



Math-Net.Ru

Общероссийский математический портал

Г. Ю. Куликов, Вложенные симметричные неявные гнездовые методы Рунге–Кутты типов Гаусса и Лобатто для решения жестких обыкновенных дифференциальных уравнений и гамильтоновых систем, *Ж. вычисл. матем. и матем. физ.*, 2015, том 55, номер 6, 986–1007

DOI: <https://doi.org/10.7868/S0044466915030114>

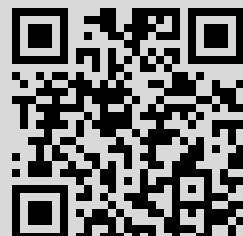
Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 185.215.176.82

6 декабря 2022 г., 17:34:00



УДК 519.622

ВЛОЖЕННЫЕ СИММЕТРИЧНЫЕ НЕЯВНЫЕ ГНЕЗДОВЫЕ МЕТОДЫ РУНГЕ–КУТТЫ ТИПОВ ГАУССА И ЛОБАТТО ДЛЯ РЕШЕНИЯ ЖЕСТКИХ ОБЫКНОВЕННЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ И ГАМИЛЬТОНОВЫХ СИСТЕМ¹⁾

© 2015 г. Г. Ю. Куликов

(CEMAT, Institute Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal)

e-mail: gkulikov@math.ist.utl.pt

Поступила в редакцию 29.07.2014 г.

Предлагается методика построения неявных гнездовых методов Рунге–Кутты для решения обыкновенных дифференциальных уравнений, которые относятся к классу монологичных формул этого типа. Их отличительной чертой является высокая практическая эффективность, которая вытекает из сохранения размерности исходной системы дифференциальных уравнений при применении, что невозможно в неявных многостадийных формулах Рунге–Кутты общего вида. С другой стороны, неявные гнездовые методы Рунге–Кутты наследуют все наиболее важные свойства общих формул этого вида, такие как A -устойчивость, симметрия, симплектичность в определенном смысле. Кроме того, они могут иметь достаточно высокие стадийные и классические порядки, а также обеспечивать плотную выдачу результатов интегрирования той же точности, что и порядок основного метода, без больших дополнительных затрат. Таким образом, гнездовые методы эффективны для численного интегрирования дифференциальных уравнений самых разных видов, к которым относятся жесткие и нежесткие задачи, а также гамильтоновы системы и обратимые уравнения.

Настоящая статья посвящена обобщению предложенных ранее гнездовых методов, основанных на квадратурных формулах Гаусса, на методы типа Лобатто. Более того, в статье представлена единая методика построения всех таких методов. Ее работоспособность продемонстрирована на вложенных примерах неявных гнездовых формул разного порядка. Все построенные методы снабжены механизмами для оценки локальной ошибки численного интегрирования и автоматической генерации переменных сеток на отрезке интегрирования за счет выбора оптимального шага. Такие вычислительные процедуры апробированы на тестовых задачах с известным решением и исследованы в сравнении со встроенными решателями системы матричных вычислений MATLAB. Библ. 73. Фиг. 3.

Ключевые слова: численные методы, решения обыкновенных дифференциальных уравнений, неявный гнездовой метод Рунге–Кутты, вложенные формулы, методы Рунге–Кутты типов Гаусса и Лобатто, автоматический контроль точности, жесткая задача Коши, гамильтоновы уравнения.

DOI: 10.7868/S0044466915030114

1. ВВЕДЕНИЕ

Минимизация вычислительных затрат, связанных с решением линейных систем, — одно из главных средств для повышения практической эффективности неявных методов Рунге–Кутты при решении жестких дифференциальных задач Коши большой размерности

$$x'(t) = g(t, x(t)), \quad t \in [t_0, t_{\text{end}}], \quad x(t_0) = x^0, \quad (1.1)$$

где правая часть $g: \mathbf{R}^{m+1} \rightarrow \mathbf{R}^m$ обладает необходимой гладкостью. Здесь и далее предполагается, что система дифференциальных уравнений (1.1) имеет единственное решение на всем отрезке

¹⁾ Работа выполнена при финансовой поддержке португальского фонда науки и технологии (Fundação para a Ciência e a Tecnologia) в рамках проекта Pest-OE/MAT/UI0822/2011 и программы Investigador FCT 2013.

интегрирования $[t_0, t_{\text{end}}]$. Следует отметить, что необходимость в решении таких задач возникает естественным образом в ряде прикладных исследований (см., например, [1]–[4]), а также при использовании более сложных вычислительных процедур, включая метод стрельбы для решения краевых задач, алгоритмы калмановской фильтрации и метод прямых для дискретизации задач математической физики (см., например, [5]–[14]). Поэтому разработке эффективных вычислительных формул и программ для решения уравнений (1.1) отводится важное место в отечественной и зарубежной вычислительной математике (см., например, [14]–[26]).

Как сказано выше, тематике удешевления вычислений при численном решении жестких дифференциальных уравнений большой размерности вида (1.1) уделялось и уделяется серьезное внимание в мировой литературе, а основной подход к реализации неявных методов Рунге–Кутты, снижающий указанные выше вычислительные издержки, заложен еще в [27] и [28]. Главная идея этих статей состоит в учете структуры матрицы коэффициентов A l -стадийного неявного метода Рунге–Кутты (РК-метода). Напомним, что все такие методы для решения задачи (1.1) могут быть записаны следующим образом:

$$t_{ki} = t_k + \tau_k c_i, \quad (1.2a)$$

$$x_{ki} = x_k + \tau_k \sum_{j=1}^l a_{ij} g(t_{kj}, x_{kj}), \quad i = 1, 2, \dots, l, \quad (1.2b)$$

$$x_{k+1} = x_k + \tau_k \sum_{j=1}^l b_j g(t_{kj}, x_{kj}), \quad k = 0, 1, \dots, K-1, \quad (1.2b)$$

где $x_0 = x^0$, а τ_k , как обычно, – размер $k+1$ -го шага численного интегрирования. Понятно, что любая РК-формула (1.2) однозначно задается вещественными коэффициентами a_{ij} , b_j и c_i , $i, j = 1, 2, \dots, l$, которые принято оформлять в виде таблицы Бутчера

$$\begin{array}{c|c} c & A \\ \hline & b^T \end{array},$$

в которой A обозначает матрицу размера $l \times l$, a , b и c есть l -мерные векторы (см., например, [20], [26]). Неявность означает, что некоторые (или все) коэффициенты в верхней треугольной части матрицы A , включая главную диагональ, отличны от нуля при любой перестановке строк и столбцов, что эквивалентно изменению порядка уравнений и переименованию неизвестных в системе (1.2б). В этом случае формулы (1.2б) определяют систему нелинейных уравнений, которые необходимо решать на каждом шаге метода (1.2). Тогда если дифференциальная задача имеет большую размерность, то затраты на решение соответствующих нелинейных уравнений могут быть очень значительными, так как нелинейная система (1.2б) в общем виде имеет размер $l \times m$, где l – число стадий в РК-формуле (1.2), а m – размер исходной дифференциальной задачи (1.1). Заметим, что возникающие нелинейные задачи требуют применения того или иного итерационного метода ньютоновского типа, а значит, приводят к большим линейным системам, решение которых по тем или иным причинам может быть неприемлемо на практике. В связи с этим и возникает необходимость в оптимизации решения таких линейных задач с целью максимального сокращения вычислительных издержек.

Подход, предложенный в [27], [28], приводит к замене решения одной большой линейной системы размера $l \times m$ на последовательное решение l линейных задач размера m , что, очевидно, менее затратно. На этом принципе основано большое число эффективных неявных формул Рунге–Кутты, которые условно можно разделить на три группы: 1) однократно диагонально-неявные методы Рунге–Кутты, т.е. ОДНРК-методы (Singly diagonally implicit Runge–Kutta methods), которые требуют, чтобы матрица коэффициентов A была нижней треугольной с ненулевыми, но одинаковыми элементами на главной диагонали (сюда же относятся и ОДНРК-формулы с первой явной стадией); 2) однократно неявные методы Рунге–Кутты (Singly implicit Runge–Kutta methods), или ОНРК-методы, которые подразумевают, что их матрица коэффициентов A имеет единственное простое вещественное собственное значение, и 3) моонеявные методы Рунге–Кутты (Mono-implicit Runge–Kutta methods), или МНРК-методы, главное достоинство которых состоит в том, что все их стадийные величины (1.2б) могут быть последовательно подставлены в формулу (1.2в) продвижения на шаг. Последнее означает, что размер нелинейной задачи, решаемой на каждом шаге МНРК-метода, совпадает с размером дифференциального уравнения (1.1) вне зави-

симости от числа стадий и порядка такого метода. Наиболее серьезным недостатком МНРК-формулы является тот факт, что их матрица Якоби, необходимость вычисления которой диктуется применением метода Ньютона в силу неявности этих РК-схем, представляет собой матричный многочлен, зависящий от матрицы Якоби правой части задачи Коши для системы (1.1). Степень этого многочлена определяется и не превосходит числа стадий в МНРК-формуле (см. процитированные ниже статьи Ката и Сингхал). Таким образом, полное вычисление матрицы Якоби МНРК-методов может быть очень затратным на практике, сводящим на нет все преимущества этого класса неявных РК-формул. Поэтому эффективное использование МНРК-схем возможно только на базе упрощенных итераций Ньютона со специально подобранной матрицей для аппроксимации упомянутой выше матрицы Якоби численного метода, примененного для дискретизации дифференциальных уравнений (1.1). По этой причине эффективной реализации МНРК-формулы уделено пристальное внимание в мировой литературе. Уже в [29]–[31] показано, что удачную аппроксимацию для матрицы Якоби МНРК-методов следует искать в виде степенной функции от некоторой подходящей матрицы. Понятно, что в этом случае затраты на вычисление матрицы Якоби сравнимы с затратами на вычисление такой матрицы для правой части системы дифференциальных уравнений (1.1), а одна итерация упрощенного метода Ньютона заключается в последовательном решении линейных систем с одной и той же матрицей коэффициентов, размер которой совпадает с размером исходной задачи (1.1).

По каждому из перечисленных выше направлений исследований известно большое число интересных статей, среди которых следует отметить [29]–[56]. Мы не будем подробно останавливаться на анализе этих работ, так как подобный анализ с перечислением достоинств этих групп методов и присущих им недостатков выполнен, например, в [18], [21], [26], [53]–[55]. В конце этого раздела укажем только на некоторые новые идеи, которые появились сравнительно недавно и имеют непосредственное отношение к эффективному решению жестких дифференциальных уравнений вида (1.1).

В основу важного направления в построении эффективных численных методов для решения жестких задач Коши для системы (1.1) большой размерности легла идея, заключающаяся в применении явных вычислительных формул с расширенными областями устойчивости. Этот подход был предложен В.И. Лебедевым (см. [18]) и в настоящее время приобрел достаточно большую популярность (см. [22], [57]–[68]). Его суть состоит в том, что для некоторых практических жестких дифференциальных уравнений вида (1.1) условие A -устойчивости применяемых численных методов, которое и приводит к их неявности, является избыточным и может быть с успехом заменено на более слабое условие спектральной устойчивости. Методы с таким свойством, как показывают упомянутые работы, можно построить и в классе явных схем, что позволяет полностью избежать решения линейных систем большой размерности и приводит к дешевым численным методам для жестких задач Коши. С другой стороны, недостатком этого подхода является тот факт, что спектральную устойчивость можно обеспечить для дифференциальных уравнений, спектр которых априорно известен и фиксирован. Этому условию удовлетворяют некоторые классы дифференциальных уравнений в частных производных при их дискретизации методом прямых (см. указанную выше литературу). Однако для нелинейных задач Коши общего вида (1.1), спектр которых в процессе численного интегрирования может меняться произвольным образом, спектрально устойчивые численные методы не эффективны.

Еще одним направлением в построении эффективных решателей для жестких дифференциальных уравнений могут служить так называемые ABC -схемы. Они предложены и исследованы впервые в [69]. Такие методы пригодны для решения произвольных жестких задач Коши для системы (1.1) в силу своей A - и L -устойчивости. Однако их относительным недостатком является достаточно низкий второй порядок сходимости и тот факт, что матрица коэффициентов возникающих линейных задач представляет собой квадратичную функцию от матрицы Якоби исходных дифференциальных уравнений. Авторы [69] правильно замечают, что за счет факторизации этой квадратичной функции нужно отказаться от возведения соответствующих матриц в квадрат, повысив тем самым эффективность вычислений. Но они зачем-то приписывают эту идею Хайреру, когда та же самая факторизация (и даже более дешевая аппроксимация степенью некоторой матрицы) уже использовалась значительно раньше при реализации МНРК-формулы в [29]–[31].

