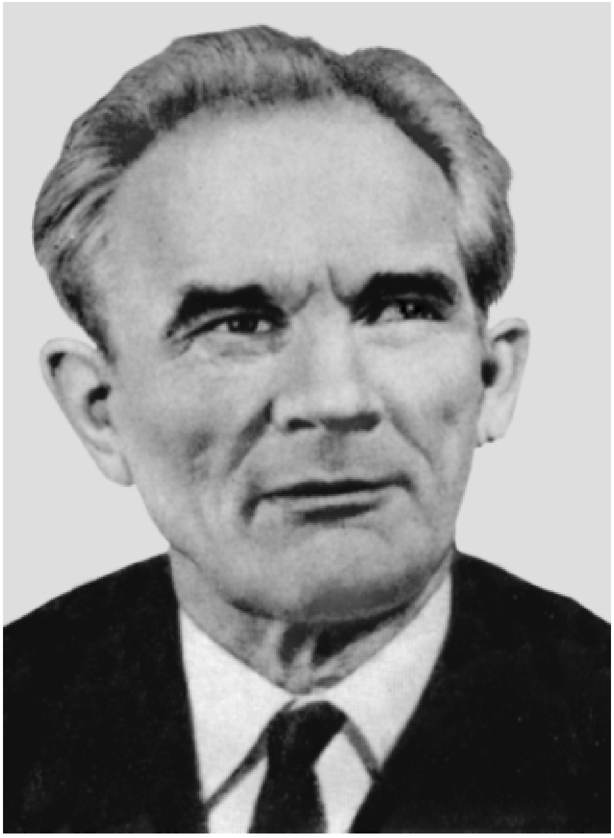
Simultaneous English language translation of the journal is available from Pleiades Publishing, Ltd.
Distributed worldwide by Springer. *Computational Mathematics and Mathematical Physics* ISSN 0965-5425.

Сдано в набор 10.04.2007 г. Подписано к печати 20.06.2007 г. Формат бумаги $60 \times 88^1/_8$
Цифровая печать Усл. печ. л. 22.0 Усл. кр.-отт. 8.1 тыс. Уч.-изд. л. 22.0 Бум. л. 11.0
Тираж 362 экз. Зак. 369

Учредители: Российская академия наук, Вычислительный центр им. А.А. Дородницына РАН

Издатель: Академиздатцентр "Наука", 117997 Москва, Профсоюзная ул., 90
Оригинал-макет подготовлен МАИК "Наука/Интерпериодика"
Отпечатано в ППП "Типография "Наука", 121099 Москва, Шубинский пер., 6

ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 2007, том 47, № 8, с. 1283–1286



К СТОЛЕТИЮ СО ДНЯ РОЖДЕНИЯ ЧЛЕНА-КОРРЕСПОНДЕНТА АН СССР ДМИТРИЯ КОНСТАНТИНОВИЧА ФАДДЕЕВА (1907–1989)

© 2007 г. М. К. Керимов

(119991 Москва, ул. Вавилова, 40, ВЦ РАН)

e-mail: comp_mat@ccas.ru

Поступила в редакцию 27.02.2007 г.

В эти дни научная математическая общественность отмечает столетие со дня рождения выдающегося математика, Лауреата Государственной премии СССР, члена-корреспондента АН СССР, многолетнего члена редколлегии Журнала вычислительной математики и математической физики профессора Дмитрия Константиновича Фаддеева (1907–1989). Все, кто знал Дмитрия Константиновича, вспоминают его как разносторонне одаренного, доступного в общении, обаятельного человека. Д.К. Фаддеев всегда был обязательным человеком. Автор этих строк, соприкасаясь с ним, особенно отмечает это его качество. Будучи членом редколлегии нашего журнала, живя в Ленинграде, являясь главой ленинградской математической школы по алгебре и вычислительной математике, будучи очень занятым человеком, он регулярно приезжал в Москву на заседания редколлегии, активно участвовал в обсуждении намеченных к опубликованию в журнале статей и своим высоким авторитетом повышал научное качество журнала. Дмитрий Константинович обычно немного раньше назначенного времени приходил в Вычислительный центр АН СССР на заседания редколлегии, заходил в мой кабинет, мы совместно с ним обсуждали кое-какие проблемы. Однажды, поскольку до заседания редколлегии оставалось некоторое время, я предложил ему какую-то книгу. Дмитрий Константинович мне ответил, что у него есть своя работа, достал из портфеля свою очередную научную статью и стал ее править, т.е., как истинный ученый, он все время думал о текущих проблемах, над которыми работал, и даже в командировках занимался математическими проблемами.

Д.К. Фаддеев родился 30 июня 1907 г. в г. Юхнове Калужской области. В 1923 г. он поступил на математический факультет Ленинградского государственного университета, где слушал лекции таких известных профессоров, как Г.М. Фихтенгольц, Б.Н. Делоне, И.М. Виноградов и др. В 1928 г. он успешно окончил университет и стал работать лаборантом в Палате мер и весов (Институт метрологии им. Д.И. Менделеева). В 1929 г. Д.К. Фаддеев поступил на композиторское отделение Ленинградской консерватории, однако совместить музыкальное образование с научной работой и обязанностями службы оказалось чересчур трудно и консерваторию со второго курса пришлось оставить. В 1930 г. началась преподавательская деятельность Дмитрия Константиновича в ленинградских вузах, с 1933 г. до своей кончины он работал в Ленинградском государственном университете, пройдя все ступени от ассистента до профессора. В 1932–1934 гг. он работал также в Математическом институте им. В.П. Стеклова АН СССР, а с 1940 г. после организации Ленинградского отделения этого института и до своей кончины работал также в Ленинградском отделении этого института. В 1935 г. Дмитрий Константинович успешно защитил диссертацию на степень доктора физико-математических наук. В 1952–1954 гг. он был деканом математико-механического факультета Ленинградского университета.

Д.К. Фаддеев является учеником известного ученого-математика Бориса Николаевича Делоне, и первые его научные работы были созданы под влиянием этого ученого. В дальнейшем были опубликованы их совместные фундаментальные работы по обратным задачам теории Галуа, вошедшие в большую монографию, посвященную кубическим иррациональностям. Много лет спустя после опубликования этой монографии я получил письмо из Америки с просьбой прислать экземпляр этой монографии для перевода на английский язык, и я вспоминаю, с какими трудностями мне удалось найти и выслать в Америку эту книгу (позже она была опубликована на английском языке в Американском математическом обществе).

В 1964 г. за выдающиеся научные достижения Д.К. Фаддеев был избран членом-корреспондентом АН СССР. С 1944 г. до своей кончины он являлся профессором Ленинградского университета. В 1930 г. Д.К. Фаддеев женился на Вере Николаевне Фаддеевой (Замятиной), ставшей не только матерью троих детей (один из них, Людвиг Дмитриевич Фаддеев, стал выдающимся специалистом по математической и теоретической физике, ныне – академик РАН, главой Санкт-Петербургской математической школы), но и известным специалистом по вычислительной математике. Этот счастливый союз двух выдающихся людей оказал большое влияние на дальнейшее развитие вычислительной математики. Дмитрий Константинович известен своими фундаментальными работами по теоретической математике (алгебра, теория чисел и др.), однако указанный союз двух ученых сделал Д.К. Фаддеева выдающимся специалистом и по вычислительной линейной алгебре. Вряд ли найдется ученый-исследователь, который не пользовался бы всемирно известной монографией этих двух ученых по методам вычислений в области линейной алгебры. При написании этой книги Дмитрий Константинович опирался на свои глубокие теоретические знания по алгебре, а Вера Николаевна, будучи тонким специалистом по методам вычислений, использовала свой большой опыт для написания вычислительной части такой хорошей книги. В 1981 г. за свои выдающиеся работы по вычислительной алгебре Д.К. Фаддеев и В.Н. Фаддеева были удостоены звания Лауреатов Государственной премии СССР.

Д.К. Фаддеев опубликовал свыше 160 научных работ. Сюда относятся его замечательные работы по теории чисел, по алгебре, по теории функций, по геометрии, по теории вероятностей, по вычислительной математике и др. Большой популярностью пользовались в свое время написанные им учебники для учащихся высших и средних учебных заведений. Эти книги долго будут служить прекрасными пособиями для молодых людей. Автор этих строк не ставил задачу изложить все его замечательные научные результаты по теоретической математике. Однако его результаты по вычислительной линейной алгебре в какой-то степени отражены в статье, посвященной столетию со дня рождения В.Н. Фаддеевой, опубликованной в нашем журнале (Ж. вычисл. матем. и матем. физ. 2007. Т. 47. № 7. С. 1).

Д.К. Фаддеев являлся одним из организаторов нашего журнала, был представителем журнала в таком крупнейшем научном центре, каким является Ленинград. Он со дня основания журнала в 1961 г. в течение 25 лет до 1986 г. был одним из активнейших и добросовестных членов редколлегии, многие его научные работы по вычислительной математике были опубликованы именно в нашем журнале. Благодаря его широкой эрудиции, требовательности и доброжелательности уровень работ по вычислительным методам алгебры, публиковавшихся в нашем журнале, всегда был высоким. Д.К. Фаддеев не жалел ни времени, ни энергии, помогая молодым авторам совершенствовать свои работы как по содержанию, так и по оформлению.

Будучи много лет университетским профессором, Дмитрий Константинович подготовил много молодых ученых-математиков. Его ученик Игорь Ростиславович Шафаревич является академиком РАН, другие его ученики – З.И. Боревич, Д.С. Горшков, А.И. Скопин – доктора физико-математических наук. Многие другие его ученики стали видными учеными и продолжают развивать его идеи.

Широко известны учебники, написанные Д.К. Фаддеевым, особенно его (совместно с И.С. Соминским) "Сборник задач по высшей алгебре" (у нас вышли 6 изданий этой книги), переведенный на некоторые иностранные языки. Известны также книги Д.К. Фаддеева для средних учебных заведений.

Человеческие качества Д.К. Фаддеева, широта его взглядов и интересов, благородство, требовательность к себе и доброжелательность к другим были соразмерны его глубокой внутренней культуре. Все это, вместе с особым сочетанием порядочности и научной одержимости, делало общение с Дмитрием Константиновичем событием для всех, кто с ним соприкасался. Его присутствие на заседаниях редколлегии нашего журнала создавало особую творческую атмосферу. Члены редколлегии нашего журнала, многочисленные авторы, сотрудники редакции и читатели навсегда сохранят светлую память о Дмитрии Константиновиче Фаддееве.

СПИСОК НЕКОТОРЫХ НАУЧНЫХ РАБОТ Д.К. ФАДДЕЕВА[1)]

О представлении суммируемых функций сингулярными интегралами в точках Лебега // Матем. сб. 1936. Т. 1(43). С. 351–368.

О преобразовании векового уравнения матрицы // Тр. Ин-та инж. пром. стр-ва. М., 1937. Т. 4. С. 78–86.

Арифметика и алгебра комплексных чисел. Л.: Учпедгиз, 1939. (Совм. с Р.О. Кузьминым.)

Теория иррациональностей третьей степени. М.: Изд-во АН СССР, 1940. 340 с. (Тр. Матем. ин-та АН СССР. М., 1940. Т. 11). (Совм. с Б.Н. Делоне.)

Линейная алгебра // Математика, ее содержание, методы и значение. Т. 3. М.: Изд-во АН СССР, 1956. С. 37–92.

Вычислительные методы линейной алгебры // Тр. III Всес. матем. съезда. М.: Изд-во АН СССР, 1958. С. 434–445. (Совм. с В.Н. Фаддеевой.)

К вопросу о верхней релаксации при решении систем линейных уравнений // Изв. вузов. Сер. Матем. 1958. № 5. С. 122–125.

О некоторых последовательностях полиномов, полезных для построения итерационных методов решения систем линейных алгебраических уравнений // Вестн. ЛГУ. Сер. Матем. 1958. Т. 7. № 2. С. 155–159.

Об обусловленности матриц // Тр. Матем. ин-та АН СССР. М., 1959. Т. 53. С. 387.

Вычислительные методы линейной алгебры. М.: Физматгиз, 1960. 656 с. (Совм. с В.Н. Фаддеевой.)

О плохо обусловленных системах линейных уравнений // Ж. вычисл. матем. и матем. физ. 1961. Т. 1. № 3. С. 412–417. (Совм. с В.Н. Фаддеевой.)

Таблицы основных унитарных представлений федоровских групп. М.–Л.: Физматгиз, 1961. 174 с. (Тр. Матем. ин-та АН СССР. М., 1961. Вып. 56.)

Вычислительные методы линейной алгебры. Изд. 2-е, доп. М.–Л.: Физматгиз, 1963. (Совм. с В.Н. Фаддеевой.)

Об эквивалентности систем целочисленных матриц // Изв. АН СССР. Сер. Матем. 1966. Т. 30. № 2. С. 449–454.

Задача масштабирования для линейных систем // Соврем. числ. методы. Материалы Международ. школы по числ. методам. Киев, 1966; М. 1968. С. 76–84. (Совм. с В.Н. Фаддеевой.)

Stability in linear algebra problems // Proc. IFIP. Congress 68. Edinburgh, 1968. V. 1. Mathematics, Software. Amsterdam: North-Holland Publ. Co., 1969. P. 33–39 (With V.N. Faddeeva.)

Естественные нормы в алгебраических процессах // Вопр. точности и эффективности вычисл. алгоритмов. Тр. Симпозиума. Т. 1. Киев, 1969. С. 122–141. (Совм. с В.Н. Фаддеевой.)

Natural norms in algebraic processes // SIAM J. Numer. Analys. 1970. V. 7. № 4. P. 520–531. (With V.N. Faddeeva.)

Сопутствующая матрица и оценивание конечной вычислительной линейной алгебры. Новосибирск, 1973. С. 4–10. (Совм. с В.Н. Фаддеевой.)

[1)] Более полный список научных работ Д.К. Фаддеева см. в журналах: Успехи матем. наук. 1958. Т. 13. Вып. 1. С. 236–238; 1968. Т. 23. Вып. 3. С. 193–195; 1979. Т. 34. Вып. 2. С. 226–228.

К вопросу о решении линейно-алгебраических систем // Ж. вычисл. матем. и матем. физ. 1979. Т. 14. № 3. С. 539–558. (Совм. с В.Н. Фаддеевой.)

Линейные алгебраические системы с прямоугольными матрицами // Соврем. числ. методы. Материалы Междунар. летней школы по числ. методам. Киев, 1966; М., 1968. С. 16–75. (Совм. с В.Н. Фаддеевой и В.Н. Кублановской.)

О решении линейно-алгебраических систем с прямоугольными матрицами // Тр. Матем. ин-та АН СССР. М., 1968. Т. 96. С. 76–92. (Совм. с В.Н. Фаддеевой и В.Н. Кублановской.)

Вычислительные методы линейной алгебры // Тр. III Всес. матем. съезда. Т. 3. М.: Изд-во АН СССР, 1958. С. 434–445. (Совм. с В.Н. Фаддеевой.)

Вычислительные методы линейной алгебры // Зап. научн. семинаров ЛОМИ АН СССР. 1975. Т. 54. С. 3–228. (Совм. с В.Н. Фаддеевой.)

Взгляд на развитие вычислительных методов линейной алгебры // Вычисл. методы линейной алгебры. Новосибирск, 1977. С. 4–14 (Совм. с В.Н. Фаддеевой.)

Параллельные вычисления в линейной алгебре // Кибернетика. 1977. Т. 6. С. 28–40. (Совм. с В.Н. Фаддеевой).

Параллельные вычисления в линейной алгебре: Препринт Р-6-81. Л.: ЛОМИ АН СССР, 1981. 47 с. (Совм. с В.Н. Фаддеевой.)

ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 2007, том 47, № 8, с. 1287–1297

УДК 519.642.8

МЕТОДЫ РЕГУЛЯРИЗАЦИИ В ГИЛЬБЕРТОВОМ ПРОСТРАНСТВЕ ДЛЯ НЕКОТОРЫХ КВАЗИВАРИАЦИОННЫХ НЕРАВЕНСТВ С ПРИБЛИЖЕННЫМИ ДАННЫМИ

© 2007 г. И. П. Рязанцева

(603600 Нижний Новгород, ул. Минина, 24, НГТУ)

e-mail: ryazantseva@waise.nntu.sci-nnov.ru

Поступила в редакцию 02.02.2007 г.

На основе операторного, непрерывного и итеративного методов регуляризации построены приближения, сильно сходящиеся в гильбертовом пространстве к решению исходного квазивариационного неравенства специального вида при приближенном задании данных. Библ. 18.

Ключевые слова: квазивариационные неравенства, операторный метод регуляризации, непрерывный метод регуляризации первого порядка, итеративный метод регуляризации.

Пусть Ω – выпуклое замкнутое множество в вещественном гильбертовом пространстве H, $A : H \longrightarrow H$, $B : H \longrightarrow H$ – некоторые операторы, при всех u и v из H удовлетворяющие следующим условиям:

а) $(Au - Av, u - v) \geq m\|u - v\|^2$, $m > 0$;

б) $\|Au - Av\| \leq L_1\|u - v\|$, $L_1 > 0$;

в) $\|Bu - Bv\| \leq L_2\|u - v\|$, $L_2 > 0$.

Здесь (u, v) – скалярное произведение элементов u и v из H.

В [1] установлено, что при

$$M = \frac{L_1L_2}{m} + L_2 < 1 \tag{1}$$

квазивариационное неравенство

$$(Ax - f, x - y) \leq 0, \quad x \in \Omega(x), \quad \forall y \in \Omega(x), \tag{2}$$

где $\Omega(x) = \Omega + Bx$, имеет единственное решение x для любого фиксированного элемента $f \in H$. Далее считаем, что (1) выполнено. Пусть данные задачи (2), а именно операторы A и B, элемент f и множество Ω, возмущены, т.е. вместо A, B, f и Ω известны, соответственно, семейства $\{A^h\}$, $\{B^h\}$, $\{f^\delta\}$, $\{\Omega^\sigma\}$, $0 \leq h \leq \bar{h}$, $0 \leq \delta \leq \bar{\delta}$, $0 \leq \sigma \leq \bar{\sigma}$, причем операторы $A^h : H \longrightarrow H$ и $B^h : H \longrightarrow H$ при всех u и v из H обладают следующими свойствами:

а') $(A^hu - A^hv, u - v) \geq m^h\|u - v\|^2$, $m^h > 0$;

б') $\|A^hu - A^hv\| \leq L_1^h\|u - v\|$, $L_1^h > 0$;

в') $\|B^hu - B^hv\| \leq L_2^h\|u - v\|$, $L_2^h > 0$;

Ω^σ – выпуклое замкнутое множество в H, $f^\delta \in H$. Кроме того,

$$m^h \longrightarrow m, \quad L_i^h \longrightarrow L_i \text{ при } h \longrightarrow 0, \quad i = 1, 2, \tag{3}$$

$$\|Au - A^hu\| \leq hg(\|u\|), \tag{4}$$

$$\|Bu - B^hu\| \leq hq(\|u\|), \tag{5}$$

$$\|f - f^\delta\| \leq \delta, \tag{6}$$

$$r_H(\Omega, \Omega^\sigma) \leq \sigma; \tag{7}$$

здесь $r_H(\Omega, \Omega^\sigma)$ – хаусдорфово расстояние в H между множествами Ω и Ω^σ, $g(s)$ и $q(s)$ – неотрицательные ограниченные (т.е. переводящие ограниченное множество в ограниченное) функции, $s \ge 0$.

Следовательно, вместо (2) имеем задачу

$$(A^h x - f^\delta, x - y) \le 0, \quad x \in \Omega_h^\sigma(x) = \Omega^\sigma + B^h x \quad \forall y \in \Omega_h^\sigma(x). \tag{8}$$

Условие разрешимости (1) для (8) в силу предположений а')–в') принимает вид

$$M^h = \frac{L_1^h L_2^h}{m^h} + L_2^h < 1. \tag{9}$$

Если величина M близка к единице, то неравенство (9) может не выполняться и при малых h. Кроме того, на практике величину h часто нельзя сделать меньше некоторого значения h_0. Значит, в предположении (1) мы в общем случае не можем гарантировать существование решения возмущенной задачи (8). По этой причине перейдем от задачи (8) к близкой к ней регуляризованной задаче следующего вида:

$$(A^h x_\lambda^\mu + \alpha x_\lambda^\mu - f^\delta, x_\lambda^\mu - y) \le 0, \quad x_\lambda^\mu \in \Omega^\sigma + \beta B^h x_\lambda^\mu = \Omega^\Delta(x_\lambda^\mu) \quad \forall y \in \Omega^\Delta(x_\lambda^\mu), \tag{10}$$

где $\Delta = (h, \sigma, \beta)$, $\beta \in (0, 1)$, $\alpha > 0$, $\lambda = (\alpha, \beta)$, $\mu = (\delta, h, \sigma)$. Условие разрешимости вида (1) для квазивариационного неравенства (10) выражается неравенством

$$M^{\tilde\Delta} = \left(\frac{L_1^h + \alpha}{m^h + \alpha} + 1\right)\beta L_2^h < 1, \quad \tilde\Delta = (h, \alpha, \beta). \tag{11}$$

Так как $L_1^h \ge m^h$, то для фиксированного h при $\alpha > 0$ и $\beta \in (0, 1)$ справедливо неравенство $M^{\tilde\Delta} < M^h$.

Заметим, что при выполнении (9) переход от (8) к задаче (10) может быть оправдан тем, что значения постоянных $M^{\tilde\Delta}$ и M^h влияют на скорость сходимости приближенных методов решения задач (8) и (10) (см., например, [1]–[3]).

Пусть уровни возмущений данных δ, h, σ квазивариационного неравенства (2) обладают свойством

$$\mu = (\delta, h, \sigma) \longrightarrow (0, 0, 0) \text{ при } \lambda = (\alpha, \beta) \longrightarrow \lambda_0 = (0, 1-) \tag{12}$$

и при этом выполнено (11). Тогда квазивариационное неравенство (10) имеет единственное решение x_λ^μ. Исследуем поведение x_λ^μ при $\lambda \longrightarrow \lambda_0$. Так как x – решение неравенства (2), то $x = u + Bx$, $u \in \Omega$. Для элемента $u \in \Omega$ в силу условия (7) найдется точка $u_\sigma \in \Omega^\sigma$ такая, что

$$\|u - u_\sigma\| \le \sigma. \tag{13}$$

Следовательно,

$$z_\lambda^\mu = u_\sigma - u + x - Bx + \beta B^h x_\lambda^\mu \in \Omega^\Delta(x_\lambda^\mu).$$

Поскольку $x_\lambda^\mu \in \Omega^\Delta(x_\lambda^\mu)$, то $x_\lambda^\mu = u_\lambda^\mu + \beta B^h x_\lambda^\mu$, $u_\lambda^\mu \in \Omega^\sigma$. Для u_λ^μ согласно условию (7) найдем элемент $v_\lambda^\mu \in \Omega$ такой, что

$$\|u_\lambda^\mu - v_\lambda^\mu\| \le \sigma. \tag{14}$$

Тогда

$$w_\lambda^\mu = v_\lambda^\mu + Bx = x_\lambda^\mu - \beta B^h x_\lambda^\mu + Bx - (u_\lambda^\mu - v_\lambda^\mu) \in \Omega(x).$$

В (2) положим $y = w_\lambda^\mu$, а в (10) положим $y = z_\lambda^\mu$, сложив полученные результаты, имеем

$$(Ax - f, x - x_\lambda^\mu + \beta B^h x_\lambda^\mu - Bx + u_\lambda^\mu - v_\lambda^\mu) + (A^h x_\lambda^\mu + \alpha x_\lambda^\mu - f^\delta, x_\lambda^\mu - x + u - u_\sigma + Bx - \beta B^h x_\lambda^\mu) \le 0.$$

Перепишем последнее неравенство в эквивалентной форме:

$$\begin{gathered}(Ax - A^h x, x - x_\lambda^\mu) + (A^h x - A^h x_\lambda^\mu, x - x_\lambda^\mu) + (f^\delta - f, x - x_\lambda^\mu) + \alpha\|x_\lambda^\mu - x\|^2 + \alpha(x, x_\lambda^\mu - x) + \\ + (Ax - A^h x + A^h x - A^h x_\lambda^\mu - f + f^\delta - \alpha(x_\lambda^\mu - x) - \alpha x, \beta B^h x_\lambda^\mu - Bx) + (Ax - f, u_\lambda^\mu - v_\lambda^\mu) + \\ + (A^h x_\lambda^\mu + \alpha x_\lambda^\mu - f^\delta, u - u_\sigma) \le 0.\end{gathered} \tag{15}$$

Зафиксируем элемент $z_0 \in H$, тогда при всех $y \in H$ верно неравенство

$$\|A^h y - f^\delta\| \le \|A^h y - A^h z_0\| + \|A^h z_0 - Az_0\| + \|Az_0 - f\| + \|f - f^\delta\|.$$

Теперь условия б'), (3), (4), (6) гарантируют существование постоянной $\bar{b} > 0$ такой, что

$$\|A^h y - f^\delta\| \le \bar{b}(\|y\| + 1) \quad \forall y \in H, \quad 0 \le h \le \bar{h}, \quad 0 \le \delta \le \bar{\delta}. \tag{16}$$

Учитывая (16), свойства а')–в') операторов A^h и B^h, предположения (4)–(7) и оценки (13), (14), из (15) выводим неравенство

$$\begin{gathered}(m^h + \alpha)\|x_\lambda^\mu - x\|^2 \le [hg(\|x\|) + \delta + \alpha\|x\|]\|x_\lambda^\mu - x\| + \\ + [hg(\|x\|) + \delta + (L_1^h + \alpha)\|x_\lambda^\mu - x\| + \alpha\|x\|]\{\beta[L_2^h\|x_\lambda^\mu - x\| + hq(\|x\|)] + (1-\beta)\|Bx\|\} + \\ + [\|Ax - f\| + \bar{b}(\|x_\lambda^\mu\| + 1) + \alpha\|x_\lambda^\mu\|]\sigma.\end{gathered} \tag{17}$$

Из предположения (11) следует, что величина

$$K_{\tilde{\Delta}} = m^h + \alpha - \beta(L_1^h + \alpha)L_2^h$$

положительна при всех $\alpha > 0$, $\beta \in (0, 1)$, $0 \le h \le \bar{h}$. Кроме того, свойства (3) величин m^h, L_1^h, L_2^h и условия (1), (12) обеспечивают справедливость соотношений

$$\lim_{\lambda \to \lambda_0} K_{\tilde{\Delta}} = m - L_1 L_2 = K > 0. \tag{18}$$

Значит, из (17) вытекает ограниченность семейства $\{x_\lambda^\mu\}$ при $\lambda \longrightarrow \lambda_0$. Теперь из (17) нетрудно получить неравенство

$$\|x_\lambda^\mu - x\|^2 \le c_1 \frac{\delta + h + \sigma + \alpha + 1 - \beta}{K_{\tilde{\Delta}}}, \tag{19}$$

где c_1 – положительная постоянная. Отсюда в силу (12) и (18) делаем вывод о сильной сходимости $\{x_\lambda^\mu\}$ при $\lambda \longrightarrow \lambda_0$ к единственному решению квазивариационного неравенства (2).

Сформулируем доказанное утверждение.

Теорема 1. *Пусть H – вещественное гильбертово пространство, операторы $A : H \longrightarrow H$, $B : H \longrightarrow H$ обладают свойствами* а)–в) *и справедливо неравенство* (1), *Ω – выпуклое замкнутое множество в H, элемент $f \in H$. Предположим, что данные квазивариационного неравенства* (2) *возмущены, т.е. вместо A, B, f и Ω известны семейства $\{A^h\}$, $\{B^h\}$, $\{f^\delta\}$ и $\{\Omega^\sigma\}$ соответственно, $0 \le h \le \bar{h}$, $0 \le \delta \le \bar{\delta}$, $0 \le \sigma \le \bar{\sigma}$, причем операторы $A^h : H \longrightarrow H$, $B^h : H \longrightarrow H$ обладают свойствами* а')–в'), *$f^\delta \in H$, Ω^σ – выпуклое замкнутое множество в H и выполнены условия* (3)–(7), (11), (12). *Тогда $x_\lambda^\mu \longrightarrow x$ при $\lambda \longrightarrow \lambda_0$, где x_λ^μ и x – единственные решения квазивариационных неравенств* (10) *и* (2) *соответственно.*

Как отмечалось в [4], [5], условие (7) близости множеств Ω и Ω^σ естественно для ограниченных множеств. Для неограниченных Ω и Ω^σ неравенство (7) следует заменить неравенством

$$\tau(R, \Omega, \Omega^\sigma) \le a(R)\sigma, \tag{20}$$

где $R \ge 0$ – некоторое число, $a(R)$ – неотрицательная непрерывная функция, $a(R) \longrightarrow +\infty$ при $R \longrightarrow +\infty$,

$$\tau(R, \Omega, \Omega^\sigma) = \max\{S(R, \Omega, \Omega^\sigma), S(R, \Omega^\sigma, \Omega)\},$$

$$S(R, \Omega, \Omega^\sigma) = \sup\{\inf\{\|u - v\| \mid v \in \Omega^\sigma\} \mid u \in \Omega^R\};$$

здесь $\Omega^R = \Omega \cap B(0, R)$, $B(0, R) = \{u \mid \|u\| \le R\}$. Следовательно, при замене (7) неравенством (20) оценки (13) и (14) примут, соответственно, вид

$$\|u - u_\sigma\| \le a(\|u\|)\sigma, \quad \|u_\lambda^\mu - v_\lambda^\mu\| \le a(\|u_\lambda^\mu\|)\sigma.$$

Поскольку $u_\lambda^\mu = x_\lambda^\mu - \beta[B^h x_\lambda^\mu - B^h x + B^h x]$, то, приняв во внимание условия в') и (5), получим

$$\|u_\lambda^\mu\| \le \|x_\lambda^\mu\| + \beta[L_2^h(\|x_\lambda^\mu\| + \|x\|) + hq(\|x\|) + \|Bx\|] \le (1 + \beta L_2^h)\|x_\lambda^\mu\| + c_2;$$

здесь положительная постоянная c_2 не зависит от λ и μ. Значит, в предположении (20) слагаемое $\|Ax - f\|\sigma$ в правой части (17) будет заменено на следующее:

$$\sigma\|Ax - f\|a((1 + \beta L_2^h)\|x_\lambda^\mu\| + c_2),$$

и при произвольном порядке роста функции $a(R)$ из полученного вместо (17) неравенства установить ограниченность семейства $\{x_\lambda^\mu\}$ при $\lambda \longrightarrow \lambda_0$ не удается. Пусть

$$a(R) \le c_3 R^s, \quad c_3 > 0, \quad s \in (0, 2]. \tag{21}$$

Так как $\sigma \longrightarrow 0$, $\beta L_2^h \longrightarrow L_2 > 0$ при $\lambda \longrightarrow \lambda_0$, то при $s \in (0, 2)$ утверждение теоремы 1, очевидно, сохраняется. Если $s = 2$, то при выводе оценки вида (19) роль $K_{\hat{\Delta}}$ будет выполнять величина

$$T(h, \sigma, \alpha, \beta) = K_{\hat{\Delta}} - c_3\sigma\|Ax - f\|(1 + \beta L_2^h)^2,$$

причем

$$\lim_{\lambda \to \lambda_0} T(h, \sigma, \alpha, \beta) = K > 0.$$

Таким образом, и в этом случае будет установлена ограниченность $\{x_\lambda^\mu\}$ при $\lambda \longrightarrow \lambda_0$, т.е. справедлива

Теорема 2. *Если в условиях теоремы* 1 *предположение* (7) *заменить на* (20) *с функцией* $a(R)$, *удовлетворяющей неравенству* (21), *то утверждение теоремы* 1 *остается справедливым.*

Замечания. 1. В [4], [5] при изучении операторного метода регуляризации для вариационных неравенств требование (21) на порядок роста функции $a(R)$ удалось снять за счет введения в регуляризованную задачу (10) вместо αx_λ^μ регуляризующего слагаемого $\alpha E^a x_\lambda^\mu$, где оператор $E^a : H \longrightarrow H$, $E^a(0) = 0$, $E^a x = x a(\|x\|)/\|x\|$ при $x \ne 0$. Для квазивариационного неравенства (2) такой прием не снимает указанную проблему, так как оператор E^a в общем случае не является сильно монотонным и липшиц-непрерывным на всем H (см. [6]), что оставляет открытым вопрос о разрешимости регуляризованного таким образом квазивариационного неравенства.

2. Опустим в условиях теоремы 1 предположение (1). Тогда из (11) при $\lambda \longrightarrow \lambda_0$ имеем $M \le 1$. В [1] при доказательстве леммы 2 для решений (2) элементов x_1 и x_2 установлено неравенство

$$m\|x_1 - x_2\|^2 \le (Ax_1 - Ax_2, Bx_1 - Bx_2),$$

что с учетом предположений б) и в) дает

$$(m - L_1 L_2)\|x_1 - x_2\|^2 \le 0. \tag{22}$$

Поскольку $M \le 1$, то $m - L_1 L_2 > 0$, поэтому из (22) следует равенство $x_1 = x_2$. Таким образом, если квазивариационное неравенство (2) имеет решение, то в наших предположениях а)–в), (11) и (12) оно единственно.

В работе [1] для квазивариационного неравенства (2) построен непрерывный метод первого порядка следующего вида:

$$\xi'(t) + \xi(t) - B\xi(t) = \mathrm{Pr}_\Omega(\xi(t) - B\xi(t) - \gamma(t)A\xi(t)), \quad \xi(t_0) = \xi_0 \in H, \tag{23}$$

сходимость которого установлена при условиях

$$0 < \gamma^0 \le \gamma(t) \le \gamma^1, \tag{24}$$

$$M_1 = \gamma^1 L_1 + L_2 \le 2, \quad M_2 = 2m\gamma^0 - L_2 - L_1\gamma^1(1 + 2L_2) = b > 0. \tag{25}$$

Здесь $\gamma(t)$ – непрерывная при $t \ge t_0$ функция, $\mathrm{Pr}_\Omega : H \longrightarrow \Omega$ – оператор метрического проектирования в H на выпуклое замкнутое множество Ω. Далее считаем, что (24) имеет место для некоторой непрерывной функции $\gamma(t)$, $t \ge t_0 \ge 0$.

Пусть данные задачи (2) возмущены следующим образом: при всех $t \ge t_0 \ge 0$ известны семейства операторов $A(t) : H \longrightarrow H$, $B(t) : H \longrightarrow H$, семейство элементов $f(t) \in H$ и семейство выпуклых замкнутых множеств $\Omega(t) \subset H$, причем верно следующее:

а$_t$) оператор $A(t)$ является сильно монотонным на H с постоянной $m(t) > 0$;

б$_t$) операторы $A(t)$ и $B(t)$ непрерывны по t и удовлетворяют на H условию Липшица с положительными постоянными $L_1(t)$ и $L_2(t)$ соответственно;

в$_t$) $m(t) \longrightarrow m$, $L_i(t) \longrightarrow L_i$ при $t \longrightarrow \infty$, $i = 1, 2$;

г$_t$) справедливы неравенства

$$\|A(t)u - Au\| \le h(t)g(\|u\|), \quad \|B(t)u - Bu\| \le h(t)q(\|u\|) \quad \forall u \in H, \tag{26}$$

$$\|f - f(t)\| \le \delta(t), \tag{27}$$

$$r_H(\Omega, \Omega(t)) \le \sigma(t); \tag{28}$$

здесь неотрицательные функции $h(t)$, $\delta(t)$ и $\sigma(t)$ являются бесконечно малыми при $t \longrightarrow \infty$, функции $g(s)$ и $q(s)$ такие же, как в условиях (4) и (5), $f(t)$ непрерывна по t, $t \ge t_0$.

Неравенства (25) для метода (23) с возмущенными данными примут вид

$$M_1(t) = \gamma(t)L_1(t) + L_2(t) \le 2, \quad M_2(t) = 2m(t)\gamma(t) - L_2(t) - L_1(t)\gamma(t)(1 + 2L_2(t)) = \bar{b}(t) > b_0 > 0. \tag{29}$$

Требования (29) могут быть нарушены и при достаточно больших t, несмотря на выполнение (25) и в$_t$).

Подправим метод (23) с тем, чтобы ослабить условия сходимости непрерывного метода первого порядка.

Сделаем дополнительное предположение:

д$_t$) пусть $\alpha(t)$ и $\beta(t)$ – положительные непрерывные функции, $\beta(t) < 1$, $t \ge t_0$, и

$$\alpha(t) \longrightarrow 0, \quad \beta(t) \longrightarrow 1 \ \text{ при } \ t \longrightarrow \infty. \tag{30}$$

Построим для квазивариационного неравенства (2) непрерывный метод следующего вида:

$$\begin{gathered} u'(t) + u(t) - \beta(t)B(t)u(t) = \mathrm{Pr}_{\Omega(t)}(u(t) - \beta(t)B(t)u(t) - \gamma(t)[A(t)u(t) + \alpha(t)u(t) - f(t)]), \\ u(t_0) = u_0 \in H, \end{gathered} \tag{31}$$

где $\mathrm{Pr}_{\Omega(t)} : H \longrightarrow \Omega(t)$ – оператор метрического проектирования в H на выпуклое замкнутое множество $\Omega(t)$. Пусть $\tilde{L}_1$, $\tilde{L}_2$, $\tilde{\alpha}$ – положительные постоянные, удовлетворяющие неравенствам

$$L_i(t) \le \tilde{L}_i, \quad i = 1, 2, \quad \alpha(t) \le \tilde{\alpha} \quad \forall t \ge t_0, \tag{32}$$

что возможно в силу наших предположений в$_t$) и (30). Тогда для отображения $F(t, z) = -z + \beta(t)B(t)z + \mathrm{Pr}_{\Omega(t)}(z - \beta(t)B(t)z - \gamma(t)[A(t)z + \alpha(t)z - f(t)])$ с учетом (32), нерастяжимости оператора $\mathrm{Pr}_{\Omega(t)}$ при фиксированном t и предположений б$_t$), д$_t$) (24) имеем

$$\|F(t, z_1) - F(t, z_2)\| \le [2 + 2\tilde{L}_2 + \gamma^1(\tilde{L}_1 + \tilde{\alpha})]\|z_1 - z_2\| \quad \forall z_1, z_2 \in H.$$

Исследуем вопрос о непрерывности оператора $\mathrm{Pr}_{\Omega(t)}$ по t, $t \ge t_0$. Пусть

$$r_H(\Omega(t), \Omega(t')) \le \tilde{\xi}(t, t'), \quad t \ge t_0, \quad t' \ge t_0, \tag{33}$$

где $\tilde{\xi}(t, t')$ – некоторая неотрицательная функция,

$$\tilde{\xi}(t, t') = \tilde{\xi}(t', t) \longrightarrow 0 \ \text{ при } \ t' \longrightarrow t. \tag{34}$$

Введем обозначения $z_t = \Pr_{\Omega(t)}z$, $z_{t'} = \Pr_{\Omega(t')}z$, $z \in H$. Тогда справедливы неравенства (см., например, [5, с. 76], [7, с. 189])

$$(z_t - z, z_t - y) \le 0 \quad \forall y \in \Omega(t), \tag{35}$$

$$(z_{t'} - z, z_{t'} - y) \le 0 \quad \forall y \in \Omega(t'). \tag{36}$$

Предположение (33) обеспечивает существование элементов $y_t \in \Omega(t)$ и $y_{t'} \in \Omega(t')$ таких, что

$$\|z_t - y_{t'}\| \le \tilde{\xi}(t, t'), \quad \|z_{t'} - y_t\| \le \tilde{\xi}(t, t'). \tag{37}$$

В (35) положим $y = y_t$, а в (36) положим $y = y_{t'}$; сложив полученные результаты, получим

$$\|z_t - z_{t'}\|^2 \le (\|z_t - z\| + \|z_{t'} - z\|)\tilde{\xi}(t, t') \le (\|z_t\| + \|y_{t'}\| + 2\|z\|)\tilde{\xi}(t, t'). \tag{38}$$

Выберем в Ω некоторый фиксированный элемент x_0; тогда в силу предположения (28) найдется элемент $y(t) \in \Omega(t)$ такой, что $\|x_0 - y(t)\| \le \sigma(t) \le \bar{\sigma}$ при всех $t \ge t_0$, т.е. $\|y(t)\| \le \|x_0\| + \bar{\sigma}$. Так как $\|z_t - z\| \le \|y - z\|$ при всех $y \in \Omega(t)$, то $\|z_t\| \le 2\|z\| + \|x_0\| + \bar{\sigma}$. Аналогично получаем оценку $\|z_{t'}\| \le 2\|z\| + \|x_0\| + \bar{\sigma}$. Теперь из (38) имеем (ср. с [8])

$$\|z_t - z_{t'}\|^2 \le 2(3\|z\| + \|x_0\| + \bar{\sigma})\tilde{\xi}(t, t'), \tag{39}$$

и свойство (34) функции $\tilde{\xi}(t, t')$ обеспечивает непрерывность оператора $\Pr_{\Omega(t)}$ по t при $t \ge t_0$. Следовательно, в наших предположениях существует единственное решение задачи Коши (31) класса $C^1[t_0, +\infty)$ (см. [9, с. 399]). Предположим, что это решение ограничено на $[t_0, +\infty)$, и исследуем его поведение при $t \longrightarrow +\infty$. От (31) перейдем к эквивалентной задаче

$$\begin{gathered}(u'(t) + \gamma(t)[A(t)u(t)) + \alpha(t)u(t) - f(t)], u'(t) + u(t) - \beta(t)B(t)u(t) - y) \le 0 \quad \forall y \in \Omega(t),\\ u'(t) + u(t) - \beta(t)B(t)u(t) \in \Omega(t).\end{gathered} \tag{40}$$

Так как решение x квазивариационного неравенства (2) представимо в виде $x = \tilde{x} + Bx$, $\tilde{x} \in \Omega$, то условие (28) гарантирует существование элемента $v(t) \in \Omega(t)$ такого, что

$$\|\tilde{x} - v(t)\| \le \sigma(t) \quad \forall t \ge t_0. \tag{41}$$

Для $u'(t) + u(t) - \beta(t)B(t)u(t) \in \Omega(t)$ согласно (28) найдется $w(t) \in \Omega$ и

$$\|u'(t) + u(t) - \beta(t)B(t)u(t) - w(t)\| \le \sigma(t) \quad \forall t \ge t_0. \tag{42}$$

Полагая в (2), умноженном на $\gamma(t)$, y равным $w(t) + Bx \in \Omega(x)$, а в (40) полагая y равным $v(t)$ и складывая полученные неравенства, имеем

$$\begin{gathered}\gamma(t)(Ax - f, u'(t) + u(t) - \beta(t)B(t)u(t) - w(t)) + (u'(t) + \gamma(t)[A(t)u(t) + \alpha(t)u(t) - f(t)], \tilde{x} - v(t)) +\\ + \gamma(t)(Ax - f, x - u(t) - u'(t) + \beta(t)B(t)u(t) - Bx) +\\ + (u'(t) + \gamma(t)[A(t)u(t) + \alpha(t)u(t) - f(t)], u'(t) + u(t) - x + Bx - \beta(t)B(t)u(t)) \le 0.\end{gathered} \tag{43}$$

Используя б$_t$), (26) и ограниченность $u(t)$, нетрудно установить (см. вывод (16)), что существуют постоянные $d_1 > 0$, $d_2 > 0$ такие, что $\|A(t)u(t)\| \le d_1$, $\|B(t)u(t)\| \le d_2$ при всех $t \ge t_0$. Тогда из (31) следует ограниченность $\|u'(t)\|$ при $t \ge t_0$.

Определим функцию $\rho(t) = \|u(t) - x\|^2/2$, тогда $\rho'(t) = (u'(t), u(t) - x)$, $\rho(t_0) = \|u_0 - x\|^2/2 = \rho_0$. Несложными преобразованиями, приняв во внимание (41), (42), от (43) придем к неравенству

$$\begin{gathered}\frac{d\rho(t)}{dt} + 2\gamma(t)\alpha(t)\rho(t) + \|u'(t)\|^2 + \gamma(t)[(Ax - A(t)x, x - u(t) - u'(t) + \beta(t)B(t)u(t) - Bx) +\\ + (A(t)x - A(t)u(t), x - u(t)) + \alpha(t)(x, u(t) - x) - (f - f(t), x - u(t)) +\\ + (A(t)x - A(t)u(t) - \alpha(t)u(t) - f + f(t), \beta(t)B(t)u(t) - Bx - u'(t))] +\\ + (u'(t), Bx - \beta(t)B(t)u(t)) \le \sigma(t)c_4, \quad c_4 > 0.\end{gathered} \tag{44}$$

Использовав б$_t$), в$_t$), (26), (30), установим некоторые вспомогательные оценки:

1) $\|\beta(t)B(t)u(t) - Bx\| \le \beta(t)[\|B(t)u(t) - B(t)x\| + \|B(t)x - Bx\|] + [1 - b(t)]\|Bx\| \le [1 - \beta(t) + h(t)]c_5 + \beta(t)L_2(t)\|u(t) - x\|$;

2) $(u'(t), \beta(t)B(t)u(t) - Bx) \le \{[1 - \beta(t) + h(t)]c_5 + \beta(t)L_2(t)\|u(t) - x\|\}\|u'(t)\| \le [1 - \beta(t) + h(t)]c_6 + \beta(t)L_2(t)[\rho(t) + \|u'(t)\|^2/2]$;

3) $(A(t)u(t) - A(t)x, \beta(t)B(t)u(t) - Bx) \le L_1(t)\|u(t) - x\|\{[1 - \beta(t) + h(t)]c_5 + \beta(t)L_2(t)\|u(t) - x\|\} \le 2\beta(t)L_1(t)L_2(t)\rho(t) + [1 - \beta(t) + h(t)]c_7$;

4) $(A(t)x - A(t)u(t), u'(t)) \le L_1(t)\|u(t) - x\|\|u'(t)\| \le L_1(t)\rho(t) + L_1(t)\|u'(t)\|^2/2$.

Здесь c_5, c_6, c_7 – положительные постоянные. Отметим, что при выводе оценок 2) и 4) было использовано числовое неравенство $ab \le a^2/2 + b^2/2$. Теперь с учетом оценок 1)–4), предположений а$_t$), (24), (26), (27) и ограниченности $u(t)$ и $u'(t)$ при $t \ge t_0$ из (44) выводим неравенство

$$\begin{gathered}\|u'(t)\|^2\left[1 - \frac{\gamma(t)L_1(t)}{2} - \frac{\beta(t)L_2(t)}{2}\right] + \frac{d\rho(t)}{dt} + \\ + \rho(t)\{2\gamma(t)[m(t) + \alpha(t)] - 2\gamma(t)\beta(t)L_1(t)L_2(t) - \gamma(t)L_1(t) - \beta(t)L_2(t)\} \le \\ \le c_8[\delta(t) + h(t) + \sigma(t) + \alpha(t) + 1 - \beta(t)], \quad c_8 > 0.\end{gathered} \tag{45}$$

Теорема 3. *Пусть H – вещественное гильбертово пространство, операторы A и B, множество Ω удовлетворяют условиям теоремы* 1, $f \in H$, *и верно условие* (1). *Предположим, что данные квазивариационного неравенства* (2) *возмущены и вместо A, B, Ω и f известны семейства* $\{A(t)\}$, $\{B(t)\}$, $\{\Omega(t)\}$ *и* $\{f(t)\}$ *при $t \ge t_0 \ge 0$ такие, что выполнены предположения* а$_t$)–д$_t$), (33), (34). *Пусть $\alpha(t)$, $\beta(t)$, $\gamma(t)$ – непрерывные положительные функции, $\beta(t) < 1$, $t \ge t_0$, и верно* (24). *Тогда существует единственное решение $u(t) \in C^1[t_0, +\infty)$ задачи Коши* (31), *и если при всех $t \ge t_0$ имеют место неравенства*

$$2 - \gamma(t)L_1(t) - \beta(t)L_2(t) = \tilde{a}(t) \ge 0, \tag{46}$$

$$2\gamma(t)[m(t) + \alpha(t)] - \gamma(t)L_1(t)[1 + 2\beta(t)L_2(t)] - \beta(t)L_2(t) = \tilde{b}(t) \ge \tilde{b}_0 > 0, \tag{47}$$

то траектория $u(t)$ при $t \longrightarrow +\infty$ стабилизируется по норме H к единственному решению x квазивариационного неравенства (2).

Доказательство. В силу предположения (46), от (45) приходим к неравенству

$$\frac{d\rho(t)}{dt} \le -\tilde{b}(t)\rho(t) + p(t), \tag{48}$$

где $p(t) = c_8[\delta(t) + h(t) + \sigma(t) + \alpha(t) + 1 - \beta(t)]$, функция $\tilde{b}(t)$ определена в (47). Теперь, применив лемму из [10, с. 264], приходим к оценке

$$\rho(t) \le \rho_0 \exp[-\eta(t)] + \int_{t_0}^{t} p(s)\exp[\eta(s) - \eta(t)]ds;$$

здесь $\eta(t) = \int_{t_0}^{t} \tilde{b}(t)dt$, $t \ge t_0$. Отсюда, в силу предположения (47) и стремления к нулю функций $\delta(t)$, $h(t)$, $\sigma(t)$, $\alpha(t)$, $1 - \beta(t)$ при $t \longrightarrow +\infty$, получим утверждение теоремы в предположении ограниченности решения задачи (31). Докажем ограниченность $u(t)$ при $t \ge t_0$. Для этого выберем в Ω произвольный, но фиксированный элемент z_0. В силу условия (28), при каждом $t \ge t_0$ найдется элемент $z(t) \in \Omega(t)$ и $z(t) \longrightarrow z_0$ при $t \longrightarrow +\infty$. Тогда из (40) имеем

$$(u'(t) + \gamma(t)[A(t)u(t) + \alpha(t)u(t) - f(t)], u'(t) + u(t) - \beta(t)B(t)u(t) - z(t)) \le 0,$$

или

$$\begin{gathered}\|u'(t)\|^2 + (u'(t), u(t)) - (u'(t), \beta(t)B(t)u(t)) + \gamma(t)A(t)u(t) + \alpha(t)u(t) - f(t), u'(t)) + \\ + \gamma(t)(A(t)u(t) + \alpha(t)u(t) - f(t), u(t)) - \gamma(t)(A(t)u(t) + \alpha(t)u(t) - f(t), \beta(t)B(t)u(t)) \le \\ \le (u'(t) + \gamma(t)[A(t)u(t) + \alpha(t)u(t) - f(t)], z(t)).\end{gathered} \tag{49}$$

Из предположений б$_t$) и (26) следует, что $\|A(t)u(t)\| \le \|A(t)u(t) - A(t)(0)\| + \|A(t)(0) - A(0)\| + \|A(0)\| \le L_1(t)\|u(t)\| + h(t)g(0) + \|A(0)\|$, т.е. $\|A(t)u(t)\| \le L_1(t)\|u(t)\| + b_1$. Аналогично получаем, что $\|B(t)u(t)\| \le L_2(t)\|u(t)\| + b_2$. Здесь и далее b_j – положительные постоянные. Теперь на основе нерастяжимости оператора проектирования $\Pr_{\Omega(t)}$ из (31) выводим неравенство $\|u'(t)\| \le b_3\|u(t)\| + b_4$. Значит,

$$|(u'(t) + \gamma(t)[A(t)u(t) + \alpha(t)u(t) - f(t)], z(t))| \le b_5\|u(t)\| + b_6.$$

Используя предположения а$_t$), б$_t$), (24) и неравенство $ab \le a^2/2 + b^2/2$, оцениваем остальные слагаемые из (49):

А) $|(u'(t), \beta(t)B(t)u(t))| \le \beta(t)\|u'(t)\|[L_2(t)\|u(t)\| + b_2] \le \beta(t)L_2(t)[\|u'(t)\|^2/2 + \|u(t)\|^2/2] + b_7\|u(t)\| + b_8$;

Б) $\gamma(t)|(A(t)u(t) + \alpha(t)u(t) - f(t), u'(t))| \le \gamma(t)L_1(t)[\|u'(t)\|^2/2 + \|u(t)\|^2/2] + b_9\gamma(t)\alpha(t)\|u(t)\|^2 + b_{10}\|u(t)\| + b_{11}$;

В) $\gamma(t)(A(t)u(t) + \alpha(t)u(t) - f(t), u(t)) \ge \gamma(t)[m(t) + \alpha(t)]\|u(t)\|^2 - b_{12}\|u(t)\|$;

Г) $\gamma(t)\beta(t)|(A(t)u(t) + \alpha(t)u(t) - f(t), B(t)u(t))| \le \gamma(t)\beta(t)L_1(t)L_2(t)\|u(t)\|^2 + b_{13}\alpha(t)\|u(t)\|^2 + b_{14}\|u(t)\| + b_{15}$.

Значит, от (49) перейдем к неравенству

$$(u'(t), u(t)) + \|u'(t)\|^2[1 - \beta(t)L_2(t)/2 - \gamma(t)L_1(t)/2] +$$
$$+\frac{\|u(t)\|^2}{2}\{2\gamma(t)[m(t) + \alpha(t)] - \beta(t)L_2(t) - \gamma(t)L_1(t)[1 + 2\beta(t)L_2(t)]\} \le \tag{50}$$
$$\le m_1[\alpha(t)\|u(t)\|^2/2 + \|u(t)\| + 1], \quad m_1 > 0.$$

Очевидны равенства

$$(u'(t), u(t)) = \frac{d}{dt}\left(\frac{\|u(t)\|^2}{2}\right) = \|u(t)\|\frac{d\|u(t)\|}{dt}.$$

Следовательно, в предположениях (46), (47) неравенство (50) сводится к следующему:

$$\|u(t)\|\frac{d\|u(t)\|}{dt} + \tilde{b}_0\|u(t)\|^2/2 \le m_1[\alpha(t)\|u(t)\|^2/2 + \|u(t)\| + 1].$$

Пусть

$$\|u(t)\| \longrightarrow \infty \text{ при } t \longrightarrow +\infty. \tag{51}$$

Тогда хотя бы при достаточно больших t имеем неравенство

$$\frac{d\|u(t)\|}{dt} \le [m_1\alpha(t) - \tilde{b}_0]\|u(t)\|/2 + m_2, \quad m_2 > 0.$$

Так как $\alpha(t) \longrightarrow 0$ при $t \longrightarrow +\infty$, то последнее неравенство и предположение (51) обеспечивают существование числа $\bar{t} \ge t_0$ такого, что при $t \ge \bar{t}$ справедливо неравенство $d\|u(t)\|/dt < 0$. Получили противоречие предположению (51). Теорема доказана полностью.

Отметим, что условия (46), (47) сходимости регуляризованного непрерывного метода (31) слабее, чем (29).

Построим дискретный вариант непрерывного метода (31). Пусть приближения к A, B, f и Ω определяются последовательностями операторов $\{A^k\}$, $\{B^k\}$, элементов $\{f^k\}$ из H и выпуклых замкнутых множеств $\{\Omega^k\}$ из H, причем при каждом натуральном k и всех u и v из H верны свойства:

а$_k$) $(A^k u - A^k v, u - v) \ge m^k\|u - v\|^2$, $m^k > 0$;

б$_k$) $\|A^k u - A^k v\| \le L_1^k\|u - v\|$, $L_1^k > 0$;

в$_k$) $\|B^k u - B^k v\| \le L_2^k\|u - v\|$, $L_2^k > 0$;

г$_k$) $m^k \longrightarrow m$, $L_i^k \longrightarrow L_i$ при $k \longrightarrow \infty$, $i = 1, 2$;

д$_k$) $\|Au - A^k u\| \le h_k g(\|u\|)$, $\|Bu - B^k u\| \le h_k q(\|u\|)$, $\|f - f^k\| \le \delta_k$, здесь функции $g(s)$ и $q(s)$ такие же, как в (4) и (5);

е$_k$) $r_H(\Omega, \Omega^k) \le \sigma_k$.

Пусть элемент u_{k+1} по u_k находится из уравнения

$$\frac{u_{k+1} - u_k}{\tau_k} + u_{k+1} - \beta_k B^k u_{k+1} = \Pr_{\Omega^k}(u_{k+1} - \beta_k B^k u_{k+1} - \gamma_k(A^k u_{k+1} + \alpha_k u_{k+1} - f^k)), \tag{52}$$

элемент $u_0 \in H$ задается.

Теорема 4. *Пусть H – вещественное гильбертово пространство, для квазивариационного неравенства* (2) *выполнены условия теоремы* 1. *Предположим, что A, B, f и Ω заданы приближенно и их приближения, определяемые последовательностями $\{A^k\}$, $\{B^k\}$, $\{f^k\}$ и $\{\Omega^k\}$, удовлетворяют предположениям* а$_k$)–е$_k$); *$\{\alpha_k\}$, $\{\beta_k\}$, $\{\gamma_k\}$, $\{\tau_k\}$ – последовательности положительных чисел, причем $\alpha_k \longrightarrow 0$, $\beta_k \longrightarrow 1$, $\beta_k \in (0, 1)$, $\tau_k \le \bar{\tau}$ при всех натуральных k. Пусть $\{\delta_k\}$, $\{h_k\}$, $\{\sigma_k\}$ являются бесконечно малыми последовательностями неотрицательных чисел и*

$$\beta_k L_2^k\left(\frac{1+\gamma_k\tau_k(L_1^k+\alpha_k)}{1+\gamma_k\tau_k(m^k+\alpha_k)}+1\right) < \frac{1+\tau_k}{\tau_k}, \tag{53}$$

$$2-\gamma_k L_1^k - \beta_k L_2^k = \tilde{a}_k \ge 0, \tag{54}$$

$$2\gamma_k(m^k+\alpha_k) - \gamma_k L_1^k(1+2\beta_k L_2^k) - \beta_k L_2^k = \tilde{b}_k \ge d_0 > 0, \tag{55}$$

$$0 < \tilde{\gamma}^0 \le \gamma_k \le \tilde{\gamma}^1, \tag{56}$$

числовой ряд

$$\sum_{k=1}^{\infty}\mu_k, \quad \mu_k = \frac{\tau_k}{1+d_0\tau_k} \tag{57}$$

расходится. Тогда из уравнения (52) *элемент u_{k+1} определяется однозначно и последовательность $\{u_k\}$ сильно сходится при $k \longrightarrow +\infty$ к единственному решению x квазивариационного неравенства* (2).

Доказательство. Прежде всего установим однозначную разрешимость уравнения (52). Для этого от (52) перейдем к эквивалентной вариационной задаче

$$\left(\frac{u_{k+1}-u_k}{\tau_k}+\gamma_k(A^k u_{k+1}+\alpha_k u_{k+1}-f^k), \frac{u_{k+1}-u_k}{\tau_k}+u_{k+1}-\beta_k B^k u_{k+1}-y\right) \le 0 \quad \forall y \in \Omega^k. \tag{58}$$

Из (52) следует, что элемент

$$\xi_k = \eta_k + u_{k+1} - \beta_k B^k u_{k+1} \in \Omega^k, \quad \eta_k = \frac{u_{k+1}-u_k}{\tau_k}.$$

Значит, элемент $v_k = \eta_k + u_{k+1} \in \Omega_k + \beta_k B^k u_{k+1} = \Omega_k^{\beta}$. Тогда $u_{k+1} = (\tau_k v_k + u_k)/(1+\tau_k)$, $\eta_k = (v_k - u_k)/(1+\tau_k)$ и (58) в новых обозначениях примет вид

$$\left(\frac{v_k-u_k}{1+\tau_k}+\gamma_k\left[A^k\left(\frac{\tau_k v_k+u_k}{1+\tau_k}\right)+\alpha_k\frac{\tau_k v_k+u_k}{1+\tau_k}-f^k\right], v_k - z\right) \le 0, \quad v_k \in \Omega_k^{\beta} \ \ \forall z \in \Omega_k^{\beta},$$

т.е. пришли к квазивариационному неравенству типа (2) относительно элемента v_k. Записав для него условие (1) однозначной разрешимости, получаем (53) (см. [1]). Следовательно, для любого $u_k \in H$ из (52) (или (58)) однозначно определяется элемент $u_{k+1} \in H$.

Из предположений б$_k$)–д$_k$) легко выводятся оценки (см. вывод (16))

$$\|A^k u_{k+1}\| \le a_1\|u_{k+1}\| + a_2, \quad \|B^k u_{k+1}\| \le a_3\|u_{k+1}\| + a_4,$$

учитывая которые, из (52) имеем

$$\|\eta_k\| \le a_5\|u_{k+1}\| + a_6.$$

Здесь и далее a_j – положительные постоянные, не зависящие от k. Следовательно,

$$\|\eta_k + \gamma_k(A^k u_{k+1} + \alpha_k u_{k+1} - f^k)\| \le a_7\|u_{k+1}\| + a_8.$$

Кроме того, в силу предположений а$_k$)–в$_k$) и (56), аналогично А)–Г) получаем такие оценки

I) $|(\eta_k, \beta_k B^k u_{k+1})| \le \beta_k L_2^k(\|\eta_k\|^2/2 + \|u_{k+1}\|^2/2) + a_9\|u_{k+1}\| + a_{10}$;

II) $\gamma_k|(A^k u_{k+1} + \alpha_k u_{k+1} - f^k, \eta_k)| \le \gamma_k L_1^k(\|\eta_k\|^2/2 + \|u_{k+1}\|^2/2) + a_{11}\alpha_k\gamma_k\|u_{k+1}\|^2 + a_{12}\|u_{k+1}\| + a_{13}$;

III) $\gamma_k(A^k u_{k+1} + \alpha_k u_{k+1} - f^k, u_{k+1})| \ge \gamma_k(m^k + \alpha_k)\|u_{k+1}\|^2 - a_{14}\|u_{k+1}\|$;

IV) $\gamma_k\beta_k|(A^k u_{k+1} + \alpha_k u_{k+1} - f^k, B^k u_{k+1})| \le \gamma_k\beta_k L_1^k L_2^k \|u_{k+1}\|^2 + a_{15}\alpha_k\|u_{k+1}\|^2 + a_{16}\|u_{k+1}\| + a_{17}$.

Легко убедиться в справедливости неравенства

$$\text{V)}\ (\eta_k, u_{k+1}) \ge \frac{\|u_{k+1}\| - \|u_k\|}{\tau_k}\|u_{k+1}\|.$$

Покажем ограниченность последовательности $\{u_k\}$. Пусть $y_k \longrightarrow y \in \Omega$, $y_k \in \Omega^k$, y – фиксированный произвольный элемент из Ω. Существование элементов y_k обеспечивает предположение e_k). Полагая $y = y_k$ в (58), с учетом оценок I)–V) и (56) из полученного неравенства подобно (50) устанавливаем справедливость неравенства

$$\frac{\|u_{k+1}\| - \|u_k\|}{\tau_k}\|u_{k+1}\| + \|\eta_k\|^2(1 - \gamma_k L_1^k/2 - \beta_k L_2^k/2) + \frac{\|u_{k+1}\|^2}{2}[2\gamma_k(m^k + \alpha_k) - \gamma_k L_1^k(1 + 2\beta_k L_2^k) - \beta_k L_2^k] \le$$
$$\le \tilde{m}(\alpha_k\|u_{k+1}\|^2/2 + \|u_{k+1}\| + 1), \quad \tilde{m} > 0.$$

В силу предположений (54) и (55), из последнего неравенства получаем

$$\frac{\|u_{k+1}\| - \|u_k\|}{\tau_k}\|u_{k+1}\| + d_0\|u_{k+1}\|^2/2 \le \tilde{m}(\alpha_k\|u_{k+1}\|^2/2 + \|u_{k+1}\| + 1).$$

Отсюда делаем вывод об ограниченности последовательности $\{u_k\}$ (см. доказательство теоремы 3).

Поскольку $y = x - Bx \in \Omega$, то, согласно предположению e_k), при каждом k найдется элемент $z_k \in \Omega^k$ такой, что

$$\|z_k - x + Bx\| \le \sigma_k.$$

Далее, для элементов $\xi_k \in \Omega^k$ найдем $w_k \in \Omega$, обладающие свойством

$$\|\eta_k + u_{k+1} - \beta_k B^k u_{k+1} - w_k\| \le \sigma_k.$$

Полагая, что в (2), умноженном на γ_k, $y = w_k + Bx \in \Omega(x)$, а в (58) – что $y = z_k$, и складывая полученные неравенства, имеем

$$\begin{aligned}\gamma_k(Ax - f, \eta_k + u_{k+1} - \beta_k B^k u_{k+1} - w_k) + \gamma_k(Ax - f, x - \eta_k - u_{k+1} + \beta_k B^k u_{k+1} - Bx) +\\ + (\eta_k + \gamma_k(A^k u_{k+1} + \alpha_k u_{k+1} - f^k), x - Bx - z_k) +\\ + (\eta_k + \gamma_k(A^k u_{k+1} + \alpha_k u_{k+1} - f^k), \eta_k + u_{k+1} - x + Bx - \beta_k B^k u_{k+1}) \le 0.\end{aligned} \tag{59}$$

Определим величину $\rho_k = \|u_k - x\|^2/2$; тогда (см. [7, с. 164])

$$\rho_{k+1} - \rho_k \le (u_{k+1} - u_k, u_{k+1} - x).$$

Теперь из (59) с учетом последнего неравенства, (54) и (55) подобно (48) выводим, что

$$\rho_{k+1} \le (1 - \lambda_k)\rho_k + \tilde{\mu} p_k \mu_k, \quad k = 1, 2, \dots, \tag{60}$$

где $\tilde{\mu} > 0$, $\mu_k = \tau_k/(1 + d_0\tau_k)$, $\lambda_k = d_0\mu_k$, $p_k = \delta_k + h_k + \sigma_k + \alpha_k + 1 - \beta_k$. Так как при наших предположениях $p_k\mu_k/\lambda_k \longrightarrow 0$ при $k \longrightarrow \infty$, то расходимость ряда (57) обеспечивает сходимость к нулю последовательности $\{\rho_k\}$ (см. [11], [12, с. 83]). Теорема доказана.

Замечания. 3. Аналогично теореме 2, условия (28) и e_k) в теоремах 3 и 4 можно заменить, соответственно, более слабыми (см. (20))

$$\tau(R, \Omega, \Omega(t)) \le a(R)\sigma(t), \quad \tau(R, \Omega, \Omega^k) \le a(R)\sigma_k,$$

где функция $a(R)$ удовлетворяет (21).

4. Неравенство (19) определяет скорость сходимости операторного метода (10). Из (47) имеем $\tilde{b}(t) \ge$ $\ge \tilde{b}_0 > 0$, поэтому $\eta(t) = \int_{t_0}^{t}\tilde{b}(t)dt \ge \tilde{b}_0(t - t_0)$. Значит, из (48) вытекает оценка

$$\rho(t) \le \rho_0 \exp[-\tilde{b}_0(t-t_0)] + \int_{t_0}^{t} p(s)\exp[\eta(s)]ds\{\exp[\eta(t)]\}^{-1}.$$

Применяя к последнему слагаемому правой части правило Лопиталя, приходим к следующей оценке скорости сходимости непрерывного метода (31):

$$\rho(t) \le \bar{C}[\exp(-\tilde{b}_0 t) + p(t)], \quad \bar{C} > 0.$$

Из (60) для итерационного метода (52) следует оценка (см. [11])

$$\rho_{k+1} \le \tilde{C}[\exp(-T_k) + p_k], \quad \tilde{C} > 0, \quad T_k = d_0\sum_{i=1}^{k}\mu_i.$$

Квазивариационные неравенства вида (2) изучались в [13]–[15], где отмечено их прикладное значение.

В работах [16]–[18] изучались методы решения квазивариационных неравенств вида, отличного от (2).

СПИСОК ЛИТЕРАТУРЫ

1. *Рязанцева И.П.* Методы первого порядка для некоторых квазивариационных неравенств в гильбертовом пространстве // Ж. вычисл. матем. и матем. физ. 2007. Т. 47. № 2. С. 189–196.
2. *Байокки К., Капело А.* Вариационные и квазивариационные неравенства. Приложения к задачам со свободной границей. М.: Наука, 1988.
3. *Красносельский М.А., Вайникко Г.М., Забрейко П.П. и др.* Приближенное решение операторных уравнений. М.: Наука, 1969.
4. *Рязанцева И.П.* Вариационные неравенства с монотонными операторами на множествах, заданных приближенно // Ж. вычисл. матем и матем. физ. 1984. Т. 24. № 6. С.932–936.
5. *Alber Ya., Ryazantseva I.* Nonlinear ill-posed problems of monotone type. Dordrecht: Spirnger, 2006.
6. *Владимиров А.А., Нестеров Ю.Е., Чеканов Ю.Н.* О равномерно выпуклых функционалах // Вестн. МГУ. Сер. 15. 1978. № 3. С. 12–23.
7. *Васильев Ф.П.* Численные методы решения экстремальных задач. М.: Наука, 1988.
8. *Alber Ya.I., Notik A.I.* On some estimates for projection operator in a Banach space // Communs Appl. Nonlinear Analys. 1995. V. 2. № 1. P. 47–56.
9. *Треногин В.А.* Функциональный анализ. М.: Наука, 1988.
10. *Васильев Ф.П.* Методы решения экстремальных задач. М.: Наука, 1981.
11. *Апарцин А.С.* К построению сходящихся итерационных процессов в гильбертовом пространстве // Тр. по прикл. матем. и кибернетике. Иркутск, 1972. С. 7–14.
12. *Васин В.В., Агеев А.Л.* Некорректные задачи с априорной информацией. Екатеринбург: Уральская издат. фирма “Наука”, 1993.
13. *Noor M.A.* The quasi-complementarity problem // J. Math. Analys. and Appl. 1988. V. 130. № 2. P. 344–353.
14. *Siddiqi A.H., Ansari Q.H.* Strongly nonlinear quasivariational inequalities // J. Math. Analys. and Appl. 1990. V. 149. № 2. P. 444–450.
15. *Siddiqi A.H., Ansari Q.H.* General strongly nonlinear variational inequalities // J. Math. Analys. and Appl. 1992. V. 166. № 2. P. 386–392.
16. *Антипин А.С.* Методы решения вариационных неравенств со связанными ограничениями // Ж. вычисл. матем. и матем. физ. 2000. Т. 40. № 9. С. 1291–1307.
17. *Антипин А.С.* Решение вариационных неравенств со связанными ограничениями с помощью дифференциальных уравнений // Дифференц. ур-ния. 2000. Т. 36. № 11. С. 1443–1451.
18. *Антипин А.С., Васильев Ф.П.* Методы регуляризации для решения неустойчивых задач равновесного программирования со связанными ограничениями // Ж. вычисл. матем. и матем. физ. 2005. Т. 45. № 1. С. 27–40.

ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 2007, том 47, № 8, с. 1298–1307

УДК 517.518.475+519.644.7

О РАЗЛОЖЕНИИ ФУНКЦИЙ ДВУХ ПЕРЕМЕННЫХ В СМЕШАННЫЕ РЯДЫ ФУРЬЕ–ЯКОБИ И ПРИЛОЖЕНИЯ ИХ К ОЦЕНКЕ ПОГРЕШНОСТИ КУБАТУРНЫХ ФОРМУЛ[1)]

© 2007 г. В. А. Абилов*, Г. А. Джалаева*, М. К. Керимов**

*(*367025 Махачкала, ул. М. Гаджиева, 43а, Дагестанский гос. ун-т;*

***119991 Москва, ул. Вавилова, 40, ВЦ РАН)*

Поступила в редакцию 16.03.2007 г.

Обсуждаются некоторые вопросы разложения функций двух переменных в смешанные ряды Фурье–Якоби. В частности, даны оценки скорости сходимости этих рядов на классах функций двух переменных, характеризующихся обобщенными модулями непрерывности. Указаны приложения этих результатов к оценке остатков некоторых смешанных кубатурных формул типа Чебышева. Библ. 9.

Ключевые слова: смешанные ряды Фурье–Якоби, вопросы сходимости, оценки остаточных членов, поперечники, условия абсолютной сходимости.

ВВЕДЕНИЕ

В ряде задач математической физики (например, при решении задачи Дирихле для шара методом разделения переменных, см. [1, с. 424]) возникают двойные ряды вида

$$\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\lambda_{nm}P_n^{(\alpha,\beta)}(x)[a_{nm}(f)\cos(my)+b_{nm}(f)\sin(my)], \tag{1}$$

где

$$P_n^{(\alpha,\beta)}(x) = \sum_{\nu=0}^{n}\binom{n+\alpha}{n-\nu}\binom{n+\beta}{\nu}\left(\frac{x-1}{2}\right)^{\nu}\left(\frac{x+1}{2}\right)^{n-\nu}$$

являются многочленами Якоби, λ_{nm}, $a_{nm}(f)$ и $b_{nm}(f)$ – некоторые коэффициенты. Для краткости назовем (1) смешанным рядом Фурье–Якобсона. Работа посвящена вопросам разложения функций двух переменных $f(x, y)$ в смешанные ряды Фурье–Якоби вида (*). В частности, получим оценки скорости сходимости таких рядов на некоторых классах функций двух переменных, характеризующихся обобщенными модулями непрерывности, построенных для этих целей, установлена связь между структурными свойствами функций двух переменных и скоростью сходимости разложений этих функций в ряды вида (1). В качестве применения полученных результатов найдены оценки некоторых смешанных кубатурных формул типа Чебышёва.

Для чтения работ необходимо знакомство с работами [2], [3] (см. также там библиографию).

1. ОПРЕДЕЛЕНИЯ. КЛАССЫ ФУНКЦИЙ

Пусть $L_2 = L_2((1-x)^\alpha(1+x)^\beta, (-1, 1)\times\mathbb{R})$ – пространство суммируемых с квадратом функций $f: (-1, 1)\times\mathbb{R} \longrightarrow \mathbb{R}$ с весом $(1-x)^\alpha(1+x)^\beta$ $(\alpha > -1, \beta > -1)$, 2π-периодических по второй переменной и с нормой

$$\|f\| = \|f\|_{L_2} = \sqrt{\int_{-1}^{1}\int_{0}^{2\pi}(1-x)^{\alpha}(1+x)^{\beta}f^2(x,y)dxdy}.$$

[1)]Работа М.К. Керимова выполнена при финансовой поддержке РФФИ (код проекта 04-01-00723).

Возьмем формулу Родрига для многочленов $P_x^{(\alpha,\beta)}$:

$$P_n^{(\alpha,\beta)}(x) = \frac{(-1)^n}{n!2^n}(1-x)^{-\alpha}(1+x)^{-\beta}\frac{d^n}{dx^n}[(1-x)^{\alpha+n}(1+x)^{\beta+n}], \quad n = 0, 1, \ldots.$$

Известно (см. [4, с. 70]), что многочлены Якоби составляют ортонормированную систему с условиями ортогональности вида

$$\int_{-1}^{1}(1-x)^{\alpha}(1+x)^{\beta}P_n^{(\alpha,\beta)}(x)P_m^{(\alpha,\beta)}(x)dx = \begin{cases} 0, & n \neq m, \\ \dfrac{2^{\alpha+\beta+1}\Gamma(\alpha+n+1)\Gamma(\beta+n+1)}{(\alpha+\beta+2n+1)\Gamma(n+1)\Gamma(\alpha+\beta+n+1)}, & n = m. \end{cases}$$

Известно также, что система функций

$$P_n^{(\alpha,\beta)}(x)\cos(my), \quad P_b^{(\alpha,\beta)}(x)\sin(my), \quad n, m = 0, 1, \ldots,$$

образует ортогональный базис в пространстве L_2, т.е. для любой функции $f \in L_2$ в топологии пространства L_2, справедливо разложение в двойной ряд (смешанный – ряд Фурье–Якобсона)

$$f(x,y) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\lambda_{nm}P_n^{(\alpha,\beta)}(x)[a_{nm}(f)\cos(my) + b_{nm}(f)\sin(my)], \tag{2}$$

где

$$\lambda_{nm} = \begin{cases} 1/2, & n \geq, \quad m = 0, \\ 1, & n \geq 0, \quad m \geq 1, \end{cases}$$

$$a_{nm}(f) = \frac{(\alpha+\beta+2n+1)\Gamma(n+1)\Gamma(\alpha+\beta+n+1)}{2^{\alpha+\beta+1}\Gamma(\alpha+n+1)\Gamma(\beta+n+1)\pi}\int_{-1}^{1}\int_{0}^{2\pi}(1-\xi)^{\alpha}(1+\xi)^{\beta}f(\xi,\eta)P_n^{(\alpha,\beta)}(\xi)\cos(m\eta)d\xi d\eta,$$

$$b_{nm} = \frac{(\alpha+\beta+2n+1)\Gamma(n+1)\Gamma(\alpha+\beta+n+1)}{2^{\alpha+\beta+1}\Gamma(\alpha+n+1)\Gamma(\beta+n+1)\pi}\int_{-1}^{1}\int_{0}^{2\pi}(1-\xi)^{\alpha}(1+\xi)^{\beta}f(\xi,\eta)P_n^{(\alpha,\beta)}(\xi)\sin(m\eta)d\xi d\eta.$$

В дальнейшем будем рассматривать случай $\alpha = \beta \geq -1/2$ и писать $P_n^{(\beta)}(x) = P_n^{(\beta,\beta)}(x)$. При $\alpha = \beta = -1/2$ многочлены $P_n^{(-1/2)}(x)$ совпадают с известными многочленами Чебышёва I рода

$$T_n(x) = \cos(n\arccos x).$$

Введем обозначения

$$S_N(f; x, y) = \sum_{n^2+m^2<N^2}\lambda_{nm}P_n^{(\alpha,\beta)}(x)[a_{nm}(f)\cos(my) + b_{nm}(f)\sin(my)],$$

$$S_N^*(f; x, y) = \sum_{n(n+2\beta+1)+m^2<N(N+2\beta+1)N^2}\lambda_{nm}P_n^{(\alpha,\beta)}(x)[a_{nm}(f)\cos(my) + b_{nm}(f)\sin(my)],$$

которые выражают так называемые “сферические” частичные суммы ряда (2).

Из уравнения замкнутости Стеклова для функции f

$$\|f\| = \left(\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\lambda_{nm}c_{nm}^2(f)\right)^{1/2}$$

следует, что справедливы соотношения

$$E_N(f) = \|f - S_N(f)\| = \left(\sum_{n^2+m^2 \geq N^2}c_{nm}^2(f)\right)^{1/2},$$

$$E_N^*(f) = \|f - S_N^*(f)\| = \left(\sum_{n(n+2\beta+1)+m^2 \ge N(N+2\beta+1)+N^2} c_{nm}^2(f)\right)^{1/2},$$

где

$$c_{nm}^2(f) = \frac{2^{\alpha+\beta+1}\Gamma(\alpha+n+1)\Gamma(\beta+n+1)\pi}{(\alpha+\beta+2n+1)\Gamma(n+1)\Gamma(\alpha+\beta+n+1)}[a_{nm}^2(f) + b_{nm}^2(f)].$$

Напомним, что величина

$$d_N(\mathbb{M}) = d_N(\mathbb{M}; L_2) = \inf_{G_N \subset L_2}\{\sup_{f \in \mathbb{M}}\{\inf_{g \in G_N}\|f - g\|\}\},$$

где последняя точная нижняя грань берется по всем подпространствам $G_N \subset L_2$ размерности $N \in \mathbb{N}$, называется N-поперечником Колмогорова множества $\mathbb{M} \subset L_2$ (см. [5, с. 186]).

Определим теперь оператор $F_h : L_2 \longrightarrow L_2$, $0 < h < 1$, равенствами

$$F_h^{(\beta)} f(x, y) = \frac{\Gamma(2\beta+1)}{\pi 2^{2\beta+1}\Gamma^2(\beta+1)}\int_{-1}^{1}\int_{-\pi}^{\pi}(1-\xi^2)^{\beta-1/2} f\left(x\cos h - \xi\sqrt{1-x^2}\sin h, y + \frac{h}{\pi}\eta\right)d\xi d\eta,$$

если $\beta > -1/2$, и

$$F_h^{(-1/2)} f(x, y) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[f\left(x\cos h + \sqrt{1-x^2}\sin h, y + \frac{h}{\pi}\eta\right) - f\left(x\cos h - \sqrt{1-x^2}\sin h, y + \frac{h}{\pi}\eta\right)\right]d\eta,$$

если $\beta = -1/2$, и назовем его оператором обобщенного сдвига.

Нетрудно показать (см. [3], [6]), что оператор F_h удовлетворяет следующим условиям:

1) $F_h(f_1 + f_2) = F_h(f_1) + F_h(f_2)$,
2) $F_h(\lambda f) = \lambda F_h(f)$, $\lambda \in \mathbb{R}$,
3) $\|F_h f\| \le M\|f\|$, $M \in \mathbb{R}_+$,
4) $\|F_h f - f\| \longrightarrow 0$, $h \longrightarrow 0+$,
5) $F_h(P_n^{(\beta)}(x)\cos(my)) = \dfrac{P_n^{(\beta)}(\cos h)\sin(mh)}{P_n^{(\beta)}(1)mh}P_n^{(\beta)}(x)\cos(my)$, $F_h(P_n^{(\beta)}(x)\sin(my)) = \dfrac{P_n^{(\beta)}(\cos h)\sin(mh)}{P_n^{(\beta)}(1)mh} \times$
$\times\, P_n^{(\beta)}(x)\sin(my)$.

Определим теперь конечные разности первого и высших порядков функции $f(x, y)$ следующим образом:

$$\Delta_h(f; x, y) = F_h f(x, y) - f(x, y) = (F_h - E) f(x, y),$$

$$\Delta_h^k(f; x, y) = \Delta_h(\Delta_h^{k-1}(f; x, y); x, y) = (F_h - E)^k f(x, y) = \sum_{i=0}^{k}(-1)^{k-i}\binom{k}{i}F_h^i f(x, y),$$

где $F_h^0 f(x, y) = f(x, y)$, $F_h^i f(x, y) = F_h(F_h^{i-1} f(x, y))$, $i = 1, 2, \ldots, k$, $k = 1, 2, \ldots$, и E – единичный оператор в пространстве L_2.

Величину

$$\Omega_k(f; \delta) = \sup_{0 < h \le \delta}\left\|\Delta_h^k(f; x, y)\right\|$$

будем называть обобщенным модулем непрерывности k-го порядка функции $f \in L_2$.

Возьмем дифференциальный оператор второго порядка

$$D = (1-x^2)\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - 2(\beta+1)\frac{\partial}{\partial x}.$$

Рассмотрим следующие классы функций:

1) $L_2^r(D)$ – класс функций $f \in L_2$, имеющих обобщенные частные производные в смысле Леви (см. [7, с. 172])

$$\frac{\partial^k}{\partial x^i \partial y^j} f(x, y), \quad i + j = k, \quad k = 0, 1, \dots,$$

принадлежащие пространству L_2, для которых $D^r f \in L_2$, $r = 0, 1, \dots$, причем

$$D^0 f = f, \quad D^r f = D(D^{r-1} f), \quad r = 1, 2, \dots, \quad L_2^0(D) = L_2;$$

2) $W_2^r(D)$ – класс функций $f \in L_2^r(D)$, для которых справедливы неравенства

$$\|D^r f\| \le 1, \quad r = 1, 2, \dots;$$

3) $W_{2,\Phi}^{r,k}(D)$ – класс функций $f \in L_2^r(D)$, для которых имеют место соотношения

$$\Omega_k(D^r f; \delta) = O(\Phi(\delta^k)), \quad r = 0, 1, \dots, \quad k = 1, 2, \dots,$$

где $\Phi(t)$ – непрерывная, монотонно возрастающая функция на $[0, +\infty)$, $\Phi(0) = 0$;

4) $KH^{(\alpha)}$ – класс функций $f \in L_2$, для которых справедлива оценка

$$\sup_{(x, y) \in [-1, 1] \times \mathbb{R}} |F_h f(x, y) - f(x, y)| \le K h^\alpha,$$

где $K > 0$, $0 < \alpha < 1$;

(Нетрудно заметить, что $KH^{(\alpha)} \subset W_{2, t^\alpha}^{0, 1}(D)$.)

5) $CW_\Phi^{r,k}(D)$ – класс непрерывно дифференцируемых функций $f(x, y)$ ($(x, y) \in [-1, 1] \times \mathbb{R}$), 2π-периодических по второй переменной и удовлетворяющих условию

$$\sup_{0 < h \le \delta} \left\{ \max_{(x, y) \in [-1, 1] \times \mathbb{R}} |\Delta_h^k(D^r f; x, y)| \right\} = O[\Phi(\delta^k)], \quad r = 0, 1, \dots, \quad k = 1, 2, \dots.$$

Очевидно, что $CW_\Phi^{r,k}(D) \in W_{2,\Phi}^{r,k}(D)$.

Сформулируем основные результаты, которые мы хотим доказать.

2. ОСНОВНЫЕ РЕЗУЛЬТАТЫ

Теорема 1. *Справедлива оценка*

$$\sup_{f \in W_{2,\Phi}^{r,k}(D)} \|f - S_N(f)\| = O(N^{-2r} \Phi(N^{-k}))$$

$$\left(\sup_{f \in W_{2,\Phi}^{r,k}(D)} \|f - S_N^*(f)\| = O(N^{-2r} \Phi(N^{-k})) \right), \quad r = 0, 1, \dots, \quad k = 1, 2, \dots,$$

где константа, входящая в $O(1)$, *зависит от* r *и* k.

Теорема 2. *Для любой функции* $f \in L_2$ *имеет место оценка*

$$\Omega_k\left(f; \frac{1}{N}\right) \ll \left(\frac{1}{N^{4k}} \sum_{l=1}^{N} l^{4k-1} E_l^2(f) \right)^{1/2}$$

$$\left(\Omega_k\left(f; \frac{1}{N}\right) \ll \left(\frac{1}{N^{4k}} \sum_{l=1}^{N} l^{4k-1} E_l^{*2}(f) \right)^{1/2} \right).$$

Теорема 3. *Пусть* $f \in L_2$. *Если ряд*

$$\sum_{l=1}^{\infty} l^{2r-1} E_l(f) \left(\sum_{l=1}^{\infty} l^{2r-1} E_l^*(f) \right), \quad r = 1, 2, \ldots,$$

сходится, то $f \in L_2^r(D)$ *и справедливо неравенство*

$$\Omega_k\left(D^r f; \frac{1}{N}\right) \ll \left(\frac{1}{N^{4k}} \sum_{l=1}^{N} l^{4(r+k)-1} E_l^2(f)\right)^{1/2} + \sum_{l=N}^{\infty} l^{2r-1} E_l(f)$$

$$\left(\Omega_k\left(D^r f; \frac{1}{N}\right) \ll \left(\frac{1}{N^{4k}} \sum_{l=1}^{R} l^{4(r+k)-1} E_l^{*2}(f)\right)^{1/2} + \sum_{l=N}^{\infty} l^{2r-1} E_l^*(f)\right).$$

Теорема 4. *При каждом фиксированном* $N = 1, 2, \ldots$ *справедливы равенства*

$$\sup_{f \in W_2^r(D)} \|f - S_N^*(f)\| = \frac{1}{[N(N + 2\beta + 1) + N^2]^r}, \quad r = 1, 2, \ldots.$$

Теорема 5. *Для любой функции* $f \in W_2^r(D)$ *справедливо предельное соотношение*

$$\lim_{N \to \infty} N^{2r} \|f - S_N(f)\| = 0$$

$$(\lim_{N \to \infty} N^{2r} \|f - S_N^*(f)\| = 0).$$

Теорема 6. *Справедливы равенства*

$$d_{m(N)+l}(W_2^r(D); L_2) = \frac{1}{[N(N + 2\beta + 1) + N^2]^r},$$

$$N = 1, 2, \ldots, \quad l = 0, 1, \ldots, \overline{m}(R) - m(R) - 1, \quad r = 1, 2, \ldots,$$

где

$$m(N) = \operatorname{card}\{(n, m) : n(n + 2\beta + 1) + m < N(N + 2\beta + 1) + N^2\},$$

$$\overline{m}(N) = \operatorname{card}\{(n, m) : n(n + 2\beta + 1) + m < N(N + 2\beta + 1) + N^2\}.$$

Теорема 7. *Если* $\Phi(\lambda t) = O(\lambda + 1)\Phi(t)$, $\lambda > 0$, *то справедлива оценка*

$$d_N(W_{2,\Phi}^{r,k}(D); L_2) \asymp N^{-r} \Phi(N^{-k/2}).$$

Теорема 8. *Пусть* $f \in KH^{(\alpha)}$. *Если* $\alpha > 1$, *то ряд*

$$\sum_{n=0}^{\infty} \sum_{m=0}^{\infty} |c_{nm}(f)|$$

сходится.

Теорема 9. *Пусть* $\alpha = \beta = -1/2$ *и*

$$\int_0^1 \Phi\left[\left(\frac{\pi t}{N}\right)^k\right] t^{2r-2} dt < +\infty,$$

$$N = \min(n, m) = 1, 2, \ldots, \quad r = 0, 1, \ldots, \quad k = 1, 2, \ldots .$$

Тогда для величины

$$\mathscr{E}_N(CW_\Phi^{r,k}(D)) = \sup_{f \in CW_\Phi^{r,k}(D)} \left| \int_{-1}^{1} \int_0^{2\pi} \frac{1}{\sqrt{1-x^2}} f(x, y) dx dy - \frac{2\pi^2}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} f\left(\cos \frac{\pi(2i-1)}{2n}, \frac{2\pi(j-1)}{m-1}\right) \right| \quad (3)$$

справедлива оценка

$$\mathcal{E}_N(CW_\Phi^{r,k}(D)) = O\left(\frac{1}{N^{2r-1}}\int_0^1 \Phi\left[\left(\frac{\pi t}{R}\right)^k\right] t^{2r-2}dt\right).$$

Следствие 1. Если $r = 1, 2, \ldots, k = 1, 2, \ldots$, то

$$\mathcal{E}_N(CW_\Phi^{r,k}(D)) = O\left(\frac{1}{N^{2r-1}}\Phi\left[\left(\frac{\pi}{N}\right)^k\right]\right).$$

Следствие 2. Если $\Phi(t) = t^\alpha$, $\alpha > 0$, то при $\alpha k + 2r > 1$ справедлива оценка

$$\mathcal{E}_N(CW_\Phi^{r,k}(D)) = O\left(\frac{1}{N^{\alpha k + 2r - 1}}\right), \quad r = 0, 1, \ldots, \quad k = 1, 2, \ldots.$$

Доказательства основаны на некоторых вспомогательных предложениях.

3. ВСПОМОГАТЕЛЬНЫЕ ПРЕДЛОЖЕНИЯ

Напомним, что многочлены Якоби $P_n^{(\beta)}(x)$ удовлетворяют дифференциальному уравнению второго порядка (см. [4, с. 114])

$$(1-x^2)P_n^\beta(x)'' - 2(\beta+1)xP_n^{(\beta)}(x)' + n(n+2\beta+1)P_n^{(\beta)}(x) = 0, \tag{4}$$

а система функций $\cos(my)$ и $\sin(my)$ – дифференциальному уравнению

$$\begin{gathered} y'' + m^2y = 0 \\ ([\cos(my)]'' + m^2\cos(my) = 0, [\sin(my)]'' + m^2\sin(my) = 0). \end{gathered} \tag{5}$$

Умножив обе части равенства (4) на $\cos(my)$, а первого из равенств (5) на $P_n^{(\beta)}(x)$ и сложив полученные равенства почленно, имеем

$$\begin{gathered} (1-x^2)P_n^\beta(x)''\cos(my) - 2(\beta+1)xP_n^{(\beta)}(x)'\cos(my) + n(n+2\beta+1)P_n^{(\beta)}(x)\cos(my) + \\ + [\cos(my)]''P_n^{(\beta)}(x) + m^2P_n^{(\beta)}(x)\cos(my) = 0, \end{gathered}$$

или

$$-DP_n^{(\beta)}(x)\cos(my) = (n(n+2\beta+1)+m^2)P_n^{(\beta)}(x)\cos(my).$$

Аналогично получаем

$$-DP_n^{(\beta)}(x)\sin(my) = (n(n+2\beta+1)+m^2)P_n^{(\beta)}(x)\sin(my).$$

Таким образом, система функций

$$P_n^{(\beta)}(x)\cos(my), \quad P_n^{(\beta)}(x)\sin(my)$$

представляет собой собственные функции дифференциального оператора D, отвечающие собственным значениям $n(n + 2\beta + 1) + m^2$.

