

**ОБЫКНОВЕННЫЕ
ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ**

УДК 519.622

**МЕТОДЫ ESDIRK ТРЕТЬЕГО И ЧЕТВЕРТОГО ПОРЯДКОВ
ДЛЯ ЖЕСТКИХ И ДИФФЕРЕНЦИАЛЬНО-АЛГЕБРАИЧЕСКИХ ЗАДАЧ**

© 2022 г. Л. М. Скворцов

105005 Москва, 2-я Бауманская, 5, МГТУ им. Н.Э. Баумана, Россия

e-mail: lm_skvo@rambler.ru

Поступила в редакцию 16.11.2021 г.
Переработанный вариант 14.12.2021 г.
Принята к публикации 11.01.2022 г.

Рассматриваются жестко точные однократно диагонально-неявные методы Рунге–Кутты с явной первой стадией (ESDIRK), предназначенные для решения жестких обыкновенных дифференциальных уравнений (ОДУ) и дифференциально-алгебраических уравнений (ДАУ). Достоинство этих методов – простая реализация, но они имеют только второй стадийный порядок, что ограничивает возможность построения эффективных методов высоких порядков. Методы ESDIRK наиболее эффективны при расчетах с невысокой точностью, достаточной для решения большинства практических задач. Поэтому в статье рассматриваются методы 3-го и 4-го порядков, позволяющие получить решение с малыми вычислительными затратами при умеренных требованиях к точности. Предложены новые методы, удовлетворяющие некоторым дополнительным условиям, что позволяет эффективно решать не только жесткие ОДУ, но и ДАУ индексов 2 и 3. Уделено внимание реализации методов с автоматическим выбором размера шага, и приведены результаты численных экспериментов. Библ. 36. Фиг. 2. Табл. 12.

Ключевые слова: обыкновенные дифференциальные уравнения, жесткая задача Коши, дифференциально-алгебраические уравнения индексов 2 и 3, диагонально-неявные методы Рунге–Кутты, ESDIRK.

DOI: 10.31857/S004446692205012X

1. ВВЕДЕНИЕ

Диагонально-неявные методы Рунге–Кутты (DIRK, см. [1]–[18]) относятся к неявным методам, которые наиболее просто реализуются. Благодаря этому методы DIRK широко применяются в прикладных вычислениях. Они используются для решения жестких ОДУ, ДАУ индексов 2 и 3 (см. [3], [4], [10]–[12], [14]), уравнений в частных производных (см. [15], [19]–[21]), а также входят в состав явно-неявных аддитивных методов (см. [21], [22]). Методы DIRK реализованы в программных продуктах MATLAB и SimInTech (см. [23]).

Один шаг метода DIRK при решении задачи Коши

$$\mathbf{y}' = \mathbf{f}(t, \mathbf{y}), \quad \mathbf{y}(t_0) = \mathbf{y}_0$$

выполняется согласно формулам

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h \sum_{i=1}^s b_i \mathbf{F}_i, \quad \mathbf{F}_i = \mathbf{f}(t_n + hc_i, \mathbf{Y}_i), \quad \mathbf{Y}_i = \mathbf{y}_n + h \sum_{j=1}^i a_{ij} \mathbf{F}_j,$$

где h – размер шага, s – число стадий, \mathbf{Y}_i и \mathbf{F}_i – стадийные значения и их производные. Таблица коэффициентов (таблица Бутчера) метода DIRK имеет вид

$$\begin{array}{c|ccc} c_1 & a_{11} & & \\ c_2 & a_{21} & a_{22} & \\ \vdots & \vdots & \vdots & \ddots \\ c_s & a_{s1} & a_{s2} & \cdots & a_{ss} \\ \hline & b_1 & b_2 & \cdots & b_s \end{array} = \begin{array}{c|c} \mathbf{c} & \mathbf{A} \\ \hline & \mathbf{b}^T \end{array}$$

(нулевые элементы матрицы \mathbf{A} обычно опускают).

Практическое применение нашли однократно диагонально- неявные методы Рунге–Кутты (SDIRK – Singly DIRK) и аналогичные методы с явной первой стадией (ESDIRK – Explicit first stage SDIRK). Методы SDIRK имеют $a_{ii} = \gamma, i = 1, \dots, s$, а методы ESDIRK имеют $a_{11} = 0, a_{ii} = \gamma, i = 2, \dots, s$. В [1] исследовалась сходимость методов SDIRK при решении жесткого уравнения Протеро–Робинсона (см. [24]). В результате был сделан вывод о преимуществе методов, имеющих $b_i = a_{si}, i = 1, \dots, s$ (такие методы Рунге–Кутты называют *жестко точными*).