Итак, с учетом общего характера задачи Коши для системы (1.1), анализа, проведенного в [53]–[55], а также всего сказанного выше мы приходим к выводу, что наиболее перспективный на данный момент подход к решению жестких задач Коши большой размерности лежит в классе МНРК-методов. Здесь мы уделим особое внимание специальному семейству таких формул Рунге–Кутты, которые называются *неявными гнездовыми*, или, сокращенно, *НГРК-методами* (Nested implicit Runge–Kutta methods). Они впервые были предложены в [50]–[54]. В следующем разделе

дадим строгое определение НГРК-методов и детальное описание их основных недостатков и достоинств.

2. НЕЯВНЫЕ ГНЕЗДОВЫЕ МЕТОДЫ РУНГЕ–КУТТЫ

По определению (см. [53], [54] или [55]), *l*-уровневым неявным гнездовым методом Рунге–Кутты (НГРК-методом), примененным для решения задачи Коши для системы (1.1), называется вычислительная формула следующего вида:

$$x_{kj}^2 = a_{j1}^2 x_k + a_{j2}^2 x_{k+1} + \tau_k (d_{j1}^2 g(t_k, x_k) + d_{j2}^2 g(t_{k+1}, x_{k+1})), \quad j = 1, 2, \quad (2.1a)$$

$$x_{kj}^i = a_{j1}^i x_k + a_{j2}^i x_{k+1} + \tau_k (d_{j1}^i g(t_k, x_k) + d_{j2}^i g(t_{k+1}, x_{k+1})) + \tau_k \sum_{m=1}^{i-1} d_{j,m+2}^i g(t_{km}^{i-1}, x_{km}^{i-1}), \quad i = 3, 4, \dots, l, \quad j = 1, 2, \dots, i, \quad (2.1б)$$

$$x_{k+1} = x_k + \tau_k \sum_{i=1}^l b_i g(t_{ki}^l, x_{ki}^l), \quad (2.1в)$$

где $x_0 = x^0$, $t_{kj}^i = t_k + \tau_k c_j^i$, а τ_k – размер $k+1$ -го шага численного интегрирования, который может меняться в процессе решения задачи (1.1) или оставаться постоянным.

Заметим, что формула (2.1б) предполагает, что число стадийных величин, используемых на каждом последующем уровне l , растет на единицу. На самом деле, это не является обязательным требованием и количество стадийных величин на каждом уровне может быть произвольным и выбирается в зависимости от тех или иных обстоятельств. Однако все НГРК-методы, построенные в разд. 3 и 4 настоящей статьи, подчиняются упомянутому требованию и полностью соответствуют формулам (2.1). Ниже дополнительно полагаем, что коэффициенты a_{j1}^i , a_{j2}^i , c_j^i и d_{jm}^i НГРК-метода (2.1) удовлетворяют условию

$$a_{j1}^i + a_{j2}^i = 1, \quad c_j^i = a_{j2}^i + \sum_{m=1}^{i-1} d_{jm}^i, \quad i = 2, 3, \dots, l, \quad j = 1, 2, \dots, i. \quad (2.2)$$

В упомянутых выше статьях показано, что метод (2.1) с коэффициентами, удовлетворяющими условию (2.2), относится к классу неявных РК-формул с таблицей Бутчера вида

$$\begin{array}{c|cccccc} 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ c^2 & d_1^2 & 0 & 0 & \dots & 0 & a^2 b^T & d_2^2 \\ c^3 & d_1^3 & D^3 & 0 & \dots & 0 & a^3 b^T & d_2^3 \\ c^4 & d_1^4 & 0 & D^4 & \dots & 0 & a^4 b^T & d_2^4 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ c^l & d_1^l & 0 & 0 & \dots & D^l & a^l b^T & d_2^L \\ 1 & 0 & 0 & 0 & \dots & 0 & b^T & 0 \\ \hline & 0 & 0 & 0 & \dots & 0 & b^T & 0 \end{array}, \quad (2.3)$$

где использованы следующие обозначения:

$$D^i = \begin{pmatrix} d_{13}^i & \dots & d_{1,i+1}^i \\ \vdots & \ddots & \vdots \\ d_{i3}^i & \dots & d_{i,i+1}^i \end{pmatrix}, \quad d_1^i = \begin{pmatrix} d_{11}^i \\ \vdots \\ d_{i1}^i \end{pmatrix}, \quad d_2^i = \begin{pmatrix} d_{12}^i \\ \vdots \\ d_{i2}^i \end{pmatrix},$$

$$a^i = \begin{pmatrix} a_{12}^i \\ \vdots \\ a_{i2}^i \end{pmatrix}, \quad c^i = \begin{pmatrix} c_1^i \\ \vdots \\ c_i^i \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_l \end{pmatrix},$$

Кстати, из таблицы (2.3) видно, что матрица коэффициентов A имеет почти нижний треугольный вид и все стадийные величины (2.1а), (2.1б) могут быть последовательно подставлены в формулу (2.1в) продвижения на шаг. Отсюда вытекает основное достоинство НГРК-методов — размерности исходной дифференциальной системы (1.1) и ее дискретной аппроксимации (2.1) совпадают вне зависимости от числа стадий и порядка НГРК-метода. Следовательно, методы (2.1) могут быть очень эффективны при решении задач Коши для систем большой размерности на практике.

Дополнительно в [50]–[54] и ниже показано, что наряду с отмеченным выше основным преимуществом НГРК-методы могут наследовать все наиболее важные для практического применения свойства неявных РК-формул, такие как A -устойчивость, жесткая точность, симметрия, симплектичность в определенном смысле. Кроме того, они могут иметь достаточно высокие стадийные и классические порядки, а также обеспечивать плотную выдачу результатов интегрирования той же точности, что и порядок базового метода, без больших дополнительных затрат. Последнее свойство объясняется возможностью построения НГРК-методов со стадийными величинами l -го уровня, аппроксимирующими точное решение с тем же порядком (или на единицу ниже), что и численное решение x_{k+1} , получаемое на выходе. Таким образом, на $k+1$ -м шаге такого метода мы можем использовать численное решение x_k , x_{k+1} , а также стадийные величины l -го уровня (или их производные), чтобы построить интерполяционный многочлен Эрмита, аппроксимирующий точное решение на локальном интервале $[t_k, t_{k+1}]$ с требуемой точностью. Другими словами, в [53] объяснено, что процедура построения всех НГРК-методов заключается в следующем. Во-первых, *стадийными величинами первого уровня* мы называем численное решение x_k и x_{k+1} в двух соседних точках сетки. Во-вторых, проводя интерполяционный многочлен Эрмита третьей степени через стадийные величины первого уровня и их производные, мы можем вычислить стадийные величины второго уровня с третьим порядком точности. Затем, используя стадийные величины первого и второго уровня, построим интерполяционный многочлен Эрмита более высокой степени и вычислим стадийные величины следующего уровня с более высокой точностью и т.д. В итоге мы всегда будем иметь интерполяционный многочлен, пригодный для плотной выдачи результатов численного интегрирования, практически бесплатно.

Итак, гнездовые методы (2.1) могут быть рекомендованы для численного интегрирования дифференциальных уравнений разных видов, к которым относятся жесткие и нежесткие задачи, а также гамильтоновы системы и обратимые уравнения. Такие методы особенно эффективны при решении задач большой размерности, когда применение неявных РК-формул общего вида ограничено из-за их трудоемкости, в том числе и для дифференциальных уравнений, возникающих при использовании метода прямых для дискретизации задач математической физики. К тому же дифференциальные системы, получаемые с помощью метода прямых, не только имеют большую размерность, но и зачастую относятся к классу очень жестких задач, для численного решения которых требуются РК-методы с высоким стадийным порядком (см., например, [18], [21], [26]). В связи с последним обстоятельством отметим, что все НГРК-методы (2.1) в силу формы их таблицы Бутчера являются жестко точными, а значит, они могут применяться для решения очень жестких дифференциальных уравнений, а также и при интегрировании дифференциально-алгебраических систем индекса 1 и выше (см. [53]).

Однако при реализации НГРК-методов, впрочем как и любых МНРК-формул, особое внимание следует уделить итерационному процессу для решения возникающих нелинейных задач из-за значительных вычислительных затрат на организацию полных итераций Ньютона. Здесь мы воспользуемся идеей достаточно точной аппроксимации матрицы Якоби численного метода, которая, как сказано выше, представляет собой матричный многочлен, зависящий от матрицы Якоби правой части задачи Коши для системы (1.1), степенью некоторой специально подобранной матрицы, сохраняющей к тому же устойчивость исходного НГРК-метода (см. подробности в [50]–[54]). В следующих разделах мы представим общий алгоритм построения НГРК-методов и приведем примеры таких вычислительных схем, основанных на квадратурных формулах Гаусса и Лобатто. С учетом неявного характера всех построенных методов особое внимание уделим их эффективной итерационной реализации, а также предложим механизмы для автоматического

управления размером шага интегрирования на основе контроля локальной ошибки численного решения. Управление глобальной ошибкой в НГРК-методах тоже возможно (см. [55], [56]), но выходит за рамки настоящей статьи. Итак, начнем прежде всего с НГРК-методов гауссовского типа.

3. НГРК-МЕТОДЫ ТИПА ГАУССА

Очевидным примером гауссовского НГРК-метода второго порядка является известное неявное правило средней точки (см. [53, разд. 3]). Поэтому в данном разделе ограничимся методикой построения численных методов вида (2.1), основанных на квадратурах Гаусса 4-го и 6-го порядков, а также их анализом для прикладных вычислений. Заметим, что такие формулы уже были представлены ранее в [50]–[54], но сейчас мы опишем единый принцип их построения на основе метода неизвестных коэффициентов, который затем распространим на формулы типа Лобатто. Дополнительно объясним, что НГРК-методы типа Гаусса более высокого порядка нельзя построить в рамках примененной методики.

3.1. Двухуровневый метод порядка 4

Принимая во внимание форму таблицы Бутчера (2.3) для любого НГРК-метода (2.1), двухуровневый гауссовский метод четвертого порядка будем искать в виде

$$\begin{array}{c|cccc} 0 & 0 & 0 & 0 & 0 \\ c_2 & a_{21} & a_{22} & a_{23} & a_{24} \\ c_3 & a_{31} & a_{32} & a_{33} & a_{34} \\ 1 & 0 & a_{42} & a_{43} & 0 \\ \hline & 0 & b_2 & b_3 & 0 \end{array}, \quad (3.1)$$

где $a_{4j} = b_j, j = 2, 3$, в силу требования жесткой точности всех НГРК-методов, c_2 и c_3 — нули смешенного многочлена Лежандра $L_2(t) = \frac{d^2}{dt^2}(t^2(1-t)^2)$, а b_2 и b_3 — веса квадратурной формулы Гаусса четвертого порядка. Очевидно, что в силу выбора абсцисс c_i и весов b_i РК-формула (3.1) удовлетворяет упрощающему условию $\mathcal{B}(4)$ (более подробную информацию относительно роли упрощающих условий при построении неявных РК-методов можно найти, например, в [18], [21] и [26]).

Поскольку мы строим численный метод вида (2.1), потребуем, чтобы были выполнены соотношения

$$a_{ij} = \tilde{a}_i b_j, \quad i = 2, 3, \quad j = 2, 3. \quad (3.2)$$

Оставшиеся свободные коэффициенты для второй и третьей стадии в таблице (3.1) определим из упрощающего условия $\mathcal{C}(3)$. С учетом условия $\mathcal{B}(4)$ и формул (3.2) эти коэффициенты должны удовлетворять линейной системе

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1/2 & 1 \\ 0 & 1/3 & 1 \end{pmatrix} \begin{pmatrix} a_{i1} \\ \tilde{a}_i \\ a_{i4} \end{pmatrix} = \begin{pmatrix} c_i \\ c_i^2/2 \\ c_i^3/3 \end{pmatrix}, \quad i = 2, 3. \quad (3.3)$$

Матрица задачи (3.3) невырождена. Следовательно, неизвестные коэффициенты $a_{i1}, \tilde{a}_i, a_{i4}, i = 2, 3$, определяются однозначно.

В итоге коэффициенты метода (3.1) удовлетворяют упрощающим условиям $\mathcal{C}(3)$ и $\mathcal{B}(4)$, а значит, из [20, теорема 7.4] вытекает, что гауссовский НГРК-метод вида (3.1) имеет 4-ый порядок сходимости. Таблица Бутчера этого метода представима в виде

0	0	0	0	0
c_2	$\frac{6(c_2 + \theta) - 5}{12}$	$\frac{1 - \theta}{2}$	$\frac{1 - \theta}{2}$	$\frac{6(c_2 + \theta) - 7}{12}$
c_3	$\frac{7 - 6(c_2 + \theta)}{12}$	$\frac{\theta}{2}$	$\frac{\theta}{2}$	$\frac{5 - 6(c_2 + \theta)}{12}$
1	0	$\frac{1}{2}$	$\frac{1}{2}$	0
	0	$\frac{1}{2}$	$\frac{1}{2}$	0

где $c_2 = (3 - \sqrt{3})/6$, $c_3 = (3 + \sqrt{3})/6$ и $\theta = 1/2 + 2\sqrt{3}/9$. Сделаем некоторые замечания относительно свойств полученного численного метода и особенностей его практической реализации.