Это вместе со свойством 5) оператора F_h позволяет доказать следующую лемму, которая лежит в основе доказательств сформулированных выше утверждений (см., например, [3]).

Лемма. *Для любой функции* $f \in L_2^r(D)$ *справедливо равенство*

$$\|\Delta_h^k(D^rf)\|^2 = \sum_{n=0}^\infty\sum_{m=0}^\infty \lambda_{nm}\left(1 - \frac{P_n^{(\beta)}(\cos h)}{P_n^{(\beta)}(1)}\frac{\sin(mh)}{mh}\right)^{2k}[n(n+2\beta+1)+m^2]^{2r}c_{nm}^2(f).$$

4. ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ 9

Ограничимся доказательством последней теоремы. Остальные утверждения доказываются аналогично соответствующим им теоремам из [2], [3].

На основании условий теоремы 9 получаем

$$f(x, y) = \sum_{\nu=0}^{\infty}\sum_{\mu=0}^{\infty}\lambda_{\nu\mu}T_\nu(x)[a_{\nu\mu}(f)\cos(\mu y) + b_{\nu\mu}(f)\sin(\mu y)],$$

где

$$T_0(x) = \frac{1}{\sqrt{\pi}}, \quad T_n(x) = \sqrt{\frac{2}{\pi}}\cos(n\arccos x), \quad n = 1, 2, \ldots,$$

есть ортонормированная система многочленов Чебышёва I рода и

$$a_{\nu\mu}(f) = \frac{1}{\pi}\int_{-1}^{1}\int_{0}^{2\pi}\frac{1}{\sqrt{1-x^2}}f(x, y)T_\nu(x)\cos(\mu y)dxdy,$$

$$b_{\nu\mu}(f) = \frac{1}{\pi}\int_{-1}^{1}\int_{0}^{2\pi}\frac{1}{\sqrt{1-x^2}}f(x, y)T_\nu(x)\sin(\mu y)dxdy.$$

Поэтому для разности в формуле (3) можно получить выражение

$$R_{nm}(f) = \int_{-1}^{1}\int_{0}^{2\pi}\frac{1}{\sqrt{1-x^2}}f(x, y)dxdy - \frac{2\pi^2}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}f\left(\cos\frac{\pi(2i-1)}{2n}, \frac{2\pi(j-1)}{m-1}\right) =$$

$$= \int_{-1}^{1}\int_{0}^{2\pi}\frac{1}{\sqrt{1-x^2}}f(x, y)dxdy - \frac{2\pi^2}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\left[\sum_{\nu=0}^{\infty}\sum_{\mu=0}^{\infty}\lambda_{\nu\mu}a_{\nu\mu}(f)T_\nu\left(\cos\frac{\pi(2i-1)}{2n}\right)\cos\mu\frac{2\pi(j-1)}{m-1} +\right.$$

$$\left. + \sum_{\nu=0}^{\infty}\sum_{\mu=0}^{\infty}\lambda_{\nu\mu}b_{\nu\mu}(f)T_\nu\left(\cos\frac{\pi(2i-1)}{2n}\right)\left(\sin\mu\frac{2\pi(j-1)}{m-1}\right)\right] =$$

$$= \int_{-1}^{1}\int_{0}^{2\pi}\frac{1}{\sqrt{1-x^2}}f(x, y)dxdy - \frac{2\pi^2}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\left[\lambda_{00}a_{00}(f)T_0\left(\cos\frac{\pi(2i-1)}{2n}\right)\cos 0\frac{2\pi(j-1)}{m-1} +\right.$$

$$+ \sum_{\nu=1}^{\infty}\lambda_{\nu0}a_{\nu0}(f)T_\nu\left(\cos\frac{\pi(2i-1)}{2n}\right)\cos 0\frac{2\pi(j-1)}{m-1} + \lambda_{\nu0}b_{\nu0}(f)T_\nu\left(\cos\frac{\pi(2i-1)}{2n}\right)\sin 0\frac{2\pi(j-1)}{m-1} +$$

$$+ \sum_{\mu=1}^{\infty}\lambda_{0\mu}a_{0\mu}(f)T_0\left(\cos\frac{\pi(2i-1)}{2n}\right)\cos\mu\frac{2\pi(j-1)}{m-1} + \lambda_{0\mu}b_{0\mu}(f)T_0\left(\cos\frac{\pi(2i-1)}{2n}\right)\sin\mu\frac{2\pi(j-1)}{m-1} +$$

$$\left. + \sum_{\nu=1}^{\infty}\sum_{\mu=1}^{\infty}\lambda_{\nu\mu}a_{\nu\mu}(f)T_\nu\left(\cos\frac{\pi(2i-1)}{2n}\right)\cos\mu\frac{2\pi(j-1)}{m-1} + \sum_{\nu=1}^{\infty}\sum_{\mu=1}^{\infty}\lambda_{\nu\mu}b_{\nu\mu}(f)T_\nu\left(\cos\frac{\pi(2i-1)}{2n}\right)\sin\mu\frac{2\pi(j-1)}{m-1}\right].$$

Так как $\lambda_{00} = 1/2$, $T_0(x) = 1/\sqrt{\pi}$, то

$$\frac{2\pi^2}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\lambda_{00}a_{00}(f)T_0\left(\cos\frac{\pi(2i-1)}{2n}\right)\cos 0\frac{2\pi(j-1)}{m-1} = \pi\sqrt{\pi}a_{00}(f) =$$

$$= \pi\sqrt{\pi}\frac{1}{\pi}\int_{-1}^{1}\int_{0}^{2\pi}\frac{1}{\sqrt{1-x^2}}f(x,y)T_0(x)dxdy = \int_{-1}^{1}\int_{0}^{2\pi}\frac{1}{\sqrt{1-x^2}}f(x,y)dxdy.$$

Введем следующие обозначения:

$$\Sigma_1 = \frac{2\pi^2}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\left[\sum_{\nu=1}^{\infty}\lambda_{\nu0}a_{\nu0}(f)T_\nu\left(\cos\frac{\pi(2i-1)}{2n}\right)\cos 0\frac{2\pi(j-1)}{m-1} + \right.$$

$$\left. + \lambda_{\nu0}b_{\nu0}(f)T_\nu\left(\cos\frac{\pi(2i-1)}{2n}\right)\sin 0\frac{2\pi(j-1)}{m-1}\right],$$

$$\Sigma_2 = \frac{2\pi^2}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\left[\sum_{\mu=1}^{\infty}\lambda_{0\mu}a_{0\mu}(f)T_0\left(\cos\frac{\pi(2i-1)}{2n}\right)\cos\mu\frac{2\pi(j-1)}{m-1} + \right.$$

$$\left. + \lambda_{0\mu}b_{0\mu}(f)T_0\left(\cos\frac{\pi(2i-1)}{2n}\right)\sin\mu\frac{2\pi(j-1)}{m-1}\right],$$

$$\Sigma_3 = \frac{2\pi^2}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\left[\sum_{\nu=1}^{\infty}\sum_{\mu=1}^{\infty}\lambda_{\nu\mu}a_{\nu\mu}(f)T_\nu\left(\cos\frac{\pi(2i-1)}{2n}\right)\cos\mu\frac{2\pi(j-1)}{m-1} + \right.$$

$$\left. + \sum_{\nu=1}^{\infty}\sum_{\mu=1}^{\infty}\lambda_{\nu\mu}b_{\nu\mu}(f)T_\nu\left(\cos\frac{\pi(2i-1)}{2n}\right)\sin\mu\frac{2\pi(j-1)}{m-1}\right].$$

Рассмотрим теперь каждое из Σ_k, $k = 1, 2, 3$.

Так как

$$\lambda_{\nu0} = 1/2, \quad T_\nu\left(\cos\frac{\pi(2i-1)}{2n}\right) = \sqrt{\frac{2}{\pi}}\cos\nu\frac{\pi(2i-1)}{2n}$$

и

$$\sum_{i=1}^{n}\cos\nu\frac{\pi(2i-1)}{2n} = \begin{cases}(-1)^s n, & \nu = 2sn,\\ 0, & \nu \ne 2sn, \quad s = 0, 1, \ldots,\end{cases}$$

то нетрудно показать, что

$$\Sigma_1 = \frac{\pi^2}{n}\sqrt{\frac{2}{\pi}}\sum_{\nu=1}^{\infty}(-1)^\nu n a_{2\nu n,0}(f) = \pi^2\sqrt{\frac{2}{\pi}}\sum_{\nu=1}^{\infty}(-1)^\nu a_{2\nu n,0}(f).$$

Замечая, что $\lambda_{0\mu} = 1$, $T_0(x) = 1/\sqrt{\pi}$, имеем

$$\Sigma_2 = \frac{2\pi^2}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\left[\sum_{\mu=1}^{\infty}a_{0\mu}(f)\frac{1}{\sqrt{\pi}}\cos\mu\frac{2\pi(j-1)}{m-1} + b_{0\mu}(f)\frac{1}{\sqrt{\pi}}\sin\mu\frac{2\pi(j-1)}{m-1}\right] =$$

$$= \frac{2\pi^2}{\sqrt{\pi}nm}\sum_{j=1}^{m}\left[\sum_{\mu=1}^{\infty}a_{0\mu}(f)\cos\mu\frac{2\pi(j-1)}{m-1} + b_{0\mu}(f)\sin\mu\frac{2\pi(j-1)}{m-1}\right] =$$

$$= \frac{2\pi^2}{\sqrt{\pi}nm}\sum_{\mu=1}^{\infty}a_{0\mu}(f)\left[\sum_{j=1}^{m}\cos\mu\frac{2\pi(j-1)}{m-1}\right] + \frac{2\pi^2}{\sqrt{\pi}nm}\sum_{\mu=1}^{\infty}a_{0\mu}(f)\left[\sum_{j=1}^{m}\sin\mu\frac{2\pi(j-1)}{m-1}\right].$$

Так как

$$\sum_{j=1}^{m}\sin\mu\frac{2\pi(j-1)}{m-1} = 0, \quad \sum_{j=1}^{m}\cos\mu\frac{2\pi(j-1)}{m-1} = \begin{cases}(-1)^{s(m-1)}m, & \mu = ms,\\ 0, & \mu \ne ms,\end{cases}$$

$s = 0, 1, \ldots$, то имеем

$$\Sigma_2 = 2\pi\sqrt{\pi}\sum_{\mu=1}^{\infty}(-1)^{\mu(m-1)}a_{0,\mu m}(f).$$

Поскольку

$$\lambda_{\nu\mu} = 1, \quad T_\nu\left(\cos\frac{\pi(2i-1)}{2n}\right) = \sqrt{\frac{2}{\pi}}\cos\nu\frac{\pi(2i-1)}{2n},$$

то получаем

$$\Sigma_3 = \frac{2\pi^2}{nm}\sqrt{\frac{2}{\pi}}\sum_{\nu=1}^{\infty}\sum_{\mu=1}^{\infty}a_{\nu\mu}(f)\left[\left(\sum_{i=1}^{n}\cos\nu\frac{\pi(2i-1)}{2n}\right)\left(\sum_{j=1}^{m}\cos\mu\frac{2\pi(j-1)}{m-1}\right)\right] +$$

$$+ \frac{2\pi^2}{nm}\sqrt{\frac{2}{\pi}}\sum_{\nu=1}^{\infty}\sum_{\mu=1}^{\infty}b_{\nu\mu}(f)\left[\left(\sum_{i=1}^{\infty}\cos\nu\frac{\pi(2i-1)}{2n}\right)\left(\sum_{j=1}^{m}\sin\mu\frac{2\pi(j-1)}{m-1}\right)\right].$$

Отсюда нетрудно увидеть, что

$$\Sigma_3 = 2\pi^2\sum_{\nu=1}^{\infty}\sum_{\mu=1}^{\infty}(-1)^{\nu+\mu}a_{2\nu n, 2\mu m}(f).$$

Поэтому окончательно получаем

$$R_{nm}(f) = \pi^2\sum_{\nu=1}^{\infty}(-1)^{\nu}a_{2\nu n, 0}(f) - 2\pi\sqrt{\pi}\sum_{\mu=1}^{\infty}(-1)^{\mu(m+1)}a_{0,\mu m}(f) - 2\pi^2\sum_{\nu=1}^{\infty}\sum_{\mu=1}^{\infty}(-1)^{\nu+\mu}a_{2\nu n, \mu m}(f).$$

Отсюда вытекает оценка

$$R_{nm}(f) = O\left(\sum_{\nu=1}^{\infty}|a_{2\nu n, 0}(f)| + \sum_{\mu=1}^{\infty}|a_{0,\mu m}(f)| + \sum_{\nu=1}^{\infty}\sum_{\mu=1}^{\infty}|a_{2\nu n, \mu m}(f)|\right) = O\left(\sum_{i^2+j^2\ge N^2}|a_{ij}(f)|\right).$$

Так как $CW_{\Phi}^{r,k}(D) \subset W_{2,\Phi}^{r,k}(D)$, то из теоремы 1 следует, что

$$\sum_{i^2+j^2\le N^2}a_{ij}^2(f) = O\left(N^{-4r}\Phi^2\left[\left(\frac{\pi}{N}\right)^k\right]\right).$$

Далее, поступая как в [6], получаем требуемое утверждение (см. также [8], [9]).

СПИСОК ЛИТЕРАТУРЫ

1. *Соболев С.Л.* Уравнения математической физики. М.: Наука, 1966.
2. *Абилов В.А., Керимов М.К.* Некоторые оценки погрешности смешанных рядов Фурье–Бесселя функций двух переменных // Ж. вычисл. матем. и матем. физ. 2006. Т. 46. № 9. С. 1545–1565.
3. *Абилов В.А., Керимов М.К.* Некоторые вопросы разложения функций в двойные ряды Фурье–Эрмита–Якоби // Ж. вычисл. матем. и матем. физ. 2004. Т. 44. № 9. С. 1596–1607.
4. *Сегё Г.* Ортогональные многочлены. М.: Физматгиз, 1962.
5. *Колмогоров А.Н.* Избранные труды. Математика и механика. М.: Наука, 1987.
6. *Абилов В.А., Керимов М.К.* Об оценках остаточных членов кратных рядов Фурье–Чебышёва и кубатурных формул чебышёвского типа // Ж. вычисл. матем. и матем. физ. 2003. Т. 43. № 5. С. 643–663.
7. *Никольский С.М.* Приближение функций многих переменных и теоремы вложения. М.: Наука, 1969.
8. *Abilov V.A.* On the convergence of multiple Fourier series and quadrature formulae // Math. Balkanica, New Series. 2002. V. 16. Fasc. 1–4. P. 73–94.
9. *Абилов В.А., Абилова Ф.В.* Об одной квадратурной формуле // Ж. вычисл. матем. и матем. физ. 2002. Т. 42. № 4. С. 451–458.

ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 2007, том 47, № 8, с. 1308–1322

УДК 519.626

ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ В ОДНОЙ ЗАДАЧЕ МАКРОЭКОНОМИКИ

© 2007 г. В. К. Булгаков, Г. Л. Шатов
(680056 Хабаровск, ул. Серышева, 47, ДВГУПС)
e-mail: blvic@rambler.ru
Поступила в редакцию 02.02.2007 г.

На основе принципа максимума Понтрягина разработан оригинальный алгоритм решения задачи оптимального управления одной задачей макроэкономики. Приводятся результаты расчетов на ЭВМ оптимального управления и оптимальной траектории развития региональной экономической системы. Для некоторой области изменения оптимального управления приводится инвариант макроэкономической системы. Библ. 7. Фиг. 9. Табл. 2.

Ключевые слова: задачи макроэкономики, метод оптимального управления, численный метод решения.

1. ПОСТАНОВКА ЗАДАЧИ

Пусть $\mu K(t)$ – износ основного капитала за год, $g(t)N(t)$ – годовой доход живого капитала, $x(t) = \mathrm{B}\dfrac{\mu K(t)}{g(t)N(t)}$ – фазовая безразмерная макроэкономическая переменная, $x(t) \in C^1[0, T]$, B – постоянная производственной B(x)-функции (см. [1]):

$$\mathrm{B}(x) = b(1-e^{-x}) + (1-b)x(1-e^{-1/x}), \tag{1}$$

где b – вторая постоянная производственной функции, $0 < b < 1$.

Траектория экономической системы описывается уравнением (см. [2])

$$dx/dt = a\mathrm{B}(x) - \lambda x - pw, \tag{2}$$

где $a, \lambda, p > 0$ – параметры макроэкономической модели; B(x) определяется по формуле (1), $w(t)$ – управление, $w \in \overline{W}$:

$$\overline{W} = \{w(t) \in C[0, T] : w(t) \in [w_1, w_2]\}, \tag{3}$$

$0 < w_1 < w_2$ – постоянные, T – горизонт планирования, $T < \infty$.

Рассмотрим задачу об определении оптимального управления $w^*(t) \in \overline{W}$ экономическим процессом (2) в следующем смысле: найти управление $w^*(t) \in \overline{W}$, которое переводит процесс (2) из одного фиксированного состояния $x(0) = x_1$ в другое фиксированное состояние $x(T) = x_2$ при условии, что интеграл благосостояния

$$J(w) = \int_0^T w^{\alpha}(t)dt \tag{4}$$

принимает наибольшее значение. Здесь $\alpha \in (0, 1)$ – эмпирическая постоянная. Математическая постановка задачи имеет вид

$$\max_{w \in \overline{W}} \int_0^T w^{\alpha}(t)dt,$$
$$dx/dt = a\mathrm{B}(x) - \lambda x - pw, \quad x(0) = x_1, \quad x(T) = x_2, \tag{5}$$
$$\mathrm{B}(x) = b(1-e^{-x}) + (1-b)x(1-e^{-1/x}).$$

Отметим, что в задаче (5) момент времени T заранее не задан.

Введем функцию Гамильтона исследуемой задачи

$$H(x, \psi, w) = w^{\alpha} + \psi[a\mathrm{B}(x) - \lambda x - pw], \tag{6}$$

гамильтонову систему уравнений

$$\begin{aligned} dx/dt &= a\mathrm{B}(x) - \lambda x - pw \\ d\psi/dt &= -[a\mathrm{B}'(x) - \lambda]\psi, \end{aligned} \tag{7}$$

где $\psi(t)$ – сопряженная к $x(t)$ переменная, $\psi(t) \in C^1[0, T]$. Обозначим через $R(x)$, $R(\psi)$ области возможных значений переменных $x(t)$, $\psi(t)$ системы (7). Пусть $R(x)$, $R(\psi) = R_1^+ = (0, \infty)$.

Решение задачи (5) получим на основе принципа максимума Понтрягина (см. [3]): если $w^*(t)$ – оптимальное управление задачи (5), а $x^*(t)$, $\psi(t)$ – соответствующие ему траектории системы (7), то функция Гамильтона (6) должна удовлетворять равенству

$$H(x^*(t), \psi(t), w^*(t)) = \sup_{w \in \overline{W}} H(x^*(t), \psi(t), w). \tag{*}$$

2. РЕШЕНИЕ ЗАДАЧИ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ

Введем постоянные $\psi_1 = p^{-1}\alpha w_2^{\alpha-1}$, $\psi_2 = p^{-1}\alpha w_1^{\alpha-1}$, $\pi = (\alpha/p)^{1/(1-\alpha)}$. Справедлива

Теорема 1. *Пусть* $w^*(t) \in \overline{W}$ *– оптимальное решение задачи* (5), *а* $x^*(t)$, $\psi(t)$ *– соответствующие ему решения гамильтоновой системы* (7). *Тогда между оптимальным управлением* $w^*(t)$, *соответствующими ему оптимальными траекториями фазовой и сопряженной переменных* $x^*(t)$, $\psi(t)$ *имеет место зависимость*

$$w^*(t) = \begin{cases} w_2 \ \text{при} \ \psi(t) \le \psi_1, \\ \pi\psi^{-1/(1-\alpha)}(t) \ \text{при} \ \psi_1 < \psi(t) < \psi_2, \\ w_1 \ \text{при} \ \psi(t) \ge \psi_2. \end{cases} \tag{8}$$

Доказательство. Рассмотрим вначале открытое множество $W = \{w(t) \in C[0, T] : w(t) \in (w_1, w_2)\}$ – внутренность множества $\overline{W}$. Известно (см. [3]), что когда область управления является открытым множеством, то рассматриваемая оптимальная задача эквивалентна задаче Лагранжа классического вариационного исчисления. Точка максимума $w^* \in W$ является стационарной точкой функции Гамильтона $\partial H/\partial w = 0$, откуда получаем

$$\psi = p^{-1}\alpha w^{*\alpha-1}. \tag{9}$$

В соотношении (9) имеем управление $w^*(t) \in W$, $\psi(t) \in \Psi$, где $\Psi = \{\psi(t) \in C^1[0, T] : \psi(t) \in (\psi_1, \psi_2)\}$. Заметим, что так как $\alpha < 1$, то $\left.\dfrac{\partial^2 H}{\partial w^2}\right|_{w^*} < 0$. Таким образом, при любых фиксированных $x \in R(x)$, $\psi \in \Psi$ функция $w(t) = \pi\psi^{-1/(1-\alpha)}(t)$ доставляет максимум функции Гамильтона (6) на множестве функций $w(t) \in W$.

Рассмотрим теперь множество $\Psi_1 = \{\psi(t) \in C^1[0, T] : \psi(t) \in (0, \psi_1]\}$. Пусть $\psi(t)$ – произвольная точка множества Ψ_1, например $\psi(t) = \psi_1 - a_1(t)$, где $a_1(t) \in C^1[0, T]$, $0 \le a_1(t) < \psi_1$. Функцию Гамильтона в рассматриваемой точке $\psi \in \Psi_1$ можно записать в форме

$$\begin{aligned} H(x, \psi, w) &= w^{\alpha} - p\psi w + \psi[a\mathrm{B}(x) - \lambda x] = w^{\alpha} - p\psi_1 w + pa_1 w + \psi[a\mathrm{B}(x) - \lambda x] = \\ &= \omega_1(w) + \varphi_1(w) + \psi[a\mathrm{B}(x) - \lambda x], \end{aligned}$$

где

$$\omega_1(w) = w^{\alpha}\left[1 - \alpha\left(\frac{w}{w_2}\right)^{1-\alpha}\right], \quad \varphi_1(w) = pa_1 w.$$

Анализ функции $\omega_1(w)$ (производных ω_{1w}, ω_{1ww}) показывает, что $\omega_1(w)$ имеет максимум при $w = w_2$. Функция $\varphi_1(w)$ имеет максимум в точке $w = w_2$ при $a_1(t) > 0$ или равна нулю, если $a_1(t) = 0$.

Таким образом, при любых фиксированных $x \in R(x)$, $\psi \in \Psi_1$ функция $w(t) = w_2$ доставляет максимум функции Гамильтона $H(x, \psi, w)$ на множестве $\overline{W}$.

Аналогично, рассмотрев $H(x, \psi, w)$ в произвольной точке $\psi(t) \in \Psi_2$, где $\Psi_2 = \{\psi(t) \in C^1[0, T] : \psi(t) \in [\psi_2, \infty)\}$, например, при $\psi(t) = \psi_2 + a_2(t)$, где $a_2(t) \in C^1[0, T]$, $0 \le a_2(t) < \infty$, получим

$$H(x, \psi, w) = w^{\alpha} - p\psi_2 w - pa_2 w + \psi[a\mathrm{B}(x) - \lambda x] = \omega_2(w) + \varphi_2(w) + \psi[a\mathrm{B}(x) - \lambda x],$$

где

$$\omega_2(w) = w^{\alpha}\left[1 - \alpha\left(\frac{w}{w_1}\right)^{1-\alpha}\right], \quad \varphi_2(w) = -pa_2 w,$$

и, проведя анализ функций $\omega_2(w)$, $\varphi_2(w)$, убедимся, что при любых фиксированных $x \in R(x)$, $\psi \in \Psi_2$ функция $w(t) = w_1$ доставляет максимум функции Гамильтона $H(x, \psi, w)$ на множестве $w(t) \in \overline{W}$. Поскольку $\Psi \cup \Psi_1 \cup \Psi_2 = \{\psi(t) \in C^1[0, T] : R(\psi) = R_1^+\}$, то рассмотрены области $R(x)$, $R(\psi)$ всех возможных значений функций $\psi(t)$, $x(t)$ и найдено, что максимум функции Гамильтона (6) имеет место при любых фиксированных $x \in R(x)$, $\psi \in R(\psi)$ на управлениях

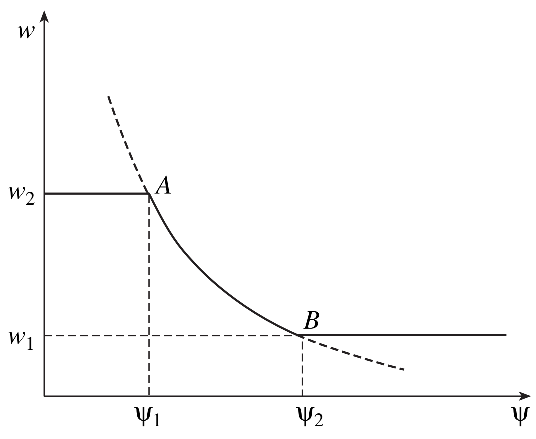
$$w(t) = \begin{cases} w_2, \text{ если } \psi(t) \in \Psi_1, \\ \pi\psi^{-\frac{1}{1-\alpha}}(t), \text{ если } \psi(t) \in \Psi, \\ w_1, \text{ если } \psi(t) \in \Psi_2. \end{cases} \tag{10}$$

Пусть $w^*(t)$ – оптимальное управление задачи (5), а $x^*(t)$, $\psi(t)$ – соответствующее ему решение системы (7). Тогда в соответствии с основным равенством принципа максимума Понтрягина (*) соотношение (10) можно записать в виде (8). Теорема доказана.

Замечание 1. Непосредственной проверкой можно убедиться, что функция $w(\psi) = \pi\psi^{-1/(1-\alpha)}$ имеет пределы $\lim\limits_{\psi \to \psi_1+} w(\psi) = w_2$, $\lim\limits_{\psi \to \psi_2-} w(\psi) = w_1$, т.е. оптимальное управление как функция переменной ψ есть непрерывная функция в точках ψ_1, ψ_2 и, следовательно, на всем множестве $R(\psi) = R_1^+$.

На фиг. 1 в плоскости (ψ, w) показан график оптимального управления w^*, определяемого зависимостью (8); он состоит из прямой w_2, кривой AB ($w^* = \pi\psi^{-1/(1-\alpha)}$) и прямой w_1.

Пусть $w(t)$ – оптимальное управление динамикой макроэкономики, определяемое зависимостями (8) теоремы 1, а $x(t)$, $\psi(t)$ – соответствующие ему решения системы (7) (звездочки у w и x для простоты опустим) с граничными условиями $x(0) = x_1$, $x(T) = x_2$.



Фиг. 1.

Введем следующие постоянные модели:

$$\gamma = \frac{\lambda}{a}, \quad \gamma_1 = \frac{pw_1}{a}, \quad \gamma_2 = \frac{pw_2}{a}, \quad \sigma = \frac{1}{a}\left(\frac{\alpha}{p^\alpha}\right)^{1/(1-\alpha)}, \quad c_0 = \sigma^{1-\alpha}, \quad \eta = \frac{1}{a}.$$

Тогда краевую задачу, определяющую оптимальные траектории $x(t)$, $\psi(t)$, можно записать в форме

$$\begin{gathered} dx/dt = F(x, \psi), \\ d\psi/dt = G(x, \psi), \quad 0 < t < T, \\ x(0) = x_1, \quad x(T) = x_2, \end{gathered} \tag{11}$$

где

$$F(x, \psi) = \begin{cases} a(\mathrm{B}(x) - \gamma x - \gamma_2), & \psi(t) \le \psi_1, \\ a(\mathrm{B}(x) - \gamma x - \sigma\psi^{-1/(1-\alpha)}), & \psi_1 < \psi(t) < \psi_2, \\ a(\mathrm{B}(x) - \gamma x - \gamma_1), & \psi(t) \ge \psi_2, \end{cases}$$

$$G(x, \psi) = a(\gamma - \mathrm{B}'(x))\psi.$$

Запишем также следующую систему уравнений, эквивалентную системе (11), в которой в качестве независимой переменной взята сопряженная переменная ψ, а функциями являются фазовая переменная x и время t:

$$\frac{dx}{d\psi} = \varphi(\psi, x), \quad \frac{dt}{d\psi} = \chi(\psi, x), \tag{12}$$

где

$$\varphi(\psi, x) = \begin{cases} \dfrac{\mathrm{B}(x) - \gamma x - \gamma_2}{\gamma - \mathrm{B}'(x)}\dfrac{1}{\psi}, & \psi \le \psi_1, \\ \dfrac{\mathrm{B}(x) - \gamma x - \sigma\psi^{-1/(1-\alpha)}}{\gamma - \mathrm{B}'(x)}\dfrac{1}{\psi}, & \psi_1 < \psi < \psi_2, \\ \dfrac{\mathrm{B}(x) - \gamma x - \gamma_1}{\gamma - \mathrm{B}'(x)}\dfrac{1}{\psi}, & \psi \ge \psi_2, \end{cases}$$

$$\chi(\psi, x) = \frac{\eta}{\gamma - \mathrm{B}'(x)}\frac{1}{\psi}.$$

Проведем предварительный анализ интегральных кривых (оптимальных траекторий $x(\psi)$) системы (12) и установим их основные свойства.

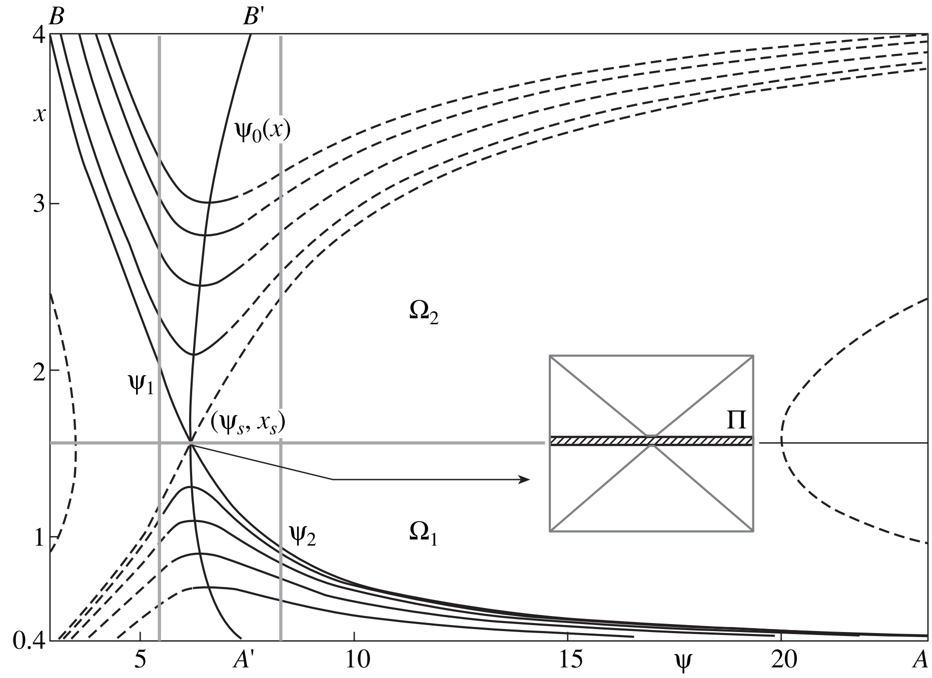
Рассмотрим на положительном ортанте $R_+(\psi, x)$ плоскости сопряженной и фазовой переменной замкнутую область $\Omega(\psi, x) = [\psi_{\min} \le \psi \le \psi_{\max}, x_{\min} \le x \le x_{\max}]$, где отрезок $[x_{\min}, x_{\max}]$ содержит в себе все возможные реальные начальные и конечные состояния x_1, x_2 экономических систем, а $\psi_{\min} = \psi(x_{\max})$, $\psi_{\max} = \psi(x_{\min})$ – соответствующие значения сопряженных переменных. Область Ω (см. фиг. 2) назовем областью реальных состояний и реальных процессов экономических систем при конечном горизонте планирования ($T < \infty$).

Заметим, что отрезок $[\psi_{\min}, \psi_{\max}]$ содержит все граничные значения ψ, необходимые для решения краевой задачи оптимального управления.

Рассмотрим в области Ω вертикальные прямые ψ_1, ψ_2. Прямые ψ_1, ψ_2 делят область Ω на три подобласти. Эти подобласти будем обозначать через Ψ_1, Ψ, Ψ_2.

В подобласти Ψ есть особая точка (ψ_s, x_s), координаты которой определяются уравнениями

$$\gamma - \mathrm{B}'(x_s) = 0, \quad \psi_s = \frac{c_0}{[\mathrm{B}(x_s) - \gamma x_s]^{1-\alpha}}.$$



Фиг. 2.

Введем в подобласти Ψ кривую $\psi_0(x)$, на которой производная $dx/d\psi = 0$ (точка (ψ_s, x_s) пока не рассматривается). Кривая $\psi_0(x)$ определяется уравнением

$$\psi_0(x) = \frac{c_0}{[\mathrm{B}(x) - \gamma x]^{1-\alpha}}. \tag{13}$$

Из первого уравнения системы (11) видно, что $\left.\frac{dx}{dt}\right|_{\psi_0} = 0$ на кривой $\psi_0(x)$, поэтому $\psi_0(x)$ – это линия стационарных состояний системы.

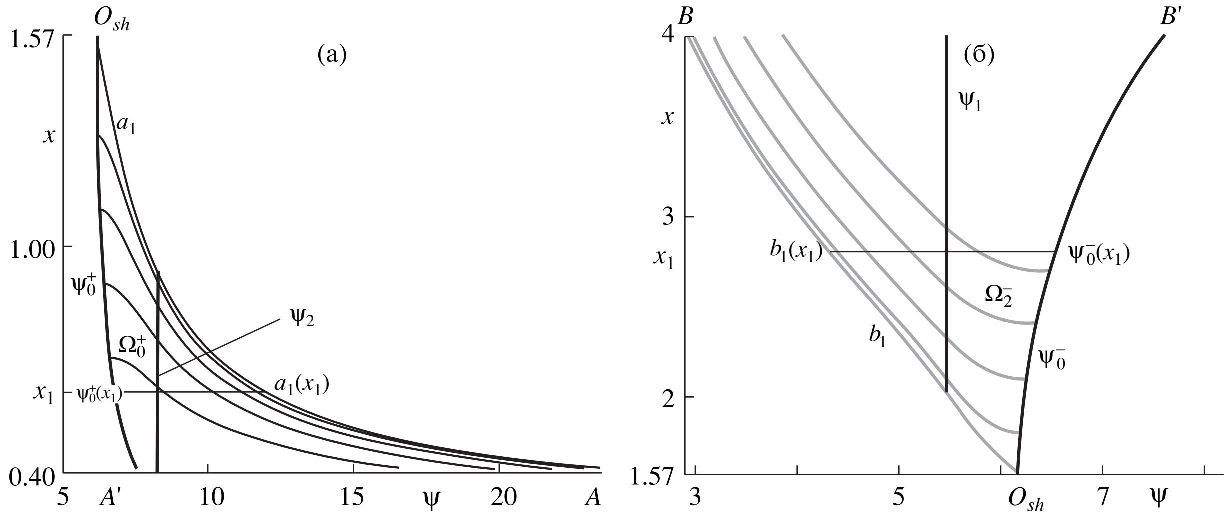
Введем окружность с центром в точке (ψ_s, x_s) с малым радиусом h, обозначив ее через O_{sh} (в наших расчетах $h = 10^{-5}$). Точка A имеет такую координату $\psi_{\max}$, что интегральная кривая первого уравнения системы (12), исходящая из точки A (кривая a_1), касается снизу окружности O_{sh} а точка B имеет такую координату $\psi_{\min}$, что интегральная кривая, исходящая из точки B (кривая b_1), касается O_{sh} сверху.

На фиг. 2 показаны интегральные кривые $\{a_i\}$, исходящие из промежутка $A'A = \{\psi_0(x_{\min}) < \psi \le \le \psi_{\max}, x = x_{\min}\}$, и кривые $\{b_i\}$, исходящие из промежутка $BB' = \{\psi_{\min} \le \psi < \psi_0(x_{\max}), x = x_{\max}\}$, посчитанные на ЭВМ для рассматриваемого примера решением системы (11), эквивалентной системе (12) для начальных данных x, ψ из $A'A$, BB', методом Рунге–Кутты четвертого порядка точности.

Рассмотрим области $\Omega_1, \Omega_2 \subset \Omega$, разделенные горизонтальной полосой П шириной $2h$; границы полосы касаются круга O_{sh} (фиг. 2).

Утверждение 1. *Интегральные кривые системы* (12), *исходящие из промежутков* $A'A$, BB', *образуют два семейства* $\{a_{i_1}\}$, $\{b_{i_2}\}$ ($\{a_{i_1}\}$ *расположены в* Ω_1, $\{b_{i_2}\}$ – *в* Ω_2). *Ни одна кривая семейств* $\{a_{i_1}\}$, $\{b_{i_2}\}$ *не пересекает полосу* П. *Здесь* $i_1 = \overline{1, n_1}$, $i_2 = \overline{1, n_2}$, n_1, n_2 – *любые целые числа.*

Справедливость утверждения 1 вытекает из фиг. 2. Следует только заметить, что для любой экономической системы на линии $\psi_0(x)$ каждая интегральная кривая $x = x(\psi)$ имеет условие ста-



Фиг. 3.

ционарности $\left.\dfrac{dx}{d\psi}\right|_{\psi_0} = 0$, вторая производная

$$\left.\frac{d^2x}{d\psi^2}\right|_{\psi_0} \begin{matrix} < 0 \text{ при } \gamma - \mathrm{B}'(x) < 0, \\ > 0 \text{ при } \gamma - \mathrm{B}'(x) > 0, \end{matrix}$$

т.е. для семейства $\{a_{i_1}\}$ точка кривой ψ_0 есть точка максимума, а для семейства $\{b_{i_2}\}$ – точка минимума, следовательно, интегральные кривые максимально приближаются к полосе П при пересечении линии ψ_0.

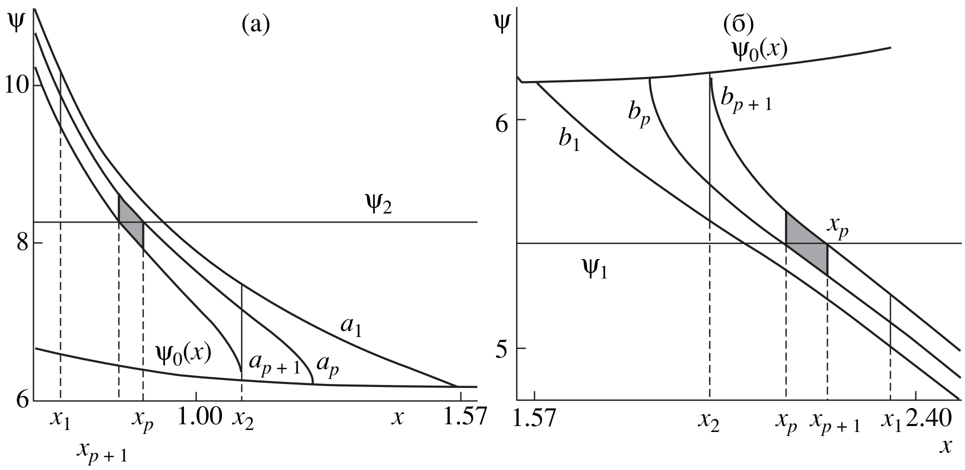
Обозначим через Ω_1^+ часть области Ω_1, ограниченную интегральной кривой a_1 (проходящей через точку A), кривой $\psi_0(x)$ и отрезком $A'A$, а через Ω_2^- – часть области Ω_2, ограниченную интегральной кривой b_1 (проходящей через точку B), кривой $\psi_0(x)$ и отрезком BB' (фиг. 3). Обозначим часть кривой $\psi_0(x)$, лежащей в области Ω_1^+, через $\psi_0^+(x)$, а в области Ω_2^- – через $\psi_0^-(x)$. Обозначим интегральные кривые $\{a_{i_1}\}$, лежащие в Ω_1^+, через $\{a_{i_1}^+\}$, а интегральные кривые $\{b_{i_2}\}$, лежащие в Ω_2^-, через $\{b_{i_2}^-\}$.

Утверждение 2. *Интегральные кривые $\{a_{i_1}^+\} \subset \Omega_1^+$ не пересекаются между собой, интегральные кривые $\{b_{i_2}^-\} \subset \Omega_2^-$ также не пересекаются между собой. Каждой точке ξ_{i_1} промежутка $A'A$ кривые $\{a_{i_1}^+\}$ ставят в соответствие только одну точку кривой $\psi_0^+(x)$. Каждой точке ξ_{i_2} промежутка BB' кривые $\{b_{i_2}^-\}$ ставят в соответствие только одну точку кривой $\psi_0^-(x)$.*

В справедливости первой части утверждения 2 можно убедиться, рассмотрев задачу Коши:

$$\frac{dx}{d\psi} = \begin{cases} \dfrac{\mathrm{B}(x) - \gamma x - \gamma_1}{\gamma - \mathrm{B}'(x)} \dfrac{1}{\psi}, & \psi \ge \psi_2, \\ \dfrac{\mathrm{B}(x) - \gamma x - \sigma\psi^{-1/(1-\alpha)}}{\gamma - \mathrm{B}'(x)} \dfrac{1}{\psi}, & \psi_0^+(x) \le \psi < \psi_2, \quad (\psi, x) \in \Omega_1^+, \end{cases} \tag{14}$$

$$x(\xi_{i_1}) = x_{\min}, \quad \xi_{i_1} \in A'A,$$



Фиг. 4.

которая определяет интегральные кривые a_{i_1}. Существование глобального решения задачи Коши (14) в области Ω_1^+ следует из леммы Вишика–Дубинского (см. [4, с. 67–68]), а из второй части теоремы 2 из [5, с. 168–169] следует единственность глобального решения задачи Коши (14) в области Ω_1^+.

Аналогично, рассмотрев задачу Коши

$$\frac{dx}{d\psi} = \begin{cases} \dfrac{\mathrm{B}(x) - \gamma x - \gamma_2}{\gamma - \mathrm{B}'(x)} \dfrac{1}{\psi}, & \psi \le \psi_1, \\ \dfrac{\mathrm{B}(x) - \gamma x - \sigma\psi^{-1/(1-\alpha)}}{\gamma - \mathrm{B}'(x)} \dfrac{1}{\psi}, & \psi_0^-(x) \ge \psi > \psi_1, \quad (\psi, x) \in \Omega_2^-, \end{cases}$$

$$x(\xi_{i_2}) = x_{\max}, \quad \xi_{i_2} \in BB',$$

которая определяет интегральные кривые b_{i_2}, убеждаемся в справедливости второй части утверждения 2.

Пусть x_1, x_2 – точки начального и конечного состояния экономической системы. Возможны два случая.

Случай А (см. фиг. 4а): $x_{\min} < x_1 < x_2 < x_s$. В этом случае, согласно утверждениям 1, 2, через сечение x_1, x_2 плоскости (x, ψ) проходят интегральные кривые (экстремали) $\{a_{i_1}\} \in \Omega_1^+$, $i_1 = \overline{1, m_1}$ (m_1 – любое целое число). Экстремаль a_1 касается круга O_{sh}, а экстремаль a_{m_1} в сечении x_2 имеет общую точку с кривой $\psi_0(x)$. Очевидно, что все экстремали семейства $\{a_{i_1}\}$, проходящие через сечения x_1 и x_2, находятся между экстремалями a_1, a_{m_1}.

Случай Б (см. фиг. 4б): $x_s < x_2 < x_1 < x_{\max}$. Здесь через сечения x_1, x_2 плоскости (x, ψ) проходят интегральные кривые (экстремали) $\{b_i\} \in \Omega_2^-$, $i_2 = \overline{1, m_2}$ (m_2 – любое целое число). Экстремаль b_1 касается круга O_{sh}, а b_{m_2} в сечении x_2 имеет общую точку с кривой $\psi_0(x)$. Все экстремали, проходящие через сечения x_1 и x_2, находятся между экстремалями b_1, b_{m_2}.

Пусть $J_1^1, \ldots, J_{m_1}^1$ – значения интеграла благосостояния на экстремалях $a_1, \ldots, a_{m_1}$, а $T_1^1, \ldots, T_{m_1}^1$ – времена перехода из состояния x_1 в состояние x_2 по экстремалям $a_1, \ldots, a_{m_1}$. Обозначим через

$J_1^2, \ldots, J_{m_2}^2$ и $T_1^2, \ldots, T_{m_2}^2$, соответственно, интегралы благосостояния и времена перехода на экстремалях $b_1, \ldots, b_{m_2}$.

Теорема 2. *Рассмотрим экономический процесс, описываемый системой уравнений* (12).

Рассмотрим $x_1, x_2 \in \Omega_1^+$ *– начальное и конечное состояния экономической системы,* $x_{\min} < x_1 < < x_2 < x_s$ (*случай* А).

Рассмотрим также $x_1, x_2 \in \Omega_2^-$ *– начальное и конечное состояния системы,* $x_s < x_2 < x_1 < x_{\max}$ (*случай* Б). *Тогда время перехода из состояния* x_1 *в состояние* x_2, $x_1, x_2 \in \Omega_1^+$, *по экстремалям* $a_1, \ldots, a_{m_1}$ *и соответствующие значения интеграла благосостояния, а также время перехода из состояния* x_1 *в состояние* x_2, $x_1, x_2 \in \Omega_2^-$, *по экстремалям* $b_1, \ldots, b_{m_2}$ *и соответствующие значения интеграла благосостояния удовлетворяют неравенствам*

$$T_{m_\delta}^\delta > \ldots > T_2^\delta > T_1^\delta, \tag{15}$$

$$J_{m_\delta}^\delta > \ldots > J_2^\delta > J_1^\delta, \tag{16}$$

где $\delta = 1$ (*для случая* А) *и* $\delta = 2$ (*для случая* Б).

Доказательство. Рассмотрим случай А. Пусть сечение x_1 находится в области $\Omega_1^+ \cup \Psi_2$ (выше прямой ψ_2, см. фиг. 4а). Рассмотрим две произвольные соседние интегральные кривые (экстремали) a_i, $i = p, p+1$. Вдоль рассматриваемых экстремалей справедливы уравнения

$$\left(\frac{dt}{dx}\right)_{a_i} = \frac{\eta}{B(x) - \gamma x - \gamma_1} \quad \text{при} \quad \psi \ge \psi_2, \tag{17}$$

$$\left(\frac{dt}{dx}\right)_{a_i} = \frac{\eta}{B(x) - \gamma x - \zeta_{a_i}(x)} \quad \text{при} \quad \psi < \psi_2, \tag{18}$$

где $\zeta_{a_i}(x) = \dfrac{\sigma}{\psi_{a_i}^{1/(1-\alpha)}(x)}$.

Из уравнения (17) следует, что на отрезке $[x_1, x_{p+1}]$ имеем

$$\left(\frac{dt}{dx}\right)_{a_p} = \left(\frac{dt}{dx}\right)_{a_{p+1}}. \tag{19}$$

Рассмотрим переход из состояния x_{p+1} в состояние x_p. В этом случае изменение времени перехода по экстремали a_p описывается уравнением (17), а по экстремали a_{p+1} – уравнением (18).

Так как

$$\lim_{\psi \to \psi_2^+} B(x) - \gamma x - \sigma\psi^{-1/(1-\alpha)} = B(x) - \gamma x - \gamma_1,$$

то правые части уравнений (17), (18) в точке x_{p+1} равны между собой, во всех остальных точках отрезка $[x_{p+1}, x_p]$ имеет место неравенство

$$\frac{\eta}{B(x) - \gamma x - \gamma_1} < \frac{\eta}{B(x) - \gamma x - \zeta_{a_{p+1}}(x)}, \tag{20}$$

поэтому имеем

$$\left(\frac{dt}{dx}\right)_{a_p} < \left(\frac{dt}{dx}\right)_{a_{p+1}}, \quad x \in (x_{p+1}, x_p]. \tag{21}$$

Рассмотрим переход из состояния x_p в состояние x_2. На отрезке $[x_p, x_2]$ изменение времени перехода по экстремалям a_p, a_{p+1} описывается уравнением (18). Для любого $x \in [x_p, x_{p+1}]$ имеет место неравенство (см. фиг. 4а) $\psi_{a_p}(x) > \psi_{a_{p+1}}(x)$, следовательно, $\zeta_{a_p}(x) < \zeta_{a_{p+1}}(x)$. В области Ω_a, ограниченной линиями $\psi_0(x)$, ψ_2, имеем также $\mathrm{B}(x) - \gamma x > \zeta_{a_p}(x)$, $\mathrm{B}(x) - \gamma x > \zeta_{a_{p+1}}(x)$. Следовательно, справедливо неравенство

$$\mathrm{B}(x) - \gamma x - \zeta_{a_p}(x) > \mathrm{B}(x) - \gamma x - \zeta_{a_{p+1}}(x), \tag{22}$$

или

$$\left(\frac{dt}{dx}\right)_{a_p} < \left(\frac{dt}{dx}\right)_{a_{p+1}}$$

на отрезке $[x_p, x_{p+1}]$. Отсюда с учетом (19), (21) и граничного условия $t(x_1) = 0$ приходим к неравенству $T^1_{p+1} > T^1_p$, а в силу того, что a_p, a_{p+1} – произвольные соседние экстремали, получаем неравенства (15) для случая А.

Используя (17), (18), для разности интегралов благосостояния $J^1_{p+1} - J^1_p$ на экстремалях a_p, a_{p+1} имеем

$$\begin{aligned} J^1_{p+1} - J^1_p = \eta\left(\int\limits_{x_{p+1}}^{x_p}\left((\pi\psi_{a_{p+1}}^{-1/(1-\alpha)}(x))^{\alpha}\frac{1}{\mathrm{B}(x)-\gamma x-\zeta_{a_{p+1}}(x)} - \frac{w_1^{\alpha}}{\mathrm{B}(x)-\gamma x-\gamma_1}\right)dx\right) + \\ + \eta\int\limits_{x_p}^{x_2}\left(\frac{(\pi\psi_{a_{p+1}}^{-1/(1-\alpha)}(x))^{\alpha}}{\mathrm{B}(x)-\gamma x-\zeta_{a_{p+1}}(x)} - \frac{(\pi\psi_{a_p}^{-1/(1-\alpha)})}{\mathrm{B}(x)-\gamma x-\zeta_{a_p}(x)}\right)dx. \end{aligned} \tag{23}$$

Подынтегральные функции первого интеграла выражения (23) удовлетворяют неравенству (20) и $(\pi\psi_{a_{p+1}}^{-1/(1-\alpha)}(x))^{\alpha} > w_1^{\alpha}$ $\forall x, x \in (x_{p+1}, x_p]$, поэтому первый интеграл больше нуля.

В силу неравенства (22) и $\psi_{a_{p+1}}^{-1/(1-\alpha)}(x) > \psi_{a_p}^{-1/(1-\alpha)}(x)$, имеющих место на отрезке $[x_p, x_2]$, приходим к выводу, что и второй интеграл больше нуля, следовательно, имеем $J^1_{p+1} > J^1_p$. Откуда приходим к неравенствам (16) для случая А.

Рассмотрим теперь случай Б (фиг. 4б). Пусть сечение x_1 находится в области $\Omega_2^+ \cup \Psi_1$ (ниже прямой ψ_1). Пусть также b_i, $i = p, p + 1$, – две произвольные соседние экстремали. Вдоль экстремалей b_i справедливы уравнения

$$\left(\frac{dx}{dt}\right)_{b_i} = -\frac{\eta}{\gamma_2 + \gamma x - \mathrm{B}(x)} \quad \text{при} \quad \psi \le \psi_1, \tag{24}$$

$$\left(\frac{dt}{dx}\right)_{b_i} = -\frac{\eta}{\zeta_{b_i}(x) + \gamma x - \mathrm{B}(x)} \quad \text{при} \quad \psi > \psi_1, \tag{25}$$

откуда для разности времен перехода $T^2_{p+1} - T^2_p$ из состояния x_1 в состояние x_2 по экстремалям b_{p+1}, b_p получаем формулу

$$\begin{aligned} T^2_{p+1} - T^2_p = \eta\int\limits_{x_p}^{x_{p+1}}\left(\frac{1}{\zeta_{b_{p+1}}(x)+\gamma x-\mathrm{B}(x)} - \frac{1}{\gamma_2+\gamma x-\mathrm{B}(x)}\right)dx + \\ + \eta\int\limits_{x_p}^{x_2}\left(\frac{1}{\zeta_{b_{p+1}}(x)+\gamma x-\mathrm{B}(x)} - \frac{1}{\zeta_{b_p}(x)+\gamma x-\mathrm{B}(x)}\right)dx. \end{aligned} \tag{26}$$

Подынтегральные функции первого интеграла в (26) на промежутке $[x_p, x_{p+1})$ удовлетворяют неравенству

$$\frac{1}{\zeta_{b_{p+1}}(x)+\gamma x-\mathrm{B}(x)} > \frac{1}{\gamma_2+\gamma x-\mathrm{B}(x)}, \tag{27}$$

поэтому на основании (27) первый интеграл больше нуля. Рассмотрим второй интеграл. Для любого $x \in [x_2, x_p]$ имеем $\psi_{b_{p+1}}(x) > \psi_{b_p}(x)$ (см. фиг. 4б), следовательно, $\zeta_{b_{p+1}}(x) < \zeta_{b_p}(x)$. В области Ω_b, ограниченной линиями ψ_1, $\psi_0(x)$, имеем также $\zeta_{b_{p+1}}(x) > \mathrm{B}(x) - \gamma x$, $\zeta_{b_p}(x) > \mathrm{B}(x) - \gamma x$; тогда получаем

$$0 < \zeta_{b_{p+1}}(x) - [\mathrm{B}(x) - \gamma x] < \zeta_{b_p}(x) - [\mathrm{B}(x) - \gamma x], \tag{28}$$

откуда следует, что и второй интеграл больше нуля, следовательно, $T^2_{p+1} > T^2_p$. В силу произвольности соседних экстремалей приходим к неравенствам (15) для случая Б.

Используя (24), (25) для разности интегралов благосостояния $J^2_{p+1} - J^2_p$ на экстремалях b_{p+1}, b_p, получаем

$$\begin{aligned} J^2_{p+1} - J^2_p = \eta\pi^\alpha \int_{x_p}^{x_{p+1}} [f_{p+1}(x) - f_p(x)]dx + \\ + \eta\pi^\alpha \int_{x_2}^{x_p} \left(\frac{1}{\dfrac{\sigma}{\psi_{b_{p+1}}(x)} - [\mathrm{B}(x) - \gamma x]\psi_{b_{p+1}}^{\alpha/(1-\alpha)}} - \frac{1}{\dfrac{\sigma}{\psi_{b_p}(x)} - [\mathrm{B}(x) - \gamma x]\psi_{b_p}^{\alpha/(1-\alpha)}(x)} \right) dx. \end{aligned} \tag{29}$$

Подынтегральные функции f_{p+1}, f_p первого интеграла в (29) монотонны, положительны. В точке x_{p+1} функции $f_{p+1}(x)$, $f_p(x)$ равны между собой. В точке x_p имеем

$$f_{p+1}(x_p) = \frac{1}{\dfrac{\sigma}{\psi_{b_{p+1}}(x_p)} - [\mathrm{B}(x_p) - \gamma x_p]\psi_{b_p}^{\alpha/(1-\alpha)}(x_p)},$$

$$f_p(x_p) = \frac{1}{\dfrac{\sigma}{\psi_{b_p}(x_p)} - [\mathrm{B}(x_p) - \gamma x_p]\psi_{b_p}^{\alpha/(1-\alpha)}(x_p)}.$$

В силу неравенства (28) и соотношения

$$\frac{\sigma}{\psi_{b_i}(x)} - [\mathrm{B}(x) - \gamma x]\psi_{b_i}^{\alpha/(1-\alpha)}(x) = \psi_{b_i}^{\alpha/(1-\alpha)}(x)\{\zeta_{b_i}(x) - [\mathrm{B}(x) - \gamma x]\}$$

имеем

$$\frac{\sigma}{\psi_{b_i}(x_p)} - [\mathrm{B}(x) - \gamma x]\psi_{b_i}^{\alpha/(1-\alpha)}(x) > 0. \tag{30}$$

Откуда с учетом того, что

$$\psi_{b_{p+1}}(x) > \psi_{b_p}(x) \quad \forall x \in [x_2, x_1], \tag{31}$$

получаем, что $f_{p+1}(x) > f_p(x)$ $\forall x \in [x_p, x_{p+1})$. Следовательно, первый интеграл больше нуля. Из неравенств (30), (31) также следует, что второй интеграл больше нуля, т.е. $J^2_{p+1} > J^2_p$. Замечание о произвольности соседних экстремалей завершает доказательство неравенств (16) для случая Б и в целом теоремы.

Экономический процесс проистекает во времени из состояния (ψ^1, x_1) в состояние (ψ^2, x_2). Задача, определяющая изменение времени, имеет вид

$$\frac{dt}{d\psi} = \frac{\eta}{\gamma - \mathrm{B}'(x)}\frac{1}{\psi}, \quad t(\psi^1) = 0.$$

Время изменяется от $t = 0$ до $t = T$. В результате различных подходов к решению краевой задачи оптимального управления оказалось, что задача довольно просто решается, если рассмотреть процесс от состояния (ψ^2, x_2) к состоянию (ψ^1, x_1), т.е. за “начальный момент” времени взять горизонт планирования – точку (ψ^2, x_2), а за конечный – точку (ψ^1, x_1). Переменную времени этого процесса обозначим через θ. Состояние системы (ψ^2, x_2) соответствует моменту $\theta = 0$, а физически истинное начальное состояние (ψ^1, x_1) становится прошлым, соответствующим времени $\theta = -T$. Поскольку $\theta = t - T$, то задача для θ имеет вид

$$\frac{d\theta}{d\psi} = \frac{\eta}{\gamma - \mathrm{B}'(x)}\frac{1}{\psi}, \quad \theta(\psi^2) = 0.$$

Теорема 3. *Пусть экономический процесс описывается системой уравнений* (12). *Пусть* x_1, x_2 – *точки начального и конечного состояния системы*, $x_1, x_2 \in \Omega$. *Тогда решение задачи оптимального управления определяется задачей Коши*:

а) *при* $x_{\min} \le x_1 < x_2 < x_\delta$ *имеем*

$$\frac{dx^*}{d\psi} = \begin{cases} \dfrac{\mathrm{B}(x^*) - \gamma x^* - \gamma_1}{\gamma - \mathrm{B}'(x^*)}\dfrac{1}{\psi}, & \psi \ge \psi_2, \\ \dfrac{\mathrm{B}(x^*) - \gamma x^* - \sigma\psi^{-1/(1-\alpha)}}{\gamma - \mathrm{B}'(x^*)}\dfrac{1}{\psi}, & \psi_0(x_2) \le \psi < \psi_2, \end{cases} \tag{32}$$

$$\frac{d\theta}{d\psi} = \frac{\eta}{\gamma - \mathrm{B}'(x^*)}\frac{1}{\psi}$$

с начальными условиями

$$\psi^2 = \psi_0(x_2) = \frac{c_0}{[\mathrm{B}(x_2) - \gamma x_2]^{1-\alpha}}, \quad x(\psi^2) = x_2, \quad \theta(\psi^2) = 0; \tag{33}$$

б) *при* $x_s < x_2 < x_1 \le x_{\max}$

$$\frac{dx^*}{d\psi} = \begin{cases} \dfrac{\mathrm{B}(x^*) - \gamma x^* - \gamma_2}{\gamma - \mathrm{B}'(x^*)}\dfrac{1}{\psi}, & \psi \le \psi_1, \\ \dfrac{\mathrm{B}(x^*) - \gamma x^* - \sigma\psi^{-1/(1-\alpha)}}{\gamma - \mathrm{B}'(x^*)}\dfrac{1}{\psi}, & \psi_0(x_2) \ge \psi > \psi_1, \end{cases} \tag{34}$$

$$\frac{d\theta}{d\psi} = \frac{\eta}{\gamma - \mathrm{B}'(x^*)}\frac{1}{\psi}$$

с начальными условиями (33).

Доказательство. Рассмотрим случай А. Пусть $x_1 < x_2$, $x_2 < x_s$. Тогда, согласно утверждению 1, оптимальная траектория находится среди интегральных кривых семейства $\{a_{i_1}\} \in \Omega_1$ (см. фиг. 4а). Если $x_1 > x_{\min}$, то по начальному состоянию x_1 строим в области Ω_1^+ промежуток $(\varphi_0^+(x_1), a_1(x_1)]$, который параллелен AA' (см. фиг. 4а). Область $\Omega_1^+(x_1) \subset \Omega_1^+$, интегральные кривые $\{a_{i_1}\} \subset \Omega_1^+(x_1)$ удовлетворяют условиям утверждения 2, согласно которому для любого x_1, $x_1 < x_2$, существует единственная интегральная кривая $a \in \{a_{i_1}\}$, соединяющая точки x_1 и x_2, $x_2 \in \psi_0^+(x)$. Это и есть интегральная кривая a_{m_1}. Согласно теореме 2, интеграл благосостояния будет наибольшим на экстремали a_{m_1}, содержащей точку (ψ^2, x_2). Координаты точки $(\psi^2, x_2) \in \psi_0^+(x)$ известны:

$$\psi^2 = \frac{c_0}{[\mathrm{B}(x_2) - \gamma x_2]^{1-\alpha}}.$$

Поэтому, решая для точки (ψ^2, x_2) задачу Коши (32), (33), находим зависимости $x^* = \tilde{x}^*(\psi)$, $\theta = f(\psi)$, а также значение T как значение решения θ со знаком минус второго уравнения системы (32) в момент, когда x^* становится равной x_1. Функция $\theta = f(\psi)$ строго монотонно убывающая, следовательно, существует f^{-1}и такая, что $\psi = f^{-1}(\theta)$, или (учитывая, что $\theta = t - T$) имеем $\psi = \psi(t)$. Далее имеем $x^* = \tilde{x}^*(\psi(t))$ или $x^* = x^*(t)$, т.е. решение задачи Коши (32), (33) определяет траектории $x^* = x^*(t)$, $\psi = \psi(t)$ для $t \in [0, T]$. Используя зависимости (8) теоремы 1, находим оптимальное управление $w^*(t)$ для $t \in [0, T]$.

В случае Б, когда $x_s < x_2 < x_1$, согласно предложению 1, оптимальная траектория находится среди семейства интегральных кривых $\{b_{i_2}\} \subset \Omega_2$ (см. фиг. 4б). Если $x_1 < x_{\max}$, то по начальному состоянию x_1 строим в области Ω_2^- промежуток $(b_1(x_1), \psi_0^-(x_1)]$, который параллелен $B'B$ (см. фиг. 4б). Область $\Omega_2^-(x_1) \subset \Omega_2^-$, интегральные кривые на $\Omega_2^-(x_1)$ удовлетворяют условиям предложения 2 и теоремы 2, поэтому, повторяя рассуждения случая А, применяя в заключение зависимости (8) теоремы 1, приходим к выводу, что решение задачи оптимального управления для случая Б определяется задачей Коши (34), (33). Теорема доказана.

Замечания. 2. Численное решение задачи Коши (32), (33) и (34), (33) на ЭВМ проводилось методом Рунге–Кутты четвертого порядка точности. На каждом шаге процедуры Рунге–Кутты считалось оптимальное управление w^* и слагаемое интеграла благосостояния. В момент завершения решения задачи Коши формируется строка времени $\{t_n\} = T + \{\theta_{N-n}\}$, $n = \overline{0, N}$, $0 \le \{t_n\} \le T$.

3. Численное решение задачи Коши (32), (33) проводится с шагом интегрирования $\Delta\psi > 0$, а задачи (34), (33) – с шагом $\Delta\psi < 0$.

3. ЧИСЛЕННЫЕ ИССЛЕДОВАНИЯ

Рассмотрим некоторые результаты расчетов оптимального управления в следующем примере экономической системы. Параметры зависимости (3), определяющие область допустимых значений управления, равны $w_1 = 2$, $w_2 = 8$. Параметры систем дифференциальных уравнений (11), (12) имеют значения $a = 0.7336$, $\gamma = 0.1711$, $\gamma_1 = 0.1874$, $\gamma_2 = 0.7495$, $\sigma = 214.5624$, $\eta = 1.3631$, $\alpha = 0.7$, $c_0 = 5.0057$, $\pi = 2.2903 \times 10^3$. Особая точка имеет координаты $\psi_s = 6.1731281$, $x_s = 1.5703512$.

Рассмотрим вначале результаты расчетов оптимального управления и соответствующих ему оптимальных траекторий для случая А: $x_{\min} \le x_1 < x_2 < x_s$.

Исходные данные вариантов x_1, x_2, соответствующие значения сопряженной переменной ψ^1, ψ^2, интеграл благосостояния, время перехода из состояния x_1 в состояние x_2 приведены в табл. 1.

На фиг. 5 показаны оптимальные траектории $x^*(t)$ между начальным и конечным состояниями системы. На фиг. 6 представлены оптимальные управления $w^*(t)$ для вариантов *1–5* (см. табл. 1).

Оптимальные траектории $x^*(t)$, оптимальные управления $w^*(t)$ имеют два характерных участка.

Участок I соответствует оптимальному потреблению, находящемуся на предельно низком уровне w_1, когда $\psi(t) \ge \psi_2$.

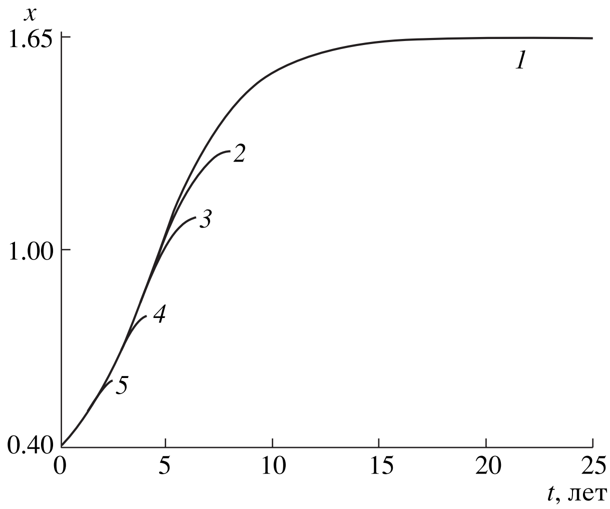
Участок II характерен тем, что в плоскости ψ, x он находится в области Ψ $(\psi_0(x) \le \psi(t) < \psi_2)$, при этом оптимальное потребление $w^*(t)$ резко увеличивается (см. фиг. 6), его величина определяется сопряженной переменной $\psi(t)$ (зависимостью (8)).

Рассмотрим теперь результаты расчетов оптимального управления, оптимальных траекторий для случая Б. Исходные данные x_1, x_2, соответствующие им значения сопряженной переменной ψ_1, ψ_2, значения интеграла благосостояния, времени перехода из состояния x_1 в состояние x_2 представлены в табл. 2.

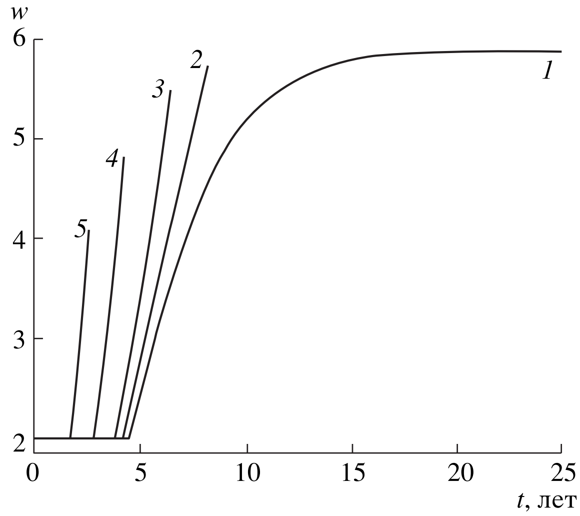
На фиг. 7 представлены оптимальные траектории $x^*(t)$. На фиг. 8 показаны оптимальные управления $w^*(t)$. Напомним, что случай Б соответствует ситуации, когда в начальном состоянии x_1 система имеет переизбыток основного капитала, его величина больше той, которую может эффективно задействовать живой капитал для производства продукции.

Таблица 1

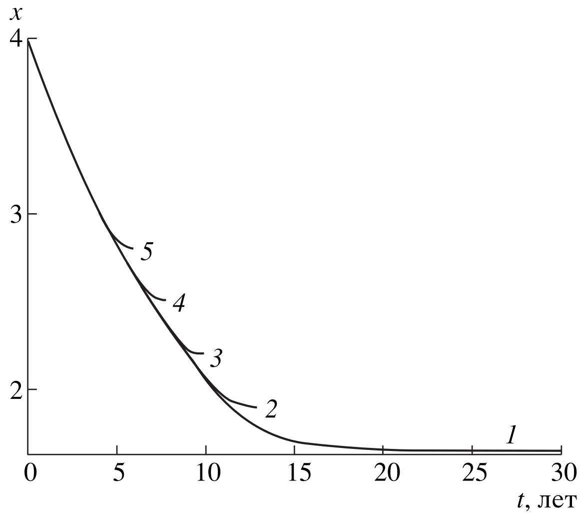
Вариант	x_1	x_2	ψ^1	ψ^2	J	T
1	0.4	1.57033	23.43058	6.17313	109.750	37.5
2	0.4	1.3	22.95956	6.19964	16.753	8.1
3	0.4	1.1	21.85935	6.26308	12.033	6.2
4	0.4	0.9	19.86746	6.38367	8.693	4.7
5	0.4	0.7	16.60768	6.59895	5.741	3.2



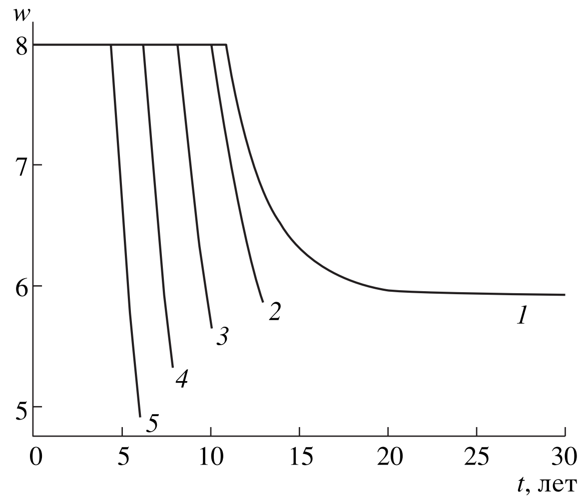
Фиг. 5.



Фиг. 6.



Фиг. 7.



Фиг. 8.

И в случае Б оптимальные траектории фазовой переменной $x^*(t)$, оптимальные управления $w^*(t)$ имеют два отмеченных выше характерных участка: I – для $w = w_2$, II – для $w^*(t)$ определяемом второй зависимостью (8).

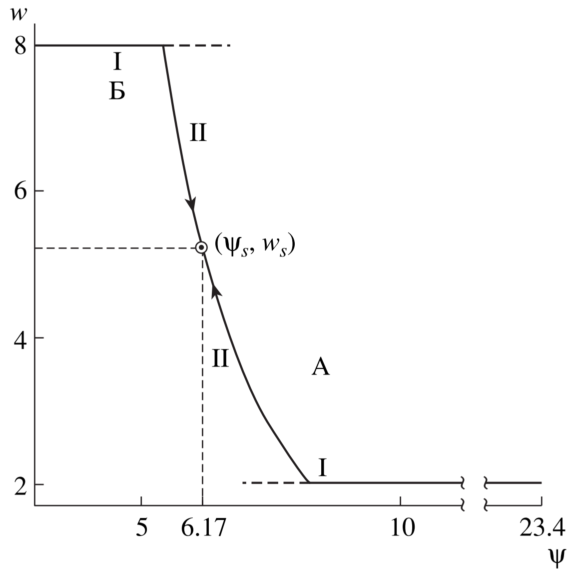
На фиг. 9 в переменных ψ, w показаны оптимальные управления $w^*(\psi)$ для вариантов *1* случаев А и Б. Здесь же показаны отмеченные участки I и II кривых.

В вариантах *1* случая А и Б фазовая переменная $x^*(t)$, сопряженная переменная $\psi(t)$ к моменту времени $t_r = 37.5$ лет (случай А), $t_r = 33.9$ лет (случай Б) достигают h-окрестности особой точки (ψ_s, x_s), при этом $x^*(t)$, $w^*(t)$ выходят с точностью до величины h на стационарные значения (см. кривые *1* на фиг. 5, 6, 7, 8):

$$x^*(t) = x_r = x_s = \underset{t}{\mathrm{const}}, \quad w^*(t) = w_r = w_s = \pi\psi_s^{-1/(1-\alpha)} = \underset{t}{\mathrm{const}}.$$

Выражая безразмерные переменные через размерные, получаем, что в h-окрестности особой точки (ψ_s, x_s) имеют место инварианты

$$\frac{K(t)}{g(t)N(t)} = \mathrm{const}, \quad \frac{W(t)}{g(t)N(t)} = \mathrm{const}. \tag{35}$$



Фиг. 9.

В простейшем случае, когда рост численности работников характеризуется темпом ν, а рост заработной платы темпом τ, можем записать

$$g(t)N(t) = g_r N_r e^{(\nu+\tau)(t-t_r)},$$

где $N_r = N_0 e^{\nu t_r}$, $g_r = g_0 e^{\tau t_r}$.

Учитывая, что при использовании производственной В(x)-функции годовой объем производства Y и инвестиции I равны (см. [7])

$$Y = C_\infty g N \mathrm{B}(x), \quad I(t) = qY(t) - W(t),$$

из инвариантов (35) следует, что макроэкономические переменные $K(t)$, $Y(t)$, $W(t)$, $I(t)$ в окрестности особой точки развиваются по "траекториям сбалансированного роста" (экспоненциальным) с темпом $\nu + \tau$, равным сумме темпа роста населения и темпа роста заработной платы.

Инварианты (35) имеют место, когда экономическая система достигает h-окрестности особой точки (ψ_s, x_s). Они отражают тот факт, что переменные x_s, w_s выходят на стационарные значения (с точностью до $O(h)$).

Рассмотрим еще один инвариант макроэкономики, имеющий другую природу. В [6] показано, что если оптимальная задача автономна (подынтегральная функция в интеграле благосостояния и правая часть уравнения (2) явно не зависят от времени), то на оптимальных траекториях функция Гамильтона постоянна:

$$H(x^*(t), \psi(t), w^*(t)) = \text{const}.$$

Таблица 2

Вариант	x_1	x_2	ψ^1	ψ^2	J	T
1	4	1.57036	2.91515	6.17313	119.66	33.9
2	4	1.8	2.96651	6.18870	44.44	10.9
3	4	2.1	3.16188	6.24917	34.94	8.5
4	4	2.4	3.46814	6.34806	28.00	6.7
5	4	2.7	3.86059	6.48277	22.08	5.5

Действительно, в нашей задаче

$$\frac{d}{dt}H(x^*(t), \psi(t), w^*(t))_{a_{i_1}(b_{i_2})} = (\alpha w^{*\alpha-1} - p\psi)\frac{dw^*}{dt} = 0,$$

где $a_{i_1} \subset \Omega_1^+$, $b_{i_2} \subset \Omega_2^-$ – оптимальные траектории между точками x_1 и x_2 (случаи А, Б).

В силу соотношений (8), для всех $t \in [0, T]$ один из сомножителей $\alpha w^{*\alpha-1} - p\psi$, dw^*/dt равен нулю, а второй ограничен.

Для управления макроэкономическим процессом наиболее содержательным является достижение экономической системой области Ψ, когда оптимальное управление (потребление) больше минимально допустимого w_1 и меньше максимально возможного w_2. Напомним, что в области Ψ_1 имеем $w^* = w_2$, а в Ψ_2 имеем $w^* = w_1$.

В области Ψ оптимальное управление и сопряженная переменная ψ удовлетворяют уравнению (9). Подставляя (9) в функцию Гамильтона (6), можем записать

$$H = \left\{w^\alpha\left[1 - \alpha + \frac{\alpha}{p}a(\mathrm{B}(x) - \gamma x)\frac{1}{w}\right]\right\}_{a_{i_1}(b_{i_2}) \subset \Psi} = \mathrm{const}.$$

Выражая безразмерные переменные через размерные и учитывая, что (см. [7]) $Y = C_\infty gNB(x)$, $p = \mu B$, $\lambda = \mu + \nu + \tau$, $a = q\mu C_\infty B$, где C_∞ – третий параметр производственной $\mathrm{B}(x)$-функции, получаем, что на части оптимальных траекторий $a_{i_1} \subset \Psi$, соединяющих начальное x_1 и конечные x_{2i_1} состояния (случай А: $x_1 < x_2 < x_s$), а также на части оптимальных траекторий $b_{i_2} \subset \Psi$ (случай Б: $x_s < x_{2i_2} < x_1$) имеет место инвариант

$$\left\{\left(\frac{W}{gN}\right)^\alpha\left[1 - \alpha + \alpha q\left(Y - \frac{\mu + \nu + \tau}{q}K\right)\frac{1}{W}\right]\right\}_{a_{i_1}(b_{i_2}) \subset \Psi} = \mathrm{const}. \tag{36}$$

Как правило, параметр $q \cong 1$ (например, для Хабаровского края $q = 0.998$).

При этом допущении ($q = 1$) макроэкономический инвариант имеет наиболее простую форму

$$\left\{\left(\frac{W}{gN}\right)^\alpha\left[1 - \alpha + \alpha(Y - (\mu + \nu + \tau)K)\frac{1}{W}\right]\right\}_{a_{i_1}(b_{i_2}) \subset \Psi} = \underset{t}{\mathrm{const}}.$$

СПИСОК ЛИТЕРАТУРЫ

1. *Булгаков В.К., Булгаков О.В.* Моделирование динамики обобщающих показателей развития региональных экономических систем России // Экономика и матем. методы. 2006. Т. 42. № 1. С. 32–49.
2. *Булгаков В.К., Стригунов В.В.* Модель и исследование макроэкономики региона на основе производственной B-функции // Вестн. Тихоокеанского гос. ун-та. Хабаровск, 2005. № 1. С. 173–196.
3. *Понтрягин Л.С., Болтянский В.Г., Гамкрелидзе Р.В., Мищенко Е.Ф.* Математическая теория оптимальных процессов. М.: Наука, 1976.
4. *Дубинский Ю.А.* Квазилинейные эллиптические и параболические уравнения любого порядка // Успехи матем. наук. 1968. Т. 23. Вып. 1. С. 45–90.
5. *Понтрягин Л.С.* Обыкновенные дифференциальные уравнения. М.: Наука, 1982.
6. *Интрилигатор М.* Математические методы оптимизации и экономическая теория. М.: Айрис-пресс, 2002.
7. *Булгаков В.К., Стригунов В.В.* Решение задачи оптимального управления динамикой региональной экономической системы для конечного горизонта планирования // Вестн. Тихоокеанского гос. ун-та. Хабаровск, 2006. № 1. С. 15–30.

ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 2007, том 47, № 8, с. 1323–1339

УДК 519.626

РАЗНОСТНАЯ АППРОКСИМАЦИЯ ЗАДАЧ ДИРИХЛЕ-НАБЛЮДЕНИЯ СЛАБЫХ РЕШЕНИЙ ВОЛНОВОГО УРАВНЕНИЯ С КРАЕВЫМИ УСЛОВИЯМИ III РОДА[1)]

© 2007 г. М. М. Потапов

(119992 Москва, Ленинские горы, МГУ, ВМК)

e-mail: mpotapov@tochka.ru

Поступила в редакцию 12.03.2007 г.

Для волнового уравнения с переменными коэффициентами и краевыми условиями II и III рода рассмотрены две взаимодвойственные задачи: задача дирихле-наблюдения со слабыми обобщенными решениями и задача управления с сильными обобщенными решениями. Для обеих задач построены конечно-разностные аппроксимации с сохранением отношения двойственности и установлена сходимость их приближенных решений в нормах соответствующих взаимно сопряженных пространств. Библ. 16.

Ключевые слова: волновое уравнение, управляемость, наблюдаемость, двойственность, конечномерная аппроксимация, сходимость.

1. ВВЕДЕНИЕ

Рассматривается пара взаимодвойственных задач граничного управления и наблюдения для волнового уравнения с переменными коэффициентами и краевыми условиями II или III рода. В задаче управления

$$\rho(x)y_{tt} = (k(x)y_x)_x, \quad 0 < t < T, \quad 0 < x < l, \tag{1.1}$$

$$-ky_x + \sigma_0 y|_{x=0} = u_0(t), \quad ky_x + \sigma_1 y|_{x=l} = u_1(t), \quad 0 < t < T, \tag{1.2}$$

$$y|_{t=0} = 0, \quad y_t|_{t=0} = 0, \quad 0 < x < l, \tag{1.3}$$

требуется соответствующим выбором управляющих воздействий $u = (u_0(t), u_1(t)) \in H = L^2(0, T) \times \times L^2(0, T)$ перевести систему (1.1)–(1.3) из нулевого начального состояния (1.3) в заданное целевое состояние $f = (f^0(x), f^1(x)) \in F = H^1(0, l) \times L^2_\rho(0, l)$:

$$y|_{t=T} = f^0(x), \quad y_t|_{t=T} = f^1(x), \quad 0 < x < l. \tag{1.4}$$

Длина $l > 0$, момент $T > 0$ и коэффициенты $\rho(x)$, $k(x)$, $\sigma_0 \ge 0$, $\sigma_1 \ge 0$ предполагаются заданными, причем

$$\rho(x) \in C^1[0, l], \quad k(x) \in C^1[0, l], \quad \rho(x) \ge \rho_* > 0, \quad k(x) \ge k_* > 0. \tag{1.5}$$

Случай граничных условий II рода, когда $\sigma_0 = \sigma_1 = 0$ в (1.2), будет рассматриваться как частный в общей постановке (1.1)–(1.3), а на отдельные его особенности мы обратим внимание по ходу дела. Двойственной по отношению к задаче управления (1.1)–(1.4) является следующая задача наблюдения:

$$\rho(x)p_{tt} = (k(x)p_x)_x, \quad 0 < t < T, \quad 0 < x < l, \tag{1.6}$$

$$-kp_x + \sigma_0 p|_{x=0} = 0, \quad kp_x + \sigma_1 p|_{x=l} = 0, \quad 0 < t < T, \tag{1.7}$$

$$p|_{t=T} = v^0(x), \quad p_t|_{t=T} = -v^1(x), \quad 0 < x < l, \tag{1.8}$$

[1)] Работа выполнена при финансовой поддержке программы “Развитие научного потенциала высшей школы” (проект РНП 2.1.1.1714).

$$p|_{x=0} = g_0(t), \quad p|_{x=l} = g_1(t), \quad 0 < t < T. \tag{1.9}$$

В этой задаче по известным дирихле-наблюдениям $g = (g_0(t), g_1(t)) \in H^* = L^2(0, T) \times L^2(0, T)$ из (1.9) требуется восстановить конечное состояние $v = (v^0(x), v^1(x)) \in F^* = L^2_\rho(0, l) \times (H^1(0, l))^*$ процесса (1.6)–(1.8), через которое, при необходимости, можно определять и другие его характеристики, решая краевую задачу (1.6)–(1.8).

Взаимодвойственная пара задач в постановке, подобной (1.1)–(1.4) и (1.6)–(1.9), рассматривалась автором ранее в [1]. В отличие от [1], здесь управления u и целевые состояния f принадлежат более гладким классам, а наблюдения g и восстанавливаемые состояния v, напротив, являются менее регулярными. При выборе управлений $u = (u_0(t), u_1(t)) \in L^2(0, T) \times L^2(0, T)$ обобщенные решения $y = y(t, x)$ задачи (1.1)–(1.3), как и в [2]–[4], принадлежат пространству $\hat{W}^1_2(Q_T)$, введенному в [5], [6], и обладают свойствами

$$\begin{gathered} y(t, \cdot) \in C([0, T]; H^1(0, l)), \quad y_t(t, \cdot) \in C([0, T]; L^2_\rho(0, l)), \\ y(\cdot, x) \in C([0, l]; H^1(\overset{\circ}{0}, T)), \quad y_x(\cdot, x) \in C([0, l]; L^2(0, T)). \end{gathered} \tag{1.10}$$

Здесь $H^1(\overset{\circ}{0}, T)$ – подпространство функций из $H^1(0, T)$, принимающих при $t = 0$ согласованные с (1.3) нулевые значения. Класс $F = H^1(0, l) \times L^2_\rho(0, l)$ целевых состояний (1.4), в свою очередь, согласован с (1.10). Решения $p = p(t, x)$ начально-краевой задачи (1.6)–(1.8) с терминальными значениями $v = (v^0(x), v^1(x)) \in L^2_\rho(0, l) \times (H^1(0, l))^*$ оказываются слабыми обобщенными решениями из класса функций с нерегулярными производными:

$$\begin{gathered} p(t, \cdot) \in C([0, T]; L^2_\rho(0, l)), \quad p_t(t, \cdot) \in C([0, T]; (H^1(0, l))^*), \\ p(\cdot, x) \in C([0, l]; L^2(0, T)), \quad p_x(\cdot, x) \in C([0, l]; (H^1(0, T))^*). \end{gathered}$$

Вопросы управляемости и минимизации соответствующего порогового момента для задач граничного управления вида (1.1)–(1.4) в функциональных классах различной степени гладкости исследованы в [7]. Явные аналитические решения таких задач, в том числе и нормальные, на достаточно протяженных временных промежутках в случае постоянных коэффициентов, граничных условий Неймана и регулярных управлений из $L^2(0, T)$ построены в [8]. В случае нерегулярных управлений для второй и третьей краевых задач с постоянными коэффициентами на временно́м промежутке критической длины явное решение задачи управления (1.1)–(1.4) приведено в [9], а для задач с переменными коэффициентами на достаточно больших временных промежутках в [1] построены и исследованы их конечномерные аппроксимации, которые могут быть использованы для отыскания приближенных решений с помощью предложенного в [10] вариационного метода.

Главной целью настоящей работы является построение устойчивых приближений к решениям обеих задач (1.1)–(1.4) и (1.6)–(1.9), а главным инструментом будет служить вариационный метод из [10] в сочетании с конечномерными аппроксимациями разностного типа. С позиций изложенного в [11] подхода к двойственным задачам управления и наблюдения как ко взаимно сопряженным операторным уравнениями вида $Au = f$ и $A^*v = g$, ключевым из условий, достаточных для применимости этого метода, является свойство корректной разрешимости сопряженного уравнения

$$\|A^*v\|^2_{H^*} \geq \mu \|v\|^2_{F^*} \quad \forall v \in F^* \tag{1.11}$$

с известным значением $\mu > 0$, являющимся одним из параметров алгоритма из [10]. В нашем случае операторы A и A^* являются линейными ограниченными взаимно сопряженными отображениями вида

$$Au = (y|_{t=T}, y_t|_{t=T}), \quad A : H \longrightarrow F, \tag{1.12}$$

$$A^*v = (p|_{x=0}, p|_{x=l}), \quad A^* : F^* \longrightarrow H^* \simeq H, \tag{1.13}$$

где $y(t, x)$ – решение задачи (1.1)–(1.3), а $p(t, x)$ – решение задачи (1.6)–(1.8).

Ниже сначала будут рассмотрены двойственные задачи (1.1)–(1.4) и (1.6)–(1.8) с двусторонними граничными управлениями и наблюдениями и для сопряженного оператора A^* получена оценка (1.11) с явным выражением для $\mu > 0$. Заметим, что из-за недостатка регулярности решения $p(t, x)$ сопряженной системы это потребует относительно бóльших по сравнению с [1] усилий. Затем оценка (1.11) будет выведена для случая односторонних управлений и наблюдений, когда управление $u_1(t)$ обращается в нуль, а значения $p|_{x=l}$ не наблюдаются, т.е. остаются неизвестными. После этого будут построены конечно-разностные аппроксимации обеих задач (1.1)–(1.4) и (1.6)–(1.9), описана процедура применения к ним вариационного метода из [10] и показано, что выполнены все условия его сходимости.

Заметим, что результаты, касающиеся вывода оценки (1.11) и не затрагивающие вопросов аппроксимации, были анонсированы в [12] и приводятся здесь в более подробном изложении.