Большое значение при решении жестких ОДУ и ДАУ высших индексов имеет также *стадийный порядок* – наибольшее целое q , для которого выполняются равенства

$$\mathbf{Ac}^{k-1} = \mathbf{c}^k/k, \quad \mathbf{b}^T \mathbf{c}^{k-1} = 1/k, \quad k = 1, \dots, q$$

(здесь и далее предполагаем покомпонентное выполнение операций возведения вектора в степень и умножения векторов). Стадийный порядок методов SDIRK не может быть выше 1-го, а методов ESDIRK – выше 2-го.

Простейший тест для методов решения ОДУ – уравнение Далквиста $y' = \lambda y$. Один шаг решения этого уравнения методом Рунге–Кутты запишется в виде $y_{n+1} = R(h\lambda)y_n$, где $R(z)$ – функция устойчивости, вычисляемая по формуле

$$R(z) = 1 + z\mathbf{b}^T (\mathbf{I} - z\mathbf{A})^{-1} \mathbf{e}, \quad \mathbf{e} = [1, \dots, 1]^T, \quad \mathbf{I} = \text{diag}(\mathbf{e}).$$

Метод называется $A(\alpha)$ -устойчивым, если $|R(z)| \leq 1$ при $|\arg(-z)| \leq \alpha$, и A -устойчивым, если $\alpha = 90^\circ$. Если при этом $R(\infty) = 0$, то метод называется $L(\alpha)$ -устойчивым либо L -устойчивым (при $\alpha = 90^\circ$). Жестко точные методы SDIRK удовлетворяют условию $R(\infty) = 0$. Потребуем выполнения этого условия также и для методов ESDIRK. Требование L -устойчивости часто является завышенным, поэтому будем рассматривать также и $L(\alpha)$ -устойчивые методы при значении α , близком к 90° .

Согласно [25], метод называется L -затухающим порядка $\mu > 0$, если его функция устойчивости удовлетворяет соотношению $|R(z)| = O(z^{-\mu})$ при $z \rightarrow \infty$. $L(\alpha)$ -устойчивый метод с порядком L -затухания $\mu > 1$ будем обозначать как $L\mu(\alpha)$ -устойчивый. На основании численных экспериментов мы убедились, что повышенный порядок L -затухания ($\mu > 1$) не дает преимуществ при решении жестких задач. Действительно, любой $L(\alpha)$ -устойчивый метод обеспечит порядок L -затухания μ , если принять μ шагов за один шаг. Однако $L2(\alpha)$ -устойчивый метод позволяет реализовать эффективный контроль ошибки при решении ДАУ индекса 3. Поэтому наряду с $L(\alpha)$ -устойчивыми методами рассмотрим также и $L2(\alpha)$ -устойчивый метод.

Тестовое сравнение методов SDIRK и ESDIRK было выполнено в [7]–[10], [14], где показано преимущество методов ESDIRK при решении жестких ОДУ и ДАУ индексов 2 и 3. Это преимущество объясняется более высоким стадийным порядком и наличием дополнительных коэффициентов матрицы \mathbf{A} при таком же числе неявных стадий. Поэтому мы рассматриваем жестко точные $L(\alpha)$ -устойчивые методы ESDIRK 2-го стадийного порядка, и в дальнейшем под методами ESDIRK подразумеваем именно такие методы.

Простейший метод ESDIRK, обладающий перечисленными свойствами, имеет 2-й порядок и задается таблицей Бутчера

$$\begin{array}{c|ccc} 0 & 0 & & \\ 2\gamma & \gamma & \gamma & \\ 1 & \beta & \beta & \gamma \\ \hline & \beta & \beta & \gamma \end{array}, \quad \gamma = 1 - \frac{\sqrt{2}}{2}, \quad \beta = \frac{\sqrt{2}}{4}.$$

Этот метод можно интерпретировать как последовательное применение правила трапеций (TR) и формулы дифференцирования назад 2-го порядка (BDF2), поэтому он получил название TR-BDF2. Благодаря простым расчетным формулам и высокой эффективности, этот метод широко применяется в практических вычислениях, и ему посвящено множество работ, среди которых [17], [18]. Однако метод TR-BDF2 эффективен только при вычислениях с низкой точностью, а для получения более точного результата следует использовать методы более высоких порядков.