Замечание 1. Легко проверить, что метод (3.1) симметричен (это следует из [20, теорема 8.7]), что, кстати, обусловило его эффективность при решении краевых задач из гидромеханики в [7]. Кроме того, он может быть полезен для решения обратимых дифференциальных уравнений произвольного вида (1.1).

Замечание 2. Нетрудно показать, что метод (3.1) сопряжен с симплектическим методом по крайней мере до 6-го порядка включительно (доказывается прямой подстановкой коэффициентов этого метода в условия (21)–(23) в [53, лемма 2]). Поэтому он может быть использован при интегрировании гамильтоновых систем на больших интервалах времени, что и подтверждается численными экспериментами на задаче Кеплера в [52].

Замечание 3. Параметр θ можно считать свободным в приведенной выше таблице Бутчера, т.е. при любом его действительном значении мы получаем численный метод 4-го порядка сходимости, но его стадийный порядок равен трем только для $\theta = 1/2 + 2\sqrt{3}/9$. При любых других значениях θ стадийный порядок будет равен двум, как и в стандартном методе Хаммера–Холлингсуорта четвертого порядка из [20, таблица 7.3].

Замечание 4. Функция устойчивости метода (3.1) представляет собой диагональную аппроксимацию Паде $R_{22}(z)$, а сам метод A -устойчив при любом θ . Следовательно, построенное семейство численных методов применимо для решения жестких задач. Более того, все эти формулы жестко точны, а значит — эффективны для решения очень жестких дифференциальных уравнений и дифференциально-алгебраических систем. Все это подтверждается численными экспериментами, проведенными в [52] и [53].

Замечание 5. Известно, что в случае жестких дифференциальных уравнений вида (1.1) для решения возникающих нелинейных задач требуются итерационные методы ньютоновского типа. Последнее означает, что для реализации построенного гауссовского НГРК-метода четвертого порядка будут эффективны следующие итерации:

$$x_{k1}^{2, \ell-1} = a_{11}^2 \bar{x}_k + a_{12}^2 x_{k+1}^{\ell-1} + \tau_k [d_{11}^2 g(t_k, \bar{x}_k) + d_{12}^2 g(t_{k+1}, x_{k+1}^{\ell-1})], \quad (3.4a)$$

$$x_{k2}^{2, \ell-1} = a_{21}^2 \bar{x}_k + a_{22}^2 x_{k+1}^{\ell-1} + \tau_k [d_{21}^2 g(t_k, \bar{x}_k) + d_{22}^2 g(t_{k+1}, x_{k+1}^{\ell-1})], \quad (3.4б)$$

$$Q_2(\tau_k J)(x_{k+1}^{\ell} - x_{k+1}^{\ell-1}) = -x_{k+1}^{\ell-1} + \bar{x}_k + \tau_k [b_1 g(t_k + c_1^2 \tau_k, x_{k1}^{2, \ell-1}) + b_2 g(t_k + c_2^2 \tau_k, x_{k2}^{2, \ell-1})], \quad (3.4в)$$

$\ell = 1, 2, \dots, N$, с начальным приближением $x_{k+1}^0 = \bar{x}_k$, где \bar{x}_k означает численное решение задачи (1.1), полученное в точке t_k за N итерационных шагов комбинированного метода (3.4), т.е. $\bar{x}_k = x_k^N$, $k = 1, 2, \dots, K$; $J = \partial_x g(t_{k+1}, x_{k+1}^0)$ есть матрица Якоби правой части дифференциального уравнения (1.1), вычисленная в точке сетки t_{k+1} для начального приближения x_{k+1}^0 ; а $Q_2(z) = (E - z/4)^2$ — матричный многочлен второй степени с единичной матрицей E размера m . Числовые значения коэффициентов комбинированного НГРК-метода (3.4) легко найти из таблицы Бутчера, приведенной выше, для $\theta = 1/2 + 2\sqrt{3}/9$, т.е. $b_1 = b_2 = 1/2$, $c_1^2 = (3 - \sqrt{3})/6$, $c_2^2 = (3 + \sqrt{3})/6$, $a_{11}^2 = a_{22}^2 = 1/2 + 2\sqrt{3}/9$, $a_{12}^2 = a_{21}^2 = 1/2 - 2\sqrt{3}/9$, $d_{11}^2 = -d_{22}^2 = (3 + \sqrt{3})/36$, $d_{12}^2 = -d_{21}^2 = (-3 + \sqrt{3})/36$.

Заметим, что в настоящей работе итерации (3.4) применяются с тривиальным предиктором, т.е. решение, вычисленное в предыдущей точке сетки, берется в качестве начального приближения для упрощенного метода Ньютона. В этом случае два итерационных шага комбинированного метода (3.4) в каждой точке сетки, т.е. $N = 2$, достаточны, чтобы обеспечить сходимость четвертого порядка для численного решения \bar{x}_k (см. [51]). Полученные в результате итераций (3.4) величины $\bar{x}_{k+1} = x_{k+1}^N, x_{k1}^{2,N}, x_{k2}^{2,N}$ (подчеркнем, что формулы для вычисления явных стадийных величин (3.4а) и (3.4б) применяются еще один раз, чтобы найти их значения, соответствующие численному решению \bar{x}_{k+1}) и решение \bar{x}_k с предыдущего шага могут быть использованы не только для построения интерполяционного многочлена Эрмита для плотной выдачи результатов численного интегрирования тоже четвертого порядка, но и для предсказания начального приближения x_{k+2}^0 для итераций (3.4) в новой точке сетки t_{k+2} . Тогда, как следует из [51, теорема 2.3], одной итерации в комбинированном методе (3.4), т.е. $N = 1$, будет достаточно для сходимости максимального четвертого порядка.

Отдельно отметим, что замена точной матрицы Якоби на ее аппроксимацию многочленом $Q_2(z) = (E - z/4)^2$ не нарушает A -устойчивости исходного двухуровневого НГРК-метода четвертого порядка (см. указанные выше работы). Подчеркнем также, что решение линейной задачи (3.4в) с матрицей коэффициентов $Q_2(\tau_k J)$ осуществляется очень эффективно в силу специального вида матрицы коэффициентов, т.е. каждый итерационный шаг метода (3.4) состоит в последовательном решении двух линейных систем уравнений с одной и той же матрицей коэффициентов $E - \tau_k J/4$.

Замечание 6. Известно, что наиболее предпочтительный способ эффективной оценки локальной ошибки методов Рунге–Кутты заключается в нахождении вложенных формул порядка, отличного от порядка сходимости исходного РК-метода (см., например, [18], [20], [26]). Для двухуровневого гауссовского НГРК-метода четвертого порядка, построенного выше, вложенную формулу будем искать в следующем общем виде:

$$\begin{array}{c|cccc} 0 & 0 & 0 & 0 & 0 \\ c_2 & \frac{6(c_2 + \theta) - 5}{12} & \frac{1 - \theta}{2} & \frac{1 - \theta}{2} & \frac{6(c_2 + \theta) - 7}{12} \\ c_3 & \frac{7 - 6(c_2 + \theta)}{12} & \frac{\theta}{2} & \frac{\theta}{2} & \frac{5 - 6(c_2 + \theta)}{12} \\ 1 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \hline & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \hline & e_1 & e_2 & e_3 & e_4 \end{array}.$$

Тогда для оценки локальной ошибки получаем

$$le_{k+1} = \tau_k [e_1 g(t_k, \bar{x}_k) + e_2 g(t_{k1}^2, x_{k1}^{2,N}) + e_3 g(t_{k2}^2, x_{k2}^{2,N}) + e_4 g(t_{k+1}, \bar{x}_{k+1})], \quad (3.5)$$

где $t_{kj}^2 = t_k + \tau_k c_j^2, j = 1, 2$ (верхние индексы означают номер уровня стадийной величины и количество итерационных шагов в методе (3.4)), а $e_i, i = 1, 2, 3, 4$, — некоторые действительные числа.

Несложно показать, что при любом выборе коэффициентов $e_i, i = 1, 2, 3, 4$, вложенное численное решение не будет A -устойчивым. В частности, если в качестве вложенного метода использовать правило трапечий, т.е. положить $e_1 = e_4 = 1/2$ и $e_2 = e_3 = -1/2$, то его функция устойчивости, которая получается в результате применения этого метода к линейному скалярному уравнению Далквиста (см. [70], [71])

$$x'(t) = \lambda x(t), \quad t \geq 0, \quad x(0) = x^0,$$

где константа λ принадлежит множеству комплексных чисел, будет иметь вид

$$R_{TR}(z) = \frac{1 + 1/2z + 1/12z^2 + 1/12z^3}{1 - 1/2z + 1/12z^2}$$

при $z = \lambda \tau_k$. Понятно, что для очень жестких задач такая формула (3.5) может существенно переоценивать ошибку, приводя к слишком малым шагам и, следовательно, к избыточным вычислительным издержкам. Чтобы избежать такой переоценки, при решении жестких дифференциальных уравнений вида (1.1) реко-

мендуется заменить оценку локальной ошибки le_{k+1} на величину, получаемую в результате решения линейной системы

$$(E - \tau_k J/4)^3 \tilde{le}_{k+1} = le_{k+1}. \quad (3.6)$$

Важно понимать, что решение задачи (3.6) лишь незначительно удорожает процедуру численного интегрирования методом (3.4), так как и в том, и в другом случае решаются линейные системы с одной и той же матрицей коэффициентов $E - \tau_k J/4$.

В то же время функция устойчивости, соответствующая модифицированной оценке локальной ошибки \tilde{le}_{k+1} , есть

$$\tilde{R}_{TR}(z) = \frac{1 - 1/4z - 5/48z^2 - 13/192z^3 + 1/128z^4 - 1/768z^5}{1 - 5/4z + 31/48z^2 - 11/192z^3 + 3/128z^4 - 1/768z^5}.$$

Из последней формулы следует, что функция устойчивости $\tilde{R}_{TR}(z)$ A -устойчива, а значит, оценка ошибки по формулам (3.5), (3.6) не будет существенно переоценивать локальную погрешность даже для очень жестких задач Коши для систем (1.1).

Кроме того, важным новшеством всех современных программ для численного интегрирования задач Коши для систем (1.1), включая известные решатели MATLAB, является использование масштабированного контроля локальной погрешности. В этом случае модифицированная оценка локальной ошибки \tilde{le}_{k+1} (или ее исходная оценка le_{k+1} для нежестких дифференциальных уравнений) масштабируется по следующей формуле:

$$|\tilde{le}_{k+1}|_{sc} = \max_{i=1,2,\dots,m} \left\{ \frac{|\tilde{le}_{k+1}^i|}{\text{atol} + \text{rtol} \cdot |x_{k+1}^i|} \right\},$$

где верхний индекс i означает i -ю компоненту соответствующего вектора, x_{k+i} — численное решение, полученное НГРК-методом (3.4) в новой точке сетки t_{k+1} , а числовые параметры atol и rtol определяют вклад абсолютной и относительной локальных ошибок в масштабированную оценку. В настоящей статье мы применяем вариант масштабированного контроля с условием $\text{rtol} = \text{atol} = \text{Tol}$, где Tol — устанавливаемый пользователем допуск на масштабированную локальную ошибку.

После этого оценка ошибки $|\tilde{le}_{k+1}|_{sc}$ проверяется на соответствие установленной точности численного интегрирования по формуле

$$|\tilde{le}_{k+1}|_{sc} \leq 1. \quad (3.7)$$

Если неравенство (3.7) не выполнено, то считается, что требуемая точность численного интегрирования не достигнута. Поэтому найденное численное решение x_{k+1} отбраковывается и перевычисляется для нового размера шага

$$\tau_k^* := \min \{ 1.5, 0.8(1/|\tilde{le}_{k+1}|_{sc})^{1/s+1} \} \cdot \tau_k,$$

где s означает порядок сходимости вложенной РК-формулы, т.е. $s = 2$ для вложенной НГРК-пары гауссовского типа, рассмотренной в настоящем разделе (в НГРК-методах, представленных ниже, порядок s меняется в зависимости от точности вложенных формул в пределах от 2 до 4). Затем модифицированная локальная ошибка \tilde{le}_{k+1} пересчитывается для нового численного решения по формулам (3.5) и (3.6). Если же неравенство (3.7) выполнено для текущего (или сокращенного) размера шага, то численное решение x_{k+1} считается достаточно точным и принимается как окончательное в точке сетки t_{k+1} , а размер для следующего шага устанавливается на основе сделанного прогноза для оптимального шага, т.е. по формуле $\tau_{k+1} = \tau_k^*$.

Заметим, что контроль точности, изложенный в замечании 6, автоматически увеличивает размер следующего шага численного интегрирования, когда оценка локальной ошибки $|\tilde{le}_{k+1}|_{sc}$ существенно меньше единицы. С другой стороны, формула для вычисления τ_k^* говорит, что коэффициент удлинения шага не может превышать 1.5.