2. КОРРЕКТНАЯ РАЗРЕШИМОСТЬ ЗАДАЧИ С ДВУСТОРОННИМИ НАБЛЮДЕНИЯМИ

В качестве базовых гильбертовых пространств функций переменных t и x возьмем, соответственно, обычное и весовое пространства Лебега $L^2(0, T)$ и $L^2_\rho(0, l)$ со скалярными произведениями

$$\langle f, g\rangle_{L^2(0,T)} = \int_0^T f(t)g(t)dt \quad \text{и} \quad \langle f, g\rangle_{L^2_\rho(0,l)} = \int_0^l \rho(x)f(x)g(x)dx.$$

Эти пространства будут отождествляться по Риссу со своими сопряженными:

$$(L^2(0, T))^* \simeq L^2(0, T), \quad (L^2_\rho(0, l))^* \simeq L^2_\rho(0, l).$$

В пространстве $H^1(0, l)$ в случае, когда $\sigma_0 + \sigma_1 > 0$, вводим скалярное произведение

$$\langle f, g\rangle_{H^1(0,l)} = \sigma_0 f(0)g(0) + \sigma_1 f(l)g(l) + \int_0^l k(x)f'(x)g'(x)dx, \tag{2.1}$$

а в случае краевых условий Неймана, когда $\sigma_0 = \sigma_1 = 0$ и билинейная форма (2.1) порождает в пространстве $H^1(0, l)$ лишь полунорму, берем вместо нее

$$\langle f, g\rangle_{H^1(0,l)} = f(0)g(0) + f(l)g(l) + \int_0^l k(x)f'(x)g'(x)dx.$$

В работе рассматриваются постановки взаимодвойственных задач с двусторонними управлениями и наблюдениями в следующих взаимно сопряженных пространствах:

$$\begin{gathered} H = L^2(0, T)\times L^2(0, T) \simeq H^*, \\ F = H^1(0, l)\times L^2_\rho(0, l), \quad F^* = L^2_\rho(0, l)\times (H^1(0, l))^*. \end{gathered} \tag{2.2}$$

В терминах пространств (2.2) интересующая нас оценка (1.11) принимает вид

$$\int_0^T (p^2(t, 0) + p^2(t, l))dt \ge \mu\left(\int_0^l \rho(x)|v^0(x)|^2dx + \|v^1\|^2_{(H^1(0,l))^*}\right). \tag{2.3}$$

Для вывода оценки (2.3) разложим пространство $H^1(0, l)$ в сумму

$$H^1(0, l) = H^1(\mathring{0}, l) + H^1(0, \mathring{l}) \tag{2.4}$$

двух гильбертовых пространств $H^1(\mathring{0}, l) = \{f(x) \in H^1(0, l) \mid f(0) = 0\}$ и $H^1(0, \mathring{l}) = \{f(x) \in H^1(0, l) \mid f(l) = 0\}$ с одинаковыми скалярными произведениями

$$\langle f, g\rangle_{H^1(\mathring{0},l)} = \langle f, g\rangle_{H^1(0,\mathring{l})} = \int_0^l k(x)f'(x)g'(x)dx. \tag{2.5}$$

Тогда справедливо следующее важное отношение двойственности (см. [13]):

$$(H^1(0,l))^* = (H^1(\overset{\circ}{0},l)+H^1(0,\overset{\circ}{l}))^* = (H^1(\overset{\circ}{0},l))^*\cap(H^1(0,\overset{\circ}{l}))^*, \tag{2.6}$$

в котором подразумевается как поэлементное совпадение, так и равенство соответствующих скалярных произведений и норм, в частности (подробности см. в [3], [4])

$$\|v^1\|^2_{(H^1(0,l))^*} = \|v^1\|^2_{(H^1(\overset{\circ}{0},l))^*\cap(H^1(0,\overset{\circ}{l}))^*} = \frac{1}{2}\|v^1\|^2_{(H^1(\overset{\circ}{0},l))^*}+\frac{1}{2}\|v^1\|^2_{(H^1(0,\overset{\circ}{l}))^*}. \tag{2.7}$$

Соотношение (2.6) дает возможность смотреть на функционалы $v^1\in(H^1(0,l))^*$ как на элементы пересечения $(H^1(\overset{\circ}{0},l))^*\cap(H^1(0,\overset{\circ}{l}))^*$. В процессе оценивания важную роль будет играть мультипликатор (см. [14]) $m(x)=m(x;\xi)$:

$$m'(x) = 1+a(x)m(x),\quad 0<x<l,\quad m(\xi)=0,$$

$$\text{где}\ \ 0<\xi<l\ \ \text{и}\ \ a(x) = \min\left\{-\frac{\rho'(x)}{\rho(x)},\frac{k'(x)}{k(x)}\right\}\ \text{при}\ \ 0\le x\le\xi,$$

$$a(x) = \max\left\{-\frac{\rho'(x)}{\rho(x)},\frac{k'(x)}{k(x)}\right\}\ \text{при}\ \ \xi\le x\le l. \tag{2.8}$$

Пороговый момент управляемости при двусторонних управлениях и наблюдениях обозначим через T_2 и определим по правилу

$$T_2 = 2\min_{0\le\xi\le l}\max_{0\le x\le l}(|m(x;\xi)|\sqrt{\rho(x)/k(x)}). \tag{2.9}$$

В итоговых оценках будем использовать мультипликатор $m_*(x)=m(x;\xi_*)$, соответствующий тому значению параметра $\xi=\xi_*$, которое доставляет минимум в (2.9).

Теорема 1. *Пусть выполняются условия* (1.5), *а значение момента T_2 определено в* (2.9). *Тогда при $T>T_2$ для решения p сопряженной задачи* (1.6)–(1.8) *справедлива оценка*

$$\|A^*v\|^2_{H^*} = \int_0^T[p^2(t,0)+p^2(t,l)]dt\ge\mu_2\left(\int_0^l\rho(x)|v^0(x)|^2dx+\|v^1\|^2_{(H^1(0,l))^*}\right) =$$
$$= \mu_2\|v\|^2_{F^*}\quad\forall v=(v^0,v^1)\in F^*, \tag{2.10}$$

в которой $\mu_2=(T-T_2)/M>0$, $M=\max\{M_0,M_1\}$, *а значения* M_0, M_1 *вычисляются по формулам*

$$M_0 = \rho(0)|m_*(0)|+\frac{T^2}{2}\left(\sigma_0\left(1+\sigma_0\frac{|m_*(0)|}{k(0)}\right)\right),$$
$$M_1 = \rho(l)m_*(l)+\frac{T^2}{2}\left(\sigma_1\left(1+\sigma_1\frac{m_*(l)}{k(l)}\right)\right). \tag{2.11}$$

Доказательство. Для избавления от неудобств, вызванных нерегулярностью компоненты $v^1\in(H^1(0,l))^*$, произведем сглаживание решения $p(t,x)$ сопряженной задачи (1.6)–(1.8) двумя способами, соответствующими разложению (2.4):

$$q(t,x) = -\int_t^T p(s,x)ds+z(x),\quad r(t,x) = -\int_t^T p(s,x)ds+w(x). \tag{2.12}$$

Здесь

$$z(x) = J_{H^1(\overset{\circ}{0},l)}v^1\in H^1(\overset{\circ}{0},l)\ \ \text{и}\ \ w(x) = J_{H^1(0,\overset{\circ}{l})}v^1\in H^1(0,\overset{\circ}{l})$$

суть образы Рисса нерегулярного элемента $v^1\in(H^1(0,l))^*$, которые при выборе в $H^1(\overset{\circ}{0},l)$ и $H^1(0,\overset{\circ}{l})$

скалярного произведения (2.5) будут связаны с v^1 следующими краевыми задачами Рисса:

$$\begin{aligned} -(k(x)z'(x))' &= \rho(x)v^1(x), \quad 0<x<l, \quad z(0)=0, \quad z'(l)=0, \\ -(k(x)w'(x))' &= \rho(x)v^1(x), \quad 0<x<l, \quad w'(0)=0, \quad w(l)=0. \end{aligned} \tag{2.13}$$

Далее будет удобнее ограничить области вариаций риссовских образов $z(x)$ и $w(x)$ всюду плотными в $H^1(\overset{\circ}{0}, l)$ и $H^1(0, \overset{\circ}{l})$ подмножествами

$$\begin{aligned} Z &= \{z(x)\in H^2(0,l) \mid z(0)=0, z'(0)=0, z'(l)=0\} \subset H^1(\overset{\circ}{0}, l), \\ W &= \{w(x)\in H^2(0,l) \mid w'(0)=0, w(l)=0, w'(l)=0\} \subset H^1(0, \overset{\circ}{l}). \end{aligned} \tag{2.14}$$

При выборе $z(x)\in Z$, $w(x)\in W$ функции $q(t, x)$ и $r(t, x)$ будут решениями задач того же типа, что и (1.6)–(1.8), но уже с регулярными конечными состояниями и частичными неоднородностями на границе:

$$\begin{gathered} \rho(x)q_{tt} = (k(x)q_x)_x, \quad 0<t<T, \quad 0<x<l, \\ -kq_x+\sigma_0 q|_{x=0} = 0, \quad kq_x+\sigma_1 q|_{x=l} = \sigma_1 z(l), \quad 0<t<T, \\ q|_{t=T} = z(x), \quad q_t|_{t=T} = v^0(x), \quad 0<x<l; \\ \rho(x)r_{tt} = (k(x)r_x)_x, \quad 0<t<T, \quad 0<x<l, \\ -kr_x+\sigma_0 r|_{x=0} = \sigma_0 w(0), \quad kr_x+\sigma_1 r|_{x=l} = 0, \quad 0<t<T, \\ r|_{t=T} = w(x), \quad r_t|_{t=T} = v^0(x), \quad 0<x<l. \end{gathered} \tag{2.15}$$

По степени гладкости функции $q(t, x)$ и $r(t, x)$ будут сильными обобщенными решениями того же класса (1.10), что и $y(t, x)$ в исходной задаче управления (1.1)–(1.3), а в таких классах получать оценки типа (1.11) удобнее, чем в менее регулярных. Умножая дифференциальное уравнение (2.15) на $q_t(t, x)$ и интегрируя по $(t, x)\in Q_t = (t, T)\times(0, l)$, получаем первое энергетическое равенство

$$\begin{gathered} E_q(t) = E_q(T) - G_1(t), \quad E_q(t) = \frac{1}{2}\int_0^l [\rho(x)q_t^2(t,x) + k(x)q_x^2(t,x)]dx, \\ G_1(t) = \frac{\sigma_0}{2}\left(\int_t^T p(s,0)ds\right)^2 + \frac{\sigma_1}{2}\left(\int_t^T p(s,l)ds\right)^2, \quad 0\le t\le T. \end{gathered} \tag{2.16}$$

При умножении дифференциального уравнения (2.15) на произведение $m_*(x)q_x(t, x)$ с последующим интегрированием по $(t, x)\in Q_t$ получаем второе энергетическое равенство

$$G_2(t) = \frac{1}{2}\iint_{Q_t} m_*[(\rho'+a\rho)q_t^2 + (ak-k')q_x^2]dsdx + \frac{1}{2}\iint_{Q_t}(\rho q_t^2 + kq_x^2)dsdx + i(T) - i(t), \quad 0\le t\le T, \tag{2.17}$$

где

$$\begin{gathered} G_2(t) = \frac{1}{2}\int_t^T\left[\frac{|m_*(0)|}{k(0)}\sigma_0^2\left(\int_s^T p(\eta,0)d\eta\right)^2 + \frac{m_*(l)}{k(l)}\sigma_1^2\left(\int_s^T p(\eta,l)d\eta\right)^2\right]ds + \\ + \frac{1}{2}\int_t^T[\rho(0)|m_*(0)|p^2(s,0) + p(l)m_*(l)p^2(s,l)]ds, \\ i(t) = \int_0^l \rho(x)m_*(x)q_t(t,x)q_x(t,x)dx, \quad 0\le t\le T. \end{gathered}$$

Первый интеграл в правой части (2.17) при выборе $a(x)$ по правилу (2.8) будет неотрицательным. Второй интеграл в силу первого энергетического равенства (2.16) будет равен

$$(T-t)E_q(T) - \int_t^T G_1(s)ds.$$

Третий и четвертый интегралы оцениваются единообразно с помощью неравенства Коши–Буняковского и (2.9):

$$|i(t)| \le \int_0^l |m|\sqrt{\rho/k}\sqrt{\rho}|q_t|\sqrt{k}|q_x|dx \le (T_2/2)E_q(t) \le (T_2/2)E_q(T), \quad 0 \le t \le T.$$

Полагая $t = 0$, получаем оценку

$$\int_0^T G_1(t)dt + G_2(0) \ge (T-T_2)E_q(T). \tag{2.18}$$

Загрубим левую часть (2.18) до обычной нормы наблюдений в пространстве $L^2(0, T)$, используя введенные в (2.10), (2.11) постоянные M_0, M_1, M:

$$\int_0^T G_1(t)dt + G_2(0) \le (M_0\|p(\cdot, 0)\|^2_{L^2(0, T)} + M_1\|p(\cdot, l)\|^2_{L^2(0, T)})/2 \le M\|A^*v\|^2_{H^*}/2, \tag{2.19}$$

и заметим, что

$$E_q(T) = (\|v^0\|^2_{L^2_\rho(0, l)} + \|z\|^2_{H^1(\overset{\circ}{0}, l)})/2.$$

В результате оценка для функции $q(t, x)$ принимает вид

$$M\|A^*v\|^2_{H^*}/2 \ge (T-T_2)(\|v^0\|^2_{L^2_\rho(0, l)} + \|z\|^2_{H^1(\overset{\circ}{0}, l)})/2. \tag{2.20}$$

Первоначально полученная для произвольных $z(x) \in Z$, оценка (2.20) распространяется по непрерывности на весь класс функций $z(x) \in H^1(\overset{\circ}{0}, l)$. Совершенно аналогично для функции $r(t, x)$ получается оценка

$$M\|A^*v\|^2_{H^*}/2 \ge (T-T_2)(\|v^0\|^2_{L^2_\rho(0, l)} + \|w\|^2_{H^1(0, \overset{\circ}{l})})/2, \tag{2.21}$$

справедливая при любых $w(x) \in H^1(0, \overset{\circ}{l})$. Перейдем от $z(x)$ и $w(x)$ к их общему риссовскому прообразу v^1. В силу (2.6), для произвольных $v^1 \in (H^1(0, l))^*$ будут выполняться обе оценки (2.20), (2.21) одновременно и после их сложения с учетом (2.7) получится итоговая оценка (2.10) из утверждения теоремы. Теорема доказана.

Оценка (2.10) означает, что за время $T > T_2$ подходящим выбором пары граничных управлений $u = (u_0(t), u_1(t)) \in L^2(0, T) \times L^2(0, T)$ систему (1.1)–(1.3) можно перевести в любую наперед заданную цель $f^0(x) \in H^1(0, l)$, $f^1(x) \in L^2_\rho(0, l)$, а в задаче (1.6)–(1.9) по паре граничных наблюдений $g = (g_0(t), g_1(t)) \in L^2(0, T) \times L^2(0, T)$ однозначно восстанавливается финальное состояние процесса $v^0(x) \in L^2_\rho(0, l)$, $v^1(x) \in (H^1(0, l))^*$.

Замечания. 1. Как видно из (2.11), постоянная μ_2 зависит от граничных коэффициентов σ_0, σ_1, но от них не зависит пороговый момент T_2, который в случае постоянных коэффициентов $\rho(x) \equiv k(x) \equiv 1$ совпадает с известным оптимальным значением $T_{2*} = \int_0^l \sqrt{\rho(x)/k(x)}\,dx = l$. Тем самым в случае двусторонних управлений и наблюдений и граничных условий Неймана, когда $\sigma_0 = \sigma_1 = 0$, полученное здесь пороговое значение T_2 является более точным по сравнению со значением из [4], где соответствующий пороговый уровень определялся с помощью комбинации двух мультипликаторов и был в $\sqrt{2}$ раз больше.

2. Задачи с двусторонними граничными управлениями и наблюдениями и краевыми условиями Дирихле в классах обобщенных решений аналогичной степени гладкости, которые рассматривались в [3], формально не укладываются в постановки (1.1)–(1.4) и (1.6)–(1.9), но для них с помощью используемой здесь техники оценивания с привлечением одного мультипликатора также возможно снижение в $\sqrt{2}$ раз полученного в [3] порогового момента управляемости–наблюдаемости до уровня, совпадающего с оптимальным значением l в случае постоянных коэффициентов.

3. ЗАДАЧИ С ОДНОСТОРОННИМИ УПРАВЛЕНИЯМИ И НАБЛЮДЕНИЯМИ

Рассмотрим задачу управления (1.1)–(1.4) в предположении, что $u_1(t) = 0$, когда управляющие воздействия сосредоточены на левом конце отрезка $x \in [0, l]$ и граничные условия (1.2) принимают вид

$$-ky_x + \sigma_0 y|_{x=0} = u_0(t), \quad ky_x + \sigma_1 y|_{x=l} = 0, \quad 0 < t < T. \tag{3.1}$$

Соответствующая ей двойственная задача наблюдения будет описываться прежней системой (1.6)–(1.8), только из двух наблюдений (1.9) останется лишь одно на левом конце отрезка:

$$p|_{x=0} = g_0(t), \quad 0 < t < T. \tag{3.2}$$

Связанный с задачей управления (1.1), (1.3), (3.1) оператор A будет действовать по прежнему правилу (1.12), а конструкция (1.13) сопряженного оператора A^* поменяется на

$$A^*v = p|_{x=0}, \quad 0 < t < T, \quad A^* : F^* \longrightarrow H^* \simeq H,$$

где $p = p(t, x)$ – решение задачи (1.6)–(1.8). Функциональные пространства по существу не меняются:

$$H \simeq H^* = L^2(0, T), \quad F = H^1(0, l) \times L^2_\rho(0, l), \quad F^* = L^2_\rho(0, l) \times (H^1(0, l))^*, \tag{3.3}$$

все различия между (2.2)–(3.3) вызваны исключительно исчезновением второго управления $u_1(t)$ и второго наблюдения $g_1(t)$. В данном случае интересующая нас оценка (1.1) примет укороченный по сравнению с (2.3) вид

$$\int_0^T p^2(t, 0)dt \ge \mu\left(\int_0^l \rho(x)|v^0(x)|^2 dx + \|v^1\|^2_{(H^1(0, l))^*}\right).$$

В отличие от случая двусторонних наблюдений здесь будут использоваться два мультипликатора, обслуживающих отдельно случаи больших и малых значений граничного коэффициента σ_1 на ненаблюдаемом конце. Правило разделения значений σ_1 на малые $0 < \sigma_1 \le \sigma_{1*}$, при которых краевое условие приближается к условию Неймана, и большие $\sigma_1 \le \sigma_{1*}$, при которых оно становится близким к условию Дирихле, будет указано ниже. В случае больших σ_1 обозначаем мультипликатор через $m_+(x)$ и определяем его из условий

$$\begin{gathered} m'_+(x) = 1 + a(x)m_+(x), \quad 0 < x < l, \quad m_+(l) = -\frac{k(l)}{\sigma_1}, \\ \text{где} \ \ a(x) = \min\left\{-\frac{\rho'(x)}{\rho(x)}, \frac{k'(x)}{k(x)}\right\}, \quad 0 \le x \le l. \end{gathered} \tag{3.4}$$

В отличие от мультипликатора $m(x)$ из (2.8), изменяющего знак во внутренней точке $\xi \in (0, l)$, все значения мультипликатора $m_+(x)$ на отрезке $[0, l]$ остаются отрицательными. Пороговый момент управляемости при односторонних управлениях и наблюдениях и больших σ_1 обозначим через $T_1^+(\sigma_1)$ и определим по правилу, аналогичному (2.9):

$$T_1^+(\sigma_1) = 2\max_{0 \le x \le l}[|m_+(x)|\sqrt{\rho(x)/k(x)}]. \tag{3.5}$$

Теорема 2. *Пусть выполняются условия* (1.5), *а значение момента* $T_1^+ = T_1^+(\sigma_1)$ *определено в* (3.5). *Тогда при* $T > T_1^+$ *для решения p сопряженной задачи* (1.6)–(1.8) *справедлива оценка*

$$\int_0^T p^2(t,0)dt \ge \mu_1^+\left(\int_0^l \rho(x)|v^0(x)|^2dx + \|v^1\|^2_{(H^1(0,l))^*}\right) = \mu_1^+\|v\|^2_{F^*} \quad \forall v \in F^*, \tag{3.6}$$

в которой

$$\mu_1^+ = \frac{T - T_1^+}{M_1^+} > 0, \quad M_1^+ = \rho(0)|m_+(0)| + \frac{T^2}{2}\left[\sigma_0\left(1 + \sigma_0\frac{|m_+(0)|}{k(0)}\right)\right].$$

Доказательство. При проведении сглаживания решения p сопряженной задачи (1.6)–(1.8) по правилам (2.12)–(2.14) первое энергетическое равенство сохранит прежний вид (2.16), а во втором энергетическом равенстве (2.17) поменяется лишь мультипликатор $m_*(x)$ на $m_+(x)$. Принципиальная разница заключается в том, что теперь в левой части итоговой оценки (3.6) не должно остаться ненаблюдаемых значений $p|_{x=l}$. Избавиться от них помогает принятое в (3.4) условие $m_+(l) = -k(l)/\sigma_1$, которое позволяет выполнить подобное соотношениям (2.18), (2.19) оценивание с привлечением только наблюдаемых значений $p|_{x=0}$:

$$(T - T_1^+)E_q(T) \le \int_0^T G_1(t)dt + G_2(0) \le M_1^+\|p(\cdot,0)\|^2_{L^2(0,T)}/2 = M_1^+\|A^*v\|^2_{H^*}/2.$$

Дальнейшие рассуждения почти дословно повторяют завершающую часть доказательства теоремы 1. Теорема 2 доказана.

Оценка (3.6) означает, что за время $T > T_1^+$ подходящим выбором граничного управления $u_0(t) \in L^2(0, T)$ систему (1.1), (1.3), (3.1) можно перевести в любую наперед заданную цель $f^0(x) \in H^1(0, l)$, $f^1(x) \in L^2_\rho(0, l)$, а в задаче (1.6)–(1.8), (3.2) по граничному наблюдению $g_0(t) \in L^2(0, T)$ однозначно восстанавливается финальное состояние процесса $v^0(x) \in L^2_\rho(0, l)$, $v^1(x) \in (H^1(0, l))^*$.

Замечание 3. Оценка (3.6) формально верна при любых $\sigma_1 > 0$, но при малых $\sigma_1 > 0$ пользоваться ею нецелесообразно, так как $T_1^+(\sigma_1) \longrightarrow +\infty$ при $\sigma_1 \longrightarrow 0$, а при больших σ_1, когда правое краевое условие в (3.1) приближается по типу к условию Дирихле, пороговое значение $T_1^+(\sigma_1)$ становится асимптотически негрубым:

$$\lim_{\sigma_1 \to +\infty} T_1^+(\sigma_1) = T_1^0 = 2\max_{0 \le x \le l}[|m_0(x)|\sqrt{\rho(x)/k(x)}]. \tag{3.7}$$

Действительно, пороговый момент T_1^0 в (3.7) определяется через мультипликатор $m_0(x)$, являющийся пределом мультипликаторов $m_+(x)$ из (3.4) при $\sigma_1 \longrightarrow +\infty$ и в случае постоянных коэффициентов $\rho(x) \equiv 1$, $k(x) \equiv 1$ в точности совпадающий с известным неулучшаемым значением $T_{1*} = 2\int_0^l \sqrt{\rho(x)/k(x)}\,dx = 2l$.

Под малыми σ_1 будем понимать значения из диапазона

$$0 < \sigma_1 < \frac{1}{2(1+\varepsilon)I_k}, \quad I_k = \int_0^l \frac{dx}{k(x)}, \tag{3.8}$$

где $\varepsilon > 0$ – некоторое фиксированное достаточно малое число. Случаю малых σ_1 сопоставим мультипликатор $m_-(x)$ вида

$$\begin{gathered} m'_-(x) = 1 + \delta_1 + a(x)m_-(x), \quad 0 < x < l, \quad m_-(l) = 0, \\ \delta_1 = 2\sigma_1(1+\varepsilon)I_k, \quad a(x) = \min\left\{-\frac{\rho'(x)}{\rho(x)}, \frac{k'(x)}{k(x)}\right\}, \quad 0 \le x \le l, \end{gathered} \tag{3.9}$$

а через него определим соответствующий критический момент $T_1^- = T_1^-(\sigma_1)$ управляемости–наблюдаемости:

$$T_1^- = \frac{2}{1-\delta_1}\max_{0\le x\le l}[|m_-(x)|\sqrt{\rho(x)/k(x)}], \quad 0<\sigma_1<[2(1+\varepsilon)I_k]^{-1}. \tag{3.10}$$

При остальных σ_1 можно формально положить

$$T_1^- = T_1^-(\sigma_1) = +\infty, \quad \sigma_1\ge[2(1+\varepsilon)I_k]^{-1}. \tag{3.11}$$

Теорема 3. *Пусть выполняются условия* (1.5), *значения* σ_1 *малы в смысле* (3.8) *и момент* $T_1^- = T_1^-(\sigma_1)$ *определен в* (3.10) *через мультипликатор* $m_-(x)$ *из* (3.9). *Тогда при* $T > T_1^-$ *для решения* p *сопряженной задачи* (1.6)–(1.8) *справедлива оценка*

$$\int_0^T p^2(t,0)dt \ge \mu_1^-\left(\int_0^l \rho(x)|v^0(x)|^2dx + \|v^1\|^2_{(H^1(0,l))^*}\right) = \mu_1^-\|v\|^2_{F^*} \quad \forall v\in F^*, \tag{3.12}$$

в которой

$$\mu_1^- = \frac{(1-\delta_1)(T-T_1^-)}{M_1^-}>0, \quad M_1^- = \rho(0)|m_-(0)| + \frac{T^2}{2}\left[(1+\delta_1)\sigma_0 + \frac{|m_-(0)|}{k(0)}\sigma_0^2 + (1+\varepsilon^{-1})\sigma_1\right]. \tag{3.13}$$

Доказательство. Снова проводим сглаживание по правилам (2.12)–(2.14) и получаем аналогичные (2.16), (2.17) энергетические равенства с участием мультипликатора $m_-(x)$. При этом в соответствии с (3.9) коэффициент 1/2 при втором двойном интеграле из (2.17) поменяется на $(1 + \delta_1)/2$. Комбинацией энергетических равенств с правилом выбора в (3.9) коэффициента $a(x)$ получаем промежуточную оценку

$$(1+\delta_1)\int_0^T G_1(t)dt + G_2(0) \ge (1+\delta_1)TE_q(T) + i(T) - i(0), \tag{3.14}$$

из которой, как и в теореме 2, нужно исключить ненаблюдаемые значения $p|_{x=l}$. Сначала с помощью неравенства Коши–Буняковского и (3.10) оценим интегралы $i(T)$ и $i(0)$ и перейдем от (3.14) к неравенству

$$(1+\delta_1)\int_0^T G_1(t)dt + G_2(0) \ge [(1+\delta_1)T - (1-\delta_1)T_1^-]E_q(T). \tag{3.15}$$

Затем займемся оценкой присутствующей в левой части (3.15) ненаблюдаемой величины

$$(1+\delta_1)\frac{\sigma_1}{2}\int_0^T\left(\int_t^T p(s,l)ds\right)^2 dt,$$

входящей в конструкцию (2.16) выражения $G_1(t)$. Для краткости введем обозначение $P(t, x) = \int_t^T p(s,x)ds$ и заметим, что в силу (2.12) имеем $P_x(t,x) = z'(x) - q_x(t,x)$; запишем формулу Ньютона–Лейбница

$$P(t,l) = P(t,0) + \int_0^l [z'(x) - q_x(t,x)]dx$$

и, фиксируя некоторое достаточно малое $\varepsilon > 0$, перейдем к оценке

$$P^2(t,l) \le (1+\varepsilon^{-1})P^2(t,0) + (1+\varepsilon)2I_k\int_0^l [|z'(x)|^2 + q_x^2(t,x)]dx.$$

Пользуясь первым энергетическим равенством (2.16) и оценкой $G_1(t) \ge \frac{\sigma_1}{2}P^2(t, l)$, получаем

$$(1+\delta_1)P^2(t, l) \le (1+\varepsilon^{-1})P^2(t, 0) + (1+\varepsilon)8I_k E_q(T). \tag{3.16}$$

Умножим обе части (3.16) на $\sigma_1/2$ и, проинтегрировав по t от 0 до T, получим

$$(1+\delta_1)\frac{\sigma_1}{2}\int_0^T P^2(t, l)dt \le (1+\varepsilon^{-1})\frac{\sigma_1}{2}\int_0^T P^2(t, 0)dt + 2\delta_1 T E_q(T). \tag{3.17}$$

Подставляя (3.17) в (3.15), приходим к неравенству

$$\left[(1+\delta_1)\sigma_0 + \frac{|m_-(0)|}{k(0)}\sigma_0^2 + (1+\varepsilon^{-1})\sigma_1\right]\int_0^T P^2(t, 0)dt + \rho(0)|m_-(0)|\int_0^T p^2(t, 0)dt \ge$$

$$\ge 2(1-\delta_1)(T - T_1^-)E_q(T) = (1-\delta_1)(T - T_1^-)(\|v^0\|^2_{L^2_\rho(0, l)} + \|z\|^2_{H^1(\overset{\circ}{0}, l)}).$$

Производя загрубление

$$\int_0^T P^2(t, 0)dt \le \frac{T^2}{2}\int_0^T p^2(t, 0)dt,$$

получаем оценку

$$\int_0^T p^2(t, 0)dt \ge \mu_1^-(\|v^0\|^2_{L^2_\rho(0, l)} + \|z\|^2_{H^1(\overset{\circ}{0}, l)}) \quad \forall v^0(x) \in L^2_\rho(0, l), \quad z(x) \in Z, \tag{3.18}$$

с постоянной μ_1^-, определенной в (3.13). Совершенно аналогично с помощью сглаживающей функции $r(t, x)$ выводится оценка

$$\int_0^T p^2(t, 0)dt \ge \mu_1^-(\|v^0\|^2_{L^2_\rho(0, l)} + \|w\|^2_{H^1(0, \overset{\circ}{l})}) \quad \forall v^0(x) \in L^2_\rho(0, l), \quad w(x) \in W. \tag{3.19}$$

Распространяя оценки (3.18), (3.19) с плотных подмножеств Z, W на $H^1(\overset{\circ}{0}, l)$, $H^1(0, \overset{\circ}{l})$, а затем складывая их, с учетом (2.7) получаем искомое соотношение (3.12). Теорема 3 доказана.

Замечание 4. Как видно из (3.8)–(3.10), $T_1^-(\sigma_1) \longrightarrow +\infty$ при $\sigma_1 \longrightarrow [2(1+\varepsilon)I_k]^{-1}$, поэтому оценка (3.12) представляет интерес лишь при достаточно малых $\sigma_1 > 0$, при которых правое краевое условие в (3.1) приближается по типу к условию Неймана, а пороговое значение $T_1^-(\sigma_1)$ становится асимптотически негрубым:

$$\lim_{\sigma_1 \to 0} T_1^-(\sigma_1) = T_1^0 = 2\max_{0 \le x \le l}[|m_0(x)|\sqrt{\rho(x)/k(x)}].$$

Здесь $m_0(x)$ – мультипликатор из замечания 3, которому в случае постоянных коэффициентов $\rho(x) \equiv 1$, $k(x) \equiv 1$ соответствует неулучшаемый пороговый момент $T_1^0 = 2l$.

Зависимость определенных в (3.5), (3.10), (3.11) пороговых значений $T_1^+(\sigma_1)$, $T_1^-(\sigma_1)$ от σ_1 является монотонной, поэтому представляется разумным разделить значения σ_1 на малые и большие находящимся на промежутке $(0, [2(1+\varepsilon)I_k]^{-1})$ корнем σ_{1*} уравнения $T_1^-(\sigma_1) = T_1^+(\sigma_1)$. Для произвольных $\sigma_1 > 0$ пороговый момент управляемости–наблюдаемости положим равным

$$T_1 = T_1^-(\sigma_1) \text{ при } 0 < \sigma_1 \le \sigma_{1*}, \quad T_1 = T_1^+(\sigma_1) \text{ при } \sigma_1 \ge \sigma_{1*}, \tag{3.20}$$

а также определим значение постоянной μ_1 в виде

$$\mu_1 = \mu_1^- \text{ при } 0 < \sigma_1 \le \sigma_{1*}, \quad \mu_1 = \mu_1^+ \text{ при } \sigma_1 \ge \sigma_{1*}. \tag{3.21}$$

Приведем итоговую оценку, составленную из (3.6) и (3.12).

Теорема 4. *Пусть выполняются условия* (1.5), *а значение момента* $T_1 = T_1(\sigma_1)$ *и постоянной* μ_1 *определены в* (3.20), (3.21). *Тогда для решения p сопряженной задачи* (1.6)–(1.8) *при* $T > T_1$ *справедлива оценка*

$$\|A^*v\|^2_{H^*} = \int_0^T p^2(t, 0)dt \ge \mu_1\left(\int_0^l \rho(x)\left|v^0(x)\right|^2 dx + \|v^1\|^2_{(H^1(0, l))^*}\right) = \mu_1\|v\|^2_{F^*} \quad \forall v = (v^0, v^1) \in F^*. \tag{3.22}$$

Оценка (3.22) означает, что за время $T > T_1$ подходящим выбором граничного управления $u_0(t) \in L^2(0, T)$ систему (1.1), (1.3), (3.1) можно перевести в любую наперед заданную цель $f^0(x) \in H^1(0, l)$, $f^1(x) \in L^2_\rho(0, l)$, а в задаче (1.6)–(1.8) по граничному наблюдению $g_0(t) \in L^2(0, T)$ из (3.2) однозначно восстанавливается соответствующее ему финальное состояние процесса $v^0(x) \in L^2_\rho(0, l)$, $v^1(x) \in (H^1(0, l))^*$.

4. РАЗНОСТНАЯ АППРОКСИМАЦИЯ ЗАДАЧ УПРАВЛЕНИЯ И НАБЛЮДЕНИЯ

Введем на отрезках $[0, T]$ и $[0, l]$ равномерные сетки с шагами $\tau = T/M$ по t, $h = l/N$ по x, узлами $t_j = j\tau$, $j = 0, 1, \ldots, M$, $x_i = ih$, $i = 0, 1, \ldots, N$, и аппроксимируем дифференциальную задачу (1.1)–(1.3) разностной схемой

$$\rho y_{\bar{t}t} = (ky_{\bar{x}})_x, \quad i = 1, 2, \ldots, N-1, \quad j = 1, 2, \ldots, M-1, \tag{4.1}$$

$$-k_1 y_{x,0} + \sigma_0 y_0 = u_0, \quad k_N y_{\bar{x},N} + \sigma_1 y_N = u_1, \quad j = 1, 2, \ldots, M-1, \tag{4.2}$$

$$y|_{j=0} = 0, \quad y_t|_{j=0} = 0, \quad i = 1, 2, \ldots, N-1. \tag{4.3}$$

Задаче наблюдения (1.6)–(1.8) сопоставим ее разностный аналог

$$\rho p_{\bar{t}t} = (kp_{\bar{x}})_x, \quad i = 1, 2, \ldots, N-1, \quad j = 1, 2, \ldots, M-1, \tag{4.4}$$

$$-k_1 p_{x,0} + \sigma_0 p_0 = 0, \quad k_N p_{\bar{x},N} + \sigma_1 p_N = 0, \quad j = 1, 2, \ldots, M-1, \tag{4.5}$$

$$p|_{j=M} = v^0, \quad p_{\bar{t}}|_{j=M} = -v^1, \quad i = 1, 2, \ldots, N-1. \tag{4.6}$$

Выберем правило вычисления коэффициентов $\rho = \rho_i = \rho(x_i)$, $k = k_i = k(x_i - 0.5h)$ и уточним размерности исходных данных: $u_0 = (u_0^1, \ldots, u_0^{M-1})$, $u_1 = (u_1^1, \ldots, u_1^{M-1})$, $v^0 = (v_1^0, \ldots, v_{N-1}^0)$, $v^1 = (v_1^1, \ldots, v_{N-1}^1)$. Будем предполагать, что выполняются следующие двусторонние условия согласования шагов сетки:

$$0 < C_1 \le \frac{\tau}{h} \le \alpha C_2, \quad C_2 = \min_{1 \le i \le N} \min\left\{\sqrt{\frac{\rho_i}{k_i}}, \sqrt{\frac{\rho_{i-1}}{k_i}}\right\}, \quad 0 < \alpha < 1, \tag{4.7}$$

в которых значения α и $C_1 \le \alpha C_2$ могут быть любыми. Как и исходные дифференциальные задачи (1.1)–(1.3) и (1.6)–(1.8), их разностные аналоги (4.1)–(4.3) и (4.4)–(4.6) находятся в отношении двойственности, которое выражается тождеством

$$\sum_{i=1}^{N-1} \rho_i(y_i^M v_i^1 + y_{\bar{t},i}^M v_i^0)h = \sum_{j=1}^{M-1} (u_0^j p_0^j + u_1^j p_N^j)\tau. \tag{4.8}$$

На базе дискретных конструкций (4.1)–(4.6) построим конечномерные взаимно сопряженные отображения $A_{\tau h}$, $A^*_{\tau h}$, действующие в тех же пространствах, что и исходные операторы A, A^*, определенные в (1.12), (1.13).

На вход оператора $A_{\tau h}$ подается произвольная пара функций $u = (u_0(t), u_1(t)) \in H = L^2(0, T) \times L^2(0, T)$, которая преобразуется в пару сеточных функций $u = (u_0, u_1)$ по правилу

$$u_0^j = \int_0^T u_0(t)e_j^{(0)}(t)dt, \quad u_1^j = \int_0^T u_1(t)e_j^{(0)}(t)dt, \quad j = 1, 2, \ldots, M-1, \tag{4.9}$$

где $e_j^{(0)}(t)$ – ступенчатые функции, равные единице внутри отрезков $t \in [t^j - 0.5\tau, t^j + 0.5\tau]$ и нулю в остальных точках промежутка $[0, T]$. Сеточная пара (u_0, u_1) подставляется в граничные условия (4.2), решается разностная задача (4.1)–(4.3) и по терминальным значениям ее решения y определяется значение

$$A_{\tau h}u = \left(\sum_{i=0}^{N} y_i^M e_i^{(1)}(x), \sum_{i=1}^{N-1} y_{\bar{t},i}^M e_i^{(0)}(x)\right) \in H^1(0, l) \times L_\rho^2(0, l), \tag{4.10}$$

где $e_i^{(1)}(x)$ – базисные сплайны первого порядка, равные $1 - h^{-1}|x - x_i|$ при $x \in [x_{i-1}, x_{i+1}]$ и нулю в остальных точках, а $e_i^{(0)}(x)$ – ступенчатые функции, равные единице при $|x - x_i| \le 0.5h$ и нулю при $|x - x_i| > 0.5h$. Присутствующие в (4.10) угловые значения y_0^M, y_N^M, формально не определяемые разностной схемой (4.1)–(4.3), находятся из условий

$$-k_1 y_{x,0}^M + \sigma_0 y_0^M = 0, \quad y_{\bar{x},N}^M + \sigma_1 y_N^M = 0, \tag{4.11}$$

продолжающих (4.2) для номера $j = M$ при $u_0^M = 0$, $u_1^M = 0$.

Оператор $A_{\tau h}^*$ применяется к произвольной паре $v = (v^0, v^1) \in F^* = L_\rho^2(0, l) \times (H^1(0, l))^*$. Сначала он преобразует ее в сеточную пару (v^0, v^1) по правилу

$$\begin{gathered} v_i^0 = \frac{1}{\rho_i h}\int_0^l \rho(x)e_i^{(0)}(x)v^0(x)dx, \quad i = 1, 2, \ldots, N-1, \\ v_i^1 = \frac{1}{\rho_i h}\langle v^1, e_i^{(1)}\rangle, \quad i = 2, 3, \ldots, N-2, \\ v_1^1 = \frac{1}{\rho_1 h}\left\langle v^1, \frac{k_1}{k_1 + h\sigma_0}e_0^{(1)} + e_1^{(1)}\right\rangle, \quad v_{N-1}^1 = \frac{1}{\rho_{N-1}h}\left\langle v^1, e_{N-1}^{(1)} + \frac{k_N}{k_N + h\sigma_1}e_N^{(1)}\right\rangle, \end{gathered} \tag{4.12}$$

где под $\langle v^1, e\rangle$ понимается значение функционала $v^1 \in (H^1(0, l))^*$ на элементе $e \in H^1(0, l)$. Затем эти сеточные функции v^0, v^1 подставляются в (4.6), решается разностная задача (4.4)–(4.6) и формируется подобное (1.13) итоговое значение

$$A_{\tau h}^* v = \left(\sum_{j=1}^{M-1} p_0^j e_j^{(0)}(t), \sum_{j=1}^{M-1} p_N^j e_j^{(0)}(t)\right) \in L^2(0, T) \times L^2(0, T). \tag{4.13}$$

Лемма 1. *Конечномерные операторы $A_{\tau h}$ и $A_{\tau h}^*$ являются взаимно сопряженными.*

Справедливость тождества $\langle A_{\tau h}u, v\rangle = \langle u, A_{\tau h}^* v\rangle$ $\forall u \in H$, $\forall v \in F^*$ из определения сопряженного оператора является следствием базового дискретного равенства (4.8) и выбранных нами способов дискретизации (4.9), (4.12) функциональных объектов и восполнения (4.10), (4.13) сеточных функций.

Лемма 2. *Операторы $A_{\tau h}$ и $A_{\tau h}^*$ являются равномерно ограниченными по τ и h.*

Доказательство. В силу известного равенства норм $\|A_{\tau h}\| = \|A_{\tau h}^*\|$ свойство равномерной ограниченности достаточно установить только для операторов $A_{\tau h}$, действующих в более гладких по сравнению с $A_{\tau h}^*$ функциональных классах. Главную роль будет играть свойство устойчивости разностной схемы с неоднородными начальными, граничными условиями и правой частью:

$$\rho y_{\bar{t}t} = (ky_{\bar{x}})_x + f, \quad i = 1, 2, \ldots, N-1, \quad j = 1, 2, \ldots, M-1, \tag{4.14}$$

$$-k_1 y_{x,0} + \sigma_0 y_0 = u_0, \quad k_N y_{\bar{x},N} + \sigma_1 y_N = u_1, \quad j = 1, 2, \ldots, M, \tag{4.15}$$

$$y\big|_{j=0} = \varphi, \quad y_t\big|_{j=0} = \psi, \quad i = 1, 2, \ldots, N-1. \tag{4.16}$$

Это свойство выражается оценками

$$\begin{gathered}\max_{1 \le j \le M} E^j \le C(\|\varphi\|_k^2 + \|\psi\|_\rho^2 + \Phi) + C(1+T)(U_0 + U_1),\\ \|y_{\bar{t},0}\|_\tau^2 + \|y_{\bar{t},N}\|_\tau^2 + \|y_{x,0}\|_\tau^2 + \|y_{\bar{x},N}\|_\tau^2 \le C(1+T)(\|\varphi\|_k^2 + \|\psi\|_\rho^2 + \Phi) + C(1+T)^2(U_0 + U_1),\end{gathered} \tag{4.17}$$

в которых $C > 0$ – некоторая постоянная, не зависящая от T, τ, h и входных данных f, u_0, u_1, φ, ψ. В (4.17) использовались следующие обозначения:

$$\begin{gathered}E^j = \frac{1}{2}(\sigma_0|y_0^j|^2 + \sigma_1|y_N^j|^2 + \|y_{\bar{t}}^j\|_\rho^2 + \|y^j\|_k^2), \quad \|y\|_\rho^2 = \sum_{i=1}^{N-1}\rho_i|y_i|^2 h,\\ \|y\|_k^2 = \sum_{i=1}^{N}k_i|y_{\bar{x},i}|^2 h, \quad \Phi = \left(\sum_{j=1}^{M-1}\left(\sum_{i=1}^{N-1}|f_i^j|^2 h\right)^{1/2}\tau\right)^2,\\ U_0 = \sum_{j=1}^{M-1}|u_0^j|^2\tau, \quad U_1 = \sum_{j=1}^{M-1}|u_1^j|^2\tau, \quad \|y\|_\tau^2 = \sum_{j=1}^{M}|y^j|^2\tau.\end{gathered} \tag{4.18}$$

В граничных условиях (4.15), как и в (4.11), подразумевается, что $u_0^M = u_1^M = 0$. У оценок (4.17) имеются прямые аналоги, означающие непрерывную зависимость от φ, ψ, u_0, u_1 и f решений соответствующей (4.14)–(4.16) дифференциальной задачи и дополняющие перечисленные в (1.10) свойства их обобщенной гладкости по t и x. Рассмотрим сначала случай $\sigma_0 + \sigma_1 > 0$, когда квадратичная форма

$$\sigma_0 f^2(0) + \sigma_1 f^2(l) + \int_0^l k(x)|f'(x)|^2 dx \tag{4.19}$$

порождает одну из эквивалентных норм в пространстве $H^1(0, l)$. Тогда имеем $\|A_{\tau h}u\|_F^2 \le CE^M$, дискретная терминальная энергия E^M оценивается сверху с помощью (4.17) при нулевых φ, ψ, Φ величиной $C(1 + T)(U_0 + U_1)$, а сумма $U_0 + U_1$ с учетом (4.9), (4.18) мажорируется нормой $\|u\|_H^2 = \|u_0\|_{L^2(0,T)}^2 + \|u_1\|_{L^2(0,T)}^2$. В результате получается искомое свойство равномерной ограниченности:

$$\|A_{\tau h}\| = \|A_{\tau h}^*\| \le C(1+T)^{1/2} \quad (\sigma_0 + \sigma_1 > 0). \tag{4.20}$$

В случае, когда $\sigma_0 = \sigma_1 = 0$ и квадратичная форма (4.19) порождает в пространстве $H^1(0, l)$ лишь полунорму, приходится отдельно оценивать отсутствующие в E^M граничные слагаемые $|y_0^M|^2$, $|y_N^M|^2$, что приводит к изменению в (4.20) характера зависимости оценочной константы от T:

$$\|A_{\tau h}\| = \|A_{\tau h}^*\| \le C(1+T)^{3/2} \quad (\sigma_0 = \sigma_1 = 0). \tag{4.21}$$

Лемма 2 доказана.

Лемма 3. *Конечномерные операторы $A_{\tau h}$ обладают свойством сильной поточечной сходимости*:

$$\|A_{\tau h}u - Au\|_F \longrightarrow 0 \quad \forall u \in H \ \text{при} \ \tau, h \longrightarrow 0. \tag{4.22}$$

Доказательство. По теореме Банаха–Штейнгауза (см. [15]), при наличии равномерной ограниченности (4.20), (4.21) свойство (4.22) достаточно установить на некотором всюду плотном в H подмножестве, например на множестве $u = (u_0(t), u_1(t)) \in \dot{C}^\infty(0, T) \times \dot{C}^\infty(0, T)$ финитных и бесконечно гладких на $(0, T)$ функций. Представим решение $y(t, x)$ задачи (1.1)–(1.3), отвечающее таким u, в виде $y(t, x) = z(t, x) + Y(t, x)$, взяв

$$z(t, x) = \frac{x(x-l)}{l}\left(\left[\frac{u_1(t)}{k(l)} - \frac{u_0(t)}{k(0)}\right]\frac{x}{l} + \frac{u_0(t)}{k(0)}\right).$$

Тогда функция $Y(t, x)$ будет решением следующей задачи с однородными начальными и граничными условиями:

$$\begin{gathered} \rho(x)Y_{tt} = (k(x)Y_x)_x + \Phi(t, x), \quad 0 < t < T, \quad 0 < x < l, \\ -kY_x + \sigma_0 Y|_{x=0} = 0, \quad kY_x + \sigma_1 Y|_{x=l} = 0, \quad 0 < t < T, \\ Y|_{t=0} = 0, \quad Y_t|_{t=0} = 0, \quad 0 < x < l, \end{gathered} \tag{4.23}$$

где $\Phi(t, x) = (k(x)z_x(t, x))_x - \rho(x)z_{tt}(t, x)$. Возьмем в пространстве $L^2_\rho(0, l)$ ортонормированный базис $\{e_k(x)\}_{k=1}^\infty$, составленный из собственных функций дифференциального оператора $Dy(x) = -(k(x)y'(x))'/\rho(x)$, удовлетворяющих на концах отрезка однородным граничным условиям $-k(0)y'(0) + \sigma_0 y(0) = 0$, $k(l)y'(l) + \sigma_1 y(l) = 0$, и аппроксимируем функцию Φ конечной суммой ее ряда Фурье

$$\Phi_K(t, x) = \sum_{k=1}^{K} \Phi_k(t)e_k(x).$$

Известно (см. [16]), что при выполнении условий (1.5) имеем $e_k(x) \in C^2[0, l]$, поэтому $\Phi_K(t, x) \in C^{\infty, 2}(\bar{Q}_T)$, а соответствующие им решения $Y_K(t, x)$ задачи (4.23) будут классическими: $Y_K(t, x) \in C^{\infty, 2}(\bar{Q}_T)$. Рассмотрим функции $y_K(t, x) = z(t, x) + Y_K(t, x)$ и введем операторы

$$A_K u = (y_K(T, x), y_{Kt}(T, x)).$$

Заметим, что, в силу финитности граничных управлений, значения операторов A и A_K можно выразить через Y и Y_K:

$$Au = (Y(T, x), Y_t(T, x)), \quad A_K u = (Y_K(T, x), Y_{Kt}(T, x)). \tag{4.24}$$

По значениям $A_K u$ построим подобные (4.10) сплайн-аппроксимации

$$A_{Kh}u = \left(\sum_{i=0}^{N} Y_K(T, x_i)e_i^{(1)}(x), \sum_{i=1}^{N-1} Y_{Kt}(T, x_i)e_i^{(0)}(x)\right) \tag{4.25}$$

и запишем оценку

$$\|A_{\tau h}u - Au\| \le \|A_{\tau h}u - A_{Kh}u\| + \|A_{Kh}u - A_K u\| + \|A_K u - Au\|. \tag{4.26}$$

Последнее слагаемое $\|A_K u - Au\|$ в (4.26) не зависит от шагов сетки и стремится к нулю при $K \longrightarrow \infty$, что следует из представлений (4.24) и сходимости правых частей дифференциальных уравнений для Y_K и Y : $\|\Phi_K - \Phi\|_{C([0, T];\, L^2_\rho(0, l))} \longrightarrow 0$ при $K \longrightarrow \infty$. Второе слагаемое $\|A_{Kh}u - A_K u\|$ в (4.26) для каждого фиксированного $K = 1, 2, \ldots$ будет стремиться к нулю при $h \longrightarrow 0$ в силу гладкости по x финальных следов $Y_K(T, x)$, $Y_{Kt}(T, x) \in C^2[0, l]$ и аппроксимационных свойств задействованных в (4.25) базисных сплайнов. При оценке первого слагаемого в (4.26), пользуясь финитностью управлений, заменим $Y_K(t, x)$ в (4.25) на $y_K(t, x)$ – классическое решение краевой задачи

$$\begin{gathered} \rho(x)y_{Ktt} = (k(x)y_{Kx})_x + \Phi_K(t, x) - \Phi(t, x), \quad 0 < t < T, \quad 0 < x < l, \\ -ky_{Kx} + \sigma_0 y_K|_{x=0} = u_0(t), \quad ky_{Kx} + \sigma_1 y_K|_{x=l} = u_1(t), \quad 0 < t < T, \end{gathered} \tag{4.27}$$

$$y_K\big|_{t=0} = 0, \quad y_{Kt}\big|_{t=0} = 0, \quad 0 < x < l.$$

Если $y^j_{Ki} = y_K(t_j, x_i)$ – сеточная функция, составленная из узловых значений решения $y_K(t, x)$ задачи (4.27), то разность $q^j_{Ki} = y^j_i - y^j_{Ki}$, $i = 0, 1, \ldots, N$, $j = 0, 1, \ldots, M$, будет решением следующей сеточной задачи типа (4.14)–(4.16):

$$\rho q_{K\bar{t}t} = (kq_{K\bar{x}})_x + (\Phi - \Phi_K)(t_j, x_i) + \alpha_K(\tau, h), \quad i = 1, 2, \ldots, N-1, \quad j = 1, 2, \ldots, M-1,$$

$$-k_1 q_{Kx,0} + \sigma_0 q_{K,0} = u^j_0 - u_0(t_j) + \gamma^0_K(h), \quad j = 0, 1, \ldots, M,$$

$$k_N q_{K\bar{x},N} + \sigma_1 q_{K,N} = u^j_1 - u_1(t_j) + \gamma^1_K(h), \quad j = 0, 1, \ldots, M,$$

$$q_K\big|_{j=0} = 0, \quad q_{Kt}\big|_{j=0} = \beta_K(\tau), \quad j = 1, 2, \ldots, N-1.$$

Здесь u^j_0, u^j_1 строятся по правилам (4.9), дополненным нулевыми концевыми значениями $u^0_0 = u^0_1 = u^M_0 = u^M_1 = 0$, а величины $\alpha_K(\tau, h)$, $\gamma^0_K(h)$, $\gamma^1_K(h)$, $\beta_K(\tau)$ для каждого фиксированного K стремятся к нулю при $\tau, h \longrightarrow 0$ в равномерной сеточной норме. Через q^j_{Ki} оценивается интересующее нас уклонение:

$$\|A_{\tau h}u - A_{Kh}u\|^2 \le C(|q^M_{K0}|^2 + |q^M_{KN}|^2 + \|q^M_K\|^2_k + \|q^M_{K\bar{t}}\|^2_\rho) + C_K\tau^2,$$

где $C, C_K > 0$ – постоянные, вторая из которых зависит от номера K. Затем применяется первая оценка из (4.17), смысл обозначений в которой описан в (4.18). В случае если оба граничных коэффициента обращаются в ноль: $\sigma_0 = \sigma_1 = 0$, применяем модификацию оценки (4.17), которая использовалась при выводе свойства (4.21). Во избежание громоздких выкладок запишем результат оценивания в предельной форме

$$\lim_{\tau, h \to 0} \sup \|A_{\tau h}u - A_{Kh}u\|^2_F \le C\|\Phi_K - \Phi\|^2_{C([0,T];\, L^2_\rho(0,l))} \quad \forall K = 1, 2, \ldots.$$

Для получения (4.22) остается перейти в (4.26) к повторному пределу $\lim\limits_{K\to\infty} \lim\limits_{\tau, h\to 0} \sup$. Лемма 3 доказана.

Лемма 4. *Конечномерные операторы* $A^*_{\tau h}$ *обладают свойством сильной поточечной сходимости*:

$$\|A^*_{\tau h}v - A^*v\|_{H^*} \longrightarrow 0 \quad \forall v \in F^* \text{ при } \tau, h \longrightarrow 0. \tag{4.28}$$

Доказательство. Равномерная ограниченность операторов $A^*_{\tau h}$ уже установлена в лемме 2, поэтому свойство сходимости (4.28) достаточно проверить на всюду плотном в F^* множестве $\mathcal{L} \times \mathcal{L}$, где в роли $\mathcal{L}$ взята линейная оболочка собственных функций $e_k(x)$, которые использовались выше при доказательстве леммы 3. При выборе $v = (v^0, v^1) \in \mathcal{L} \times \mathcal{L}$ решения $p(t, x)$ сопряженной задачи (1.6)–(1.8) будут классическими, причем $p(t, x) \in C^{\infty,2}(\bar{Q}_T)$ и значения $A^*v = (p(t, 0), p(t, l))$ оператора A^* окажутся бесконечно гладкими. Их сплайн-аппроксимации

$$A^*_\tau v = \left(\sum_{j=1}^{M-1} p(t_j, 0)e^{(0)}_j(t), \sum_{j=1}^{M-1} p(t_j, l)e^{(0)}_j(t)\right)$$

не зависят от шага h и обладают сходимостью вида

$$\|A^*_\tau v - A^*v\|_{H^*} \longrightarrow 0 \text{ при } \tau \longrightarrow 0. \tag{4.29}$$

Введем сеточную функцию $q^j_i = p^j_i - p(t_j, x_i)$, $i = 0, 1, \ldots, N$, $j = 0, 1, \ldots, M$, и оценим уклонение:

$$\|A^*_{\tau h}v - A^*_\tau v\|^2_{H^*} = \sum_{j=1}^{M-1}(|q^j_0|^2 + |q^j_N|^2)\tau \le C(|q^M_0|^2 + |q^M_N|^2 + \|q_{\bar{t},0}\|^2_\tau + \|q_{\bar{t},N}\|^2_\tau) \le C(h^2 + \|q_{\bar{t},0}\|^2_\tau + \|q_{\bar{t},N}\|^2_\tau). \tag{4.30}$$

Составим разностную задачу для $q = q_i^j$:

$$\rho q_{\bar{t}t} = (kq_{\bar{x}})_x + \alpha(\tau, h), \quad i = 1, 2, \ldots, N-1, \quad j = 1, 2, \ldots, M-1,$$

$$-k_1 q_{x,0} + \sigma_0 q_0 = \gamma^0(h), \quad k_N q_{\bar{x},N} + \sigma_1 q_N = \gamma^1(h), \quad j = 0, 1, \ldots, M, \tag{4.31}$$

$$q\big|_{j=M} = v_i^0 - v^0(x_i), \quad q_{\bar{t}}\big|_{j=M} = -v_i^1 + v^1(x_i) + \beta(\tau), \quad i = 1, 2, \ldots, N-1.$$

В (4.31) величины $\alpha(\tau, h)$, $\gamma^0(h)$, $\gamma^1(h)$, $\beta(\tau)$ стремятся к нулю при $\tau, h \longrightarrow 0$ в равномерной сеточной норме, а значения v_i^0, v_i^1 строятся по правилам (4.12), так что при фиксированных $v = (v^0, v^1) \in \in \mathcal{L} \times \mathcal{L}$ разности $v_i^0 - v^0(x_i)$ и $v_i^1 - v^1(x_i)$ также будут стремиться к нулю при $h \longrightarrow 0$ в равномерной сеточной норме. Заметим, что обе оценки из (4.17) справедливы и для разностной задачи (4.14)–(4.16) с обратным течением дискретного времени, т.е., в частности, и для задачи (4.31). Используя второе неравенство из (4.17) для продолжения оценки (4.30) и учитывая (4.29), получаем искомое предельное соотношение (4.28). Лемма 4 доказана.

Установленных в теореме 1 и леммах 1–4 свойств вполне достаточно для построения устойчивых приближений к решениям обеих взаимодвойственных задач граничного управления (1.1)–(1.4) и наблюдения (1.6)–(1.9) с помощью вариационного метода из [10] подобно тому, как это было сделано в [1]–[3]. В соответствии с [10] приближенные решения $\tilde{u}_*$, $\tilde{v}_*$ задач управления и наблюдения ищутся в истокообразном виде:

$$\tilde{u}_* = A_{\tau h}^* J_F^{-1} \tilde{v}, \quad \tilde{v} \in \mathcal{L}_h \subset F, \quad \tilde{v}_* = J_F^{-1} A_{\tau h} \tilde{u}, \quad \tilde{u} \in \mathcal{L}_\tau \subset H^* \simeq H, \tag{4.32}$$

где $J_X : X^* \longrightarrow X$ – оператор Рисса, устанавливающий изоморфное соответствие между элементами сопряженного и основного гильбертовых пространств, а $J_X^{-1} : X \longrightarrow X^*$ – обратный к нему оператор. Об отождествлениях по Риссу пространств $(L^2(0, T))^* \simeq L^2(0, T)$ и $(L_\rho^2(0, l))^* \simeq L_\rho^2(0, l)$ мы договорились с самого начала, а риссовский изоморфизм $J_{H^1(0, l)}$ между $(H^1(0, l))^*$ и $H^1(0, l)$ можно определить правилом

$$J_{H^1(0, l)} v^1 = \frac{1}{2}\left(J_{H^1(\mathring{0}, l)} v^1 + J_{H^1(0, \mathring{l})} v^1\right) \quad \forall v^1 \in (H^1(0, l))^*,$$

которое согласуется с разложением (2.4), представлением (2.6) и равенством норм (2.7). Конструкции изоморфизмов Рисса для подпространств $H^1(\mathring{0}, l)$ и $H^1(0, \mathring{l})$ описаны в (2.13). Области вариаций источников $\tilde{v}$ и $\tilde{u}$ в (4.32) ограничены конечномерными подпространствами $\mathcal{L}_h = \mathcal{L}_h^{(1)} \times \mathcal{L}_h^{(0)}$ и $\mathcal{L}_\tau = \mathcal{L}_\tau^{(0)} \times \mathcal{L}_\tau^{(0)}$, образованными линейными оболочками $\mathcal{L}_h^{(1)}$, $\mathcal{L}_h^{(0)}$ и $\mathcal{L}_\tau^{(0)}$ систем базисных сплайнов $e_i^{(1)}(x)$, $e_i^{(0)}(x)$ и $e_j^{(0)}(t)$ соответственно. Как видно из (4.10) и (4.13), эти линейные оболочки содержат в себе значения операторов $A_{\tau h}$ и $A_{\tau h}^*$. Вместо точного целевого состояния $f = (f^0(x), f^1(x))$ в задаче управления и точного сигнала $g = (g_0(t), g_1(t))$ в задаче наблюдения допускается задание некоторых их приближений $\tilde{f} \in F$ и $\tilde{g} \in H$, от которых требуется лишь асимптотическая близость к оригиналам:

$$\|\tilde{f} - f\|_F \longrightarrow 0, \quad \|\tilde{g} - g\|_H \longrightarrow 0 \text{ при } \tau, h \longrightarrow 0. \tag{4.33}$$

Элементы–источники $\tilde{v}$, $\tilde{u}$ из (4.32) являются приближенными (с точностью ε) решениями следующих конечномерных задач квадратичной минимизации с ограничениями на норму:

$$\tilde{v} \in \mathcal{L}_h, \quad \|\tilde{v}\|_F \le r, \quad I(\tilde{v}) \le \inf_{v \in \mathcal{L}_h, \|v\|_F \le r} I(v) + \varepsilon, \tag{4.34}$$

$$\tilde{u} \in \mathcal{L}_\tau, \quad \|\tilde{u}\|_H \le r^*, \quad I^*(\tilde{u}) \le \inf_{u \in \mathcal{L}_\tau, \|u\|_H \le r^*} I^*(u) + \varepsilon, \tag{4.35}$$

где $I(v) = \|A^*_{\tau h} J_F^{-1} v\|^2_H - 2\langle \tilde{f}, J_F^{-1} v\rangle$, $I^*(u) = \|A_{\tau h} u\|^2_F - 2\langle \tilde{g}, u\rangle$. Радиусы шаров r и r^* в (4.34) и (4.35) следует выбирать, с учетом указанного в теореме 1 значения постоянной μ_2, из условий

$$r \geq \|f\|_F/\mu_2, \quad r^* \geq \|g\|_H/\mu_2. \tag{4.36}$$

Сформулируем основной результат данного раздела о сходимости приближенных решений.

Теорема 5. *Пусть выполняются условия* (1.5), *условие управляемости–наблюдаемости* $T > T_2$ *со значением* T_2 *из* (2.9), *условие согласования шагов сетки* (4.7), *а приближенные данные* $\tilde{f}$ *и* $\tilde{g}$ *асимптотически точны в смысле* (4.33). *Тогда в задаче управления* (1.1)–(1.4) *для любых целевых состояний* $f^0(x) \in H^1(0, l)$, $f^1(x) \in L^2_\rho(0, l)$ *приближения* $\tilde{u}_*$, *построенные по правилам* (4.32), (4.34), (4.36), *будут сильно сходиться при* $\tau, h, \varepsilon \longrightarrow 0$ *к нормальному решению* u_* *задачи управления* $\|\tilde{u}_* - u_*\|_H \longrightarrow 0$. *При тех же условиях в задаче наблюдения* (1.6)–(1.9) *для всевозможных реализаций граничных режимов* $g_0(t) \in L^2(0, T)$, $g_1(t) \in L^2(0, T)$ *элементы* $\tilde{v}_*$, *определенные условиями* (4.32), (4.35), (4.36), *сходятся к единственному решению* v_*, *порождающему состоявшееся наблюдение* g : $\|\tilde{v}_* - v_*\|_F \longrightarrow 0$.

Замечание 5. При разностной аппроксимации задач (1.1), (1.3), (1.4), (3.1) и (1.6)–(1.8), (3.2) с одним граничным управлением и одним граничным наблюдением не возникает никаких принципиальных отличий от рассмотренного случая двусторонних управлений и наблюдений. Все изменения сведутся к удалению управляющей компоненты $u_1(t)$ и наблюдаемой компоненты $g_1(t)$, соблюдению нового требования $T > T_1$ к длине временнóго промежутка с пороговым моментом T_1 из (3.20) и замене в (4.36) значения оценочной константы μ_2 на μ_1 из (3.21). Во избежание повторов воздержимся от описания соответствующих аппроксимирующих конструкций и вычислительных процедур.

Автор выражает глубокую благодарность Ф.П. Васильеву за полезные обсуждения и замечания.

СПИСОК ЛИТЕРАТУРЫ

1. *Потапов М.М.* Метод прямых в задачах граничного управления и наблюдения для гиперболического уравнения с краевыми условиями второго и третьего рода // Вестн. Моск. ун-та. Сер. 15. Вычисл. матем. и кибернетика. 1996. № 2. С. 35–41.
2. *Потапов М.М.* Аппроксимация задачи дирихле-управления и двойственной задачи с нерегулярными Нейман-наблюдениями для волнового уравнения // Докл. РАН. 2006. Т. 408. № 5. С. 596–600.
3. *Потапов М.М.* Приближенное решение задач дирихле-управления для волнового уравнения в классах Соболева и двойственных к ним задач наблюдения // Ж. вычисл. матем. и матем. физ. 2006. Т. 46. № 12. С. 2191–2208.
4. *Потапов М.М.* Наблюдаемость нерегулярных решений задачи Неймана для волнового уравнения с переменными коэффициентами // Докл. РАН. 2007. Т. 412. № 6. С. 747–752.
5. *Ильин В.А.* Граничное управление процессом колебаний на двух концах в терминах обобщенного решения волнового уравнения с конечной энергией // Дифференц. ур-ния. 2000. Т. 36. № 11. С. 1523–1528.
6. *Ильин В.А.* Граничное управление процессом колебаний на одном конце при закрепленном втором конце в терминах обобщенного решения волнового уравнения с конечной энергией // Дифференц. ур-ния. 2000. Т. 36. № 12. С. 1670–1686.
7. *Komornik V.* Exact controllability and stabilization. The multiplier method. Chichester: John Wiley and Sons; Paris: Masson, 1994.
8. *Ильин В.А., Моисеев Е.И.* Оптимизация граничных управлений колебаниями струны // Успехи матем. наук. 2005. Т. 60. Вып. 6(366). С. 89–114.
9. *Знаменская Л.Н.* Управление упругими колебаниями. М.: Физматлит, 2004.
10. *Потапов М.М.* Устойчивый метод решения линейных уравнений с неравномерно возмущенным оператором // Докл. РАН. 1999. Т. 365. № 5. С. 596–598.
11. *Васильев Ф.П.* О двойственности в линейных задачах управления и наблюдения // Дифференц. ур-ния. 1995. Т. 31. № 11. С. 1893–1900.
12. *Потапов М.М.* Наблюдаемость нерегулярных решений третьей краевой задачи для волнового уравнения с переменными коэффициентами // Докл. РАН. 2007. Т. 414. № 6. С. 738–742.
13. *Берг Й., Лефстрем Й.* Интерполяционные пространства. Введение. М.: Мир, 1980.
14. *Ho L.F.* Exact controllability of the one-dimensional wave equation with locally distributed control // SIAM J. Control. and Optimizat. 1990. V. 28. № 3. P. 733–748.
15. *Треногин В.А.* Функциональный анализ. М.: Наука, 1980.
16. *Михайлов В.П.* Дифференциальные уравнения в частных производных. М.: Наука, 1983.

ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 2007, том 47, № 8, с. 1340–1349

УДК 519.622.2

ВТОРАЯ ГАМИЛЬТОНОВА СТРУКТУРА ЧАСТНОГО СЛУЧАЯ УРАВНЕНИЙ ЛОТКИ–ВОЛЬТЕРРА[1)]

© 2007 г. Ю. В. Бибик

(119991 Москва, ул. Вавилова, 40, ВЦ РАН)

e-mail: yvbibik@ccas.ru

Поступила в редакцию 04.04.2006 г.
Переработанный вариант 28.12.2006 г.

Рассмотрен частный случай уравнений Лотки–Вольтерра, для которого удается найти вторую гамильтонову структуру, дополнительную к уже известной. Вид нового гамильтониана позволяет найти решение в квадратурах, что и является главной особенностью рассматриваемого частного случая. Как следствие этого, выражение для периода также может быть представлено в квадратурах. Уравнения движения в новых переменных допускают механическую аналогию с колебаниями массы на нелинейной пружине. Библ. 4. Фиг. 6.

Ключевые слова: уравнения Лотки–Вольтерра, гамильтонова структура, интегрируемость в квадратурах.

1. ВВЕДЕНИЕ

Исследуется дополнительная к уже известной (см. [1], [2]) гамильтонова структура, присущая уравнениям Лотки–Вольтерра и возникающая при определенном подборе параметров. Новая структура по ряду причин удобнее для анализа системы Лотки–Вольтерра, чем известная. Она позволяет проинтегрировать уравнения движения в квадратурах и допускает механическую аналогию с колебанием тела единичной массы на нелинейной пружине. Выявление наличия у исследуемой динамической системы гамильтоновой структуры в ряде случаев значительно облегчает задачу ее исследования. Например, может быть использовано уравнение Гамильтона–Якоби для определения производящей функции, обеспечивающей каноническое преобразование от прежнего набора канонически сопряженных координат и импульсов к новым, в которых интегрирование проводится тривиально. В случае наличия нескольких пар координат и импульсов часто помогает использование метода разделения переменных. К сожалению, уравнение Гамильтона–Якоби далеко не всегда удается решить в квадратурах даже в случае одной пары координат и импульсов или после проведения разделения переменных. Система уравнений Лотки–Вольтерра, исследуемая в данной статье, обладает хорошо известной гамильтоновой структурой именно такого вида. Система используется для описания динамики эволюции видов в математической экологии и описания динамики некоторых химических реакций.

2. ПОЛУЧЕНИЕ ВТОРОЙ ГАМИЛЬТОНОВОЙ СТРУКТУРЫ

Рассмотрим частный случай системы Лотки–Вольтерра. Неизвестные фунуции обозначим через a и x, где a – численность жертв, x – численность хищников. Исследуемая система имеет вид

$$da/dt = \alpha_1 a - \gamma_1 ax, \tag{1}$$

$$dx/dt = -\alpha_2 x + \gamma_2 ax. \tag{2}$$

К уравнениям Лотки–Вольтерра может быть применен гамильтонов формализм (см. [3]). После перехода к переменным $q = v_1 = \ln(a)$, $p = v_2 = \ln(x)$ гамильтониан принимает вид

$$H = \alpha_1 v_2 + \alpha_2 v_1 - \gamma_1 e^{v_2} - \gamma_2 e^{v_1}. \tag{3}$$

Знание гамильтониана позволяет в данном случае определить фазовые траектории системы, но неясно, как исследовать динамику, поскольку вид гамильтониана приводит к уравнению Га-

[1)] Работа выполнена при финансовой поддержке РФФИ (код проекта 04-01-00309).

мильтона–Якоби, которое трудно исследовать. Попытаться построить производящую функцию перехода к каноническим координатам непосредственно при помощи формулы

$$S(I, q) = \int_{q_0}^{q} p dq,$$

где интеграл берется вдоль кривой $M_{h(I)}$, определяемой условием

$$H(p, q) = h = h(I),$$

затруднительно в силу той же причины – неудобного вида гамильтониана.

Тем не менее в одном частном случае можно построить более удобные формулы для исследования динамики, опираясь на вид известного гамильтониана, используя его в качестве первого интеграла системы.

Изменением масштаба времени и переменных a и x параметры α_1, γ_1 и γ_2 всегда могут быть приведены к единице. Потребуем, чтобы выполнялось условие $\alpha_2 = 1$.

Как обычно, при переходе к новой гамильтоновой структуре делается подходящая замена переменных. Для системы Лотки–Вольтерра соответствующая замена достаточно проста. Необходимо перейти от переменных a и x к переменным ξ и η, где

$$\xi = a + x, \quad \eta = a - x.$$

Сложим уравнения (1) и (2) и получим

$$\xi_t = \eta. \tag{4}$$

Далее вычтем из уравнения (1) уравнение (2) и получим

$$\eta_t = \xi - 2xy = \xi - \frac{1}{2}(\xi^2 - \eta^2). \tag{5}$$

Воспользуемся теперь известным законом сохранения (3):

$$H = v_2 + v_1 - e^{v_2} - e^{v_1} = \mathrm{const},$$

или

$$\ln(a) + \ln(x) - a - x = \ln(ax) - a - x = C,$$

из которого следует, что

$$ax = \frac{A}{2}e^{a+x} = \frac{A}{2}e^{\xi}.$$

$$A = 2e^{C}.$$

Подставив это выражение в (5), получим

$$\eta_t = \xi - Ae^{\xi}. \tag{6}$$

Уравнения (4), (6) теперь можно записать в гамильтоновой форме:

$$\xi_t = \eta = \partial H/\partial \eta. \tag{7}$$

$$\eta_t = \xi - Ae^{\xi} = -\partial H/\partial \xi. \tag{8}$$

Возможность использования новой гамильтоновой структуры в данном случае основана на двух совпадениях. Во-первых, алгебраическая структура правых частей уравнения Лотки–Вольтерра позволяет после указанной замены переменных получить в уравнении для ξ_t правую часть, зависящую только от η и, более того, линейную по η, что обеспечивает появление в новом гамильтониане выражения, квадратичного по обобщенному импульсу η. Во-вторых, закон сохранения позволяет в уравнении для η_t сделать правую часть зависящей только от ξ и записать соответствующую правую часть как производную гамильтониана по обобщенной координате ξ со знаком минус, что и требовалось. Искомый гамильтониан имеет вид

$$G = \frac{1}{2}\eta^2 - \frac{1}{2}\xi^2 + Ae^{\xi}. \tag{9}$$

Поскольку полученный гамильтониан квадратичен по импульсу η, уравнения Гамильтона можно свести к одному уравнению относительно координаты ξ и оно будет иметь ньютонов вид. Убедимся в этом. Уравнения Гамильтона имеют вид (7), (8).

Продифференцируем уравнение (7) по t и воспользуемся уравнением (8). Получим

$$d^2\xi/dt^2 = d\eta/dt = \xi - Ae^{\xi}. \tag{10}$$

Уравнение (10) представляет собой уравнение колебаний тела единичной массы на нелинейной пружине.

Теперь мы можем сформулировать основной результат данной работы.

Предложение 1. *Система уравнений* (1), (2) *при* $\alpha_1 = \alpha_2 = \gamma_1 = \gamma_2 = 1$ *имеет кроме гамильтоновой структуры с гамильтонианом* H (см. [1], [2]) *вторую гамильтонову структуру с гамильтонианом*

$$G = \frac{1}{2}\eta^2 - \frac{1}{2}\xi^2 + Ae^{\xi}$$

и уравнениями движения (7), (8), *которые сводятся к одному интегрируемому в квадратурах уравнению второго порядка*, *описывающему нелинейные колебания и имеющему вид* (10).

Заметим, что полученный гамильтониан зависит от интеграла движения $A = 2e^H$. Сам он, естественно, тоже является интегралом движения, но всегда равен нулю. Убедиться в этом факте легко. Поскольку $A = 2axe^{-(a+x)}$, то имеет место формула

$$A = 2axe^{-\xi}.$$

С другой стороны, имеем

$$ax = \frac{1}{4}(\xi^2 - \eta^2).$$

Поэтому

$$A = \frac{1}{2}(\xi^2 - \eta^2)e^{-\xi}. \tag{11}$$

Подставив формулу (11) в (9), получим

$$G = \frac{1}{2}\eta^2 - \frac{1}{2}\xi^2 + \frac{1}{2}(\xi^2 - \eta^2)e^{-\xi}e^{\xi} = 0. \tag{12}$$

Таким образом, имеет место

Предложение 2. *Гамильтониан* G *всегда равен нулю.*

Натяжение пружины, естественно, тоже зависит от интеграла движения A.

Следует отметить, что, вообще говоря, речь идет о системе гамильтоновых структур, сопоставляемых данной задаче и параметризуемых первым интегралом A. Структура траекторий в фазовых плоскостях для каждой из рассматриваемых гамильтоновых структур отличается от структуры траекторий в фазовой плоскости рассматриваемой задачи, но фазовая траектория, соответствующая данной величине A, совпадает с одной из траекторий в фазовой плоскости для гамильтониана $G(A)$, и мы исследуем движение по ней с помощью этого гамильтониана. Поэтому более правильно было бы вести речь о системе квадратичных по импульсу гамильтонианов, ассоциированных с данной задачей, но для краткости мы говорим о второй гамильтоновой структуре. На результатах, полученных с помощью формул (7), (8), это никак не сказывается. Интересно то, что для каждой ассоциированной с данной задачей гамильтоновой структуры с гамильтонианом $G(A)$ мы выбираем в качестве исследуемой фазовую траекторию, для которой $G(A) = 0$, как это следует из предложения 2.

3. ПОЛУЧЕНИЕ УРАВЕНЕНИЯ ДИНАМИКИ И ВЫРАЖЕНИЯ ДЛЯ ПЕРИОДА В КВАДРАТУРАХ

Вернемся к вопросу об интегрировании уравнения (10). Проблема интегрирования уравнения нелинейных колебаний подробно освещена, например, в [3].

Вид гамильтониана позволяет решить уравнение Гамильтона–Якоби в квадратурах, но проще воспользоваться уравнениями (7), (9), (12) и заменить η на $d\xi/dt$. Тогда имеет место уравнение

$$G = \frac{1}{2}\left(\frac{d\xi}{dt}\right)^2 - \frac{1}{2}\xi^2 + Ae^{\xi}. \tag{13}$$

Как следствие уравнения (3.1) имеет место уравнение

$$\xi_t^2 = \xi^2 - 2Ae^{\xi},$$

которое можно решить методом разделения переменных (см. [3]). Применив этот метод, получим

$$\frac{d\xi}{\sqrt{\xi^2 - 2Ae^{\xi}}} = dt,$$

или

$$\int_{\xi_*}^{\xi} \frac{d\zeta}{\sqrt{\zeta^2 - 2Ae^{\zeta}}} = t, \tag{14}$$

если $\xi_* < \xi$ и движение происходит вправо в сторону увеличения ξ (t – время движения от ξ_* до ξ), и

$$\int_{\xi}^{\xi_*} \frac{d\zeta}{\sqrt{\zeta^2 - 2Ae^{\zeta}}} = t, \tag{15}$$

если $\xi_* > \xi$, а движение происходит влево в сторону уменьшения ξ.

Пусть ξ_1 и ξ_2 – корни уравнения

$$\xi_{1,2}^2 - 2Ae^{\xi_{1,2}} = 0. \tag{16}$$

Тогда имеем выражение для периода колебаний T (см. [3]), которое описывает

Предложение 3. *Для периода колебаний T имеет место следующая формула*:

$$2\int_{\xi_1}^{\xi_2} \frac{d\zeta}{\sqrt{\zeta^2 - 2Ae^{\zeta}}} = T. \tag{17}$$

4. МЕТОД РАСТЯНУТЫХ ПАРАМЕТРОВ

Уравнение для периода колебаний (17) и уравнения (14), (15) для зависимости ξ от t зависят от интеграла движения $A = 2e^H$. Он принимает максимальное значение $A_m = 2e^{-2}$ на стационарном решении. В этом легко убедиться, если вспомнить, что $e^H = axe^{-(a+x)} = (ae^{-a})(xe^{-x})$ и функция $f(z) = ze^{-z}$ достигает максимума при $z = 1$, где $f(1) = e^{-1}$. Значения $a = 1$ и $x = 1$ соответствуют как раз стационарному решению, в чем можно непосредственно убедиться подстановкой этих значений в (1), (2). Будем рассматривать малые колебания, описываемые уравнением (10). Им будет соответствовать интеграл движения A, мало отличающийся от A_m. Запишем A в виде

$$A = A_m - \delta, \tag{18}$$

где δ мало́. Теперь попробуем найти зависимость $\omega(A) = 2\pi/T(A)$ в виде $\omega(\delta)$ и зависимость $\xi(t, A)$ в виде $\xi(t, \delta)$ с помощью теории возмущений, используя малость параметра δ.

Применим для анализа уравнения (10) одну из разновидностей методов теории возмущений – метод растянутых параметров (см. [4]). Его суть заключается в корректировке частоты основной гармоники системы по ходу решения уравнений в каждом порядке теории возмущений. В качестве малого параметра α будет выступать разность $\xi_2 - \xi_0$ или $\xi_1 - \xi_0$, где величины ξ_1 и ξ_2 определяются из уравнения (16), а величина ξ_0 – из уравнения

$$\xi_0 = Ae^{\xi_0}. \tag{19}$$

Такой выбор малого параметра удобнее в техническом отношении, затем он будет связан с параметром δ.

Для определенности в качестве малого параметра выберем разность $\xi_2 - \xi_0$.

Если использовать механическую аналогию для описания уравнения (10), то величина ξ_0 соответствует дну потенциальной ямы, в которой колеблется частица с координатой ξ, а точки ξ_1 и ξ_2 – точки поворота, в них скорость частицы обращается в нуль и вся энергия является потенциальной.

Уравнение (10) нужно преобразовать с помощью замены

$$\xi = \xi_0 + x,$$

и оно примет вид

$$x_{tt} = x + \xi_0 - Ae^{\xi_0}e^x. \qquad (20)$$

После разложения в ряд по x функции e^x и группировки выражения в правой части (20) по степеням x получим

$$x_{tt} = (1-\xi_0)x - \xi_0\left(\frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \ldots\right). \qquad (21)$$

Далее для удобства полагаем, что колебания начинаются в одной из точек поворота, например из ξ_2. Тогда имеем

$$x|_{t=0} = \alpha, \quad x_t|_{t=0} = 0.$$

Теперь сделаем замену переменных

$$x = \alpha z.$$

Начальные условия примут вид

$$z|_{t=0} = 1, \quad z_t|_{t=0} = 0,$$

а уравнение (21) запишется в форме

$$z_{tt} = (1-\xi_0)z - \xi_0\left(\alpha\frac{z^2}{2} + \alpha^2\frac{z^3}{6} + \alpha^3\frac{z^4}{24} + \ldots\right). \qquad (22)$$

На этом этапе следует учесть зависимость величины ξ_0 от параметра α:

$$\xi_0 = \xi_{00} + \alpha\xi_{01} + \alpha^2\xi_{02} + \ldots. \qquad (23)$$

Схематично опишем метод получения параметров ξ_{0i}, $i = 0, 1, \ldots$.

Вспомним, что параметр A представи́м в виде (18).

Будем искать величины $\xi_0(A)$ и $\xi_2(A)$ в виде

$$\xi_0 = \xi_{00}^- + \delta\xi_{01}^- + \delta^2\xi_{02}^-\ldots, \qquad (24)$$

$$\xi_2 = \xi_{20} + \sqrt{\delta}\xi_{21} + \delta\xi_{22} + \ldots. \qquad (25)$$

Величина ξ_{20} соответствует экстремуму функции $\frac{\xi_2^2}{2}e^{-\xi_2}$, в котором она принимает значение A_m, поэтому величина δ представляет собой квадратичную функцию от $\xi_2 - \xi_{20}$ при малых значениях $\xi_2 - \xi_{20}$. В силу этого величина ξ_2 находится в виде ряда по степеням $\sqrt{\delta}$. Зависимость ξ_2 и ξ_1 от $A = A_m - \delta$ представлена жирной линией на фиг. 1.

Подставим разложения (23)–(25) в уравнения (19), (16). Приравнивая выражения при одинаковых степенях параметра δ, после несложных, но громоздких вычислений получаем

$$\xi_{00}^- = 2, \quad \xi_{01}^- = e^2, \quad \xi_{21} = e\sqrt{2}.$$

Вспомним теперь, что $\alpha = \xi_2 - \xi_0$. В самом грубом приближении, удерживая выражение порядка $\sqrt{\delta}$ и опуская величины порядка δ и выше, получаем

$$\alpha = e\sqrt{2}\sqrt{\delta}. \qquad (26)$$

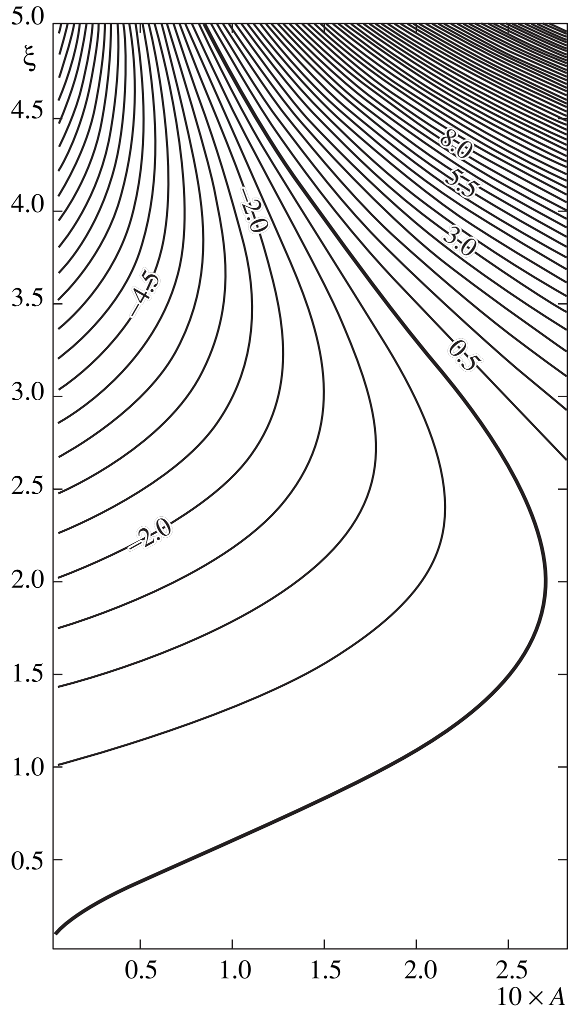
Обращая эту формулу, получаем при малых δ

$$\delta = \alpha^2/(2e^2). \qquad (27)$$

Комбинируя теперь формулы (24) и (27), получим зависимость ξ_0 от α:

$$\xi_{00} = 2, \quad \xi_{01} = 0, \quad \xi_{02} = 1/2.$$

Далее для удобства вместо обозначения зависимой переменной z опять будем использовать обозначение x.



Фиг. 1.

Теперь можно сформулировать

Предложение 4. *Решение уравнения* (22) *с точностью до членов второго порядка по* α *имеет вид*

$$x = x_0 + \alpha x_1 + \alpha^2 x_2, \quad x_0 = \cos s,$$

$$x_1 = -\frac{1}{2} + \frac{1}{3}\cos s + \frac{1}{6}\cos(2s),$$

$$x_2 = -\frac{1}{3} + \frac{55}{288}\cos s + \frac{1}{9}\cos(2s) + \frac{1}{32}\cos(3s),$$

$$t = s(1 + \alpha\omega_1 + \alpha^2\omega_2 + \ldots), \quad s = \tilde{\omega}t,$$

$$\tilde{\omega} = 1 - \alpha\omega_1 - \alpha^2(\omega_2 + \omega_1^2) - \ldots, \quad \omega_1 = 0, \quad \omega_2 = 1/24.$$

Формула (26) *в выбранном приближении дает выражение* α *через* δ.

Доказательство. Будем решать уравнение (22) методом растянутых параметров. Для этого сделаем замену переменных

$$t = s(1 + \alpha\omega_1 + \alpha^2\omega_2 + \ldots)$$

и используем разложение в ряд

$$x = \sum_{n=0}^{\infty} \alpha^n x_n.$$

Получим уравнение

$$\frac{d^2}{ds^2}(x_0 + \alpha x_1 + \alpha^2 x_2 + \alpha^3 x_3 + \ldots) =$$

$$= (1 + \alpha\omega_1 + \alpha^2\omega_2 + \ldots)^2\Big[(1 - \xi_{00}) - \alpha\xi_{01} - \alpha^2\xi_{02} - \ldots)(x_0 + \alpha x_1 + \alpha^2 x_2 + \ldots) -$$

$$-(\xi_{00} + \alpha\xi_{01} + \alpha^2\xi_{02} + \ldots)\left(\frac{\alpha}{2}(x_0 + \alpha x_1 + \alpha^2 x_2 + \ldots)^2 + \frac{\alpha^2}{6}(x_0 + \alpha x_1 + \alpha^2 x_2 + \ldots)^3 + \ldots\right)\Big].$$

Используем тот факт, что

$$(1 + \alpha\omega_1 + \alpha^2\omega_2 + \ldots)^2 = 1 + \alpha(2\omega_1) + \alpha^2(2\omega_2 + \omega_1^2) +$$

$$+ \alpha^3(2\omega_3 + 2\omega_2\omega_1) + \alpha^4(2\omega_4 + 2\omega_3\omega_1 + \omega_2^3) + \ldots.$$

Кроме того, получим выражения для степеней x:

$$(x_0 + \alpha x_1 + \alpha^2 x_2 + \ldots)^2 = x_0^2 + \alpha(2x_1x_0) + \alpha^2(2x_2x_0 + x_1^2) +$$

$$+ \alpha^3(2x_3x_0 + 2x_2x_1) + \alpha^4(2x_4x_0 + 2x_3x_1 + x_2^2) + \ldots,$$

$$(x_0 + \alpha x_1 + \alpha^2 x_2 + \ldots)^3 = x_0^3 + \alpha(3x_1x_0^2) + \alpha^2(3x_2x_0^2 + 3x_0x_1^2) + \alpha^3(3x_3x_0^2 + 6x_0x_1x_2 + x_1^3) + \ldots.$$

Теперь уже легко получить уравнения в нескольких порядках по α, группируя члены с одинаковыми степенями α.

Первые три уравнения имеют вид

$$x_{0tt} + x_0 = 0,$$

$$x_{1tt} + x_1 + (2\omega_1 + \xi_{01})x_0 + x_0^2 = 0, \tag{28}$$

$$x_{2tt} + x_2 + (2\omega_2 + \omega_1^2 + \xi_{02} + \omega_1\xi_{01})x_0 + (2\omega_1 + \xi_{01})x_1 + 2x_0x_1 + (4\omega_1 + \xi_{01})\frac{x_0^2}{2} + \frac{1}{3}x_0^3 = 0. \tag{29}$$

Решение первого из этих трех уравнений с учетом начальных условий имеет вид

$$x_0 = \cos s.$$

Подставим его в уравнение (28) и используем произвол в выборе параметра ω_1 для уничтожения вековых членов. После несложных, но громоздких вычислений получим

$$\omega_1 = 0,$$

$$x_1 = -\frac{1}{2} + \frac{1}{3}\cos s + \frac{1}{6}\cos(2s).$$

Подставим полученные решения в уравнение (29) и потребуем уничтожения вековых членов. Тогда получим уравнение для ω_2:

$$2\omega_2 + \xi_{02} = 7/12,$$

откуда следует, что

$$\omega_2 = 1/24.$$

Функция x_2 тогда примет вид

$$x_2 = -\frac{1}{3} + \frac{55}{288}\cos s + \frac{1}{9}\cos(2s) + \frac{1}{32}\cos(3s).$$

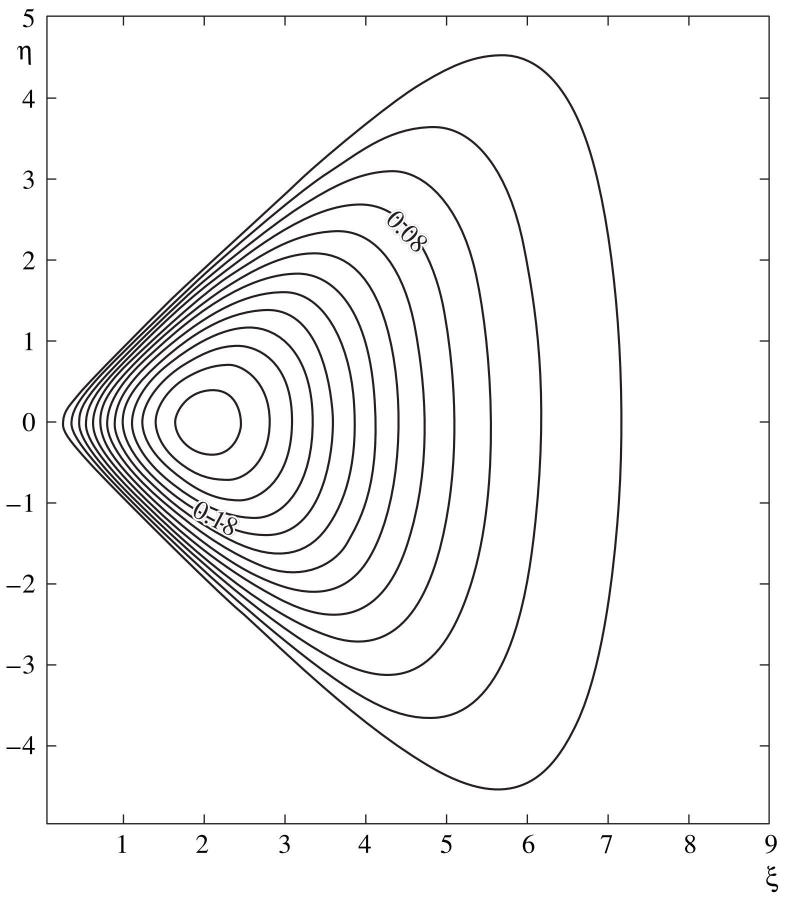
5. АНАЛИЗ РИСУНКОВ

На фиг. 2 приведен график функции $A = \frac{1}{2}(\xi^2 - \eta^2)e^{-\xi}$. Линии уровня на этом графике представляют собой траекторию, по которой движется на плоскости точка с координатами ξ и η, отображающая состояние системы с заданным параметром A, который и соответствует высоте линии уровня.