Четырехстадийный метод ESDIRK 3-го порядка можно получить, задав γ равным одному из корней многочлена $1 - 9\gamma + 18\gamma^2 - 6\gamma^3$. Задав $\gamma = 0.4358\dots$, получаем L -устойчивый метод, а задав

$\gamma = 0.1589\dots$, получаем более точный, но $L(75.6^\circ)$ -устойчивый метод. Такие методы рассматривались в [4]–[7]. Последующие исследования показали, что можно построить более эффективные методы, увеличив число стадий. Поэтому в настоящей статье рассматриваются пяти- и шести-стадийные методы порядков 3 и 4.

2. ФУНКЦИЯ УСТОЙЧИВОСТИ

Построение метода ESDIRK обычно начинается с выбора значения диагонального элемента γ . Пусть $r = s - 1$ – число неявных стадий, а p – порядок метода. Примем $p = r - 1$ и $R(\infty) = 0$, тогда функция устойчивости метода ESDIRK однозначно определяется значением γ и имеет вид

$$R(z) = \frac{1}{1 - \gamma z} + \sum_{i=1}^{r-1} D_i(\gamma) \frac{z^i}{(1 - \gamma z)^{i+1}}, \quad D_i(\gamma) = \sum_{j=0}^i \frac{(-\gamma)^j}{(i-j)!} \binom{i}{j}. \quad (2.1)$$

Функция (2.1) аппроксимирует экспоненту с порядком $p = r - 1$. Разложив выражение $e^z - R(z)$ в ряд Тейлора, получаем

$$e^z - R(z) = \frac{C_r}{r!} z^r + \frac{C_{r+1}}{(r+1)!} z^{r+1} + O(z^{r+2}), \quad (2.2)$$

где коэффициенты C_r и C_{r+1} совпадают с одними из коэффициентов погрешности метода. Заметим, что при $\gamma = 0$, $r \leq 5$ получаем функцию устойчивости явного p -стадийного метода порядка $p = r - 1$, тогда $C_r = C_{r+1} = 1$.

При выборе подходящих значений γ исходим из того, что метод должен иметь достаточно большой сектор устойчивости (примем $\alpha > 75^\circ$). Это требование задает ограничение величины γ снизу. Кроме того, должна обеспечиваться достаточно высокая точность решения уравнения Далквиста (зададим ограничения $|C_r| < 1$, $|C_{r+1}| < 1$, т.е. точность должна быть заведомо выше, чем у явного метода порядка $r - 1$, полученного при $\gamma = 0$). Из условия $0 \leq c_i \leq 1$ получаем также $\gamma \leq 0.5$. Эти требования ограничивают величину γ сверху. Исходя из перечисленных требований, получаем ограничения значения γ для пятистадийных методов в виде $0.117 < \gamma < 0.263$, а для шестистадийных методов – в виде $0.143 < \gamma < 0.289$.

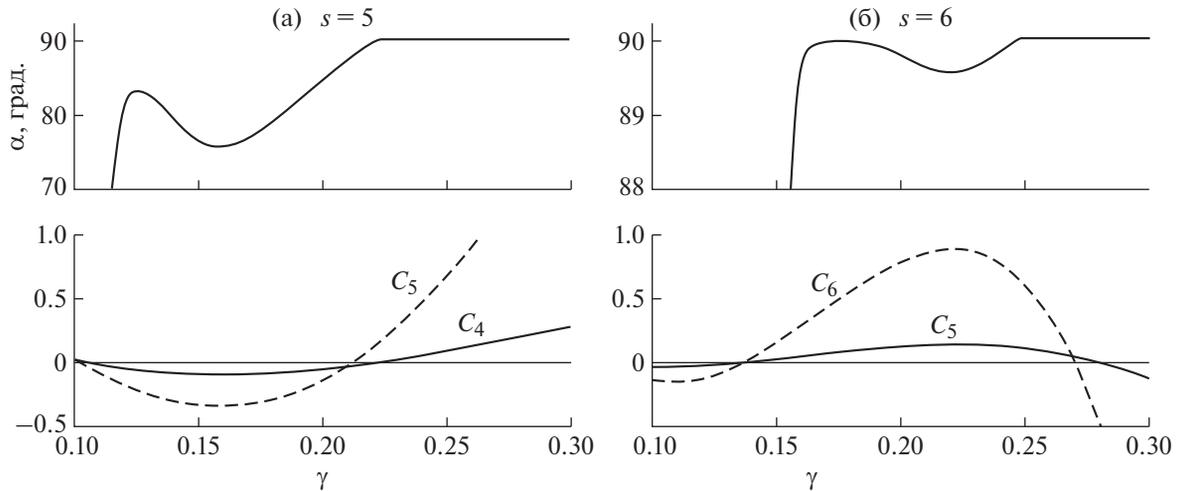
Исследуем зависимости угла α и коэффициентов C_r , C_{r+1} от γ . Для пятистадийных методов ($r = 4$) с функцией устойчивости (2.1) получаем

$$C_4 = 24D_4(\gamma), \quad C_5 = 1 - 200\gamma^2 + 1200\gamma^3 - 1800\gamma^4 + 480\gamma^5.$$