Описанный масштабированный контроль точности численного интегрирования является стандартной опцией практически всех современных численных методов с переменным шагом

интегрирования (см., например, [20]). Он не только значительно повышает эффективность численного решения задач Коши для систем (1.1), но также обезразмеривает компоненты локальной ошибки и тем самым исключает сравнение разных физических единиц, что важно при решении прикладных задач. Поэтому все представленные в статье НГРК-методы с переменным шагом используют именно масштабированный контроль локальной ошибки для управления точностью численного решения дифференциальных уравнений (1.1).

3.2. Трехуровневый метод порядка 6

Опять же принимая во внимание форму таблицы Бутчера (2.3), трехуровневый гауссовский НГРК-метод шестого порядка будем искать в виде

$$\begin{array}{c|cccccccc}
 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 c_2 & a_{21} & 0 & 0 & a_{24} & a_{25} & a_{26} & a_{27} & \\
 c_3 & a_{31} & 0 & 0 & a_{34} & a_{35} & a_{36} & a_{37} & \\
 c_4 & a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} & a_{47} & \\
 c_5 & a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} & a_{57} & \\
 c_6 & a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} & a_{67} & \\
 1 & 0 & 0 & 0 & a_{74} & a_{75} & a_{76} & 0 & \\
 \hline
 & 0 & 0 & 0 & b_4 & b_5 & b_6 & 0 &
 \end{array} \quad (3.8)$$

и потребуем, чтобы коэффициенты этого метода (3.8) удовлетворяли следующим условиям:

1) $a_{7j} = b_j, j = 4, 5, 6$;

2) c_2 и c_3 — нули смещенного многочлена Лежандра $L_2(t) = \frac{d^2}{dt^2}(t^2(1-t)^2)$;

3) c_4, c_5 и c_6 — нули смещенного многочлена Лежандра $L_3(t) = \frac{d^3}{dt^3}(t^3(1-t)^3)$;

4) b_4, b_5 и b_6 — веса квадратурной формулы Гаусса шестого порядка.

Очевидно, что последние два пункта в приведенном списке требований гарантируют выполнение упрощающего условия $\mathcal{B}(6)$.

Как и для двухуровневого гауссовского НГРК-метода четвертого порядка, наложим на коэффициенты в таблице Бутчера (3.8) дополнительные условия

$$a_{ij} = \tilde{a}_i b_j, \quad i = 2, \dots, 6, \quad j = 4, 5, 6. \quad (3.9)$$

Оставшиеся свободные коэффициенты во второй и третьей строках определим из упрощающего условия $\mathcal{C}(3)$, а в строках с номером $i = 4, 5, 6$ — из упрощающего условия $\mathcal{C}(5)$ с учетом формул (3.9). При этом коэффициенты a_{i1} , \tilde{a}_i и a_{i7} , $i = 2, 3$, должны удовлетворять уравнениям (3.3), а остальные легко найти, решив линейную систему

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & c_2 & c_3 & 1/2 & 1 \\ 0 & c_2^2 & c_3^2 & 1/3 & 1 \\ 0 & c_2^3 & c_3^3 & 1/4 & 1 \\ 0 & c_2^4 & c_3^3 & 1/5 & 1 \end{pmatrix} \begin{pmatrix} a_{i1} \\ a_{i2} \\ a_{i3} \\ \tilde{a}_i \\ a_{i7} \end{pmatrix} = \begin{pmatrix} c_i \\ c_i^2/2 \\ c_i^3/3 \\ c_i^4/4 \\ c_i^5/5 \end{pmatrix}, \quad i = 4, 5, 6. \quad (3.10)$$

Матрицы коэффициентов линейных задач (3.3) и (3.10) невырождены, поэтому искомый численный метод определяется единственным образом. В результате таблица Бутчера трехуровневого гауссовского НГРК-метода примет вид

0	0	0	0	0	0	0	0
c_2	$\frac{c_3}{6}$	0	0	$\frac{35-40c_3}{108}$	$\frac{56-64c_3}{108}$	$\frac{35-40c_3}{108}$	$\frac{c_3-1}{6}$
c_3	$\frac{c_2}{6}$	0	0	$\frac{35-40c_2}{108}$	$\frac{56-64c_2}{108}$	$\frac{35-40c_2}{108}$	$\frac{c_2-1}{6}$
c_4	$\frac{20c_6-3}{200}$	$\frac{18c_3-9}{100} + \theta$	θ	$\frac{v_1}{360} - \frac{5\theta}{9}$	$\frac{v_1}{225} - \frac{8\theta}{9}$	$\frac{v_1}{360} - \frac{5\theta}{9}$	$\frac{3-20c_4}{200}$
c_5	$\frac{1}{32}$	$\frac{\sqrt{27}}{32}$	$-\frac{\sqrt{27}}{32}$	$\frac{5}{36}$	$\frac{2}{9}$	$\frac{5}{36}$	$-\frac{1}{32}$
c_6	$\frac{20c_4-3}{200}$	$-\theta$	$\frac{18c_2-9}{100} - \theta$	$\frac{v_2}{360} - \frac{5\theta}{9}$	$\frac{v_2}{225} - \frac{8\theta}{9}$	$\frac{v_2}{360} - \frac{5\theta}{9}$	$\frac{3-20c_4}{200}$
1	0	0	0	$\frac{5}{18}$	$\frac{4}{9}$	$\frac{5}{18}$	0
	0	0	0	$\frac{5}{18}$	$\frac{4}{9}$	$\frac{5}{18}$	0

где $c_2 = (3 - \sqrt{3})/6$, $c_3 = (3 + \sqrt{3})/6$, $c_4 = (5 - \sqrt{15})/10$, $c_5 = 1/2$, $c_6 = (5 + \sqrt{15})/10$, $v_1 = 120c_4 - 18c_3 - 1$, $v_2 = 120c_6 - 18c_2 - 1$ и $\theta = (36c_3 - 18c_2 - 9)/200$.

Сделаем некоторые замечания относительно свойств полученного численного метода и особенностей его практической реализации.

Замечание 7. Упорядочиванием строк в таблице Бутчера для трехуровневого гауссовского НГРК-метода по возрастанию коэффициентов c_i , $i = 1, 2, \dots, 7$, и с использованием [20, теорема 8.7] убеждаемся, что построенный НГРК-метод симметричен, что делает его привлекательным для решения обратимых дифференциальных уравнений вида (1.1).

Замечание 8. Трехуровневый гауссовский НГРК-метод (3.8) сходится с шестым порядком и имеет третий стадийный порядок (см. строгое доказательство этого факта в [53, теорема 1]). Кстати, замечание 7 позволяет значительно упростить опубликованное ранее доказательство. Действительно, тот факт, что порядок НГРК-метода (3.8) не меньше пяти, устанавливается тривиально. Тогда, поскольку порядок симметричного метода не может быть нечетным, сразу же заключаем, что его порядок равен шести. Сходимость же более высокого порядка, очевидно, невозможна в силу отсутствия упрощающего условия $\mathcal{B}(7)$, которое не может быть удовлетворено для квадратурной формулы Гаусса шестого порядка (см. приведенный выше список требований к построенному НГРК-методу).

Замечание 9. Параметр θ можно считать свободным в приведенной выше таблице Бутчера, т.е. при любом его действительном значении мы получаем численный метод 6-го порядка сходимости, но точность его стадийных величин 3-го уровня имеет пятый порядок только для $\theta = (36c_3 - 18c_2 - 9)/200$. При любых других значениях θ точность этих стадийных величин будет на единицу ниже. В целом по своим свойствам этот метод очень похож на стандартный метод Кунцмана–Бутчера шестого порядка (см. [20, таблица 7.4]), но значительно дешевле последнего. Поэтому ожидается, что построенный трехуровневый гауссовский НГРК-метод (3.8) также применим для численного интегрирования гамильтоновых систем на больших временных интервалах, что подтверждается численными экспериментами на задаче Кеплера в [54].

Замечание 10. К сожалению, описанную выше процедуру построения гауссовских НГРК-методов 4-го и 6-го порядков нельзя продолжить дальше. Это следует из того, что для построения потенциального НГРК-метода 8-го порядка необходимо добавить еще 4 стадии, соответствующие нулям смещенного мно-

гочлена Лежандра $L_4(t) = \frac{d^4}{dt^4} (t^4(1-t)^4)$, потребовать выполнения упрощающего условия $\mathcal{B}(8)$ для коэф-

фициентов искомого метода и определить недостающие коэффициенты a_{ij} из упрощающего условия $\mathcal{C}(6)$ или даже $\mathcal{C}(7)$. Однако последнее сделать не удастся по причине вырожденности линейных систем, аналогичных (3.3) и (3.10). Если же требовать выполнения только упрощающего условия $\mathcal{C}(5)$, то нельзя удо-

влетворить условиям порядка для деревьев вида $[[[\bullet, \bullet, \bullet]]]$ (здесь и далее использованы понятия и обозначения, принятые в теории условий порядка РК-формулы, основанной на замене соответствующих частных производных корневыми деревьями; см. подробности в [20, разд. II.2] или [26, глава 3]).

Замечание 11. Функция устойчивости метода (3.8) представляет собой диагональную аппроксимацию Паде $R_{33}(z)$, а сам трехуровневый гауссовский НГРК-метод A -устойчив при любом θ . Следовательно, построенное семейство численных методов применимо для решения жестких задач. Более того, все эти формулы жестко точны, а значит — эффективны для решения очень жестких дифференциальных уравнений и дифференциально-алгебраических систем. Это подтверждается численными экспериментами, проведенными для различных классов тестовых задач в [54].

Замечание 12. Упрощенные итерации Ньютона для реализации трехуровневого гауссовского НГРК-метода шестого порядка имеют вид

$$x_{kj}^{2,\ell-1} = a_{j1}^2 \bar{x}_k + a_{j2}^2 x_{k+1}^{\ell-1} + \tau_k (d_{j1}^2 g(t_k, \bar{x}_k) + d_{j2}^2 g(t_{k+1}, x_{k+1}^{\ell-1})), \quad j = 1, 2, \quad (3.11a)$$

$$x_{kj}^{3,\ell-1} = a_{j1}^3 \bar{x}_k + a_{j2}^3 x_{k+1}^{\ell-1} + \tau_k (d_{j1}^3 g(t_k, \bar{x}_k) + d_{j2}^3 g(t_{k+1}, x_{k+1}^{\ell-1})) + \\ + \tau_k \sum_{m=1}^2 d_{j,m+2}^3 g(t_{km}^2, x_{km}^{2,\ell-1}), \quad j = 1, 2, 3, \quad (3.11b)$$

$$Q_3(\tau_k J)(x_{k+1}^\ell - x_{k+1}^{\ell-1}) = -x_{k+1}^{\ell-1} + \bar{x}_k + \tau_k \sum_{i=1}^3 b_i g(t_{ki}^3, x_{ki}^{3,\ell-1}), \quad \ell = 1, 2, \dots, N, \quad (3.11b)$$

где матричный многочлен $Q_3(z) = (E - z/6)^3$, а остальные обозначения имеют тот же самый смысл, что и в методе (3.4). Из последней таблицы Бутчера нетрудно вычислить значения постоянных коэффициентов комбинированного метода (3.11) при $\theta = (36c_3 - 18c_2 - 9)/200$, т.е. $b_1 = b_3 = 5/18$, $b_2 = 4/9$, $c_1^2 = (3 - \sqrt{3})/6$, $c_2^2 = (3 + \sqrt{3})/6$, $a_{11}^2 = a_{22}^2 = 1/2 + 2\sqrt{3}/9$, $a_{12}^2 = a_{21}^2 = 1/2 - 2\sqrt{3}/9$, $d_{11}^2 = -d_{22}^2 = (3 + \sqrt{3})/36$, $d_{12}^2 = -d_{21}^2 = (-3 + \sqrt{3})/36$, $c_1^3 = (5 - \sqrt{15})/10$, $c_2^3 = 1/2$, $c_3^3 = (5 + \sqrt{15})/10$, $a_{11}^3 = a_{32}^3 = (125 + 39\sqrt{15})/250$, $a_{12}^3 = a_{31}^3 = (125 - 39\sqrt{15})/250$, $a_{21}^3 = a_{22}^3 = 1/2$, $d_{11}^3 = -d_{32}^3 = (7 + 2\sqrt{15})/200$, $d_{12}^3 = -d_{31}^3 = (-7 + 2\sqrt{15})/200$, $d_{13}^3 = -d_{34}^3 = (18\sqrt{15} + 15\sqrt{3})/1000$, $d_{14}^3 = -d_{33}^3 = (18\sqrt{15} - 15\sqrt{3})/1000$, $d_{21}^3 = -d_{22}^3 = 1/32$, $d_{23}^3 = -d_{24}^3 = 3\sqrt{3}/32$.