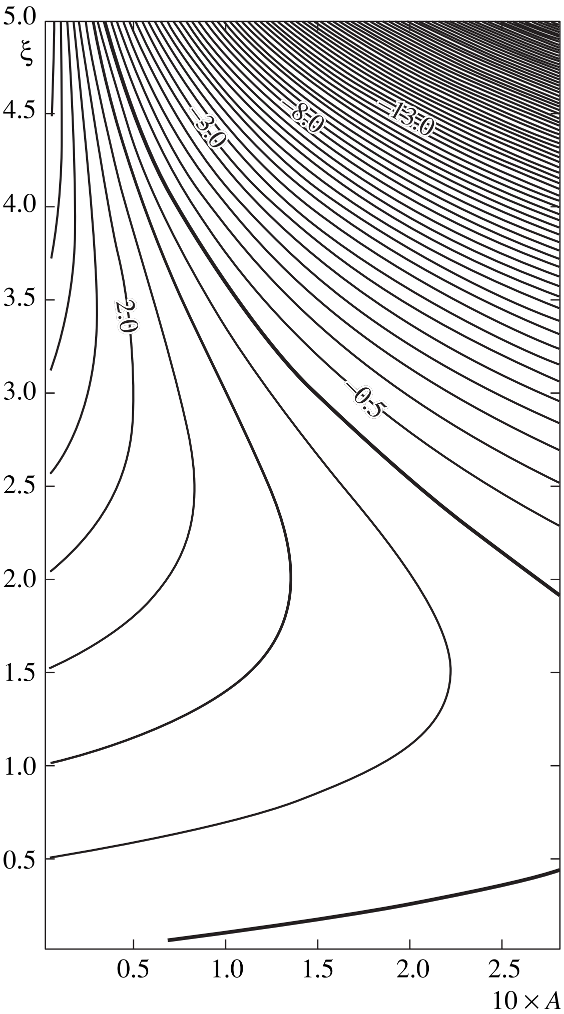
На фиг. 1 приведен график функции $\frac{1}{2}\xi^2 - Ae^{-\xi}$. При заданном A эта функция описывает потенциальную яму, в которой движется точка с координатой ξ, описывающая систему. График изображает сразу весь набор потенциальных ям, соответствующих разным значениям параметра A. Нули этой функции соответствуют точкам поворота, где происходит остановка движения изображающей систему точки с координатой ξ и начинается движение в обратную сторону.

Жирная линия на этом графике представляет собой линию нулевого уровня и является графиком зависимости точек поворота ξ_1, ξ_2 от параметра A. При максимальном значении параметра A они смыкаются в точке ξ с координатой 2, соответствующей стационарному состоянию системы.

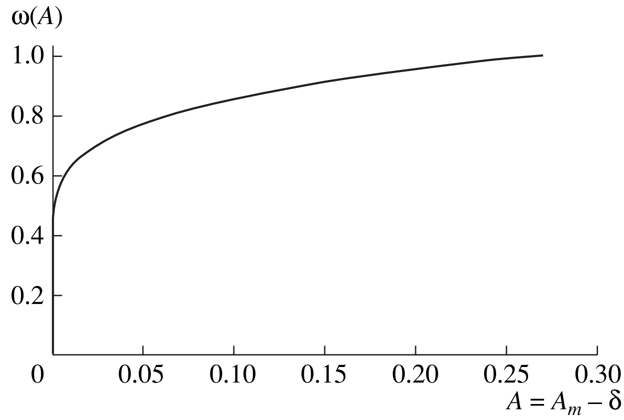
На фиг. 3 изображен график функции $\xi - Ae^{-\xi}$, представляющей собой силу, вызывающую колебания точки с координатой ξ. График изображает сразу весь набор сил, соответствующих разным значениям параметра A. Нули этой функции соответствуют дну потенциальных ям, изображенных на фиг. 1. Жирная линия является линией нулевого уровня. Одна из ветвей этой линии лежит между точками поворота, другая лежит вне области колебаний и не имеет отношения к процессу колебаний. При данном A координата первой ветви этой линии соответствует точке ξ_0 и разность $\xi_2 - \xi_0$ соответствует малому параметру α, используемому в разд. 4.



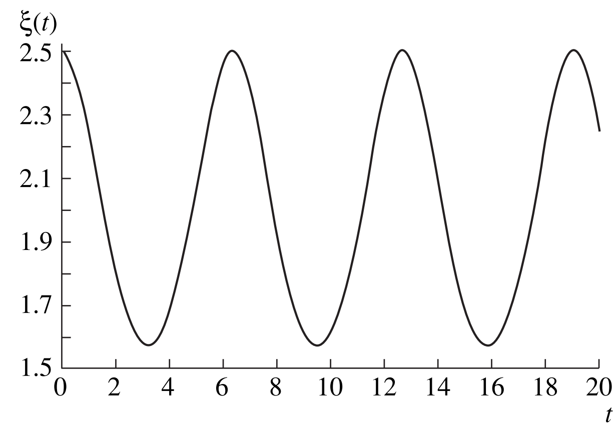
Фиг. 2.



Фиг. 3.



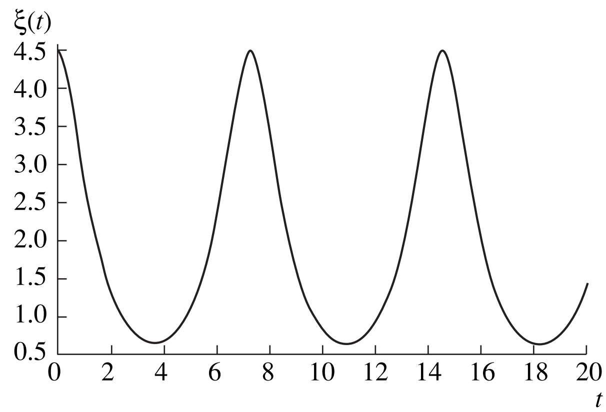
Фиг. 4.



Фиг. 5.

На фиг. 4 приведена кривая зависимости частоты колебаний $\omega = 2\pi/T$ от параметра A, полученная численным решением уравнения (17).

На фиг. 5 и фиг. 6 представлены графики изменения величины $\xi(t)$ со временем для двух значений A. Кривая на фиг. 5 соответствует изменению величины $\xi(t)$ для случая, когда $A = 0.256$. Кривая на фиг. 6 соответствует случаю $A = 0.112$. Из графиков видно, что с уменьшением A форма кривых все сильнее отклоняется от синусоидальной. Кроме того, видно, что система в окрестности минимума ξ проводит большее количество времени, чем в окрестности максимума, и эта тенденция усиливается с уменьшением A.



Фиг. 6.

6. ЗАКЛЮЧЕНИЕ

Уравнения Лотки–Вольтерра обладают хорошо известной гамильтоновой структурой (см. [2]). Она позволяет изучить поведение траекторий в фазовой плоскости и восстановить динамику в квадратурах. К сожалению, на практике для исследования движения по найденным фазовым траекториям эта гамильтонова структура не очень удобна, поскольку не позволяет выразить производную одной из исследуемых динамических переменных по времени через нее саму в элементарных функциях. Попытка найти другие скобки Пуассона и другой гамильтониан в данном случае не приведет к успеху, поскольку она будет приводить к такому же первому интегралу системы. Построить каноническую замену переменных также трудно по той причине, что задача нахождения производящей функции канонических преобразований является очень сложной. Тем не менее существует подход, позволяющий получить приемлемые выражения для исследования динамики рассматриваемого частного случая системы Лотки–Вольтерра. Он основан на идее исследования движения по каждой из выбранных фазовых траекторий с помощью своего гамильтониана и на том факте, что разные гамильтоновы системы могут иметь совпадающие фазовые траектории, причем движение по ним может происходить одинаково. Несложная замена переменных и использование известного первого интеграла – гамильтониана позволяет привести уравнение движения к формально гамильтоновой форме с той разницей, что гамильтониан зависит от выбранной фазовой траектории. Таким образом, каждой фазовой траектории сопоставляется гамильтонова система, имеющая, вообще говоря, другую структуру фазовой плоскости, чем исходная, но данная фазовая траектория у них общая и характер движения по ней совпадает. Замечательно то, что каждый из полученного набора гамильтонианов квадратичен по обобщенному импульсу. Гамильтоновы системы, квадратичные по обобщенному импульсу, хорошо изучены и широко используются в классической механике. Квадратичность по обобщенному импульсу позволяет развить для исследования динамики данной системы полезную аналогию с колебаниями массы на нелинейной пружине. Период колебаний и связь между координатой точечной массы на пружине и временем определяется обычным образом (см. [3]). Преимущества, которые дает при исследовании динамики использование квадратичных по обобщенному импульсу гамильтонианов, получены в работе без обращения к сложной задаче построения производящей функции канонических преобразований.

СПИСОК ЛИТЕРАТУРЫ

1. *Алексеев В.В.* Биофизика сообществ живых организмов // Успехи физ. наук. 1976. Т. 120. Вып. 4. С. 647–675.
2. *Вольтерра В.* Математическая теория борьбы за существование. М.: Наука, 1976.
3. *Ландау Л.Д., Лифшиц Е.М.* Механика. М.: Наука, 1988.
4. *Найфе А.* Методы возмущений. М.: Мир, 1976.

ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 2007, том 47, № 8, с. 1350–1355

УДК 519.624.1

О ВЫЧИСЛЕНИИ СОБСТВЕННЫХ ЗНАЧЕНИЙ ЗАДАЧИ ШТУРМА–ЛИУВИЛЛЯ С ФРАКТАЛЬНЫМ ИНДЕФИНИТНЫМ ВЕСОМ[1)]

© 2007 г. А. А. Владимиров

(119992 Москва, Ленинские горы, МГУ, механико-матем. ф-т)

e-mail: vladimi@mech.math.msu.su

Поступила в редакцию 17.11.2006 г.
Переработанный вариант 26.02.2007 г.

Предлагается эффективный метод вычисления собственных значений граничной задачи

$$-y'' - \lambda\rho y = 0, \quad y(0) = y(1) = 0,$$

где $\rho \in \overset{\circ}{W}{}_2^{-1}[0, 1]$ – обобщенная производная некоторой самоподобной функции $P \in L_2[0, 1]$. Библ. 6.

Ключевые слова: задача Штурма–Лиувилля, метод вычисления собственных значений, фрактальная индефинитная весовая функция.

1. ВВЕДЕНИЕ

1. Рассмотрим граничную задачу

$$-y'' - \lambda\rho y = 0, \tag{1}$$

$$y(0) = y(1) = 0, \tag{2}$$

где $\lambda \in \mathbb{C}$ – спектральный параметр. В случае когда вес ρ представляет собой равномерно положительную непрерывную функцию, эта задача является классической. Однако она представляет интерес и в той существенно менее изученной ситуации, когда в роли веса ρ выступает обобщенная функция класса $\overset{\circ}{W}{}_2^{-1}[0, 1]$. Для случая неотрицательности весовой функции операторные модели такой сингулярной задачи указывались, например, в [1], [2]. Более общая ситуация, связанная с допущением веса произвольного знака, рассматривалась в [3], [4].

Известно (см., например, [1, ф-ла (11.7)]), что в классическом случае асимптотика считающей функции $\mathcal{N}$ собственных значений задачи (1), (2) при $\lambda \longrightarrow +\infty$ имеет вид

$$\mathcal{N}(\lambda) \sim \left(\frac{1}{\pi}\int_0^1 \sqrt{\rho}\, d\mu\right)\lambda^{1/2},$$

где $d\mu$ – линейная мера Лебега. Сингулярные задачи не допускают такого единообразного описания. Поэтому для более детального исследования необходимо вводить различные ограничения на вид сингулярного веса. Наиболее изученной является ситуация, когда вес $\rho \in \overset{\circ}{W}{}_2^{-1}[0, 1]$ представляет собой обобщенную (в смысле Соболева) производную некоторой самоподобной (фрактальной) функции $P \in L_2[0, 1]$. Именно такая ситуация и рассматривалась в [2]–[4] (обсуждение приложений полученных при этом результатов в теории случайных процессов может быть найдено в [5]). Подробное описание класса квадратично суммируемых самоподобных функций будет дано в п. 3 разд. 2.

Полученные на указанном пути результаты могут быть кратко охарактеризованы следующим образом. Каждой самоподобной функции $P \in L_2[0, 1]$ ставится в соответствие неотрицательное число $D < 2$, называемое *спектральным порядком* этой функции (см. [3, § 3]). В случае вы-

[1)] Работа выполнена при финансовой поддержке РФФИ (код проекта 07-01-00283), гранта поддержки ведущих научных школ НШ-5247.2006.1 и INTAS, грант № 05-1000008-7883.

полнения неравенства $D > 0$ асимптотика считающей функции положительных собственных значений задачи (1), (2) с весом $\rho \rightleftharpoons P'$ имеет при $\lambda \longrightarrow +\infty$ вид

$$\mathcal{N}(\lambda) \sim s(\ln\lambda)\lambda^{D/2}, \tag{3}$$

где s – некоторая периодическая непрерывная функция (зависящая от выбора функции P).

2. В настоящее время, по всей видимости, неизвестен способ, который позволял бы охарактеризовывать поведение фигурирующей в асимптотике (3) функции s на основе общих соображений. В частности, мы не располагаем критерием, позволяющим определять, для каких весов ρ функция s вырождается в константу. Между тем сколь угодно точные оценки функций s для различных конкретных весов ρ могут быть установлены экспериментально на основе численного определения собственных значений задачи (1), (2). Примеры применения такого подхода были приведены в [3, § 5]. В частности, там было отмечено, что если вес представляет собой обобщенную производную известной канторовой лестницы, то функция s заведомо не является постоянной. Ниже будет изложен вычислительный метод, применение которого позволяет сделать указанные теоретические выводы.

2. СВЕДЕНИЕ ЗАДАЧИ К КОНЕЧНОМЕРНОЙ

1. Как и в [3], [4], под *индексом инерции* $\mathrm{ind}F$ эрмитова оператора F, действующего в некотором гильбертовом пространстве $\mathfrak{H}$, будем понимать точную верхнюю грань размерностей конечномерных подпространств $\mathfrak{M} \subseteq \mathfrak{H}$, удовлетворяющих условию

$$\exists\varepsilon > 0 \quad \forall x \in \mathfrak{M} : \langle Fx, x\rangle \le -\varepsilon\|x\|^2_{\mathfrak{H}}. \tag{4}$$

Имеет место следующее

Утверждение 1. *Пусть $\mathfrak{H}_1$ и $\mathfrak{H}_2$ – два гильбертовых пространства, а $\mathfrak{H}$ – прямая сумма $\mathfrak{H}_1 \oplus \mathfrak{H}_2$ пространств $\mathfrak{H}_1$ и $\mathfrak{H}_2$. Пусть также $F : \mathfrak{H} \longrightarrow \mathfrak{H}$ – ограниченный эрмитов оператор с блочно-матричным представлением*

$$F = \begin{pmatrix} A & B^* \\ B & C \end{pmatrix},$$

в котором оператор $C : \mathfrak{H}_2 \longrightarrow \mathfrak{H}_2$ равномерно положителен. Тогда выполняется равенство

$$\mathrm{ind}F = \mathrm{ind}(A - B^*C^{-1}B). \tag{5}$$

Доказательство. Ввиду ограниченной обратимости оператора C, для оператора F можно рассмотреть факторизацию Фробениуса–Шура $F = H^*F_0H$, где через H и F_0 обозначены операторы с блочно-матричными представлениями:

$$H \rightleftharpoons \begin{pmatrix} 1 & 0 \\ C^{-1}B & 1 \end{pmatrix}, \quad F_0 \rightleftharpoons \begin{pmatrix} A - B^*C^{-1}B & 0 \\ 0 & C \end{pmatrix}.$$

При этом оператор H обладает ограниченным обратным вида

$$H^{-1} = \begin{pmatrix} 1 & 0 \\ -C^{-1}B & 1 \end{pmatrix},$$

а для любого вектора $x \in \mathfrak{H}$ выполняются равенства

$$\langle Fx, x\rangle = \langle H^*F_0Hx, x\rangle = \langle F_0(Hx), (Hx)\rangle.$$

Тогда для любого конечномерного подпространства $\mathfrak{M} \subseteq \mathfrak{H}$, удовлетворяющего условию (4), подпространство $\mathfrak{M}_0 \rightleftharpoons H\mathfrak{M}$ имеет ту же размерность и удовлетворяет условию

$$\exists\varepsilon > 0 \quad \forall x \in \mathfrak{M}_0 : \langle F_0x, x\rangle \le -\varepsilon\|x\|^2_{\mathfrak{H}}, \tag{6}$$

а для любого конечномерного подпространства $\mathfrak{M}_0 \subseteq \mathfrak{H}$, удовлетворяющего условию (6), подпространство $\mathfrak{M} \rightleftharpoons H^{-1}\mathfrak{M}_0$ имеет ту же размерность и удовлетворяет условию (4). Таким образом, выполняется равенство $\operatorname{ind}F = \operatorname{ind}F_0$. Ввиду положительности оператора C это означает выполнение равенства (5).

2. В дальнейшем через $\mathfrak{H}$ будем обозначать пространство Соболева $\overset{\circ}{W}{}_2^1[0, 1]$, снабженное скалярным произведением

$$\langle y, z\rangle \rightleftharpoons \int_0^1 y'\bar{z}'d\mu.$$

Простым следствием теоремы 4.1 из [3] является

Утверждение 2. *Пусть F – пучок действующих в пространстве $\mathfrak{H}$ линейных операторов, удовлетворяющий тождеству*

$$\langle F(\lambda)y, y\rangle = \int_0^1 (|y'|^2 + \lambda P(|y|^2)')d\mu \quad \forall\lambda \in \mathbb{R} \quad \forall y \in \mathfrak{H}. \tag{7}$$

Пусть также $\{\nu_n\}_{n=1}^{\infty}$ – последовательность (возможно, частичная) сосчитанных в порядке возрастания положительных собственных значений граничной задачи (1), (2). *Тогда для любых натурального числа $n \geq 1$ и вещественного числа $\lambda \in (0, \nu_n)$ выполняется неравенство* $\operatorname{ind}F(\lambda) < n$, *а для любых натурального числа $n \geq 1$ и вещественного числа $\lambda \in (\nu_n, +\infty)$ выполняется неравенство* $\operatorname{ind}F(\lambda) \geq n$.

Таким образом, при наличии в нашем распоряжении метода вычисления достаточно точных оценок индексов инерции операторов из пучка F мы можем вычислять оценки собственных значений граничной задачи (1), (2) на основе метода деления отрезка.

3. Пусть теперь S – набор из натурального числа $N > 1$ и вещественных чисел $a_k > 0$, d_k и β_k, где $k = 1, 2, \ldots, N$, удовлетворяющих соотношениям

$$\sum_{k=1}^{N} a_k = 1, \quad \theta_S \rightleftharpoons \sqrt{\sum_{k=1}^{N} a|d_k|^2} < 1.$$

Как и в [3], с набором S связывается семейство $\{G_k\}_{k=1}^{N}$ действующих в пространстве $L_2[0, 1]$ линейных операторов вида

$$G_k\chi_{(\zeta, \xi)} = \chi_{(\alpha_k + a_k\zeta, \alpha_k + a_k\xi)} \quad \forall(\zeta, \xi) \subseteq [0, 1],$$

где через $\{\alpha_k\}_{k=1}^{N+1}$ обозначен набор чисел вида $\alpha_1 = 0$, $\alpha_{k+1} = \alpha_k + a_k$, а через χ_I – индикаторы промежутков I. При этом на основе семейства операторов $\{G_k\}_{k=1}^{N}$ конструируется нелинейный, вообще говоря, оператор G_S вида

$$G_S f \rightleftharpoons \sum_{k=1}^{N} (d_k G_k f + \beta_k \chi_{(\alpha_k, \alpha_{k+1})}) \quad \forall f \in L_2[0, 1]. \tag{8}$$

Этот оператор является сжимающим, и введенная ранее величина θ_S представляет собой его коэффициент сжатия (см. [3, лемма 3.1]).

Сжимающие операторы, допускающие представление в виде (8), мы называем *операторами подобия*. Функции $f \in L_2[0, 1]$, являющиеся неподвижными точками операторов подобия, называются *самоподобными*. Набор чисел S, которому отвечает оператор подобия G_S, оставляющий неподвижной некоторую заранее фиксированную функцию $f \in L_2[0, 1]$, называются набором *параметров самоподобия* функции f.

4. С произвольно фиксированным набором S параметров самоподобия функции P мы в дальнейшем будем связывать конечномерное подпространство $\mathfrak{H}_{S,1}$ пространства $\mathfrak{H}$, обладающее базисом $\{y_{S,k}\}_{k=1}^{N-1}$ вида

$$y_{S,k}(x) = \begin{cases} (x-\alpha_k)/a_k & \text{при } x \in [\alpha_k, \alpha_{k+1}], \\ (\alpha_{k+2}-x)/a_{k+1} & \text{при } x \in [\alpha_{k+1}, \alpha_{k+2}], \\ 0 & \text{иначе}. \end{cases}$$

При этом через $\mathfrak{H}_{S,2}$ будет обозначаться ортогональное дополнение $\mathfrak{H} \ominus \mathfrak{H}_{S,1}$ подпространства $\mathfrak{H}_{S,1}$. Нетрудно видеть, что справедливо следующее

Утверждение 3. *Подпространство* $\mathfrak{H}_{S,2}$ *имеет вид* $\mathfrak{H}_{S,2} = \{y \in \mathfrak{H} \mid \forall k = 2, 3, \ldots, N : y(\alpha_k) = 0\}$.

Введем теперь в рассмотрение три пучка A_S, B_S и C_S линейных операторов, значениями которых $A_S(\lambda) : \mathfrak{H}_{S,1} \longrightarrow \mathfrak{H}_{S,1}$, $B_S(\lambda) : \mathfrak{H}_{S,1} \longrightarrow \mathfrak{H}_{S,2}$ и $C_S(\lambda) : \mathfrak{H}_{S,2} \longrightarrow \mathfrak{H}_{S,2}$ при произвольно фиксированном $\lambda \in \mathbb{R}$ являются элементы блочно-матричного представления

$$F(\lambda) = \begin{pmatrix} A_S(\lambda) & B_S^*(\lambda) \\ B_S(\lambda) & C_S(\lambda) \end{pmatrix},$$

определенного условием (7) оператора $F(\lambda)$. Из утверждения 1 следует

Утверждение 4. *Пусть даны два вещественных числа* $\lambda > 0$ *и* $\varepsilon > 0$, *удовлетворяющих неравенствам* $\varepsilon \geq 2\|B_S(\lambda)\|^2$ *и* $\|C_S(\lambda) - 1\| < 1/2$. *Тогда выполняются неравенства* $\mathrm{ind}A_S(\lambda) \leq \mathrm{ind}F(\lambda)$ *и* $\mathrm{ind}F(\lambda) \leq \mathrm{ind}(A_S(\lambda) - \varepsilon)$.

Утверждение 4 позволяет свести задачу вычисления оценок индекса инерции оператора $F(\lambda)$ к допускающей непосредственное решение на ЭВМ задаче вычисления оценок индексов инерции конечномерных операторов $A_S(\lambda)$ и $A_S(\lambda) - \varepsilon$. Однако при этом встает вопрос об области применимости утверждения 4 и о степени точности получаемых на его основе оценок величины $\mathrm{ind}F(\lambda)$. Изучению этого вопроса будет посвящен следующий раздел.

3. ИССЛЕДОВАНИЕ ТОЧНОСТИ КОНЕЧНОМЕРНЫХ ПРИБЛИЖЕНИЙ

1. Имеет место следующее

Утверждение 5. *Пусть даны набор* S *параметров самоподобия функции* P *и вещественное число* $\lambda > 0$. *Тогда выполняются неравенства* $\|B_S(\lambda)\| \leq \lambda\theta_S\|P\|_{L_2[0,1]}$ *и* $\|C_S(\lambda) - 1\| \leq \lambda\theta_S\|P\|_{L_2[0,1]}$.

Доказательство. Заметим, что полуторалинейные формы операторов $B_S(\lambda)$ и $C_S(\lambda) - 1$ удовлетворяют тождествам

$$\langle B_S(\lambda)y, z\rangle = \lambda\int_0^1 P(y\bar{z})'d\mu \quad \forall y \in \mathfrak{H}_{S,1}, \quad \forall z \in \mathfrak{H}_{S,2},$$

$$\langle (C_S(\lambda) - 1)y, z\rangle = \lambda\int_0^1 P(y\bar{z})'d\mu \quad \forall y, z \in \mathfrak{H}_{S,2}.$$

При этом функция P в правых частях выписанных тождеств может быть заменена функцией $\hat{P}_S \in L_2[0, 1]$ вида $\hat{P}_S \rightleftharpoons \Sigma_{k=1}^N d_k G_k P$, удовлетворяющей очевидному равенству $\|\hat{P}_S\|_{L_2[0,1]} = \theta_S\|P\|_{L_2[0,1]}$. Заметим также, что для любых двух функций $y, z \in \mathfrak{H}$ выполняются соотношения

$$\|(\bar{y}z)'\|_{L_2[0,1]} \leq \|\bar{y}z'\|_{L_2[0,1]} + \|\bar{y}'z\|_{L_2[0,1]} \leq \|y\|_{C[0,1]}\|z'\|_{L_2[0,1]} + \|z\|_{C[0,1]}\|y'\|_{L_2[0,1]} \leq$$
$$\leq \frac{\|y'\|_{L_2[0,1]}}{2}\|z'\|_{L_2[0,1]} + \frac{\|z'\|_{L_2[0,1]}}{2}\|y'\|_{L_2[0,1]} = \|y\|_{\mathfrak{H}}\|z\|_{\mathfrak{H}}.$$

С учетом сделанных замечаний доказываемое утверждение тривиальным образом выводится из неравенства Коши–Буняковского.

2. Имеет место следующее

Утверждение 6. *Пусть даны набор S параметров самоподобия функции P и натуральное число* $m > 1$. *Тогда оператор* G_S^m *является оператором подобия с неподвижной точкой P*.

Утверждение 7. *Пусть дано вещественное число* $\varepsilon > 0$. *Тогда существует набор S параметров самоподобия функции P, удовлетворяющий неравенству* $\theta_S < \varepsilon$.

Утверждение 8. *Пусть даны набор S параметров самоподобия функции P и два вещественных числа* $\lambda > 0$ *и* $\varepsilon > 0$, *удовлетворяющих неравенствам* $\varepsilon \ge 2\lambda^2\theta_S^2\|P\|^2_{L_2[0,1]}$ *и* $\lambda\theta_S\|P\|_{L_2[0,1]} < 1/2$. *Тогда выполняются неравенства* $\operatorname{ind}A_S(\lambda) \le \operatorname{ind}F(\lambda)$ *и* $\operatorname{ind}F(\lambda) \le \operatorname{ind}(A_S(\lambda) - \varepsilon)$.

Утверждение 7 следует из утверждения 6. .Утверждение 8 следует из утверждений 4 и 5.

Утверждения 7 и 8 указывают способ вычисления оценок индекса инерции оператора $F(\lambda)$ при произвольно фиксированном значении $\lambda > 0$.

3. Пусть теперь ν_n – имеющее произвольно фиксированный номер $n \ge 1$ положительное собственное значение задачи (1), (2).

Заметим, что при любом значении $\lambda \in (0, \nu_n)$ существует вещественное число $\delta \in (0, 1)$, удовлетворяющее неравенству $\lambda < (1 - \delta)\nu_n$. В таком случае для любых набора S параметров самоподобия функции P и вещественного числа $\varepsilon \in (0, \delta]$, удовлетворяющих условиям утверждения 8, выполняются соотношения

$$\operatorname{ind}(A_S(\lambda) - \varepsilon) \le \operatorname{ind}(A_S(\lambda) - \delta) \le \operatorname{ind}(F(\lambda) - \delta) = \operatorname{ind}F\left(\frac{\lambda}{1-\delta}\right) < n.$$

Из утверждений 7 и 5 следует, что при любом значении $\lambda \in (\nu_n, +\infty)$ оператор $F(\lambda) - 1$ является вполне непрерывным. Поэтому существует вещественное число $\delta \in (0, 1/2)$, удовлетворяющее равенству $\operatorname{ind}(F(\lambda) + \delta) = \operatorname{ind}F(\lambda)$ (см. [6, п. 95]). В таком случае для любого набора S параметров самоподобия функции P, удовлетворяющего неравенству $2\lambda^2\theta_S^2\|P\|^2_{L_2[0,1]} \le \delta$, выполняются соотношения

$$\operatorname{ind}A_S(\lambda) \ge \operatorname{ind}(A_S(\lambda) + \delta - B_S(\lambda)^*C_S^{-1}(\lambda)B_S(\lambda)) \ge \operatorname{ind}(F(\lambda) + \delta) = \operatorname{ind}F(\lambda) \ge n.$$

Таким образом, получаемые на основе утверждений 7 и 8 оценки величин $\operatorname{ind}F(\lambda)$ позволяют для любых двух вещественных чисел $\lambda_1 > 0$ и $\lambda_2 > \lambda_1$ установить верность одного из неравенств $\lambda_1 \le \nu_n$ или $\lambda_2 \ge \nu_n$. Иначе говоря, они позволяют находить при помощи метода деления отрезка сколь угодно точные оценки собственного значения ν_n.

4. НЕПОСРЕДСТВЕННОЕ ВЫЧИСЛЕНИЕ ОЦЕНОК

1. Использование условия самоподобия функции P позволяет выписывать рекуррентные формулы для ее степенных моментов

$$\mathbf{P}_q \rightleftharpoons \int_0^1 Px^q d\mu, \quad q = 0, 1, 2, \ldots.$$

В частности, употребляемые в дальнейшем моменты нулевой и первой степеней имеют вид

$$\mathbf{P}_0 = \frac{\sum_{k=1}^N a_k\beta_k}{1 - \sum_{k=1}^N a_k d_k}, \quad \mathbf{P}_1 = \frac{\sum_{k=1}^N a_k\left(\frac{a_k\beta_k}{2} + \alpha_k d_k\mathbf{P}_0 + \alpha_k\beta_k\right)}{1 - \sum_{k=1}^N a_k^2 d_k}.$$

2. Прямым просчетом устанавливается

Утверждение 9. *Пусть даны набор S параметров самоподобия функции P и два вещественных числа* $\lambda > 0$ *и* $\varepsilon \ge 0$. *Тогда матрица квадратичной формы оператора* $A_S(\lambda) - \varepsilon$ *в базисе* $\{y_{S,k}\}_{k=1}^{N-1}$ *является трехдиагональной и имеет элементы*

$$\langle (A_S(\lambda) - \varepsilon)y_{S,k}, y_{S,k}\rangle = (1 - \varepsilon)[a_k^{-1} + a_{k+1}^{-1}] + \lambda[2d_{k+1}(\mathbf{P}_1 - \mathbf{P}_0) + 2d_k\mathbf{P}_1 + \beta_k - \beta_{k+1}],$$

$$\langle (A_S(\lambda) - \varepsilon) y_{S,k}, y_{S,k-1} \rangle = -(1-\varepsilon) a_k^{-1} - \lambda d_k (2\mathbf{P}_1 - \mathbf{P}_0).$$

Сигнатура матрицы из утверждения 9 может теперь быть вычислена различными способами, например, как число перемен знака в ряде главных миноров этой матрицы.

3. Машинная программа на языке РЕФАЛ, вычисляющая оценки положительных собственных значений задачи (1), (2) на основе изложенной схемы, может быть найдена по адресу http://www.math.msu.su/labs/spectrallab/soft/Dirichlet.ref.

СПИСОК ЛИТЕРАТУРЫ

1. *Кац И.С., Крейн М.Г.* О спектральных функциях струны / В кн.: Ф. Аткинсон. Дискретные и непрерывные граничные задачи. М.: Мир, 1968. С. 613–733.
2. *Solomyak M., Verbitsky E.* On a spectral problem related to self-similar measures // Bull. London Math. Soc. 1995. V. 27. P. 242–248.
3. *Владимиров А.А., Шейпак И.А.* Самоподобные функции в пространстве $L_2[0, 1]$ и задача Штурма–Лиувилля с сингулярным индефинитным весом // Матем. сб. 2006. Т. 197. № 11. С. 13–30.
4. *Владимиров А.А., Шейпак И.А.* Индефинитная задача Штурма–Лиувилля для некоторых классов самоподобных весов // Тр. Матем. ин-та РАН. М., 2006. Т. 255. С. 88–98.
5. *Назаров А.И.* Логарифмическая асимптотика малых уклонений для некоторых гауссовских процессов в L_2-норме относительно самоподобной меры // Зап. научн. семинара ПОМИ РАН. СПб., 2004. Т. 311. С. 190–213.
6. *Рисс Ф., Сёкефальви-Надь Б.* Лекции по функциональному анализу. М.: Мир, 1979.

ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 2007, том 47, № 8, с. 1356–1364

УДК 519.633.8

О ФОРМИРОВАНИИ РЕЗКИХ ПЕРЕХОДНЫХ СЛОЕВ В ДВУМЕРНЫХ МОДЕЛЯХ РЕАКЦИЯ–ДИФФУЗИЯ[1)]

© 2007 г. В. Т. Волков, Н. Е. Грачёв, Н. Н. Нефёдов, А. Н. Николаев

(111992 Москва, Ленинские горы, МГУ, физ. ф-т)

e-mail: volkov@phys.msu.ru; grachev_nick@mail.ru

Поступила в редакцию 09.01.2007 г.

Исследован вопрос о том, как в сингулярно возмущенном параболическом уравнении, рассматриваемом в пространственно-двумерном случае, из начальной функции достаточно общего вида формируется решение с резким переходным слоем. На основе асимптотического анализа получены оценки времени формирования контрастной структуры. Приведены также результаты численного эксперимента. Библ. 6. Фиг. 5.

Ключевые слова: двумерные модели реакция–диффузия, асимптотический метод решения, образование контрастных структур в решениях, численное исследование.

ВВЕДЕНИЕ

Рассмотрим нелинейное параболическое уравнение типа реакция–диффузия

$$\varepsilon^2(\Delta u - u_t) = f(u, x, y, t). \tag{1}$$

Уравнения и системы уравнений такого вида описывают различные физические, биологические и химические объекты и используются в качестве математических моделей процессов фазового разделения, в задачах массо- и теплопереноса, в задачах химической кинетики и других. Физическая природа параметра ε может быть разнообразна, однако характерной особенностью подобных задач часто является наличие в системе процессов с сильно отличающимися характерными временными масштабами, например малая диффузия или быстрая реакция между компонентами, что позволяет во многих случаях считать ε малым параметром. В этом приближении уравнение (1) рассматривалось многими авторами (см. [1]–[4]). В [5] был изучен вопрос о формировании решений в виде контрастной структуры для (1) в одномерном случае.

Одним из примеров может служить явление фазового разделения, для описания которого используется уравнение

$$\alpha\varepsilon^2\frac{\partial\varphi}{\partial t} = \varepsilon^2\Delta\varphi - \frac{\partial F}{\partial\varphi},$$

которое носит название уравнения Аллена – Кана (см. [6]). Здесь α – постоянная, φ – параметр порядка – функция, значения которой заключены между 0 и 1, причем $\varphi = 0$ соответствует неупорядоченной (жидкой) фазе, а $\varphi = 1$ – упорядоченной (твердой). Величина F – известная функция переменной φ, называемая плотностью свободной энергии. Для задач фазового разделения типичным является наличие двух минимумов у функции плотности свободной энергии, что как раз и означает возможность устойчивого сосуществования двух фаз.

В настоящей работе изучается вопрос о том, как из начальной функции достаточно общего вида, заданной для уравнения (1), с течением времени формируется контрастная структура, т.е. решение с резким внутренним переходным слоем.

1. ПОСТАНОВКА ЗАДАЧИ

Рассмотрим уравнение (1) в двумерной области D с гладкой границей Γ, начальным условием

$$u(x, y, 0, \varepsilon) = u^0(x, y) \text{ в } \overline{D} \tag{2}$$

1) Работа выполнена при частичной финансовой поддержке РФФИ (код проекта 05-01-00465).

и с краевым условием II рода (условие непроницаемости границы)

$$\frac{\partial u(x, y, t, \varepsilon)}{\partial n} = 0, \quad t > 0, \quad (x, y) \in \Gamma, \tag{3}$$

где производная берется по внешней нормали к границе.

Потребуем, чтобы (2) и (3) удовлетворяли условию согласования

$$\frac{\partial u^0(x, y)}{\partial n} = 0, \quad (x, y) \in \Gamma. \tag{3'}$$

Функции $f(u, x, y, t)$ и $u^0(x, y)$ далее предполагаются достаточно гладкими.

Если в уравнении (1) положить $\varepsilon = 0$, то получим вырожденное уравнение

$$f(u, x, y, t) = 0, \quad (x, y) \in \bar{D}, \quad t \geq 0. \tag{4}$$

Предположим, что функция $f(u, x, y, t)$ имеет нелинейность кубического типа, что характерно для многих прикладных задач, т.е. пусть выполнены следующие условия.

Условие A_1. Существуют такие функции $\underline{u}$, $\bar{u}$, что уравнение (4) имеет в области $G = \{\underline{u} \leq u \leq \bar{u}, (x, y) \in \bar{D}, t > 0\}$ три корня относительно u : $u = \varphi_i(x, y, t)$, $i = 0, 1, 2$, причем $\underline{u} < \varphi_1(x, y, t) < \varphi_0(x, y, t) < \varphi_2(x, y, t) < \bar{u}$ при $(x, y, 0) \in \bar{D}$, $t > 0$, $f_u(\varphi_i(x, y, t), x, y, t) > 0$ для $i = 1, 2$ и $(x, y) \in \bar{D}$, $t > 0$;

Условие A_2. Существует гладкая замкнутая кривая $h_0 \in D$ такая, что

$$\underline{u} < u^0(x, y) < \varphi_0(x, y, 0), \quad (x, y) \in D_e \cup \Gamma.$$

$$u^0(x, y) = \varphi_0(x, y, 0), \quad (x, y) \in h_0$$

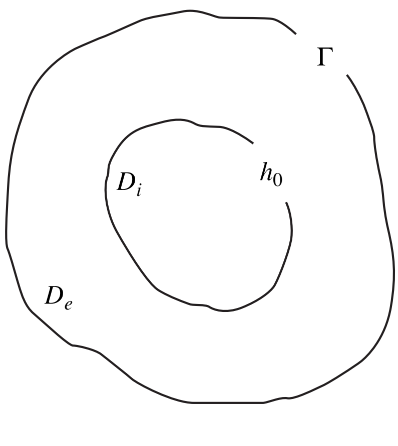
$$\varphi_0(x, y, 0) < u^0(x, y) < \bar{u}, \quad (x, y) \in D_i,$$

причем $D_e \cup D_i \cup h_0 = D$ (см. фиг. 1).

В дальнейшем будет доказано, что именно в окрестности кривой h_0 наблюдается формирование резкого переходного слоя, причем образование контрастной структуры происходит за короткий промежуток времени порядка $O(\varepsilon^2|\ln\varepsilon|)$. Если таких линий несколько, то появится несколько переходных слоев, а формирование каждого из них можно описывать так же, как формирование слоя в окрестности кривой h_0.

2. АСИМПТОТИКА РЕШЕНИЯ НА ВРЕМЕННО́М ПРОМЕЖУТКЕ $0 \leq t \leq A\varepsilon^2|\ln\varepsilon|$

Отметим, что вопрос о существовании решения задачи (1)–(3) при условиях A_1 и A_2 решается довольно просто. Постоянные функции $\underline{u}$ и $\bar{u}$ являются, соответственно, нижним и верхним решениями задачи (1)–(3), и, следовательно, существует единственное ее решение $u(x, y, t, \varepsilon)$, при-



Фиг. 1.

чем выполняются неравенства

$$\underline{u} < u(x, y, t, \varepsilon) < \bar{u} \quad \text{при} \quad (x, y, t) \in \bar{D}.$$

Для построения асимптотики указанного решения на временно́м промежутке $0 \le t \le A\varepsilon^2|\ln\varepsilon|$ (число A уточним ниже) сделаем замену переменных $t = \varepsilon^2\tau$. Тогда задача (1)–(3) примет вид

$$\varepsilon^2\Delta\bar{u} - \bar{u}_\tau = f(\bar{u}, x, y, \varepsilon^2\tau), \quad (x, y) \in D, \quad \tau > 0, \tag{5}$$

$$\bar{u}(x, y, 0, \varepsilon) = u^0(x, y), \quad (x, y) \in D, \tag{6}$$

$$\frac{\partial\bar{u}(x, y, \tau, \varepsilon)}{\partial n} = 0, \quad (x, y) \in \Gamma, \quad \tau > 0, \tag{7}$$

где введено обозначение $\bar{u}(x, y, \tau, \varepsilon) = u(x, y, \varepsilon^2\tau, \varepsilon)$.

При $\varepsilon = 0$ из (5), (6) получим

$$\begin{aligned} -\tilde{u}_\tau &= f(\tilde{u}, x, y, 0), \quad \tau > 0, \\ \tilde{u}(x, y, 0) &= u^0(x, y), \quad (x, y) \in \bar{D}, \end{aligned} \tag{8}$$

здесь переменные x и y входят как параметры.

Из условий A_1 и A_2 следует, что для любой точки $(x, y) \in \bar{D}$ решение $\tilde{u}(x, y, \tau)$ задачи (8) является монотонной функцией τ, причем

$$\lim_{\tau\to\infty}\tilde{u}(x, y, \tau) = \begin{cases} \varphi_1(x, y, 0), & (x, y) \in D_i, \\ \varphi_2(x, y, 0), & (x, y) \in D_e, \end{cases}$$
$$\tilde{u}(x, y, \tau) = \varphi_0(x, y, 0), \quad (x, y) \in h_0, \quad \tau \ge 0. \tag{9}$$

Введем теперь в окрестности кривой h_0 следующую локальную систему координат (r, s): переменная r – это расстояние от точки M вблизи переходного слоя до точки N на линии h_0, вычисленное по нормали к h_0, выходящей из N, попадающей в M и направленной внутрь области D, а переменная s – это координата, описывающая положение точки N на кривой h_0.

Возьмем любое достаточно малое положительное число δ. Тогда, в силу (9), для любого числа $\eta > 0$ найдется такое $\tau_\delta = \tau_\delta(\eta)$, что при всех $\tau = \tau_\delta(\eta)$ будут выполнены неравенства

$$|\tilde{u}(x, y, \tau) - \varphi_1(x, y, 0)| \le \eta \tag{10а}$$

при $(x, y) \in \bar{D}^-$, где D^- – область, ограниченная в выбранной локальной системе координат линией $r = -\delta$ (обозначим ее через h_0^-) и границей Γ, а также

$$|\tilde{u}(x, y, \tau) - \varphi_2(x, y, 0)| \le \eta \tag{10б}$$

при $(x, y) \in \bar{D}^+$, где D^+ – область, ограниченная в выбранной локальной системе координат линией $r = \delta$ (обозначим ее через h_0^+).

Далее положим $D_\delta = \bar{D}^+ \cup \bar{D}^-$ (см. фиг. 2). Пусть

$$\min\{\min_{(x, y) \in D_\delta} f_u(\varphi_1(x, y, 0), x, y, 0), \min_{(x, y) \in D_\delta} f_u(\varphi_2(x, y, 0), x, y, 0)\} = 2m. \tag{11}$$

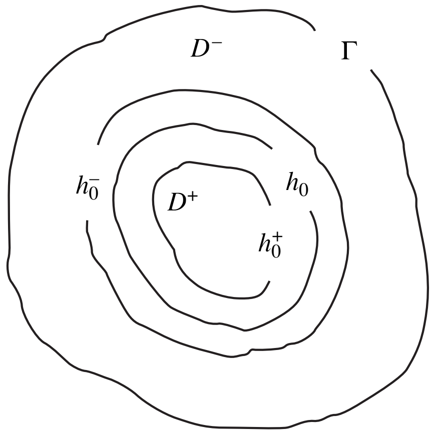
Из условия A_1 ясно, что $m > 0$. Поэтому существует число $\eta_0 > 0$ такое, что для любой точки $(x, y) \in \bar{D}_\delta$ и всех u, удовлетворяющих неравенствам

$$|u - \varphi_1(x, y, 0)| \le \eta_0 \quad \text{и} \quad |u - \varphi_2(x, y, 0)| \le \eta_0, \tag{12}$$

выполняется оценка

$$f_u(u, x, y, 0) \ge m.$$

Отсюда следует, что предельный переход в (9) имеет экспоненциальный характер. Более точно, справедлива



Фиг. 2.

Лемма 1. *Пусть* m *и* η_0 – *числа, определенные в* (11) *и* (12). *Тогда для любого достаточно малого положительного числа* δ *существует такое* $\tau_\delta > 0$, *что выполняются условия*

$$|\tilde{u}(x, y, \tau) - \varphi_1(x, y, 0)| \le \eta_0 e^{-m(\tau-\tau_\delta)}, \quad (x, y) \in \overline{D}^-, \quad \tau \ge \tau_\delta, \tag{13}$$

$$|\tilde{u}(x, y, \tau) - \varphi_2(x, y, 0)| \le \eta_0 e^{-m(\tau-\tau_\delta)}, \quad (x, y) \in \overline{D}^+, \quad \tau \ge \tau_\delta. \tag{14}$$

Доказательство. Для любого $\delta > 0$ выберем такое $\tau_\delta = \tau_\delta(\eta)$, для которого выполнены неравенства (10) при $\eta = \eta_0$, и представим правую часть уравнения (8) при $(x, y) \in \overline{D}^-$ в виде $f(\tilde{u}(x, y, \tau), x, y, 0) = f(\varphi_1(x, y, 0), x, y, 0) + f_u^*(\tilde{u}(x, y, \tau) - \varphi_1(x, y, 0)) = f_u^*(\tilde{u}(x, y, \tau) - \varphi_1(x, y, 0))$, где f_u^* обозначает производную, взятую в промежуточной точке $(\tilde{u}(x, y, \tau) + \theta(\varphi_1(x, y, 0) - \tilde{u}(x, y, \tau)), x, y, 0)$, $0 < \theta < 1$. Уравнение (8) запишем в виде

$$\frac{\partial}{\partial\tau}(\tilde{u}(x, y, \tau) - \varphi_1(x, y, 0)) = f_u^*(\tilde{u}(x, y, \tau) - \varphi_1(x, y, 0)). \tag{15}$$

В силу (10) и (12) имеем $|u - \varphi_1(x, y, 0)| \le \eta_0$, $f_u^* \ge m$ при $(x, y) \in \overline{D}^-$, $\tau \ge \tau_\delta$. Отсюда и из (15) следует оценка (13). Оценка (14) может быть получена таким же образом. Лемма 1 доказана.

3. ОЦЕНКИ ПРОИЗВОДНЫХ

Выведем также важные для дальнейшего оценки производных решения уравнения (8) по декартовым координатам (x, y) и локальным координатам (ρ, l), введенным вблизи границы Γ. Пусть $\tilde{f}_u(x, y, \tau) = f_u(\tilde{u}(x, y, \tau), x, y, 0)$. Так как $\underline{u} \le \tilde{u} \le \bar{u}$ при всех $(x, y) \in \overline{D}$, $\tau \ge 0$, а $f_u(\tilde{u}, \rho, l, 0)$ – ограниченная функция при $\underline{u} \le u \le \bar{u}$, то существует такое число $p > 0$, что

$$-\tilde{f}_u(x, y, \tau) \le p \quad \forall (x, y) \in \overline{D}, \quad \tau \ge 0. \tag{16}$$

Заметим, что координаты (x, y) входят в уравнение (8) как параметры. Продифференцировав (8) по x и положив $\tilde{f}_x(x, y, \tau) = f_x(\tilde{u}(x, y, \tau), x, y, 0)$, получим следующую задачу для $\tilde{u}_x$:

$$\frac{\partial\tilde{u}_x}{\partial\tau} = -\tilde{f}_u(x, y, \tau)\tilde{u}_x - \tilde{f}_x(x, y, \tau), \quad (x, y) \in D, \quad \tau > 0,$$

$$\tilde{u}_x(x, y, 0) = \frac{\partial u_0(x, y)}{\partial x}, \quad (x, y) \in \overline{D},$$

решение которой можно выписать точно. Далее, с учетом (16) нетрудно доказать оценку

$$|\tilde{u}_x(x, y, \tau)| \le c_1 e^{p\tau} + c_2 \int_0^\tau e^{p(\tau-\xi)} d\xi \le c e^{p\tau}, \quad (x, y) \in \bar{D}, \quad \tau \ge 0, \tag{17а}$$

где учтено, что $\left|\dfrac{\partial u^0(x, y)}{\partial x}\right| \le c_1$, $|\tilde{f}_x(x, y, \tau)| \le c_2$ при $(x, y) \in \bar{D}$, $\tau \ge 0$. Буквами c и c_i здесь и далее обозначаются подходящие положительные числа, не зависящие от ε.

Аналогичное неравенство справедливо и для производной по второй координате y:

$$|\tilde{u}_y(x, y, \tau)| \le c e^{p\tau}, \quad (x, y) \in \bar{D}, \quad \tau \ge 0. \tag{17б}$$

Из (17а) и (17б) непосредственно вытекает оценка

$$\left|\frac{\partial \tilde{u}(x, y, \tau)}{\partial n}\right| \le c e^{p\tau}, \quad (x, y) \in \Gamma, \quad \tau \ge 0. \tag{17}$$

Далее, дифференцируя (8) дважды по x, получаем

$$\frac{\partial \tilde{u}_{xx}}{\partial \tau} = -\tilde{f}_u(x, y, \tau)\tilde{u}_{xx} - [\tilde{f}_{uu}(x, y, \tau)\tilde{u}_x^2 + 2\tilde{f}_{ux}(x, y, \tau)\tilde{u}_x + \tilde{f}_{xx}(x, y, \tau)], \quad (x, y) \in D, \quad \tau > 0,$$

$$\tilde{u}_{xx}(x, y, 0) = \frac{\partial^2 u^0(x, y)}{\partial^2 x}, \quad (x, y) \in D,$$

откуда с учетом

$$\left|\frac{\partial^2 u_0(x, y)}{\partial^2 x}\right| \le c_3, \quad \left|\tilde{f}_{uu}(x, y, \tau)\tilde{u}_x^2 + 2\tilde{f}_{ux}(x, y, \tau)u_x + \tilde{f}_{xx}(x, y, \tau)\right| \le c_4 e^{2p\tau}$$

имеем

$$|\tilde{u}_{xx}(x, y, \tau)| \le c_3 e^{p\tau} + c_4 \int_0^\tau e^{2p\xi} e^{p(\tau-\xi)} d\xi \le c e^{2p\tau}, \quad (x, y) \in \bar{D}, \quad \tau \ge 0. \tag{17а}$$

Действуя аналогично, получаем оценку $|\tilde{u}_{xx}(x, y, \tau)| \le c e^{2p\tau}$, а также

$$|\Delta_{xy}\tilde{u}(x, y, \tau)| \le c e^{p\tau}, \quad (x, y) \in \bar{D}, \quad \tau \ge 0. \tag{18}$$

Кроме того, нам понадобятся оценки производных функции $\tilde{u}(x, y, \tau)$ – решения уравнения (8) – по координате вдоль границы области D. Для этого в окрестности границы Γ введем локальную систему координат (ρ, l) так, что ρ – координата вдоль нормали к границе внутрь области D, а l – координата той точки на кривой Γ, из которой эта нормаль выпущена. Записав теперь уравнение (8) в локальных координатах (ρ, l):

$$-\tilde{u}_\tau = f(\tilde{u}, \rho, l, 0), \quad \tau > 0, \quad \tilde{u}(\rho, l, 0) = u^0(\rho, l),$$

и проделав процедуру, аналогичную описанной выше, получим

$$|\tilde{u}_{\rho\tau}(0, l, \tau)| \le c e^{p\tau}, \quad \left|\frac{\partial}{\partial l}\tilde{u}_\rho(0, l, \tau)\right| \le c e^{2p\tau}, \quad \left|\frac{\partial^2}{\partial l^2}\tilde{u}_\rho(0, l, \tau)\right| \le c e^{3p\tau}, \quad \tau \ge 0, \tag{19}$$

где l – координата вдоль кривой Γ.

При помощи оценок (17)–(19) доказывается

Лемма 2. *Пусть выполнены условия* A_1, A_2, *и пусть* A *– любое число из интервала* $(0, 1 - p)$, *где* p *определено неравенством* (16). *Тогда для решения* $\bar{u}(x, y, \tau, \varepsilon)$ *задачи* (5)–(7) *имеет место представление*

$$\bar{u}(x, y, \tau, \varepsilon) = \tilde{u}(x, y, \tau) + O(\varepsilon^{1-pA}), \quad (x, y) \in \bar{D}, \quad 0 \le \tau \le \tau_A = A|\ln\varepsilon|. \tag{20}$$

Доказательство. Введем функцию

$$U(x, y, \tau, \varepsilon) = \tilde{u}(x, y, \tau) + R(\rho, l, \tau, \varepsilon),$$

где

$$R(\rho, l, \tau, \varepsilon) = -\varepsilon\sigma(\rho)\frac{\partial\tilde{u}(x, y, \tau)}{\partial n}\bigg|_{(x, y) \in \Gamma} e^{-\rho/\varepsilon} = \varepsilon\sigma(\rho)\tilde{u}_\rho(0, l, \tau)e^{-\rho/\varepsilon}.$$

Здесь $\sigma(\rho)$ – достаточно гладкая срезающая функция, равная единице при $\rho \in [0, \delta]$, а при $\rho \ge 2\delta$ – нулю, δ – достаточно малое положительное число.

Рассмотрим выражение $L_\varepsilon U \equiv \varepsilon^2\Delta U - U_\tau - f(U, x, y, \varepsilon^2\tau)$, где $\Delta U = \Delta_{xy}\tilde{u} + \Delta_{\rho l}R$. Здесь

$$\Delta_{\rho l}R = \frac{\partial^2 R}{\partial\rho^2} + k(\rho, l)\frac{\partial R}{\partial\rho} + m(\rho, l)\frac{\partial^2 R}{\partial l^2} + h(\rho, l)\frac{\partial R}{\partial l},$$

а $k(\rho, l)$, $m(\rho, l)$ и $h(\rho, l)$ – известные достаточно гладкие функции. Вычислив необходимые производные и сохранив слагаемые порядка ε, с учетом (18), (19) получим

$$L_\varepsilon U = \varepsilon\sigma(\rho)\tilde{u}_\rho(0, l, \tau)e^{-\rho/\varepsilon} - \tilde{u}_\tau - \varepsilon\sigma(\rho)(\tilde{u}_\rho(0, l, \tau))_\tau e^{-\rho/\varepsilon} - f(\tilde{u}, x, y, 0) -$$
$$-\varepsilon f_u(\tilde{u}(x, y, \tau) + \theta R(\rho, l, \tau, \varepsilon), x, y, 0)\sigma(\rho)\tilde{u}_\rho(0, l, \tau)e^{-\rho/\varepsilon} + O(\varepsilon^2), \quad 0 < \theta < 1.$$

Далее, используя оценки (17) и уравнение (8), нетрудно показать, что

$$L_\varepsilon U \equiv \varepsilon^2\Delta U - U_t - f(U, x, y, \varepsilon^2\tau) = O(\varepsilon e^{p\tau}), \quad (x, y) \in \bar{D}, \quad 0 < \tau < \tau_A, \quad pA < 1.$$

Кроме того,

$$\frac{\partial U(x, y, \tau, \varepsilon)}{\partial n} = 0, \quad (x, y) \in \Gamma, \quad \tau > 0,$$

а в силу дополнительного условия из (8) и условия согласования (3) имеем $U(x, y, 0, \varepsilon) = u^0(x, y)$, $(x, y) \in \bar{D}$.

Сделаем теперь в (5)–(7) замену переменных

$$\bar{u}(x, y, \tau, \varepsilon) = U(x, y, \tau, \varepsilon) + w(x, y, \tau, \varepsilon)e^{p\tau}. \tag{21}$$

Заметим, что

$$L_\varepsilon u = L_\varepsilon U + (\varepsilon^2\Delta w - w_t - pw)e^{p\tau} - [f(U + we^{p\tau}, x, y, \varepsilon^2\tau) - f(U, x, y, \varepsilon^2\tau)] = 0,$$

а также

$$f(U + we^{p\tau}, x, y, \varepsilon^2\tau) - f(U, x, y, \varepsilon^2\tau) = we^{p\tau}\int_0^1 f_u(U + \xi we^{p\tau}, x, y, \varepsilon^2\tau)d\xi > -p.$$

Тогда для определения $w(x, y, \tau, \varepsilon)$ получим задачу

$$\begin{gathered}\varepsilon^2\Delta w - w_\tau - (p + H(x, y, \tau, \varepsilon))w = -e^{-p\tau}L_\varepsilon U = O(\varepsilon), \quad (x, y) \in D, \quad 0 < \tau \le \tau_A,\\ \frac{\partial w(x, y, \tau, \varepsilon)}{\partial n} = 0, \quad (x, y) \in \Gamma, \quad 0 < \tau \le \tau_A, \quad w(x, y, 0, \varepsilon) = 0, \quad (x, y) \in \bar{D},\end{gathered} \tag{22}$$

где введено обозначение

$$H(x, y, \tau, \varepsilon) = \int_0^1 f_u(U + \xi we^{p\tau}, x, y, \varepsilon^2\tau)d\xi.$$

Так как коэффициент $p + H(x, y, \tau, \varepsilon)$ в уравнении (22) положителен, в силу принципа максимума имеем

$$w(x, y, \tau, \varepsilon) = O(\varepsilon), \quad (x, y) \in \bar{D}, \quad 0 \le \tau \le \tau_A,$$

и, учитывая (21), получаем

$$\bar{u}(x, y, \tau, \varepsilon) = U(x, y, \tau, \varepsilon) + O(\varepsilon)e^{p\tau} = U(x, y, \tau, \varepsilon) + O(\varepsilon^{1-pA}),$$
$$(x, y) \in \bar{D}, \quad 0 \le \tau \le \tau_A(\varepsilon) = A\ln|\varepsilon|.$$

В силу того, что из (17) и (22) следует оценка

$$R(\rho, l, \tau, \varepsilon) = O(\varepsilon e^{p\tau}) = O(\varepsilon^{1-pA}),$$

имеем

$$U(x, y, \tau, \varepsilon) = \tilde{u}(x, y, \tau) + R(\rho, l, \tau, \varepsilon) = \tilde{u}(x, y, \tau) + O(\varepsilon^{1-pA}),$$

откуда получаем (19). Лемма 2 доказана.

Из доказанных лемм 1 и 2 вытекает следующее утверждение о формировании двумерной контрастной структуры.

Теорема 1. *Пусть выполнены условия* A_1, A_2, *и пусть* m *и* p – *числа из* (12) *и* (16), *а* δ – *любое достаточно малое число. Положим* $A = 1/(p + m)$, $r = m/(p + m)$. *Тогда при достаточно малых* ε *для решения* $u(x, y, t, \varepsilon)$ *задачи* (1)–(3) *в момент времени* $t = t_A(\varepsilon) = A\varepsilon^2|\ln\varepsilon|$ *справедливы представления*

$$u(x, y, t_A(\varepsilon), \varepsilon) = \varphi_1(x, y, 0) + O(\varepsilon^r), \quad (x, y) \in \bar{D}^-,$$
$$u(x, y, t_A(\varepsilon), \varepsilon) = \varphi_2(x, y, 0) + O(\varepsilon^r), \quad (x, y) \in \bar{D}^+. \tag{23}$$

Доказательство. В силу леммы 2,

$$u(x, y, t_A(\varepsilon), \varepsilon) = \bar{u}(x, y, \tau_A(\varepsilon), \varepsilon) = \tilde{u}(x, y, \tau_A) + O(\varepsilon^{1-pA}) = \tilde{u}(x, y, \tau_A) + O(\varepsilon^r),$$

а в силу леммы 1 получаем

$$|\tilde{u}(x, y, \tau_A) - \varphi_1(x, y, 0)| \le \eta_0 e^{-m(\tau_A - \tau_\delta)} = O(\varepsilon^{mA}) = O(\varepsilon^r), \quad (x, y) \in \bar{D}^-.$$

Поэтому имеем

$$u(x, y, t_A(\varepsilon), \varepsilon) = \varphi_1(x, y, 0) + O(\varepsilon^r), \quad (x, y) \in \bar{D}^-.$$

Аналогично устанавливается второе равенство в (23). Теорема 1 доказана.

Равенства (23) показывают, что в момент времени $t_A(\varepsilon) = A\varepsilon^2|\ln\varepsilon|$ решение $u(x, y, t, \varepsilon)$ задачи (1)–(3) вне δ-окрестности замкнутой кривой h_0 отличаются от $\varphi_1(x, y, 0)$ в подобласти $\bar{D}^-$ и от $\varphi_2(x, y, 0)$ в подобласти $\bar{D}^+$ на величину порядка $O(\varepsilon^r)$, т.е. к этому времени контрастная структура с переходным слоем в окрестности кривой h_0 сформировалась.

4. СРАВНЕНИЕ С РЕЗУЛЬТАТАМИ ЧИСЛЕННОГО ЭКСПЕРИМЕНТА

Численное решение задачи (1)–(3) производилось методом конечных разностей при различных значениях малого параметра ε и для двух вариантов правой части уравнения (1).

Область D в численном эксперименте представляет собой квадрат со сторонами 0.8, а начальное условие в обоих случаях было выбрано в виде

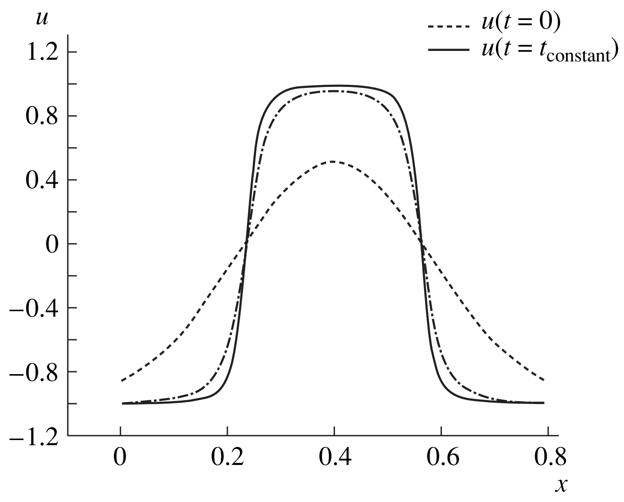
$$u^0(x, y) = 0.75(\exp\{-15[x - 0.4]\} + \exp\{-15[y - 0.4]\}) - 1. \tag{24}$$

В первом случае функция $f(u, x, y, t)$ была взята в виде

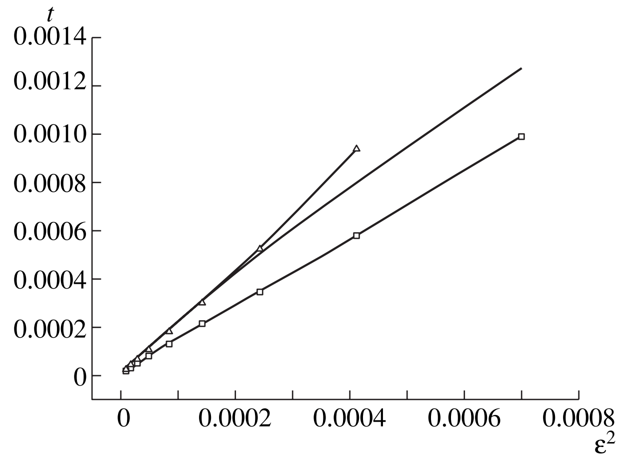
$$f(u, x, y, t) = u(u^2 - 1), \tag{25}$$

т.е. не зависящей явно от пространственных переменных.

На фиг. 3 изображено сечение области D при $y = 0.4$, параллельное оси x. Штриховыми линиями показана эволюция решения задачи от начального условия (24) к контрастной структуре типа ступеньки при $\varepsilon = 7 \times 10^{-3}$. Штрихпунктирной линией обозначено промежуточное значение. Время выхода на контрастную структуру составило $t_{\text{contrast}} = 11.03 \times 10^{-5}$ с. Это время вычислялось в соответствии с условиями теоремы 1. Численный счет останавливался, когда решение прибли-



Фиг. 3.



Фиг. 4.

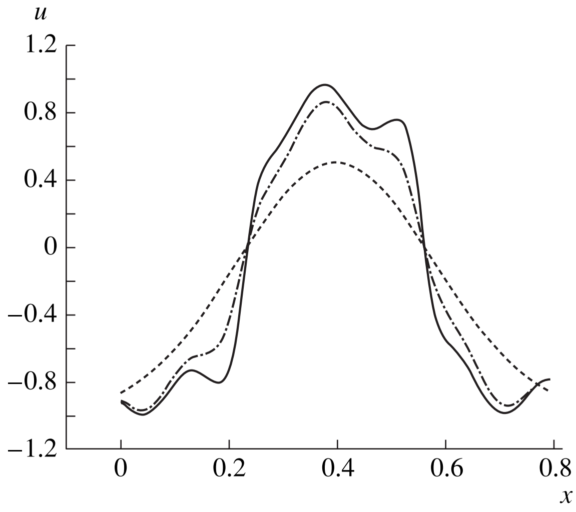
жалось при $\varepsilon^r = 0.0837$ к устойчивым корням вырожденного уравнения (4), в данном случае $\varphi_1(x, y, 0) = 1$ и $\varphi_2(x, y, 0) = -1$, причем в δ-окрестности неустойчивого корня $\varphi_0(x, y, 0) = 0$ сравнение на близость не производилось. В численных расчетах принималось $\delta = 0.05$, а значения констант, вычисленные в соответствии с теоремой 1, в данном случае таковы: $r = 0.5$ и $A = 0.5$.

Теорема 1 утверждает, что в исследуемом случае контрастная структура образуется не позже момента времени $t_A(\varepsilon) = A\varepsilon^2|\ln\varepsilon| = 12.18 \times 10^{-5}$ с, что прекрасно согласуется с результатами моделирования.

Для более точной иллюстрации доказанной выше теоремы 1 были проделаны вычисления для нескольких значений параметра ε в диапазоне от $\varepsilon = 3 \times 10^{-3}$ до $\varepsilon = 3 \times 10^{-2}$. Кривая зависимости времени образования контрастной структуры от квадрата малого параметра представлена на фиг. 4.

На фиг. 4 сплошной линией обозначена зависимость $t_A(\varepsilon) = A\varepsilon^2|\ln\varepsilon|$, предсказанная теоремой 1, а значки представляют собой результаты численного эксперимента для различных значений ε: квадратики – время выхода на контрастную структуру при $\delta = 0.1$, треугольники – при $\delta = 0.05$.

Видно, что при достаточно малых ε теоретическая и экспериментальная кривые практически совпадают для $\delta = 0.05$. При больших же ε отмечается существенное расхождение с теорией, что говорит о том, что в этом случае следует выбирать другое δ.



Фиг. 5.

Действительно, при $\delta = 0.1$ для всех ε из выбранного диапазона численный график остается ниже теоретической кривой, т.е. в эксперименте контрастная структура образовывалась раньше, чем в момент $t_A(\varepsilon) = A\varepsilon^2|\ln\varepsilon|$, что соответствует теоретическим результатам.

Во втором численном эксперименте полагаем $f(u, x, y, t) = u(u^2 - 0.8 - 0.2\sin(12\pi x))$, т.е. теперь, в отличие от (25), функция $f(u, x, y, t)$ явно зависит от пространственной переменной x. На фиг. 5 изображено сечение области, аналогичное рассмотренному ранее на фиг. 3.

В данном случае в областях D^+ и D^- (см. фиг. 2) решение близко не к постоянным уровням, а зависящим от переменной x и являющимся устойчивыми нулями функции $f(x, y, t)$.

Время формирования контрастной структуры при $\varepsilon = 7 \times 10^{-3}$ составило $t_{\text{contrast}} = 5.51 \times 10^{-5}$ c, тогда как теоретический результат дает оценку сверху $t_A = 19.65 \times 10^{-5}$ c.

СПИСОК ЛИТЕРАТУРЫ

1. *Васильева А.Б., Бутузов В.Ф., Нефедов Н.Н.* Контрастные структуры в сингулярно возмущенных задачах // Фундаментальная и прикл. матем. 1998. Т. 4. № 3. С. 799–851.
2. *Нефёдов Н.Н.* Асимптотический метод дифференциальных неравнеств в исследовании периодических контрастных структур: существование, асимптотика, устойчивость // Дифференц. ур-ния. 2000. Т. 36. № 2. С. 262–269.
3. *Nefedov N.N., Radzinus M., Schneider K.R., Vaseil'eva A.B.* Change of the type of contrast structures in parabolic Neumann Problems // Ж. вычисл. матем. и матем. физ. 2005. Т. 45. № 1. С. 41–55.
4. *Бутузов В.Ф., Неделько И.В.* О глобальной области влияния устойчивых решений с внутренними слоями // Матем. сб. 2001. Т. 192. № 5. С. 13–52.
5. *Бутузов В.Ф., Нефедов Н.Н., Шнайдер К.Р.* О формировании и распространении резких переходных слоев в параболических задачах // Вестн. МГУ. Сер. 3. Физ. Астрон. 2005. № 1. С. 9–13.
6. *Allen S.M., Cahn J.W.* A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening // Acta Metal. 1979. V. 27. P. 1085–1095.

ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 2007, том 47, № 8, с. 1365–1377

УДК 519.633.9

ОБРАТНЫЕ ЗАДАЧИ ОПРЕДЕЛЕНИЯ ИСТОЧНИКА И КОЭФФИЦИЕНТА В ЭЛЛИПТИЧЕСКОМ УРАВНЕНИИ В ПРЯМОУГОЛЬНИКЕ[1)]

© 2007 г. В. В. Соловьёв

(115409 Москва, Каширское ш., 31, МИФИ)

e-mail: solovevv@mail.ru

Поступила в редакцию 17.12.2006 г.
Переработанный вариант 26.12.2006 г.

Исследуются обратные задачи определения источника и коэффициента в эллиптическом уравнении, заданном в прямоугольнике. В качестве дополнительной информации о решении прямой задачи (переопределении) считается известным след ее решения на отрезке внутри прямоугольника. Для изучаемых обратных задач получены достаточные условия существования и единственности (глобальные). Исследование проводится в классе непрерывно дифференцируемых функций, производные которых удовлетворяют условию Гёльдера. Библ. 16.

Ключевые слова: обратные задачи, эллиптическое уравнение.

1. ПОСТАНОВКА ЗАДАЧИ

Пусть $0 < \alpha < 1$, $l > 0$, $q_1 < 0 < q_2$, где l, q_1, q_2 – фиксированные числа. На плоскости с декартовыми координатами (x, y) рассмотрим прямоугольник $\Pi = \{(x, y) : 0 < x < l, q_1 < y < q_2\}$. Боковые границы прямоугольника будем обозначать через $T_1 = \{0\} \times (q_1, q_2)$, $T_2 = \{l\} \times (q_1, q_2)$.

Для формулировки обратных задач, изучаемых в данной работе, определим некоторые функциональные пространства. Для этого будем пользоваться общепринятыми обозначениями пространств непрерывных функций, удовлетворяющих условию Гёльдера (см., например, [1, с. 57], [2, с. 29]). Как обычно, если $\Omega \subset \mathbb{R}^2$ – ограниченная область, то $C^{2+\alpha}(\Omega)$ – пространство функций, имеющих внутри Ω все непрерывные производные до второго порядка включительно, при этом вторые производные предполагаются локально-непрерывными по Гёльдеру с показателем α. Определим пространства функций:

$$\overline{C}^{2+\alpha}(\Pi) = C^{2+\alpha}(\Pi) \cap C(\overline{\Pi}) \cap C_x^2(\Pi \cup T_1 \cup T_2), \quad \overline{C}^{2+\alpha}(0, l) = C^{2+\alpha}(0, l) \cap C^2[0, l],$$

$$\overline{C}^{\alpha}(0, l) = C^{\alpha}(0, l) \cap C[0, l], \quad C_-[0, l] = \{c(x) : c \in C[0, l], c(x) \le 0, x \in [0, l]\},$$

$$\overline{C}_-^{\alpha}(0, l) = C^{\alpha}(0, l) \cap C_-[0, l], \quad C_-^{\alpha}[0, l] = C^{\alpha}[0, l] \cap C_-[0, l], \quad \overline{C}^{\alpha}(\Pi) = C^{\alpha}(\Pi) \cap C(\overline{\Pi}).$$

Всюду далее знаком нормы без дополнительных индексов обозначается обычная sup-норма.

В прямоугольнике $\overline{\Pi}$ рассмотрим две тесно связанные между собой обратные задачи.

Задача 1. Определить пару функций $(u, f) \in \overline{C}^{2+\alpha}(\Pi) \times \overline{C}^{\alpha}(0, l)$ из условий

$$-\Delta u(x, y) = c(x)u(x, y) + f(x)h(x, y) + g(x, y), \quad (x, y) \in \Pi, \quad u(x, y) = \varphi(x, y), \quad (x, y) \in \partial\Pi, \tag{1}$$

$$u(x, 0) = \chi(x), \quad x \in [0, l]. \tag{2}$$

Задача 2. Определить пару функций $(u, c) \in \overline{C}^{2+\alpha}(\Pi) \times \overline{C}_-^{\alpha}(0, l)$ из условий

$$-\Delta u(x, y) = c(x)u(x, y) + g(x, y), \quad (x, y) \in \Pi, \quad u(x, y) = \varphi(x, y), \quad (x, y) \in \partial\Pi, \tag{3}$$

$$u(x, 0) = \chi(x), \quad x \in [0, l]. \tag{4}$$

[1)] Работа выполнена при финансовой поддержке РФФИ (код проекта 06-01-00401).

В условиях (1), (2) величина $c \le 0$, h, q, φ, χ – заданные достаточно гладкие функции. В условиях (3), (4) величины g, φ, χ – заданные достаточно гладкие функции. Приведем здесь для полноты изложения и дальнейших ссылок формулировки теорем существования и единственности для прямых задач (1), (3), следующих из [1, с. 108, с. 112, с. 107] (там даны результаты для более общих областей и более общих уравнений).

Теорема 0. *Пусть* $c \in C_-^{\alpha}[0, l]$, $g \in C^{\alpha}(\overline{\Pi})$, $\varphi \in C(\partial\Pi)$. *Тогда существует единственное решение прямой задачи* (3) *в классе функций* $u \in C^{2+\alpha}(\Pi) \cap C(\overline{\Pi})$. *При этом если* $\varphi \in C^{2+\alpha}(T_1) \cap C^{2+\alpha}(T_2)$, *то* $u \in C^{2+\alpha}(\Pi \cup T_1 \cup T_2)$, *откуда следует* $u(\cdot, 0) \in C^{2+\alpha}[0, l] \subset \overline{C}^{2+\alpha}(0, l)$.

Теорема 00. *Пусть* $c \in \overline{C}_-^{\alpha}(0, l)$, $g \in \overline{C}^{\alpha}(\Pi)$, $\varphi \in C(\partial\Pi)$. *Тогда существует единственное решение прямой задачи* (3) *в классе функций* $u \in C^{2+\alpha}(\Pi) \cap C(\overline{\Pi})$.

Обратные задачи вида (1), (2) и (3), (4) изучались для различных типов уравнений многими авторами (подробную библиографию и историю вопроса (см. в [3]–[10]).

Задача 1 изучалась автором в работе [11] для случая области с гладкой границей в предположении $c = 0$. Для нее там доказана теорема существования и единственности при некоторых ограничениях на заданные функции и область (прямоугольник требованиям, наложенным на область в [11], не удовлетворяет).

В [16] задача 1 рассматривалась для областей существенно более общего вида, в более узком классе функций $(u, f) \in C^{2+\alpha}(\Pi \cup T_1 \cup T_2) \cap C(\overline{\Pi}) \times C^{\alpha}[0, l]$. Для задачи 1 в этом классе функций там доказана справедливость альтернативы Фредгольма и – во второй части работы – доказаны теоремы, дающие различные достаточные условия для однозначной разрешимости задачи 1. Доказанные в этой работе теоремы носят характер разрешимости в “малом”, т.е. в качестве условий разрешимости выступают требования малости l или большой по модулю величины коэффициента c. Как и в [16] при изучении задачи 1, для функции f строится некоторое уравнение II рода в функциональном пространстве. В настоящей статье оно получается способом, отличным от способа в [16], при этом появляется возможность рассмотреть это уравнение в пространстве $C[0, l]$ вместо $C^{\alpha}[0, l]$. При изучении полной непрерывности оператора, входящего в построенное уравнение II рода, в [16] используются оценки Шаудера, а в настоящей статье – оценки “$1 + \alpha$”, специфические для $n = 2$, т.е. используется принципиально другая методика изучения задачи 1 по сравнению с [16]. Все это позволяет получить для задачи 1 глобальные условия однозначной разрешимости (без требования малости). Кроме того, предлагаемая методика дает возможность получить достаточные условия однозначной разрешимости коэффициентной задачи 2.

2. ФОРМУЛИРОВКА И ДОКАЗАТЕЛЬСТВО АЛЬТЕРНАТИВЫ ФРЕДГОЛЬМА ДЛЯ ЗАДАЧИ 1

Определим следующее линейное множество троек функций:

$$G = \{(g, \varphi, \chi) : g \in C^{\alpha}(\overline{\Pi}), \varphi \in C(\partial\Pi) \cap C^{2+\alpha}(T_1) \cap C^{2+\alpha}(T_2), \chi \in \overline{C}^{2+\alpha}(0, l),$$
$$\varphi(0, 0) = \chi(0), \varphi(l, 0) = \chi(l)\}.$$

Заметим, что условия $\varphi(0, 0) = \chi(0)$, $\varphi(l, 0) = \chi(l)$ являются необходимыми для существования решения прямой задачи (1) в рассматриваемом классе функций и носят характер условий согласования исходных данных.

Теорема 1 (альтернатива Фредгольма). *Пусть* $c \in C_-^{\alpha}[0, l]$, h, $h_{yy} \in C^{\alpha}(\overline{\Pi})$, $|h(x, 0)| \ge h_0 > 0$, $x \in [0, l]$. *Тогда для задачи* 1 *верно одно из двух утверждений*:

1) *задача* 1 *имеет единственное решение для любой тройки функций* $(g, \varphi, \chi) \in G$;

2) *при* $g = 0$, $\varphi = 0$, $\chi = 0$ *задача* 1 *имеет нетривиальное решение*.

Доказательство. Определим необходимые для дальнейших рассмотрений функциональные пространства:

$$\overline{C}_0^{2+\alpha}(0, l) = \{\chi : \chi \in \overline{C}^{2+\alpha}(0, l), \chi(0) = \chi(l) = 0\}.$$

Сформулируем задачу 1 при $g = 0$, $\varphi = 0$, $\chi \in \overline{C}_0^{2+\alpha}(0, l)$.

Задача 1_0. Для произвольной функции $\chi \in \overline{C}_0^{2+\alpha}(0, l)$ определить пару функций $(u, f) \in \overline{C}^{2+\alpha}(\overline{\Pi}) \times \overline{C}^{\alpha}(0, l)$ из условий

$$-\Delta u = cu + fh, \quad (x, y) \in \Pi, \quad u(x, y) = 0, \quad (x, y) \in \partial\Pi, \tag{5}$$

$$u(x, 0) = \chi(x), \quad x \in [0, l]. \tag{6}$$

Лемма 1. *Теорема* 1 *верна для задачи* 1 *тогда и только тогда, когда она верна для задачи* 1_0, *т.е. верно одно из двух утверждений*:

1) *задача* 1_0 *имеет единственное решение для любой функции* $\chi \in \overline{C}_0^{2+\alpha}(0, l)$;

2) *при* $\chi = 0$ *задача* 1_0 *имеет нетривиальное решение.*

Доказательство этой и последующих лемм см. далее в специальном разделе.

Из леммы 1 следует, что для доказательства теоремы 1 достаточно доказать альтернативу Фредгольма для задачи 1_0, что и будет проведено далее. Кроме того, так как $|h(x, 0)| \ge h_0 > 0$, то без ограничения общности будем считать далее, что $h(x, 0) \equiv 1$.

Для дальнейшего изучения сформулированных обратных задач потребуется некоторое обобщение понятия решения прямой задачи (3) на случай разрывных краевых условий (разрывы I рода у функции φ) в точках $A(0, q_1)$, $B(0, q_2)$, $C(l, q_2)$, $D(l, q_1)$). Пусть функция $\varphi(x, y)$ имеет вид

$$\varphi(0, y) = \nu_1(y), \quad \varphi(l, y) = \nu_2(y), \quad q_1 < y < q_2, \quad \varphi(x, q_2) = \mu_2(x), \quad \varphi(x, q_1) = \mu_1(x), \quad 0 < x < l.$$

Будем предполагать, что $\nu_1, \nu_2 \in C[q_1, q_2]$, $\mu_1, \mu_2 \in C[0, l]$. Тогда если выполнены условия согласования $\nu_1(q_2) = \mu_2(0)$, $\mu_2(l) = \nu_2(q_2)$, $\mu_1(l) = \nu_2(q_1)$, $\nu_1(q_1) = \mu_1(0)$, то функция $\varphi \in C(\partial\Pi)$ и решение прямой задачи (3), как известно (в силу того что A, B, C, D – регулярные точки границы), строится, например, методом Перрона (см. [1, с. 105]). Если эти условия не выполнены, то решение краевой задачи (3) понимается в обобщенном смысле. Идейная сторона такого обобщения понятия решения изложена, например, в [12, с. 107]. Так как автору не удалось найти в литературе необходимую точную ссылку, то приведем для полноты необходимое утверждение о прямой задаче с доказательством. Итак, пусть указанные выше условия согласования в точках A, B, C, D, вообще говоря, не выполнены. Заменим функции μ_1, μ_2, ν_1, ν_2 в окрестности угловых точек, например, линейными функциями так, чтобы условия согласования в угловых точках были выполнены, при этом будем предполагать что в самих точках A, B, C, D эти линейные функции обращаются в ноль. Кроме того, в силу справедливости теоремы 00 и линейности задачи (3) без ограничения общности при определении понятия обобщения решения задачи (3), можно считать уравнение в условиях (3) однородным. В соответствии со сказанным выше определим при достаточно больших n функции $\nu_1^{(n)}$, $\mu_2^{(n)}$, $\nu_2^{(n)}$, $\mu_1^{(n)}$. Для примера выпишем функцию $\mu_1^{(n)}$ (остальные функции строятся аналогично):

$$\mu_1^{(n)}(y) = \left\{\mu_1\left(\frac{1}{n}\right)x,\ 0 \le x \le \frac{1}{n};\ \mu_1(x),\ \frac{1}{n} \le x \le l - \frac{1}{n};\ \mu_1\left(l - \frac{1}{n}\right)n(q_2 - y),\ l - \frac{1}{n} \le x \le l\right\}.$$

Таким образом, в качестве краевых условий будем рассматривать функции $\varphi_n(x, y)$:

$$\varphi_n(0, y) = \nu_1^{(n)}(y), \quad \varphi_n(x, q_2) = \mu_2^{(n)}(x), \quad \varphi_n(l, y) = \nu_2^{(n)}(y), \quad \varphi_n(x, q_1) = \mu_1^{(n)}(x).$$

Ясно, что $\varphi_n \in C(\partial\Pi)$ и $\|\varphi_n\| \le \max\{\|\nu_1\|, \|\nu_2\|, \|\mu_1\|, \|\mu_2\|\}$ при достаточно больших n. Будем далее полагать $\Pi^* = \overline{\Pi} \backslash \{A, B, C, D\}$. Рассмотрим задачу определения функции $u(x, y)$ из условий

$$-\Delta u(x, y) = c(x)u(x, y), \quad (x, y) \in \Pi, \tag{7}$$

$$u(0, y) = \nu_1(y), \quad u(l, y) = \nu_2(y), \quad q_1 < y < q_2, \quad u(x, q_1) = \mu_2(x), \quad u(x, q_2) = \mu_2(x), \quad 0 < x < l. \tag{8}$$

Задаче (7), (8) поставим в соответствие последовательность функций $u^{(n)}(x, y)$, где $u^{(n)}$ – решения краевых задач

$$-\Delta u^{(n)}(x, y) = c(x)u^{(n)}(x, y), \quad (x, y) \in \Pi, \quad u^{(n)}(x, y) = \varphi_n(x, y), \quad (x, y) \in \partial\Pi. \tag{9}$$

Определение. *Решением задачи* (7), (8) *назовем функцию*

$$u(x, y) = \lim_{n \to \infty} u^{(n)}(x, y), \quad (x, y) \in \Pi^*.$$

Лемма 2. *Пусть* $c \in \bar{C}^{\alpha}_{-}(0, l)$. *Тогда существует решение задачи* (7), (8) *в смысле указанного определения*, $u \in C^2(\Pi) \cap C(\Pi^*)$ *и удовлетворяет в* Π *уравнению* (7). *При этом верны оценки, следующие из принципа максимума*: $\|u\| \le \max\{\|\nu_1\|, \|\nu_2\|, \|\mu_1\|, \|\mu_2\|\}$.

Следствие. Пусть $c \in \bar{C}^{\alpha}_{-}(0, l)$, $g \in \bar{C}^{\alpha}(\Pi)$, $\nu_1, \nu_2 \in C[q_1, q_2]$, $\mu_1, \mu_2 \in C[0, l]$; тогда существует решение $u \in C^2(\Pi) \cap C(\Pi^*)$ задачи (3) в смысле указанного выше определения.

Доказательство прямо следует из леммы 2 и теоремы 00.

Для дальнейшего изучения обратной задачи 1_0 рассмотрим задачу определения тройки функций $(u, v, f) \in \bar{C}^{2+\alpha}(\Pi) \times (C^2(\Pi) \cap C(\Pi^*)) \times \bar{C}^{\alpha}[0, l]$ из условий

$$-\Delta u(x, y) = c(x)u(x, y) + f(x)h(x, y), \quad (x, y) \in \Pi, \quad u(x, y) = 0, \quad (x, y) \in \partial\Pi, \tag{10}$$

$$-\Delta v(x, y) = c(x)v(x, y) + f(x)h_{yy}(x, y), \quad (x, y) \in \Pi, \tag{11}$$

$$\begin{gathered} v(x, q_1) = -f(x)h(x, q_1), \quad v(x, q_2) = -f(x)h(x, q_2), \quad 0 < x < l, \\ v(0, y) = v(l, y) = 0, \quad q_1 < y < q_2, \end{gathered} \tag{12}$$

$$f(x) = -\chi''(x) - c(x)\chi(x) - v(x, 0), \quad 0 \le x \le l. \tag{13}$$

Связь задач 1_0 и (10)–(13) устанавливает

Лемма 3. *Пусть* (u, f) – *решение задачи* 1_0; *тогда* $u_{yy} \in C^2(\Pi) \cap C(\Pi^*)$ *и тройка функций* $(u, v, f) = (u, y_{yy}, f)$ – *решение задачи* (10)–(13). *Обратно*: *если* (u, v, f) – *решение задачи* (10)–(13) *из указанного класса*, *то* $v = u_{yy}$ *и пара функций* (u, f) – *решение задачи* 1_0.

Из леммы 3 следует, что для доказательства теоремы 1 достаточно доказать альтернативу Фредгольма для задачи (10)–(13). Для этого определим линейный оператор A с областью определения $D(A) = \bar{C}^{\alpha}(0, l) \subset C[0, l]$ и областью значений в $C[0, l]$ по правилу $(Af)(x) = v(x, 0)$, где $v(x, y)$ – решение прямой задачи (11)–(13) (существующее в силу леммы 2). Заметим, что в силу принципа максимума (так как $c \le 0$) для прямой задачи (11), (12) выполнена оценка $\|v(\cdot, 0)\| \le C_1\|f\|$, где постоянная C_1 не зависит от f. Отсюда следует, что $v(\cdot, 0)$ равномерно непрерывно зависит от $f \in \bar{C}^{\alpha}(0, l)$ в sup-норме. Так как $\bar{C}^{\alpha}(0, l)$ всюду плотно в $C[0, l]$, то (см., например, [13, с. 81]) существует непрерывное продолжение оператора A с $D(A)$ на все пространство $C[0, l]$. Будем обозначать это продолжение через $\tilde{A}(f)$, $f \in C[0, l]$.

Лемма 4. $\tilde{A}(f)$ – *вполне непрерывный оператор*, *при этом для любой функции* $f \in C[0, l]$ *выполнено включение* $\tilde{A}(f) \in \bar{C}^{\alpha}(0, l)$.

Заметим, что равенство (13) можно записать в виде уравнения II рода в банаховом пространстве $C[0, l]$ с вполне непрерывным оператором $\tilde{A}$:

$$f + \tilde{A}f = \psi, \tag{14}$$

где функция $\psi = -\chi'' - c\chi \in \bar{C}^{\alpha}(0, l)$. Если $f \in C[0, l]$ – решение уравнения (14), то в силу леммы 4, выполнено включение $f \in \bar{C}^{\alpha}(0, l)$. Но тогда, решив прямую задачу (11), (12) (ее решение существует в силу следствия леммы 2) и прямую задачу (10) (см. теорему 00), получим, в силу леммы 3, что пара функций (u, f) – решение обратной задачи 1_0. Из леммы 3 следует, что и, обратно, если (u, f) – решение обратной задачи 1_0, то f – решение уравнения (14). Но для уравнения (14) верна альтернатива Фредгольма (см. [13, с. 203]). Отсюда следует, что она верна и для задачи 1_0. Теорема 1 доказана.

3. ТЕОРЕМЫ ЕДИНСТВЕННОСТИ И СУЩЕСТВОВАНИЯ ДЛЯ ЗАДАЧИ 1

Из теоремы 1 следует, что для доказательства существования единственного решения задачи 1 достаточно доказать единственность решения задачи 1.

Лемма 5. *Пусть* $h, h_{yy} \in C^\alpha(\overline{\Pi})$, $f \in \overline{C}^\alpha(0, l)$, $c \in \overline{C}_-^\alpha(0, l)$, $h(x, y) \geq 0$, $h_{yy} \leq 0$, $(x, y) \in \Pi$. *Тогда если* $u \in C^2(\Pi) \cap C(\overline{\Pi})$ – *решение прямой задачи* (1) *при* $g \equiv 0$, $\varphi \equiv 0$ *и выполнено условие* $u(x, 0) = 0$, $x \in (0, l)$, *то* $f(x)h(x, 0) = 0$, $x \in (0, l)$.

Теорема 2 (существования и единственности решения задачи 1). *Пусть* h, $h_{yy} \in C^\alpha(\overline{\Pi})$, $c \in$ $\in C_-^\alpha[0, l]$, $h(x, 0) \geq h_0 > 0$, $x \in [0, l]$, $h(x, y) \geq 0$, $h_{yy} \leq 0$ $(x, y) \in \Pi$. *Тогда для любой тройки функций* $(g, \varphi, \chi) \in G$ *существует единственное решение обратной задачи* (1), (2) *в классе функций* $(u, f) \in$ $\in \overline{C}^{2+\alpha}(\Pi) \times \overline{C}^\alpha(0, l)$.