На фиг. 1а приведены графики зависимостей α , C_4 и C_5 от γ . В общем случае при выборе соответствующих коэффициентов такие методы имеют 3-й порядок. А если задать γ равным одному из 4 значений, удовлетворяющих условию $C_4 = 0$, то метод может иметь 4-й порядок. Наилучшая точность обеспечивается при наименьшем из этих значений ($\gamma = 0.1064\dots$), но такой метод не является даже $L(0)$ -устойчивым. Таким образом, выбор γ сводится к компромиссу между точностью и устойчивостью.

В первых пяти строках табл. 1 приведены характеристики функции устойчивости для некоторых значений γ , пригодных для построения пятистадийных $L(\alpha)$ -устойчивых методов 3-го порядка. Значение $\gamma = 0.125$ примерно соответствует локальному максимуму зависимости $\alpha(\gamma)$.

Близкое к этому значение $\gamma = 0.1288\dots$ (один из корней многочлена $1 - 12\gamma + 36\gamma^2 - 24\gamma^3$) обеспечивает 2-й порядок L -затухания. Метод будет L -устойчивым, если $0.2236\dots \leq \gamma \leq 0.5728\dots$ (см. [2]). Значение $\gamma = 0.225$ близко к левой границе этого интервала и использовалось в двух методах из [13]. А значение $\gamma = 0.5728\dots$ (правая граница интервала) является единственным, при котором метод L -устойчив и имеет 4-й порядок. В результате сравнения характеристик мы выбрали значение $\gamma = 0.2204\dots$, при котором метод имеет 4-й порядок (см. строку 1 в табл. 2). Метод с таким значением γ был предложен в [7] и является наиболее эффективным среди пятистадийных $L(\alpha)$ -устойчивых методов. Он не является L -устойчивым, но ниже будет показано, что даже при решении жесткой задачи с чисто мнимым спектром матрицы Якоби он показывает приемлемые результаты. В табл. 1 (строки 3, 4, 6, 7) приведены также значения γ методов, построенных в разд. 4, 5 и удовлетворяющих дополнительным условиям порядка для ДАУ индексов 2 и 3.



Фиг. 1.

Рассмотрим теперь функцию $R(z)$ шестистадийных методов 4-го порядка. В этом случае коэффициенты погрешности в (2.2) находим по формулам

$$C_5 = 120D_5(\gamma), \quad C_6 = 1 - 450\gamma^2 + 4800\gamma^3 - 16200\gamma^4 + 17280\gamma^5 - 3600\gamma^6.$$

На фиг. 1б приведены графики зависимостей α , C_5 и C_6 от γ на интересующем нас интервале, а в строках 2–5 табл. 2 приведены характеристики функции устойчивости при некоторых значениях γ . В пределах этого интервала метод будет L -устойчивым при $\gamma \geq 0.2479\dots$ (см. [2]). Наиболее удобно близкое к граничному значение $\gamma = 0.25$, которое использовалось при построении методов 4-го порядка в [2], [7]–[10], [13], [14], [21]. В [2] рекомендовалось также значение $\gamma = 4/15 = 0.2666\dots$, при котором C_5 и C_6 малы. Отметим, что при $0.164 \leq \gamma \leq 0.191$ коэффициенты погрешности невелики, а $\alpha > 89.9^\circ$. Значение $\gamma = 1/6$ – наиболее удобное из этого интервала. Методы с таким γ предлагались в [9], [10], [14]. Близкое к этому значение $\gamma = 0.1744\dots$ (один из корней многочлена $1 - 20\gamma + 120\gamma^2 - 240\gamma^3 + 120\gamma^4$) обеспечивает 2-й порядок L -затухания.