Заметим, что числовые значения коэффициентов 3-го уровня d_{ij}^3 , соответствующие трехуровневому НГРК-методу шестого порядка, удовлетворяющему условию $\mathcal{C}(3, 5)$, слегка отличаются в [54] от приведенных выше. Во избежание возможной путаницы подчеркнем, что НГРК-метод с указанными выше свойствами единственен, а вся разница возникает из-за различной индексации этих коэффициентов. В настоящей работе их индексация соответствует формуле (3.11), тогда как в [54] коэффициенты d_{ij}^3 пронумерованы в соответствии с видом их таблицы Бутчера. Индексация коэффициентов, принятая в настоящей статье, представляется более естественной, так как соответствует общему определению НГРК-методов, данному формулой (2.1).

Как и для НГРК-метода (3.4) четвертого порядка, упрощенные итерации Ньютона (3.11) проводятся с тривиальным предиктором. Поэтому для достижения максимального шестого порядка сходимости будет достаточно трех итерационных шагов в каждой точке сетки. Опять же в силу условия $\mathcal{C}(3, 5)$, которое означает, что стадийные значения второго уровня вычислены с третьим порядком точности, а стадийные величины третьего уровня — с пятым, полученные в результате итераций (3.11) величины $\bar{x}_{k+1} = x_{k+1}^N$, $x_{k1}^{3,N}$, $x_{k2}^{3,N}$, $x_{k3}^{3,N}$ (подчеркнем, что формулы для вычисления явных стадийных величин (3.11a) и (3.11b) применяются еще один раз, чтобы найти их значения, соответствующие численному решению \bar{x}_{k+1}) и решение \bar{x}_k с предыдущего шага могут быть использованы не только для построения интерполяционного многочлена Эрмита для плотной выдачи результатов численного интегрирования тоже шестого порядка, но и для предсказания начального приближения x_{k+2}^0 для итераций (3.11) в новой точке сетки t_{k+2} . Тогда, как следует из [51, теорема 2.3], одной итерации в комбинированном методе (3.11), т.е. $N = 1$, будет достаточно для сходимости максимального шестого порядка.

Кроме того, важно подчеркнуть, что замена точной матрицы Якоби трехуровневого НГРК-метода гауссовского типа шестого порядка с условием $\mathcal{C}(3, 5)$ на матричный многочлен $Q_3(\tau_k J)$ обладает тем преимуществом, что подходит для решения жестких задач, с одной стороны, а с другой — допускает эффективное решение линейных систем (3.11в), т.е. их решение осуществляется через последовательное решение трех линейных задач с одной и той же матрицей коэффициентов $E - \tau_k J/6$ (см. более подробное разъяснение этих вопросов и результаты численных экспериментов в [54]).

Замечание 13. Из представленной выше процедуры построения видно, что часть коэффициентов гауссовских НГРК-методов (3.4) и (3.11) совпадает. Поэтому эти методы можно рассматривать как вложенную пару, которая имеет таблицу коэффициентов

0	0	0	0	0	0	0	0
c_2	$\frac{c_3}{6}$	0	0	$\frac{35-40c_3}{108}$	$\frac{56-64c_3}{108}$	$\frac{35-40c_3}{108}$	$\frac{c_3-1}{6}$
c_3	$\frac{c_2}{6}$	0	0	$\frac{35-40c_2}{108}$	$\frac{56-64c_2}{108}$	$\frac{35-40c_2}{108}$	$\frac{c_2-1}{6}$
c_4	$\frac{20c_6-3}{200}$	$\frac{18c_3-9}{100} + \theta$	θ	$\frac{v_1}{360} - \frac{5\theta}{9}$	$\frac{v_1}{225} - \frac{8\theta}{9}$	$\frac{v_1}{360} - \frac{5\theta}{9}$	$\frac{3-20c_4}{200}$
c_5	$\frac{1}{32}$	$\frac{\sqrt{27}}{32}$	$-\frac{\sqrt{27}}{32}$	$\frac{5}{36}$	$\frac{2}{9}$	$\frac{5}{36}$	$-\frac{1}{32}$
c_6	$\frac{20c_4-3}{200}$	$-\theta$	$\frac{18c_2-9}{100} - \theta$	$\frac{v_2}{360} - \frac{5\theta}{9}$	$\frac{v_2}{225} - \frac{8\theta}{9}$	$\frac{v_2}{360} - \frac{5\theta}{9}$	$\frac{3-20c_4}{200}$
1	0	0	0	$\frac{5}{18}$	$\frac{4}{9}$	$\frac{5}{18}$	0
	0	0	0	$\frac{5}{18}$	$\frac{4}{9}$	$\frac{5}{18}$	0
	e_1	e_2	e_3	e_4	e_5	e_6	e_7

Тогда для оценки локальной ошибки получаем очевидную формулу

$$le_{k+1} = \tau_k \left[\frac{1}{2} g(t_{k1}^2, x_{k1}^{2,N}) + \frac{1}{2} g(t_{k2}^2, x_{k2}^{2,N}) - \frac{5}{18} g(t_{k1}^3, x_{k1}^{3,N}) - \frac{4}{9} g(t_{k2}^3, x_{k2}^{3,N}) - \frac{5}{18} g(t_{k3}^3, x_{k3}^{3,N}) \right]. \quad (3.12)$$

Другими словами, формуле (3.12) соответствуют следующие коэффициенты в последней строке таблицы Бутчера: $e_1 = 0$, $e_2 = e_3 = 1/2$, $e_4 = e_6 = -5/18$, $e_5 = -4/9$ и $e_7 = 0$. К сожалению, выбранная вложенная формула опять не \mathcal{A} -устойчива, поскольку функция устойчивости вложенного решения имеет вид

$$R_{\text{ENIRK4}}(z) = \frac{1 + 1/2z + 1/10z^2 + 1/120z^3 - 1/720z^5}{1 - 1/2z + 1/10z^2 - 1/120z^3}.$$

Как и раньше, чтобы избежать существенной переоценки ошибки при решении очень жестких дифференциальных уравнений (1.1), заменим оценку локальной ошибки le_{k+1} на ее модифицированную величину, получаемую в результате решения линейной системы

$$(E - \tau_k J/6)^2 \tilde{le}_{k+1} = le_{k+1}. \quad (3.13)$$

Опять же решение задачи (3.13) лишь незначительно удорожает процедуру численного интегрирования методом (3.11), так как и для продвижения на шаг, и для вычисления локальной ошибки решаются линейные системы с одной и той же матрицей коэффициентов $E - \tau_k J/6$. В то же время функция устойчивости, соответствующая модифицированной оценке локальной ошибки \tilde{le}_{k+1} , есть

$$\tilde{R}_{\text{ENIRK4}}(z) = \frac{1 + 1/6z - 7/180z^2 - 1/90z^3 - 1/864z^5}{1 - 5/6z + 53/180z^2 - 1/18z^3 + 1/180z^4 - 1/4320z^5}.$$

Нетрудно показать, что функция устойчивости $\tilde{R}_{\text{ENIRK4}}(z)$ ограничена для любых λ из комплексной полуплоскости с неположительной действительной частью и для любого размера шага τ_k . Последнее означает,

что формула (3.13) будет давать ограниченную оценку локальной погрешности даже для очень жестких задач Коши для системы (1.1).

Напомним, что для управления точностью численного интегрирования задачи (1.1) в настоящей статье используется масштабированный контроль локальной погрешности. Следовательно, модифицированная оценка локальной ошибки $\tilde{l}e_{k+1}$ (или ее исходная оценка le_{k+1} для нежестких дифференциальных уравнений) должна быть промасштабирована и проконтролирована, как указано в замечании 6 в конце предыдущего раздела.

4. НГРК-МЕТОДЫ ТИПА ЛОБАТТО

Очевидным примером лобаттовского НГРК-метода второго порядка является хорошо известное правило трапеций (см. [53, разд. 2]). Поэтому в данном разделе ограничимся построением численного метода вида (2.1), основанного на квадратуре Лобатто 4-го порядка, с использованием методики, разработанной для построения гауссовских НГРК-методов из разд. 3. Кроме того, проведем анализ этого метода, дадим рекомендации по его практическому применению и объясним, почему приведенная методика не позволяет построить НГРК-методы типа Лобатто более высокого порядка.

Из [21, разд. IV.5] известно, что абсциссы $c_i, i = 1, 2, \dots, s$, квадратурной формулы Лобатто порядка $2s - 2$ являются нулями многочлена $L_s(t) = \frac{d^{s-2}}{dt^{s-2}}(t^{s-1}(1-t)^{s-1})$. Отсюда имеем, что $c_1 = 0$, $c_s = 1$ и коэффициенты двухуровневого лобаттовского НГРК-метода четвертого порядка надо искать в виде

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ c_2 & a_{21} & a_{22} & a_{23} \\ 1 & a_{31} & a_{32} & a_{33} \\ \hline & b_1 & b_2 & b_3 \end{array}. \quad (4.1)$$

Дополнительно потребуем, чтобы выполнялись следующие условия:

- 1) $a_{22} = \tilde{a}_2 b_2$;
- 2) $a_{2j} = \tilde{a}_2 b_j + \tilde{a}_{2j}, j = 1, 3$;
- 3) $a_{3j} = b_j, j = 1, 2, 3$;
- 4) c_2 — нуль многочлена $L_3(t) = \frac{d}{dt}(t^2(1-t)^2)$, отличный от 0 и 1, т.е. $L_3(c_2) = 0$;
- 5) $b_j, j = 1, 2, 3$, — веса квадратурной формулы Лобатто 4-го порядка.

Для нахождения недостающих коэффициентов в таблице Бутчера (4.1) потребуем выполнения упрощающего условия $\mathcal{C}(3)$, которое обеспечивает третий порядок точности для стадийной величины второго уровня. Поскольку упрощающее условие $\mathcal{B}(4)$ удовлетворено благодаря использованию весов и абсцисс квадратурной формулы Лобатто 4-го порядка (см. приведенный выше список требований), то оставшиеся коэффициенты должны подчиняться линейной системе

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1/2 & 1 \\ 0 & 1/3 & 1 \end{pmatrix} \begin{pmatrix} \tilde{a}_{21} \\ \tilde{a}_2 \\ \tilde{a}_{23} \end{pmatrix} = \begin{pmatrix} c_2 \\ c_2^2/2 \\ c_2^3/3 \end{pmatrix}. \quad (4.2)$$

Матрица коэффициентов линейной задачи (4.2) невырождена, а значит и двухуровневый лобаттовский НГРК-метод, удовлетворяющий всем приведенным выше условиям, существует и единственен. Его таблица коэффициентов выглядит следующим образом:

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ 1/2 & 5/24 & 1/3 & -1/24 \\ 1 & 1/6 & 2/3 & 1/6 \\ \hline & 1/6 & 2/3 & 1/6 \end{array}.$$

Сделаем некоторые замечания относительно свойств полученного численного метода и особенностей его практической реализации.

Замечание 14. Сравнивая полученную таблицу Бутчера с [18, таблица (3.3.20)], убеждаемся, что построенный двухуровневый лобаттовский НГРК-метод совпадает с классическим трехстадийным методом Лобатто IIIA. Следовательно, он сходится с четвертым порядком, имеет третий стадийный порядок и является симметричным. Последнее делает его эффективным для решения обратимых задач Коши для уравнений вида (1.1).

Замечание 15. В силу совпадения с методом Лобатто IIIA, функция устойчивости полученной НГРК-формулы представляет собой диагональную аппроксимацию Паде $R_{22}(z)$. Поэтому этот метод применим для решения жестких дифференциальных уравнений (1.1), а также и очень жестких задач из-за его жесткой точности.

Замечание 16. Прямая подстановка показывает, что коэффициенты построенного двухуровневого лобаттовского НГРК-метода четвертого порядка обращают равенства (21)–(23) в [53, лемма 2] в тождества. Поэтому этот НГРК-метод сопряжен с симплектическим методом по меньшей мере до шестого порядка, а значит, он пригоден для интегрирования гамильтоновых систем на больших интервалах времени.

Замечание 17. К сожалению, процедуру, примененную для построения двухуровневого лобаттовского НГРК-метода четвертого порядка, нельзя продолжить дальше. Это следует из того, что для построения потенциального НГРК-метода 6-го порядка необходимо добавить еще 2 стадии, соответствующие нулям

многочлена $L_4(t) = \frac{d^2}{dt^2} (t^3(1-t)^3)$, отличным от 0 и 1, потребовать выполнения упрощающего условия $\mathcal{B}(6)$

для коэффициентов искомого метода и определить недостающие коэффициенты a_{ij} из упрощающего условия $\mathcal{C}(4)$. Однако последнее сделать невозможно по причине вырожденности линейной системы, аналогичной (4.2). Если же ограничиться только упрощающим условием $\mathcal{C}(3)$, то нельзя удовлетворить условиям порядка для корневых деревьев вида $[[\cdot, \cdot, \cdot]]$.