Доказательство. В силу линейности задачи 1 и справедливости теоремы 1, достаточно доказать, что задача об определении пары функций $(u, f) \in \overline{C}^{2+\alpha}(\Pi) \times \overline{C}^\alpha(0, l)$ из условий

$$-\Delta u = cu + fh, \quad (x, y) \in \Pi, \quad u(x, y) = 0, \quad (x, y) \in \partial\Pi, \quad u(x, 0) = 0, \quad x \in [0, l], \tag{15}$$

имеет только тривиальное решение. Но из леммы 5 для задачи (15) получим равенство $f(x)h(x, 0) = 0$, $x \in (0, l)$. Тогда в силу условия $h(x, y) \geq h_0 > 0$ получаем, что $f \equiv 0$. Но тогда и $u \equiv 0$. Теорема 2 доказана.

Рассмотрим второй случай, когда для задачи 1 можно указать достаточные условия существования единственного решения. Будем предполагать, что $h(x, y) \equiv h(y)$, $h, h_{yy} \in C^\alpha[q_1, q_2]$. Для дальнейших формулировок введем следующие используемые далее обозначения. Рассмотрим на отрезке $[0, l]$ задачу Штурма–Лиувилля

$$L[z](x) = -z''(x) - c(x)z(x) = \lambda z(x), \quad 0 < x < l, \quad z(q_1) = z(q_2) = 0. \tag{16}$$

Известно (см. [14, с. 202]), что собственные значения $\{\lambda_n\}$ задачи (16) образуют монотонно возрастающую последовательность положительных чисел, а соответствующие им собственные функции $\{\varphi_n(x)\}$, $\varphi_k \in C^2[0, l]$, образуют ортогональный базис в $L_2(0, l)$. Пусть $g \in C[q_1, q_2]$ – произвольная функция. Будем обозначать через $w_k(y)$ решение краевой задачи

$$-w_k''(y) + \lambda_k w_k(y) = g(y), \quad y \in [q_1, q_2], \quad w_k(q_1) = w_k(q_2) = 0. \tag{17}$$

Известно, что решение задачи (17) выражается через функцию Грина $G(y, s; \lambda_k)$, имеющую вид

$$G(y, s; \lambda_k) = \begin{cases} \dfrac{\operatorname{sh}\sqrt{\lambda_k}(s - q_2)\operatorname{sh}\sqrt{\lambda_k}(y - q_1)}{\operatorname{sh}\sqrt{\lambda_k}(q_2 - q_1)}, & q_1 < y < s, \\[2ex] \dfrac{\operatorname{sh}\sqrt{\lambda_k}(s - q_1)\operatorname{sh}\sqrt{\lambda_k}(y - q_2)}{\operatorname{sh}\sqrt{\lambda_k}(q_2 - q_1)}, & s < y < q_2. \end{cases}$$

При этом $w_k(y)$ представляется в виде

$$w_k(y) = -\int_{q_1}^{q_2} G(y, s; \lambda_k)g(s)ds.$$

Определим при $\lambda > 0$ функцию $\kappa(\lambda)$ по правилу

$$\kappa(\lambda) = -\int_{q_1}^{q_2} G(0, s; \lambda)h(s)ds.$$

Теорема 3. *Задача* 1 *имеет единственное решение для любых функций* $(g, \varphi, \chi) \in G$ *тогда и только тогда, когда нули функции* $\kappa(\lambda)$ *не совпадают с* λ_k. *Если при некотором* k *выполнено равенство* $\kappa(\lambda_k) = 0$, *то задача* 1 *при* $g = 0$, $\varphi = 0$, $\chi = 0$ *имеет нетривиальное решение, при этом* $f = \varphi_k$.

Доказательство. В силу альтернативы Фредгольма, для доказательства существования единственного решения задачи 1 достаточно доказать единственность решения задачи 1_0 в условиях теоремы 3. Пусть (u, f) – решение задачи 1_0. Тогда условия (5), (6) можно записать в виде

$$-u_{yy}(x, y) + L[u](x, y) = f(x)h(y), \quad 0 < x < l, \quad q_1 < y < q_2, \tag{18}$$

$$u(0, y) = u(l, y) = 0, \quad q_1 \le y \le q_2, \quad u(x, q_1) = u(x, q_2) = 0, \quad 0 \le x \le l, \tag{19}$$

$$u(x, 0) = \chi(x), \quad 0 \le x \le l. \tag{20}$$

Умножим (18) на $\varphi_k(x)$ и проинтегрируем полученное равенство от 0 до по x. Тогда для функции

$$U_k(y) = \int_0^l u(x, y)\varphi_k(x)dx = (u, \varphi_k)(y)$$

из условий (18)–(20) получим задачу

$$\begin{gathered} -U_k''(y) + \lambda_k U_k(y) = (f, \varphi_k)h(y) = f_k h(y), \quad 0 < y < l, \\ U_k(q_1) = U_k(q_2) = 0, \quad U_k(0) = (\chi, \varphi_k) = \chi_k. \end{gathered} \tag{21}$$

Из (21), пользуясь функцией Грина, получаем условия для определения f_k:

$$f_k \kappa(\lambda_k) = \chi_k. \tag{22}$$

Из (22) следует, что если $\kappa(\lambda_k) \ne 0$, то коэффициенты Фурье функции $f(x)$ по системе $\{\varphi_k\}$, образующей базис в $L_2(q_1, q_2)$, определяются однозначно, т.е. задача 1_0 при $\chi = 0$ имеет только тривиальное решение. Если при некотором k выполнено равенство $\kappa(\lambda_k) = 0$, то непосредственно проверяется, что пара функций (u, φ_k), где u – решение прямой задачи (5), удовлетворяет задаче 1_0 при $\chi = 0$. Теорема 3 доказана.

Следствие 1. Пусть h, $h_{yy} \in C^\alpha[q_1, q_2]$, $h(y)$ не меняет знак на $[q_1, q_2]$, $h(0) \ne 0$, тогда задача (1) имеет единственное решение для любой тройки функций $(g, \varphi, \chi) \in G$.

Доказательство. Заметим, что если $h(y)$ не меняет знак на $[q_1, q_2]$, $h(0) \ne 0$, то $\kappa(\lambda) \ne 0$ для любых $\lambda > 0$. Отсюда, используя теорему 6, получаем утверждение следствия.

Следствие 2. Пусть (u, f) – решение задачи 1_0, $u(x, 0) = \chi(x)$. Тогда если $\kappa(\lambda_k) \ne 0$, то неизвестная функция f может быть представлена в виде сходящегося в $L_2(0, l)$ ряда:

$$f(x) = \sum_1^\infty \frac{(\chi, \varphi_k)}{\kappa(\lambda_k)}\varphi_k(x). \tag{23}$$

Доказательство следует из (22).

4. ТЕОРЕМА ЕДИНСТВЕННОСТИ И СУЩЕСТВОВАНИЯ ДЛЯ ЗАДАЧИ 2

Задачу 2 будем рассматривать в предположении, что $\varphi(x, q_1) = \varphi(x, q_2) = 0$, $x \in [0, l]$, т.е. рассмотрим задачу определения пары функций $(u, c) \in \bar{C}^{2+\alpha}(\Pi) \times \bar{C}_-^\alpha(0, l)$ из условий

$$\begin{gathered} -\Delta u(x, y) = c(x)u(x, y) + g(x, y), \quad (x, y) \in \Pi, \\ u(0, y) = \nu_1(y), \quad u(l, y) = \nu_2(y), \quad y \in [q_1, q_2], \quad u(x, q_1) = u(x, q_2) = 0, \quad x \in [0, l], \end{gathered} \tag{24}$$

$$u(x, 0) = \chi(x), \quad x \in [0, l]. \tag{25}$$

Теорема 4 (единственности для задачи 2). *Пусть* g, $g_{yy} \in C^\alpha(\bar{\Pi})$, ν_i, $\nu_i'' \in C[q_1, q_2]$, $\nu_i(y) \ge 0$, $\nu_i''(y) \le 0$, $y \in (q_1, q_2)$ $i = 1, 2$, $g(x, y) \ge 0$, $g_{yy}(x, y) \le 0$, $(x, y) \in \Pi$, *хотя бы одна из функций* g, ν_i *не является тождественным нулем. Тогда задача* 2 *не может иметь двух различных решений.*

Доказательство (от противного). Пусть существуют два различных решения $(u^{(2)}, c^{(2)})$, $(u^{(1)}, c^{(1)})$ задачи (24), (25). Тогда их разность $(u, c) = (u^{(2)} - u^{(1)}, c^{(2)} - c^{(1)})$ – решение обратной задачи

$$\begin{gathered} -\Delta u = c^{(2)}u + cu^{(1)}, \quad (x, y) \in \Pi, \quad u(0, y) = u(l, y) = 0, \quad y \in [q_1, q_2], \\ u(x, q_1) = u(x, q_2) = 0, \quad x \in [0, l], \quad u(x, 0) = 0, \quad x \in [0, l]. \end{gathered} \tag{26}$$

Но функция $w(x, y) = u_{yy}^{(1)}(x, y)$ в силу леммы 3 удовлетворяет прямой задаче

$$-\Delta w = c^{(1)}w + g_{yy}, \quad (x, y) \in \Pi, \quad w(0, y) = \nu_1''(y), \quad w(l, y) = \nu_2''(y), \quad y \in (q_1, q_2),$$
$$w(x, q_1) = -g(x, q_1), \quad w(x, q_2) = -g(x, q_2), \quad x \in (0, l). \tag{27}$$

Из принципа максимума для функции $u^{(1)}$ следует неравенство $u^{(1)}(x, y) \geq 0$, $(x, y) \in \Pi$. Кроме того, так как одна из функций g, ν_1, ν_2 не есть тождественный нуль, то из строгого принципа максимума (см. [1, с. 41]) следует, что $u^{(1)}(x, 0) = \chi(x) > 0$, $x \in (0, l)$. Для прямой задачи (27) из принципа максимума следует, что $w(x, y) \leq 0$, $(x, y) \in \Pi$. Но тогда применяя лемму 5 к обратной задаче (26), получаем, что $c \equiv 0$, т.е. $u \equiv 0$. Противоречие. Теорема 4 доказана.

Для формулировки теоремы существования для задачи 2 определим функцию $\bar{w} \in C^2(\Pi) \cap C(\Pi^*)$ как решение прямой задачи

$$-\Delta \bar{w} = g_{yy}(x, y), \quad (x, y) \in \Pi, \quad \bar{w}(0, y) = \nu_1''(y), \quad \bar{w}(l, y) = \nu_2''(y), \quad y \in (q_1, q_2),$$
$$\bar{w}(x, q_1) = \bar{w}(x, q_2) = 0, \quad x \in (0, l). \tag{28}$$

Теорема 5. *Пусть* $\nu_1, \nu_2 \in C^2[0, l]$, $g, g_{yy} \in C^\alpha(\overline{\Pi})$, $\chi \in \overline{C}^{2+\alpha}(0, l)$, $\nu_i(y) \geq 0$, $\nu_i''(y) \leq 0$, $y \in (q_1, q_2)$, $g(x, y) \geq 0$, $g_{yy}(x, y) \leq 0$, $(x, y) \in \Pi$, $\chi(x) \geq \chi_0 > 0$, $x \in [0, l]$, *выполнены условия согласования* $\chi(0) = \nu_1(0)$, $\chi(l) = \nu_2(0)$, $\nu_1(q_1) = \nu_1(q_2) = \nu_2(q_1) = \nu_2(q_2) = 0$. *Тогда если выполнено неравенство*

$$-\chi''(x) - g(x, 0) \leq \bar{w}(x, 0), \quad x \in (0, l), \tag{29}$$

где $\bar{w}(x, y)$ *– решение задачи* (28), *то задача* 2 *имеет, и притом единственное, решение.*

Замечание. Теорема 5 накладывает большое количество ограничений на задаваемые функции ν_1, ν_2, g, χ. Возникает вопрос о существовании такого набора функций, которые удовлетворяют всем условиям теоремы 5 (непустота класса). Приведем соответствующий пример. Пусть произвольные функции $\chi \in \overline{C}^{2+\alpha}(0, l)$, $\nu_1, \nu_2 \in C^2[q_1, q_2]$ удовлетворяют условиям $\chi(x) \geq \chi_0 > 0$, $x \in [0, l]$, $\nu_1(y) \geq 0$, $\nu_2(y) \geq 0$, $\nu_1''(y) \leq 0$, $\nu_2''(y) \leq 0$, $\nu_1(q_1) = \nu_1(q_2) = \nu_2(q_1) = \nu_2(q_2) = 0$, $\nu_1(0) = \chi(0)$, $\nu_2(0) = \chi(l)$. Ясно, что такие функции существуют. Пусть $r \in C^{2+\alpha}(\overline{\Pi})$ – произвольная функция, удовлетворяющая условиям $r_{yy} \leq 0$, $(x, y) \in \Pi$. Пусть $g(x, y) = Q + r(x, y)$, где $Q > 0$ – некоторое число, которое выберем дальше. Покажем, что при достаточно большом Q набор функций ν_1, ν_2, χ, g удовлетворяет всем условиям теоремы 5. Рассмотрим функцию $\bar{w} = Qw_1 + w_2$, где для функций $\bar{w}$, w_1, w_2 имеют место задачи

$$-\Delta \bar{w} = Qr_{yy}, \quad (x, y) \in \Pi, \quad \bar{w}(0, y) = \nu_1''(y), \quad \bar{w}(l, y) = \nu_2''(y), \quad y \in (q_1, q_2),$$
$$\bar{w}(q_1, x) = -Q - r(x, q_1), \quad \bar{w}(q_2, x) = -Q - r(x, q_2), \quad x \in (0, l), \tag{30}$$

$$-\Delta w_1 = 0, \quad (x, y) \in \Pi, \quad w_1(0, y) = w_1(l, y) = 0, \quad q_1 < y < q_2,$$
$$w_1(x, q_1) = w_1(x, q_2) = -1, \quad 0 < x < l, \tag{31}$$

$$-\Delta w_2 = r_{yy}, \quad (x, y) \in \Pi, \quad w_2(0, y) = \nu_1''(y), \quad w_2(l, y) = \nu_2''(y), \quad y \in (q_1, q_2),$$
$$w_2(x, q_1) = -r(x, q_1), \quad w_2(x, q_2) = -r(x, q_2), \quad 0 < x < l. \tag{32}$$

Покажем, что при достаточно больших значениях параметра Q набор функций g, ν_1, ν_2, χ удовлетворяет всем условиям теоремы 4. Заметим, что, в силу строгого принципа максимума, для функции w_1 выполнено неравенство $-1 < -\delta \leq w_1(x, 0) \leq 0$, $\delta > 0$ – некоторое фиксированное число. Тогда будет справедлива цепочка неравенств

$$-\chi''(x) - g(x, 0) - \bar{w}(x, 0) = -\chi''(x) - Q[1 + w_1(x, 0)] - [r(x, 0) + w_2(x, 0)] \leq$$
$$\leq -\chi''(x) - [r(x, 0) + w_2(x, 0)] - Q(1 - \delta) \leq 0.$$

Последнее неравенство будет всегда выполнено при достаточно больших значениях параметра $Q > 0$, т.е. будет выполнено неравенство (29).

Доказательство теоремы 5. Рассмотрим вспомогательную задачу об определении тройки функций $(u, v, c) \in \bar{C}^{2+\alpha}(\Pi) \times (C^{2+\alpha}(\Pi) \cap C(\Pi^*)) \times \bar{C}^{\alpha}(0, l)$ из условий

$$-\Delta u(x, y) = c(x)u(x, y) + g(x, y), \quad (x, y) \in \Pi, \quad u(0, y) = \nu_1(y), \quad u(l, y) = \nu_2(y), \quad y \in [q_1, q_2],$$
$$u(x, q_1) = u(x, q_2) = 0, \quad x \in [0, l], \tag{33}$$

$$-\Delta v(x, y) = c(x)v(x, y) + g_{yy}(x, y), \quad (x, y) \in \Pi, \quad v(0, y) = \nu_1''(y), \quad v(l, y) = \nu_2''(y), \quad y \in (q_1, q_2),$$
$$v(x, q_1) = -g(x, q_1), \quad v(x, q_2) = -g(x, q_2), \tag{34}$$

$$c(x) = [-\chi''(x) - g(x, 0) - v(x, 0)]/\chi(x), \quad 0 \le x \le l. \tag{35}$$

Задачу 2 и задачу (33)–(35) связывает следующая

Лемма 6. *Пусть* (u, c) *– решение задачи* 2; *тогда* (u, u_{yy}, c) *– решение задачи* (33)–(35). *Обратно*: *пусть* (u, v, c) *– решение задачи* (33)–(35). *Тогда* $v = u_{yy}$ *и пара функций* (u, c) *– решение задачи* 2.

Доказательство леммы 6 дословно совпадает с доказательством леммы 3 (отличие только в том, что $g(x, y) = f(x)h(x, y)$, что не меняет доказательства).

Для доказательства теоремы 5 определим на множестве $\bar{C}_{-}^{\alpha}(0, l) \subset C_{-}[0, l]$ нелинейный оператор со значениями в $C[0, l]$ по правилу $B(c)(x) = [-v(x, 0) - \chi''(x) - g(x, 0)]/\chi(x)$, $x \in [0, l]$, где $v(x, y)$ – решение прямой задачи (34). Из оценки $v(x, 0)$, следующей из принципа максимума (см. [1, с. 39]), вытекает, что $\|B(c_2) - B(c_1)\| \le K\|c_2 - c_1\|$, где K не зависит от $c_1, c_2 \in \bar{C}_{-}^{\alpha}(0, l)$. Поэтому $B(c)$ равномерно непрерывен на $C_{-}[0, l]$ и задан на всюду плотном в $C_{-}[0, l]$ множестве $\bar{C}_{-}^{\alpha}(0, l)$. Отсюда следует, что его можно однозначно продолжить по непрерывности на все множество $C_{-}[0, l]$ (см., например, [13, с. 83]). Продолженный таким образом на $C_{-}[0, l]$ оператор будем обозначать через $\tilde{B}(c)$. При этом для $\tilde{B}(c)$ будут выполнены оценки, следующие из принципа максимума для задачи (34), так как они выполняются для допредельных функций.

Лемма 7. *В условиях теоремы* 5 *оператор* $\tilde{B}(c)$ *вполне непрерывен*, *при этом для любой функции* $c \in C_{-}[0, l]$ *выполнено включение* $\tilde{B}(c) \in \bar{C}_{-}^{\alpha}(0, l)$.

Рассмотрим равенство (35) как уравнение II рода в полном метрическом пространстве $C_{-}[0, l]$ относительно неизвестной функции c:

$$c = \tilde{B}(c). \tag{36}$$

Если уравнение (36) имеет решение $c \in C_{-}[0, l]$, то, в силу леммы 7, выполнено включение $c \in \bar{C}_{-}^{\alpha}(0, l)$. Но тогда, решив прямую задачу (33), получим, в силу леммы 6, решение обратной задачи 2. Обратно: для любого решения задачи 2, в силу леммы 6, функция $c(x)$ будет решением уравнения (36). Отсюда следует, что для доказательства теоремы 5 достаточно доказать существование решения уравнения (36). Для этого рассмотрим в банаховом пространстве $C[0, l]$ множество

$$P = \{c : c \in C[0, l], -[\chi''(x) + g(x, 0)]/\chi(x) \le c(x) \le 0, x \in [0, l]\}.$$

Из принципа максимума для задач (28), (34) и из условия теоремы 5 следует, что $\tilde{B}(P) \subset P$, т.е. оператор $\tilde{B}$ вполне непрерывный, отображающий выпуклое замкнутое множество P в банаховом пространстве $C[0, l]$ в себя. Но тогда из известного принципа Шаудера (см. [13, с. 228]) следует, что $\tilde{B}$ имеет неподвижную точку и уравнение (36) имеет решение $c \in C_{-}[0, l]$. Но, в силу леммы 7, $c \in \bar{C}_{-}^{\alpha}(0, l)$. Тогда, в силу предыдущих рассуждений, задача 2 имеет решение, которое, согласно теореме 4, будет единственно. Теорема 5 доказана.

5. ДОКАЗАТЕЛЬСТВА ВСПОМОГАТЕЛЬНЫХ УТВЕРЖДЕНИЙ

Доказательство леммы 1. Пусть теорема 1 верна для задачи 1 и для этой задачи реализуется первый случай альтернативы. Тогда при $q = 0$, $\varphi = 0$, в силу условий согласования и теоремы 0, будет выполнено включение $\chi \in \bar{C}_0^{2+\alpha}(0, l)$. Но тогда для задачи 1_0 также будет выполнен первый случай альтернативы Фредгольма. Если для задачи 1 реализуется второй случай альтернативы, т.е. существует нетривиальное решение задачи 1, то оно будет нетривиальным решением задачи 1_0. Обратно, пусть альтернатива Фредгольма верна для задачи 1_0, для которой реализуется первый случай альтернативы. В силу линейности задачи 1, отсюда следует единственность ее решения (если оно существует). Существование решения задачи 1 в этом случае следует из формулы

$$(u, f) = (u_1, 0) + (u_2, f). \tag{37}$$

В (37) величина u_1 – решение следующей прямой задачи (существующее по теореме 0):

$$-\Delta u_1(x, y) = c(x)u_1(x, y) + g(x, y), \quad (x, y) \in \Pi, \quad u_1(x, y) = \varphi(x, y), \quad (x, y) \in \partial\Pi. \tag{38}$$

Пара функций (u_2, f) – решение следующей обратной задачи 1_0:

$$-\Delta u_2(x, y) = c(x)u_2(x, y) + f(x)h(x, y), \quad (x, y) \in \Pi, \quad u_2(x, y) = 0, \quad (x, y) \in \partial\Pi,$$

$$u_2(x, 0) = \chi(x) - u_1(x, 0) = \chi_1(x), \quad x \in [0, l].$$

Из теоремы 0 и краевых условий для задачи 1 следует, что $\chi_1 \in \overline{C}^{2+\alpha}(0, l)$. Непосредственно проверяется, что пара функций (u, f), определенная по формуле (37), удовлетворяет всем необходимым условиям и является решением задачи 1. Лемма 1 доказана.

Доказательство леммы 2. Покажем сначала, что указанный в определении поточечный предел существует для любой точки $(x_0, y_0) \in \Pi^*$. В силу линейности задачи (7), (8) и теоремы 00, без ограничения общности считаем, что, например, $\nu_2 \equiv 0$, $\mu_1 \equiv \mu_2 \equiv 0$, $\nu_1(y) = 0$, $y \le 0$, $\nu_1 \in C[q_1, q_2]$. Рассмотрим последовательность краевых задач

$$\begin{gathered}-\Delta u^{(n)} = cu^{(n)}, \quad (x, y) \in \Pi, \quad u^{(n)}(0, y) = \nu_1^{(n)}(y), \quad u^{(n)}(l, y) = 0, \quad y \in [q_1, q_2], \\ u^{(n)}(x, q_1) = u^{(n)}(x, q_2) = 0, \quad x \in [0, l].\end{gathered} \tag{39}$$

В силу теоремы 00, решение задачи (39) существует в классе $C^{2+\alpha}(\Pi) \cap C(\overline{\Pi})$. Покажем, что для любого $\rho > 0$ последовательность функций $u^{(n)}$ сходится равномерно на множестве $\overline{\Pi}_\rho = \{(x, y) : (x, y) \in \overline{\Pi}, x^2 + (y - q_2)^2 \ge \rho^2\}$. Для этого воспользуемся критерием Коши сходимости в банаховом пространстве $C(\overline{\Pi}_\rho)$. Рассмотрим последовательность функций $w^{(n, m)} = u^{(m)} - u^{(n)}$, $m > n$. Для функций $w^{(n, m)}$ верны следующие краевые задачи:

$$-\Delta w^{(n, m)} = cw^{(n, m)}, \quad (x, y) \in \Pi, \quad w^{(n, m)}(0, y) = \nu^{(m)}(y) - \nu^{(n)}(y) = \nu^{(n, m)}(y),$$

$$w^{(n, m)}(l, y) = 0, \quad y \in [q_1, q_2], \quad w^{(n, m)}(x, q_1) = w^{(n, m)}(x, q_2) = 0, \quad x \in [0, l].$$

Пусть $\nu^+(y) = \max\{0, \nu^{(n, m)}(y)\}$, $\nu^-(y) = \nu^+(y) - \nu(y)$, $y \in [q_1, q_2]$. Рассмотрим краевые задачи для уравнения Лапласа:

$$\begin{gathered}-\Delta w_\pm(x, y) = 0, \quad (x, y) \in \Pi, \quad w_\pm(0, y) = \nu^\pm(y), \quad w_\pm(l, y) = 0, \quad y \in [q_1, q_2], \\ w_\pm(x, q_1) = w_\pm(x, q_2) = 0, \quad x \in [0, l].\end{gathered} \tag{40}$$

Из принципа максимума для задач (39), (40) следуют неравенства

$$-w_-(x, y) \le w^{(n, m)} \le w_+(x, y), \quad (x, y) \in \overline{\Pi}.$$

Для любых (x, y), $x > 0$, определим функции $\overline{w}_\pm(x, y)$ по формуле Пуассона (считаем, что $\nu^\pm$ доопределены нулем вне отрезка $[q_1, q_2]$):

$$\overline{w}_\pm(x, y) = \frac{x}{\pi}\int\limits_{-\infty}^{+\infty} \nu^\pm(\xi)\frac{d\xi}{x^2 + (\xi - y)^2}.$$

Известно (см. [15, с. 76]), что функции $\overline{w}_\pm$ удовлетворяют следующим краевым задачам:

$$-\Delta\overline{w}_\pm(x, y) = 0, \quad x > 0, \quad y \in \mathbb{R}, \quad \overline{w}_\pm(0, y) = \nu^\pm(y), \quad y \in \mathbb{R}.$$

Из принципа максимума для краевых задач (40) получаем оценки

$$w_+(x,y) \le \bar{w}_+(x,y), \quad w_-(x,y) \le \bar{w}_-(x,y), \quad (x,y) \in \Pi.$$

Из этих оценок следует цепочка неравенств

$$\left|w^{(n,m)}(x,y)\right| \le \max\{w^+(x,y), w^-(x,y)\} \le$$

$$\le \max\left\{\frac{x}{\pi}\int_{-\infty}^{+\infty}\nu^+(\xi)\frac{d\xi}{x^2+(\xi-y)^2}, \frac{x}{\pi}\int_{-\infty}^{+\infty}\nu^-(\xi)\frac{d\xi}{x^2+(\xi-y)^2}\right\} \le \frac{4l\|\nu\|}{\pi n\rho^2}, \quad (x,y) \in \bar{\Pi}_\rho.$$

В пояснении нуждается только последнее неравенство цепочки, следующее из неравенства треугольника при $n > 2/\rho$. Отсюда следует равномерная сходимость последовательности функций $u^{(n)}$ на множестве $\bar{\Pi}_\rho$, $\rho > 0$, откуда имеем $u \in C(\Pi^*)$. Пусть теперь $\bar{\Pi}_k = \{(x,y) : 1/k \le x \le l - 1/k$, $q_1 + 1/k \le y \le q_2 - 1/k\}$, $\bar{\Pi}_k \subset \Pi$ при достаточно больших k. Заметим, что, в силу известных результатов (см., например, [1, с. 97]) в каждом из прямоугольников $\bar{\Pi}_k$ будут выполнены оценки Шаудера $\|u^{(n)}\|^{2+\alpha} \le C_2(k)\|u^{(n)}\| \le C_2(k)\|u\|$. Отсюда, используя лемму Арцела, стандартным образом (см., например, [1, с. 30]) получаем, что $u \in C^{2+\alpha}(\Pi)$ и удовлетворяет задаче (7), (8). Так как допредельная функция удовлетворяет оценке, следующей из принципа максимума, то, переходя к пределу в неравенстве, получаем требуемую оценку. Лемма 2 доказана.

Доказательство леммы 3. Рассмотрим прямую задачу (11), (12) при фиксированном $f \in C^\alpha[0,l]$. Доказательство того, что $v = u_{yy}$, где u – решение прямой задачи (10), дословно совпадает с доказательством леммы 6 из [11]. В [11] коэффициент $c \equiv 0$ и дополнительно предполагалось, что $f(0) = f(l) = 0$. Но если $c = c(x) \le 0$, то указанное доказательство повторяется без изменений. Условие $f(0) = f(l) = 0$ в лемме 6 из [11] использовалось только в одном месте – утверждении о существовании решения прямой задачи (11), (12). Но это существование вытекает из следствия к лемме 2. Таким образом, если u – решение прямой задачи (10), v – решение прямой задачи (11), (12), то $v = u_{yy}$. Отсюда, из условия (6) и уравнения в условиях (5), записанного при $y = 0$, следует, что (u, v, f) – решение задачи (10)–(13). Покажем, что если (u, v, f) – решение задачи (10)–(13), то (u, f) – решение задачи 1_0. Для этого достаточно показать, что $u(x,0) = \chi(x)$. Докажем от противного. Пусть $u(x,0) = \chi_1(x)$. Тогда, в силу того что $v = u_{yy}$, для задачи (10) вместе с условием (13) будет выполнено условие

$$f(x) = -\chi_1''(x) - c(x)\chi_1(x) - v(x,0), \quad x \in [0,l]. \tag{41}$$

Вычитая (41) из (13) получаем, что функция $\psi = \chi_1 - \chi$ удовлетворяет краевой задаче для обыкновенного дифференциального уравнения

$$-\psi''(x) - c(x)\psi(x) = 0, \quad x \in [0,l], \quad \psi(0) = \psi(l) = 0.$$

Отсюда в силу условия $c \le 0$ получаем, что $\psi \equiv 0$. Получили противоречие. Лемма 3 доказана.

Доказательство леммы 4. Пусть сначала полагаем $f \in \bar{C}^\alpha(0,l)$. Тогда имеем $(Af)(x) = w(x,0)$, где $w \in C^2(\Pi) \cap C(\Pi^*)$ – решение прямой задачи

$$\begin{gathered} -\Delta w(x,y) = c(x)w(x,y) + f(x)h_{yy}(x,y), \quad (x,y) \in \Pi, \quad w(x,q_1) = -f(x)h(x,q_1), \\ w(x,q_2) = -f(x)h(x,q_2), \quad x \in (0,l), \quad w(0,y) = w(l,y) = 0, \quad y \in (q_1,q_2). \end{gathered} \tag{42}$$

Представим решение задачи (42) в виде $w = w^{(1)} + w^{(2)}$, где $w^{(1)}$, $w^{(2)}$ – решения краевых задач

$$\begin{gathered} -\Delta w^{(1)} = cw^{(1)}, \quad (x,y) \in \Pi, \quad w^{(1)}(x,q_1) = -f(x)h(x,q_1), \quad w^{(1)}(x,q_2) = -f(x)h(x,q_2), \quad x \in (0,l), \\ w^{(1)}(0,y) = w^{(1)}(l,y) = 0, \quad y \in (q_1,q_2), \end{gathered} \tag{43}$$

$$\begin{gathered} -\Delta w^{(2)} = cw^{(2)} + f(x)h_{yy}(x,y), \quad (x,y) \in \Pi, \quad w^{(2)}(x,q_1) = w^{(2)}(x,q_2) = 0, \quad x \in (0,l), \\ w^{(2)}(0,y) = w^{(2)}(l,y) = 0, \quad y \in (q_1,q_2). \end{gathered} \tag{44}$$

Определим линейные операторы A_1, A_2, $D(A_i) = \bar{C}^\alpha(0,l) \subset C[0,l]$, $i = 1, 2$, по правилу

$$(A_1 f)(x) = w^{(1)}(x,0), \quad (A_2 f)(x) = w^{(2)}(x,0), \quad x \in [0,l].$$

Так как для каждой из функций $w^{(1)}$, $w^{(2)}$ имеют место оценки, следующие из принципа максимума, то каждый из операторов A_1, A_2 имеет продолжение с $\bar{C}^{\alpha}(0, l)$ на все пространство $C[0, l]$. Будем обозначать эти продолжения через $\tilde{A}_1$, $\tilde{A}_2$. Очевидно, что $\tilde{A} = \tilde{A}_1 + \tilde{A}_2$. В соответствии с оценкой Шаудера вплоть до границы, из [1, с. 96] следует оценка (постоянная C_3 не зависит от f)

$$\|u^{(1)}(\cdot, 0)\| + \|u_x^{(1)}(\cdot, 0)\| \le C_3\|f\|.$$

Отсюда следует (по теореме Арцела), что $\tilde{A}_1$ – вполне непрерывный оператор и для любой функции $f \in C[0, l]$ выполнено включение $\tilde{A}_1 f \in \bar{C}^{\alpha}(0, l)$.

В случае оператора $\tilde{A}_2$ определим отрезки $l_k = [a_k, b_k]$, $a_{k+1} < a_k < b_k < b_{k+1}$, $(0, l) = \bigcup [a_k, b_k]$. Пусть $\|w^{(2)}(\cdot, 0)\|_k = \max\{|w^{(2)}(x, 0)|, x \in l_k\}$. Тогда если $f \in \bar{C}^{\alpha}(0, l)$, то в силу оценки из [1, с. 279] получим

$$\|w(\cdot, 0)\|_k + \|w_x(\cdot, 0)\|_k \le C_4(k)(\|c(\cdot)w^{(2)}(\cdot, 0)\| + \|f(\cdot)h_{yy}(\cdot, \cdot)\|) \le C_5(k)\|f\|. \tag{45}$$

Последнее неравенство в оценке (45) следует из принципа максимума для задачи (44), при этом $C_5(k)$ не зависит от f. Но тогда если $f \in C[0, l]$, $\{f_n\} \subset \bar{C}^{\alpha}(0, l)$, $f_n \longrightarrow f$, то из (45) получаем оценку

$$\|\tilde{A}_2 f\|_k + \|(\tilde{A}_2 f)_x\|_k \le C_5(k)\|f\|. \tag{46}$$

Из (46) и определения оператора $\tilde{A}_2$ следует, что для любой функции $f \in C[0, l]$ имеет место включение $\tilde{A}_2 f \in C^{\alpha}(0, l)$. Покажем, что $\tilde{A}_2$ – вполне непрерывный оператор. Для этого рассмотрим произвольную ограниченную в $C[0, l]$ последовательность $\{f_n\}$, $\|f_n\| \le C_6\}$. Определим последовательность $\{g_n\}$ в $C[0, l]$ по правилу $\tilde{A}_2 f_n = g_n$. Покажем, что из последовательности $\{g_n\}$ можно выбрать сходящуюся в $C[0, l]$ подпоследовательность. В силу оценки (46), последовательность $\{g_n\}$ на каждом из отрезков l_k равномерно ограничена и равностепенно непрерывна. На отрезке l_1, по теореме Арцела, из последовательности $\{g_n\}$ выделяем сходящуюся подпоследовательность g_{n_1}. На отрезке l_2 из подпоследовательности g_{n_1} выделяем сходящуюся подпоследовательность g_{n_2}, и т.д. Тогда "диагональная" последовательность g_{n_n} сходится на интервале $(0, l)$ к некоторой функции $g \in C(0, l)$. Кроме того, в силу оценки (46) выполнено включение $g \in C^{\alpha}(0, l)$. Покажем, что $g \in C[0, l]$. Для этого заметим, что в силу принципа максимума верна оценка

$$|(\tilde{A} f_n)(x)| \le \bar{w}(x, 0), \quad x \in (0, l). \tag{47}$$

В (47) функция $\bar{w} \in C(\Pi) \cap C^2(\Pi)$ – решение краевой задачи

$$-\Delta\bar{w}(x, y) = c(x)\bar{w}(x, y) + C_6\|h_{yy}\|, \quad (x, y) \in \Pi, \quad \bar{w}(x, q_1) = \bar{w}(x, q_2) = 0, \quad x \in [0, l],$$
$$\bar{w}(0, y) = \bar{w}(l, y) = 0, \quad y \in [q_1, q_2].$$

Из (47) следует, что существуют пределы

$$\lim_{x \to 0}(\tilde{A} f)(x) = \lim_{x \to l}(\tilde{A} f)(x) = 0,$$

т.е. $g \in \bar{C}^{\alpha}(0, l)$. Непосредственная проверка (с использованием неравенства треугольника в $C[0, l]$ и неравенства (47)) показывает, что $g_{n_n} \longrightarrow g$ в $C[0, l]$. Лемма 4 доказана.

Доказательство леммы 5 (от противного). Пусть $f \not\equiv 0$, u – решение задачи 1 с $g \equiv 0$, $\varphi \equiv 0$ и выполнено условие $u(x, 0) = 0$, $x \in [0, l]$. Представим функцию f в виде $f = f^+ - f^-$, где $f^+(x) = \max\{0, f(x)\}$, $f^-(x) = f^+(x) - f(x)$, $x \in [0, l]$. Ясно, что $f^+, f^- \in \bar{C}^{\alpha}(0, l)$. Пусть $u_{\pm}$ – решения следующих прямых задач (существующие по теореме 00):

$$-\Delta u_{\pm}(x, y) = c(x)u_{\pm}(x, y) + f^{\pm}(x)h(x, y), \quad (x, y) \in \Pi, \quad u_{\pm}(x, y) = 0, \quad (x, y) \in \partial\Pi. \tag{49}$$

Из условия $u(x, 0) = 0$ следует, что для функций $u_\pm(x, 0)$ выполнены условия

$$u_\pm(x, 0) = \chi(x), \quad x \in [0, l]. \tag{50}$$

Рассмотрим задачи об определении функций $w_\pm \in C^2(\Pi) \cap C(\Pi^*)$ из условий

$$\begin{gathered} -\Delta w_\pm = c(x)w_\pm(x, y) + f^\pm(x)h_{yy}(x, y), \quad (x, y) \in \Pi, \quad w_\pm(0, y) = w_\pm(l, y) = 0, \quad y \in (q_1, q_2), \\ w_\pm(x, q_1) = -f^\pm(x)h(x, q_1), \quad w_\pm(x, q_2) = -f^\pm(x)h(x, q_2), \quad x \in (0, l). \end{gathered} \tag{51}$$

Используя уравнение в (49) и условие (50), для функций $w_\pm(x, 0)$ получаем условия

$$-\chi''(x) - w_\pm(x, 0) = c(x)\chi(x) + f^\pm(x)h(x, 0), \quad x \in (0, l). \tag{52}$$

В силу строгого принципа максимума, для задачи (49) получаем, что $\chi(x) > 0$, $x \in (0, l)$. Пусть $x_0 \in (0, l)$ – точка, в которой функция χ имеет максимум на отрезке $[0, l]$. Тогда из необходимого условия максимума следует, что выполнено неравенство $\chi''(x_0) \le 0$. Кроме того, из строгого принципа максимума, примененного к задачам (51), следует, что $w(x_0, 0) < 0$ (так как $h \not\equiv 0$). Тогда из (52) следует, что $f^\pm(x_0) > 0$. Получили противоречие. Лемма 5 доказана.

Доказательство леммы 7. Полагаем сначала $c \in \bar{C}_-^\alpha(0, l)$, функция $v \in C^2(\Pi) \cap C(\Pi^*)$ – решение краевой задачи (существующее по лемме 2)

$$\begin{gathered} -\Delta v(x, y) = c(x)v(x, y) + g_{yy}(x, y), \quad (x, y) \in \Pi, \quad v(0, y) = \nu_1''(y), \quad v(l, y) = \nu_2''(y), \quad y \in (q_1, q_2), \\ v(x, q_1) = -g(x, q_1), \quad v(x, q_2) = -g(x, q_2), \quad x \in (0, l). \end{gathered} \tag{53}$$

Тогда, по определению, выполнено равенство

$$B(c)(x) = [-v(x, 0) - \chi''(x) - g(x, 0)]/\chi(x), \quad x \in [0, l].$$

Аналогично доказательству леммы 5, пусть $(0, l) = \bigcup l_k$, $l_k = [a_k, b_k] \subset (0, l)$. Тогда, в силу оценки из [1, с. 277], получаем цепочку неравенств (пояснения ниже)

$$\begin{gathered} \|v_x(\cdot, 0)\|_k + \|v(\cdot, 0)\|_k \le C_7(k)(\|v\| + [(q_2 - q_1)^2 + l^2]\|cv + g_{yy}\|) \le \\ \le C_7(k)(\max\{\|\nu_1''\|, \|\nu_2''\|, \|g(\cdot, q_1)\|, \|g(\cdot, q_2)\|\}) + \\ + (\exp\{l\} - 1)\|g_{yy}\| + [(q_2 - q_1)^2 + l^2](\|c\|\|v\| + \|g_{yy}\|) \le C_7(k)(C_8\|c\| + C_9). \end{gathered} \tag{54}$$

В цепочке (54) первое неравенство – прямое следствие оценки из [1, с. 277], записанной для задачи (53). Второе неравенство цепочки следует из принципа максимума (см. [1, с. 39]), в третьем неравенстве еще раз использован принцип максимума и проведены элементарные преобразования (C_8, C_9, конечно, не зависят от k и c). Тогда из оценки (54) получаем (переходом к пределу в неравенстве), что для любой функции $c \in C_-[0, l]$ выполнено включение $\tilde{B}(c) \in \bar{C}^\alpha(0, l)$. При этом, в силу условия (29), из принципа максимума следует, что $\tilde{B}(c)(x) \le 0$, $x \in (0, l)$, т.е. $(\tilde{B})(c) \in \bar{C}_-^\alpha(0, l)$. Покажем, что из оценки (54) следует полная непрерывность оператора $\tilde{B}$. В самом деле, пусть $M \subset C_-[0, l]$ – ограниченное множество, т.е. для любой функции $c \in M$ выполнено неравенство $-m \le c(x) \le 0$, $x \in (0, l)$, $m > 0$. Пусть$\{c_n\} \subset M$. Рассмотрим последовательность $b_n = \tilde{B}(c_n)$. В силу оценки (54) и теоремы Арцелла, множество функций $\{b_n(x)\}$ равномерно ограничено и равностепенно непрерывно на отрезке l_1. Выбираем сходящуюся подпоследовательность $\{b_{n_1}\}$. На отрезке l_2 из последовательности $\{b_{n_1}\}$ выбираем сходящуюся подпоследовательность, и т.д. Тогда диагональная последовательность $\{b_{n_n}\}$ будет сходиться в $C(0, l)$ к некоторой функции $b \in C_-^\alpha(0, l)$. Для функции $v(x, 0)$, в силу принципа максимума, будут выполнены оценки

$$\underline{v}(x, 0) \le v(x, 0) \le \bar{v}(x, 0), \tag{55}$$

где функции $\underline{v}$, $\overline{v} \in C^2(\Pi) \cap C(\Pi^*)$ удовлетворяют задаче

$$-\Delta\overline{v}(x, y) = -m\overline{v}(x, y) + g_{yy}(x, y), \quad (x, y) \in \Pi, \quad \overline{v}(0, y) = \nu_1''(y), \quad \overline{v}(l, y) = \nu_2''(y), \quad y \in (q_1, q_2),$$
$$\overline{v}(x, q_1) = -g(x, q_1), \quad \overline{v}(x, q_2) = -g(x, q_2), \quad x \in (0, l). \tag{56}$$

$$-\Delta\underline{v}(x, y) = g_{yy}(x, y), \quad (x, y) \in \Pi,$$
$$\underline{v}(0, y) = \nu_1''(y), \quad \underline{v}(l, y) = \nu_2''(y), \quad y \in (q_1, q_2), \quad \underline{v}(x, q_1) = -g(x, q_1), \tag{57}$$
$$\underline{v}(x, q_2) = -g(x, q_2), \quad x \in (0, l).$$

Из оценки (55) и условий (56), (57) следует, что для функции $b(x)$ верна оценка

$$-\overline{v}(x, 0) - \chi''(x) - g(x, 0) \le b(x) \le -\underline{v}(x, 0) - \chi''(x) - g(x, 0), \quad x \in (0, l). \tag{58}$$

Из (58) следует, что $b \in C_-[0, l]$. Но тогда $b \in \overline{C}_-^{\alpha}(0, l)$, т.е. оператор $\tilde{B}(c)$ вполне непрерывен.

Автор выражает благодарность всем участникам семинара по обратным задачам на механико-математическом факультете МГУ под руководством В.А. Садовничего и А.И. Прилепко за заинтересованное обсуждение изложенных выше результатов.

СПИСОК ЛИТЕРАТУРЫ

1. *Гилбарг Д., Трудингер Н.* Эллиптические дифференциальные уравнения с частными производными второго порядка. М.: Наука, 1989.
2. *Ладыженская О.А., Уральцева Н.Н.* Линейные и квазилинейные уравнения эллиптического типа. М.: Наука, 1973.
3. *Алифанов О.М.* Идентификация процессов теплообмена летательных аппаратов (введение в теорию обратных задач). М.: Наука, 1979.
4. *Аниканов Ю.Е.* Некоторые методы исследования многомерных обратных задач. Новосибирск: Наука, 1978.
5. *Денисов А.М.* Введение в теорию обратных задач. М.: Наука, 1994.
6. *Романов В.Г.* Обратные задачи математической физики. М.: Наука, 1984.
7. *Тихонов А.Н., Леонов А.С., Ягола А.Г.* Нелинейные некорректные задачи. М.: Наука, 1995.
8. *Belov Yu.Ya.* Inverse problems for partial differential equations. Utrecht etc.: VSP 2002.
9. *Isakov V.* Inverse problems for partial differential equations. New York etc. Springer 1998.
10. *Prilepko A.I., Orlovsky D.G., Vasin I.A.* Methods for solving inverse problems in mathematical physics. New York, Basel: Marcel Dekker Inc, 2000.
11. *Соловьёв В.В.* Обратные задачи определения источника для уравнения Пуассона на плоскости // Ж. вычисл. матем. и матем. физ. 2004. Т. 44. № 5. С. 862–871.
12. *Миранда К.* Уравнения с частными производными эллиптического типа. М.: Изд-во иностр. лит., 1957.
13. *Хатсон В., Пим Дж.* Приложения функционального анализа и теории операторов. М.: Мир, 1983.
14. *Колдингтон Э.А., Левинсон Н.* Теория обыкновенных дифференциальных уравнений. М.: Изд-во иностр. лит., 1958.
15. *Бицадзе А.В.* Некоторые классы уравнений в частных производных. М.: Наука, 1981.
16. *Соловьев В.В.* Обратные задачи для эллиптических уравнений на плоскости. I // Дифференц. ур-ния 2006. Т. 42. № 8. С. 1106–1114.

ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 2007, том 47, № 8, с. 1378–1386

УДК 519.642.5

О СХОДИМОСТИ ЧИСЛЕННЫХ МЕТОДОВ РЕШЕНИЯ БИЛИНЕЙНОГО УРАВНЕНИЯ ВОЛЬТЕРРА I РОДА[1)]

© 2007 г. А. С. Апарцин

(664033 Иркутск, ул. Лермонтова, 130, Ин-т систем энергетики СО РАН)

e-mail: apartsyn@isem.sei.irk.ru

Поступила в редакцию 02.02.2007 г.

Дано обоснование сходимости квадратурных методов (правых и средних прямоугольников) численного решения билинейного уравнения Вольтерра I рода. Приводятся также численные результаты для тестовых примеров. Библ. 11. Табл. 5.

Ключевые слова: билинейное интегральное уравнение Вольтерра I рода, квадратуры правых и средних прямоугольников, сходимость, функция Ламберта, саморегуляризация.

ВВЕДЕНИЕ

В [1], [2] рассматривались численные методы решения билинейного уравнения Вольтерра I рода

$$\int_0^t K_1(t, s)\varphi(s)ds + \int_0^t\int_0^t K_2(t, s_1, s_2)\varphi(s_1)\varphi(s_2)ds_1ds_2 = f(t), \quad t \in [0, T], \tag{1}$$

которое возникает, когда нелинейная динамическая система типа вход–выход аппроксимируется квадратичным полиномом Вольтерра, ядра $K_1(t, s)$, $K_2(t, s_1, s_2)$ (K_2 симметрично по s_1, s_2) идентифицированы тем или иным способом (например, по методике из [3]) и ставится задача об определении такого входного сигнала $\varphi(t)$, которому соответствует заданный выход $f(t)$. На тестовых примерах было показано, что, как и в линейном случае, методы правых и средних прямоугольников сходятся по шагу сетки h с порядками $O(h)$ и $O(h^2)$ соответственно, а в случае возмущения исходных данных в метрике C они обладают саморегуляризующим свойством.

В этой работе дадим доказательство их сходимости. Хотя основные этапы доказательства сохраняются теми же, что и в линейном случае (см. [4]), их реализация существенно усложняется из-за того, что система сеточных уравнений относительно вектора ошибки содержит и само сеточное решение, а это требует установления его равномерной по шагу сетки ограниченности.

1. АНАЛИЗ СХОДИМОСТИ

Относительно исходных данных в (1) помимо нужной гладкости будем предполагать, что $K_1(t, t) \neq 0\ \forall t \in [0, T]$ и $f(0) = 0$. Введем обозначения

$$L_1 = \max_{0 \le s \le t \le T} |K'_{1_t}(t, s)| \ge 0, \tag{2}$$

$$L_2 = \max_{0 \le s_1, s_2 \le t \le T} |K'_{2_t}(t, s_1, s_2)| \ge 0, \tag{3}$$

$$M_2 = \max_{0 \le s \le t \le T} |K_2(t, t, s)| > 0, \tag{4}$$

$$F = \max_{0 \le t \le T} |f'(t)| > 0, \tag{5}$$

$$\min_{0 \le t \le T} |K_1(t, t)| = k > 0. \tag{6}$$

[1)] Работа выполнена при финансовой поддержке РФФИ (код проекта 05-01-00336).

Не уменьшая общности, положим (для сокращения записи) $k = 1$. В [5]–[7] установлена точная верхняя оценка гарантированной для данных (k, L_1, M_2, L_2, F) области существования непрерывного решения $\bar{\varphi}(t)$ уравнения (1).

Так, при $L_1 = L_2 = 0$ имеем $\bar{\varphi}(t) \in C_{[0, T]}$, если $T < T^*$, где

$$T^* = \frac{1}{4M_2F}. \tag{7}$$

При $L_1 \neq 0$, $L_2 = 0$ имеем

$$T^* = \frac{L_1 + 2M_2F}{L_1^2}\ln\left(1 + \frac{L_1}{2M_2F}\right) - \frac{1}{L_1}. \tag{8}$$

Если, к примеру, $L_1 = F = 2M_2 = 1$, то $T^* = 2\ln 2 - 1 \approx 0.3863$. При $L_1 \neq 0$ и $L_2 \neq 0$ величина T^* зависит от знака дискриминанта квадратного трехчлена $L_2x^2 + L_1x + F$.

Например, при $D = L_1^2 - 4L_2F = 0$ имеем

$$T^* = b - \frac{2M_2}{L_2}(1 + \ln(a\sqrt{L_2})), \tag{9}$$

где

$$a = \frac{L_2 + M_2L_1}{2M_2L_2}, \quad b = \frac{M_2}{L_2}\ln F + \frac{2(L_2 + M_2L_1)}{L_2L_1}. \tag{10}$$

Так, если $F = 2M_2 = L_2 = 1$, $L_1 = 2$, то $D = 0$, $a = b = 2$ и, согласно (9), получаем $T^* = 1 - \ln 2 \approx 0.3069$. Формулы для T^* при $D \gtrless 0$ можно найти в [5]. Ниже предполагается, что условие $T < T^*$ выполнено.

Вводим сетку узлов $t_i = ih$, $i = \overline{1, n}$, $nh = T$ и, аппроксимируя интегралы в (1) квадратурами правых прямоугольников, записываем в очевидных обозначениях сеточный аналог уравнения (1):

$$h\sum_{j=1}^{i} K_{1_{i,j}}\varphi_j^h + h^2\sum_{j=1}^{i}\sum_{k=1}^{i} K_{2_{i,j,k}}\varphi_j^h\varphi_k^h = f_i, \quad i = \overline{1, n}. \tag{1^h}$$

Для получения сеточной аппроксимации функции $\bar{\varphi}(t)$ в i-м узле из $(1)^h$ имеем квадратное уравнение относительно φ_i^h, вещественность корней которого гарантирует неравенство $T < T^*$, а выбор нужного корня определяется условием

$$\varphi_1^h \xrightarrow[h \to 0]{} \bar{\varphi}(0) = \frac{f'(0)}{K_1(0, 0)}.$$

Обозначим через $\bar{\varphi}_i \equiv \bar{\varphi}(t_i)$, $i = \overline{1, n}$, каркас точного решения, а через ε^h – вектор ошибки сеточного решения:

$$\varepsilon^h = \{\varepsilon_i^h\} \equiv \{\bar{\varphi}_i - \varphi_i^h\}, \quad i = \overline{1, n}.$$

Так как вектор $\{\bar{\varphi}_i\}$, $i = \overline{1, n}$, удовлетворяет системе

$$h\sum_{j=1}^{i} K_{1_{i,j}}\bar{\varphi}_j + h^2\sum_{j=1}^{i}\sum_{k=1}^{i} K_{2_{i,j,k}}\bar{\varphi}_j\bar{\varphi}_k = f_i - R_i(\bar{\varphi}), \quad i = \overline{1, n}, \tag{11}$$

где

$$R_i(\bar{\varphi}) = \int_0^{t_i} K_1(t_i, s)\bar{\varphi}(s)ds + \int_0^{t_i}\int_0^{t_i} K_2(t_i, s_1, s_2)\bar{\varphi}(s_1)\bar{\varphi}(s_2)ds_1ds_2 - \left(h\sum_{j=1}^{i} K_{1_{i,j}}\bar{\varphi}_j + h^2\sum_{j=1}^{i}\sum_{k=1}^{i} K_{2_{i,j,k}}\bar{\varphi}_j\bar{\varphi}_k\right)$$

есть суммарная погрешность квадратуры по $[0, t_i]$ и $[0, t_i] \times [0, t_i]$, то, вычитая $(1)^h$ из (11), получаем относительно ε^h следующую систему линейных алгебраических уравнений (СЛАУ):

$$h\sum_{j=1}^{i} K_{1_{i,j}}\varepsilon_j^h + h^2\sum_{j=1}^{i}\sum_{k=1}^{i} K_{2_{i,j,k}}(\bar{\varphi}_j\varepsilon_k^h + \varphi_k^h\varepsilon_j^h) = -R_i(\bar{\varphi}), \quad i = \overline{1, n}. \tag{12}$$

Как отмечено во Введении, СЛАУ (12) от аналогичной СЛАУ в линейном случае принципиально отличает вхождение в (12) векторов $\{\bar{\varphi}_i\}$ и $\{\varphi_i^h\}$, $i = \overline{1, n}$. Для получения оценки погрешности численного решения необходимо предварительно оценить $|\bar{\varphi}(t)|$ и $|\varphi_i^h|$, $i = \overline{1, n}$.

При различных соотношениях констант L_1, L_2, M_2, F в [5]–[7] получены оценки вида

$$|\bar{\varphi}(t)| \le \Phi(T), \quad T < T^*. \tag{13}$$

Так, при $L_1 = L_2 = 0$ имеем

$$\Phi(T) = \frac{F}{\sqrt{1 - 4M_2FT}}.$$

При $L_1 \ne 0$, $L_2 = 0$ получаем

$$\Phi(T) = -\frac{L_1}{2M_2}\frac{W\left(-\frac{2M_2F}{L_1 + 2M_2F}\exp\frac{L_1^2T - 2M_2F}{L_1 + 2M_2F}\right)}{1 + W\left(-\frac{2M_2F}{L_1 + 2M_2F}\exp\frac{L_1^2T - 2M_2F}{L_1 + 2M_2F}\right)}. \tag{14}$$

В (14) W – главная вещественная ветвь функции Ламберта (см. [8], [9]).

При $L_1^2 - 4L_2F = 0$ имеем

$$\Phi(T) = \frac{aL_2}{2M_2}\frac{1}{W\left(-1, -a\sqrt{L_2}\exp\left[-\frac{L_2}{2M_2}(b - T)\right]\right)}\frac{1}{1 + W\left(-1, -a\sqrt{L_2}\exp\left[-\frac{L_2}{2M_2}(b - T)\right]\right)}. \tag{15}$$

В (15) a и b те же, что и в (10), а $W(-1, \ldots)$ – вторая вещественная ветвь функции Ламберта, определенная на $\left[-\frac{1}{e}, 0\right]$.

Замечание 1. Функция $\Phi(T)$ играет в теории билинейных уравнений Вольтерра I ту же роль, что и экспонента в линейной теории.

Чтобы показать, что $|\varphi_i^h|$ также удовлетворяют неравенству типа (13), нам понадобится следующая

Лемма 1. *Пусть X – множество решений числового неравенства*

$$ax \le b + cx^2, \quad x, a, b, c > 0.$$

Тогда при $a^2 > 4bc$ имеем

$$X = X_1 \cup X_2,$$

где

$$X_1 = \left\{x : x \le x^* = \frac{a - \sqrt{a^2 - 4bc}}{2c}\right\},$$

$$X_2 = \left\{x : x \ge x^{**} = \frac{a + \sqrt{a^2 - 4bc}}{2c}\right\},$$

а при $a^2 \le 4bc$ получаем $X = R_+$.

Доказательство очевидно и следует из геометрических соображений.

Теорема 1. *Справедливо неравенство*

$$\left|\varphi_i^h\right| \le \Phi(T), \quad i = \overline{1, n}. \tag{16}$$

Доказательство. Вычитая $(i-1)$-ю строку из i-й строки уравнения $(1)^h$, переходя к оценке по модулю и учитывая обозначения (2)–(5), получаем неравенство

$$\left|\varphi_i^h\right| \le F + L_1 h\sum_{j=1}^{i-1}\left|\varphi_j^h\right| + hM_2(\varphi_i^h)^2 + 2M_2\left|\varphi_i^h\right| h\sum_{j=1}^{i-1}\left|\varphi_j^h\right| + L_2\left(h\sum_{j=1}^{i-1}\left|\varphi_j^h\right|\right)^2, \quad i = \overline{2, n}.$$

Обозначим через x_i^* меньший корень сеточного уравнения

$$x_i = F + L_1 h\sum_{j=1}^{i-1}x_j + hM_2x_i^2 + 2M_2x_ih\sum_{j=1}^{i-1}x_j + L_2\left(h\sum_{j=1}^{i-1}x_j\right)^2, \quad i = \overline{2, n}. \tag{17}$$

Согласно лемме 1, при достаточно малом h имеем

$$\left|\varphi_i^h\right| \le x_i^*, \quad i = \overline{2, n}, \tag{18}$$

и

$$\left|\varphi_1^h\right| \le x_1^* = \frac{1-\sqrt{1-4M_2Fh}}{2M_2h}.$$

Если ввести замену

$$\theta_i = h\sum_{j=1}^{i}x_j, \quad \theta_0 = 0, \quad i = \overline{1, n}, \tag{19}$$

то получим

$$x_i = \frac{\theta_i - \theta_{i-1}}{h}, \quad i = \overline{1, n}, \tag{20}$$

и подстановка (19), (20) в (17) дает

$$\frac{\theta_i - \theta_{i-1}}{h} = \frac{F + L_1\theta_{i-1} + L_2\theta_{i-1}^2}{1 - M_2(\theta_i + \theta_{i-1})}, \quad \theta_0 = 0, \quad i = \overline{1, n}. \tag{21}$$

Разностная схема (21) аппроксимирует задачу Коши

$$\theta'(t) = \frac{F + L_1\theta(t) + L_2\theta^2(t)}{1 - 2M_2\theta(t)}, \quad \theta(0) = 0, \quad 0 < t \le T, \tag{22}$$

решением $\theta^*(t)$ которой и является (см. [5]–[7]) функция Ламберта или (при $L_1 = L_2 = 0$) функция

$$\theta^*(t) = \frac{1-\sqrt{1-4M_2Ft}}{2M_2},$$

а $\theta^{*\prime}(t) = \Phi(t)$, $t \in [0, T]$. Поскольку $x_i^* = (\theta_i^* - \theta_{i-1}^*)/h$, то справедливость теоремы 1 в силу (18) будет следовать из неравенства

$$x_i^* \le \theta^{*\prime}(t_i), \quad i = \overline{1, n}. \tag{23}$$

В частном случае, когда $L_1 = L_2 = 0$, погрешность разностной схемы (21) равна нулю, так как

$$\theta_i^* = \theta^*(t_i) = \frac{1-\sqrt{1-4M_2Fih}}{2M_2}, \quad i = \overline{1, n},$$

откуда, в силу монотонного возрастания $\theta^{*\prime}(t)$, имеем

$$x_i^* = \frac{\theta^*(t_i) - \theta^*(t_{i-1})}{h} = \theta^{*\prime}(\xi) < \theta^{*\prime}(t_i), \quad \xi \in [t_{i-1}, t_i],$$

и (23) выполняется. Справедливость (23) в общем случае следует тогда из явности числителя в (21). Теорема 1 доказана.

Замечание 2. Численные результаты показывают, что при замене в числителе (21) хотя бы в одном слагаемом θ_{i-1} на θ_i неравенство (23) может нарушаться.

Еще одно вспомогательное утверждение содержит

Лемма 2. *При достаточной гладкости исходных данных в* (1) *справедлива оценка*

$$|R_i(\bar{\varphi}) - R_{i-1}(\bar{\varphi})| \le ch^2, \quad i = \overline{2, n}, \quad c = \mathrm{const}.$$

Доказательство элементарно и вытекает из стандартной оценки погрешности квадратуры правых прямоугольников.

Теперь можно сформулировать основную теорему сходимости.

Теорема 2. *Пусть*

$$\gamma(T) = 2h_0 M_2 \Phi(T) + 2M_2 T\Phi(T) < 1. \tag{24}$$

Тогда при достаточной гладкости исходных данных в (1) *и* $h \le h_0$ *справедлива оценка*

$$\max_{1 \le i \le n} |\varepsilon_i^h| = O(h). \tag{25}$$

Доказательство. Заменяя i в (12) на $i - 1$, находим разность между соседними строками СЛАУ (12) с учетом симметрии ядра K_2 по второму и третьему аргументам:

$$\begin{gathered} h\varepsilon_i^h + h\sum_{j=1}^{i-1}(K_{1_{i,j}} - K_{1_{i-1,j}})\varepsilon_j^h + h^2 K_{2_{i,i,i}}(\bar{\varphi}_i + \varphi_i^h)\varepsilon_i^h + h^2(\bar{\varphi}_i + \varphi_i^h)\sum_{j=1}^{i-1} K_{2_{i,i,j}}\varepsilon_j^h + h^2\varepsilon_i^h\sum_{j=1}^{i-1}K_{2_{i,i,j}}(\bar{\varphi}_j + \varphi_j^h) + \\ + h^2\sum_{j=1}^{i-1}\sum_{k=1}^{i-1}(K_{2_{i,j,k}} - K_{2_{i-1,j,k}})(\bar{\varphi}_k + \varphi_k^h)\varepsilon_j^h = R_{i-1}(\bar{\varphi}) - R_i(\bar{\varphi}), \quad i = \overline{2, n}. \end{gathered} \tag{26}$$

Разделим (26) на h, перейдем к оценке по модулю и, учтя (2)–(5), (13), (16), получим

$$|\varepsilon_i^h| \le \frac{ch}{1 - 2hM_2\Phi(T) - 2M_2T\Phi(T)} + \frac{L_1 + 2M_2\Phi(T) + 2L_2T\Phi(T)}{1 - 2hM_2\Phi(T) - 2M_2T\Phi(T)} h\sum_{j=1}^{i-1}|\varepsilon_j^h|, \quad i = \overline{2, n}. \tag{27}$$

Кроме того, при $i = 1$ непосредственно из (12) имеем

$$|\varepsilon_1^h| \le \frac{|R_1(\bar{\varphi})|/h}{1 - 2hM_2\Phi(T)},$$

и так как $|R_1(\bar{\varphi})| \le ch^2$, то (27) выполняется и для $i = 1$.

Окончательно, усиливая (27) с учетом (24) и применяя разностный аналог леммы Гронуолла–Беллмана, приходим к неравенству

$$|\varepsilon_i^h| \le \frac{ch}{1 - \gamma(T)} e^{\frac{(L_1 + 2M_2\Phi(T) + 2L_2T\Phi(T))T}{1 - \gamma(T)}}, \quad i = \overline{1, n},$$

из которого следует (25).

Замечание 3. Если ввести сетку узлов $t_{i-1/2} = (i - 1/2)h$ и потребовать непрерывность $K_{1_{ts^2}}'''$, $K_{2_{ts_i^2}}'''$ и $\bar{\varphi}''(t)$, то приведенная техника обеспечивает доказательство сходимости метода средних прямоугольников с порядком $O(h^2)$. Детали доказательства опускаем. Аналогичный результат справедлив и для метода интегрирования произведения (product integration method), рассмотренного в линейном случае в [10] и особенно эффективного, если ядра K_1 и K_2 сильно осциллируют.

2. ЧИСЛЕННЫЕ РЕЗУЛЬТАТЫ

В качестве примера рассмотрим два билинейных уравнения, отличающихся лишь правой частью:

$$\int_0^t (1-(t-s))\varphi(s)ds - \frac{1}{2}\left(\int_0^t \varphi(s)ds\right)^2 = t - t^2, \tag{28}$$

$$\int_0^t (1-(t-s))\varphi(s)ds - \frac{1}{2}\left(\int_0^t \varphi(s)ds\right)^2 = t. \tag{29}$$

Легко проверить, что (единственным) непрерывным решением уравнения (28) является $\bar{\varphi}(t) = 1$, $t \in [0, T]$ $\forall T < \infty$, при этом $k = 1$, $L_1 = 1$, $M_2 = 1/2$, $L_2 = 0$ и, так как $\max\limits_{t \in [0,1]} |1 - 2t| = 1$, $F = 1$, если $T \le 1$.

Если исходные данные в (1) характеризовать пятеркой параметров k, L_1, M_2, L_2, F, то уравнению (28) соответствует набор (1, 1, 1/2, 0, 1). В [7] введено понятие мажорантного уравнения для (1). В частности, если $L_2 = 0$, то мажорантное уравнение имеет вид

$$\int_0^t (k - L_1(t-s))\varphi(s)ds - M_2\left(\int_0^t \varphi(s)ds\right)^2 = Ft. \tag{30}$$

Такое название объясняется тем, что точным решением уравнения (30) является мажоранта $\Phi(t)$ для $\bar{\varphi}(t)$, $t \in [0, T]$, $T < T^*$. Таким образом, (29) есть мажорантное уравнение для (28), причем (см. (14) и (8))

$$\Phi(t) = -\frac{W\left(-\frac{1}{2}e^{(t-1)/2}\right)}{1 + W\left(-\frac{1}{2}e^{(t-1)/2}\right)}, \quad t \in [0, T],$$

$$T^* = 2\ln 2 - 1 \approx 0.3863.$$

Так как $W(z)$ обратная к функции $z = We^W$, то $W\left(-\frac{1}{2}e^{-1/2}\right) = -\frac{1}{2}$, $\Phi(0) = 1$ $(=F)$, $\Phi(T^*) = \infty$. Функция $\Phi(t)$ – производная от решения $\theta^*(t)$ задачи Коши (см. (22))

$$\theta'(t) = \frac{1 + \theta(t)}{1 - \theta(t)}, \quad \theta(0) = 0, \quad 0 < t \le 2\ln 2 - 1,$$

$$\theta^*(t) = -2W\left(-\frac{1}{2}e^{(t-1)/2}\right) - 1.$$

Применительно к (28) система $(1)^h$ сводится к рекурсии

$$\theta_i = 1 - \sqrt{1 + \theta_{i-1}^2 - 2(1+h)\theta_{i-1} - 2(h - (2i-1)h^2)}, \tag{31}$$

$$\varphi_i^h = \frac{\theta_i - \theta_{i-1}}{h}, \quad i = \overline{1, n}, \quad \theta_0 = 0. \tag{32}$$

В табл. 1 приведены результаты расчетов по (31), (32) для $h = 0.1$, $h = 0.05$, $h = 0.01$.

Численные результаты показывают, что условие (24) теоремы 2, накладывающее ограничение на величину T, отражает существо дела. При $T \longrightarrow 1$ возникает погранслой ошибки сеточного решения. Неравенство (24) выполняется при $h_0 = 0.1$ и $T < 0.29$. Если, например, $T = 0.2$, то $\|\varepsilon^{h=0.05}\|_{C_h} = \max\limits_{1 \le i \le n}\left|\varepsilon_i^{h=0.05}\right| = 0.04256$, $\|\varepsilon^{h=0.01}\|_{C_h} = 0.00949$, что дает линейную сходимость по шагу сетки.

Таблица 1

t_i	$\varphi_i^{h\,=\,0.1}$	$\varphi_i^{h\,=\,0.05}$	$\varphi_i^{h\,=\,0.01}$
0.1	0.944614	0.969371	0.993302
0.2	0.924532	0.957443	0.990515
0.3	0.893969	0.938625	0.985950
0.4	0.845776	0.907417	0.977956
0.5	0.767188	0.852688	0.962672
0.6	0.636304	0.751265	0.929893
0.7	0.421267	0.557343	0.848292
0.8	0.095975	0.205154	0.610776
0.9	–0.317630	–0.310235	–0.006406
1.0	–0.734764	–0.835726	–0.943465

Таблица 2

t_i	$\varphi_i^{h\,=\,0.05}$	$\varphi_i^{h\,=\,0.01}$	$\bar{\varphi}_i = \Phi(t_i)$
0.1	1.142588	1.227300	1.251235
0.2	1.484858	1.642044	1.689001
0.3	2.188318	2.614438	2.763269
0.4	6.752919	1–0.04i	$\bar{\varphi}(0.3862) = \infty$

Применение квадратуры средних прямоугольников к (28) порождает рекурсию

$$\theta_i = \frac{2-h}{2} - \sqrt{\frac{(2-h)^2}{4} + \theta_{i-1}^2 - (2+h)\theta_{i-1} - 2(h-(2i-1)h^2)}, \tag{33}$$

$$\varphi_{i-1/2}^h = \frac{\theta_i - \theta_{i-1}}{h}, \quad i = \overline{1, n}, \quad \theta_0 = 0, \tag{34}$$

и в силу нулевой погрешности аппроксимации интегралов на точном решении, как и следовало ожидать, при любых T и h имеем $\theta_i = ih$ и $\varphi_{i-1/2}^h = \bar{\varphi}(t_{i-1/2}) = 1$, $i = \overline{1, n}$, $nh = T$.

Для мажорантного уравнения (29) вместо (31), (33) имеем, соответственно,

$$\theta_i = 1 - \sqrt{1 + \theta_{i-1}^2 - 2(1+h)\theta_{i-1} - 2h}, \quad \theta_0 = 0, \quad i = \overline{1, n}, \tag{31'}$$

$$\theta_i = \frac{2-h}{2} - \sqrt{\frac{(2-h)^2}{4} + \theta_{i-1}^2 - (2+h)\theta_{i-1} - 2h}, \quad \theta_0 = 0, \quad i = \overline{1, n}. \tag{33'}$$

В табл. 2 наряду с расчетами по (31)', (32) с $h = 0.05$ и $h = 0.01$ приведены значения точного решения $\bar{\varphi}(t_i) = \Phi(t_i)$.

Появление при $t = 0.4$ комплексного корня – следствие нарушения неравенства $T < 0.3862 = T^*$. Если ограничиться, как и для (28), значением $T = 0.2$, то получим $\|\varepsilon^{h\,=\,0.05}\|_{C_h} = 0.20414$, $\|\varepsilon^{h\,=\,0.01}\|_{C_h} = 0.46957$ и линейная сходимость имеет место.

Аналогичные данные для квадратуры средних прямоугольников (расчеты по (33)', (34)) содержатся в табл. 3.

При $T = 0.275$ имеем $\left\|\varepsilon^{h\,=\,\frac{1}{20}}\right\|_{C_h} = 0.02748$, $\left\|\varepsilon^{h\,=\,\frac{1}{60}}\right\|_{C_h} = 0.00297$, что соответствует квадратичной скорости сходимости.

Таблица 3

$t_{i-1/2}$	$\varphi_{i-1/2}^{h=1/20}$	$\varphi_{i-1/2}^{h=1/60}$	$\bar{\varphi}_{i-1/2} = \Phi(t_{i-1/2})$
0.075	0.179086	1.177151	1.176911
0.175	1.556047	1.550974	1.550346
0.275	2.387831	2.363315	2.360349
0.375	6.51–4.72i	9.815176	8.751915

Таблица 4

h	$\lVert\varepsilon_{rr}^{h}\rVert_{C_h}$	$\lVert\varepsilon_{mr}^{h}\rVert_{C_h}$
1/16	0.025235	1.67981×10^{-4}
1/32	0.014055	0.42012×10^{-4}
1/64	0.007412	0.10503×10^{-4}
1/128	0.003805	0.02626×10^{-4}
1/256	0.001928	0.00656×10^{-4}
1/512	0.000970	0.00164×10^{-4}

Таблица 5

δ	$h_{\text{opt}}(rr)$	$\lVert\tilde{\varepsilon}_{rr}^{h_{\text{opt}}}\rVert_{C_h}$	$h_{\text{opt}}(mr)$	$\lVert\tilde{\varepsilon}_{mr}^{h_{\text{opt}}}\rVert_{C_h}$
10^{-3}	0.075331	0.051597	0.276567	0.006879
10^{-4}	0.018333	0.019052	0.136363	0.001782
10^{-5}	0.006147	0.006224	0.063447	0.000429
10^{-6}	0.002006	0.001989	0.029079	0.000098
10^{-7}	0.000627	0.000631	0.013530	0.000022

Приведем еще результаты расчетов для уравнения

$$\int_0^t (1+t-s)\varphi(s)ds + \int_0^t\int_0^t (1+s_1+s_2)\varphi(s_1)\varphi(s_2)ds_1 ds_2 = f(t) = \\ = -\sin t - \cos t + 2 + 2\cos t \sin t + 2t - \cos^2 t - 2t\cos^2 t \tag{35}$$

с "хорошими" по сравнению с (28) ядрами K_1 и K_2. Точное решение уравнения (35) имеет вид $\bar{\varphi}(t) = \cos t$, $t \in [0, T]$ $\forall T < \infty$. Хотя (35) соответствует пятерка параметров (1, 1, 3, 0, 3.7), как показывает табл. 4 значений погрешностей для квадратур правых (ε_{rr}^h) и средних (ε_{mr}^h) прямоугольников при $T = 1$, погранслой ошибок в этом примере отсутствует.

Для иллюстрации саморегуляризующего эффекта процедуры дискретизации зададим пилообразное возмущение правой части (35):

$$\tilde{f}(t_i) = f(t_i) + (-1)^i\delta, \quad i = \overline{1, n}, \quad nh = T.$$

При фиксированном δ оптимизация по h унимодальных функций $\lVert\tilde{\varepsilon}_{rr}^h\rVert_{C_h}$ и $\lVert\tilde{\varepsilon}_{mr}^h\rVert_{C_h}$ осуществлялась методом Фибоначчи с десятью испытаниями (расчеты проведены Е.В. Марковой).

Видно, что сохраняются те же асимптотические оценки, что и в линейном случае (см. [4]): для правых прямоугольников имеем $h_{\text{opt}} \asymp \delta^{1/2}$, $\lVert\tilde{\varepsilon}_{rr}^{h_{\text{opt}}}\rVert_{C_h} \asymp \delta^{1/2}$, а для средних $h_{\text{opt}} \asymp \delta^{1/3}$, $\lVert\tilde{\varepsilon}_{mr}^{h_{\text{opt}}}\rVert_{C_h} \asymp \delta^{2/3}$.

Замечания. 4. Расчеты показывают, что рассмотренные в данной работе численные методы могут быть применены и для решения трилинейных уравнений, порождаемых кубичным полиномом Вольтерра, однако для обоснования их сходимости потребуется использовать аппарат соответствующих "ламбертообразных" функций и их сеточных аналогов.

5. Если в (1) положить $K_2 \equiv 0$, так что $M_2 = 0$, то (30) является мажорантным и для линейного уравнения (1); в этом случае имеем $|\bar{\varphi}(t)| \le \Phi(t) \equiv Fk^{-1}\exp\{L_1k^{-1}t\}$, $t \in [0, T]$, $T < \infty$.

6. В некоторых моделях динамики численности популяций важную роль играет управляющее воздействие, совпадающее с правой частью Ft мажорантных уравнений. Так, задача Коши

$$\dot{\xi}(t) = k\xi(t) - Ft, \quad t > 0, \quad \xi(0) = \varepsilon > 0, \quad k, F > 0,$$

моделирует поведение неустойчивого нулевого состояния динамической системы, характерное, например, для развития эпидемии, когда эффективность борьбы возрастает пропорционально времени t. Несложный анализ показывает, что при $F < \varepsilon k^2$ экспоненциальный рост $\xi(t)$ сохраняется. При $F = \varepsilon k^2$ рост становится линейным, а если $F > \varepsilon k^2$, то в момент времени

$$T^* = -\frac{1}{k}W\left(-1, -\frac{(F - \varepsilon k^2)e^{-1}}{F}\right) - 1$$

имеем $\xi(T^*) = 0$. Подробнее об этом см. [7], [11].

СПИСОК ЛИТЕРАТУРЫ

1. *Апарцин А.С., Маркова Е.В.* О численном решении билинейного уравнения Вольтерра I рода // Тр. XII Байкальской междунар. конф. Иркутск, 2001. Т. 4. С. 20–24.
2. *Apartsyn A.S., Markova E.V.* On numerical solution of the multylinear Volterra equations of the first kind // Proc. Internat. Conf. Comput. Math. Part 2. Novosibirsk, 2001. P. 322–326.
3. *Апарцин А.С.* О новых классах линейных многомерных уравнений I рода типа Вольтерра // Изв. вузов. Математика. 1995. № 11. С. 28–41.
4. *Апарцин А.С.* Неклассические уравнения Вольтерра I рода: теория и численные методы. Новосибирск: Наука, 1999.
5. *Апарцин А.С.* О билинейных уравнениях Вольтерра I рода // Оптимизация, управление, интеллект. 2004. Т. 2. № 8. С. 20–28.
6. *Апарцин А.С.* О полилинейных уравнениях Вольтерра I рода // Автоматика и телемехан. 2004. № 2. С. 118–125.
7. *Апарцин А.С.* К теории полилинейных уравнений Вольтерра I рода // Оптимизация, управление, интеллект. 2005. Т. 1. № 9. С. 5–27.
8. *Corless R.M., Gonnet G.H., Hare D.E.G., Jeffrey D.J.* Lambert's *W* function in maple // Maple Techn. Newsletter, 1993. № 9.
9. *Corless R.M., Gonnet G.H., Hare D.E. et al.* On the Lambert *W* function // Advanc. Comput. Maths. 1996. Vol. 5.
10. *Linz P.* Product integration method for Volterra integral equations of the first kind // BIT. 1971. V. 11. P. 413–421.
11. *Апарцин А.С., Щербинин М.С.* Моделирование динамики численности популяции на базе уравнений, мажорирующих билинейные уравнения Вольтерра I рода // Оптимизация, управление, интеллект. 2005. Т. 2. № 10. С. 37–44.

ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 2007, том 47, № 8, с. 1387–1401

УДК 519.6:531.33

СОСТАВНЫЕ КОМПАКТНЫЕ СХЕМЫ ВЫСОКОГО ПОРЯДКА ДЛЯ МОДЕЛИРОВАНИЯ ТЕЧЕНИЙ ВЯЗКОГО ГАЗА[1)]

 А. Д. Савельев

(119991 Москва, ул. Вавилова, 40, ВЦ РАН)

e-mail: savel@ccas.ru

Поступила в редакцию 15.11.2004 г.
Переработанный вариант 02.10.2006 г.

Рассмотрены дифференциальные схемы до 7-го порядка включительно для численного описания законов сохранения на криволинейных сетках. Они представляют собой комбинацию симметричных компактных разностей и ориентированных в соответствии с направлением характеристик операторов диффузного типа. Рассмотрены спектральные свойства схем, проведены расчеты ряда модельных задач. Разработанный метод численного интегрирования уравнений Навье–Стокса с применением двухпараметрической модели турбулентности применяется для моделирования двумерных течений вязкого газа. Библ. 21. Фиг. 11.

Ключевые слова: компактные аппроксимации, схемы высокого порядка, криволинейные координаты, турбулентные течения вязкого газа.

ВВЕДЕНИЕ

Постоянно повышающиеся требования к уровню научно-технических разработок делают актуальной задачу совершенствования используемых численных методов. В настоящее время они применяются при исследовании проблем турбулентности, аэродинамики, теплообмена, акустики и т.д. Улучшить точность проводимых расчетов призваны методы высокого порядка аппроксимации. Поскольку повышение порядка неизбежно ведет к укрупнению сеточного шаблона, одним из перспективных направлений разработки численных алгоритмов является использование компактных аппроксимаций. Для задач аэрогазодинамики ориентированные против потока компактные разности третьего порядка успешно применяются с середины 70-х годов (см. [1]).

В настоящее время существуют два подхода к применению компактных разностей высокого порядка на неразнесенных расчетных сетках. В первом случае это симметричные разности высокого порядка, дополненные различными фильтрами для подавления схемных осцилляций (см. [2]–[4]). Несмотря на кажущуюся простоту, реализация данного подхода при решении задач с достаточно сильными градиентами параметров потока может представлять серьезную проблему. Выбор механизма демпфирования паразитных осцилляций в ряде случаев составляет предмет отдельного исследования.

Другой подход состоит в использовании несимметричных компактных аппроксимаций, ориентированных против потока (см. [1], [5], [6]). Разработанный изначально для схем третьего порядка, он применяется для получения схем произвольного порядка на основе мультиоператорного принципа. В схемах этого класса содержится естественный диссипативный механизм подавления схемных осцилляций. К сожалению, использование подобных схем в областях с криволинейными границами вызывает вопросы, связанные, например, с точностью сохранения невозмущенного потока. Альтернативный вариант применения несимметричных компактных разностей предложен в [7] как консервативный метод конечных объемов высокого порядка аппроксимации.

Большого различия между двумя упомянутыми подходами нет. В [8] убедительно показано, что известные разности, ориентированные против потока, легко могут быть получены из центральных путем добавления операторов, учитывающих направление характеристик. В [9] такие схемы на основе компактных аппроксимаций высокого порядка были построены. При этом учитывались следующие аспекты моделирования течений вязкого сжимаемого газа компактными аппроксимациями.

[1)] Работа выполнена при финансовой поддержке РФФИ (код проекта 05-01-00584).

Преодоление сеточных негладкостей. Узлы разностных сеток, адаптированных к геометрии обтекаемого тела, как правило, распределяются в физической плоскости неравномерно. Зачастую нельзя обойтись без их сильного сгущения в направлении твердой поверхности, кривизны и скоса координатных линий. Влияние данных факторов на разностное решение сильно зависит от способа вычисления метрических коэффициентов. Другой аспект проведения расчетов на криволинейных сетках состоит в проблеме сохранения невозмущенного потока. Появление даже очень слабых возмущений параметров течения в сверхзвуковом набегающем потоке следует признать очень неприятным явлением. Как показано в [4], для преодоления упомянутых трудностей необходимо при определении метрических коэффициентов расчетной области использовать те же дифференциальные операторы, что и для аппроксимации уравнений законов сохранения. В трехмерном случае этого недостаточно, однако и здесь существуют подходы, позволяющие устранить нежелательные моменты, вызванные криволинейными координатами.

Эффективность диссипативного механизма. Симметричные разностные операторы не содержат никакого механизма подавления высокочастотных гармоник, не имеющих физического смысла. Введение в уравнения членов типа искусственной вязкости не всегда удобно, поскольку не позволяет проводить сквозной счет разрывных решений. С другой стороны, использование ограничителей потока вида [10] может приводить к искажению результатов, например, к отсутствию отрыва пограничного слоя там, где эксперимент фиксирует его наличие. В этом плане схемы вида [5], [6], [11], обладающие естественными диссипативными свойствами за счет ориентации разностей против потока, имеют заметные преимущества. Применяющийся в [7] механизм монотонизации решения также представляет собой ориентацию конечных разностей с учетом направления характеристик. Можно сделать вывод о преимуществах схем, ориентированных против потока, хотя способы реализации данного подхода могут отличаться.

Разностное представление диффузных членов. С повышением порядка аппроксимации схем, описывающих конвективные составляющие уравнений, встает вопрос о более точном представлении вязких членов. Привлекательный на первый взгляд алгоритм вычисления вязких членов с использованием симметричных компактных разностей успешно применялся в [4]. Однако опыт [9] показал, что это оправдывает себя лишь при сглаживании полей физических переменных по методу искусственной вязкости. В данном случае был использован традиционный подход, когда необходимые производные параметров потока определяются в полуцелых узлах сетки. Применение интерполяционных формул высокого порядка дало хорошие результаты.

Ниже рассмотрены разностные схемы, являющиеся модификацией [9], и представлены результаты расчетов ряда течений вязкого газа.

1. ПОСТАНОВКА ЗАДАЧИ

Рассматривается задача обтекания двумерного (плоского) тела турбулентным потоком сжимаемого газа с постоянным отношением удельных теплоемкостей. Решение определяется путем численного интегрирования осредненных по Рейнольдсу нестационарных двумерных уравнений Навье–Стокса. Задание турбулентных составляющих осуществляется в терминах добавочной турбулентной вязкости. Обезразмеренные традиционным способом по параметрам набегающего потока на бесконечности и характерному линейному размеру H, они имеют следующий вид в декартовой системе координат (x, y):

$$\frac{\partial \mathbf{U}(\boldsymbol{\varphi})}{\partial t} + \frac{\partial \mathbf{F}(\boldsymbol{\varphi})}{\partial x} + \frac{\partial \mathbf{G}(\boldsymbol{\varphi})}{\partial y} = \mathrm{Re}^{-1}\left(\frac{\partial[\mathbf{F}_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}_x) + \mathbf{F}_2(\boldsymbol{\varphi}, \boldsymbol{\varphi}_y)]}{\partial x} + \frac{\partial[\mathbf{G}_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}_x) + \mathbf{G}_2(\boldsymbol{\varphi}, \boldsymbol{\varphi}_y)]}{\partial y}\right), \tag{1.1}$$

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ e \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho u v \\ (e+p)u \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} \rho v \\ \rho u v \\ \rho v^2 + p \\ (e+p)v \end{bmatrix}, \quad \boldsymbol{\varphi} = \begin{bmatrix} \rho \\ u \\ v \\ h \end{bmatrix},$$

$$\mathbf{F}_1 = \begin{bmatrix} 0 \\ (4/3)\mu u_x \\ \mu v_x \\ (4/3)\mu u u_x + \mu v v_x + \mu \mathrm{Pr}^{-1} h_x \end{bmatrix}, \quad \mathbf{F}_2 = \begin{bmatrix} 0 \\ -(2/3)\mu v_y \\ \mu u_y \\ -(2/3)\mu u v_y + \mu v u_y \end{bmatrix},$$

$$\mathbf{G}_1 = \begin{bmatrix} 0 \\ \mu v_x \\ -(2/3)\mu u_x \\ \mu u v_x - (2/3)\mu v u_x \end{bmatrix}, \quad \mathbf{G}_2 = \begin{bmatrix} 0 \\ \mu u_x \\ (4/3)\mu v_y \\ \mu u u_x + (4/3)\mu v v_y + \mu \mathrm{Pr}^{-1} h_y \end{bmatrix},$$

$$p = \rho[(\gamma - 1)\gamma^{-1} h + (2/3)q^2], \quad e = \rho[\gamma^{-1} h + 0.5(u^2 + v^2) + q^2],$$

$$\mu = \mu_l + \mu_t, \quad \mu \mathrm{Pr}^{-1} = \mu_l \mathrm{Pr}_l^{-1} + \mu_t \mathrm{Pr}_t^{-1}, \quad \mathrm{Re} = \rho_\infty u_\infty H \mu_{l\infty}^{-1}.$$

Здесь t – время, ρ – плотность, u и v – компоненты вектора скорости в направлениях x и y, p – давление, h – энтальпия, e – полная энергия, $q = k^{1/2}$ (k – кинетическая энергия турбулентности), γ – отношение удельных теплоемкостей, μ_l и μ_t – коэффициенты молекулярной и вихревой вязкостей, $\mathrm{Pr}_l = 0.72$ и $\mathrm{Pr}_r = 0.9$ – ламинарное и турбулентное числа Прандтля, Re – число Рейнольдса. Система дополняется зависимостью молекулярной вязкости от энтальпии в виде формулы Сатерленда (см. [12]).

Для расчета турбулентных характеристик используется дифференциальная двухпараметрическая модель (см. [13]). Ее уравнения имеют вид

$$\frac{\partial \mathbf{U}_t(\boldsymbol{\varphi}, \boldsymbol{\varphi}_t)}{\partial t} + \frac{\partial \mathbf{F}_t(\boldsymbol{\varphi}, \boldsymbol{\varphi}_t)}{\partial x} + \frac{\partial \mathbf{G}_t(\boldsymbol{\varphi}, \boldsymbol{\varphi}_t)}{\partial y} = \mathrm{Re}^{-1}\left(\frac{\partial \mathbf{F}_{t1}(\boldsymbol{\varphi}_{t_x})}{\partial x} + \frac{\partial \mathbf{G}_{t2}(\boldsymbol{\varphi}_{t_y})}{\partial y}\right) + \mathbf{H}_t,$$

$$\mathbf{U}_t = \begin{bmatrix} \rho q \\ \rho \nu \end{bmatrix}, \quad \mathbf{F}_t = \begin{bmatrix} \rho u q \\ \rho u \nu \end{bmatrix}, \quad \mathbf{G}_t = \begin{bmatrix} \rho v q \\ \rho v \nu \end{bmatrix}, \quad \boldsymbol{\varphi}_t = \begin{bmatrix} q \\ \nu \end{bmatrix}, \qquad (1.2)$$

$$\mathbf{F}_{t1} = \begin{bmatrix} (\mu_l + \mu_t \mathrm{Pr}_q^{-1}) q_x \\ (\mu_l + \mu_t \mathrm{Pr}_\nu^{-1}) \nu_x \end{bmatrix}, \quad \mathbf{G}_{t2} = \begin{bmatrix} (\mu_l + \mu_t \mathrm{Pr}_q^{-1}) q_y \\ (\mu_l + \mu_t \mathrm{Pr}_\nu^{-1}) \nu_y \end{bmatrix}, \quad \mathbf{H}_t = \begin{bmatrix} H_q \\ H_\nu \end{bmatrix},$$

где $\nu = (qL^{-1})^{1/2}$, L – масштаб турбулентности. Вектор источниковых членов $\mathbf{H}$ имеет компоненты

$$H_q = \frac{\rho q}{2}\left(\frac{C_\mu f_\mu S}{\nu^2} - \frac{2}{3}D - \nu^2\right), \quad H_\nu = \frac{\rho \nu}{2}\left[\alpha\left(\frac{C_\mu S}{\nu^2} - \frac{2}{3}D\right) - \beta \nu^2\right].$$

Здесь

$$S = 2u_x^2 + (u_y^2 + v_x^2) + 2v_y^2 - \frac{2}{3}D^2, \quad D = u_x + v_y, \quad \alpha = \alpha_1 f_\mu + \alpha_2, \quad f_\mu = 1 - \exp(\beta_\mu \mathrm{Re} \rho q n \mu^{-1}),$$

n – расстояние по нормали от поверхности, а $C_\mu = 0.09$, $\alpha_1 = 0.555$, $\alpha_2 = 0.065$, $\beta = 0.92$, $\beta_\mu = 0.0075$, $\mathrm{Pr}_q = 1$ и $\mathrm{Pr}_\nu = 1.3$ – константы. Коэффициент турбулентной вязкости определяется по формуле

$$\mu_t = \mathrm{Re} C_\mu f_\mu \rho \left(\frac{q}{\nu}\right)^2.$$

Граничные условия задаются следующим образом. На входной границе при сверхзвуковой скорости набегающего потока фиксируются параметры течения на бесконечности. При дозвуковой скорости течения задаются полное давление, температура торможения и угол вектора скорости, а четвертый параметр определяется на основе характеристических соотношений для одномерного нестационарного течения невязкого газа. На входной границе также фиксируются значения турбулентных параметров q_∞ и ν_∞. На выходных границах ставятся условия гладкого вытекания; если же течение дозвуковое, то задается статическое давление. На твердых поверхностях плотность ρ_w определяется из уравнения неразрывности, ставятся условия прилипания $u_w = v_w = 0$ и значение энтальпии (температуры) поверхности h_w. Для турбулентных параметров используются условия $q_w = 0$ и $\partial\nu/\partial n = 0$. Также возможно находить ν_w путем решения на границе второго уравнения системы (1.2).