Замечание 18. Упрощенные итерации Ньютона для реализации двухуровневого лобаттовского НГРК-метода четвертого порядка имеют вид

$$x_{k1}^{2, \ell-1} = a_{11}^2 \bar{x}_k + a_{12}^2 x_{k+1}^{\ell-1} + \tau_k [d_{11}^2 g(t_k, \bar{x}_k) + d_{12}^2 g(t_{k+1}, x_{k+1}^{\ell-1})], \quad (4.3a)$$

$$Q_2(\tau_k J)(x_{k+1}^{\ell} - x_{k+1}^{\ell-1}) = -x_{k+1}^{\ell-1} + \bar{x}_k + \tau_k [b_1 g(t_k, \bar{x}_k) + b_2 g(t_{k1}^2, x_{k1}^{2, \ell-1}) + b_3 g(t_{k+1}, x_{k+1}^{\ell-1})], \quad (4.3b)$$

$\ell = 1, 2, \dots, N$, где матричный многочлен $Q_2(z) = (E - z/4)^2$, а остальные обозначения имеют тот же самый смысл, что и в методе (3.4). Из последней таблицы Бутчера нетрудно найти значения постоянных коэффициентов комбинированного метода (4.3), т.е. $b_1 = b_3 = 1/6$, $b_2 = 2/3$, $c_1^2 = 1/2$, $a_{11}^2 = a_{12}^2 = 1/2$, $d_{11}^2 = -d_{12}^2 = 1/8$.

Как и раньше, упрощенные итерации Ньютона (4.3) проводятся с тривиальным предиктором. Более того, двухуровневые гауссовский и лобаттовский НГРК-методы имеют одинаковый порядок сходимости, а аппроксимации матриц Якоби в них совпадают. Поэтому два итерационных шага комбинированного метода (4.3) в каждой точке сетки, т.е. $N = 2$, достаточны, чтобы обеспечить сходимость четвертого порядка для получения численного решения \bar{x}_k . Далее, найденные в результате итераций (4.3) векторы $\bar{x}_{k+1} = x_{k+1}^N, x_{k1}^{2, N}$ (подчеркнем, что формула для вычисления явной стадийной величины (4.3a) применяется еще один раз, чтобы найти ее значение, соответствующее численному решению \bar{x}_{k+1}) и решение \bar{x}_k с предыдущего шага могут быть использованы не только для построения интерполяционного многочлена Эрмита для плотной выдачи результатов численного интегрирования тоже четвертого порядка, но и для предсказания начального приближения x_{k+2}^0 для итераций (4.3) в новой точке сетки t_{k+2} . Тогда, как сле-

дует из [51, теорема 2.3], одной итерации в комбинированном методе (4.3), т.е. $N = 1$, будет достаточно для сходимости максимального четвертого порядка.

Отдельно отметим, что замена точной матрицы Якоби на ее аппроксимацию многочленом $Q_2(z) = (E - z/4)^2$ не нарушает A -устойчивости двухуровневого лобаттовского НГРК-метода, построенного выше. Это вытекает из того, что функция устойчивости метода (4.3) с тривиальным предиктором и при $N = 1$ имеет вид

$$R(z) = \frac{(1 + 1/4z)^2}{(1 - 1/4z)^2}.$$

Следовательно, она A -устойчива, т.е. не превосходит по модулю единицы для всех комплексных z с неположительной действительной частью. Также напомним, что решение линейной задачи (4.3б) с матрицей коэффициентов $Q_2(\tau_k J)$ осуществляется очень эффективно, так как каждый итерационный шаг метода (4.3) состоит в последовательном решении двух линейных систем уравнений с одной и той же матрицей коэффициентов $E - \tau_k J/4$.

Замечание 19. Для оценки локальной ошибки можно использовать вложенную пару НГРК-методов лобаттовского типа порядков 2 и 4, т.е. пару следующего вида:

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ 1/2 & 5/24 & 1/3 & -1/24 \\ 1 & 1/6 & 2/3 & 1/6 \\ \hline & 1/6 & 2/3 & 1/6 \\ \hline & e_1 & e_2 & e_3 \end{array}.$$

В этом случае для вычисления локальной ошибки имеем формулу

$$le_{k+1} = \tau_k \left[\frac{1}{3} g(t_k, \bar{x}_k) - \frac{2}{3} g(t_{k1}^2, x_{k1}^{2,N}) + \frac{1}{3} g(t_{k+1}, \bar{x}_{k+1}) \right]. \quad (4.4)$$

Другими словами, формуле (4.4) соответствуют следующие коэффициенты в последней строке таблицы Бутчера: $e_1 = e_3 = 1/3$, $e_2 = -2/3$. К сожалению, выбранная вложенная формула опять не A -устойчива, поскольку функция устойчивости вложенного решения имеет вид

$$R_{TR} = \frac{1 + 1/2z + 1/12z^2 + 1/12z^3}{1 - 1/2z + 1/12z^2}.$$

Заметим, что эта функция устойчивости совпадает с функцией устойчивости вложенного правила трапеций для двухуровневого гауссовского метода четвертого порядка. Тогда модификация оценки локальной ошибки le_{k+1} через решение линейной системы (3.6) полностью снимает эту проблему при приемлемых затратах на указанную трансформацию. Кроме того, полученная оценка должна быть промасштабирована и проконтролирована, как объяснено выше при обсуждении гауссовского НГРК-метода четвертого порядка (см. замечание 6).

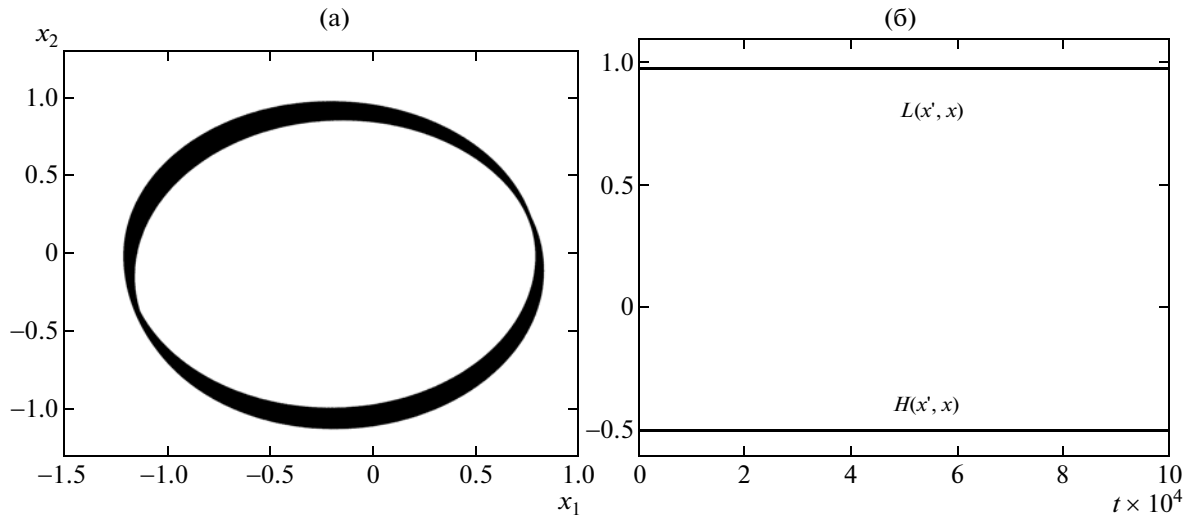
Далее, проверим работоспособность всех построенных НГРК-формул на жестких задачах Коши и гамильтоновых уравнениях.

5. ЧИСЛЕННЫЕ ПРИМЕРЫ

Начнем с гамильтоновых систем. Такие задачи характеризуются наличием первых интегралов, сохранение которых при дискретизации крайне важно, так как обычные численные методы не имеют для этого специального механизма. Последнее обусловило развитие особых численных методов, которые, кроме всего прочего, сохраняют дискретные аналогии известных инвариантов гамильтоновых систем, что существенно ограничивает накопление ошибки дискретизации в процессе численного решения и позволяет интегрировать гамильтоновы уравнения на достаточно больших интервалах времени. В настоящее время разработан богатый арсенал таких вычислительных схем, который в должной мере представлен, например, в монографиях [24], [25].

Стандартным тестом для проверки численных методов, ориентированных на численное интегрирование гамильтоновых систем, является известная задача Кеплера

$$x_1''(t) = \frac{-x_1(t)}{(x_1^2(t) + x_2^2(t))^{3/2}}, \quad x_2''(t) = \frac{-x_2(t)}{(x_1^2(t) + x_2^2(t))^{3/2}} \quad (5.1)$$



Фиг. 1.

с начальными значениями $x_1(0) = 1 - e$, $x_1'(0) = 0$, $x_2(0) = 0$, $x_2'(0) = \sqrt{1+e}/\sqrt{1-e}$. Параметр e является свободным и в нашем примере он равен 0.2. Задача (5.1) имеет два первых интеграла: полную энергию $H(x'(t), x(t))$ и угловой момент $L(x'(t), x(t))$, где $x(t) = (x_1(t), x_2(t))^T$. Точное решение задачи Кеплера известно. Оно периодически и представляет собой эллиптическую орбиту на фазовой плоскости при выбранных начальных значениях (см., например, [24], [25]).

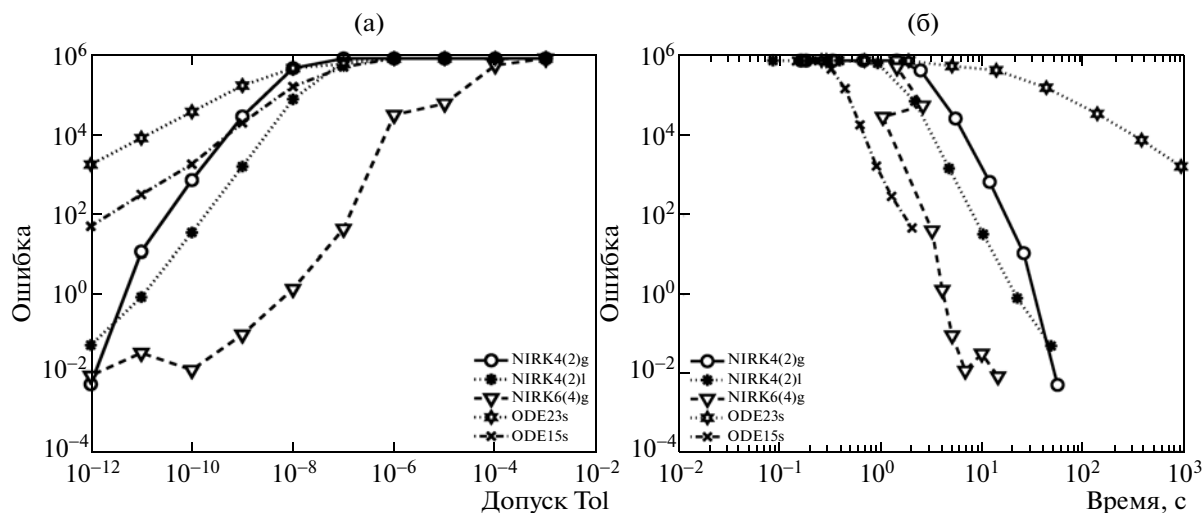
Численные методы, ориентированные на решение гамильтоновых уравнений, обязаны достаточно точно интегрировать задачу (5.1) на протяженных интервалах времени, скажем от нуля и до 10^{+5} , без критической потери точности. В [52] и [54] продемонстрировано, что приведенные выше гауссовские НГРК-методы 4-го и 6-го порядков удовлетворяют этому условию. Сейчас покажем, что лобаттовский НГРК-метод (4.3) тоже сохраняет полную энергию и угловой момент задачи Кеплера, а значит, пригоден для решения гамильтоновых систем. Надо только иметь в виду, что приведенный выше механизм автоматического управления размером шага интегрирования разрушает симплектичность численного метода (см. объяснение этого явления в [25, разд. VIII. 1]). Поэтому проверку нашего метода мы осуществляем на постоянной равномерной сетке с шагом $\tau_k = 0.1$ на интервале $[0, 10^{+5}]$. Кроме того, итерационная погрешность, которая возникает при применении итераций (4.3), также может приводить к существенному накоплению ошибки численного интегрирования. Тогда, чтобы исключить ее негативное влияние, увеличим число итераций N (с тривиальным предиктором) в каждой точке сетки до 10. Вычисления проведем на MATLAB. В итоге полученное численное решение, а также полную энергию $H(x'(t), x(t))$ и угловой момент $L(x'(t), x(t))$ изобразим на левом и правом графиках фиг. 1 соответственно.

Фиг. 1 показывает, что лобаттовский НГРК-метод (4.3) прекрасно справляется с поставленной задачей. Численное решение (фиг. 1а) достаточно точно аппроксимирует решение задачи (5.1) (см., например, [24], [25]), а полная энергия и угловой момент (фиг. 1б) сохраняют свое постоянное значение на всем интервале наблюдения. Далее, продемонстрируем, что построенные НГРК-методы пригодны для решения жестких задач для системы (1.1).