Для расчета обтекания тел с криволинейной поверхностью на основе исходных уравнений в строгой консервативной форме осуществляется переход к обобщенной системе координат (см. [14]). Область течения в физической плоскости (x, y) проецируется на единичный квадрат расчетной плоскости (ξ, η) с помощью отображения общего вида $\xi = \xi(x, y)$, $\eta = \eta(x, y)$, $0 \le \xi \le 1$, $0 \le \eta \le 1$. При этом якобиан преобразования имеет вид $J^{-1} = x_\xi y_\eta - x_\eta y_\xi$, где x_ξ, x_η, y_η, y_ξ – метрические коэффициенты.

Записанная в криволинейной системе координат (ξ, η), система (1.1) имеет вид

$$\frac{\partial\tilde{\mathbf{U}}}{\partial t} + \frac{\partial\tilde{\mathbf{F}}}{\partial\xi} + \frac{\partial\tilde{\mathbf{G}}}{\partial\eta} = \mathrm{Re}^{-1}\left\{\frac{\partial[\tilde{\mathbf{F}}_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\xi) + \tilde{\mathbf{F}}_2(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\eta)]}{\partial\xi} + \frac{\partial[\tilde{\mathbf{G}}_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\xi) + \tilde{\mathbf{G}}_2(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\eta)]}{\partial\eta}\right\}, \tag{1.3}$$

$$\tilde{\mathbf{U}} = J^{-1}\mathbf{U}, \quad \tilde{\mathbf{F}} = \mathbf{F}y_\eta - \mathbf{G}x_\eta, \quad \tilde{\mathbf{G}} = \mathbf{G}x_\xi - \mathbf{F}y_\xi,$$

$$\tilde{\mathbf{F}}_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\xi) = \mathbf{F}_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\xi)y_\eta^2 - [\mathbf{F}_2(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\xi) + \mathbf{G}_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\xi)]y_\xi y_\eta + \mathbf{G}_2(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\xi)y_\xi^2,$$

$$\tilde{\mathbf{F}}_2(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\eta) = -\mathbf{F}_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\eta)x_\eta y_\eta + \mathbf{F}_2(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\eta)x_\xi y_\eta + \mathbf{G}_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\eta)x_\eta y_\xi - \mathbf{G}_2(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\eta)x_\xi y_\xi,$$

$$\tilde{\mathbf{G}}_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\xi) = -\mathbf{F}_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\xi)x_\eta y_\eta + \mathbf{F}_2(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\xi)x_\eta y_\xi + \mathbf{G}_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\xi)x_\xi y_\eta - \mathbf{G}_2(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\xi)x_\xi y_\xi,$$

$$\tilde{\mathbf{G}}_2(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\eta) = \mathbf{F}_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\eta)x_\eta^2 - [\mathbf{F}_2(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\eta) + \mathbf{G}_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\eta)]x_\xi x_\eta + \mathbf{G}_2(\boldsymbol{\varphi}, \boldsymbol{\varphi}_\eta)x_\xi^2.$$

Уравнения системы (1.2) в криволинейной системе координат выглядят следующим образом:

$$\frac{\partial\tilde{\mathbf{U}}_t}{\partial t} + \frac{\partial\tilde{\mathbf{F}}_t}{\partial\xi} + \frac{\partial\tilde{\mathbf{G}}_t}{\partial\eta} = \mathrm{Re}^{-1}\left\{\frac{\partial[\tilde{\mathbf{F}}_{t1}(\boldsymbol{\varphi}_{t\xi}) + \tilde{\mathbf{F}}_{t2}(\boldsymbol{\varphi}_{t\eta})]}{\partial\xi} + \frac{\partial[\tilde{\mathbf{G}}_{t1}(\boldsymbol{\varphi}_{t\xi}) + \tilde{\mathbf{G}}_2(\boldsymbol{\varphi}_{t\eta})]}{\partial\eta}\right\} + \tilde{\mathbf{H}}_t, \tag{1.4}$$

$$\tilde{\mathbf{U}}_t = J^{-1}\mathbf{U}_t, \ \tilde{\mathbf{F}}_t = \mathbf{F}_t y_\eta - \mathbf{G}_t x_\eta, \ \tilde{\mathbf{G}}_t = \mathbf{G}_t x_\xi - \mathbf{F}_t y_\xi, \ \tilde{\mathbf{H}}_t = J^{-1}\mathbf{H}_t,$$

$$\tilde{\mathbf{F}}_t(\boldsymbol{\varphi}_{t\xi}) = \mathbf{F}_{t1}(\boldsymbol{\varphi}_{t\xi})y_\eta^2 + \mathbf{G}_{t2}(\boldsymbol{\varphi}_{t\xi})y_\xi^2,$$

$$\tilde{\mathbf{F}}_{t2}(\boldsymbol{\varphi}_{t\eta}) = -\mathbf{F}_{t1}(\boldsymbol{\varphi}_{t\eta})x_\eta y_\eta - \mathbf{G}_{t2}(\boldsymbol{\varphi}_{t\eta})x_\xi y_\xi,$$

$$\tilde{\mathbf{G}}_{t1}(\boldsymbol{\varphi}_{t\xi}) = -\mathbf{F}_{t1}(\boldsymbol{\varphi}_{t\xi})x_\eta y_\eta - \mathbf{G}_{t2}(\boldsymbol{\varphi}_{t\xi})x_\xi y_\xi,$$

$$\tilde{\mathbf{G}}_{t2}(\boldsymbol{\varphi}_{t\eta}) = \mathbf{F}_{t1}(\boldsymbol{\varphi}_\eta)x_\eta^2 + \mathbf{G}_{t2}(\boldsymbol{\varphi}_{t\eta})x_\xi^2.$$

Запись исходных уравнений сохранения в строгой консервативной форме предполагает выполнение в двумерном случае соотношений

$$\frac{\partial x_\xi}{\partial\eta} - \frac{\partial x_\eta}{\partial\xi} = 0, \quad \frac{\partial y_\eta}{\partial\xi} - \frac{\partial y_\xi}{\partial\eta} = 0. \tag{1.5}$$

При конечно-разностной дискретизации эти равенства должны безусловно выполняться с целью сохранения на криволинейной сетке невозмущенного потока.

2. ЧИСЛЕННЫЙ МЕТОД

Разностные операторы. Рассмотрим аппроксимацию первой производной $\partial u/\partial x$ параметра u на сетке $\omega_h = \{x_j = jh,\ h = \text{const}\}$. Будем использовать оператор сдвига $T_m u_j = u_{j+m} = u(x_j + mh)$ и известные обозначения

$$E = T_0, \quad \Delta_+ = T_1 - E, \quad \Delta_- = E - T_{-1},$$

$$\Delta_0 = \Delta_+ + \Delta_- = T_1 - T_{-1}, \quad \Delta_2 = \Delta_+ - \Delta_- = T_1 - 2E + T_{-1}.$$

С их помощью разностный аналог исходной формулы можно представить, например, хорошо известным выражением первого порядка:

$$u'_j = \frac{1}{2h}(\Delta_0 - s\Delta_2)u_j.$$

Первое слагаемое (центральная разность) здесь служит для аппроксимации производной, а второе является добавкой диффузного типа для подавления паразитных составляющих решения. Знак параметра s задается в соответствии с направлением потока, а его величина определяет степень интенсивности диссипативного механизма. При $|s| = 1$ имеет место односторонняя разность, представляемая операторами Δ_+u и Δ_-u. В случае $s = 0$ разностная формула приобретает второй порядок аппроксимации. При этом она полностью утрачивает диссипативные свойства, что делает ее непригодной к практическому применению.

Используя принцип представления разностного оператора в виде двух составляющих, для аппроксимации первой производной можно построить формулы высокого порядка. В общем виде они выглядят следующим образом:

$$u'_j = \frac{1}{h}\left(\frac{\Delta_0^{(k)}}{a} + (-1)^m s\frac{\Delta_{2m}}{b}\right)u_j,$$

где $k = 4, 6, 8$ – порядок аппроксимации, $m = 2, 3, 4$ – коэффициент при степени четной разности, $a, b = \text{const}$.

В качестве первой составляющей используются известные симметричные компактные разности

$$C_4 = \left(1 + \frac{\Delta_2}{6}\right)^{-1}\frac{\Delta_0}{2h}, \quad C_6 = \left(1 + \frac{\Delta_2}{5}\right)^{-1}\left(1 + \frac{\Delta_2}{30}\right)\frac{\Delta_0}{2h}, \quad C_8 = \left(1 + \frac{2\Delta_2}{7} + \frac{\Delta_2^2}{70}\right)^{-1}\left(1 + \frac{5\Delta_2}{42}\right)\frac{\Delta_0}{2h}.$$

На пятиточечном шаблоне возможно также использование схемы

$$C_{86} = \left[1 + \frac{\gamma + (5\gamma - 1)\Delta_2}{30}\Delta_2\right]^{-1}\left[1 + \left(\gamma - \frac{1}{6}\right)\Delta_2\right]\frac{\Delta_0}{2h}.$$

В определенном диапазоне значений параметра γ она обладает спектральными свойствами, несколько лучшими по сравнению с другими схемами.

Для приведенных схем справедливы следующие соотношения:

$$C_4 = \frac{\partial}{\partial x} - \frac{h^4}{180} + O(h^6), \quad C_6 = \frac{\partial}{\partial x} + \frac{h^6}{2100} + O(h^8),$$

$$C_8 = \frac{\partial}{\partial x} + \frac{h^8}{44100} + O(h^{10}), \quad C_{86} = \frac{\partial}{\partial x} + \frac{(2 - 7\gamma)h^6}{1260} + O(h^8).$$

Использование компактных аппроксимаций требует постановки граничных условий. Пусть точка b находится на левой границе. Возможно применение как несимметричных

$$u'_b = \frac{1}{12h}(-25E + 48T_1 - 36T_2 + 16T_3 - 3T_4)u_b,$$

$$u'_b = \frac{1}{6h}(E + 3T_1)^{-1}(-17E + 9T_1 + 9T_2 - T_3)u_b,$$

$$u'_b = \frac{1}{12h}(E + 8T_1 + 6T_2)^{-1}(-43E - 80T_1 + 108T_2 + 16T_3 - T_4)u_b,$$

$$u'_b = \frac{1}{12h}(E + 12T_1 + 18T_2 + 4T_3)^{-1}(-47E - 192T_1 + 108T_2 + 128T_3 + 3T_4)u_b,$$

так и симметричных разностей

$$u'_b = \frac{1}{6h}(E + T_1)^{-1}(-T_{-1} - 9E + 9T_1 + T_2)u_b,$$

$$u_b' = \frac{1}{3h}(E + 9T_1 + 9T_2 + T_3)^{-1}(-11E - 27T_1 + 27T_2 + T_3)u_b,$$

в том числе и формулы C_4. При этом пятиточечные прогонки, реализующие схемы C_8 и C_{86}, используют по два граничных условия.

Особенность предлагаемого подхода состоит в способе представления диффузных составляющих. Обычно при повышении степени производной увеличивается и сеточный шаблон, на котором записывается ее разностный аналог. Так, в [4] для неявного задания 8-й разности в операторе искусственной вязкости применяется 9-точечный шаблон, а 10-й требует 11-точечного. В [9] с целью использования в шаблоне не более 5 точек сеточные аналоги с 4-й по 8-ю производных представляются в виде комбинации компактного и обычного операторов. При этом для снижения вычислительных затрат в компактном операторе применяется та же неявная часть, что и в симметричной разности, аппроксимирующей первую производную. Вместе с тем 8-е разности можно получить на основе вторых производных Δ_2, последовательно вычисляя сначала 4-е, а затем и 6-е. Применение 3-точечных прогонок позволяет избежать трудности с определением четных разностей в граничных узлах и использовать симметричные операторы.

Искомые разностные операторы могут быть, в частности, представлены в виде

$$\Delta_4 = \left[1 - \left(1 + \frac{\Delta_2}{6}\right)^{-1}\right]\frac{\Delta_2}{6}, \quad \Delta_6 = \left[1 - \left(1 + \frac{\Delta_2}{12}\right)^{-1}\right]\frac{\Delta_4}{12}, \quad \Delta_8 = \left[1 - \left(1 + \frac{\Delta_2}{12}\right)^{-1}\right]\frac{\Delta_6}{12}.$$

Для них справедливы соотношения

$$\Delta_4 = h^4\frac{\partial^4}{\partial x^4} - \frac{h^8}{720} + O(h^{10}), \quad \Delta_6 = h^6\frac{\partial^6}{\partial x^6} - \frac{h^{10}}{180} + O(h^{12}), \quad \Delta_8 = h^8\frac{\partial^8}{\partial x^8} - \frac{7h^{12}}{720} + O(h^{14}).$$

На их основе строятся демпфирующие составляющие разностных схем

$$D_4 = \frac{s}{b_4}\Delta_4, \quad D_6 = -\frac{s}{b_6}\Delta_6, \quad D_8 = \frac{s}{b_8}\Delta_8,$$

константы b_4, b_6 и b_8 в которых с учетом опыта [9] имеют значения $b_4 = 1/24$, $b_6 = 1/150$ и $b_8 = 1/1500$. Знак параметра s задается с учетом направления характеристик течения. На отдельных участках расчетной области, где формируются достаточно сильные градиенты параметров потока, возможно также локальное применение оператора

$$D_2 = -\frac{s}{b_2}\Delta_2, \quad b_2 = 2.$$

Значения производных на границах могут быть определены с использованием односторонних разностей. Так, 4-е разности можно получить с использованием операторов вида

$$u_b^4 = [(2E - T_1)^{-1}\Delta_2 - 2E + 5T_1 + 4T_2 - T_3]\frac{u_b}{2}$$

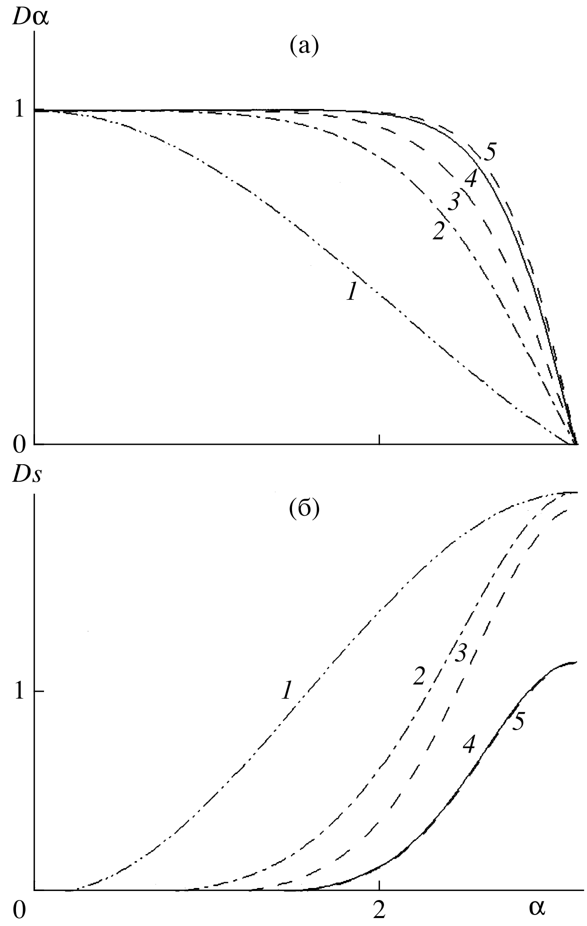
или после расчета четвертых производных во внутренних узлах

$$u_b^4 = 2\Delta_2 u_{b+1}^4 - \Delta_2 u_{b+2}^4.$$

Значения 6-х и 8-х разностей на границе определяются аналогично.

В результате описанной процедуры во всех узлах расчетной области вычисляются четные разности искомых функций со второй по восьмую включительно. Это дает возможность использования той или иной демпфирующей добавки как во всем поле течения, так и на отдельных его участках. При ссылке на разностную схему, выбиравшуюся для конкретного расчета, будем использовать символьное обозначение как симметричного компактного, так и диссипативного операторов. Например, $C_{86}D_8$ – это компактная разность C_{86} с демпфирующей добавкой D_8. Порядок аппроксимации конкретной схемы зависит от выбора обоих составляющих. При этом он за счет D_4–D_8 не ниже 3 и не выше 7.

Спектральные свойства операторов. Особенности разностных операторов исследуют, рассматривая дисперсионные и диссипативные характеристики разностного представления аналогичных одномерных уравнений [2].



Фиг. 1.

Рассмотрим линейное скалярное уравнение переноса

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0.$$

Не прибегая к разностной дискретизации временной производной и полагая решение в узле x_j на сетке ω_h имеющим вид $U(t)\exp(ikhj)$, получаем обыкновенное дифференциальное уравнение

$$\frac{dU}{dt} + \omega(\alpha)U = 0,$$

где i – мнимая единица, k – волновое число, $\alpha = kh$. Каждая из описанных схем имеет свою комплексную функцию $\omega(\alpha)$; их можно получить, используя фурье-представление операторов T_m, Δ_0 и Δ_2:

$$T_m = \exp(i\alpha m), \quad \Delta_0 = 2i\sin\alpha, \quad \Delta_2 = 2(\cos\alpha - 1).$$

Для схемной фазовой скорости справедливо выражение $D_a = \mathrm{Re}[\omega(\alpha)/(i\alpha)]$, D_a называют дисперсией разностной аппроксимации. Другая важная величина, ответственная за затухание гармоник, – схемная диссипация $D_s = \mathrm{Re}[\omega(\alpha)]$.

Фазовая скорость D_a (а) и диссипация D_s (б) рассмотренных схем в зависимости от волнового числа α представлены, соответственно, на фиг. 1а и 1б. Линии *1* получены по схеме первого порядка при значении $s = 0.5$, *2* – для C_4D_4 ($s = 1/24$), *3* – для C_6D_6 ($s = 1/150$), *4* – для C_8D_8, *5* – для $C_{86}D_8$ (обе при $s = 1/1500$).

Отличием данных аппроксимаций от компактных схем с разностями против потока (см. [1], [5]) является независимость их фазовой скорости от значения параметра s, причем значения D_a нигде не превышают 1. Величину $|s|$ для схем C_4D_4–C_8D_8 можно рассматривать как норму внутреннего диссипативного механизма, а ее знак, выбранный с учетом направления потока, обеспечивает устойчивость разностных схем.

На фиг. 1б представлены зависимости диссипации D_s схем от волнового числа α. Видно, что C_4D_4 и C_6D_6 обладают достаточно высокими способностями к подавлению паразитных осцилляций. У C_8D_8 уровень диссипации значителен только для коротких длин волн. Тем не менее проведенные расчеты показали ее пригодность для моделирования течений вязкого газа, в которых отсутствуют сильные градиенты параметров потока.

Разностный алгоритм. Рассмотрим разностное представление системы (1.3) на сетке $\omega_{\xi\eta} = \{\xi_i = ih_\xi,\ \eta_j = ih_\eta,\ h_\xi,\ h_\eta = \mathrm{const}\}$. Симметричные операторы C_4–C_8 будем применять непосредственно к векторам $\tilde{\mathbf{F}}$ и $\tilde{\mathbf{G}}$, а добавки диффузного типа D_4–D_8 – с учетом направления характеристик к составляющим этих векторов. Посредством процедуры, подробно описанной в [15], векторы $\tilde{\mathbf{F}}$ и $\tilde{\mathbf{G}}$ представляются каждый в виде суммы трех векторов $\tilde{\mathbf{f}}_k$, $\tilde{\mathbf{g}}_k$, $k = 1, 2, 3$, каждый в соответствии с количеством собственных значений λ_k^ξ и λ_k^η якобиевых матриц $\partial\tilde{\mathbf{F}}/\partial\mathbf{U}$ и $\partial\tilde{\mathbf{G}}/\partial\mathbf{U}$:

$$\mathbf{f}_1 = \frac{\rho\lambda_1^\xi(\gamma-1)}{\gamma}\begin{bmatrix}1\\ u\\ v\\ \dfrac{u^2+v^2}{2}\end{bmatrix},\quad \mathbf{g}_1 = \frac{\rho\lambda_1^\eta(\gamma-1)}{\gamma}\begin{bmatrix}1\\ u\\ v\\ \dfrac{u^2+v^2}{2}\end{bmatrix},$$

$$\mathbf{f}_{2,3} = \frac{\rho\lambda_{2,3}^\xi}{2\gamma}\begin{bmatrix}1\\ u \mp a_\xi y_\eta(x_\eta^2+y_\eta^2)^{-1}\\ v \pm a_\xi x_\eta(x_\eta^2+y_\eta^2)^{-1}\\ h + \dfrac{u^2+v^2}{2} \mp W_\xi a_\xi(x_\eta^2+y_\eta^2)^{-1}\end{bmatrix},$$

$$\mathbf{g}_{2,3} = \frac{\rho\lambda_{2,3}^\eta}{2\gamma}\begin{bmatrix}1\\ u \pm a_\eta y_\xi(x_\xi^2+y_\xi^2)^{-1}\\ v \mp a_\eta x_\xi(x_\xi^2+y_\xi^2)^{-1}\\ h + \dfrac{u^2+v^2}{2} \mp W_\eta a_\eta(x_\xi^2+y_\xi^2)^{-1}\end{bmatrix},$$

где $W_\xi = uy_\eta - vx_\eta$, $a_\xi = a(x_\eta^2 + y_\eta^2)^{1/2}$, $\lambda_1^\xi = W_\xi$, $\lambda_{2,3}^\xi = W_\xi \mp a_\xi$, $W_\eta = vx_\xi - uy_\xi$, $a_\eta = a(x_\xi^2 + y_\xi^2)^{1/2}$, $\lambda_1^\eta = W_\eta$, $\lambda_{2,3}^\eta = W_\eta \mp a_\eta$, $a = (\gamma p/\rho)^{1/2}$.

Теперь запись конвективных членов системы (1.3) выглядит следующим образом:

$$\mathbf{L}_v\boldsymbol{\varphi}_{ij} = C_{n,\xi}\tilde{\mathbf{F}} + C_{n,\eta}\tilde{\mathbf{G}} + \sum_{m=1}^{3}(D_{k,\xi}\tilde{\mathbf{f}}_m + D_{k,\eta}\tilde{\mathbf{g}}_m) + O(h_\xi^n + h_\eta^n + \mathrm{Re}^{-1}h_\xi^{k-1} + \mathrm{Re}^{-1}h_\eta^{k-1}),$$

где $n, k = 4, 6, 8$.

Поскольку основой разностного представления первых производных являются симметричные операторы C_4–C_8, а несимметричные составляющие D_4–D_8 являются лишь фильтрами высокочастотных осцилляций, разностное представление Δ_4 в них может быть несколько видоизменено. Для расчетов на криволинейных сетках Δ_4 может быть представлено в виде

$$\frac{1}{6}\left[1 - \left(1 + \frac{\Delta_2}{6}\right)^{-1}\right]\delta^{(6)}I^{(6)}M_k\delta^{(6)}\tilde{\mathbf{p}}_k^0,$$

где M_k – матрица метрических коэффициентов, а $\tilde{\mathbf{p}}_k^0$ – вектор $\tilde{\mathbf{p}}_k$, свободный от метрических коэффициентов и содержащий только газодинамические переменные. Возможно также использование выражения

$$\Delta_4\tilde{\mathbf{p}}_k = \left[1-\left(1+\frac{\Delta_2}{6}\right)^{-1}\right]S_k\frac{\Delta_2}{6}\mathbf{U},$$

где $S_k = \partial\tilde{\mathbf{p}}_k/\partial\mathbf{U}$, $\tilde{\mathbf{p}}_k = \tilde{\mathbf{f}}_k, \tilde{\mathbf{g}}_k$.

Трехточечная компактная разность шестого порядка $\delta^{(6)}$ и граничное условие для него четвертого порядка $\delta^{(4)}$ выглядят следующим образом:

$$\delta^{(6)} = \left(1+\frac{9\Delta_2}{80}\right)^{-1}\frac{189(T_{1/2}-T_{-1/2})+17(T_{3/4}-T_{-3/4})}{240h}+\frac{61h^6}{358400}+O(h^8),$$

$$\delta^{(4)} = \left(1+\frac{\Delta_2}{24}\right)^{-1}(T_{1/2}-T_{-1/2})-\frac{17h^4}{5760}+O(h^6).$$

Оператор трехточечной компактной интерполяции $I^{(6)}$ и соответствующее граничное условие $I^{(4)}$ имеют вид

$$I^{(6)} = \left(1+\frac{3\Delta_2}{16}\right)^{-1}\frac{[15(T_{-1/2}+T_{1/2})+T_{-3/2}+T_{3/2}]}{32}+\frac{h^6}{2048}+O(h^8),$$

$$I^{(4)} = (1+T_1)^{-1}\frac{T_{-1/2}+6T_{1/2}+T_{3/2}}{4}+\frac{h^4}{128}+O(h^6).$$

Для описания повторных разностей вязких членов системы (1.3) используются операторы $\delta^{(6)}$ и $I^{(6)}$, а смешанных производных – симметричные разности C_4–C_8. Разностное представление вязких членов имеет вид

$$\mathbf{L}_d\boldsymbol{\varphi}_{ij} = \mathrm{Re}^{-1}(\delta_\xi^{(6)}I_\xi^{(6)}P_1\delta_\xi^{(6)}\boldsymbol{\varphi}+C_{n,\xi}P_2C_{n,\eta}\boldsymbol{\varphi}+C_{n,\eta}Q_1C_{n,\xi}\boldsymbol{\varphi}+\delta_\eta^{(6)}I_\eta^{(6)}Q_2\delta_\eta^{(6)}\boldsymbol{\varphi})+$$
$$+O\left\lfloor \mathrm{Re}^{-1}(h_\xi^n+h_\xi^6+h_\eta^n+h_\eta^6)\right\rfloor,$$

где $P_1 = \partial\tilde{\mathbf{F}}_1/\partial\boldsymbol{\varphi}$, $P_2 = \partial\tilde{\mathbf{F}}_2/\partial\boldsymbol{\varphi}$, $Q_1 = \partial\tilde{\mathbf{G}}_1/\partial\boldsymbol{\varphi}$ и $Q_2 = \partial\tilde{\mathbf{G}}_2/\partial\boldsymbol{\varphi}$.

Интегрирование уравнений (1.3) и (1.4) по временнóй переменной возможно двумя способами. В том случае, когда расчет осуществляется на более или менее равномерной разностной сетке, для решения системы

$$\mathbf{U}^{n+1}-\mathbf{U}^n = -\tau^{-1}(\mathbf{L}_v\varphi_{ij}^n-\mathbf{L}_d\varphi_{ij}^n)$$

лучше всего подходит классический явный метод Рунге–Кутты четвертого порядка. Если же решаемая задача требует использования сетки с сильным сгущением узлов в физической плоскости, больше подходят неявные методы, например метод Гаусса–Зейделя релаксации по линиям. Для стационарной задачи целесообразно использовать метод первого порядка по времени, а для нестационарной – второго. Примеры как явного, так и неявного алгоритмов можно найти, например, в [4].

Неявная разностная система для решения системы (1.3) имеет свойство диагонального преобладания, и для ее совместного решения с неявными граничными условиями используются векторные прогонки с размером блока 4×4. Система, полученная на основе уравнений турбулентности (1.4), решается аналогично векторными прогонками с блоком 2×2.

Метрические коэффициенты x_ξ, x_η, y_ξ, y_η, использующиеся в уравнениях, вычисляются на основе симметричных компактных разностей C_4–C_8. Это обеспечивает разностное выполнение соотношений (1.5), необходимых для сохранения невозмущенного потока на криволинейной сетке. При вычислении производных в источниковых членах системы (1.4) также применяются разностные операторы C_4–C_8. В этом нет прямой необходимости, и выбор данных операторов основан лишь на соображениях общего подхода к аппроксимации всех производных исходной системы уравнений.

3. МЕТОДИЧЕСКИЕ РАСЧЕТЫ

Перенос вихря в потоке невязкого газа. Схемы высокого порядка предназначены прежде всего для численных исследований течений со сложной внутренней структурой. Из соображений экономии ресурсов практические расчеты, как правило, выполняются на ограниченном количестве узлов разностной сетки. Несмотря на это, результаты вычислений должны адекватно отражать физическую реальность. Получить представление о разрешающей способности схем C_4D_4–C_8D_8 можно на примере задачи о перемещении вихря в невязком газе.

Для задания его структуры использовались формулы из [16]. В них радиальная составляющая скорости вихря равняется нулю, а азимутальная определяется по формуле

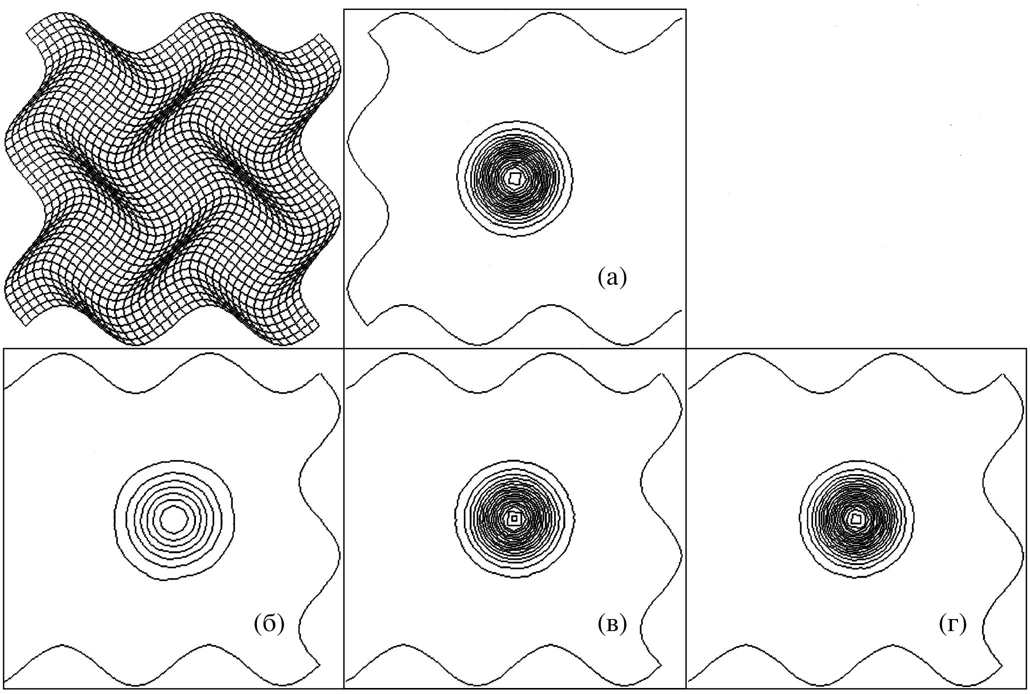
$$w = \frac{1}{2\pi}\exp\left[\frac{1-(r/r_c)^2}{2}\right],$$

где $r = (x^2 + y^2)^{1/2}$ – расстояние от центра вихря, $r_c = \text{const}$. Плотность, компоненты вектора скорости в декартовой системе координат и энтальпия задаются в виде

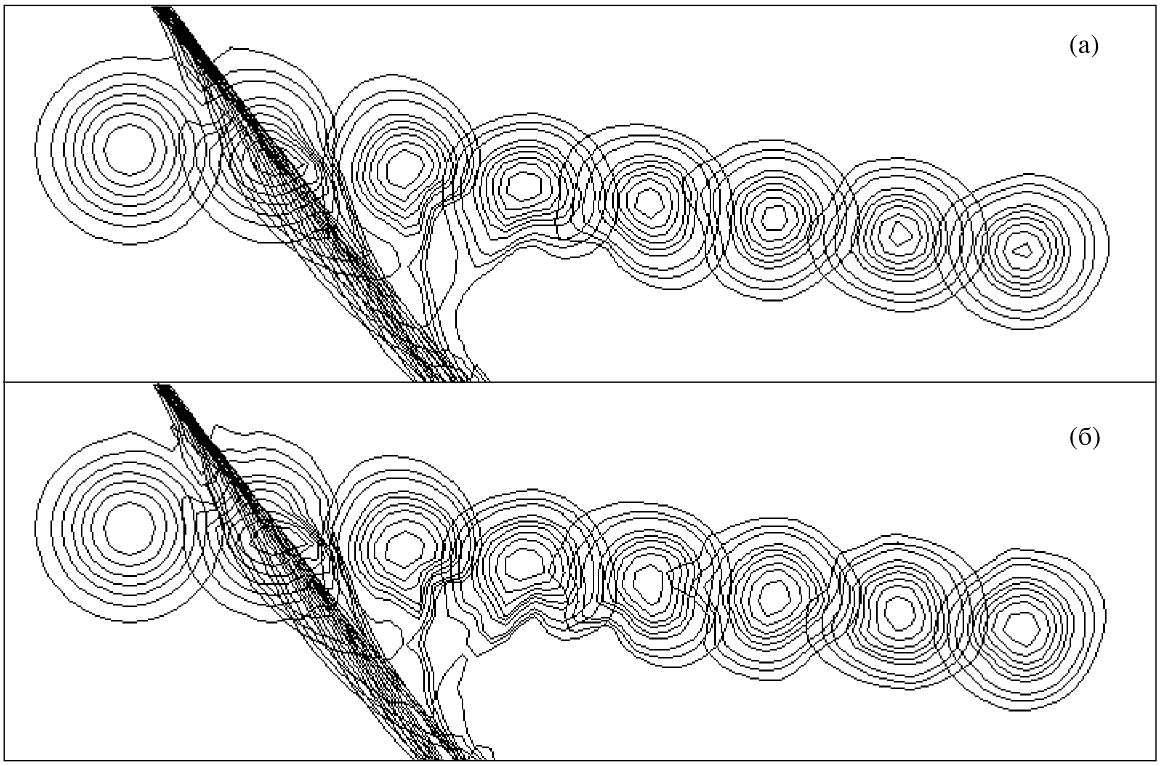
$$\rho = 1-\frac{\gamma-1}{2\gamma}w^2, \quad u = u_0-\frac{y}{r_c}w, \quad v = \frac{x}{r_c}w, \quad h = \frac{\gamma}{\gamma-1}\rho^{\gamma-1}.$$

Данная задача рассматривалась в [16], [17] для случая декартовой сетки и периодических граничных условий. В [4] проводился расчет передвижения вихря по деформированной сетке. В данном случае использовалась сильно искривленная сетка размером 441×41 узел, состоящая из 11 последовательно расположенных блоков. При этом $u_0 = 1$. В процессе расчета вихрь перемещался на расстояние $100r_c$. На фиг. 2 представлены блок расчетной сетки, линии $p = \text{const}$ при его начальном положении (фиг. 2а), а также изобары, соответствующие конечному положению вихря: (б) – для C_4D_4 ($s = 1/24$), (в) – для C_6D_6 ($s = 1/150$) и (г) – для C_8D_8 ($s = 1/1500$). Нанесены 19 изолиний с равным интервалом. Наиболее сильное размывание вихря имеет место при использовании схемы C_4D_4. Схемы C_6D_6 и C_8D_8 дают близкие результаты, однако C_8D_8 выглядит предпочтительнее, поскольку при ее использовании вихрь лучше сохраняет свою первоначальную форму.

Взаимодействие вихря с косым скачком уплотнения. В качестве еще одного теста рассматривалась задача о прохождении вихря через скачок уплотнения в невязком газе. Число Маха набегающего потока равнялось 1.5, угол наклона скачка 51.3°. Расчет проводился на равномерной



Фиг. 2.



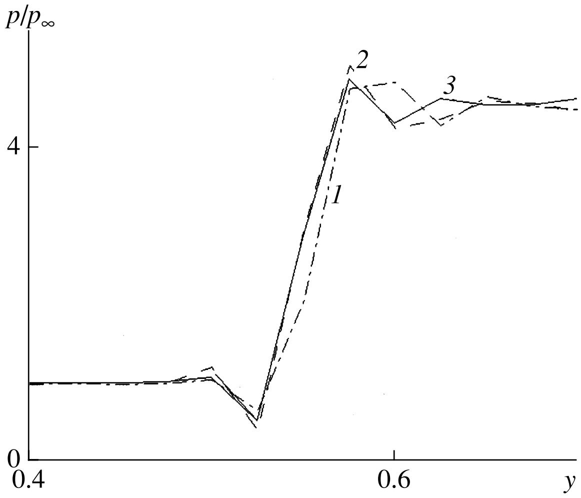
Фиг. 3.

сетке с количеством узлов 95 × 61. На фиг. 3 представлены линии равного давления, соответствующие нескольким моментам, отстоящим друг от друга на одинаковый временной интервал. Результаты расчетов с использованием схем C_4D_4 (фиг. 3а) и C_8D_8 (фиг. 3б) очень близки. Имеющиеся отличия в мелких деталях следует отнести прежде всего за счет большей диссипации C_4D_4. Данные же, полученные по схеме C_6D_6, практически совпадают с результатами C_8D_8.

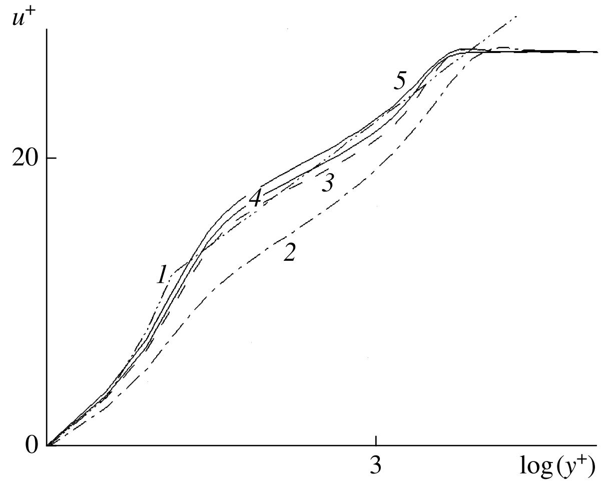
Формирование косого скачка уплотнения. Для проверки гипотезы о возможности управления диссипативными свойствами предложенных схем решалась задача о формировании косого скачка уплотнения в невязком сверхзвуковом потоке. Расчеты проводились на равномерной прямоугольной сетке размером 55 × 41 узел. Число Маха набегающего потока равнялось 4. На основном участке верхней границы фиксировались параметры течения, соответствующие скачку уплотнения с углом наклона 30°. При переходе через скачок, формировавшийся в процессе решения, давление повышалось в 4.5 раза.

Рассматривались схемы C_4D_4, C_6D_6 и C_8D_8. Из них только C_4D_4 позволяет получить решение без внесения локальных изменений в диссипативную составляющую. В двух других случаях такие поправки оказались необходимыми для успешного завершения расчета. Они вносились только в точках наибольших градиентов перед профилем скачка и за ним. Причем если в случае C_6 достаточно было в этих точках заменить D_6 на D_4, то в случае C_8 эта поправка перед профилем скачка на 50% состояла из D_4 и на 50% – из D_2.

Полученные профили относительного давления p/p_∞ представлены на фиг. 4. Линия *1* получена по схеме C_4D_4, *2* – по C_6D_6, и в ближайшей окрестности скачка с использованием оператора D_4, *3* – по C_8D_8, в окрестности скачка с применением D_4, в том числе перед скачком на 50% с D_4 и на 50% с D_2. Как видно, давление непосредственно перед профилем скачка существенно ниже, чем в набегающем потоке. Именно для преодоления подобных участков необходимо локальное усиление диссипативных свойств используемых схем. Предлагаемый метод позволяет это проделывать, хотя сам механизм коррекции используемых диссипативных операторов требует дальнейшей отработки.



Фиг. 4.



Фиг. 5.

4. МОДЕЛИРОВАНИЕ ТУРБУЛЕНТНЫХ ТЕЧЕНИЙ ВЯЗКОГО ГАЗА

Пограничный слой на плоской пластине. Рассматривалось формирование турбулентного пограничного слоя на плоской пластине при числе Маха набегающего потока 0.25 и числе Рейнольдса 10^7. Температура поверхности принималась равной 300°K. Расчетная сетка включала 61×51 узлов, существенно сгущенных к твердой поверхности.

На фиг. 5 представлены профили пограничного слоя в координатах u^+ и $\log(y^+)$, где $u^+ = u/u_\tau$, $y^+ = yu_\tau\rho_w(\mathrm{Re}\mu_w)^{-1}$, а $u_\tau = [\mu_w(\mathrm{Re}\rho_w)^{-1}\partial u/\partial y]^{1/2}$ – динамическая скорость. Там же приведена теоретическая зависимость из [12] (линия *1*), с которой результаты расчетов достаточно хорошо согласуются.

Следует отметить, что данные, полученные с использованием схемы C_8 и оператора D_6 (линия *4*), располагаются ближе к C_6D_6 (линия *3*), чем C_8D_8 (линия *5*). Несколько большие отличия от теории наблюдается у результатов по схеме C_4D_4 (линия *2*), однако здесь надо учитывать, что в данном случае описания вязких членов осуществлялось посредством традиционных операторов второго порядка. Следует признать, что при расчетах течений вязкого газа с помощью одной модели турбулентности, но разными методами возможно получение несколько отличающихся друг от друга результатов.

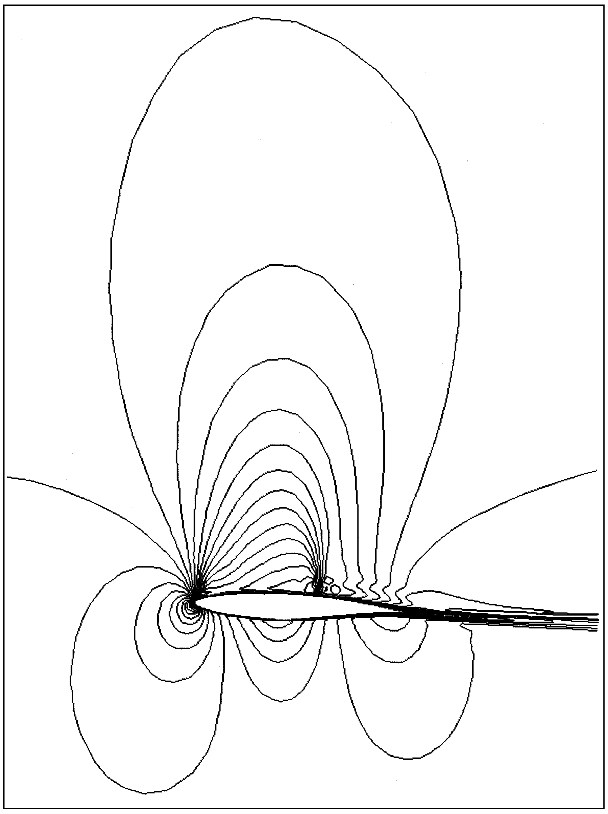
Аэродинамический профиль RAE 2822. Задача рассматривалась при числе Маха набегающего потока 0.725 и угле атаки 2.92°. Число Рейнольдса, посчитанное по длине хорды c, составляло 6.5×10^6. Поверхность профиля считалась теплоизолированной. В расчете использовалась сетка типа O размером в 135×81 узлов.

Полученное поле изолиний числа М представлено на фиг. 6. Всего построено 30 изолиний с постоянным шагом 0.035. Виден скачок уплотнения на верхней поверхности профиля, ближний след, области разгона и торможения потока.

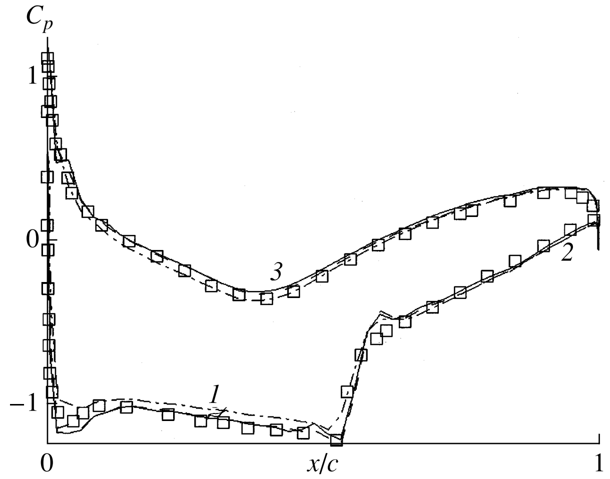
Распределения коэффициента давления $C_p = 2(p - p_\infty)/(\rho_\infty u_\infty)$ на поверхности профиля, полученные на основе схем C_4D_4 (линия *1*), C_6D_6 (линия *2*) и C_8D_8 (линия *3*), а также отмеченные маркерами экспериментальные данные из [19] приведены на фиг. 7. Все три решения хорошо согласуются с экспериментом. Данные, полученные по схемам C_6D_6 и C_8D_8, практически не различаются между собой. Однако при использовании C_8D_8 потребовалось в несколько раз уменьшить шаг интегрирования по времени.

Аэродинамический профиль Джонса. Данный случай сложнее предыдущего, экспериментальные данные для него отсутствуют. Число Маха набегающего потока $M_\infty = 0.85$, число Рейнольдса $\mathrm{Re} = 9 \times 10^6$, угол атаки 2°. Режим течения характеризуется разгоном потока до сверхзвуковых скоростей и формированием замыкающих их скачков уплотнения. В основании одного из них на верхней поверхности происходит отрыв пограничного слоя. На нижней поверхности в зоне присоединения скачка наблюдается утолщение пограничного слоя, но отрыва нет.

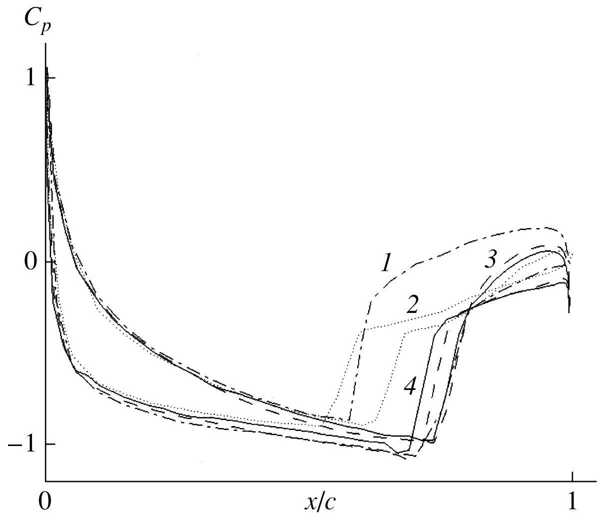
На фиг. 8 представлены распределения коэффициента давления C_p на поверхности профиля. Линия *3* соответствует схеме C_4D_4, *4* – схеме C_6D_6. Там же приведены данные из [19], полученные



Фиг. 6.



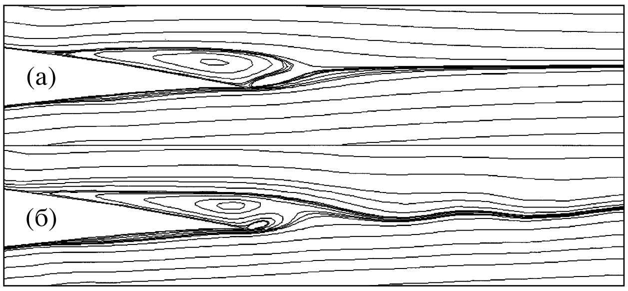
Фиг. 7.



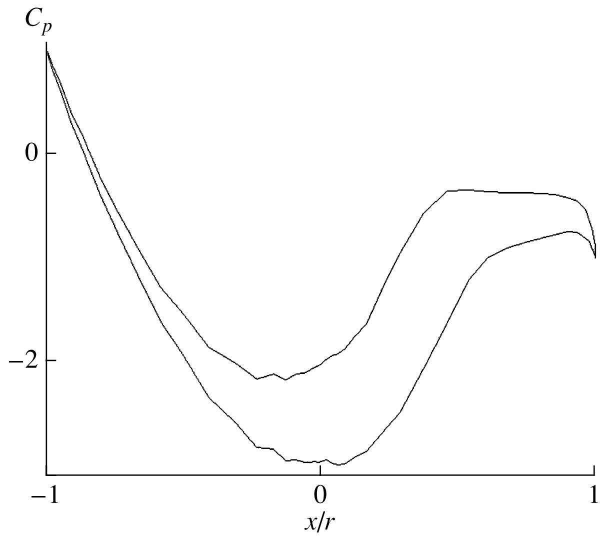
Фиг. 8.

на основе моделей турбулентности Кокли [20] (линия *1*) и Кинга [21] (линия *2*). Если на переднем участке профиля полученные результаты достаточно хорошо согласуются с данными других авторов, то ближе к задней кромке наблюдаются значительные отличия. Они касаются положения скачка уплотнения и протяженности срывной зоны. По-видимому, здесь проявляются различия в используемых турбулентных моделях.

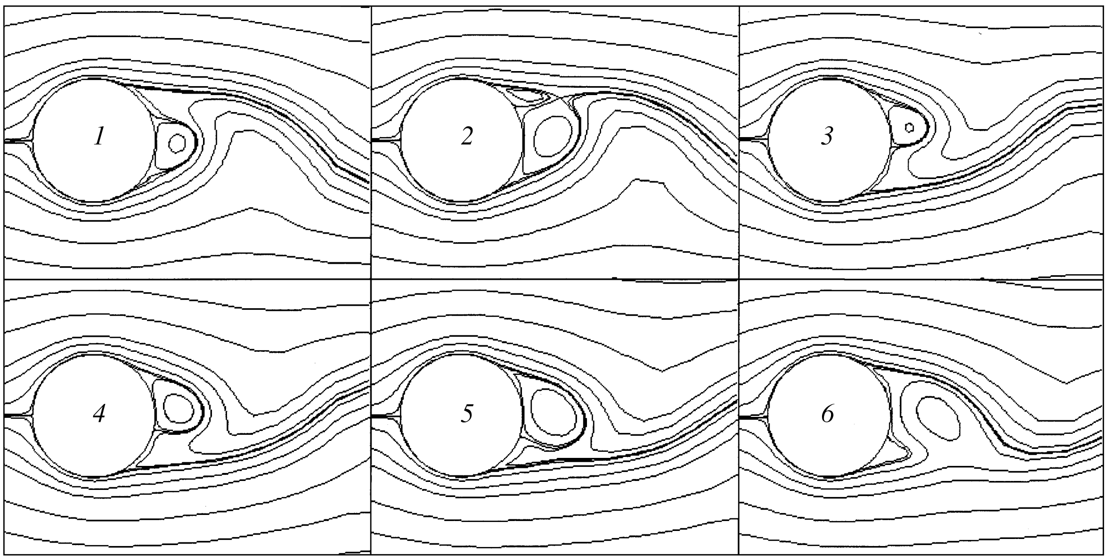
На фиг. 9 приведены поля изолиний функции тока $\psi = \int \rho\, u dy - \rho v dx$ в окрестности задней кромки профиля, полученные на основе схем C_4D_4 (линии (а)) и C_6D_6 (линии (б)). В целом результаты похожи. Однако схема C_6D_6 более детально описывает течение как в зоне отрыва, так и в



Фиг. 9.



Фиг. 10.



Фиг. 11.

следе за профилем. Видны дополнительная зона циркуляции на задней кромке и волнообразные колебания потока за ней.

Поперечное обтекание кругового цилиндра. Отрывное поперечное обтекание бесконечно длинного цилиндра рассматривалось при числе Маха набегающего потока $M_\infty = 0.25$ и числе Рейнольдса, посчитанном по его радиусу r, $\mathrm{Re} = 5 \times 10^6$. На поверхности цилиндра задавалась температура торможения. Расчетная сетка содержала 125×81 узлов. Минимальное расстояние между узлами составляло 10^{-5}. Расчет велся по схеме C_6D_6.

Обтекание носит нестационарный характер. Симметричная в начале расчета картина течения постепенно приобретает колебательный характер со срывом вихрей попеременно сверху и снизу донной части цилиндра. На фиг. 10 представлено характерное распределение коэффициента давления C_p на поверхности цилиндра. Образующийся перепад давления способен раскачивать различные элементы конструкций в направлении, поперечном набегающему потоку.

На фиг. 11 приведены линии тока, полученные для шести последовательных моментов времени за один период колебания. Виден процесс образования вихрей в донной области и срыва их то сверху, то снизу. Так, имеющий место в момент времени *1* вихрь в нижней части цилиндра в момент времени *2* отрывается. При этом сверху образуется новый вихрь, который постепенно растет и отрывается в момент времени *6*. Снизу при этом зарождается новый вихрь.

В заключение следует отметить, что предложенные разностные схемы обладают высоким порядком аппроксимации, малыми фазовыми ошибками и гибкостью в использовании диссипативного механизма. Приведенные результаты свидетельствуют о возможности проведения с их помощью качественных расчетов сложных турбулентных течений вязкого газа.

СПИСОК ЛИТЕРАТУРЫ

1. *Толстых А.И.* О методе численного решения уравнений Навье–Стокса сжимаемого газа в широком диапазоне чисел Рейнольдса // Докл. АН СССР. 1973. Т. 210. № 1. С. 48–51.
2. *Lele S.K.* Compact finite difference schemes with spectral-like resolution // J. Comput. Phys. 1992. V. 102. P. 16–42.
3. *Lele S., Lele S.K., Moin P.* Direct numerical simulation of isotropic turbylence interacting with a shock wave // J. Fluid. Mech. 1993. V. 251. P. 533–562.
4. *Visbal M.R., Gaitonde D.V.* On the use of high-order finite-difference shcemes on curvilinear and deforming meshes // J. Comput. Phys. 2002. V. 181. P. 155–185.
5. *Tolstykh A.I.* High accuracy non-centered compact schemes for fluid dynamics applicatins. Singapore: World Scienct., 1994.
6. *Толстых А.И.* Мультиоператорные схемы произвольного порядка, использующие нецентрированные компактные аппроксимации // Докл. РАН. 1999. Т. 366. № 3. С. 319–322.
7. *Гаранжа В.А., Коньшин В.Н.* Численные алгоритмы для течений вязкой жидкости, основанные на консервативных компактных схемах высокого порядка аппроксимации // Ж. вычисл. матем. и матем. физ. 1999. Т. 39. № 8. С. 1378–1392.
8. *Pulliam T.H.* Arfiticial dissipation models for the Euler equations // AIAA. Journal. 1986. V. 24. № 12. P. 1931–1940.
9. *Савельев А.Д.* О разностных схемах высокого порядка с составными стабилизирующими добавками. М.: ВЦ РАН, 2003.
10. *Charcravathy S.R., Osher S.* A new class of high accuracy TVD schemes for hyperbolic consservation laws: AIAA Paper № 85-0363, 1985. 11 p.
11. *Cockburn B., Shu C.W.* Nonlinearly stable compact schemes for shock calculations // SIAM J. Numer. Analys. 1994. V. 31(3). P. 607–627.
12. *Лойцянский Л.Г.* Механика жидкости и газа. М.: Наука, 1987.
13. *Савельев А.Д.* Расчеты течений вязкого газа на основе (q–ν)-модели турбулентности // Ж. вычисл. матем. и матем. физ. 2003. Т. 43. № 4. С. 589–600.
14. *Стегер Ж.Л.* Неявный конечно-разностный метод для расчета двумерного течения около тел произвольной формы // Ракетная техн. и космонавтика. 1978. Т. 16. № 7. С. 51–60.
15. *Steger J.L., Warming R.F.* Flux vector spliting in the inviscid gas dynamic equations with applications to finite difference methods // J. Comput. Phys. 1981. V. 40. P. 263–293.
16. *Yee H.C., Sandham N.D., Djomehri M.J.* Low dissipation high order shock-capturing methods using characteristic-based filters // J. Comput. Phys. 1999. V. 150. P. 199–238.
17. *Липавский М.В., Толстых А.И.* О сравнительной эффективности схем с нецентрированными компактными аппроксимациями // Ж. вычисл. матем. и матем. физ. 1999. Т. 39. № 10. С. 1705–1720.
18. *Сетлс Ж.С., Вэс И.Е., Богдонов С.М.* Подробная структура турбулентного пограничного слоя, оторвавшегося под действием скачка уплотнения перед углом сжатия // Ракетная техн. и космонавтика. 1976. Т. 14. № 12. С. 55–63.
19. *Holst T.J.* Viscous transonic airfoil workshop. Compendium of results: AIAA Paper № 87-1460, 1987. 32 p.
20. *Coacley T.J.* Turbulence modeling method for the compressible Navier–Stokes equations: AIAA Paper № 83-1693, 1983. 9 p.
21. *Johnson D.A., King L.S.* A mathematically simple turbulence closure model for attached and separated boundary layers // AIAA Journal. 1985. V. 23. № 11. P. 1684–1692.

ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 2007, том 47, № 8, с. 1402–1412

УДК 519.634

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ТРЕХМЕРНОЙ ЗАДАЧИ ЭВОЛЮЦИИ ПОВЕРХНОСТИ РАЗДЕЛА ЖИДКОСТЕЙ РАЗЛИЧНОЙ ВЯЗКОСТИ И ПЛОТНОСТИ В НЕОДНОРОДНОМ ГРУНТЕ[1)]

© 2007 г. Д. Н. Никольский

(302026 Орел, ул. Комсомольская, 95, Орловский гос. ун-т)

e-mail: nikolskydn@mail.ru

Поступила в редакцию 12.01.2007 г.

Рассмотрена трехмерная задача об эволюции границы раздела жидкостей различной вязкости и плотности в случае “поршневого” вытеснения. Задача сведена к численному решению системы, состоящей из интегрального и дифференциального уравнений, методом дискретных особенностей. Исследована “практическая сходимость” численной схемы, и произведено сопоставление результатов численного счета с аналитическим решением. Библ. 8. Фиг. 3. Табл. 2.

Ключевые слова: задача сопряжения, фильтрация в неоднородном грунте, эволюция поверхности сопряжения, численный метод дискретных особенностей, исследование сходимостей.

1. ВВЕДЕНИЕ

Рассмотрена трехмерная задача “поршневого” вытеснения (модель Лейбензона–Маскета) одной жидкости другой в неоднородном грунте в случае пренебрежения капиллярными силами. Математически задача сводится к решению системы, состоящей из интегрального и дифференциального уравнений. Численно эта система решена методом дискретных особенностей, обоснованным в работах И.К. Лифанова. Для полученной численной схемы исследована “практическая сходимость” и проведено сопоставление результатов численного счета с аналитическим решением, полученным для первого шага по времени.

Построенная математическая модель может быть использована для решения практических задач, связанных с первичной разработкой нефтяных месторождений и мониторингом миграции загрязнения в грунтах сложной геологической структуры.

В отличие от [1], где рассматривалась аналогичная задача, здесь учтена неоднородность грунта; в отличие от [2], во-первых, приводится численная схема, во-вторых, полученная система интегральных и дифференциальных уравнений позволяет построить более простую численную схему (нет необходимости аппроксимировать производные от искомой плотности квазипотенциала двойного слоя).

2. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ

Рассмотрим пространственную фильтрацию в неоднородном изотропном и безграничном грунте. Пусть область течения $D \in \mathbb{R}^3$ безгранична. Резкая подвижная поверхность σ_t делит область D на части D^- (внутренняя) и D^+ (внешняя). Область D^+ занята жидкостью с вязкостью μ_1 и плотностью ρ_1, область D^- – с вязкостью и плотностью μ_2 и ρ_2. Нормаль к поверхности σ_t направлена в область D^+.

Поверхность σ_t считаем гладкой. Неоднородность грунта определяется коэффициентом проницаемости грунта $K(M) \in C^{(1)}(D)$. Возбудителями течения в области D являются скважины.

Согласно [3], течение жидкостей в области D удовлетворяет закону Дарси и уравнениям неразрывности

$$\mathbf{v}(M, t) = K(M)\nabla\varphi(M, t), \quad \varphi(M, t) = -\frac{1}{\mu}(p - \rho(\mathbf{e}_g, \mathbf{r}_M)), \tag{1}$$

1)Работа выполнена при финансовой поддержке РФФИ (код проекта 06-01-96303).

$$\nabla \mathbf{v}(M, t) = 0. \tag{2}$$

Здесь $\mathbf{v}(M, t)$ – скорость фильтрации, $\varphi(M, t)$ – квазипотенциал скорости фильтрации, p – гидростатическое давление, $\mathbf{e}_g$ – единичный вектор, указывающий на направление ускорения свободного падения.

Перемещение частиц жидкости, находящихся на поверхности σ_t, описывается дифференциальным уравнением (см. [4])

$$\frac{d\mathbf{r}_M}{dt} = \frac{\mathbf{v}^+(M, t) + \mathbf{v}^-(M, t)}{2}. \tag{3}$$

Законы (1)–(3) записаны в безразмерных величинах. Характерные величины при этом имеют вид

$$P_0 = \rho_0 g L_0, \quad V_0 = \frac{K_0\rho_0 g}{\mu_0}, \quad T_0 = \frac{mL_0\mu_0}{K_0\rho_0 g}.$$

Здесь g – модуль ускорения свободного падения, m – пористость грунта.

На поверхности σ_t выполняются условия непрерывности давлений и неразрывности фильтрационного потока (см. [3])

$$p^+(M, t) = p^-(M, t), \quad v_n^+(M, t) = v_n^-(M, t), \quad M \in \sigma_t. \tag{4}$$

Здесь значки + и – введены для обозначения предельных величин соответствующих функций при их приближении к σ_t из областей D^+ и D^- соответственно.

Из (1) и (2) следует, что квазипотенциал скорости фильтрации $\varphi(M, t)$ удовлетворяет уравнению эллиптического типа

$$\nabla_M(K(M)\nabla_M\varphi(M, t)) = 0. \tag{5}$$

Квазипотенциал φ будем искать в классе функций $C^{(2)}(D)$.

Ограничимся рассмотрением метагармонической (гармонической) серии сред (см. [5]), когда $\sqrt{K(M)}$ удовлетворяет метагармоническому (гармоническому при $\beta = 0$) уравнению

$$\Delta_M\sqrt{K(M)} - \beta^2\sqrt{K(M)} = 0. \tag{6}$$

В этом случае фундаментальное решение уравнения (5) имеет вид (см. [5])

$$\Phi(M, N) = \frac{U(r_{NM}) + G(M, N)}{\sqrt{K(M)K(N)}}, \tag{7}$$

причем $U(r_{NM}) = \dfrac{e^{-\beta r_{NM}}}{4\pi r_{NM}}$ представляет собой решение уравнения

$$\Delta_M U(r_{NM}) = \beta^2 U(r_{NM}), \tag{8}$$

а функция $G(M, N)$ конечная и регулярная. Когда область D безгранична, то имеем $G = 0$.

Условия (4) для квазипотенциала скорости фильтрации примут вид

$$\begin{gathered}\mu_1\varphi^+(M, t) - \mu_2\varphi^-(M, t) = (\rho_1 - \rho_2)(\mathbf{e}_g, \mathbf{r}_M),\\ \left(K(M)\frac{\partial\varphi(M, t)}{\partial n}\right)^+ = \left(K(M)\frac{\partial\varphi(M, t)}{\partial n}\right)^-, \quad M \in \sigma_t.\end{gathered} \tag{9}$$

Квазипотенциал скорости фильтрации $\varphi(M, t)$ представим в виде суммы:

$$\varphi(M, t) = \varphi_0(M, t) + \varphi_*(M, t), \tag{10}$$

где $\varphi_0(M, t)$ – квазипотенциал невозмущенного течения, вызванного работой системы скважин (выписывается для конкретной зависимости коэффициента $K(M)$), $\varphi_*(M, t)$ – квазипотенциал возмущения, вызванного поверхностью σ_t.

Квазипотенциалы $\varphi_0(M, t)$ и $\varphi_*(M, t)$ должны удовлетворять уравнению (5). С учетом (10) условия (9) примут вид

$$\begin{gathered}\mu_1\varphi_*^+(M, t) - \mu_2\varphi_*^-(M, t) = (\mu_1 - \mu_2)\varphi_0(M, t) + (\rho_1 - \rho_2)(\mathbf{e}_g, \mathbf{r}_M), \\ \left(K(M)\frac{\partial\varphi_*(M, t)}{\partial n}\right)^+ = \left(K(M)\frac{\partial\varphi_*(M, t)}{\partial n}\right)^-, \quad M \in \sigma_t.\end{gathered} \tag{11}$$

С учетом (1) и (10) уравнение (3) примет вид

$$\begin{gathered}d\mathbf{r}_M/dt = \mathbf{v}_0(M, t) + \mathbf{v}_*(M, t), \quad M \in \sigma_t, \\ \mathbf{v}_0(M, t) = K(M)\nabla\varphi_0(M, t), \\ \mathbf{v}_*(M, t) = K(M)\frac{(\nabla\varphi_*(M, t))^+ + (\nabla\varphi_*(M, t))^-}{2}.\end{gathered} \tag{12}$$

В начальный момент времени $t = 0$ положение поверхности раздела жидкостей σ_0 задано параметрически в виде

$$\mathbf{r}_{M0} = \mathbf{r}_0(\kappa_1, \kappa_2), \quad \kappa_1, \kappa_2 \text{ – параметры.} \tag{13}$$

В бесконечности для квазипотенциала φ_* выполняются условия регулярности

$$\varphi_*(M, t) = O\left(\frac{1}{r_{MN}}\right), \quad |v_*(M, t)| = |K(M)\nabla\varphi_*(M, t)| = O\left(\frac{1}{r_{MN}^2}\right) \tag{14}$$

при $r_{MN} \longrightarrow \infty$ $(N \in D)$. Условия (14) обеспечивают единственность решения задачи сопряжения (5) и (11) для квазипотенциала возмущения $\varphi_*(M, T)$ (см. [6]).

Задача об эволюции поверхности σ_t состоит в том, что по заданным источникам (стокам) течения $\varphi_0(M, t)$ и начальному положению поверхности (13) необходимо найти ее положения в моменты времени $t > 0$. Для этого нужно решить задачу сопряжения (5) и (11) для квазипотенциала $\varphi_*(M, t)$ с учетом (14), а затем решить дифференциальное уравнение (12) при заданном начальном условии (13).

3. ОСНОВНАЯ СИСТЕМА ИНТЕГРАЛЬНОГО И ДИФФЕРЕНЦИАЛЬНОГО УРАВНЕНИЙ

Квазипотенциал возмущения $\varphi_*(M, t)$ ищем в виде квазипотенциала двойного слоя, распределенного по σ_t с плотностью $g(M, t) \in H$ ($g(M, t)$ удовлетворяет условию Гёльдера):

$$\begin{gathered}\varphi_*(M, t) = \iint_{\sigma_t} g(N, t)\Omega(M, N)d\sigma_N, \quad M \notin \sigma_t, \\ \Omega(M, N) = K(N)(\nabla_N\Phi(M, N), \mathbf{n}_N).\end{gathered} \tag{15}$$

Несложно заметить, что ядро $\Omega(M, N)$ в (15) представимо в виде

$$\Omega(M, N) = \sqrt{\frac{K(N)}{K(M)}}(\nabla_N U, \mathbf{n}_N) - \sqrt{\frac{K(N)}{K(M)}}\frac{(\nabla_N K(N), \mathbf{n}_N)}{2K(N)}U(M, N) + K(N)\left(\nabla_N\frac{G(M, N)}{\sqrt{K(M)K(N)}}, \mathbf{n}_N\right),$$

а значит, скачок квазипотенциала $\varphi_*(M, t)$ при переходе через σ_t обусловлен только асимптотическим приближением в точке N вида

$$\Omega(M, N) \sim \frac{(\mathbf{r}_{NM}, \mathbf{n}_N)}{r_{NM}^3} \text{ при } r_{NM} \longrightarrow 0,$$

что позволяет воспользоваться предельными значениями потенциала двойного слоя для уравнения Лапласа (см. [7])

$$\varphi_*^{\pm}(M, t) = \iint_{\sigma_t} g(N, t)\Omega(M, N)dl_N \pm \frac{g(M, t)}{2}, \quad M \in \sigma_t. \tag{16}$$

Потенциал $\varphi(M, t)$ в виде (10) с учетом свойств φ_0 и (15) удовлетворяет уравнению (5), условию регулярности (14) и второму граничному условию из (11). Подставим (10) в первое условие из (11) и с учетом (16) получим

$$g(M,t) - 2\lambda\int\limits_{\sigma_t} g(N,t)\Omega(M,N)dl_N = 2\lambda\varphi_0(M,t) + \alpha(\mathbf{e}_g, \mathbf{r}_M), \quad M \in \sigma_t,$$
$$\lambda = \frac{\mu_2 - \mu_1}{\mu_2 + \mu_1}, \quad \alpha = \frac{2(\rho_1 - \rho_2)}{\mu_2 + \mu_1}. \tag{17}$$

Воспользовавшись (7), представим скорость возмущения следующим образом:

$$\begin{aligned}\mathbf{v}_*(M,t) = {} & \sqrt{K(M)}\iint\limits_{\sigma_t}\sqrt{K(N)}g(N,t)\nabla_M(\nabla_N U(r_{NM}), \mathbf{n}_N)d\sigma_N - \\ & - \nabla_M\sqrt{K(M)}\iint\limits_{\sigma_t}\sqrt{K(N)}g(N,t)(\nabla_N U(r_{NM}), \mathbf{n}_N)d\sigma_N - \\ & - \frac{\sqrt{K(M)}}{2}\iint\limits_{\sigma_t} g(N,t)(\nabla_N\sqrt{K(N)}, \mathbf{n}_N)\nabla_M U(r_{NM})d\sigma_N + \\ & + \nabla_M\sqrt{K(M)}\iint\limits_{\sigma_t} g(N,t)(\nabla_N\sqrt{K(N)}, \mathbf{n}_N)U(r_{NM})d\sigma_N + \\ & + K(M)\iint\limits_{\sigma_t} K(N)g(N,t)\nabla_M\left(\nabla_N\frac{G(M,N)}{\sqrt{K(M)K(N)}}, \mathbf{n}_N\right)d\sigma_N, \quad M \notin \sigma_t.\end{aligned} \tag{18}$$

Используя известные предельные значения для потенциалов простого и двойного слоя и их производных, а также тот факт, что интеграл со слабой сингулярностью типа $\dfrac{1}{r_{NM}}$ не испытывает скачка на σ_t (см. [7]), имеем для (18)

$$\begin{aligned}\mathbf{v}_*^{\pm}(M,t) = {} & \sqrt{K(M)}\iint\limits_{\sigma_t}\sqrt{K(N)}g(N,t)\nabla_M(\nabla_N U(r_{NM}), \mathbf{n}_N)d\sigma_N - \\ & - \nabla_M\sqrt{K(M)}\iint\limits_{\sigma_t}\sqrt{K(N)}g(N,t)(\nabla_N U(r_{NM}), \mathbf{n}_N)d\sigma_N + \\ & + \frac{\sqrt{K(M)}}{2}\iint\limits_{\sigma_t} g(N,t)(\nabla_N\sqrt{K(N)}, \mathbf{n}_N)\nabla_M U(r_{NM})d\sigma_N + \\ & + \nabla_M\sqrt{K(M)}\iint\limits_{\sigma_t} g(N,t)(\nabla_N\sqrt{K(N)}, \mathbf{n}_N)U(r_{NM})d\sigma_N + \\ & + K(M)\iint\limits_{\sigma_t} K(N)g(N,t)\nabla_M\left(\nabla_N\frac{G(M,N)}{\sqrt{K(M)K(N)}}, \mathbf{n}_N\right)d\sigma_N \pm \\ & \pm \sqrt{K(M)}\frac{\tilde{\nabla}(g(M,t)\sqrt{K(M)})}{2} \pm \frac{\nabla_M K(M)}{4}g(M,t) \mp \frac{(\nabla_M K(M), \mathbf{n}_M)}{4}g(M,t)\mathbf{n}_M, \quad M \in \sigma_t.\end{aligned} \tag{19}$$

Здесь $\tilde{\nabla}_M\, g(M, t)$ – касательная составляющая градиента функции $g(M, t)$ (см. [7]).

С учетом (10) и (19) уравнение (12) примет вид

$$\frac{d\mathbf{r}_M}{dt} = \mathbf{v}_0(M, t) + \sqrt{K(M)}\iint\limits_{\sigma_t}\sqrt{K(N)}g(N, t)\nabla_M(\nabla_N U(r_{NM}), \mathbf{n}_N)d\sigma_N -$$
$$-\nabla_M\sqrt{K(M)}\iint\limits_{\sigma_t}\sqrt{K(N)}g(N, t)(\nabla_N U(r_{NM}), \mathbf{n}_N)d\sigma_N +$$
$$+\frac{\sqrt{K(M)}}{2}\iint\limits_{\sigma_t}g(N, t)(\nabla_N\sqrt{K(N)}, \mathbf{n}_N)\nabla_M U(r_{NM})d\sigma_N + \tag{20}$$
$$+\nabla_M\sqrt{K(M)}\iint\limits_{\sigma_t}g(N, t)(\nabla_N\sqrt{K(N)}, \mathbf{n}_N)U(r_{NM})d\sigma_N +$$
$$+K(M)\iint\limits_{\sigma_t}K(N)g(N, t)\nabla_M\left(\nabla_N\frac{G(M, N)}{\sqrt{K(M)K(N)}}, \mathbf{n}_N\right)d\sigma_N, \quad M \in \sigma_t.$$

Таким образом, для нахождения σ_t имеем систему, состоящую из интегрального уравнения (17) и уравнения (20) при заданном начальном условии (13).

4. ПЕРЕХОД К КОНТУРНОМУ ИНТЕГРАЛУ

Перед построением численной схемы необходимо доказать вспомогательную теорему. Пусть имеется некоторая поверхность Σ, опирающаяся на контур L.

Теорема. *Для функции* $U(r_{NM})$ *из фундаментального решения* (7) *уравнения* (5) *вытекает тождество*

$$\iint\limits_{\Sigma}\nabla_M(\nabla_M U(r_{NM}), \mathbf{n}_N)d\sigma_N = -\oint\limits_L[\nabla_M U(r_{NM}), d\mathbf{r}_N] - \iint\limits_{\Sigma}\beta^2 U(r_{NM})\mathbf{n}_N d\sigma_N. \tag{21}$$

Доказательство. Применим известную теорему Стокса к скалярной функции

$$\iint\limits_{\Sigma}[d\boldsymbol{\sigma}_N, \nabla_M U(M, N)] = \oint\limits_L d\mathbf{r}_N U(M, N).$$

Действуя оператором ∇_M на левую и правую части последнего выражения, получаем

$$\iint\limits_{\Sigma}[\nabla_M, [d\boldsymbol{\sigma}_N, \nabla_M U(M, N)]] = \oint\limits_L[\nabla_M, d\mathbf{r}_N U(M, N)]. \tag{22}$$

Несложно доказать следующие соотношения:

$$\begin{gathered}\nabla_M(\nabla_N U(M, N), d\boldsymbol{\sigma}_N) = -[\nabla_M, [d\boldsymbol{\sigma}_N, \nabla_N U(M, N)]] + \\ + d\boldsymbol{\sigma}_N(\nabla_M, \nabla_N U(M, N)) + [d\boldsymbol{\sigma}_N, [\nabla_M, \nabla_N U(M, N)]],\end{gathered} \tag{23}$$

$$[\nabla_M, U(M, N)d\mathbf{r}_N] = [\nabla_M U(M, N), d\mathbf{r}_N]. \tag{24}$$

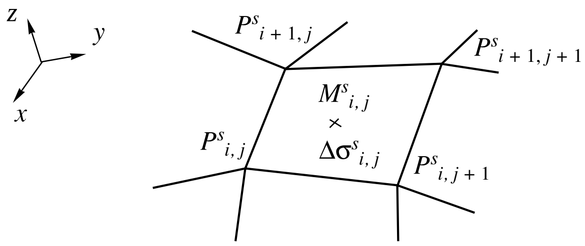
Учитывая, что $U(M, N) = U(r_{NM})$ в нашем случае (а значит, $\nabla_M U(r_{NM}) = -\nabla_N U(r_{NM})$), а также уравнение (8), для (23) получаем

$$\nabla_M(\nabla_N U(M, N), d\boldsymbol{\sigma}_N) = -[\nabla_M, [d\boldsymbol{\sigma}_N, \nabla_N U(M, N)]] - \beta^2 U(r_{NM})d\boldsymbol{\sigma}_N. \tag{25}$$

Тогда, с учетом (24) и (25), преобразованная теорема Стокса (22) примет вид (21). Отметим, что (21) при $\beta = 0$ представляет собой известную теорему Ампера.

5. ЧИСЛЕННОЕ РЕШЕНИЕ

Пусть выбрана правая система координат. Построим численную схему для уравнений (17) и (20), воспользовавшись методом дискретных особенностей для решения (17), и разовьем метод



Фиг. 1.

замкнутых вихревых рамок для вычисления гиперсингулярного интеграла в (20). Для этого поверхность σ_t, известную в некоторый момент времени t^s, $s = 0, 1, \ldots$, представим элементарными площадками $\Delta\sigma^s_{m,k}$, $m = 0, 1, \ldots, n_1 - 2$, $k = 0, 1, \ldots, n_2 - 2$, $s = 0, 1, \ldots$, приблизительно одинаковой площади (фиг. 1).

Края этих площадок задают точки $P^s_{i,j}$, $i = 0, 1, \ldots, n_1 - 1$, $j = 0, 1, \ldots, n_2 - 1$, $s = 0, 1, \ldots$. Приблизительно в центре каждой площадки $\Delta\sigma^s_{m,k}$ расположим расчетные точки $M^s_{m,k}$, $i = 0, 1, \ldots, n_1 - 2$, $j = 0, 1, \ldots, n_2 - 2$, $s = 0, 1, \ldots$, вычислив их координаты по формуле

$$M^s_{m,k} = \frac{1}{4}(P^s_{m,k} + P^s_{m+1,k} + P^s_{m+1,k+1} + P^s_{m,k+1}).$$

В интегральном уравнении (17) заменим интеграл суммой:

$$g^s_{m,k} - 2\lambda \sum_{\substack{i=0 \\ m,k \neq i,j}}^{n_1-2} \sum_{j=0}^{n_2-2} g^s_{i,j}\Omega_{m,k,i,j}\Delta\sigma^s_{i,j} = 2\lambda\varphi^s_{0m,k} + \alpha(\mathbf{e}_g, \mathbf{r}_{m,k}),$$
$$m = 0, 1, \ldots, n_1 - 2, \quad k = 0, 1, \ldots, n_2 - 2, \tag{26}$$

которая представляет собой систему линейных алгебраических уравнений.

Отметим, что в ходе численных расчетов для системы (26) было установлено диагональное преобладание, что позволило решать эту систему методом простой итерации. В этом методе каждое новое приближение находится по формуле

$$g^{s,p}_{m,k} = 2\lambda \sum_{\substack{i=0 \\ \mathrm{m,k} \neq \mathrm{i,j}}}^{n_1-2} \sum_{j=0}^{n_2-2} g^{s,p-1}_{i,j}\Omega_{m,k,i,j}\Delta\sigma^s_{i,j} + 2\lambda\varphi^s_{0m,k} + \alpha(\mathbf{e}_g, \mathbf{r}_{m,k}),$$
$$m = 0, 1, \ldots, n_1 - 2, \quad k = 0, 1, \ldots, n_2 - 2, \quad p = 1, 2, \ldots, J, \tag{27}$$

или, в матричном виде,

$$g^{s,p} = B^{s,p-1} + C^s, \quad s = 0, 1, \ldots, \quad p = 1, 2, \ldots, J \text{ при } \|B^{s,p-1}\| < 1. \tag{28}$$

Здесь J – число итераций, которое определяется из условия (см. [8])

$$\frac{\|g^J - g^{J-1}\|_1}{\|g^J\|_1} < \frac{1 - \|B\|_1}{\|B\|_1}\varepsilon, \tag{29}$$

где ε – заданная точность.

Выбор метода простой итерации в данной работе объясняется простотой его распараллеливания.

Считаем, что плотность потенциала двойного слоя $g^s_{m,k}$ и коэффициент $K(M)$ не меняются в пределах каждой площадки $\Delta\sigma^s_{m,k}$, $m = 0, 1, \ldots, n_1 - 2$, $k = 0, 1, \ldots, n_2 - 2$, $s = 0, 1, \ldots$. Тогда разностная схема для (20) с учетом (21) примет вид

$$\frac{\Delta \mathbf{r}^s_{m,k}}{\Delta t^s} = \mathbf{v}^s_{0m,k} - \sqrt{K_{m,k}} \sum_{i=0}^{n_1-2} \sum_{j=0}^{n_2-2} \sqrt{K_{i,j}} g^s_{i,j} \oint_{L^s_{i,j}} \nabla_M U(r_{NM}) d\mathbf{r}_N +$$
$$+ \sqrt{K_{m,k}} \sum_{\substack{i=0 \\ m,k \neq i,j}}^{n_1-2} \sum_{j=0}^{n_2-2} \sqrt{K_{i,j}} g^s_{i,j} \iint_{\Delta\sigma^s_{i,j}} [(\nabla_N \ln\sqrt{K(N)}, \mathbf{n}_N) \nabla_N U(r_{NM}) - \nabla_M \ln\sqrt{K(N)} (\nabla_N U(r_{NM}), \mathbf{n}_N)] d\sigma_N +$$
$$+ \sqrt{K_{m,k}} \sum_{\substack{i=0 \\ m,k \neq i,j}}^{n_1-2} \sum_{j=0}^{n_2-2} \sqrt{K_{i,j}} g^s_{i,j} \iint_{\Delta\sigma^s_{i,j}} [(\nabla_N \ln\sqrt{K(N)}, \mathbf{n}_N) \nabla_M \ln\sqrt{K(M)} - \beta^2 \mathbf{n}_N] U(r_{NM}) d\sigma_N + \tag{30}$$
$$+ K_{m,k} \sum_{\substack{i=0 \\ m,k \neq i,j}}^{n_1-2} \sum_{j=0}^{n_2-2} K_{i,j} g^s_{i,j} \iint_{\Delta\sigma^s_{i,j}} \nabla_M \left(\nabla_N \frac{G(M,N)}{\sqrt{K(M)K(N)}} \right) d\sigma_N,$$

$$\Delta \mathbf{r}^s_{m,k} = \mathbf{r}^{s+1}_{m,k} - \mathbf{r}^s_{m,k}, \quad \Delta t^s = t^{s+1} - t^s,$$
$$m = 0, 1, \ldots, n_1 - 2, \quad k = 0, 1, \ldots, n_2 - 2, \quad s = 0, 1, \ldots.$$

Уточним формулы (27) и (30) для конкретной зависимости коэффициента проводимости $K(M)$:

$$K(M) = e^{2\beta z_M}. \tag{31}$$

(Нетрудно убедиться, что $\sqrt{K}$ является метагармонической функцией.)

С учетом (31) из (27) и (30) получаем

$$g^{s,p}_{m,k} = 2\lambda \sum_{\substack{i=0 \\ m,k \neq i,j}}^{n_1-2} \sum_{j=0}^{n_2-2} g^{s,p-1}_{i,j} \Omega_{m,k,i,j} \Delta\sigma^s_{i,j} + 2\lambda \varphi^s_{0m,k} + \alpha(\mathbf{e}_g, \mathbf{r}_{m,k}),$$

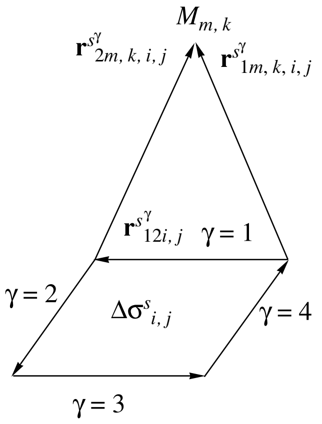
$$\Omega_{m,k,i,j} = \frac{1}{4\pi} \sqrt{\frac{K_{i,j}}{K_{m,k}}} \left(\frac{(1 + \beta r^s_{m,k,i,j}) e^{-\beta r^s_{m,k,i,j}} (\mathbf{r}^s_{m,k,i,j}, \mathbf{n}_{i,j})}{(r^s_{m,k,i,j})^3} - \frac{\beta(\mathbf{k}, \mathbf{n}_{i,j}) e^{-\beta r^s_{m,k,i,j}}}{r^s_{m,k,i,j}} \right), \tag{32}$$

$$m = 0, 1, \ldots, n_1 - 2, \quad k = 0, 1, \ldots, n_2 - 2, \quad p = 1, 2, \ldots, J, \quad s = 0, 1, \ldots;$$

$$\frac{\Delta \mathbf{r}^s_{m,k}}{\Delta t^s} = \mathbf{v}^s_{0m,k} + \mathbf{v}^s_{*m,k},$$

$$\mathbf{v}^s_{*m,k} = -\sqrt{K_{m,k}} \sum_{i=0}^{n_1-2} \sum_{j=0}^{n_2-2} \sqrt{K_{i,j}} g^s_{i,j} \sum_{\gamma=1}^{4} \mathbf{V}^{s\gamma}_{1m,k,i,j} + \sqrt{K_{m,k}} \sum_{\substack{i=0 \\ m,k \neq i,j}}^{n_1-2} \sum_{j=0}^{n_2-2} \sqrt{K_{i,j}} g^s_{i,j} (\beta \mathbf{V}^s_{2m,k,i,j} + \beta^2 \mathbf{V}^s_{3m,k,i,j}) \Delta\sigma^s_{i,j},$$

$$\mathbf{V}^{s\gamma}_{1m,k,i,j} = \frac{\mathbf{e}^s_{1m,k,i,j}}{4\pi r^{s\gamma}_{0m,k,i,j}} \int_{\theta_1^{s\gamma}}^{\theta_2^{s\gamma}} (\sin\theta + \beta r^{s\gamma}_{0m,k,i,j}) e^{-\frac{\beta r^{s\gamma}_{0m,k,i,j}}{\sin\theta}} d\theta,$$



Фиг. 2.

$$
\begin{gathered}
r^{\gamma}_{0m,k,i,j} = \frac{\left|[\mathbf{r}^{s\gamma}_{12i,j}, \mathbf{r}^{s\gamma}_{1m,k,i,j}]\right|}{r^{s\gamma}_{12i,j}}, \quad \mathbf{e}^{s}_{1m,k,i,j} = \frac{[\mathbf{r}^{s\gamma}_{1m,k,i,j}, \mathbf{r}^{s\gamma}_{12i,j}]}{\left|[\mathbf{r}^{s\gamma}_{1m,k,i,j}, \mathbf{r}^{s\gamma}_{12i,j}]\right|}, \\
\theta_{\zeta} = \arcsin\frac{\left|[\mathbf{r}^{s\gamma}_{12i,j}, \mathbf{r}^{s\gamma}_{\zeta m,k,i,j}]\right|}{r^{s\gamma}_{12i,j} r^{s\gamma}_{\zeta m,k,i,j}}, \quad \zeta = 1, 2, \\
\mathbf{V}^{s}_{2m,k,i,j} = \frac{(1 + \beta r^{s}_{m,k,i,j}) U^{s}_{m,k,i,j}}{(r^{s}_{m,k,i,j})^2}[\mathbf{n}^{s}_{i,j}, [\mathbf{r}^{s}_{m,k,i,j}, \mathbf{k}]], \\
\mathbf{V}^{s}_{3m,k,i,j} = U^{s}_{m,k,i,j}[\mathbf{k}, [\mathbf{k}, \mathbf{n}^{s}_{i,j}]], \quad U^{s}_{m,k,i,j} = \frac{e^{-\beta r^{s}_{m,k,i,j}}}{4\pi r^{s}_{m,k,i,j}}, \\
m = 0, 1, \dots, n_1 - 2, \quad k = 0, 1, \dots, n_2 - 2, \quad s = 0, 1, \dots .
\end{gathered} \tag{33}
$$

Интеграл в (33) не выражается через элементарные функции, поэтому вычисляется численно по формуле Симпсона. Векторы $\mathbf{r}^{s\gamma}_{1m,k,i,j}$, $\mathbf{r}^{s\gamma}_{2m,k,i,j}$, $\mathbf{r}^{s\gamma}_{12i,j}$ изображены на фиг. 2 (вихревая рамка).

Итак, для численного решения задачи об эволюции поверхности раздела жидкостей в неоднородном грунте, проводимость которого характеризуется функцией (31), имеем систему линейных алгебраических уравнений (32) и разностный аналог дифференциального уравнения движения поверхности раздела жидкостей (33). Для нахождения смещения расчетных точек $M^s_{m,k}$, $m = 0, 1, \dots, n_1 - 2$, $k = 0, 1, \dots, n_2 - 2$, в каждый момент времени t^s, $s = 0, 1, \dots$, нужно найти плотность $g^s_{m,k}$ в этих точках путем решения системы (26), а затем по формулам (30) вычислить смещения $\Delta\mathbf{r}^s_{m,k}$ точек $M^s_{m,k}$.

Отметим, что при $\beta = 0$ полученная дискретная система совпадает с результатом из [1].

6. СОПОСТАВЛЕНИЕ С АНАЛИТИЧЕСКИМ РЕШЕНИЕМ

Рассмотрим случай, когда поверхность раздела жидкостей σ_0 в начальный момент времени $t = 0$ совпадает с плоскостью $x = 0$, а влияние силы тяжести мало́: $\alpha = 0$ (напорная фильтрация).

Скважину моделируем единичным стоком, потенциал которого имеем вид

$$
\varphi_0(M) = \frac{q}{4\pi}\sqrt{\frac{K(M_1)}{K(M)}}\frac{e^{-\beta r_{M_1M}}}{r_{M_1M}}, \quad q = \frac{\Pi}{K_{M_1}}, \tag{34}
$$

где q – приведенный дебит (расход на единицу проводимости грунта).

Для поставленной задачи на первом шаге по времени $t = 0$ несложно выписать аналитическое выражение, определяющее скорость перемещения поверхности σ_0. Действительно, потенциалы

$$\varphi_1(M) = \varphi_0(M) - \lambda\tilde{\varphi}_0(M), \quad \varphi_2(M) = (1+\lambda)\varphi_0(M),$$
$$\tilde{\varphi}_0(M) = \frac{q}{4\pi}\sqrt{\frac{K(M_1)}{K(M)}}\frac{e^{-\beta r_{\tilde{M}_1 M}}}{r_{\tilde{M}_1 M}}, \quad q = \frac{\Pi}{K_{M_1}}, \quad \tilde{M}_1 = (-1, 0, 0) \tag{35}$$

удовлетворяют уравнению (5), граничным условием (9) и описывают течения жидкостей с вязкостями μ_1 и μ_2 соответственно.

Из (3) с учетом (35) скорость перемещения точек поверхности σ_0 получим в виде

$$\mathbf{v}_a(M, 0) = \mathbf{v}_0(M, 0) + K(M)\frac{\nabla_M\varphi_1(M) + \nabla_M\varphi_2(M)}{2}, \quad M \in \sigma_0. \tag{36}$$

Для оценки ошибки вычислений введем величину

$$\eta(M) = \left|\frac{v(M)}{v_a(M)} - 1\right| \times 100\%.$$

В ходе численных расчетов было замечено, что ошибка η принимает наибольшее значение в центре плоскости σ_0 (напротив скважины). Исключение представляют только крайние точки плоскости (образуют рамку) при $\beta > 0$ и крайняя полоса плоскости (область между двумя рамками) при $\beta = 0$ (см. [1]). Погрешность в них связана с обрывом σ_0, поэтому здесь и далее устраняем ее путем наложения ограничений на скорость возмущения:

$$\mathbf{v}_*(x_M, y_M, z_M, t) = \mathbf{v}_{\min} \text{ для } |y_M| \in [|y_{\min}|, L_y], \quad |z_M| \in [|z_{\min}|, L_z]. \tag{37}$$

Здесь $y_{\min}$, $z_{\min}$ – координаты точки, в которой модуль скорости возмущения $v_*(M, t)$ уменьшается в 1000 раз при движении от центра плоскости σ_0 к ее краям вдоль одной из координатных осей Oy и Oz соответственно; L_y, L_z – абсолютные значения соответствующих координат крайних точек плоскости.

Для выполнения численных расчетов с учетом (37) плоскость заменим квадратом $x = 0$, $y \in [-3, 3]$ и $z \in [-3, 3]$. (Из-за малости z-й компоненты скорости по сравнению с модулем скорости несимметрия практически незаметна.) Точность ε из (29) выберем равной $\varepsilon = 10^{-10}$.

Введем обозначение $\eta = \eta(0, 0, 0)$. В табл. 1 представлена зависимость ошибки η, % от $n = n_1 - 1 = n_2 - 1$ для различных значений λ. Для определения выберем $\beta = 1$.

Анализируя табл. 1, замечаем, что с ростом n ошибка η уменьшается, а значит, численное решение приближается к аналитическому.

7. УТОЧНЕНИЕ ЧИСЛЕННОЙ СХЕМЫ

При вычислении предпоследнего слагаемого в (30) для $m = i$ и $k = j$ из сумм выбрасывается слагаемое, которое является конечным. Даже грубый учет этого слагаемого позволяет в несколько раз уменьшить погрешность вычислений. Покажем это.

Таблица 1

λ	n			
	10	20	40	60
1	12.4	5.4	2.4	1.6
0.5	7.9	3.4	1.6	1.1
–0.5	14.1	5.9	2.6	1.7

Таблица 2

λ	n			
	10	20	40	60
1	5.8	1.9	0.7	0.4
0.5	3.7	1.2	0.4	0.3
–0.5	6.4	2.1	0.7	0.4

Вычислим слагаемое, отброшенное из сумм предпоследнего слагаемого в (30):

$$\mathbf{v}_{*3m,\,k} = K_{m,\,k} g^s_{m,\,k} \int\limits_{\Delta\sigma^2_{i,\,j}}\!\!\int \beta^2[\mathbf{k}, [\mathbf{k}, \mathbf{n}_N]] U(r_{NM}) \Delta\sigma_N,$$

$$m = 0, 1, \ldots, n_1 - 1, \quad k = 0, 1, \ldots, n_2 - 1.$$

Заменим площадку $\Delta\sigma^s_{m,\,k}$ на круг радиуса

$$R = \frac{r_{P^s_{m,\,k},\,P^s_{m+1,\,k}} + r_{P^s_{m+1,\,k},\,P^s_{m+1,\,k+1}} + r_{P^s_{m+1,\,k+1},\,P^s_{m,\,k+1}} + r_{P^s_{m,\,k+1},\,P^s_{m,\,k}}}{8}.$$

Тогда в полярной системе координат ρ, θ с центром в $M^s_{m,\,k}$ интеграл $\mathbf{v}_{*3m,\,k}$ может быть вычислен точно по формуле

$$\mathbf{v}_{*3m,\,k} = K_{m,\,k} g^s_{m,\,k} \frac{\beta^2[\mathbf{k}, [\mathbf{k}, \mathbf{n}^s_{m,\,k}]]}{4\pi} \int\limits_0^R \int\limits_0^{2\pi} \frac{e^{-\beta\rho}}{\rho} \rho\, d\theta\, d\rho = K_{m,\,k} g^s_{m,\,k} \beta[\mathbf{k}, [\mathbf{k}, \mathbf{n}^s_{m,\,k}]] \frac{1 - e^{-\beta R}}{2}. \tag{38}$$

Проведем численный эксперимент, описанный в разд. 6, с учетом (38). Его результаты отображены в табл. 2 при $\beta = 1$.

Из табл. 2 видно, что уточнение численной схемы привело к уменьшению погрешности вычислений.

8. ИССЛЕДОВАНИЕ “ПРАКТИЧЕСКОЙ СХОДИМОСТИ”

Решим задачу об эволюции σ_t к единичной скважине (34). Выберем за характерное расстояние d расстояние от скважины до σ_0. Скважину расположим на оси z, т.е. $M_1(0, 0, 1)$. Для определенности полагаем $\lambda = 0.5$, $\alpha = 0.1$ и $\beta = 1$. Первоначальную поверхность раздела моделируем квадратом $z = 0$, $x \in [-5, 5]$, $y \in [-5, 5]$. Параметры квадрата выбраны с учетом (37).

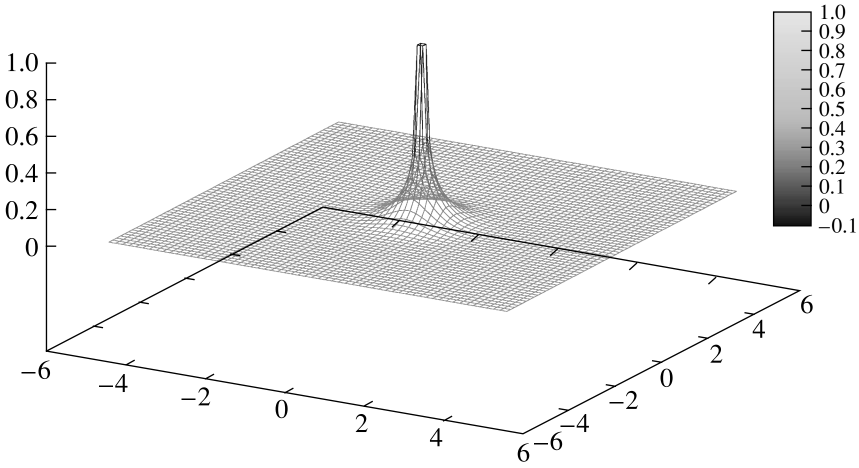
За характерное выберем время достижения поверхностью σ_t контура скважины при $\lambda = 0$, $\alpha = 0$ и $\beta = 0$:

$$T_0 = \frac{4\pi d^3}{q}.$$

Представим зависимость времени T (времени достижения поверхностью σ_t контура скважины) от числа точек разбиения поверхности $n = n_1 - 1 = n_2 - 1$ и шага по времени Δt (величина $\eta_{\Delta t} = |T/T' - 1| \times 100\%$, где T – время из текущего столбца, T' – время из предыдущего столбца, введена для характеристики “практической сходимости”, см. [7]):

n	10	20	40	60
$T_{\Delta t} = 0.003$	0.477	0.432	0.405	0.381
$\eta_{\Delta t} = 0.003$, %	–	9.4	6.3	5.9
$T_{\Delta t} = 0.0015$	0.474	0.4305	0.4065	0.3840
$\eta_{\Delta t} = 0.0015$, %	–	9.2	5.9	5.5

Анализируя этот вывод, замечаем, что величина $\eta_{\Delta t}$ уменьшается с ростом n. Это означает, что имеет место “практическая сходимость”.



Фиг. 3.

Конечное положение (при $t = T$) поверхности раздела жидкостей σ_t представлено на фиг. 3 ($n = 60$ и $\Delta t = 0.003$).

СПИСОК ЛИТЕРАТУРЫ

1. *Никольский Д.Н.* Математическое моделирование трехмерной задачи эволюции границы раздела жидкостей различной вязкости и плотности в однородном и безграничном грунте // Вычисл. методы и программирование. 2006. Т. 7. Разд. 1. С. 236–242; http://num-meth.srcc.msu.su/zhurnal/tom_2006/v7r129.html.
2. *Пивень В.Ф., Буравлев И.В.* Исследование трехмерной задачи эволюции границы раздела жидкостей к несовершенной скважине // Тр. XII Междунар. симпозиума МДОЗМФ-2005. Харьков–Херсон, 2005. С. 275–278.
3. *Голубева О.В.* Курс механики сплошных сред. М.: Высш. школа, 1972.
4. *Lifanov I.K., Nikolsky D.N., Piven' V.F.* Mathematical modeling of the work of the system of wells in a layer with the exponential law of permeability variation and the mobile liquid interface // Russ. J. Numer. Analys. Math. Modeling. 2002. V. 17. № 4. P. 381–391.
5. *Пивень В.Ф.* Фундаментальные решения уравнений физических процессов, протекающих в неоднородных средах // Тр. Междунар. школы-семинара МДОЗМФ. Вып. 3. Орел, 2004. С. 43–53.
6. *Пивень В.Ф.* Единственность решения граничных задач сопряжения физических процессов в неоднородной среде // Тр. X Междунар. симпозиума МДОЗМФ-2001. Херсон, 2001. С. 265–269.
7. *Лифанов И.К.* Метод сингулярных интегральных уравнений и численный эксперимент. М.: ТОО Янус, 1995.
8. *Рябенький В.С.* Введение в вычислительную математику. М.: Физматлит, 2000.

ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 2007, том 47, № 8, с. 1413–1422

УДК 519.65

ОПРЕДЕЛЕНИЕ ХИМИЧЕСКОГО СОСТАВА И СТРУКТУРЫ НЕОДНОРОДНОЙ СРЕДЫ МЕТОДОМ РЕНТГЕНОВСКОЙ ТОМОГРАФИИ[1)]

© 2007 г. В. Г. Назаров

(690041 Владивосток, ул. Радио, 7, Ин-т прикл. матем. ДВО РАН)

e-mail: naz@iam.dvo.ru

Поступила в редакцию 02.05.2006 г.
Переработанный вариант 26.03.2007 г.

Рассматривается вопрос об определении химического состава неоднородного тела, состоящего из нескольких однородных частей, методом рентгеновской томографии. На первом этапе с помощью индикатора неоднородностей определяется внутренняя структура тела. Затем, используя полученную информацию, при некоторых дополнительных предположениях о свойствах частей предлагается метод частичного или полного определения химического состава каждой части тела. Математически задача сводится к решению уравнения переноса излучения и систем линейных алгебраических уравнений. На основе метода компьютерного моделирования выполнен ряд численных экспериментов. Результаты расчетов иллюстрируются серией графиков и томограмм. Библ. 11. Фиг. 3. Табл. 1.

Ключевые слова: уравнение переноса излучения, рентгеновская томография, численный метод решения задачи определения химического состава среды.

1. ВВЕДЕНИЕ

Определение химического состава вещества радиографическим методом является интересной задачей с точки зрения теории и, кроме того, имеет большую практическую значимость. Среди последних работ, посвященных этому вопросу, отметим [1], [2]. Существуют патенты, связанные с практической реализацией различных радиографических методов нахождения химического состава (см., например, [3], [4]). В данной работе рассматривается следующая

Задача. Определить внутреннюю структуру и химический состав неоднородного тела, состоящего из конечного числа однородных по составу частей и занимающего в $\mathbb{R}^3$ ограниченную выпуклую область G, по результатам просвечивания этого тела рентгеновским излучением на диапазоне энергии $[E_\alpha, E_\beta]$.

Такая постановка задачи представляется оправданной в случаях, когда, например, речь идет о нахождении внутреннего устройства промышленного изделия, разборка которого на части невозможна.

Ввиду большого количества различных условий, которые ставятся при решении задачи (условия на гладкость границ частей тела, входящее излучение, структуру сечений тела плоскостями, химический состав частей), мы сочли целесообразным приводить эти условия по ходу изложения решения.

На первом этапе решения задачи кратко излагается способ нахождения границ неоднородностей с помощью индикатора неоднородностей. Подробное описание этого метода можно найти в [5, 6]. Основная часть работы посвящена второму этапу решения задачи – способу определения химического состава неоднородностей. При этом с учетом информации, полученной на первом этапе, посредством энергетического сканирования тела по конечному набору специально выбранных прямых производится определение химических элементов, входящих в состав каждой неоднородности, находятся их массовые доли и плотности материалов, составляющих неоднородности. Полученные теоретические результаты иллюстрируются описанием одного из проведенных численных экспериментов.

1) Работа выполнена при государственной поддержке научных исследований, проводимых ведущими научными школами РФ (грант НШ-9004.2006.1) и в рамках гранта № 06-II-СУ-01-001 конкурса интеграционных проектов ДВО РАН с научными учреждениями Сибирского и Уральского отделений РАН.

2. НАХОЖДЕНИЕ СТРУКТУРЫ ТЕЛА

В работах [5], [6] изучалась следующая задача рентгеновской томографии. Пусть неоднородная среда занимает некоторую ограниченную выпуклую область G в трехмерном пространстве, состоящую из конечного числа попарно непересекающихся подобластей, заполненных различными материалами, и подвергается рентгеновскому облучению. Считается, что энергия фотонов сравнительно невелика и процесс переноса излучения хорошо описывается следующим стационарным интегродифференциальным уравнением переноса (см. [5]):

$$\omega \cdot \nabla_r f(r, \omega, E) + \mu(r, E) f(r, \omega, E) = \frac{1}{4\pi}\int_\Omega k(r, \omega \cdot \omega', E) f(r, \omega', E) d\omega' + J(r, \omega, E). \tag{1}$$

Здесь r – точка трехмерного евклидова пространства $\mathbb{R}^3$ с ортонормированным базисом e_1, e_2, e_3, $\Omega = \{\omega \in \mathbb{R}^3 : |\omega| = 1\}$, $\omega, \omega' \in \Omega$, E – энергия излучения, $f(r, \omega, E)$ – плотность потока фотонов в точке r, движущихся в направлении ω и имеющих энергию E, $\mu(r, E)$ – коэффициент полного взаимодействия фотонов с веществом, $k(r, \omega\omega', E)$ – индикатриса рассеяния, $J(r, \omega, E)$ – плотность внутренних источников излучения. В записи $\omega \cdot \nabla_r$ и $\omega \cdot \omega'$ точка означает скалярное произведение векторов из $\mathbb{R}^3$. При упругом рассеянии фотонов их энергия не меняется, поэтому переменная E в этом управлении присутствует в качестве параметра. В дальнейшем для нас будет важна зависимость коэффициентов μ и k от энергии, поэтому эта переменная в уравнении сохранена.

Функции $\mu_s(r, E)$ и $\mu_a(r, E)$, определенные равенствами

$$\mu_s(r, E) = \frac{1}{4\pi}\int_\Omega k(r, \omega \cdot \omega', E) d\omega', \quad \mu_a = \mu(r, E) - \mu_s(r, E),$$

называются коэффициентом рассеяния и коэффициентом поглощения энергии соответственно.

Считаем, что на границах подобластей по крайней мере один из коэффициентов μ, k, J уравнения (1) имеет разрыв I рода по пространственной переменной. Далее эти подобласти будем также называть неоднородностями. Задача состоит в нахождении границ подобластей (неоднородностей) с помощью уравнения переноса при условии, что на границе области известна плотность выходящего из нее излучения $H(\eta, \omega)$, $\eta \in \partial G$, $\omega \in \Omega$.

Для нахождения границ использовалась одна из версий индикатора неоднородностей (см. [6]). Остановимся вкратце на сути метода. Пусть P – подпространство в $\mathbb{R}^3$, порожденное векторами e_1, e_2, пересечение $D = G \cap P$ не пусто и $D_1, \ldots, D_q$ – совокупность всех плоских областей в D, которые получаются при пересечении плоскости P с неоднородностями из G, $q \geq 2$:

$$\bar{D} = \bigcup_{i=1}^{q} \bar{D}_i, \quad D_0 = \bigcup_{i=1}^{q} D_i, \quad D_i \cap D_j = \varnothing, \quad i \neq j, \quad i, j = 1, 2, \ldots, q.$$

Пусть $\Omega_1 = \Omega \cap P$, $\nabla_2 = e_1\dfrac{\partial}{\partial r_1} + e_2\dfrac{\partial}{\partial r_2}$, $d(r, \omega)$ – расстояние от точки $r \in D_0$ до ∂G в направлении ω. Функция $V(r, E)$, определенная равенством

$$V(r, E) = \left|\nabla_2 \int_{\Omega_1} f(r + d(r, \omega)\omega, \omega, E) d\omega\right|, \quad r \in D_0, \quad \omega \in \Omega_1, \tag{2}$$

называется *индикатором неоднородностей*. Заметим, что под интегралом в (2) стоит плотность выходящего из G излучения в точках $r + d(r, \omega)\omega \in P \cap \partial G$, а сам интеграл берется по направлениям, лежащим в P. При выполнении некоторых естественных условий гладкости на коэффициенты уравнения (1) и границы неоднородностей функция $V(r, E)$ непрерывна по r (см. [6]).

Пусть z – точка контакта двух соседних областей D_i и D_j, $1 \leq i < j \leq q$, $z \in D \cap \partial D_i \cap \partial D_j$, в точке z кривые ∂D_i, ∂D_j являются гладкими, $n(z)$ – нормаль к ∂D_i в точке z; тогда существует число $\delta > 0$ такое, что для $r = z + tn(z)$, $0 < |t| < \delta$, справедливо равенство

$$V(r, E) = 2|M(z, E)||\ln|r - z|| + F(r, E), \tag{3}$$

где M и F – некоторые ограниченные функции. Таким образом, наличие логарифмической особенности у функции V позволяет найти те точки множества ∂D_0, в которых $M(z, E) \neq 0$ при данной энергии E.

В [5], [6] выводятся формулы для функции M. Ввиду громоздкости приводить их здесь мы не будем. Отметим лишь следующее. В случае когда внутренние источники излучения в среде отсутствуют и $J = 0$, а излучение $f(z, \omega, E)$ близко к изотропному в точке z для энергии E, величина $M(z, E)$ пропорциональна скачку $[\![\mu_a(z, E)]\!]$ коэффициента поглощения в точке $z \in D \cap \partial D_i \cap \partial D_j$, который, по определению, равен

$$[\![\mu_a(z, E)]\!] = \lim_{r' \to z}\mu_a(r', E) - \lim_{r'' \to z}\mu_a(r'', E), \quad r' \in D_i, \quad r'' \in D_j.$$

Во многих случаях энергию E можно подобрать так, чтобы скачок $[\![\mu_a(z, E)]\!]$, а вместе с ним и функция $M(z, E)$ были достаточно большими для того, чтобы можно было успешно найти множество ∂D_0. Таким образом будет найдена часть структуры области G, лежащая в плоскости P. Используя набор различных секущих плоскостей можно найти границы всех искомых неоднородностей в G.

3. ОПРЕДЕЛЕНИЕ ХИМИЧЕСКОГО СОСТАВА

Предлагаемый способ определения химического состава основан на том, что для различных химических элементов коэффициенты их полного взаимодействия имеют разрывы при различных значениях энергии E. Далее речь всегда будет идти о коэффициентах $\mu(E)$ и $\mu_a(E)$, относящихся к однородным (по пространственной переменной) веществам или химическим элементам, поэтому аргумент r у μ и μ_a опущен. Для более подробного изучения зависимости μ и μ_a от энергии E использовались таблицы из [7], которые содержат данные для 92 химических элементов от водорода до урана включительно и 48 веществ сложного химического состава, представляющих интерес для дозиметрии. Данные представлены в виде сеточных функций, определенных для ряда значений энергии из промежутка 1 кэВ–20 МэВ. Для получения информации об элементах с атомными номерами от $Z = 93$ до $Z = 100$ нами использовалась база данных из [8]. Была написана компьютерная программа, анализирующая данные о 100 химических элементах, основные результаты работы которой состоят в следующем.

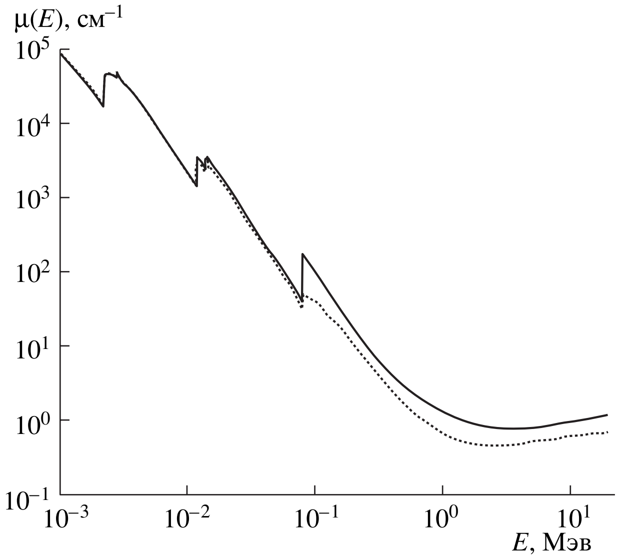
1. На промежутке 1 кэВ–20 МэВ функции $\mu(E)$ и $\mu_a(E)$ для каждого элемента непрерывны всюду, кроме конечного числа точек, в которых они одновременно имеют разрыв I рода. Такие значения энергии мы будем также называть скачками. Отметим, что наличие скачков у функций $\mu(E)$, $\mu_a(E)$ следует из физических моделей взаимодействия излучения с веществом, на основе которых строились таблицы в [7]. При этом величины скачков могут быть как очень маленькими, так и большими. Для элементов с номерами Z меньше 11 функции $\mu(E)$ и $\mu_a(E)$ непрерывны на рассматриваемом промежутке энергии.

2. Для разных элементов все точки разрыва E' функций $\mu(E)$, $\mu_a(E)$ всегда различны. Общее число всех различных скачков (для всех элементов) составляет 566. Самый первый скачок (в порядке возрастания по энергии) имеет место при $E' = 1.002$ кэВ (для фермия), а самый последний – при $E' = 143.1$ кэВ (тоже для фермия).

3. Для всех веществ в точках разрыва E' коэффициенты $\mu(E)$, $\mu_a(E)$ всегда возрастают, при этом величина $\mu^*(E') = \mu(E' + 0)/\mu(E' - 0)$ – отношение правого предельного значения коэффициента μ к левому предельному значению в точке E' меняется в пределах от 1.0044 (для рения) до 12.02 (для магния). Если рассмотреть множество скачков, в которое от каждого элемента взят только один скачок с максимальным значением $\mu^*(E')$, то наименьшее значение в нем составит $\mu^*(E') = 3.20$ для фермия. Это обстоятельство (достаточно большая величина $\mu^*(E')$) позволяет, как станет ясно из дальнейшего, утверждать, что нахождение любого химического элемента в составе материала во многих случаях может быть проведено успешно. Заметим, что все предельные значения $\mu(E' + 0)$ и $\mu(E' - 0)$, использованные здесь при расчетах, брались непосредственно из [7], поэтому точность этих расчетов напрямую зависит от точности данных в таблицах работы [7].

На фиг. 1 в качестве иллюстрации представлены графики зависимости $\mu(E)$ и $\mu_a(E)$ для золота. На горизонтальной оси отложено значение энергии E, Мэв, на вертикальной – значения $\mu(E)$, см$^{-1}$ (сплошная кривая), и $\mu_a(E)$, см$^{-1}$ (штриховая кривая). Масштаб по обеим осям логарифмический.

Для определения химического состава тела, внутреннюю структуру которого уже считаем найденной, используется наличие скачков у коэффициента $\mu(E)$ и индивидуальность набора таких скачков для каждого химического элемента. Далее будем считать, что выполняются следующие условия:



Фиг. 1.

а) область G подвергается внешнему облучению, плотность потока которого $h(\xi, \omega, E)$ на поверхности G непрерывна по всем переменным:

б) функция h имеет коллимированный характер по (ξ, ω), так что $h(\xi, \omega, E)$ отлично от нуля лишь для (ξ, ω), мало отличающихся от некоторых заданных (ξ_0, ω_0);

в) рассеянием в G можно пренебречь;

г) в G нет внутренних источников излучения ($J = 0$).

Пусть $r_0 \in D_0$, $\xi_0 = r_0 - d(r_0, -\omega_0)\omega_0$, $\eta_0 = r_0 + d(r_0, \omega_0)\omega_0$. Пара (r_0, ω_0) определяет прямую $\hat{l}_0$, лежащую в плоскости P, которая проходит через точку r_0 в направлении ω_0 и пересекает ∂D в точках ξ_0 и η_0. Пусть $H(\eta_0, \omega_0, E)$ есть плотность выходящего из G излучения в точке $\eta_0 \in \partial G$. Прямая $\hat{l}_0$ имеет непустые пересечения с некоторыми из плоских областей $D_1, \ldots, D_q$.

Обозначим через $A_1, \ldots, A_q$ вещества, соответствующие областям $D_1, \ldots, D_q$, так что каждая область D_i заполнена веществом A_i, $i = 1, 2, \ldots, q$, и пусть $\mu_1(E), \ldots, \mu_q(E)$ – коэффициенты полного взаимодействия веществ $A_1, \ldots, A_q$. Если $X_{i1}, \ldots, X_{ir}$ – все химические элементы, входящие в состав вещества A_i, $\mu_{xi1}(E), \ldots, \mu_{xir}(E)$ – соответствующие коэффициенты полного взаимодействия, $\rho_{xi1}, \ldots, \rho_{xir}$ – плотности, $w_{i1}, \ldots, w_{ir}$ – массовые доли элементов $X_{i1}, \ldots, X_{ir}$ в составе A_i и ρ_i – плотность вещества A_i, то, согласно [7], справедлива формула

$$\mu_i(E) = \rho_i \sum_{k=1}^{r} w_{ik} \frac{\mu_{xik}(E)}{\rho_{xik}}. \tag{4}$$

Отсюда видно, что множество всех разрывов функции $\mu_i(E)$ совпадает с объединением множеств всех разрывов функции $\mu_{xi1}(E), \ldots, \mu_{xir}(E)$. Функция $h(\xi_0, \omega_0, E)$ непрерывна по E, поэтому функция $H(\eta_0, \omega_0, E)$ будет иметь разрыв при некотором E' тогда и только тогда, когда на пути движения фотона, выходящего из G в точке η_0 и движущегося в направлении ω_0, встречался химический элемент, коэффициент $\mu(E)$ которого имел разрыв при $E = E'$. Проведя измерения функции $H(\eta_0, \omega_0, E)$ на диапазоне энергии $[E_\alpha, E_\beta]$, содержащем скачки функций $\mu(E)$ для всех химических элементов ($Z = 1$–100), мы по разрывам $H(\eta_0, \omega_0, E)$ найдем все элементы, которые встречались на траектории движения фотона. Для краткости множество всех таких химических элементов будем называть *спектром прямой* $\hat{l}_0$. Проведем такие измерения по всевозможным прямым, пересекающим D_0, и попытаемся по результатам измерений указать для каждой области D_i соответствующий ей набор химических элементов $X_{i1}, \ldots, X_{ir}$.

Сначала рассмотрим ситуацию, когда точность измерения выходящего из тела излучения $H(\eta_0, \omega_0, E)$ является невысокой (например, по причине технических трудностей измерения предельных значений $H(\eta_0, \omega_0, E' + 0)$, $H(\eta_0, \omega_0, E' - 0)$ в точках разрыва E'), так что мы с уверенностью можем говорить лишь о наличии разрывов у функции H при некоторых E'.

Введем следующие обозначения:

$A = \{A_1, \ldots, A_q\}$ – множество всех рассматриваемых веществ.

$\mathbb{L}$ – множество всех прямых $\hat{l}$ на P, для которых $\hat{l} \cap D_0 \neq \varnothing$,

D_i^+ – множество всех $\hat{l}$ из $\mathbb{L}$ таких, что $\hat{l} \cap D_i \neq \varnothing$,

D_i^- – множество всех $\hat{l}$ из $\mathbb{L}$ таких, что $\hat{l} \cap D_i = \varnothing$,

A_i^+ – максимальное подмножество в A такое, что для любого A_k из A_i^+ справедливо включение $D_i^+ \subset D_k^+$ (т.е. A_k включается в A_i^+ тогда и только тогда, когда всякая прямая, пересекающая D_i, пересекает и D_k),

A_i^- – максимальное подмножество в A такое, что для любого A_k из A_i^- множество $D_k^+ \backslash D_i^+$ непусто (т.е. A_k включается в A_i^- тогда и только тогда, когда существует прямая, которая пересекает D_k и не пересекает D_i).

Ясно, что множества D_i^+ и A_i^+ непусты для любого $i = 1, 2, \ldots, q$ и $A_i \in A_i^+$, а множества D_i^- и A_i^- могут оказаться пустыми для некоторых i. Множество A_i^+ состоит из всех веществ, которые входят в спектр каждой прямой из D_i^+, а множество A_i^- состоит из всех веществ, которые входят в спектр по крайней мере одной прямой из D_i^-.

Утверждение 1. *Для любых $i, j = 1, 2, \ldots, q$ множества A_i^+ и A_i^- можно построить с помощью конечного числа прямых из $\mathbb{L}$.*

Доказательство. Для множества D_i сначала выберем из набора $\{D_1, \ldots, D_q\}$ все множества D_k, для которых $D_k^+ \backslash D_i^+ \neq \varnothing$. Для таких D_k выберем по одной прямой $\hat{l}_k$ из $D_k^+ \backslash D_i^+$ и найдем ее спектр. Объединение всех этих спектров совпадает с A_i^-. Далее, из набора $\{D_1, \ldots, D_q\}$ выберем все множества D_t, для которых $D_i^+ \backslash D_t^+ \neq \varnothing$. Для таких D_t выберем по одной прямой $\hat{l}_t$ из $D_i^+ \backslash D_t^+$ и найдем ее спектр. Пересечение всех этих спектров совпадает с A_i^+. Утверждение 1 доказано.

Сформулируем следующие два условия.

Условие 1. Для любых $i, j = 1, 2, \ldots, q$, $i \neq j$ множества A_i и A_j не содержат одинаковых химических элементов.

Условие 2. Для любых $i, j = 1, 2, \ldots, q$, $i \neq j$ существует прямая $\hat{l}$ из $\mathbb{L}$ такая, что $\hat{l}$ имеет непустое пересечение с одним, и только одним, из множеств D_i, D_j.

Легко видеть, что условие 2 эквивалентно следующему:

$$D_i^+ \Delta D_j^+ = (D_i^+ \backslash D_j^+) \cup (D_j^+ \backslash D_i^+) \neq \varnothing, \quad i, j = 1, 2, \ldots, q, \quad i \neq j. \tag{5}$$

Утверждение 2. *Пусть выполнены условия* 1 *и* 2, *тогда верно следующее*:

1) *справедливы равенства* $A_i^+ \backslash A_i^- = A_i$, $i = 1, 2, \ldots, q$;

2) *множества A_i^+ и A_i^-, а вместе с ним и химический состав всех областей $D_1, \ldots, D_q$ могут быть найдены путем просвечивания области D_0 по конечному числу прямых из $\mathbb{L}$.*

Доказательство. Ясно, что $A_i \in A_i^+ \backslash A_i^-$. Если для $j \neq i$ будет $A_j \notin A_i^-$, то $D_j^+ \backslash D_i^+ = \varnothing$ и, в силу (5), $D_i^+ \backslash D_j^+ \neq \varnothing$, но тогда $A_j \notin A_i^+$, поэтому $A_i \notin A_i^+ \backslash A_i^-$ и равенство доказано. Второе утверждение следует из утверждения 1. Утверждение 2 доказано.

Если A_i – химический элемент, то химический состав области D_i можно считать найденным. В противном случае при использовании изложенного метода мы можем только перечислить все химические элементы, входящие в состав D_i.

Если каждое сечение области G необходимым нам набором секущих плоскостей удовлетворяет приведенным выше условиям 1 и 2, то можно найти химический состав всех частей тела.

Условие 1 довольно обременительно, но если в точках разрыва по энергии E' величины $H(\eta_0, \omega_0, E' + 0)$ и $H(\eta_0, \omega_0, E' - 0)$ могут быть определены с приемлемой точностью, то от него можно отказаться и для каждой подобласти D_i найти перечень всех находящихся в ней химических элементов, их массовые доли и плотность материала в этой подобласти. Рассмотрим эту ситуацию.

Далее используем следующие обозначения:

$X_1, \ldots, X_N$ – перечень всех химических элементов, присутствующих в D_0,

$\mu_{x1}(E), \ldots, \mu_{xN}(E)$ – коэффициенты полного взаимодействия для $X_1, \ldots, X_N$.

Для каждого X_i выберем какую-нибудь энергию E_i, на которой функция $\mu_{xi}(E)$ имеет скачок; тогда, в силу ранее сказанного, имеем

$$\begin{aligned} &\mu_{xi}(E_j + 0) - \mu_{xi}(E_j - 0) > 0, \quad i = j, \\ &\mu_{xi}(E_j + 0) - \mu_{xi}(E_j - 0) = 0, \quad i \neq j, \end{aligned} \tag{6}$$

$\rho_{x1}, \ldots, \rho_{xN}$ – плотности $X_1, \ldots, X_N$ соответственно, $\rho_1, \ldots, \rho_N$ – плотности (в общем случае сложных по составу) веществ из областей $D_1, \ldots, D_q$, $\mu_1(E), \ldots, \mu_q(E)$ – коэффициенты полного взаимодействия в областях $D_1, \ldots, D_q$, w_{kj} – массовая доля химического элемента X_j в области D_k, $k = 1, 2, \ldots, q$, $j = 1, 2, \ldots, N$, $\hat{l}$ – прямая из $\mathbb{L}$, l_k – линейная мера множества $\hat{l} \cap D_k$, $k = 1, 2, \ldots, q$, $h(E)$, $H(E)$ – плотности входящего в D и выходящего из D излучения в точках ∂D и для направления, определяемого прямой $\hat{l}$ для энергии E.

В этом случае справедливы равенства для правых и левых предельных значений:

$$H(E_j \pm 0) = h(E_j \pm 0)\exp\left(-\sum_{k=1}^{q} l_k\mu_k(E_j \pm 0)\right), \quad j = 1, 2, \ldots, N. \tag{7}$$

С учетом сказанного для функций $\mu_k(E)$ имеем (см. формулу (4))

$$\mu_k(E_j \pm 0) = \rho_k \sum_{i=1}^{N} w_{ki}\frac{\mu_{xi}(E_j \pm 0)}{\rho_{xi}}, \quad k = 1, 2, \ldots, q, \quad j = 1, 2, \ldots, N. \tag{8}$$

Введем обозначения $H_j^{\pm} = H(E_j \pm 0)$, $h_j^{\pm} = h(E_j \pm 0)$, $\mu_{kj}^{\pm} = \mu_k(E_j \pm 0)$, $\mu_{xij}^{\pm} = \mu_{xi}(E_j \pm 0)$, $i, j = 1, 2, \ldots, N$, $k = 1, 2, \ldots, q$. Тогда, в силу непрерывности h по E, из (7) получим

$$\frac{H_j^+}{H_j^-} = \exp\left(-\sum_{k=1}^{q} l_k(\mu_{kj}^+ - \mu_{kj}^-)\right), \quad j = 1, 2, \ldots, N,$$

или

$$\sum_{k=1}^{q} l_k(\mu_{kj}^+ - \mu_{kj}^-) = \ln(H_j^-/H_j^+), \quad j = 1, 2, \ldots, N. \tag{9}$$

Из (8) и (6) имеем

$$\mu_{kj}^+ - \mu_{kj}^- = \rho_k w_{kj}\frac{(\mu_{xjj}^+ - \mu_{xjj}^-)}{\rho_{xj}}, \quad j = 1, 2, \ldots, N. \tag{10}$$

Подставляя (10) в (9), получаем

$$\sum_{k=1}^{q} l_k\rho_k w_{kj} = \gamma_j, \quad j = 1, 2, \ldots, N, \tag{11}$$

где

$$\gamma_j = \frac{\rho_{xj}\ln(H_j^-/H_j^+)}{\mu_{xjj}^+ - \mu_{xjj}^-}, \quad j = 1, 2, \ldots, N, \tag{12}$$

$$\sum_{j=1}^{N} w_{kj} = 1, \quad w_{kj} \ge 0, \quad k = 1, 2, \ldots, q. \tag{13}$$

Прокомментируем формулы (11)–(13). Для каждого химического элемента X_j величина $\mu^+_{xjj} - \mu^-_{xjj} = \mu_{xj}(E_j + 0) - \mu_{xj}(E_j - 0)$ всегда положительна. Величины H^-_j и H^+_j считаем известными из результатов просвечивания тела вдоль прямой $\hat{l}$, величины ρ_{xj}, μ^+_{xjj}, μ^-_{xjj}, $j = 1, 2, \ldots, N$, известны из таблиц работы [7], величины l_k, $k = 1, 2, \ldots, q$, считаем известными из результатов решения задачи о нахождении границ неоднородностей. Величины ρ_k и w_{kj}, $k = 1, 2, \ldots, q$, $j = 1, 2, \ldots, N$, нам неизвестны. В итоге мы имеем $N + q$ уравнений (11), (13) относительно $Nq + q$ неизвестных. В случае $q = 1$ система (11), (13) легко разрешима. В этом случае $l_1\rho_1 w_{1j} = \gamma_j$, $j = 1, 2, \ldots, N$, откуда

$$\rho_1 = \frac{1}{l_1}\sum_{j=1}^{N}\gamma_j, \quad w_{1j} = \gamma_j\left(\sum_{i=1}^{N}\gamma_i\right)^{-1}. \tag{14}$$

Для решения задачи при $q \ge 2$ сделаем следующее. Из $\mathbb{L}$ выберем q различных прямых $\hat{l}_1, \ldots, \hat{l}_q$ и обозначим через l_{tk} линейную меру множества $\hat{l}_t \cap D_k$, $t, k = 1, 2, \ldots, q$. Проведем процедуру "просвечивания" тела вдоль каждой прямой $\hat{l}_t$ для значений энергии $E_j \pm 0$, $j = 1, 2, \ldots, N$, и определим величины γ_{tj}, $t = 1, 2, \ldots, q$, $j = 1, 2, \ldots, N$, так же как мы это делали раньше для прямой $\hat{l}$ (первый индекс у γ_{tj} – номер прямой $\hat{l}_t$, а второй – номер энергии E_j). Рассмотрим возникающую при этом систему уравнений

$$\sum_{k=1}^{q} l_{tk}\rho_k w_{kj} = \gamma_{tj}, \quad t = 1, 2, \ldots q, \quad j = 1, 2, \ldots, N, \tag{15}$$

где

$$\gamma_{tj} = \frac{\rho_{xj}\ln(H^-_{tj}/H^+_{tj})}{\mu^+_{xjj} - \mu^-_{xjj}}, \quad H^{\pm}_{tj} = H_t(E_j \pm 0), \quad t = 1, 2, \ldots, q, \quad j = 1, 2, \ldots, N,$$

относительно неизвестных ρ_k, w_{kj}. Совокупная система уравнений (15), (13) содержит $Nq + q$ уравнений относительно $Nq + q$ неизвестных. Справедлива следующая

Теорема. *Если $q \times q$-матрица $L = \|l_{tk}\|$ не вырождена, то система уравнений и условий* (15), (13) *разрешима относительно переменных* ρ_k, w_{kj} *и ее решение единственно.*

Доказательство. Будем рассматривать (15) как систему линейных уравнений относительно Nq неизвестных произведений $\rho_k w_{kj}$, которые упорядочим следующим образом: ($\rho_1 w_{11}$, $\rho_2 w_{21}$, … …, $\rho_q w_{q1}$, $\rho_1 w_{12}$, …, $\rho_q w_{q2}$, …, $\rho_1 w_{1N}$, …, $\rho_q w_{qN}$). Уравнения расположим так, чтобы при фиксированном j сначала менялся индекс t от 1 до q, а после этого j возрастал на 1 и все повторялось. Тогда основная матрица, соответствующая системе (15), будет иметь диагонально-блочный вид, при котором на диагонали будет расположено N матриц L:

$$\begin{pmatrix} L & & 0 \\ & \ddots & \\ 0 & & L \end{pmatrix}.$$

Детерминант этой матрицы равен $|L|^q$ и поэтому не вырожден. Значит, система (15) имеет единственное решение $\rho_k w_{kj} = a_{kj}$, $k = 1, 2, \ldots, q$, $i = 1, 2, \ldots, N$. В силу условий (13) получим

$$\rho_k = \sum_{j=1}^{N}\rho_k w_{kj} = \sum_{j=1}^{N} a_{kj}, \quad k = 1, 2, \ldots, q, \tag{16}$$

откуда

$$w_{kj} = a_{kj}/\rho_k = a_{kj}\left(\sum_{i=1}^{N} a_{ki}\right)^{-1}, \quad k = 1, 2, \ldots, q, \quad j = 1, 2, \ldots, N. \tag{17}$$

С другой стороны, системе (15) удовлетворяет по крайней мере один набор конкретных данных ρ'_k, w'_{kj}, $k = 1, 2, \ldots, q$, $j = 1, 2, \ldots, N$, удовлетворяющий также системе уравнений и условий (13), поэтому $\rho_k w_{kj} = \rho'_k w'_{kj} = a_{kj}$. Значит, ρ'_k и w'_{kj} удовлетворяют равенствам (16) и (17), поэтому $\rho_k = \rho'_k$ и $w_{kj} = w'_{kj}$ для $k = 1, 2, \ldots, q$, $j = 1, 2, \ldots, N$. Теорема доказана.

Замечание. Используя диагонально-блочный вид системы (15), ее решение a_{kj}, $k = 1, 2, \ldots, q$, $j = 1, 2, \ldots, N$, можно записать в сравнительно компактной форме. Пусть $L(k, j)$ – матрица, которая получается из L путем замены элементов ее k-го столбца $l_{1k}, \ldots, l_{qk}$ на числа $\gamma_{1j}, \ldots, \gamma_{qj}$ соответственно и $|L(k, j)|$ – детерминант матрицы $L(k, j)$, тогда, как легко видеть,

$$a_{kj} = \rho_k w_{kj} = \frac{|L(k, j)||L|^{q-1}}{|L|^q} = \frac{|L(k, j)|}{|L|}.$$

Из доказанной теоремы также следует, что при выполнении перечисленных выше условий задача определения химического состава разрешима и ее решение единственно.

4. ВЫЧИСЛИТЕЛЬНЫЕ ЭКСПЕРИМЕНТЫ

Проиллюстрируем полученные результаты на примере одного из проведенных численных экспериментов. В этом эксперименте область G представляла собой шар радиуса $R = 1$ мкм с центром в начале координат и содержала подобласти G_1 и G_2. При этом G_2 была частью однополостного гиперболоида, задаваемого условиями

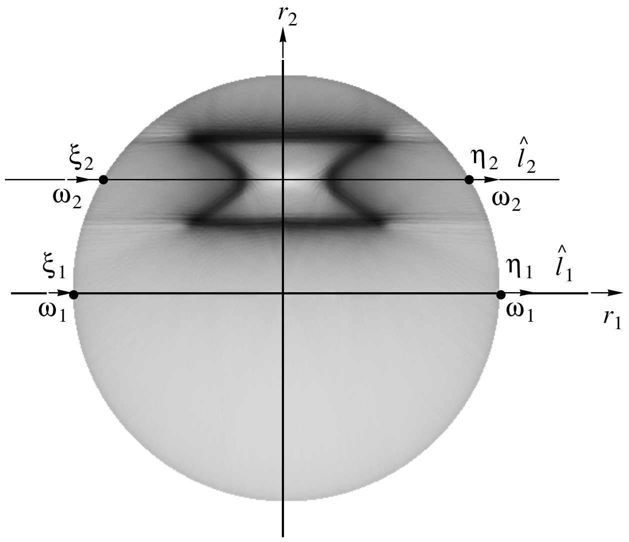
$$\frac{(r_1/R)^2}{(0.2)^2} - \frac{(r_2/R - 0.5)^2}{(0.1)^2} + \frac{(r_3/R)^2}{(0.2)^2} - 1 < 0, \quad 0.3 < r_2/R < 0.7, \tag{18}$$

а $G_1 = G \backslash G_2$. Подобласть G_1 была заполнена фосфидом алюминия AlP, а G_2 – алюмоселенидом меди $CuAlSe_2$. Таким образом, области G_1 и G_2 соответствовали различным однородным веществам, не содержащим внутренних источников излучения.

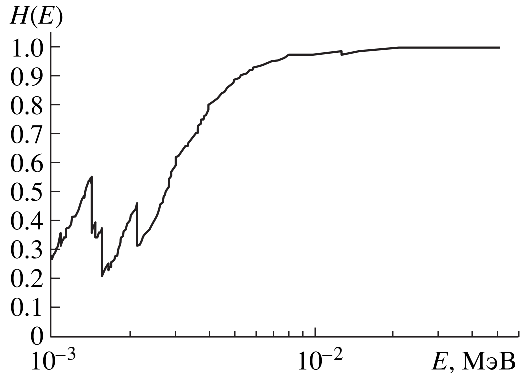
На первом этапе эксперимента решалась задача определения плотности H выходящего из G излучения при заданных коэффициентах уравнения (1) и плотности падающего на G внешнего потока излучения $h(E) = 1$. Энергия E при этом была фиксированной, значения коэффициентов μ и k в G_1 и G_2 вычислялись с помощью таблиц из [7] (их можно также найти на информационном ресурсе [10]). На окружности $D = G \cap P$ задавалось некоторое семейство точек η_i, $i = 1, 2, \ldots, N_\rho$, и выходящих из G направлений $\omega_k \in P$, $k = 1, 2, \ldots, N_\varphi$. Во всех экспериментах значения сеточных параметров составляли $N_\rho = 121$, $N_\varphi = 180$. Уравнение (1) решалось численно, при этом находились величины $f(\eta_i, \omega_k, E)$. Для нахождения функции f применялась одна из версий метода Монте-Карло, называемая методом сопряженных блужданий (см. [11]). Для каждой точки (η_i, ω_k) производился розыгрыш 1000 сопряженных траекторий, в которых учитывалось 10 актов рассеяния.

На втором этапе решалась задача томографии об определении части границы множества (18), лежащей в плоскости P. Для этого использовался индикатор неоднородности (2). Напомним, что под знаком интеграла в (2) находится плотность выходящего потока излучения, найденная на первом этапе. Вычисления проводились следующим образом. В плоскости P рассматривался квадрат с вершинами в точках (–1, –1, 0), (1, –1, 0), (1, 1, 0), (–1, 1, 0) (размерность длины в микронах здесь и далее опущена), содержащий сечение шара G. На этом квадрате задавалась равномерная сетка размером 401×401, и в узлах этой сетки, принадлежащих сечению шара G, вычислялся индикатор (2). Всем остальным узлам приписывалось значение 0. Полученная таким образом числовая матрица представлялась в виде изображения, на котором бóльшим значениям функция $V(r, E)$ соответствовал более темный цвет. Расчеты проводились при $E = 1.5$ кэВ. Результаты этой работы представлены на фиг. 2.

На третьем этапе эксперимента сначала производился выбор прямых $\hat{l}_1$ и $\hat{l}_2$ из $\mathbb{L}$. В качестве $\hat{l}_1$ была взята прямая, проходящая через точку $\eta_1 = (1, 0, 0)$ в направлении $\omega_1 = (1, 0, 0)$, а прямая $\hat{l}_2$ проходила через точку $\eta_2 = (\sqrt{3}/2, 1/2, 0)$ в том же направлении $\omega_2 = (1, 0, 0)$. При этом $l_{11} = 2$, $l_{12} = 0$, $l_{21} = \sqrt{3} - 0.4$, $l_{22} = 0.4$. Таким образом, матрица L была невырожденной. В диапазоне энергии $[E_\alpha, E_\beta]$, где $E_\alpha = 1$ кэВ, $E_\beta = 144$ кэВ, рассматривалось 566 точек $E_1 < \ldots < E_{566}$, соответствующих всем имеющимся скачкам коэффициента μ для всех химических элементов и проводилось



Фиг. 2.



Фиг. 3.

"просвечивание" области G вдоль прямых $\hat{l}_1$ и $\hat{l}_2$ на этих значениях энергии для получения спектров этих прямых. Иными словами, путем численного решения уравнения (1) находились значения плотности выходящего излучения $H(\eta_1, \omega_1, E_i - 0)$, $H(\eta_1, \omega_1, E_i + 0)$, $i = 1, \ldots, 566$, для прямой $\hat{l}_1$ и $H(\eta_2, \omega_2, E_i - 0)$, $H(\eta_2, \omega_2, E_i + 0)$, $i = 2, \ldots, 566$ для прямой $\hat{l}_2$. В обоих случаях плотность входящего излучения $h(\xi, \omega, E)$ не зависела от энергии и задавалась функцией типа "шапочка" с максимальным значением 1, т.е. была коллимирована в некоторой малой окрестности соответствующей точки (ξ_1, ω_1) для прямой $\hat{l}_1$ и точки (ξ_2, ω_2) для прямой $\hat{l}_2$.

Таким образом, было найдено, что спектр $\hat{l}_1$ состоит из алюминия и фосфора, а спектр $\hat{l}_2$ – из алюминия, меди, селена и фосфора. На фиг. 3 показана зависимость от энергии функции $H(\eta_2, \omega_2, E)$ для прямой $\hat{l}_2$. Общее количество скачков у H равнялось 8, из них по одному скачку приходилось на Al и P, два – на Cu и три – на Se. Аналогично выглядел график функции $H(\eta_1, \omega_1, E)$ для $\hat{l}_1$.

После этого найденный перечень химических элементов упорядочивался: $X_1 = \mathrm{Al}$, $X_2 = \mathrm{P}$, $X_3 = \mathrm{Cu}$, $X_4 = \mathrm{Se}$ – и каждому X_i ставилась в соответствие какая-нибудь одна энергия $E_{j(i)}$, на которой данный химический элемент X_i имел хорошо выраженный скачок. После выполнения очевидных переобозначений ($E_{j(i)}$ переобозначено как E_i) были выбраны следующие значения энергий: $E_1 = 1.5596$ кэВ (для Al), $E_2 = 2.1455$ кэВ (для P), $E_3 = 8.9789$ кэВ (для Cu) и $E_4 = 1.6539$ кэВ (для Se).

Таблица

Значения	ρ_1(AlP)	w_{11}(Al)	w_{12}(P)	w_{13}(Cu)	w_{14}(Se)
Точное	2.420	0.4655578	0.5344423	0.0	0.0
Расчет	2.420	0.4655576	0.5344424	0.0	0.0
	ρ_2($CuAlSe_2$)	w_{21}(Al)	w_{22}(P)	w_{23}(Cu)	w_{24}(Se)
Точное	4.700	0.1086005	0.0	0.2557723	0.6356272
Расчет	4.72137	0.1093943	0.002568157	0.2534516	0.6345859

Количество траекторий при нахождении функции H на этих значениях энергии составляло 100000. После этого вычислялись величины γ_{ij}, входящие в правую часть системы уравнений (15).

На последнем, четвертом этапе эксперимента решалась система уравнений (15), (13) и находились искомые неизвестные ρ_k, ω_{kj}, $k = 1, 2$, $j = 1, 2, 3, 4$. Результаты этих расчетов приведены в таблице. Первые три строки таблицы относятся к подобласти G_1, заполненной AlP, содержат точные и расчетные значения плотности материала ρ (г/см^3) и массовые доли каждого химического элемента, присутствующего в G. Последние три строки таблицы относятся к подобласти G_2, заполненной $CuAlSe_2$. Как видно из этой таблицы, расчетные и точные значения искомых величин хорошо согласуются.

5. ЗАКЛЮЧЕНИЕ

Полученные результаты позволяют утверждать, что предлагаемый метод может быть использован для нахождения химического состава и структуры неоднородных тел малого размера. Недостатком данного метода является то обстоятельство, что хорошая точность при нахождении искомых величин может быть достигнута только для тел, размеры которых не превышают нескольких микрон. Использование больших значений энергии зондирующего излучения, при которых коэффициенты полного взаимодействия для веществ быстро убывают, оказывается невозможным ввиду отсутствия скачков у коэффициентов для таких энергий.

СПИСОК ЛИТЕРАТУРЫ

1. *Найденов С.В., Рыжиков В.Д.* Об определении химического состава методом мультиэнергетической радиографии // Письма в ЖТФ. 2002. Т. 28. Вып. 9. С. 6–13.
2. *Зотьев Д.В., Филиппов М.Н., Ягола А.Г.* Об одной обратной задаче количественного рентгеноспектрального микроанализа // Вычисл. методы и программирование. 2003. Т. 4. С. 26–32.
3. *Федосеев В.М.* Рентгеновский способ обнаружения вещества по значению его эффективного атомного номера: Патент РФ № 2095795. 1997.
4. *Румянцев А.Н., Мостовой В.И., Сухоручкин В.К., Яковлев Г.В.* Способ обнаружения и неразрушающего анализа веществ, содержащих ядра легких элементов: Патент РФ № 2095796. 1997.
5. *Аниконов Д.С., Ковтанюк А.Е., Прохоров И.В.* Использование уравнения переноса в томографии. М.: Логос, 2000.
6. *Anikonov D.S., Nazarov V.G., Prokhorov I.V.* Poorly visible media in X-ray tomography. Utrecht, Boston: VSP, 2002.
7. *Hubbell J.H., Seltzer S.M.* Tables of X-ray mass attenuation coefficients and mass energy-absorption coefficients 1 Kev to 20 Mev for elements $Z = 1$ to 92 and 48 additional substances of dosimetric interest. 1995.
8. *Berger M.J., Hubell J.H., Seltzer S.M. et al.* XCOM: Photon cross section database // Nat. Inst. Standards and Technol. Gaithersburg, MD, 2005. (version 1.3. [Online] Available: http://physics.nist.gov/xcom).
9. *Аниконов Д.С., Назаров В.Г.* Классификация неоднородных сред в томографии на основе показателя их контрастности // Оптика и спектроскопия. 2005. Т. 99. № 4. С. 674–679.
10. *Аниконов Д.С., Ковтанюк А.Е. и др.* База данных радиационных характеристик веществ, представляющих интерес в рентгенодиагностике. http://sxray.iam.dvo.ru/
11. *Марчук Г.И., Михайлов Г.А., Назарлиев М.А. и др.* Метод Монте-Карло в атмосферной оптике. Новосибирск: Наука, 1976.

ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 2007, том 47, № 8, с. 1423–1427

УДК 519.7

ОБ ОДНОМ МЕТОДЕ ПОСТРОЕНИЯ РАСПОЗНАЮЩЕГО АЛГОРИТМА В АЛГЕБРЕ НАД МНОЖЕСТВОМ ВЫЧИСЛЕНИЯ ОЦЕНОК[1)]

М. Ю. Романов

(141700 Долгопрудный, М. о., Институтский пер., 9, МФТИ)

e-mail: mromanov@ccas.ru

Поступила в редакцию 24.01.2007 г.

Предлагается метод построения алгоритма в алгебре над множеством вычисления оценок в алгебраическом расширении наименьшей степени. Библ. 5.

Ключевые слова: алгоритм распознавания, алгебра над множеством вычисления оценок.

В настоящей работе рассматривается метод построения распознающего алгоритма в алгебре над множеством вычисления оценок. Будем использовать обозначения, применяемые в [1].

Рассмотрим алгоритмы вычисления оценок (АВО). Такие алгоритмы составляются из распознающего оператора и решающего правила. Распознающий оператор вычисляет оценки близости объектов к классам, а решающее правило на основе этих оценок классифицирует объекты.

Для распознающих операторов (РО) можно ввести операции сложения, умножения на константу и операторного умножения таким образом, что это будут операции над оценками, т.е. над значениями операторов. Эти операции индуцируют соответствующие операции над алгоритмами распознавания. Над множеством распознающих операторов $B^* = \{B_1, B_2, \ldots, B_n\}$ введем линейное замыкание

$$L(B^*) = \{c_1B_1 + c_2B_2 + \ldots + c_nB_n\},$$

а также алгебраическое замыкание k-й степени

$$U^k(B^*) = L\{B_1 \cdot B_2 \cdot \ldots \cdot B_s : B_1, B_2, \ldots, B_s \in B^*, 1 \le s \le k\}.$$

Множество алгоритмов, построенное на алгебраическом замыкании k-й степени над некоторым набором РО, будем называть алгебраическим расширением k-й степени.

Для краткости степенью алгоритма A будем называть степень алгебраического расширения, в котором построен этот алгоритм. Иначе говоря, это можно понимать как степень полинома над операторами $\{B_k\}$, который является распознающим оператором алгоритма A.

Для применения корректного алгоритма A важной характеристикой является его степень. При решении практических задач имеет смысл рассматривать полиномы операторов небольшой степени, поэтому актуальной является проблема минимизации степени полинома.

В работе [2] показано, что для задачи с обучающей выборкой из q объектов и с l классами существует алгоритм степени не выше

$$[\log_2 ql]. \tag{1}$$

Однако при решении конкретных задач его степень может оказаться существенно ниже. В настоящей работе рассматривается метод построения полиномиального распознающего оператора наименьшей степени, при этом исходный набор распознающих операторов считается заданным.

Опишем формальную постановку задачи построения распознающего алгоритма. Введем обучающую выборку

$$\tilde{S}_m = \{S_i = (a_{i1}, \ldots, a_{in}), i = \overline{1, m}\}$$

[1)] Работа выполнена при финансовой поддержке РФФИ (коды проектов 05-01-00718, 06-07-89299) и гранта Президиума РФ по поддержке ведущих научных школ НШ-5833.2006.1.

и контрольную выборку

$$\tilde{S}^q = \{S^i = (b_{i1}, \ldots, b_{in}), i = \overline{1, q}\}.$$

Начальная информация I_0 задается обучающей выборкой и информационными векторами $\tilde{\alpha}(S_i)$, $i = \overline{1, m}$. Здесь

$$\tilde{\alpha}(S_i) = (\alpha_{ij} = P_j(S_i) \equiv (S_i \in K_j), j = \overline{1, l}).$$

Таким образом,

$$I_0 = \{\tilde{S}_m, \tilde{\alpha}(S_i), i = \overline{1, m}\}.$$

Задача распознавания $Z = (I_0, S^q)$ определяется начальной информацией I_0 и контрольной выборкой $\tilde{S}^q$. Для каждого контрольного вектора необходимо найти информационный вектор $\tilde{\beta}(S^i) = (\beta_{ij} = P_j(S^i) \equiv (S^i \in K_j), j = \overline{1, l})$.

Рассмотрим множество алгоритмов $\{A\}$ решения задачи $Z = (I_0, S^q)$, представленных в виде произведения операторов $A = B \cdot C$. Здесь оператор B определяет матрицу оценок $B(Z) = \|\Gamma_{ij}\|_{q \times l}$, где Γ_{ij} – действительные числа, а оператор C – матрицу ответов $C(\|\Gamma_{ij}\|_{q \times l}) = \|\beta_{ij}\|_{q \times l}$, где $\beta_{ij} \in \{0, 1, \Delta\}$.

Для задачи Z и заданного множества алгоритмов $\{A\}$ при выполнении некоторых условий можно построить алгоритм A^* (из алгебраического замыкания множества $\{A\}$ некоторой степени k), определяющий значения $P_j(S^i)$ без ошибок (см. [3], [4]).

Будем использовать стандартные пороговые решающие правила

$$C(\|\Gamma_{ij}\|_{q \times l}) = \|C(\Gamma_{ij})\|_{q \times l}, \quad C(\Gamma_{ij}) = \begin{cases} 0, & \Gamma_{ij} \le c_1, \\ 1, & \Gamma_{ij} \ge c_2, \quad 0 < c_1 < c_2, \\ \Delta, & c_1 < \Gamma_{ij} < c_2. \end{cases}$$

Элементы матрицы ответов можно разбить на два множества: $\|\beta_{uv}\|_{q \times l} = M_0 \cup M_1$, где $M_i = \{(u, v) : \beta_{uv} = i\}$, q – число объектов, l – число классов.

Оператор B с матрицей оценок $B(Z) = \|\Gamma_{uv}\|_{q \times l}$ называется *допустимым* для задачи Z, если существует хотя бы одна пара $(u, v) \in M_1$ такая, что для всех $(i, j) \in M_0$ справедливо неравенство

$$\Gamma_{uv}(B) > |\Gamma_{ij}(B)|.$$

Такая пара (u, v) называется *отмеченной* в B, и множество отмеченных в B пар обозначим через $M(B)$. Фактически это означает, что отмеченные пары "отделены" от всех пар (i, j) таких, что объект S^i не принадлежит классу K_j. Поэтому при некоторых порогах c_1, c_2 алгоритм $B \cdot C$ будет давать правильные ответы для всех пар, отмеченных в B.

Допускается альтернативный способ определения отмеченных точек. Пара (u, v) называется отмеченной в B при некоторой фиксированной величине δ, если для всех $(i, j) \in M_0$ имеет место неравенство

$$\Gamma_{uv}(B) \ge |\Gamma_{ij}(B)| + \delta.$$

При таком определении все дальнейшие рассуждения остаются справедливыми.

Пусть

$$\Gamma^0_{\max}(B) = \max_{(i, j) \in M_0} |\Gamma_{ij}(B)|,$$

$$\Gamma^1_{\min}(B) = \min_{(i, j) \in M(B)} |\Gamma_{ij}(B)|.$$

Введем обозначения

$$\Gamma(B) = [\Gamma^1_{\min}(B)]^{-1}, \quad Q(B) = \Gamma^0_{\max}(B)/\Gamma^1_{\min}(B).$$

Для оператора B можно ввести оператор $B' = \Gamma(B) \cdot B$. Определив в задаче $Z = (I_0, S^q)$ элементы матрицы оценок в виде $B'(Z) = \|\Gamma'_{ij}\|_{q \times l}$, получим $\Gamma'_{ij}(B) = \Gamma(B) \cdot \Gamma_{ij}(B)$.

В [1] доказана

Лемма 1. *Если* (u, v) *отмечена в* B, *то имеем* $\Gamma'_{ij}(B) \geq 1$; *если* (u, v) *не отмечена в* B, *то имеем* $\Gamma'_{ij}(B) \leq Q(B) < 1$.

Тогда, учитывая, что матрица оценок $B'(Z)$ получается из $B(Z)$ поэлементным умножением, получаем, что пара (u, v) является отмеченной в B' тогда и только тогда, когда она отмечена в B.

Будем рассматривать множество операторов $\{A_k : A_k = B_k \cdot C, k = \overline{1, n}\}$, где каждый оператор B_k дает матрицу оценок $B_k(Z) = \|\Gamma_k^{uv}\|_{q \times l}$. Предположим, что для исходного набора операторов уже выполнено описанное выше преобразование, т.е. для каждого оператора B_k выполняется такое свойство:

$$\text{если } (u, v) \text{ отмечена в } B_k, \text{ то } \Gamma_k^{uv} \geq 1, \text{ иначе } \Gamma_k^{uv} < 1. \tag{2}$$

Система $\{B_k\}$ называется *базисной* для Z, если $M_1 = \bigcup_{k = \overline{1, n}} M(B_k)$.

Тогда, по теореме 1 из [1], существует такой набор степеней $\{x_k, k = \overline{1, n}\}$, что для базисной системы $\{B_k\}$, удовлетворяющей свойству (2), алгоритм

$$A = \left\{ \sum_{k = \overline{1, n}} B_k^{x_k} \right\} \cdot C(c_1, c_2) \tag{3}$$

является корректным для задачи Z.

Для того чтобы алгоритм удовлетворял условию корректности, для набора $\{x_k\}$ должны выполняться следующие условия:

1) для любой пары $(u, v) \in M_0$ выполняется неравенство $\Gamma_{uv} = \sum_{k=1}^{n} (\Gamma_k^{uv})^{x_k} \leq c_1$;

2) для любой пары $(u, v) \in M_1$ выполняется неравенство $\Gamma_{uv} = \sum_{k=1}^{n} (\Gamma_k^{uv})^{x_k} \geq c_2$.

Пользуясь введенными обозначениями, можно сформулировать оптимизационную задачу нахождения степеней операторов $\{B_k\}$ в алгоритме A из формулы (3) так, чтобы степень алгоритма была минимальной:

$$\begin{gathered} \forall (u, v) \in M_0 : \sum_{k=1}^{n} (\Gamma_k^{uv})^{x_k} \leq c_1, \\ \forall (u, v) \in M_1 : \sum_{k=1}^{n} (\Gamma_k^{uv})^{x_k} \geq c_2, \\ \max_{k = \overline{1, n}} x_k \longrightarrow \min. \end{gathered} \tag{4}$$

В задаче (4) имеются n переменных и $q \times l$ ограничений.

Введем следующие обозначения: $y_k = e^{x_k}$, $\gamma_k^{uv} = \ln \Gamma_k^{uv}$, $\varphi^{uv}(\tilde{y}) = \sum_{k=1}^{n} y_k^{\gamma_k^{uv}}$. Тогда задача (4) примет следующий вид:

$$\begin{gathered} \forall (u, v) \in M_0 : \varphi^{uv}(\tilde{y}) \leq c_1, \\ \forall (u, v) \in M_1 : \varphi^{uv}(\tilde{y}) \geq c_2, \\ \max_{k = \overline{1, n}} y_k \longrightarrow \min. \end{gathered} \tag{5}$$

Используя ограничение сверху степени алгоритма, приведенное в формуле (1), можно утверждать, что для всякого решения $\tilde{y}$ оптимизационной задачи (5) выполняются неравенства

$$y_i \leq M, \quad i = \overline{1, n},$$

где в качестве M можно взять, например,

$$M = \exp([\log_2 ql]).$$

Тогда решение задачи не изменится, если к множеству ограничений добавить условия $y_i \le M$, $i = \overline{1, n}$. При этом множество ограничений запишется в виде

$$U = \{\tilde{y} : \forall(u, v) \in M_0 : \varphi^{uv}(\tilde{y}) \le c_1;\ \forall(u, v) \in M_1 : \varphi^{uv}(\tilde{y}) \ge c_2;\ \ y_i \le M,\ i = \overline{1, n}\}.$$

Таким образом, задача (5) эквивалентна задаче

$$R = \{\tilde{y} \in U, \max_{k = \overline{1, n}} y_k \longrightarrow \min\}.$$

Рассмотрим для каждого непустого множества индексов

$$I = \{i_1, i_2, \ldots, i_m\} \subset \{1, 2, \ldots, n\}$$

вспомогательную задачу

$$R_I = \{\tilde{y} \in U, y_{i_1} = \ldots = y_{i_m}, y_{i_1} \longrightarrow \min\}.$$

Обозначим переменные $y_{i_1}, \ldots, y_{i_m}$ через z_1, а переменные y_i, $i \in \{1, \ldots, n\} \backslash I$, – через $z_2, z_3, \ldots, z_{n-m+1}$. Тогда задача $\bar{R}_I$ может быть записана в виде

$$R_I = \{\tilde{z} \in U_I, z_1 \longrightarrow \min\}.$$

Здесь, вводя $\varphi_I^{uv}(\tilde{z})$ равным значению φ^{uv} от соответствующего $\tilde{y}$, полагаем

$$U_I = \{\tilde{z} : \forall(u, v) \in M_0 : \varphi_I^{uv}(\tilde{z}) \le c_1;\ \forall(u, v) \in M_1 : \varphi_I^{uv}(\tilde{z}) \ge c_2;\ \ z_i \le M\}.$$

Таким образом, из исходной задачи получен набор из $2^n - 1$ вспомогательных задач. К их решению применимы такие методы, как метод множителей Лагранжа, метод проекции градиента и метод линеаризации (см. [5]).

В частности, если для метода линеаризации представить U_I в виде $U_I = \{\tilde{z} : \psi_{uv}(\tilde{z}) \le 0, u = \overline{1, q}$, $v = \overline{1, l}$; $z_i \le M$, $i = \overline{1, n - m + 1}\}$, где

$$\psi_{uv}(\tilde{z}) = \begin{cases} \varphi_I^{uv}(\tilde{z}) - c_1 & \text{при } (u, v) \in M_0, \\ -\varphi_I^{uv}(\tilde{z}) + c_2 & \text{при } (u, v) \in M_1, \end{cases}$$

то получим последовательность оптимизационных задач вида

$$\psi_{uv}(\tilde{z}^k) + \langle \psi'_{uv}(\tilde{z}^k), (\tilde{z} - \tilde{z}^k)\rangle \le 0, \quad u = \overline{1, q}, \quad v = \overline{1, l},$$

$$0 \le z_i \le M, \quad i = \overline{1, n - m + 1},$$

$$\frac{1}{2}|\tilde{z} - \tilde{z}^k|^2 + \beta_k(z_1 - z_1^k) \longrightarrow \min.$$

В результате образуется последовательность задач с линейными ограничениями и квадратичным функционалом. При этом $(k + 1)$-е приближение $\tilde{z}^{k+1}$ выбирается как решение k-й оптимизационной задачи. Так как функция ψ_{uv} является полиномом, то ее производная выписывается аналитически.

В результате для каждого набора индексов I получаем решение $\tilde{z}_I$ задачи $\bar{R}_I$ и, соответственно, набор $\tilde{y}_I$, являющийся решением задачи R_I.

Теорема. *Существует I такое, что $\tilde{y}_I$ является решением задачи R.*

Доказательство. Пусть $\tilde{y}$ – решение задачи R, причем $\max\limits_{k=\overline{1,n}} \tilde{y}_k = y_j$. Пусть $I = \{i : y_i = y_j, i = \overline{1, n}\}$. Тогда $\tilde{y}$ удовлетворяет ограничениям задачи R_I и условию $\max\limits_{k=\overline{1,n}} \tilde{y}_k = y_j \longrightarrow \min$ на U. Следовательно, $\tilde{y}$ является решением задачи R_I.

Из доказательства следует, что кандидатами на решение задачи R могут быть только те решения задач R_I, для которых $\max\limits_{k=\overline{1,n}} (\tilde{y}_I)_k = (\tilde{y}_I)_{i_1}$, т.е. максимальными компонентами являются именно те, которые отобраны в набор I. Таким образом, достаточно среди полученных кандидатов выбрать то решение, для которого $(\tilde{y}_I)_{i_1}$ принимает минимальное значение. Полученный вектор будет давать решение задачи (5), и из него обратной подстановкой переменных легко получить решение исходной задачи (4).

Описанный метод поддается эффективному распараллеливанию. При реализации численного метода возможно сокращение перебора за счет исключения вспомогательных задач, заведомо не дающих решение исходной.

СПИСОК ЛИТЕРАТУРЫ

1. *Журавлёв Ю.И., Исаев И.В.* Построение алгоритмов распознавания, корректных для заданной контрольной выборки // Ж. вычисл. матем. и матем. физ. 1979. Т. 19. № 3. С. 726–738.
2. *Рудаков К.В.* Алгебраическая теория универсальных и локальных ограничений для алгоритмов распознавания: Дис. … докт. физ.-матем. наук. М.: ВЦ РАН, 1992.
3. *Журавлёв Ю.И.* Корректные алгебры над множеством некорректных (эвристических) алгоритмов. I // Кибернетика. 1977. № 4. С. 5–17.
4. *Журавлёв Ю.И.* Корректные алгебры над множеством некорректных (эвристических) алгоритмов. III // Кибернетика. 1978. № 2. С. 35–43.
5. *Васильев Ф.П.* Численные методы решения экстремальных задач. М.: Наука, 1988.

ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 2007, том 47, № 8, с. 1428–1454

УДК 519.6:519.710.24

ПРЕДСТАВЛЕНИЕ И ОБНАРУЖЕНИЕ ЗНАНИЙ В ЭКСПЕРТНЫХ СИСТЕМАХ ДЛЯ ЗАДАЧ РАСПОЗНАВАНИЯ ОБРАЗОВ[1)]

© 2007 г. О. М. Васильев, Д. П. Ветров, Д. А. Кропотов

(119992 Москва, Ленинские Горы, МГУ, ВМиК)

e-mail: ovasiliev@inbox.ru; vetrovd@yandex.ru; dkropotov@yandex.ru

Поступила в редакцию 01.09.2006 г.
Переработанный вариант 21.02.2007 г.

Предложен новый подход к проектированию нечетких экспертных систем. Подробно рассмотрены вопросы представления знаний и формирования высказываний средствами нечеткой логики, а также описана модель нечетких рассуждений. Основное внимание уделено вопросам автоматического получения знаний (нечетких правил вывода) по множеству прецедентов. В частности, введены различные критерии качества правил и предложен алгоритм их генерации (метод эффективных сужений). Описаны возможности расширения вида допустимых правил путем введения операции нечеткой дизъюнкции. Также исследованы возможности последующей оптимизации найденных правил. Представлены результаты экспериментов, демонстрирующие ценность предлагаемых подходов. Библ. 47. Фиг. 10. Табл. 2.

Ключевые слова: распознавание образов, поиск закономерностей в данных, нечеткая логика, экспертные системы.

ВВЕДЕНИЕ

Теория распознавания образов оперирует с задачами, связанными с предсказанием набора зависимых переменных по множеству независимых переменных, доступных для измерений и оценок, в ситуациях, когда известна некоторая эмпирическая информация об исследуемой взаимосвязи (см. [1, 2]). Для решения подобных задач разработано большое количество методов (см. [3]). Большинство из них (например, нейтронные сети, см. [4]) используют концепцию "черного ящика", т.е. просто выдают результат предсказания для данного набора признаков. Таким образом, связь между переменными внутри процесса остается неясной. Хотя для ряда задач одного прогноза оказывается достаточно, но зачастую возникают ситуации, когда необходимы дополнительные знания об исследуемом процессе (см. [5]). Последнее становится особенно актуальным в тех случаях, когда значения независимых переменных могут быть модифицированы исследователем. Тогда приходят к так называемой задаче data mining. После извлечения необходимых знаний можно вернуться к прогнозу зависимых переменных, принимая во внимание полученную информацию. В широком поле прикладных задач подобную дополнительную информацию получают при помощи экспертов данной прикладной области. В связи с этим необходимо, чтобы знания, извлекаемые автоматически при помощи компьютера в процессе data mining, были сформулированы в тех же терминах, что и информация, получаемая от экспертов. Таким образом, эксперт будет в состоянии оценить, добавить или отвергнуть часть полученной дополнительной информации. Системы подобного рода, в которых важную роль играют экспертные дополнительные знания об исследуемой области, получили название экспертных систем.

Исследования в области экспертных систем берут свое начало с середины 50-х годов прошлого столетия и уже пережили несколько этапов развития. За это время был реализован ряд успешных проектов по созданию экспертных систем, работающих на практике (см. [6, 7]). Однако большинство из них следует отнести к области нечеткого логического управления. Дело в том, что в задачах управления база знаний системы (совокупность управляющих правил) зачастую состоит из небольшого количества правил, непосредственно подсказываемых предметной областью. Поэтому основной проблемой при разработке таких систем становится реализация нечетких рассуждений или выводов. В этой области существует большое количество наработок,

[1)] Работа выполнена при финансовой поддержке РФФИ (коды проектов 05-07-90333, 06-01-00492, 06-01-08045, 07-01-00211), ИНТАС (YS 04-83-2942, 04-77-7036) и программы ОМН РАН № 02.

которые приводят к успешным результатам (см. [8–10]). При построении экспертных систем, ориентированных на решение задач распознавания и прогнозирования, проблема построения базы данных (представления знаний в системе), организация базы знаний (получение знаний от экспертов предметной области) и получение модели вычисления прогноза играют ключевую роль. Необходимость решения комплекса сложных проблем приводит к ограниченному распространению экспертных систем распознавания и прогноза на сегодняшний день.

В области экспертных систем выделяют три основные проблемы:

1) представление знаний;

2) использование знаний;

3) приобретение знаний.

Наиболее естественной формой **представления знаний** является совокупность утверждений типа "ЕСЛИ..., ТО..." (см. [9]). При этом условия в посылке правила (антецеденте), а также в заключении правила (выводе) должны формулироваться достаточно просто, чтобы эксперт мог относительно легко проинтерпретировать то или иное правило. Для задач распознавания образов в качестве условий, входящих в правила, традиционно выбираются набор отрезков (значение признака лежит в определенном диапазоне $a \le x \le b$, см. [11, 12]) или совокупность нечетких подмножеств значений признаков (см. [13]). Нечеткие подмножества предоставляют возможность задавать правила на естественном языке, т.е. в форме, наиболее приближенной к представлению знаний эксперта. В настоящей статье исследуются различные способы представления знаний в виде совокупности нечетких подмножеств, рассматриваются вопросы генерации и оптимизации представления.

Использование знаний для прогнозирования, как правило, представляет собой различные схемы голосования по набору закономерностей (правил, см. [3, 11]). В случае представления знаний с помощью нечетких подмножеств здесь подключается развитый аппарат нечеткой логики. При этом в конечном итоге механизм нечеткого вывода также может быть проинтерпретирован с точки зрения схемы голосования (см. [6]). В статье рассматриваются различные способы голосования и настройки весов правил.

Наиболее трудоемкой является задача **приобретения знаний**. Для решения этой проблемы требуются как талант исследователя находить общий язык с экспертами, так и специальные автоматические алгоритмы извлечения знаний из эмпирической информации. Алгоритмы автоматической генерации знаний, как правило, связаны с решением многопараметрических оптимизационных задач, в которых целевая функция обладает большим количеством локальных оптимумов (см. [11]). Более того, в случае экспертных систем возникает ряд дополнительных требований к множеству искомых закономерностей. Среди них можно отметить следующие:

а) относительно небольшое количество информативных закономерностей, для того чтобы эксперт мог просмотреть, оценить и/или изменить набор правил;

б) небольшое количество условий в посылках правил;

в) правила должны быть сформулированы относительно *исходных* признаков, выражения типа $\sqrt[3]{\text{Вес}}$ или $\sin(\text{Рост})$ являются недопустимыми.

Для поиска набора информативных закономерностей в экспертных системах обычно используют алгоритмы на базе нечетких нейтронных сетей (neuro-fuzzy approach, см. [14, 15]), генетических алгоритмов (см. [16], [17], 13]), а также гибридные подходы с использованием кластерного анализа (см. [18]). К сожалению, эти подходы обладают рядом недостатков:

– нечеткие нейронные сети (ННС) генерируют огромное количество правил относительно низкой информативности, что соответствует ситуации переобучения, а также не позволяет интерпретировать и редактировать набор правил;

– высокий порядок генерируемых правил для ННС;

– большая зависимость ННС от начального приближения;

– значительное время обучения ННС и генетических алгоритмов;

– необходимость определения числа кластеров в алгоритмах, использующих кластеризацию;

– большое количество параметров, настраиваемых пользователем, в генетических алгоритмах.

Можно отметить также такие алгоритмы генерации знаний, как решающие списки (см. [19]) и подходы с использованием процедуры адаптивной коррекции или бустинга (boosting, см. [20–23]). В этих алгоритмах применяется последовательная схема использования правил, что на практике значительно затрудняет или делает невозможной интерпретацию решений для эксперта.

Большое количество исследований посвящено алгоритмам с использованием решающих деревьев (см. [5, 24–26]). Эти алгоритмы обладают высокими точностными показателями и часто используются на практике. Однако интерпретация древовидной структуры в виде набора правил часто приводит к большому количеству “длинных” правил, посылки которых очень похожи друг на друга (см. [24]). Это также затрудняет интерпретацию решения.

В данной статье предлагается алгоритм генерации знаний, лишенный большинства упомянутых недостатков. Он строит небольшой набор компактных информативных правил, которые могут быть успешно использованы при решении задач.

Работа состоит из пяти разделов. В разд. 1 рассматриваются способы представления знаний в виде совокупности нечетких подмножеств. В разд. 2 рассматриваются вопросы использования знаний, описывается модель получения прогноза в случае нечетких закономерностей. Разд. 3 посвящен проблеме генерации знаний. В нем предлагаются автоматические алгоритмы поиска закономерностей, а также исследуются свойства алгоритмов. Разд. 4 посвящен вопросам оптимизации модели, поиска начального представления, настройки весов правил, поиска границ множеств в посылках правил и пр. В разд. 5 приводятся результаты экспериментов по применению разрабатываемых технологий для решения реальных прикладных задач, подводятся итоги и делаются выводы.

В дальнейшем будут использоваться следующие обозначения. Предположим, что имеется d действительных признаков (независимых переменных) и одна зависимая переменная, принимающая значения на конечном множестве $\{1, 2, \ldots, l, \Delta\}$ для задачи распознавания образов (классификации) с l классами или на множестве действительных чисел для задачи восстановления регрессии (прогнозирования). Прецедентная (эмпирическая) информация (обучающая выборка) представляет собой совокупность пар $\{\mathbf{x}^i, y^i\}_{i=1}^{q}$, где $\mathbf{x}^i \in \mathbb{R}^d$, $y^i \in \{1, 2, \ldots, l\}$, или $y^i \in \mathbb{R}$.

1. ПРЕДСТАВЛЕНИЕ ЗНАНИЙ

Особенностью экспертных систем является представление знаний в легко интерпретируемых терминах. Как уже отмечалось выше, наиболее доступной формой представления знаний являются выражения, сформулированные в лингвистическом виде. Такие знания формулируются на естественном и, следовательно, нечетком языке. В работе [9] предлагается формулировать знания в форме продукционных правил вида “ЕСЛИ..., ТО...”, при этом посылка правила представляет собой логическое выражение относительно нечетких подмножеств для признаков.

Определение 1. Пусть $\mathcal{I}$ – произвольное множество, называемое *полным пространством*. *Нечетким подмножеством* M полного пространства $\mathcal{I}$ называется множество упорядоченных пар

$$\{(x, \mu_M(x)) \mid x \in \mathcal{I}, \mu_M : \mathcal{I} \longrightarrow [0, 1]\}.$$

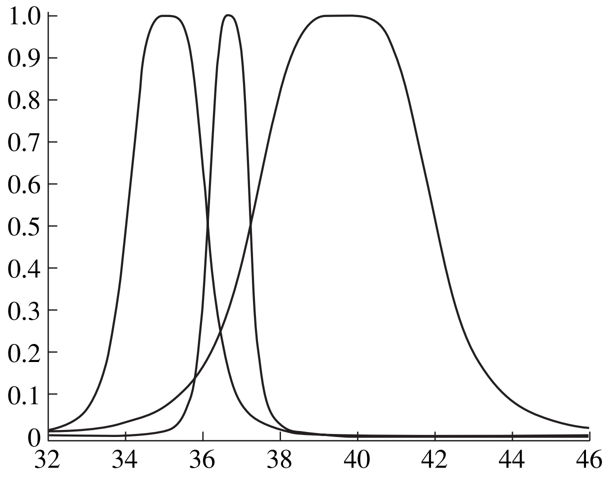
При этом, отображение $\mu_M(\cdot)$ называется *характеристической функцией* нечеткого подмножества M.

Рассмотрим произвольный числовой признак (один из тех, относительно которых сформулирована прецедентная информация). С точки зрения эксперта, область его изменения может быть описана вполне упорядоченным семейством подмножеств, каждому из которых соответствует некоторое лингвистическое значение, сформулированное на естественном языке. В качестве примера на фиг. 1 изображены характеристические функции трех нечетких подмножеств признака “температура тела пациента”. Соответствующие им лингвистические значения (слева направо): “НИЗКАЯ”, “НОРМАЛЬНАЯ”, “ВЫСОКАЯ”.

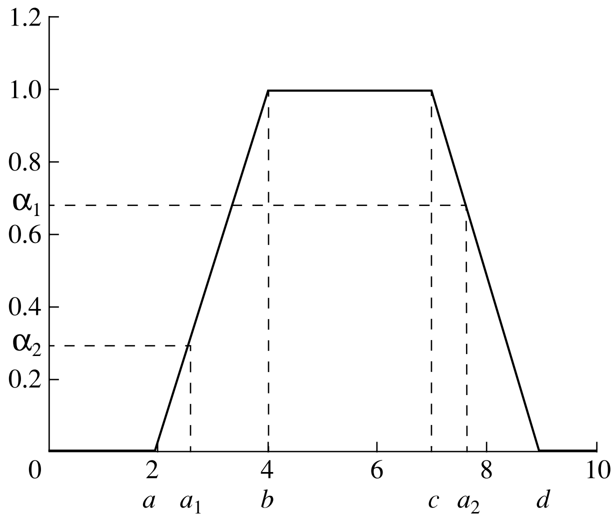
В общем случае можно предположить, что с точки зрения эксперта существует некоторое разбиение области значений признака, определяющее условные границы между различными состояниями. Это положение отражает следующее

Определение 2. *Экспертной интерпретацией признака* $i \in \{1, 2, \ldots, d\}$ для разбиения области его значений $a_i^1 < \ldots < a_i^{n_i}$ будем называть множество всех нечетких подмножеств числовой прямой с условными границами в соседствующих точках заданного разбиения:

$$\mathfrak{T}_i = \left\{M_i^j, j = \overline{1, n_i - 1} \mid \mu_{M_i^j}(\cdot) = \mu(x; a_i^j, a_i^{j+1}) \ \forall x \in \mathbb{R}\right\}.$$



Фиг. 1.



Фиг. 2.

При этом

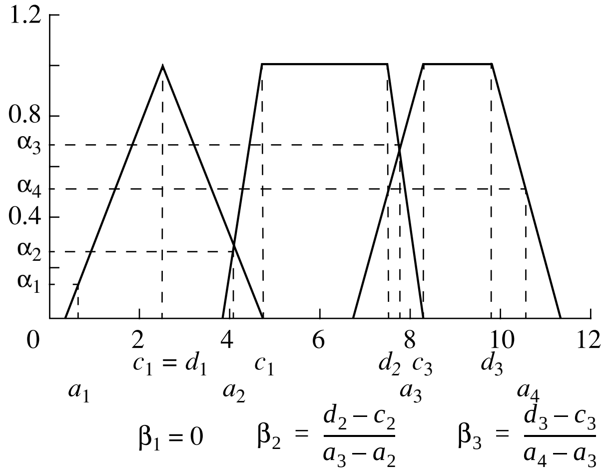
$$\forall j = \overline{1, n_i - 1} \ \exists x_* \in [a_i^j, a_i^{j+1}] : \mu_{M_i^j}(x_*; a_i^j, a_i^{j+1}) = \max_x [\mu_{M_i^j}(x; a_i^j, a_i^{j+1})].$$

Здесь выражение $\mu(x; a, b)$ обозначает характеристическую функцию нечеткого подмножества с условными границами a и b. Вид характеристической функции $\mu(\cdot)$ можно выбирать поразному. В дальнейшем будут рассматриваться семейства трапециевидных и колоколообразных функций. Связь между функцией нечеткого подмножества и условными границами будет уточнена ниже.

1.1. Семейство трапециевидных функций

В качестве базовой формы нечеткого подмножества можно рассматривать равнобедренную трапецию (фиг. 2). В частном случае она может иметь форму прямоугольника (если $a = b$, $c = d$), т.е. задавать четкое множество, а также треугольника (если $b = c$).

Определение 3. $\{\boldsymbol{\alpha}, \boldsymbol{\beta}\}$-*покрытием* признака для разбиения области его значений $a_1 < \ldots < a_n$ будем называть совокупность нечетких подмножеств $\{M_i\}_{i=1}^{n-1}$ с характеристическими трапецие-



Фиг. 3.

видными функциями (фиг. 2) такую, что верно следующее:

$$\alpha_1, \ldots, \alpha_n \in (0, 1), \quad \beta_1, \ldots, \beta_{n-1} \in [0, 1), \tag{1.1}$$

$$\mu_{M_i}(a_{i+1}) = \mu_{M_{i+1}}(a_{i+1}) = \alpha_{i+1} \quad \forall i = \overline{1, n-2}, \tag{1.2}$$

$$\mu_{M_1}(a_1) = \alpha_1, \quad \mu_{M_{n-1}}(a_n) = \alpha_n, \tag{1.3}$$

$$\exists x_* \in [a_i, a_{i+1}] : \mu_{M_i}(x_*) = \max_x[\mu_{M_i}(x)] \quad \forall i \in \overline{1, n-1}, \tag{1.4}$$

$$\frac{\left|\{x \in \mathbb{R} \mid \mu_{M_i}(x) = 1\}\right|}{a_{i+1} - a_i} = \beta_i \quad \forall i \in \overline{1, n-1}. \tag{1.5}$$

Понятие $\{\boldsymbol{\alpha}, \boldsymbol{\beta}\}$-покрытия проиллюстрировано на фиг. 3.

Справедлива следующая

Теорема 1. *$\{\boldsymbol{\alpha}, \boldsymbol{\beta}\}$-покрытие для разбиения $a_1 < \ldots < a_n$ определено однозначно.*

Доказательство. Трапеция определяется однозначно посредством четырех параметров – точек излома трапеции: a, b, c, d (фиг. 2). Рассмотрим $\{\boldsymbol{\alpha}, \boldsymbol{\beta}\}$-покрытие для заданного разбиения $\{a_i\}_{i=1}^n$. Требование (1.4) в определении покрытия означает, что для нечеткого подмножества M_i вершина трапеции располагается внутри интервала $[a_i, a_{i+1}]$. Отсюда, учитывая требования (1.1) и (1.2) в определении покрытия, можно получить, что $a_i \le b \le c \le a_{i+1}$, где b, c – соответствующие параметры характеристической функции для M_i.

Для простоты рассмотрим отдельно интервал разбиения (обозначим его через $[a_1, a_2]$) и соответствующее ему нечеткое подмножество – трапецию. Левая сторона трапеции представляет собой прямую линию, проходящую через три точки соответственно: $(0, a)$, (α_1, a_1), $(1, b)$ (фиг. 2). Это условие дает следующую систему уравнений:

$$0 = ka + b_1,$$

$$\alpha_1 = ka_1 + b_1,$$

$$1 = kb + b_1.$$

Здесь коэффициенты k, b_1 – угол наклона и параметр сдвига прямой.

Аналогично, правая сторона трапеции есть прямая линия, проходящая через точки $(0, d)$, (α_2, a_2), $(1, c)$ (фиг. 2). Ввиду равнобедренности трапеции угол наклона прямой $k_1 = -k$. Таким образом, получаем следующую систему:

$$0 = -kd + b_2,$$

$$\alpha_2 = -ka_2 + b_2,$$
$$1 = -kc + b_2.$$

Здесь b_2 – параметр сдвига прямой.

Условие (1.5) в определении покрытия приводит к равенству

$$c - b = \beta(a_2 - a_1).$$

Объединяя все полученные соотношения в единую систему, получаем

$$\begin{aligned} 0 &= ka + b_1, \\ \alpha_1 &= ka_1 + b_1, \\ 1 &= kb + b_1, \\ 0 &= -kd + b_2, \\ \alpha_2 &= -ka_2 + b_2, \\ 1 &= -kc + b_2, \\ c - b &= \beta(a_2 - a_1). \end{aligned}$$

Данная система представляет собой систему из семи линейных уравнений относительно семи неизвестных с невырожденным определителем и приводит к однозначному решению для a, b, c, d:

$$\begin{aligned} a &= a_1 - \frac{\alpha_1(1-\beta)(a_2-a_1)}{2-\alpha_1-\alpha_2}, \\ b &= a_1 + \frac{(1-\alpha_1)(1-\beta)(a_2-a_1)}{2-\alpha_1-\alpha_2}, \\ c &= a_2 - \frac{(1-\alpha_2)(1-\beta)(a_2-a_1)}{2-\alpha_1-\alpha_2}, \\ d &= a_2 + \frac{\alpha_2(1-\beta)(a_2-a_1)}{2-\alpha_1-\alpha_2}. \end{aligned}$$

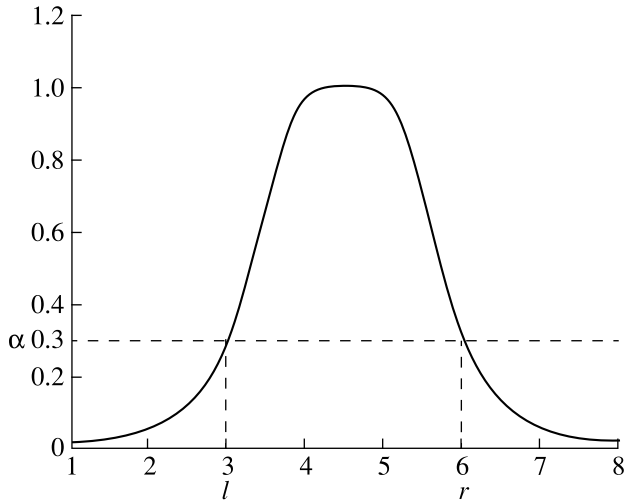
Теорема доказана.