Одним из наиболее важных тестовых примеров для проверки численных методов, ориентированных на решение жестких задач, является известное уравнение Ван дер Поля (см. [21]). Оно имеет осциллирующее решение (см., например, [72, рис. 1]), которое подчиняется следующим дифференциальным уравнениям первого порядка:

$$x_1'(t) = x_2(t), \quad x_2'(t) = \mu[(1 - x_1^2(t))x_2(t) - x_1(t)], \quad \mu = 10^{+6}, \quad (5.2)$$

с начальными условиями $x(0) = (2, 0)^T$. Жесткость задачи (5.2) определяется параметром μ , который в настоящей работе равен 10^{+6} , что соответствует жесткому варианту этой задачи (см. дальнейшие объяснения в [21]). Тестирование проведем на MATLAB, в том числе и для сравнения с известными жесткими решателями этой системы ODE23s и ODE15s. Для единообразия проце-



Фиг. 2.

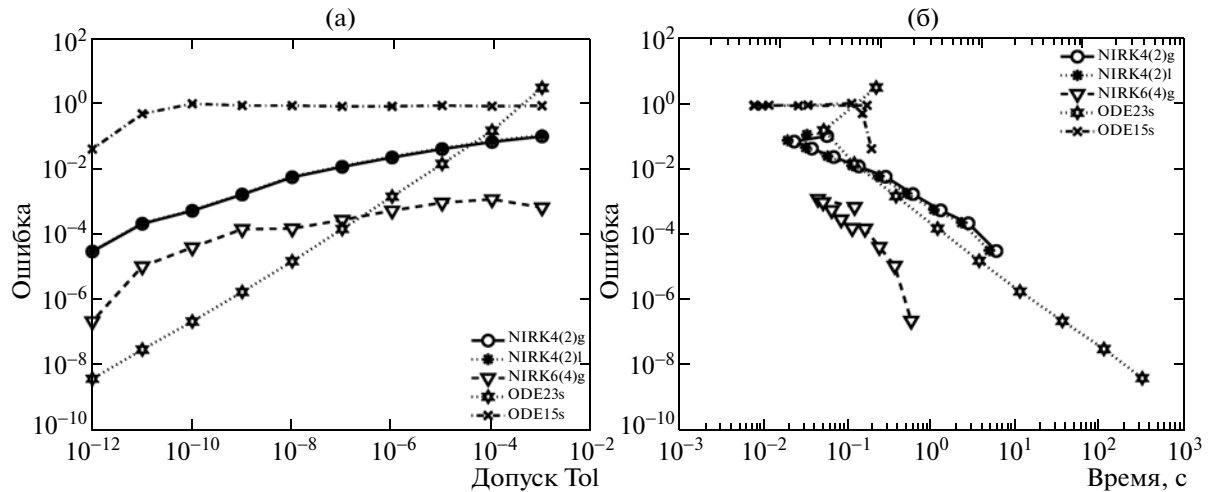
дур контроля во всех тестируемых методах мы устанавливаем следующие опции в решателях MATLAB: options = odeset('AbsTol', Tol, 'RelTol', Tol, 'MaxStep', 0.1).

Таким образом, точность численного решения определяется единым допуском Tol на локальную ошибку, который принимает следующие значения в наших экспериментах: $Tol = 10^{-3}$, 10^{-4} , ..., 10^{-12} . Как и в решателях MATLAB ODE23s и ODE15s, в НГРК-методах мы тоже ограничиваем максимальный размер шага интегрирования константой 0.1. Для обозначения результатов, соответствующих НГРК-методам, в статье приняты сокращения: NIRK4(2)g означает численный метод (3.4) вместе с вложенной формулой для оценки локальной ошибки и процедурой автоматического управления размером шага интегрирования из п. 3.1; NIRK6(4)g — численный метод (3.11) вместе с вложенной формулой для оценки локальной ошибки и процедурой автоматического управления размером шага из п. 3.2; NIRK4(2)l — численный метод (4.3) вместе с вложенной формулой для оценки локальной ошибки и процедурой автоматического управления размером шага из разд. 4. Во всех перечисленных решателях число итерационных шагов в каждой точке сетки увеличено, а также использован интерполяционный многочлен второй степени для предсказания начальных приближений x_{k+1}^0 . Это сделано, чтобы ослабить влияние итерационной погрешности на работу механизма автоматического управления локальной ошибкой, а значит, повысить эффективность численного интегрирования.

Уравнение Ван дер Поля (5.2) решается на интервале $[0, t_6]$, где $t_6 = 1.614286811415814$, а численное решение, вычисленное с высокой точностью в конечной точке этого интервала, взято из [72, табл. 8]. Указанное численное решение используется в качестве точного, чтобы оценить реальную ошибку каждого метода в конечной точке отрезка интегрирования в векторной норме $l_\infty = \|\cdot\|_\infty$. Заметим, что значение $t_6 = 1.614286811415814$ выбрано в качестве точки оценки точности не случайно. Оно соответствует импульсному поведению решения задачи (5.2), в котором погрешность численного метода существенно возрастает (см., например, [52], [54] и [72]). Поэтому преимущества и недостатки каждого решателя видны там наиболее выпукло.

Результаты этого численного эксперимента представлены на фиг. 2, где левый график (фиг. 2а) показывает зависимость достигнутой точности численного интегрирования от использованного значения допуска Tol, а правый график (фиг. 2б) — ту же самую ошибку, но относительно времени (в сек), затраченного на интегрирование уравнения Ван дер Поля (5.2). Таким образом, фиг. 2а дает диаграмму точности, а фиг. 2б показывает эффективность каждого решателя, принимавшего участие в этом эксперименте. Заметим, что фиг. 2 и 3 настоящей статьи приведены в логарифмическом масштабе.

Итак, диаграмма точности для задачи (5.2) говорит, что встроенные программы MATLAB ODE23s и ODE15s значительно проигрывают по точности вычислений решателям NIRK4(2)g, NIRK6(4)g и NIRK4(2)l, предложенным в настоящей работе, особенно при высоких требованиях к точности вычислений, т.е. при $Tol \leq 10^{-10}$. Следует также подчеркнуть, что наивысшую точ-



Фиг. 3.

ность (т.е. самое точное численное решение) дает метод NIRK6(4)g практически для всех значений допуска Tol на масштабированную локальную ошибку, принятых в этом эксперименте. Кроме того, отметим, что несмотря на то, что программы NIRK4(2)g и NIRK4(2)l базируются на вложенных парах одного и того же порядка, второй метод в целом находит более точное численное решение, чем первый, за исключением Tol = 10^{-12} . С другой стороны, диаграмма эффективности для задачи (5.2) показывает, что самым эффективным по точности, достигнутой за фиксированное время, является решатель MATLAB ODE15s. Второй по эффективности метод NIRK6(4)g, затем NIRK4(2)l и NIRK4(2)g. Наконец, программа ODE23s оказалась не только самой неэффективной, но и самой неточной из всех, протестированных в этом эксперименте (см. фиг. 2). Далее, если мы примем во внимание, что максимальная точность, достигнутая ODE15s при Tol = 10^{-12} , есть $4.981 \times 10^{+1}$, что нельзя признать удовлетворительным, то становится понятно, что наилучшим методом из рассмотренных в этой статье для численного интегрирования уравнения Ван дер Поля (5.2) будет NIRK6(4)g, затем NIRK4(2)l и NIRK4(2)g.

В качестве последнего теста возьмем наиболее трудную задачу из [73]:

$$x_1'(t) = \mu(x_2^2(t) - x_1(t)) + 2x_1(t)/x_2(t), \quad x_2'(t) = x_1(t) - x_2^2(t) + 1, \quad (5.3a)$$

$$x_3'(t) = -50(x_2(t) - 2)x_3(t), \quad \mu = 10^{+6}, \quad (5.3b)$$

с начальными условиями $x(0) = (1, 1, \exp\{-25\})^T$. Задача (5.3) решается на интервале $[0, 2]$, а ее жесткость определяется параметром μ . При $\mu = 10^{+6}$ эта задача демонстрирует жесткое поведение и является хорошим тестом для проверки наших методов. Кроме всего прочего, задача (5.3) обладает точным решением в замкнутой форме

$$x_1(t) = (t+1)^2, \quad x_2(t) = t+1, \quad x_3(t) = \exp\{-25(t-1)^2\},$$

которое не зависит от жесткости и может быть использовано для нахождения максимальной ошибки численного метода на отрезке интегрирования, тоже в векторной норме $l_\infty = \|\cdot\|_\infty$.

Результаты последнего теста представлены на фиг. 3. Интересно отметить, что для последней задачи результаты и выводы несколько отличаются от того, что было сказано в отношении уравнения Ван дер Поля (5.2). Во-первых, диаграмма точности показывает, что теперь ODE23s — один из самых точных методов для решения задачи (5.3) (см. фиг. 3a). Это и понятно, так как он имеет наименьший порядок из рассмотренных в этой статье, а значит, его управление размером шага при фиксированном допуске Tol выбирает самые маленькие шаги, что положительно сказывается на точности. Во-вторых, диаграмма эффективности показывает, что он и достаточно эффективен, хотя в этом смысле самым лучшим является NIRK6(4)g (см. фиг. 3b). Все это сильно контрастирует с результатами для задачи (5.2), где решатель MATLAB ODE23s показал наихудшее поведение. Если программа ODE15s была самой эффективной для решения уравнения Ван дер Поля (5.2), то для задачи (5.3) она, наоборот, наименее эффективна. Более того, диаграмма точ-

ности показывает, что ODE15s начинает сходиться только при очень высоких требованиях к точности вычислений, т.е. при $\text{Tol} < 10^{-10}$. При больших допусках на масштабированную локальную ошибку точность решения задачи (5.3) практически не меняется и не зависит от Tol. Таким образом, мы заключаем, что точность и эффективность встроенных решателей системы матричных вычислений MATLAB для жестких дифференциальных уравнений полностью определяется конкретной задачей Коши (1.1) и программа, которая хорошо интегрирует одну задачу, может быть совершенно неэффективной для другой. В этом смысле метод NIRK6(4)g выгодно отличается от ODE23s и ODE15s. Он демонстрирует высокую точность и эффективность для обеих жестких тестовых задач (см. фиг. 2 и 3). Кроме того, он пригоден и для решения гамильтоновых систем на больших интервалах времени (см. дальнейшее объяснение и тесты в [54]). Суммируя все сказанное выше, мы приходим к однозначному выводу, что этот метод является наиболее предпочтительным среди рассмотренных в нашей статье для практического решения жестких и гамильтоновых уравнений вида (1.1), включая задачи большой размерности.

6. ЗАКЛЮЧЕНИЕ

В настоящей работе мы изложили единую методику построения НГРК-методов типов Гаусса и Лобатто и построили конкретные вычислительные схемы порядков 4 и 6. Дополнительно мы показали, что в рамках представленного подхода построить такие методы более высокого порядка невозможно. Заметим, что отличительной чертой НГРК-методов является высокая практическая эффективность, которая вытекает из сохранения размерности исходной системы дифференциальных уравнений при их дискретизации НГРК-методом, что невозможно в неявных многостадийных РК-формулах общего вида (1.2). С другой стороны, НГРК-методы наследуют все наиболее важные свойства неявных РК-формул, такие как A -устойчивость, симметрия, симплектичность в определенном смысле. Более того, они могут иметь достаточно высокие стадийные и классические порядки, а также обеспечивать плотную выдачу результатов интегрирования той же точности, что и порядок основного метода, без больших дополнительных затрат. Таким образом, НГРК-методы эффективны для численного интегрирования дифференциальных уравнений разных видов, к которым относятся жесткие и нежесткие задачи, а также гамильтоновы системы и обратимые уравнения.

Особое внимание в статье уделено эффективной реализации построенных вычислительных схем на основе автоматического подбора достаточно оптимального размера шага интегрирования, т.е. во всех предложенных методах в качестве оптимального выбирается максимальный шаг, при котором масштабированная оценка локальной ошибки не превышает заданный предел Tol. С целью эффективного вычисления локальной ошибки для всех построенных НГРК-методов найдены соответствующие вложенные формулы. Более того, проведен анализ их функций устойчивости и предложены модификации, которые позволяют применять вложенные пары НГРК-методов для решения даже очень жестких дифференциальных уравнений. На примерах показано, что все НГРК-методы типов Гаусса и Лобатто пригодны для интегрирования гамильтоновых систем на больших интервалах времени, а также дают хорошие результаты при решении жестких задач Коши для систем вида (1.1). Более того, сравнительный анализ показал, что в целом программы NIRK4(2)g, NIRK6(4)g и NIRK4(2)1 превосходят по эффективности жесткие решатели MATLAB, ODE23s и ODE15s, а наилучшим методом среди рассмотренных в статье можно считать NIRK6(4)g, в том числе и из-за его высокого порядка сходимости.