Таким образом, для определения экспертной интерпретации для заданного разбиения можно воспользоваться понятием $\{\boldsymbol{\alpha}, \boldsymbol{\beta}\}$-покрытия. Для этого необходимо определить $2n - 1$ параметров, где n – число элементов разбиения. Как было указано выше, на практике элементы разбиения представляют собой примерные границы между различными состояниями, задаваемыми экспертом, а введение нечетких подмножеств отражает степень уверенности эксперта относительно выбранных границ. Поэтому можно предположить, что коэффициенты $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ отражают особенности знания эксперта о значениях признака в целом, а не об отдельных нечетких подмножествах. Это означает, что можно положить $\alpha_1 = \ldots = \alpha_n = \alpha$, $\beta_1 = \ldots = \beta_{n-1} = \beta$. Таким образом, пространство оптимизируемых параметров существенно сокращается.

Характеристическая функция нечетких подмножеств $\{\boldsymbol{\alpha}, \boldsymbol{\beta}\}$-покрытия имеет следующий вид:

$$\mu_I(x; l, r, \alpha, \beta) = \begin{cases} 0, \quad x < l - \dfrac{\alpha(r-l)(1-\beta)}{2(1-\alpha)}, \\ 2\dfrac{(x-l)(1-\alpha)}{(r-l)(1-\beta)} + \alpha, \quad l - \dfrac{\alpha(r-l)(1-\beta)}{2(1-\alpha)} \le x < l + \dfrac{(1-\beta)(r-l)}{2}, \\ 1, \quad l + \dfrac{(1-\beta)(r-l)}{2} \le x < r - \dfrac{(1-\beta)(r-l)}{2}, \\ 2\dfrac{(r-x)(1-\alpha)}{(r-l)(1-\beta)} + \alpha, \quad r - \dfrac{(1-\beta)(r-l)}{2} \le x < r + \dfrac{\alpha(r-l)(1-\beta)}{2(1-\alpha)}, \\ 0, \quad x \ge r + \dfrac{\alpha(r-l)(1-\beta)}{2(1-\alpha)}. \end{cases}$$

Здесь параметры l, r обозначают левую и правую условную границу соответственно.



Фиг. 4.

1.2. Семейство колоколообразных бесконечно дифференцируемых функций

Семейство трапециевидных функций может оказаться не очень удобным в тех случаях, когда необходимо проводить оптимизацию по условным границам разбиений. Дело в том, что трапеция не является дифференцируемой функцией в точках излома (которые соответствуют границам разбиений), что может осложнить применение градиентных методов оптимизации. Для того чтобы преодолеть эту проблему, введем семейство колоколообразных функций. Форму нечетких подмножеств данного семейства задают следующие характеристические функции (фиг. 4):

$$\mu_{\mathrm{II}}(x;\, l, r, \alpha, \beta) = \frac{1}{1 + \left(\dfrac{1}{\alpha} - 1\right)\left(\dfrac{x - (l + r)/2}{(r - l)/2}\right)^{2\beta}}, \quad \alpha \in (0, 1), \quad \beta \in \mathbb{N}, \quad l, r, x \in \mathbb{R}, \quad l < r. \tag{1.6}$$

Параметры l и r определяют *условные границы* нечеткого подмножества ($\mu_{\mathrm{II}}(l;\, l, r, \alpha, \cdot) = \mu_{\mathrm{II}}(r;\, l, r, \alpha, \cdot) = \alpha$); β показывает “степень нечеткости” подмножества. Предлагаемое семейство является достаточно гибким. В частности, при стремлении β к бесконечности получаем характеристическую функцию классического множества – отрезка $[l, r]$. Основным достоинством предлагаемого семейства является аналитичность μ_{II} по всем своим аргументам (x, l и r).

По аналогии с $\{\boldsymbol{\alpha}, \boldsymbol{\beta}\}$-покрытием можно также ввести понятие покрытия для характеристических функций (1.6).

Определение 4. $[\boldsymbol{\alpha}, \boldsymbol{\beta}]$-*покрытием* признака для разбиения области его значений $a_1 < \ldots < a_n$ будем называть совокупность нечетких подмножеств $\{M_i\}_{i=1}^{n-1}$ с характеристическими колоколообразными функциями $\mu_{M_i}(x) = \mu_{\mathrm{II}}(x;\, a_i, a_{i+1}, \alpha_i, \beta_i)$. При этом $\alpha_1, \ldots, \alpha_{n-1} \in (0, 1)$, $\beta_1, \ldots, \beta_{n-1} \in \mathbb{N}$.

Для $[\boldsymbol{\alpha}, \boldsymbol{\beta}]$-покрытия значения параметров $\boldsymbol{\alpha}$ определяют значения характеристических функций принадлежности в точках разбиения. В том случае, когда отсутствуют веские причины задавать значения данного параметра, то разумным выбором является значение 0.5. По аналогии с $\{\boldsymbol{\alpha}, \boldsymbol{\beta}\}$-покрытием значения параметров β в $[\boldsymbol{\alpha}, \boldsymbol{\beta}]$-покрытии можно трактовать как степень уверенности эксперта относительно выбранных границ разбиения. Потому можно положить одно и то же значение β для всех нечетких подмножеств покрытия, т.е. $\beta_1, \ldots, \beta_{n-1} = \beta$.

Характеристические функции нечетких подмножеств в семействах трапециевидных и колоколообразных функций определяются относительно условных границ l и r. В дальнейшем обозначение $\mu(x;\, l, r)$ может восприниматься в смысле I или II без ограничения корректности построений.

2. МОДЕЛЬ НЕЧЕТКИХ РАССУЖДЕНИЙ

Предположим, что для каждого признака $i \in \{1, 2, \ldots, d\}$ задана экспертная интерпретация $\mathfrak{T}_i$ (база данных системы). Она может определяться с помощью $\{\boldsymbol{\alpha}, \boldsymbol{\beta}\}$- или $[\boldsymbol{\alpha}, \boldsymbol{\beta}]$-покрытия. Пусть также известен набор правил, связывающих нечеткие подмножества входных и выходных переменных (база знаний системы). Каждое правило может быть представлено следующим образом:

$$\text{“ЕСЛИ } x_{i_1} \in M_{i_1}^{j_1} \ \& \ldots \& \ x_{i_r} \in M_{i_r}^{j_r}, \text{ ТО } y \in N^k\text{”}. \tag{2.1}$$

Здесь $M_{i_1}^{j_1}, \ldots, M_{i_r}^{j_r}, N^k$ – нечеткие подмножества соответствующих переменных. Множество $\mathrm{Sump}(R) = \{M_{i_1}^{j_1}, \ldots, M_{i_r}^{j_r}\}$ будем называть *посылкой* правила R (2.1), метку класса в задаче классификации (индекс нечеткого подмножества вещественной оси в задаче восстановления регрессии) $\mathrm{Res}(R) = N^k$ – *заключением* правила R; число $\mathrm{Ord}(R) = r$ – *порядком* правила R. Такие правила могут быть получены от эксперта или сгенерированы по прецедентной информации (см. разд. 3).

Рассмотрим задачу проведения нечетких рассуждений (классификации новых объектов) с использованием базы данных и базы знаний. К сожалению, на настоящий момент не существует единых подходов по отношению к тому, какие вычисления необходимо проделать в программе, где встроены функции принадлежности, чтобы смоделировать нечеткий вывод. Существует более 100 методов преобразования нечетких выводов на лингвистическом уровне в вычислении (см. [6]). Какому именно методу отдать предпочтение, определяется прежде всего практическими результатами использования экспертных систем. В связи с этим представляется разумным принять стратегию выбора методов, лучше всего зарекомендовавших себя на практике.

Для определения нечеткого вывода необходимо задать способ вычисления функций принадлежности антецедента правила (нечеткой конъюнкции) и вывода (нечеткой импликации), а также определить композиционное правило вывода.

Нечеткая конъюнкция традиционно задается t-нормой [см. [7]). В настоящей статье будет использоваться операция min, которая является типичной t-нормой.

Знание эксперта (правило) $\mathrm{Sump}(R) \longrightarrow \mathrm{Res}(R)$ отражает некоторое нечеткое причинное отношение предпосылки и заключения, т.е.

$$R = \mathrm{Sump}(R) \longrightarrow \mathrm{Res}(R).$$

Посылку R можно рассматривать как нечеткое подмножество на прямом произведении $X \times Y$ полного пространства предпосылок X и полного пространства заключений Y. Таким образом, процесс получения результата вывода B' с использованием данных наблюдения A' и знания $\mathrm{Sump}(R) \longrightarrow \mathrm{Res}(R)$ можно представить в виде формулы

$$B' = A' \bullet R = A' \bullet (\mathrm{Sump}(R) \longrightarrow \mathrm{Res}(R)).$$

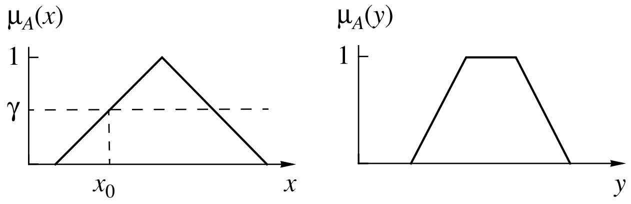
Здесь знак $\bullet$ является композиционным правилом нечеткого вывода, а знак $\longrightarrow$ в правиле $\mathrm{Sump}(R) \longrightarrow \mathrm{Res}(R)$ – нечеткой импликацией.

На практике нередки случаи, когда из-за особенностей системы информацию с достаточно хорошей точностью получить не удается либо нет возможности установить устройство измерения показателя, что вынуждает оценивать его по косвенным характеристикам. В подобных случаях удобно принимать за информационное наблюдение представление с помощью нечеткого подмножества. Поэтому в более общем случае наблюдение A' представляет собой нечеткое подмножество, а закон композиции $\bullet$ становится отдельным предметом исследования. Однако для систем распознавания можно положить, что значения входных переменных известны в виде конкретных значений, вследствие чего изучение этого вопроса опускается.

2.1. Выбор способа введения нечеткой импликации

Так же как и для операций НЕ, И, ИЛИ, расширение понятия импликации до нечеткого случая приводит к семействам обобщений. Однако в данном случае не удается построить аксиоматическую базу для расширения (по аналогии с t-нормой и t-конормой), вследствие чего обобщение данного понятия носит скорее эвристический характер.

Рассмотрим два нечетких подмножества A и B с их функциями принадлежности $\mu_A(x)$ и $\mu_B(y)$ (см. фиг. 5). С целью обоснованного выбора определения нечеткой импликации $\mu_{A \to B}(x, y)$ было



Фиг. 5.

проведено обширное исследование (см. [10]). В качестве критерия выбора была взята выполнимость некоторых схем нечетких рассуждений. Среди них можно выделить такие:

$$\frac{\text{если } X \text{ есть } A, \text{ то } Y \text{ есть } B;\ X \text{ есть } A}{Y \text{ есть } B};$$

$$\frac{\text{если } X \text{ есть } A, \text{ то } Y \text{ есть } B;\ X \text{ есть очень } A}{Y \text{ есть } B};$$

$$\frac{\text{если } X \text{ есть } A, \text{ то } Y \text{ есть } B;\ X \text{ есть более или менее } A}{Y \text{ есть более или менее } B};$$

$$\frac{\text{если } X \text{ есть } A, \text{ то } Y \text{ есть } B;\ X \text{ есть не-}A}{Y \text{ неизвестно}};$$

$$\frac{\text{если } X \text{ есть } A, \text{ то } Y \text{ есть } B;\ Y \text{ есть не-}B}{X \text{ есть не-}A};$$

$$\frac{\text{если } X \text{ есть } A, \text{ то } Y \text{ есть } B;\ Y \text{ есть } B}{X \text{ неизвестно}}.$$

Для исследования были взяты все известные в литературе по нечетким множествам определения нечеткой импликации с добавлением нескольких своих вариантов. В результате проверок был сделан вывод, что можно использовать в основном следующие правила (R_c, R_s, R_g – обозначения, введенные авторами, см. [10]):

(а) правило R_c (правило Мамдани):

$$\mu_{A \to B}(x, y) = \min(\mu_A(x), \mu_B(y));$$

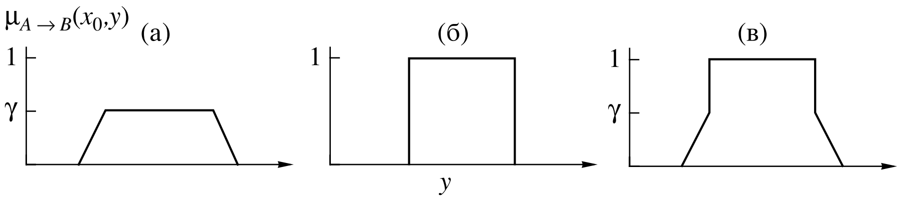
(б) правило R_s:

$$\mu_{A \to B}(x, y) = \begin{cases} 1, & \mu_A(x) \le \mu_B(y), \\ 0, & \mu_A(x) > \mu_B(y); \end{cases}$$

(в) правило R_g:

$$\mu_{A \to B}(x, y) = \begin{cases} 1, & \mu_A(x) \le \mu_B(y), \\ \mu_B(y), & \mu_A(x) > \mu_B(y). \end{cases}$$

Смысл каждого из этих выводов для примера, представленного фиг. 5, дается на фиг. 6. Как нетрудно видеть, правила вывода R_s и R_g не учитывают степень принадлежности $\gamma = \mu_A(x_0)$ в явном виде, а используют эти значения лишь для “усечения нечеткости”. Таким образом, в этом случае на выходе мы получаем практически четкие подмножества с уровнем принадлежности, равным 1, причем чем меньше γ, тем больше носитель подмножества вывода. Для правила вывода R_c значение γ играет большую роль. В связи с этим для задач распознавания представляется разумным использовать именно правило R_c для построения модели нечеткого вывода по правилам.



Фиг. 6.

2.2. Модель вычисления прогноза

Пусть имеется m лингвистических правил $\{R_i\}_{i=1}^{m}$ вывода вида (2.1). Согласно выводам предыдущего раздела, в качестве функции принадлежности нечеткой импликации выберем функцию, предложенную Мамдани (см. [8]):

$$\mu_{\mathrm{Sump}(R)\to\mathrm{Res}(R)}(x, y) = \min(\mu_{\mathrm{Sump}(R)}(x), \mu_{\mathrm{Res}(R)}(y)).$$

Тогда с учетом того, что в качестве нечеткой конъюнкции используется функция min, для правила вывода типа (2.1) соответствующая функция принадлежности будет иметь вид

$$\mu_{R_i}(\mathbf{x}, y) = \min\left(\mu_{M_{i_1}^{j_1}}(x_{i_1}), \ldots, \mu_{M_{i_r}^{j_r}}(x_{i_r}), \mu_{N^k}(y)\right).$$

Для группы правил вывода с одним и тем же результирующим нечетким подмножеством вычисление функции принадлежности происходит по формуле

$$\mu_{R_{N^k}}(\mathbf{x}, y) = \frac{\sum\limits_{\{i\,:\,\mathrm{Res}(R_i) = N^k\}} \mu_{R_i}(\mathbf{x}, y)}{\sum\limits_{\{i\,:\,\mathrm{Res}(R_i) = N^k\}} 1}, \qquad (2.2)$$

которая несколько отличается от классического взятия максимума (см. [6]). Однако авторы считают (2.2) более гибким средством, в большей степени учитывающим индивидуальность каждого правила вывода. Предположим далее, что помимо самих правил вывода известна также информация о весах этих правил. Эти веса могут быть заданы экспертом или получены в результате оптимизации по обучающей выборке (см. разд. 4). В качестве весов также могут выступать значения эффективностей правил (см. разд. 3). Обозначим вес i-го правила через w_i. Тогда (2.2) преобразуется к виду

$$\mu_{R_{N^k}}(\mathbf{x}, y) = \frac{\sum\limits_{\{i\,:\,\mathrm{Res}(R_i) = N^k\}} w_i\mu_{R_i}(\mathbf{x}, y)}{\sum\limits_{\{i\,:\,\mathrm{Res}(R_i) = N^k\}} w_i}.$$

Пусть в пространстве прогнозной переменной определено K нечетких подмножеств $\{N^k\}_{k=1}^{K}$. Обозначим через med(N^k) медиану k-го нечеткого подмножества, т.е.

$$\mathrm{med}(N^k) = \frac{a_k + a_{k+1}}{2}.$$

Здесь a_k, a_{k+1} – соответствующие элементы разбиения в экспертной интерпретации пространства результата.

На последнем этапе остается определить по нечеткому подмножеству вывода точечное значение прогноза. Для этого в основном используются следующие методы (см. [6, 7, 8, 10]):

$$z = \arg\max_{k=\overline{1,K}} \mu_{R_{N^k}}(\mathbf{x}, \mathrm{med}(N^k)), \qquad (2.3)$$

$$z = \frac{\sum_{k=1}^{K} \operatorname{med}(N^k)\mu_{R_{N^k}}(\mathbf{x}, \operatorname{med}(N^k))}{\sum_{k=1}^{K} \mu_{R_{N^k}}(\mathbf{x}, \operatorname{med}(N^k))}. \tag{2.4}$$

Дефаззификация по формуле (2.3) получила название дефаззификации по моде, в то время как формула (2.4) приводит к методу центра тяжести (ЦТ). Выражение (2.3) используется при построении экспертной системы для решения задачи классификации, т.е. когда необходимо разделить генеральную совокупность на непересекающиеся классы. Формула (2.4) используется в том случае, когда необходимо построить непрерывный прогноз.

Таким образом, описана следующая модель вычисления прогноза:

$$\mathfrak{M}[\{a_i^j\}_{i=0,\, j=1}^{d,\, n_i}, \{\alpha_i\}_{i=1}^{d}, \{\beta_i\}_{i-1}^{d}, \{R_i\}_{i=1}^{m}, \{w_i\}_{i=1}^{m}]. \tag{2.5}$$

Здесь $\{a_i^j\}_{i=0,\, j=1}^{d,\, n_i}$ – разбиения множеств значений каждого из d признаков и пространства результата, $\{\alpha_i\}_{i=1}^{d}$, $\{\beta_i\}_{i=1}^{d}$ – параметры $\{\boldsymbol{\alpha}, \boldsymbol{\beta}\}$- или $[\boldsymbol{\alpha}, \boldsymbol{\beta}]$-покрытия экспертной интерпретации признаков, $\{R_i\}_{i=1}^{m}$ – набор правил вывода, $\{w_i\}_{i=1}^{m}$ – соответствующие веса правил.

Значения параметров модели $\mathfrak{M}$ можно находить путем оптимизации по множеству прецедентов $\{\mathbf{x}^i, y^i\}_{k=1}^{q}$ с помощью выбранного функционала качества. Конкретные функционалы качества и методы их оптимизации будут рассмотрены в разд. 4.

3. АВТОМАТИЧЕСКОЕ ПРИОБРЕТЕНИЕ ЗНАНИЙ

Рассмотрим продукционное правило вывода (2.1):

$$\text{“ЕСЛИ } x_{i_1} \in M_{i_1}^{j_1} \;\&\; \dots \;\&\; x_{i_r} \in M_{i_r}^{j_r}, \text{ ТО } y \in N^k\text{”}.$$

Множество всех возможных правил данного вида $\mathfrak{R}^* = \{R\}$ будем называть *универсальной базой знаний*. Посылку правила $\operatorname{Sump}(R) = \{M_{i_1}^{j_1}, \dots, M_{i_r}^{j_r}\}$ можно рассматривать как нечеткое подмножество полного пространства $\mathbb{R}^r$. Характеристическая функция этого нечеткого подмножества, в дальнейшем называемая *степенью принадлежности объекта посылке*, в соответствии с выбранной характеристической функцией нечеткой конъюнкции имеет вид

$$\nu_R(x) = \min\Big(\mu_{M_{i_1}^{j_1}}(x_{i_1}), \dots, \mu_{M_{i_r}^{j_r}}(x_{i_r})\Big).$$

Характеристическая функция заключения правила имеет вид

$$\mu_{N^k}(y) = \begin{cases} 1 & \text{при } y = c_k, \\ 0 & \text{иначе}, \end{cases} \qquad c_k \in \{1, 2, \dots, l\},$$

в случае задачи классификации и

$$\mu_{N^k}(y) = \mu(y; l_k, r_k), \quad l_k < r_k,$$

в случае задачи восстановления регрессии.

Проблема приобретения знаний заключается в том, чтобы выделить подмножество $\mathfrak{R}$ универсальной базы знаний. При этом правила, составляющие $\mathfrak{R}$ должны удовлетворять некоторому критерию значимости, основанному на прецедентной информации. В следующем подразделе формулируется ряд критерием значимости, заданные предикатами C на единичном квадрате числовой плоскости. Важными характеристиками, при помощи которых можно сформулировать многие известные критерии значимости, являются репрезентативность и эффективность.

Определение 5. *Репрезентативностью* правила R называется величина

$$\operatorname{rep}(R) = \frac{\sum_{k=1}^{q} \nu_R(x^k)}{q}. \tag{3.1}$$

Определение 6. *Эффективностью* правила R называется величина

$$\operatorname{eff}(R) = \frac{\sum_{k=1}^{q} \min(\nu_R(x^k), \mu_{\operatorname{Res}(R)}(y^k))}{\operatorname{rep}(R)q}.$$

Если все множества в посылке правила четкие, то тогда репрезентативность правила является долей объектов обучающей выборки, удовлетворяющих посылке правила, а эффективность – доля объектов, удовлетворяющих правилу целиком, среди объектов, удовлетворяющих посылке правила. Задача автоматического приобретения знаний по прецедентной информации заключается в нахождении всех правил $R \in \mathfrak{R}^*$ таких, что некоторый предикат значимости правила равен единице, т.е. $C(\operatorname{rep}(R), \operatorname{eff}(R)) = 1$.

3.1. Критерии значимости правил

Задачей процесса генерации правил вывода является поиск правил, которые по возможности обладают высокими значениями репрезентативности и эффективности. Более формально: правило $R \in \tilde{\mathfrak{R}} \subset \mathfrak{R}$ является значимым, если $C(\operatorname{rep}(R), \operatorname{eff}(R)) = 1$, где $C : [0, 1]^2 \longrightarrow \{0, 1\}$.

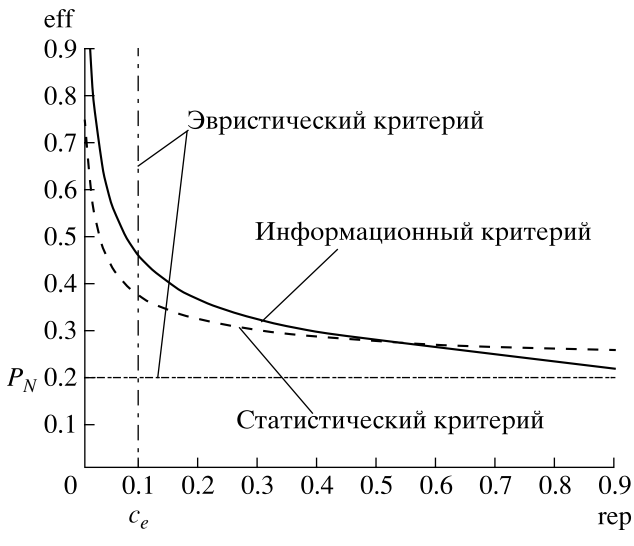
Эвристический критерий значимости. В наиболее простой форме предикат значимости правила задается фиксированными порогами для значений эффективности и репрезентативности.

Определение 7. Пусть даны числа $0 < c_r \le 1$ и $0 < c_e \le 1$. *Эвристическим критерием* будем называть следующий предикат на единичном квадрате числовой плоскости (фиг. 7):

$$C^h(v, w) = \begin{cases} 1, \text{ если } v \ge c_r \text{ и } w \ge c_e, \\ 0 \text{ иначе}, \end{cases} \quad (v, w) \in [0, 1]^2.$$

Правило R будем называть *значимым по эвристическому критерию*, если $C^h(\operatorname{rep}(R), \operatorname{eff}(R)) = 1$.

Однако константные пороги для значений эффективности и репрезентативности не совсем точно отражают суть значимости правила. Понятно, что для низких значений репрезентативно-



Фиг. 7.

сти значимое правило должно обладать высоким значением эффективности. С другой стороны, для высоких значений репрезентативности значение эффективности немного выше априорной вероятности появления класса уже будет свидетельствовать о значимости правила. Следующий критерий значимости принимает во внимание указанные замечания.

Статистический критерий значимости основывается на технике проверки статистических гипотез. Предположим, что все множества в посылке правила четкие. Тогда правило R является незначимым, если информация о том, что объект удовлетворяет посылке правила $\mathrm{Sump}(R)$, не дает информации о его принадлежности заключению этого правила $\mathrm{Res}(R) = N$:

$$\mathbb{P}(y \in N \mid \mathbf{x} \in \mathrm{Sump}(R)) = \mathbb{P}(y \in N).$$

Обозначим оценку априорной вероятности заключения N по обучающей выборке через P_N:

$$P_N = \frac{\sum_{k=1}^{q} \mu_N(y^k)}{q}. \tag{3.2}$$

Если гипотеза верна, тогда случайная величина $\mathrm{rep}(R)\mathrm{eff}(R)q$ представляет собой $n = \mathrm{rep}(R)q$ испытаний Бернулли с вероятностью успеха P_N и может быть приближена нормальным распределением с центром $\mathrm{rep}(R)qP_N$ и дисперсией $\mathrm{rep}(R)qP_N(1 - P_N)$, т.е.

$$\frac{\mathrm{eff}(R) - P_N}{\sqrt{\frac{P_N(1 - P_N)}{\mathrm{rep}(R)q}}} \sim \mathcal{N}(0, 1).$$

Выбирая z, соответствующее достаточно высокому уровню значимости, имеем доверительный интервал, в котором принимается гипотеза

$$\frac{\mathrm{eff}(R) - P_N}{\sqrt{\frac{P_N(1 - P_N)}{\mathrm{rep}(R)q}}} \in (-z, z).$$

В области $(-\infty, -z]$ расположены антиправила, из принадлежности объекта посылке которых следует сделать вывод о непринадлежности объекта их заключению. Таким образом, правило признается значимым, если

$$\frac{\mathrm{eff}(R) - P_N}{\sqrt{\frac{P_N(1 - P_N)}{\mathrm{rep}(R)q}}} > z.$$

Определение 8. Пусть дан уровень значимости $\alpha > 0$. *Статистическим критерием* будем называть следующий предикат на единичном квадрате числовой плоскости (фиг. 7):

$$C^s(v, w) = \begin{cases} 1, \text{ если } \dfrac{w - P_N}{\sqrt{\dfrac{P_N(1 - P_N)}{vq}}} \ge z_\alpha \\ 0 \text{ иначе,} \end{cases} \quad (v, w) \in [0, 1]^2.$$

Здесь z_α – квантиль стандартного нормального распределения, отвечающая уровню значимости α. Правило R будем называть *значимым по статистическому критерию*, если $C^s(\mathrm{rep}(R), \mathrm{eff}(R)) = 1$.

Статистический критерий сформулирован для случая, когда все множества в посылке правила четкие. В этом случае рассмотрение испытаний Бернулли и аппроксимация нормальным распределением являются вполне законными. Данный критерий можно использовать и в нечетком случае, однако тогда его следует считать эвристикой. Тем не менее в практических приложениях его использование часто приводит к хорошим результатам.

Информационный критерий значимости основывается на понятиях энтропии и информационного выигрыша. Напомним, что энтропией дискретного распределения называется математическое ожидание количества информации; при этом если вероятность исхода равна $P(w)$, то количество информации, связанное с ним, равно $-\log(P(w))$. Таким образом, энтропию обучающей выборки до получения информации о существовании правила R с заключением N, учитывая соотношение (3.2), определяем по формуле

$$H_1 = -P_N\log(P_N) - (1-P_N)\log(1-P_N).$$

После получения информации о существовании правила R энтропия принимает более сложный вид:

$$H_2(\mathrm{rep}(R), \mathrm{eff}(R)) = -\mathrm{rep}(R)\{\mathrm{eff}(R)\log[\mathrm{eff}(R)] + [1-\mathrm{eff}(R)]\log[1-\mathrm{eff}(R)]\} -$$

$$-(1-\mathrm{rep}(R))\left[\frac{P_N - \mathrm{rep}(R)\mathrm{eff}(R)}{1-\mathrm{rep}(R)}\log\left(\frac{P_N - \mathrm{rep}(R)\mathrm{eff}(R)}{1-\mathrm{rep}(R)}\right) +\right.$$

$$\left.+\frac{1-\mathrm{rep}(R)[1-\mathrm{eff}(R)] - P_N}{1-\mathrm{rep}(R)}\log\left(\frac{1-\mathrm{rep}(R)[1-\mathrm{eff}(R)] - P_N}{1-\mathrm{rep}(R)}\right)\right].$$

Так же как и в рассуждениях для статистического критерия, при $\mathrm{eff}(R) < P_N$, значимым объявляется антиправило. Для устранения этой проблемы достаточно домножить информационный выигрыш на $\mathrm{sign}(\mathrm{eff}(R) - P_N)$.

Определение 9. Пусть дано число $c_i > 0$. *Информационным критерием* будем называть следующий предикат на единичном квадрате числовой плоскости (фиг. 7):

$$C^i(v, w) = \begin{cases} 1, & \text{если} \ \ \mathrm{sign}(w - P_N)(H_1 - H_2(v, w)) \geq c_i, \\ 0 & \text{иначе}, \end{cases} \qquad (v, w) \in [0, 1]^2.$$

Правило R будем называть *значимым по информационному критерию*, если $C^i(\mathrm{rep}(R), \mathrm{eff}(R)) = 1$.

На фиг. 7 представлены различные критерии значимости. Правила, лежащие выше обозначенных кривых, признаются значимыми по соответствующему критерию. Возможные значения эффективности и репрезентативности удовлетворяют соотношению $vw \leq P_N$.

Критерий значимости Колмогорова–Смирнова. Рассмотрим распределение значений характеристической функции правила R:

$$G(t) = \mathbb{P}\{\mu_R(x) < t\}.$$

Можно построить эмпирические функции распределения $G'(t)$ и $G''(t)$, используя объекты $y_i \in$ $\in \mathrm{Res}(R)$ и $y_i \notin \mathrm{Res}(R)$ соответственно. Тогда для определения значимых различий между распределениями может быть использован непараметрический статистический тест Колмогорова–Смирнова.

Заметим, что, в отличие от предыдущих критериев значимости, данный критерий сформулирован и обоснован для применения в нечетком случае.

3.2. *Общие требования к предикатам значимости правил*

За исключением эвристического критерия значимости, предикаты для остальных критериев задаются в форме

$$C(v, w) = \begin{cases} 1, & \text{если} \ \ F(v, w) \geq F_0, \\ 0 & \text{иначе}, \end{cases}$$

где $F : [0, 1]^2 \longrightarrow \mathbb{R}$ – функция критерия и F_0 – некоторый порог. Чем больше значение функции критерия F, тем более значимым является правило. Можно сформулировать следующее требование к функциям критерия для “разумных” критериев значимости правил.

Определение 10. Функция критерия F *удовлетворяет условию монотонности*, если $\forall(v_0, w_0) : F(v_0, w_0) \geq F_0$ выполняется условие

$$F(v, w) \geq F(v_0, w_0) \ \ \forall(v, w) > (v_0, w_0). \qquad (3.3)$$

Для выполнения условия (3.3) необходимо и достаточно, чтобы $\forall(\nu_0, w_0) : F(\nu_0, w_0) \ge F_0$ выполнялись условия

$$F(\nu, w_0) \ge F(\nu_0, w_0) \quad \forall \nu > \nu_0, \tag{3.4}$$

$$F(\nu_0, w) \ge F(\nu_0, w_0) \quad \forall w > w_0, \tag{3.5}$$

где (3.4) означает, что с ростом репрезентативности с сохранением значения эффективности правило не становится менее значимым, а (3.5) – что для фиксированного уровня репрезентативности увеличение значения эффективности только увеличивает значимость правила.

Теорема 2. *Функции критерия для критерия значимости Колмогорова–Смирнова в четком случае и для статистического, информационного критериев значимости в общем случае удовлетворяют условию* (3.3).

Доказательство. Для доказательства (3.3) необходимо и достаточно проверить выполнение условий (3.4) и (3.5).

1. Для статистического критерия функция критерия определяется выражением

$$\frac{w - P_N}{\sqrt{\frac{P_N(1 - P_N)}{\nu q}}}.$$

При фиксированном значении репрезентативности ν знаменатель является постоянной величиной и функция критерия возрастает с ростом эффективности w. С другой стороны, при фиксированном значении w числитель является константой, а знаменатель убывает с ростом ν. Это означает, что функция критерия возрастает с ростом ν.

2. Для информационного критерия функция критерия определяется выражением

$$\nu H(w) + (1 - \nu) H\left(\frac{P_N - \nu w}{1 - \nu}\right), \tag{3.6}$$

где $H(x) = x\log(x) + (1 - x)\log(1 - x)$. Вычисляя производную выражения (3.6) по w, получаем

$$\nu \log\left(\frac{w}{1 - w}\right) - \nu \log\left(\frac{P_N - \nu w}{1 - \nu - P_N + \nu w}\right) = \nu \log\left(\frac{w - \nu w - w P_N + \nu w^2}{P_N - \nu w - w P_N + \nu w^2}\right).$$

Исключая из рассмотрения антиправила, т.е. принимая $w \ge P_N$, получаем, что производная всюду неотрицательна. Следовательно, функция монотонно неубывает по w. Аналогичные рассуждения для ν приводят к тому, что функция (3.6) монотонно неубывает по ν.

3. В четком случае функция критерия для критерия Колмогорова–Смирнова имеет вид

$$p/P - n/N, \tag{3.7}$$

где P и N – количество своих и чужих объектов в выборке, а p и n – количество своих и чужих объектов, выделяемых правилом. При фиксированном значении репрезентативности $\mathrm{rep}(R) = (p + n)/(P + N)$ увеличение эффективности $\mathrm{eff}(R) = p/(p + n)$ эквивалентно увеличению p. При этом значение n уменьшается так, что $p + n = \mathrm{const}$. Тогда функция критерия (3.7) возрастает.

При фиксированном значении эффективности имеем $p = Cn$, где $C \in (0, 1)$ – некоторая константа. В этом случае увеличение значения репрезентативности эквивалентно увеличению p. При этом функция критерия (3.7) также возрастает.

3.3. Характеристика предиката, определяющего критерий значимости

Для дальнейших построений необходимо ввести некоторые дополнительные обозначения.

Определение 11. Число $\alpha \ge 0$ будем называть *характеристикой* предиката $C : [0, 1]^2 \longrightarrow \{0, 1\}$, если выполнено следующее:

$\forall(\nu, w) \in [0, 1]^2 : C(\nu, w) = 1 \Rightarrow \nu w \ge \alpha$;

$\forall \varepsilon > 0\ \exists(\nu, w) \in [0, 1]^2 : C(\nu, w) = 1$ и $\nu w < \alpha + \varepsilon$.

Определение 12. Предикат $C^\alpha : [0, 1]^2 \longrightarrow \{0, 1\}$ с характеристикой α будем называть *максимальным*, если для любого другого предиката $C' : [0, 1]^2 \longrightarrow \{0, 1\}$ с характеристикой α выполнено условие $\forall(\nu, w) \in [0, 1] : C'(\nu, w) = 1 \Rightarrow C^\alpha(\nu, w) = 1$.

Теорема 3. *Для любого* $\alpha \in [0, 1]$ *существует единственный максимальный предикат с характеристикой* α:

$$C^{\alpha}(v, w) = \begin{cases} 1, & \text{если } \; vw \geq \alpha, \\ 0 & \text{иначе,} \end{cases} \quad (v, w) \in [0, 1]^2. \tag{3.8}$$

Доказательство. *Существование* следует из того, что C^{α}, определяемый формулой (3.8), удовлетворяет определению максимального предиката.

Единственность. Пусть C' – произвольный максимальный предикат с характеристикой α. Тогда, согласно определению максимального предиката,

$$\forall (v, w) \in [0, 1]^2 : \begin{cases} C'(v, w) \Rightarrow C^{\alpha}(v, w), \\ C^{\alpha}(v, w) \Rightarrow C'(v, w), \end{cases}$$

т.е. $C'(v, w) \Leftrightarrow C^{\alpha}(v, w)$ $\forall (v, w) \in [0, 1]^2$. Утверждение доказано.

3.4. Метод эффективных сужений

Определение 13. Правило R_2 будем называть *сужением* правила R_1 и обозначать $R_2 \subset R_1$, если выполнены следующие условия:

1) $\mathrm{Res}(R_1) = \mathrm{Res}(R_2)$;

2) $\mathrm{Sump}(R_1) \subset \mathrm{Sump}(R_2)$.

Теорема 4. *Если* $R' \supset R$, *то выполнены неравенства*

$$\mathrm{rep}(R') \geq \mathrm{rep}(R),$$

$$\mathrm{rep}(R')\mathrm{eff}(R') \geq \mathrm{rep}(R)\mathrm{eff}(R).$$

Доказательство. Допустим, что $\mathrm{Sump}(R') = \{M_{i_1}^{j_1}, \ldots, M_{i_r}^{j_r}\}$, а посылка правила R отличается от посылки R' на одно множество $\mathrm{Sump}(R) = \mathrm{Sump}(R') \cup \{M_{i_{r+1}}^{j_{r+1}}\}$. Тогда

$$\mathrm{rep}(R') = \{\text{опр. } (3.1)\} = \frac{1}{q}\sum_{k=1}^{q} \min\left(\mu_{M_{i_1}^{j_1}}(x_{i_1}^k), \ldots, \mu_{M_{i_r}^{j_r}}(x_{i_r}^k)\right) \geq$$

$$\geq \frac{1}{q}\sum_{k=1}^{q} \min\left(\mu_{M_{i_1}^{j_1}}(x_{i_1}^k), \ldots, \mu_{M_{i_r}^{j_r}}(x_{i_r}^k), \mu_{M_{i_{r+1}}^{j_{r+1}}}(x_{i_{r+1}}^k)\right) = \mathrm{rep}(R).$$

Соответственно,

$$\mathrm{rep}(R')\mathrm{eff}(R') = \frac{1}{q} \sum_{k = \overline{1, q};\; y^k = \mathrm{Res}(R')} \min\left(\mu_{M_{i_1}^{j_1}}(x_{i_1}^k), \ldots, \mu_{M_{i_r}^{j_r}}(x_{i_r}^k)\right) \geq$$

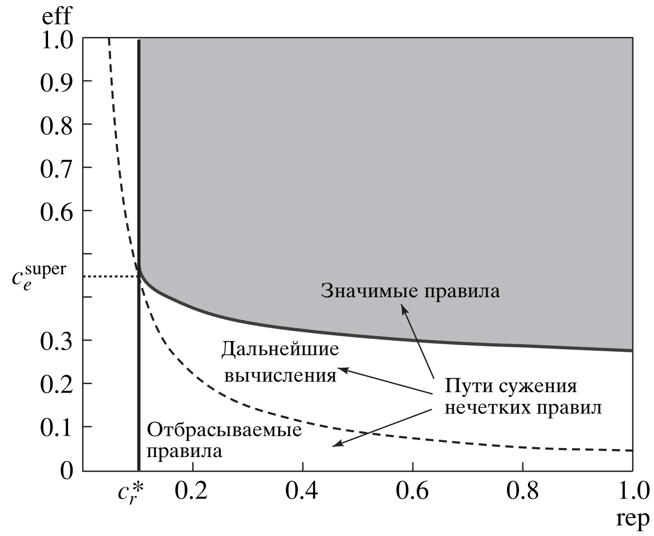
$$\geq \frac{1}{q} \sum_{k = \overline{1, q};\; y^k = \mathrm{Res}(R')} \min\left(\mu_{M_{i_1}^{j_1}}(x_{i_1}^k), \ldots, \mu_{M_{i_r}^{j_r}}(x_{i_r}^k), \mu_{M_{i_{r+1}}^{j_{r+1}}}(x_{i_{r+1}}^k)\right) = \{\mathrm{Res}(R) = \mathrm{Res}(R')\} = \mathrm{rep}(R)\mathrm{eff}(R).$$

Доказательство для случая произвольного порядка правила R проводится аналогично. Утверждение доказано.

Пусть заданы экспертные интерпретации всех признаков $\mathfrak{T} = \{\mathfrak{T}_1, \ldots, \mathfrak{T}_d\}$, заключение правил N и предикат C с характеристикой $\alpha \geq \alpha_0 > 0$. Пусть

$$c_r^* = \inf\{v \in [0, 1] \mid \exists w \in [0, 1] : C(v, w) = 1\}.$$

Следующий алгоритм позволяет находить все значимые правила минимального порядка.



Фиг. 8.

Шаг 1. Построим все возможные правила первого порядка для заданной экспертной интерпретации признаков $\mathfrak{T}$:

$$\mathfrak{R}' := \{R \in \mathfrak{R}^* \mid \mathrm{Res}(R) = N, \mathrm{Sump}(R) = M_i^{j_i}, j_i = \overline{1, n_i}, i = \overline{1, d}\}.$$

Шаг 2. Вычислим репрезентативность и эффективность каждого правила и исключим из рассмотрения нерепрезентативные:

$$\mathfrak{R}' := \{R \in \mathfrak{R}' \mid \mathrm{rep}(R) \geq c_r^*\}.$$

Шаг 3. Исключим из полученного множества правила, не удовлетворяющие максимальному критерию с характеристикой α_0:

$$\mathfrak{R}' := \{R \in \mathfrak{R}' \mid C^{\alpha_0}(\mathrm{rep}(R), \mathrm{eff}(R)) = 1\}.$$

Шаг 4. Если все правила исчерпаны, то алгоритм заканчивает работу. В противном случае, если правило удовлетворяет критерию C, то оно является значимым и заносится в результирующее множество $\tilde{\mathfrak{R}}$, иначе – участвует в процессе сужений:

$$\tilde{\mathfrak{R}} := \tilde{\mathfrak{R}} \cup \{R \in \mathfrak{R}' \mid C(\mathrm{rep}(R), \mathrm{eff}(R)) = 1\}$$

$$\mathfrak{R}' := \{R \in \mathfrak{R}' \mid C(\mathrm{rep}(R), \mathrm{eff}(R)) = 0\}.$$

Шаг 5. Остальные правила (если они есть) используются в процессе сужения следующим образом. Объединение посылок любых двух правил, которые сужаются к некоторому правилу большего порядка, и есть в точности посылка нового правила, если порядок последнего на единицу больше порядка исходных:

$$\mathfrak{R}' := \{R \in \mathfrak{R}^* \mid \mathrm{Res}(R) = N, \mathrm{Sump}(R) = \mathrm{Sump}(R_1) \cup \mathrm{Sump}(R_2), R_1, R_2 \in \mathfrak{R}',$$
$$\mathrm{Ord}(R) = \mathrm{Ord}(R_1) + 1\ \forall R_+ : R \subset R_+, \mathrm{Ord}(R_+) = \mathrm{Ord}(R) - 1 \Rightarrow R_+ \in \mathfrak{R}'\}.$$

Шаг 6. Если все правила обработаны, т.е. $\mathfrak{R}' = \emptyset$, то конец работы алгоритма, иначе переходим к шагу 2.

На фиг. 8 представлена схема работы метода эффективных сужений. Заштрихованная область – область координат значимых правил. Вертикальная черная линия пересекает ось абсцисс в точке c_r^*. Все точки, в которых значение максимального предиката для заданного критерия равно единице, расположены выше штриховой линии.

Теорема 5 (основная). *Метод эффективных сужений строит все значимые правила минимального порядка, т.е.*

$$\mathfrak{R} = \{R \in \mathfrak{R}^* \mid C(\text{rep}(R), \text{eff}(R)) = 1\ \forall R' \supset R : C(\text{rep}(R'), \text{eff}(R')) = 0\}.$$

Доказательство. Рассмотрим произвольное минимальное по порядку значимое правило R, пусть $\text{Ord}(R) = r$. Докажем, что оно войдет в результирующее множество $\mathfrak{R}$ на r-й итерации шага 4. Если это не так, то найдется правило меньшего (или такого же) порядка $\exists R' \supset R$, $\text{Ord}(R') = r' \le r$, для которого R является сужением (или $R' = R$), исключаемое из $\mathfrak{R}''$ на r'-й итерации шага 2 или 3. Рассмотрим оба случая.

1. R' исключено из $\mathfrak{R}''$ в результате шага 2, т.е. $\text{rep}(R') < c_r^*$. Тогда, на основании утверждения 4, имеем $\text{rep}(R) < c_r^*$. Согласно определению точной нижней грани числового множества,

$$\forall \varepsilon \in (0, c_r^*]\ \forall w \in [0, 1] : C(c_r^* - \varepsilon, w) = 0,$$

а значит, в частности, $C(\text{rep}(R), \text{eff}(R)) = 0$, что противоречит значимости правила R.

2. R' исключено из $\mathfrak{R}''$ в результате шага 3, т.е. $\text{rep}(R')\text{eff}(R') < \alpha_0 \le \alpha$. Аналогично предыдущему рассуждению, на основании теоремы 4 имеем $\text{rep}(R)\text{eff}(R) < \alpha$. Следовательно, ввиду того что C имеет характеристику α имеем $C(\text{rep}(R), \text{eff}(R)) = 0$, что также противоречит значимости правила R. Утверждение доказано.

Замечание 1. Доказанное утверждение обосновывает применимость метода эффективных сужений для произвольного предиката с положительной характеристикой. В случае если характеристика предиката равна нулю и $c_r^* = 0$, шаги 2 и 3 не имеют смысла и алгоритм вырождается в тривиальный перебор.

Теорема 6. *Предикаты C^h, C^s и C^i имеют положительные характеристики.*

Доказательство. 1. Докажем, что характеристика C^h равна $c_r c_e$:

$$\forall (v, w) \in [0, 1]^2 : C^h(v, w) = 1 \Rightarrow \text{rep} \ge c_r \text{ и } w \ge c_e \Rightarrow vw \ge c_r c_e = \alpha,$$

$$\forall \varepsilon > 0\ \exists v = c_r, \quad w = c_e : C^h(v, w) = 1 \text{ и } vw = c_r c_e = \alpha < \alpha + \varepsilon.$$

Значит, $\alpha = c_r c_e$ является характеристикой предиката C^h согласно определению (11).

2. Докажем, что характеристика C^s положительна. Предположим противное: $\alpha = 0$. Заметим, что $z \ge z_0 > 0$. Положим

$$\varepsilon = \varepsilon_0 < \min\left\{P_c, \left[z_0\sqrt{\frac{P_c(1-P_c)}{q}}/(1-P_c)\right]^2 = r_0\right\}^2;$$

тогда, по определению (11),

$$\exists (v, w) \in [0, 1]^2 : vw < \varepsilon \text{ и } \frac{w - P_N}{\sqrt{\frac{P_N(1-P_N)}{vq}}} \ge z \Rightarrow (w < P_N \vee v < r_0) \text{ и } \frac{w - P_N}{\sqrt{\frac{P_N(1-P_N)}{vq}}} \ge z \Rightarrow z < z_0.$$

Таким образом, получаем противоречие. Следовательно, характеристика C^s является положительной.

3. Докажем, что характеристика C^i положительна. Предположим противное: $\alpha = 0$. Заметим, что $c_i \ge c_0 > 0$ и $\lim\limits_{v \to +0}(H_1 - H_2(v, w)) = 0$, т.е. $\exists \delta : 0 < v < \delta \Rightarrow H_1 - H_2 < c_0$. Положим $\varepsilon = \delta P_N$; тогда, по определению (11), $\exists (v, w) \in [0, 1]^2 : vw < \varepsilon$ и $w > P_N$ и $H_1 - H_2(v, w) > c_i \Rightarrow c_i < c_0$. Противоречие. Утверждение доказано.

Замечание 2. Чем выше нижняя оценка характеристики предиката, тем больше незначимых в перспективе правил исключаются из рассмотрения на ранних этапах метода эффективных сужений. Получение точных значений характеристик является важной теоретической задачей, имеющей прямое отношение к практике. В настоящей работе сформулированы и доказаны теоремы существования положительных характеристик для известных критериев: эвристического, статистического и информационного.

3.5. Дизъюнктивная модификация метода эффективных сужений

В дальнейшем правила, посылка которых является конъюнкцией нечетких подмножеств, будем называть элементарными. В некоторых задачах закономерности, удовлетворяющие критерию значимости, имеют существенно “непрямоугольную” форму. При этом такие закономерности могут быть приближены множеством элементарных правил, каждое из которых, вообще говоря, не удовлетворяет предикату C. Такие правила назовем *сложными*, или *дизъюнктивными*.

Определение 14. Рассмотрим правила R_1, R_2 и их посылки $\mathrm{Sump}_1 = \{ M_{i_1(1)}^{j_1(1)}, \ldots, M_{i_{r(1)}(1)}^{j_{r(1)}(1)} \}$ и $\mathrm{Sump}_2 = \{ M_{i_1(2)}^{j_1(2)}, \ldots, M_{i_{r(2)}(2)}^{j_{r(2)}(2)} \}$. Соответственно, ν_1 и ν_2 – степени принадлежности этим посылкам. Данные правила будем называть *расположенными по соседству*, если найдется объект $\mathbf{x} \in \mathbb{R}^d$ такой, что

$$\mu_{M_{i_k(1)}^{j_k(1)}}(x_{i_k(1)}) \geq \alpha_{i_k(1)}, \quad k = \overline{1, r(1)},$$

$$\mu_{M_{i_k(2)}^{j_k(2)}}(x_{i_k(2)}) \geq \alpha_{i_k(2)}, \quad k = \overline{1, r(2)}.$$

Здесь $\boldsymbol{\alpha}$ задает соответствующее покрытие в экспертной интерпретации признаков.

Фактически это определение означает, что у соседних правил “ядра”, определяемые условными границами соответствующих разбиений, пересекаются. В практических приложениях разумным является использование критерия значимости правил, определяемого пересечением предикатов C^h и C^s (фиг. 8). При этом ордината пересечения левой и нижней границ области значимых правил (обозначим ее c_e^{super}, и назовем порогом суперэффективности) обладает интересным свойством: результат объединения любого правила с эффективностью, превосходящей порог суперэффективности, и произвольного значимого правила является значимым дизъюнктивным правилом.

В настоящей работе предлагается алгоритм генерации подобных дизъюнктивных правил, являющийся расширением схемы МЭС. Он позволяет отыскивать значимые правила, состоящие из расположенных по соседству элементарных (некоторые, но не все).

Данная модификация касается только шага 3, вместо которого выполняется шаг 3'.

Шаг 3'. Вычислим репрезентативность и эффективность каждого правила. Если $C^{\alpha_0}(\cdot, \cdot) = 0$ и $\mathrm{eff}(R) < c_e^{\mathrm{super}}$, то правило исключается из дальнейших рассмотрений. Если $C^{\alpha_0}(\cdot, \cdot) = 0$, а его эффективность $\mathrm{eff}(R) \geq c_e^{\mathrm{super}}$ и при этом в базе знаний уже содержится правило, являющееся соседним по отношению к рассматриваемому, то тогда правило заносится в базу результирующих правил.

Модифицированная схема МЭС позволяет получить, вообще говоря, более точные решения по сравнению с МЭС за счет большего покрытия множества прецедентов правилами. Результаты испытаний на некоторых практических задачах приведены в разд. 5. Отметим, что данная модификация метода эффективных сужений решает задачу поиска всех дизъюнктивных правил приближенно. Любой метод получения точного решения имеет более высокую алгоритмическую сложность.

4. ОПТИМИЗАЦИЯ В МОДЕЛИ АЛГОРИТМОВ

Рассматриваемая задача синтеза экспертной системы распознавания образов с использованием прецедентной информации, как отмечалось выше, допускает построение решения в три этапа: задание экспертной интерпретации каждого признака, получение набора правил (знаний) и выбор наилучшего алгоритма в модели (2.5) (уточнение весов правил, параметров покрытий, границ разбиений и пр.). Решение каждого из этих этапов может проводиться как экспертом, так и автоматически (а также совместно). Рассмотрим последовательно вопросы поиска и оптимизации значений для каждой из групп параметров модели (2.5).

4.1. Первичное формирование границ нечетких подмножеств

В [16], [17], [13] задача поиска оптимального разбиения, а также параметров покрытия нечеткими подмножествами решается в одной связке с генерацией множества правил. В данной работе понятие экспертной интерпретации предлагается разделить на понятие разбиения области значений признаков и понятие набора покрывающих нечетких подмножеств. Первое из этих понятий является независимым и определяет трактовку экспертом самого понятия признака. Второе понятие отвечает за меру неточности в определении экспертной интерпретации и вполне может оптимизироваться с учетом прецедентной информации и обнаруженного набора правил вывода. В качестве покрытия можно выбрать $\{\boldsymbol{\alpha}, \boldsymbol{\beta}\}$- или $[\boldsymbol{\alpha}, \boldsymbol{\beta}]$-покрытия и оптимизировать функционал качества по параметрам $\boldsymbol{\alpha}$ и $\boldsymbol{\beta}$.

Пусть требуется построить некоторое прямоугольное разбиение пространства признаков $\{\{a_i^j\}_{j=1}^{n_i}\}_{i=1}^{d}$. Обозначим проекцию множества объектов обучающей выборки на i-й признак через множество $\{x_i^k\}_{k=1}^{q}$. Пусть также дано требуемое количество интервалов для каждого признака: n_i, $i = \overline{1, d}$. Тогда для нахождения первичного разбиения $a_i^1 < \ldots < a_i^{n_i}$ i-го признака разумно использовать результат одномерной кластеризации. Для решения этой задачи может быть задействован, например, метод равномерной массы (проводить разбиение таким образом, чтобы в каждой группе было примерно одинаковое число объектов, или в случае, если объекты имеют индивидуальные веса, группы объектов должны иметь примерно одинаковые веса). В простейшем виде этот метод таков.

Шаг 1. Упорядочим множество $\{x_i^k\}_{k=1}^{q}$: $x_i^{k_1} \leq \ldots \leq x_i^{k_q}$.

Шаг 2. Положим

$$a_i^1 = x_i^{k_1} - \frac{x_i^{k_q} - x_i^{k_1}}{2(n_i - 2)}, \quad a_i^{n_i} = x_i^{k_q} + \frac{x_i^{k_q} - x_i^{k_1}}{2(n_i - 2)}.$$

Шаг 3. Для каждого j от 2 до $n_i - 1$ вычислим $t = \left\lfloor \frac{(j-1)q}{n_i - 1} \right\rfloor$ и положим $a_i^j = \frac{x_i^{k_t} + x_i^{k_{t+1}}}{2}$. Конец работы алгоритма.

Также для решения задачи первичных границ могут быть использованы другие методы кластеризации, например K-средних (см. [3], [27]), иерархическая группировка (см. [3], [27]), восстановление смеси нормальных распределений (см. [28], [29]) и др.

Значения параметров $\{n_i\}_{i=1}^{d}$ (число кластеров для каждого признака) могут быть определены экспертом вручную путем визуального оценивания качества кластеризации в одномерном пространстве. Если количество признаков велико и визуальное решение требует значительных временных затрат, могут быть использованы методы автоматического определения числа кластеров. Среди них можно выделить байесовский подход для смеси нормальных распределений (см. [28], [29]), байесовский подход для смеси распределений Стьюдента (см. [30]), а также подход, основанный на многоразовой кластеризации пар подвыборок из исходной выборки с последующим построением гистограммы значений критерия сравнения двух кластеризаций (см. [31], [32]).

Разбиения множеств значений признаков могут последовательно уточняться в процессе исследования данной прикладной области. Например, объекты, неправильно распознанные на очередной итерации исследования, могут менять свое влияние на разбиение значений в будущем. В [23] предлагается использование процедуры адаптивной коррекции или бустинга (см. [21]) для формирования границ разбиений.

4.2. Оптимизация весов правил

Пусть известны экспертные интерпретации всех признаков $\mathfrak{T} = \{\mathfrak{T}_1, \ldots, \mathfrak{T}_d\}$ и построена база знаний $\mathfrak{R} = \{R \in \mathfrak{R}^* \mid C(\mathrm{rep}(R), \mathrm{eff}(R)) = 1\} = \{R_i\}_{i=1}^{m}$. Требуется провести оптимизацию значений весов правил $\{w_i\}_{i=1}^{m}$ с целью минимизации вероятности ошибки алгоритма на произвольном объекте признакового пространства. Однако эта величина недоступна для измерения, а ее оценивание представляет собой отдельную математическую проблему (см. [33]). В связи с этим рассматривают альтернативные функционалы качества в терминах обучающей выборки и модели алгоритмов.

Случай задачи классификации. Предположим, что решается задача распознавания образов (классификации). Тогда решающее правило (2.3) может рассматриваться как выбор класса, которому соответствует максимальная оценка. В качестве такой оценки выступает величина $\mu_{R_{N^k}}(\mathbf{x}, \mathrm{med}(N^k))$ для класса k. Введем следующую функцию правдоподобия данных и модели (заданного набора весов правил):

$$p(\mathbf{y} \mid \mathbf{w}) = \prod_{i=1}^{q}\left[\frac{\lambda + \mu_{R_{N^{y^i}}}(\mathbf{x}^i, \mathrm{med}(N^{y^i}))}{\lambda l + \sum_{c=1}^{l}\mu_{R_{N^c}}(\mathbf{x}^i, \mathrm{med}(N^c))}\right].$$

Здесь $\lambda > 0$ – произвольный положительный, заранее заданный параметр.

Известный метод максимума правдоподобия приводит к следующей оптимизационной задаче на значения весов правил:

$$\begin{gathered}\log p(\mathbf{y} \mid \mathbf{w}) \longrightarrow \max_{\mathbf{w}},\\ w_i \geq 0, \quad i = \overline{1, m}.\end{gathered} \tag{4.1}$$

Однако простая максимизация правдоподобия (4.1) может приводить к существенному переобучению алгоритма распознавания. Для того чтобы избежать данной ситуации, используются различные методы регуляризации обучения (см. [4], [29], [34], [35]). В данной статье предлагается использовать так называемый байесовский подход (см. [4], [29], [34], [36]–[38]). В этом подходе регуляризация задается при помощи априорного распределения в пространстве оптимизируемых параметров. Функция принятия решений в решающем правиле (2.3) может трактоваться, как взвешенная линейная комбинация функций принадлежности правилам. Таким образом, данная модель близка по своему виду к обобщенным линейным моделям, которые используются в таких популярных алгоритмах распознавания, как метод опорных векторов (см. [39], [40]), метод релевантных векторов (см. [38]), предсказание с помощью гауссовских процессов (см. [34], [41]) и др. Обычно при применении байесовского подхода в обобщенных линейных моделях в качестве априорного распределения используется нормальное распределение (см. [37], [38]) или двухстороннее распределение Лапласа (см. [42]). В задаче (4.1) веса являются неотрицательными величинами. Поэтому в качестве априорного распределения для весов правил предлагается использовать одностороннее нормальное распределение (4.2) или экспоненциальное распределение (4.3):

$$p(w_i \mid \alpha_i) = \sqrt{\frac{2\alpha_i}{\pi}}\exp\left(-\frac{1}{2}\alpha_i w_i^2\right), \quad i = \overline{1, m}, \tag{4.2}$$

$$p(w_i \mid \alpha_i) = \alpha_i \exp(-\alpha_i w_i), \quad i = \overline{1, m}. \tag{4.3}$$

Заметим, что здесь веса правил предполагаются независимыми случайными величинами, т.е. априорное распределение в пространстве весов есть произведение априорных распределений для каждого весового коэффициента. При этом набор параметров $\{\alpha_i\}_{i=1}^m$ определяет конкретный вид априорного распределения для каждого веса.

Для подбора гиперпараметров (параметров модели) $\{\alpha_i\}_{i=1}^m$ в байесовском подходе используется процедура максимизации обоснованности (правдоподобия модели, см. [36], evidence см. [38]):

$$p(\mathbf{y} \mid \boldsymbol{\alpha}) = \int p(\mathbf{y} \mid \mathbf{w}) p(\mathbf{w} \mid \boldsymbol{\alpha}) d\mathbf{w} \longrightarrow \max_{\boldsymbol{\alpha}}. \tag{4.4}$$

Формально в байесовском подходе для гиперпараметров также необходимо определить априорное распределение и использовать процедуру максимизации правдоподобия гипермодели. Такая процедура введения гиперпараметров очередного уровня может быть продолжена до тех пор, пока на каком-то этапе в качестве априорного распределения не выбирается равномерное распределение в области изменения параметров. В большинстве практических задач обычно ограничиваются одним уровнем.

Интеграл (4.4) не может быть вычислен аналитически. Поэтому при практическом применении процедуры максимизации обоснованности часто используется так называемое приближение Лапласа – замена логарифма подынтегральной функции параболой, что влечет за собой приближение подынтегральной функции гауссианой, интеграл от которой является легко вычислимым выражением (см. [37], [38]). При этом значения, соответствующие максимуму подынтегральной функции (регуляризованного функционала), являются искомыми весами правил.

Для случая априорного распределения типа (4.2) или (4.3) в [37], [38] получена эффективная итерационная процедура максимизации лапласовского приближения обоснованности для произвольного вида функции правдоподобия данных и модели $p(\mathbf{y} \mid \mathbf{w})$. Для ее получения необходимо приравнять к нулю значения производных логарифма обоснованности по параметрам и использовать прием из (см. [37]).

Случай задачи восстановления регрессии. Для задачи восстановления регрессии и решающего правила (2.4) введем следующую функцию правдоподобия данных и модели:

$$p(\mathbf{y} \mid \mathbf{w}) = (\sqrt{2\pi}\sigma)^{-m}\exp\left(-\frac{1}{2\sigma^2}\|\mathbf{y}-\hat{\mathbf{y}}\|^2\right). \tag{4.5}$$

Здесь $\hat{y}^i = M[\ldots](\mathbf{x}^i)$ – результат предсказания выходной переменной с помощью экспертной системы M для объекта обучения $\mathbf{x}^i$, а σ – некоторый положительный параметр, определяющий уровень шума в данных.

Так же как и в случае задачи классификации, для подбора значений весов правил здесь может быть использована простая процедура максимума правдоподобия или процедура регуляризации в рамках байесовского подхода. Заметим, что в случае регрессии использование байесовского подхода позволяет также оценить значение параметра σ при помощи максимизации обоснованности.

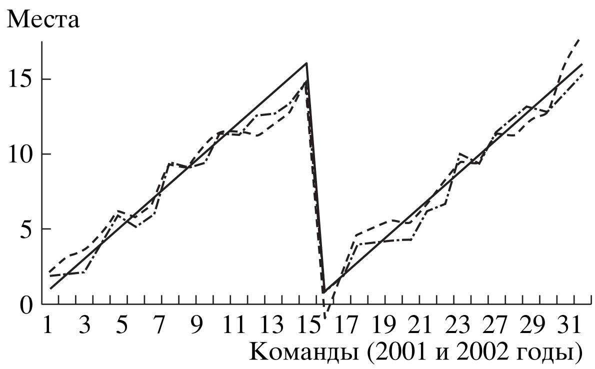
В более общем случае функция правдоподобия $p(\mathbf{y} \mid \mathbf{w})$ выбирается в виде $\exp[-\Sigma_{i=1}^q l(y^i, \hat{y}^i)]$, где $l : \mathbb{R}^2 \longrightarrow \mathbb{R}$ – некоторая функция потерь. Выбор квадратичной функции потерь (4.5) не всегда является адекватным, так как в этом случае большие отклонения штрафуются слишком сильно. В [43] предложена универсальная функция потерь, которая не штрафует близкие значения, средние отклонения штрафует квадратично, а большие отклонения – линейно. Кроме того, эта функция является аналитической. Значения параметров, регулирующих границы между близкими, средними и большими отклонениями, могут быть найдены при помощи байесовского подхода. При этом процедура совершенно аналогична случаю с квадратичной функцией потерь.

4.3. Оптимизация границ разбиений

Пусть заданы экспертные интерпретации признаков $\mathfrak{T}$ и множество правил $\mathfrak{R}$. Тогда имеется модель алгоритмов $\mathfrak{M} = \mathfrak{M}_{\mathfrak{R}}[\{\{a_i^j\}_{j=1}^{n_i}\}_{i=1}^{d}; \{w_m\}_{m=1}^{N}]$. На практике условные границы разбиений могут быть выбраны неадекватно, например если они выбираются как результат одномерной кластеризации данных. Введение покрытий признаков нечеткими подмножествами позволяет оптимизировать некоторый функционал качества Q по границам разбиений. Получается задача оптимизации с линейными ограничениями:

$$\begin{gathered} Q \longrightarrow \max_{\mathbf{a}}, \\ a_i^1 < \ldots < a_i^{n_i}, \quad i = \overline{1, d}. \end{gathered} \tag{4.6}$$

В качестве функционала качества выступает правдоподобие данных и модели. Как и во многих других нетривиальных моделях алгоритмов распознавания, рассматриваемый функционал качества является многоэкстремальным. Для бесконечно дифференцируемого семейства характеристических функций (1.6) обобщенный градиент вычисляется аналитически. Таким образом, для поиска приближенного решения поставленной задачи можно воспользоваться любым известным методом оптимизации, учитывающим обобщенный градиент.



Фиг. 9.

Итак, полностью изложена схема построения решения задачи распознавания образов с заданным функционалом качества. Результатом является алгоритм распознавания, принимающий решение о классификации объектов на основании знаний, сформулированных в терминах эксперта. Экспертная интерпретация признаков и правила, сформулированные относительно нее, могут быть синтезированы автоматически по заданной прецедентной информации. Впоследствии эксперт имеет возможность верифицировать и модифицировать полученные правила.

5. ЭКСПЕРИМЕНТЫ

Предложенная модель нечеткой экспертной системы (ЭС) была протестирована на нескольких задачах.

Сначала рассмотрим результаты тестирования в качестве системы для получения непрерывного прогноза. Первая задача заключается в определении мест команд в российском футбольном чемпионате за два года по турнирной таблице. Данные представляют собой таблицу с 32 объектами и 6 признаками. В качестве признаков выступают, соответственно, итоговое место в чемпионате, число побед, число ничьих, число поражений, количество забитых мячей и количество пропущенных мячей.

Необходимо определить итоговое положение команды в таблице (предсказать значение первого признака). Результаты испытаний сравнивались с множественной линейной регрессией. На фиг. 9 представлены результаты прогнозирования. Сплошная линия – точные значения, штрихпунктирная – предсказание ЭС, штриховая – предсказание линейной регрессии. Команды были отсортированы в соответствии с их итоговым положением. Несмотря на высокую корреляцию между признаками и прогнозируемой переменной (линейная зависимость между числом побед, ничьих, поражений и количеством набранных очков), для линейной регрессии среднеквадратичное отклонение составило 38.056, в то время как для ЭС оно равнялось 21.936. Кроме того, система ЭС генерировала небольшое количество легко интерпретируемых правил. Вот, например, некоторые правила, которые нашла система для футбольного чемпионата:

ЕСЛИ Пропущенных мячей = Немного И Поражений = Мало, ТО Место = Высокое

ЕСЛИ Побед = Немало И Забитых мечей = Много И Ничьих =

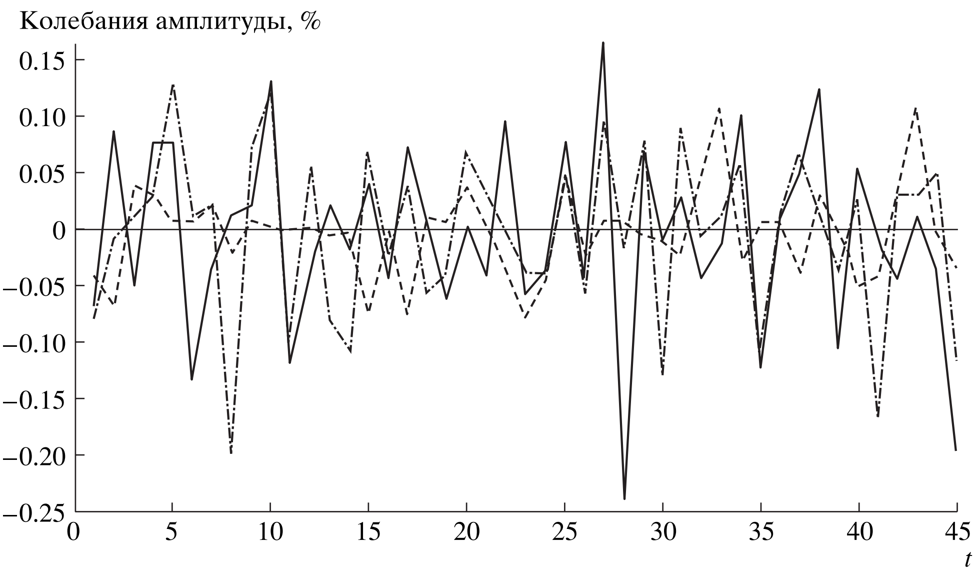
= Несколько, ТО Место = Призовое

ЕСЛИ Забитых мячей = Мало, ТО Место = Низкое

ЕСЛИ Пропущенных мячей = Немного И Ничьих = Много, ТО Место = Невысокое

Знания подобного рода особенно полезны в ситуации, когда значения некоторых признаков можно контролировать. Учитывая степень влияния этих признаков на скрытое значение прогноза, мы можем эффективно управлять процессом путем модификации значений соответствующих признаков.

Также был проведен ряд экспериментов по сравнению описанной нечеткой системы с альтернативным подходом системы MatLab fuzzy logic toolbox. Последняя представляет собой развитие так называемого нейро-нечеткого подхода (см. [14], [15], [44]). Она также позволяет генерировать автоматически нечеткие правила с использованием прецедентной информации и опреде-



Фиг. 10.

лять формы нечетких множеств. Для исследования была выбрана следующая задача: предсказание колебаний амплитуды магнитного поля в ускоряющем элементе “cavity” клистрона. Необходимые данные были получены для гамбургского линейного ускорителя в институте DESY (город Гамбург, Германия). Данная задача имеет важное практическое значение, так как точность работы ускорителя в значительной мере подвержена многочисленным внешним воздействиям (например, колебаниям земной поверхности из-за движения автомобилей, нестабильности геомагнитной обстановки, наличием различного рода шумов, состоянием атмосферы и пр.). Предсказание амплитуды колебаний магнитного поля позволяет применять нейтрализующие воздействия для предотвращения колебаний с целью добиться более надежной работы системы. В качестве исходного признакового пространства в данном случае выступали значения колебаний в других ускоряющих элементах того же клистрона. Результаты работы ЭС и MatLab на контрольной таблице представлены на фиг. 10 (колебания магнитного поля в % по времени). Сплошная линия – точные значения, штриховая – предсказание ЭС, пунктирная – MatLab. Система MatLab склонна к переобучению, несмотря на использование дополнительной выборки для его предотвращения, а также в значительной мере хуже улавливает основные пики колебаний и общую тенденцию.

Для случая задач классификации система ЭС сравнивалась с некоторыми распространенными методами распознавания – линейным дискриминантом Фишера (LDF), q-ближайших соседей (QNN), тестовым алгоритмом (TA), комитетом гиперплоскостей (LM), методом опорных векто-

Таблица 1

Метод распознавания	Меланома (малая выборка)	Фонемы (большая выборка)	Почка (средняя выборка, много классов)
MLP	65.6	78.2	77.5
LDF	56.3	77.4	77.5
TA	62.5	65.5	65.7
LM	50.0	77.2	79.3
SVM	59.4	77.4	83.1
QNN	62.5	84.7	80.3
ЭС	66.6	77.5	76.5

Таблица 2

Задача	МЭС	ДМЭС	ДМЭС$_Q$
AUSTRALIAN	19 ± 2.40	18.9 ± 1.41	18.42 ± 1.90
BUPA	43.61 ± 4.71	44.56 ± 4.64	38.16 ± 4.59
CLEVELAND	19.45 ± 2.10	19.08 ± 1.90	18.73 ± 1.82
CREDIT	18.94 ± 1.93	19.51 ± 1.77	19.26 ± 2.37
HEPATITIS	40.23 ± 5.22	39.91 ± 5.35	40.34 ± 5.26
HUNGARY	18.8 ± 1.74	18.93 ± 1.79	17.9 ± 1.83
Оценка	13	14	9

ров (SVM) многослойным перцептроном (MLP). Результаты работы (точность на контрольной выборки в %) представлены в табл. 1.

В качестве задач выступали:

МЕЛАНОМА. Определение степени заболевания по данным амбулаторного наблюдения; 33 признака, 3 класса. Всего 48 объектов для обучения, 32 – для контроля.

ФОНЕМЫ. Определение типа фонемы по ряду фильтрационных признаков сигнала; 6 признаков, 2 класса. Всего 2200 объектов для обучения, 1404 – для контроля.

ПОЧКА. Определение степени заболевания по данным амбулаторного наблюдения; 8 признаков, 7 классов. Всего 170 объектов для обучения, 150 – для контроля.

Также были проведены эксперименты по сравнению предлагаемых методов автоматической генерации знаний по прецедентной информации. Для сравнения их качества было произведено 120 экспериментов на основе 6 задач из UCI-репозитория (см. [45]). Данные для каждой задачи случайным образом разбивались на обучающую (33% всех объектов) и контрольную выборку (67% всех объектов) двадцатью различными способами. Для каждого из сравниваемых методов по обучающей выборке автоматически строился алгоритм классификации. Количество ошибок на контрольной выборке, усредненное по различным разбиениям, заносилось в соответствующую методу и задаче ячейку табл. 2. Столбец МЭС содержит результаты испытаний экспертной системы распознавания образов без оптимизации функционала качества, для которой правила генерировались методом эффективных сужений. В следующем столбце – аналогичные результаты экспертной системы распознавания образов без оптимизации функционала качества, основанной на дизъюнктивной модификации метода эффективных сужений. Столбец ДМЭС$_Q$ содержит результаты экспертной системы распознавания образов, также основанной на дизъюнктивной модификации метода эффективных сужений; веса правил и экспертные интерпретации признаков оптимизировались с точки зрения рассматриваемого функционала качества.

Для сравнения результатов в каждой строчке методам присваивалась оценка: 1 – лучший метод, 2 – следующий, 3 – худший. Суммы оценок в столбцах записаны в последней строке табл. 2.

В результате можно сделать несколько выводов. Метод эффективных сужений в основном работает не хуже, чем его дизъюнктивная модификация. Ввиду того что множество правил последнего метода всегда больше, ДМЭС склонен к переобучению. Однако оптимизация экспертных интерпретаций признаков и весов правил приводит к нужному улучшению результатов. Отсутствие переобучения на этапе оптимизации может быть объяснено наличием ограничений в формуле (4.6), связанных с наличием порядка между нечеткими подмножествами, составляющими экспертную интерпретацию каждого признака.

6. ЗАКЛЮЧИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

Данные испытаний показали, что предложенный метод построения нечеткой экспертной системы может быть использован для поддержки принятия решений в тех случаях, когда не существует строгой математической модели изучаемой области знаний, но имеется некоторая априорная информация и количество прецедентов не позволяет использовать нелинейные статистические методы. По-видимому, данный метод не стоит применять в тех случаях, когда необходимо получить абсолютно точное значение прогноза и/или существует строгая математическая модель явления.

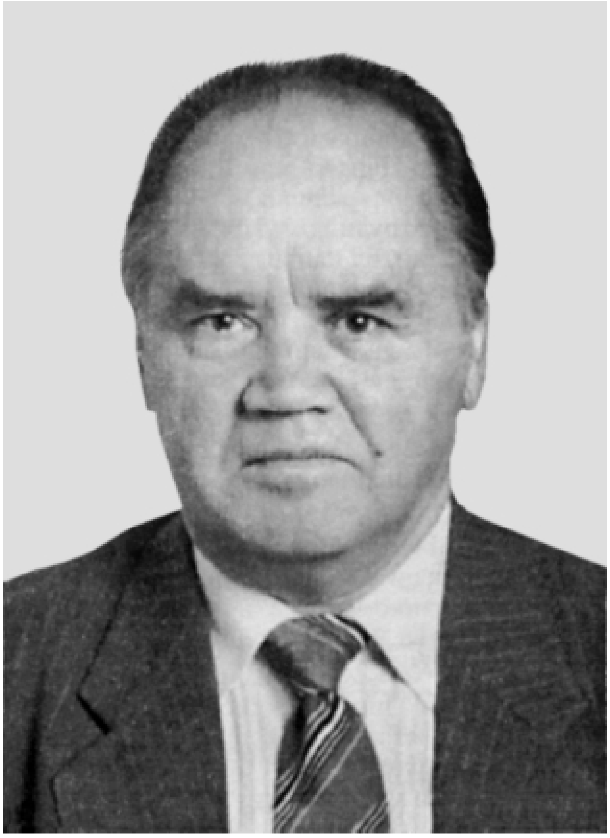
Следует заметить, что оптимизация набора нечетких правил, по сути, представляет собой подбор колоколообразной ядровой функции. Как уже отмечалось выше при оптимизации весов правил, предлагаемая модель близка к обобщенным линейным моделям, где в качестве элементарных признаков фигурируют нечеткие правила. Использование $\{\boldsymbol{\alpha}, \boldsymbol{\beta}\}$- и $[\boldsymbol{\alpha}, \boldsymbol{\beta}]$-покрытий фактически означает подбор колоколообразной ядровой функции. Решение подобной проблемы в ядровых методах распознавания, таких как методы опорных и релевантных векторов, является очень трудоемкой задачей (см. [46], [47]). В данном случае ее удается избежать за счет конструирования набора правил из элементарных блоков, задаваемых экспертом либо определяемых в результате кластеризации.

СПИСОК ЛИТЕРАТУРЫ

1. *Заде Л.* Размытые множества и их применение в распознавании образов и кластер-анализе // Классификация и кластер. М.: Мир, 1980. С. 208–247.
2. *Журавлев Ю.И.* Избранные научные труды. М.: Магистр, 1998.
3. *Журавлев Ю.И., Рязанов В.В., Сенько О.В.* РАСПОЗНАВАНИЕ. Математические методы. Программная система. Практические применения. М.: Фазис, 2006.
4. *Bishop C.M.* Neural networks for pattern recognition. Oxford: Univ. Press, 1995.
5. *Breiman L., Friedman J.H., Olshen R.A., Stone C.J.* Classification and regression trees. Belmont, CA: Wadsworth Internal. Group, 1984.
6. *Перфильева И.* Приложения теории нечетких множеств // Итоги науки и техн. Сер. Теория вероятностей. Матем. статистика. Теоретич. кибернетика. М.: ВИНИТИ, 1990. Т. 29. С. 83–151.
7. *Тэрано Т., Асаи К., Сугено М.* Прикладные нечеткие системы. М.: Мир, 1993.
8. *Mamdani E.* Advances in the linguistic synthesis of fuzzy controllers // Proc. 6th Internat. Symp. Multiple-Values Logic, 1976. P. 196–202.
9. *Заде Л.* Понятие лингвистической переменной и его применение к принятию приближенных решений. М.: Мир, 1980.
10. *Fukami S., Mizumoto M., Tanaka K.* Some considerations of fuzzy conditional inference // Fuzzy Sets and Systems. 1980. V. 3. P. 243–273.
11. *Разанов В.В., Сенько О.В.* О некоторых моделях голосования и методах их оптимизации // Распознавание, классификация, прогноз. 1990. Т. 3. С. 106–145.
12. *Воронцов К.В.* Лекции по логическим алгоритмам классификации; http://www.ccas.ru/voron/download/LogicAlgs.pdf.2006.
13. *Паклин Н.Б.* Адаптивные модели нечеткого вывода для идентификации нелинейных зависимостей в сложных системах: Дис. … канд. техн. наук. Ижевск: ИжГГУ, 2004.
14. *Ojala T.* Neuro-fuzzy systems in control. Master of science thesis. Tampere. Finland, 1994.
15. *Jang J.-S.R., Sun C.-T., Mizutani E.* Neuro-fuzzy and soft computing. Upper Saddle River, NJ: Prentice-Hall, 1997.
16. *Ishibuchi H., Nozaki K., Yamamoto N., Tanaka C.* Construction of fuzzy classification systems with rectangular fuzzy rules using genetic algorithms // Fuzzy Sets and System. 1994. V. 65. № 2/3. P. 237–253.
17. *Inoue H., Kamei K., Inoue K.* Rule pairing methods for crossover in GA for automatic generation of fuzzy control rules; citeseer.ist.psu.edu/200265.html.
18. *Gomez-Scarmeta A.F., Jimenez F.* Generating and tuning fuzzy rules using hybrid systems // Proc. 6th IEEE Internat. Conf. Fuzzy System. 1997. V. 1. P. 247–252.
19. *Rivest R.L.* Learning decision lists // Mach. Learn. 1987. V. 2. № 3. P. 229–246.
20. *Cohen W.W.* Fast effective rule induction // Proc. 12th Internat. Conf. Mach. Learn. CA: Morgan Kaufmann, 1995. P. 151–163.
21. *Freund Y., Schapire R.E.* Experiments with a new boosting algorithm // Proc. 13th Internat. Conf. Mach. Learn., 1996. P. 148–156.
22. *Cohen W.W., Singer Y.* A simple, fast, and effective rule learner // Proc. 16th Nat. Conf. Artificial Intelligence, 1999.
23. *Ветров Д.П., Кропотов Д.А.* Использование методов Boosting для генерации знаний // Тр. XII Всерос. конф. “Матем. методы распознавания образов”. М.: Макс-Пресс, 2005. С. 48–51.
24. *Quinlan J.R.* C4.5: Programs for machine learning. San Mateo, CA: Morgan Kaufmann, 1993.
25. *Breslow L.A., Aha D.W.* Simplifying decision trees: a survey // Knowledge Eng. Rev. 1997. V. 12. № 1. P. 1–40. siteseer.ist.psu.edu/breslow96simplifying.html.
26. *Дюличева Ю.Ю.* Модели коррекции редуцированных бинарных решающих деревьев: Дис. канд. физ.-матем. наук. Симферополь: ТавРГУ, 2004.
27. *Дуда Р., Харт П.* Распознавание образов и анализ сцен. М.: Мир, 1976.

28. *Corduneanu A., Bishop C.* Variational model selection for mixture distributions // Artificial Intelligence and Statistics. Morgan Kaufmann, 2001. P. 27–34.
29. *Шумский С.А.* Байесова регуляризация обучения // Лекции по нейроинформатике. Ч. 2. М.: МИФИ, 2002. С. 30–93.
30. *Bishop C.M., Svensen M.* Robust Bayesian mixture modeling // Proc. ESANN, 2004.
31. *Ben-Hur A., Elisseeff A., Guyon I.* A stability based method for discovering structure in clustered data // Proc. Symposium on Biocomputing. Lihue, Hawaii, 2002. P. 6–17.
32. *Kuncheva L.* Combining pattern classifiers: methods and algorithms. Wiley, 2004.
33. *Гуров С.И.* Оценка надежности классифицирующих алгоритмов. М.: Издат. отд. ВМиК МГУ, 2002.
34. *MacKay D.J.C.* Information theory, inference, and learning algorithms. Cambridge Univ. Press, 2003.
35. *Kropotov D.A., Tolstov I.V., Vetrov D.P.* Decision trees regularization based on stability principle // Pattern Recognition and Image Analysis. 2005. V. 15. № 1. P. 107–109.
36. *Berger J.O.* Statistical decision theory and bayesian analysis. Berlin etc: Springer, 1985.
37. *MacKay D.J.C.* Bayesian interpolation // Neural Computation. 1992. V. 4. № 3. P. 415–447.
38. *Tipping M.E.* Sparse bayesian learning and the relevance vector machine // J. Mach. Learn. Res. 2001. V. 1. P. 211–244.
39. *Burges C.* A tutorial on support vector machines for pattern recognition // Data Mining and Knowledge Discovery. 1998. V. 2. P. 121–167.
40. *Vapnik V.N.* Statistical learning theory. Wiley, 1998.
41. *Williams C.K.I.* Prediction with gaussian processes: from linear regression to linear prediction and beyond // Learning in Graphical Models. MIT. 1999. P. 599–621.
42. *Williams P.M.* Bayesian regularization and pruning using a laplace prior // Neural Computation. 1995. V. 7. № 1. P. 117–143.
43. *Chu W.* Bayesian approach to support vector machines: PhD thesis. Nat. Univ. of Singapore, 2003.
44. *Дьяконов В., Круглов В.* Математические пакеты расширения MATLAB. Специальный справочник. М.: Питер, 2001.
45. *Murphy P., Aha D.* UCI repository of machine learning databases // Univ. California, Dept. Informat. and Comput. Sci. California: Irvine, 1996; http://www.ics.uci.edu/~mlearn/MLRepository.html.
46. *Kropotov D.A., Ptashko N.O., Vetrov D.P.* The use of bayesian framework for kernel selection in vector machines classifiers // Progress in Pattern Recognition, Image Analysis and Applic. LNCS 3773. Berlin etc: Springer, 2005. P. 252–261.
47. *Kropotov D.A., Vetrov D.P., Ptashko N.O., Vasiliev O.M.* The use of stability principle for kernel determination in relevance vector machines // ICONIP2006, Part I, LNCS 423. Berlin etc: Springer, 2006. P. 727–736.

ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 2007, том 47, № 8, с. 1455–1456



ПАМЯТИ АКАДЕМИКА ВАЛЕНТИНА ВАСИЛЬЕВИЧА ВОЕВОДИНА (1934–2007)

© 2007 г. М. К. Керимов

(119991 Москва, ул. Вавилова, 40, ВЦ РАН)

e-mail: comp_mat@ccas.ru

С глубоким прискорбием сообщаем, что 27 января 2007 года от сердечной недостаточности скончался известный российский ученый-математик, академик Российской Академии Наук Валентин Васильевич Воеводин. Наука потеряла своего активно работающего представителя, семья – любимого мужа и отца. Это стало невосполнимой утратой для сотрудников и друзей, которые работали с ним в Институте вычислительной математики РАН, для членов редколлегии Журнала вычислительной математики и математической физики. Не стало одного из авторитетнейших экспертов по вычислительным методам линейной алгебры, по методам параллельных вычислений на ЭВМ. Мы потеряли веселого, жизнерадостного, доступного для бесед, доброжелательного к людям человека. Многие специалисты, обращавшиеся к нему за консультациями, уходили от него вполне удовлетворенными и благодарными этому простому в общении человеку. Валентин Васильевич с 1977 года до своей кончины являлся одним из активнейших членов редколлегии Журнала вычислительной математики и математической физики. С сожалением приходится отмечать, что совсем недавно я написал статью к его 70-летию, а теперь приходится говорить о нем такие скорбные слова.

В.В. Воеводин родился 22 марта 1934 года в селе Шилово Рязанской области в семье служащих. Отец его, Василий Никитович Воеводин, был инженером по технике безопасности. Мать, Антонина Петровна, работала машинисткой. В 1952 г. В.В. Воеводин окончил школу № 643 города Москвы с золотой медалью и поступил на механико-математический факультет Московского государственного университета им. М.В. Ломоносова (МГУ). В 1957 году он окончил МГУ по кафедре вычислительной математики и получил диплом математика с отличием. С 1956 по

1980 г. В.В. Воеводин работал в Научно-исследовательском центре МГУ, пройдя путь от старшего лаборанта до директора. С 1969 по 1979 г. он был директором этого Центра и во многом способствовал становлению и развитию этого важного для науки и страны нового научного учреждения. С 1980 г. до своей кончины Валентин Васильевич работал главным научным сотрудником Института вычислительной математики РАН в г. Москва. Фактически его можно считать одним из создателей этих двух научных учреждений.

В 1962 г. В.В. Воеводин защитил диссертацию на степень кандидата физико-математических наук по теме “Решение полной проблемы собственных значений степенными методами”, а в 1969 г. – докторскую диссертацию по теме “Ошибки округления и устойчивость в прямых методах линейной алгебры”. Обе диссертации Валентин Васильевич защитил на ученом совете МГУ. Своими учителями он считал академиков Андрея Николаевича Тихонова (1906–1999) и Гурия Ивановича Марчука.

В 1987 г. В.В. Воеводин был избран членом-корреспондентом АН СССР по специальности “математика”, а в 2000 г. – академиком РАН по специальности “математика, в том числе вычислительная математика”.

В.В. Воеводин опубликовал около 140 научных работ, из них 13 монографий и учебников. На его учебных пособиях воспитана плеяда известных математиков. Среди его непосредственных учеников 6 докторов и 25 кандидатов физико-математических наук. В.В. Воеводин – Лауреат премии Правительства Российской Федерации в области образования, он был награжден рядом правительственных наград. Последние опубликованные его работы были посвящены именно математическому образованию[1)]. Вся научная деятельность В.В. Воеводина была подчинена главной цели его жизни: постановке и решению различных математических проблем, связанных с эффективным решением прикладных задач на вычислительных системах различной архитектуры. Основные его профессиональные интересы были связаны со следующими направлениями математики и ее приложений:

1. Разработка численных методов (главным образом вычислительной алгебры).

2. Исследование ошибок округления и устойчивость вычислительных методов.

3. Информационная структура алгоритмов.

4. Математические модели в вычислительных процессах.

5. Программное обеспечение процессов вычисления.

6. Подготовка высококвалифицированных кадров математиков.

Более подробно о деятельности Валентина Васильевича сказано в моей статье, посвященной его 70-летию[2)]. Там же приводится список его основных научных работ.

В.В. Воеводин был большим любителем природы. Он вместе с семьей и сотрудниками обошел вдоль и поперек подмосковные леса. В последние годы он “осел” в дальней подмосковной деревне, там собственными силами построил свою “фазенду” и летние отпуска проводил вместе с семьей на даче. До последних дней своей жизни он оставался бодрым, жизнерадостным, инициативным. Все свои служебные и научные обязанности ученого он выполнял аккуратно и с душой. Никто из его друзей не мог предположить, что Валентин Васильевич так рано уйдет из жизни.

Члены редколлегии Журнала вычислительной математики и математической физики (его первая серьезная научная работа была опубликована именно в этом журнале), сотрудники редакции, много лет общавшиеся с ним, многочисленные авторы журнала, которые остались перед ним в долгу за бескорыстную помощь, глубоко скорбят по случаю его безвременной кончины, будут долго помнить его светлый образ. Валентина Васильевича физически нет в живых, но его научное наследие долго будет вдохновлять молодых исследователей.

[1)] *В.В. Воеводин.* Параллельные вычисления и математическое образование // Матем. в высшем образовании. 2005. № 5. С. 9–26.
В.В. Воеводин. Вычислительная математика и структура алгоритмов. М.: Изд-во МГУ, 2006. 112 с.
В.В. Воеводин. Вглядываясь в прошлое. Посвящается памяти Ивана Семеновича Березина. М.: НИВЦ МГУ, 1955–2007.

[2)] *М.К. Керимов.* К семидесятилетию академика Валентина Васильевича Воеводина // Ж. вычисл. матем. и матем. физ. 2004. Т. 44. № 11. С. 1923–1927.