СПИСОК ЛИТЕРАТУРЫ

1. Гайтон А. Физиология кровообращения: Минутный объем сердца и его регуляция. М.: Медицина, 1969.
2. Гродинз Ф. Теория регулирования и биологические системы. М.: Мир, 1966.
3. Марчук Г.И. Математические методы в иммунологии. Вычислительные методы и эксперименты. М.: Наука, 1991.
4. Нерпетер В. Расчет электрических цепей на персональной ЭВМ. М.: Энергоатомиздат, 1991.
5. Konyukhova N.B., Lima P.M., Morgado M.L., Soloviev M.B. Bubbles and droplets in nonlinear physics models: Analysis and numerical simulation of singular nonlinear boundary value problems // Comp. Maths. Math. Phys. 2008. V. 48. № 11. P. 2018–2058.
6. Lima P.M., Konyukhova N.B., Chemetov N.V., Sukov A.I. Analytical-numerical investigation of bubble-type solutions of nonlinear singular problems // J. Comput. Appl. Math. 2006. V. 189. P. 260–273.

7. Kulikov G.Yu., Lima P.M., Morgado M.L. Analysis and numerical approximation of singular boundary value problems with the p-Laplacian in fluid mechanics // J. Comput. Appl. Math. 2014. V. 262. P. 87–104.
8. Frogerais P., Bellanger J.-J., Senhadji L. Various ways to compute the continuous-discrete extended Kalman filter // IEEE Trans. Automat. Control. 2012. V. 57. № 4. P. 1000–1004.
9. Kulikov G.Yu., Kulikova M.V. Accurate numerical implementation of the continuous-discrete extended Kalman filter // IEEE Trans. Automat. Control. 2014. V. 59. № 1. P. 273–279.
10. Kulikov G.Yu., Kulikova M.V. Accurate state estimation in the Van der Vusse reaction // Proc. 2014 IEEE Multi-Conference on Systems and Control. 2014. P. 759–764.
11. Kulikov G.Yu., Kulikova M.V. The accurate continuous-discrete extended Kalman filter for continuous-time stochastic systems // Russian J. Numer. Anal. Math. Modelling. 2015. V. 30. № 6.
12. Kulikova M.V., Kulikov G.Yu. Square-root accurate continuous-discrete extended Kalman filter for target tracking // Proc. 52-nd IEEE Conference on Decision and Control. 2013. P. 7785–7790.
13. Kulikova M.V., Kulikov G.Yu. Adaptive ODE solvers in extended Kalman filtering algorithms // J. Comput. Appl. Math. 2014. V. 262. P. 205–216.
14. Калиткин Н.Н. Численные методы. М.: Наука, 1978.
15. Штеттер Х. Анализ методов дискретизации для обыкновенных дифференциальных уравнений. М.: Мир, 1978.
16. Ракитский Ю.В., Устинов С.М., Черноруцкий И.Г. Численные методы решения жестких систем. М.: Наука, 1979.
17. Бахвалов И.С., Жидков Н.П., Кобельков Г.М. Численные методы. М.: Наука, 1987.
18. Деккер К., Вервер Я. Устойчивость методов Рунге–Кутты для жестких нелинейных дифференциальных уравнений. М.: Мир, 1988.
19. Арушанян О.Б., Залеткин С.Ф. Численное решение обыкновенных дифференциальных уравнений на Фортране. М.: Изд-во МГУ, 1990.
20. Хайпер Э., Нерсетт С., Ваннер Г. Решение обыкновенных дифференциальных уравнений. Нежесткие задачи. М.: Мир, 1990.
21. Хайпер Э., Ваннер Г. Решение обыкновенных дифференциальных уравнений. Жесткие и дифференциально-алгебраические задачи. М.: Мир, 1999.
22. Лебедев В.И. Функциональный анализ и вычислительная математика. М.: Физматлит, 2005.
23. Шалашин В.И., Кузнецов Е.Б. Метод продолжения решения по параметру и наилучшая параметризация в прикладной математике и механике. М.: Эдиториал УРСС, 1999.
24. Sanz-Serna J.M., Calvo M.P. Numerical Hamilton problems. London: Chapman & Hall, 1994.
25. Hairer E., Lubich C., Wanner G. Geometric numerical integration: structure preserving algorithms for ordinary differential equations. Berlin: Springer-Verlag, 2002.
26. Butcher J.C. Numerical methods for ordinary differential equations. Chichester: John Wiley & Sons, 2008.
27. Butcher J.C. On the implementation of implicit Runge–Kutta methods // BIT. 1976. V. 16. P. 237–240.
28. Bickart T.A. An efficient solution process for implicit Runge–Kutta methods // SIAM J. Numer. Anal. 1977. V. 14. P. 1022–1027.
29. Cash J.R. On a class of implicit Runge–Kutta procedures // J. Inst. Math. Appl. 1977. V. 19. P. 455–470.
30. Cash J.R. On a note of the computational aspects of a class of implicit Runge–Kutta procedures // J. Inst. Maths Applies. 1977. V. 20. P. 425–441.
31. Cash J.R., Singhal A. Mono-implicit Runge–Kutta formulae for the numerical integration of stiff differential systems // IMA J. Numer. Anal. 1982. V. 2. P. 211–227.
32. Alexander R. Diagonally implicit Runge–Kutta methods for stiff ODEs // SIAM J. Numer. Anal. 1977. V. 14. P. 1006–1024.
33. Alt R. Methodes A-stables pour l'integration de systemes differentiels mal conditionnes: PhD thesis. Paris: Universite Paris, 1971.
34. Crouzeix M. Sur l'approximation des equations differentielles operationnelles lineaires par de methodes de Runge–Kutta: PhD thesis. Paris: Universite Paris, 1975.
35. Kurdi M. Stable high order methods for time discretization of stiff differential equations: PhD thesis. California: University of California, 1974.
36. Скворцов Л.М. Диагонально неявные FSAL-методы Рунге–Кутты для жестких и дифференциально-алгебраических систем // Матем. моделирование. 2002. Т. 14. № 2. С. 3–17.
37. Kvaerno A. Singly diagonally implicit Runge–Kutta methods with an explicit first stage // BIT. 2004. V. 44. P. 489–502.
38. Скворцов Л.М. Диагонально-неявные методы Рунге–Кутты для жестких задач // Ж. вычисл. матем. и матем. физ. 2006. Т. 46. № 12. С. 2209–2222.
39. Скворцов Л.М. Однократно неявные диагонально расширенные методы Рунге–Кутты четвертого порядка // Ж. вычисл. матем. и матем. физ. 2014. Т. 54. № 5. С. 755–765.
40. Скворцов Л.М., Козлов О.С. Эффективная реализация диагонально-неявных методов Рунге–Кутты // Матем. моделирование. 2014. Т. 26. № 1. С. 96–108.
41. Norsett S.P. Semi-explicit Runge–Kutta methods. Report № 6/74. Trondheim: Dept. of Math., University of Trondheim, 1974.
42. Prothero A., Robinson A. On the stability and accuracy of one-step methods for solving stiff systems of ordinary differential equations // Math. Comp. 1974. V. 24. P. 145–162.

43. Burrage K. A special family of Runge–Kutta methods for solving stiff differential equations // BIT. 1978. V. 18. P. 22–41.
44. Burrage K., Butcher J.C., Chipman F.H. An implementation of singly implicit Runge–Kutta methods // BIT. 1980. V. 20. P. 326–340.
45. Norsett S.P. Runge–Kutta methods with multiple real eigenvalue only // BIT. 1976. V. 16. P. 388–393.
46. Butcher J.C., Cash J.R. Towards efficient Runge–Kutta methods for stiff systems // SIAM J. Numer. Anal. 1990. V. 27. P. 753–761.
47. Butcher J.C., Chen D.J.L. A new type of singly implicit Runge–Kutta methods // Appl. Numer. Math. 2000. V. 34. P. 179–188.
48. Norsett S.P., Wolfbrandt A. Attainable order of rational approximations to the exponential function with only real poles // BIT. 1977. V. 17. P. 200–208.
49. Van Bokhoven W.M.G. Efficient higher order implicit one-step methods for integration of stiff differential equations // BIT. 1980. V. 20. P. 34–43.
50. Kulikov G.Yu., Shindin S.K. On a family of cheap symmetric one-step methods of order four // Computational Science – ICCS 2006. 6th International Conference, Reading, UK, May 28–31, 2006. Proc. Part I. Lecture Notes in Computer Science. 2006. V. 3991. P. 781–785.
51. Kulikov G.Yu., Merkulov A.I., Shindin S.K. Asymptotic error estimate for general Newton-type methods and its application to differential equations // Russian J. Numer. Anal. Math. Modelling. 2007. V. 22. № 6. P. 567–590.
52. Kulikov G.Yu., Shindin S.K. Numerical tests with Gauss-type nested implicit Runge–Kutta formulas // Computational Science – ICCS 2007. 7th International Conference, Beijing, China, May 27–30, 2007. Proc. Part I. Lecture Notes in Computer Science. 2007. V. 4487. P. 136–143.
53. Kulikov G.Yu., Shindin S.K. Adaptive nested implicit Runge–Kutta formulas of Gauss type // Appl. Numer. Math. 2009. V. 59. P. 707–722.
54. Kulikov G.Yu. Automatic error control in the Gauss-type nested implicit Runge–Kutta formula of order 6 // Russian J. Numer. Anal. Math. Modelling. 2009. V. 24. № 2. P. 123–144.
55. Куликов Г.Ю., Кузнецов Е.Б., Хрусталева Е.Ю. О контроле глобальной ошибки в неявных гнездовых методах Рунге–Кутты гауссовского типа // Сиб. ж. вычисл. матем. 2011. Т. 14. № 3. С. 245–259.
56. Kulikov G.Yu. Cheap global error estimation in some Runge–Kutta pairs // IMA J. Numer. Anal. 2013. V. 33. № 1. P. 136–163.
57. Лебедев В.И. Как решать явными методами жесткие системы дифференциальных уравнений // Вычисл. процессы и системы. Т. 8. М.: Наука, 1991. С. 237–291.
58. Лебедев В.И., Медовиков А.А. Явный метод второго порядка точности для решения жестких систем обыкновенных дифференциальных уравнений // Изв. высш. учебн. заведений. Матем. 1998. № 9. С. 55–63.
59. Лебедев В.И. Явные разностные схемы для решения жестких задач с комплексным или разделимым спектром // Ж. вычисл. матем. и матем. физ. 2000. Т. 40. № 12. С. 1801–1812.
60. Verwer J.G. Explicit Runge–Kutta methods for parabolic partial differential equations // Appl. Numer. Math. 1996. V. 22. № 1–3. P. 359–379.
61. Sommeijer B.P., Shampine L.F., Verwer J.G. RKC: An explicit solver for parabolic PDEs // J. Comput. Appl. Math. 1997. V. 88. № 2. P. 315–326.
62. Medovikov A.A. High order explicit methods for parabolic equations // BIT. 1998. V. 38. № 1–3. P. 372–390.
63. Abdulle A., Medovikov A.A. Second order Chebyshev methods based on orthogonal polynomials // Numer. Math. 2001. V. 90. № 1. P. 1–18.
64. Скворцов Л.М. Явный многошаговый метод численного решения жестких дифференциальных уравнений // Ж. вычисл. матем. и матем. физ. 2007. Т. 47. № 6. С. 959–967.
65. Martin-Vaquero J., Janssen B. Second-order stabilized explicit Runge–Kutta methods for stiff problems // Comput. Phys. Commun. 2009. V. 180. № 10. P. 1802–1810.
66. Скворцов Л.М. Явные многошаговые методы с расширенными областями устойчивости // Ж. вычисл. матем. и матем. физ. 2010. Т. 50. № 9. С. 1539–1549.
67. Скворцов Л.М. Явные стабилизированные методы Рунге–Кутты // Ж. вычисл. матем. и матем. физ. 2011. Т. 57. № 7. С. 1236–1250.
68. Скворцов Л.М. Простой способ построения многочленов устойчивости для явных стабилизированных методов Рунге–Кутты // Матем. моделирование. 2011. Т. 23. № 1. С. 81–86.
69. Булатов М.В., Тыглиян А.В., Филиппов С.С. Об одном классе одношаговых одностадийных методов для жестких систем обыкновенных дифференциальных уравнений // Ж. вычисл. матем. и матем. физ. 2011. Т. 57. № 7. С. 1251–1265.
70. Dahlquist G. Stability and error bounds in the numerical integration of ordinary differential equations. Trans. Royal Inst. of Technology. Stockholm, 1959. № 130.
71. Dahlquist G. A special stability problem for linear multistep methods // BIT. 1963. V. 3. P. 27–43.
72. Kulikov G.Yu., Shindin S.K. One-leg variable-coefficient formulas for ordinary differential equations and local-global step size control // Numer. Algorithms. 2006. V. 43. P. 99–121.
73. Kulikov G.Yu., Weiner R. Global error estimation and control in linearly-implicit parallel two-step peer W-methods // J. Comput. Appl. Math. 2011. V. 236. P. 1226–1239.