3. ТОЧНОСТЬ

Главным показателем точности методов численного решения ОДУ является порядок аппроксимации. Условия, обеспечивающие порядок p метода Рунге–Кутты, можно представить в виде

$$e(T_{ij}) = 0, \quad i = 1, \dots, p, \quad j = 1, \dots, v_i, \tag{3.1}$$

где $e(T_{ij})$ – коэффициенты погрешности метода; T_{ij} – корневые деревья порядка i , соответствующие этим коэффициентам; v_i – число различных деревьев порядка i . Для $i \leq 6$ имеем $v_1 = v_2 = 1$,

Таблица 1. Характеристики функций устойчивости методов 3-го порядка

№	s	γ	Устойчивость	C_4	C_5
1	5	0.125	$L(83.12^\circ)$	-0.0566	-0.206
2	5	0.1288	$L2(82.90^\circ)$	-0.0691	-0.233
3	5	0.1815	$L(79.35^\circ)$	-0.0800	-0.272
4	5	0.2164	$L(88.81^\circ)$	-0.0107	0.076
5	5	0.225	$L(90^\circ)$	0.0130	0.207
6	6	1/6	$L2(88.91^\circ)$	-0.0185	-0.0185
7	6	0.2	$L(90^\circ)$	0.0096	0.170

Таблица 2. Характеристики функций устойчивости методов 4-го порядка

№	s	γ	Устойчивость	C_5	C_6
1	5	0.2204	$L(89.55^\circ)$	0.135	0.878
2	6	1/6	$L(89.95^\circ)$	0.059	0.367
3	6	0.1744	$L2(89.97^\circ)$	0.076	0.476
4	6	0.25	$L(90^\circ)$	0.102	0.590
5	6	4/15	$L(90^\circ)$	0.050	0.109

$v_3 = 2, v_4 = 4, v_5 = 9, v_6 = 20$. Условия порядка до 5-го включительно вместе с соответствующими деревьями приведены в [14], [26].

В общем случае все коэффициенты погрешности различны, но если стадийный порядок больше 1, то число различных коэффициентов сокращается. Для методов 2-го стадийного порядка справедливы равенства

$$e(T_{31}) = e(T_{32}); \quad e(T_{41}) = e(T_{42}), \quad e(T_{43}) = e(T_{44}); \\ e(T_{51}) = e(T_{52}) = e(T_{55}), \quad e(T_{53}) = e(T_{54}), \quad e(T_{56}) = e(T_{57}), \quad e(T_{58}) = e(T_{59}).$$

Для методов ESDIRK из условия 1-го стадийного порядка получаем

$$a_{i1} = c_i - \sum_{j=2}^{i-1} a_{ij} - \gamma, \quad i = 2, \dots, s. \quad (3.2)$$

Коэффициенты a_{i1} не входят во все остальные условия, поэтому их вычисляют в последнюю очередь. Из условия 2-го стадийного порядка имеем

$$c_2 = 2\gamma, \quad a_{i2} = \frac{1}{4\gamma} \left[c_i^2 - 2 \left(\sum_{j=3}^{i-1} a_{ij} c_j + \gamma c_i \right) \right], \quad i = 3, \dots, s. \quad (3.3)$$

Остальные коэффициенты находим исходя из условий порядка (3.1) для $2 < i \leq p$, необходимого условия $L(\alpha)$ -устойчивости $R(\infty) = 0$ и некоторых дополнительных условий (например, условий порядка для ДАУ индексов 2 и 3).

Основная трудность при построении методов высоких порядков (4 и выше) заключается в необходимости обеспечить выполнение большого числа алгебраических условий. Учет диагональной формы матрицы \mathbf{A} позволяет упростить эти условия, в результате они становятся не сложнее, чем для явных методов. Упрощенные условия порядка для методов ESDIRK до 5-го порядка включительно приведены в [9], [14]. Пусть выполняются условия 2-го стадийного порядка (3.2), (3.3). Тогда для обеспечения 3-го порядка метода дополнительно должно выполняться условие

$$\sum_{i=2}^{s-1} b_i c_i^2 = \frac{1}{3} - \gamma, \quad (3.4)$$

а для обеспечения 4-го порядка должны выполняться также и условия

$$\sum_{i=2}^{s-1} b_i c_i^3 = \frac{1}{4} - \gamma, \quad \sum_{i=4}^{s-1} b_i \left[\sum_{j=3}^{i-1} a_{ij} \left(\sum_{k=2}^{j-1} a_{jk} c_k \right) \right] = \frac{1}{24} - \frac{1}{2}\gamma + \frac{3}{2}\gamma^2 - \gamma^3. \quad (3.5)$$

Кроме выполнения условий порядка потребуем выполнения необходимого условия $L(\alpha)$ -устойчивости, которое запишется в виде

$$R(\infty) = 1 - \tilde{\mathbf{b}}^T \tilde{\mathbf{A}}^{-2} \tilde{\mathbf{c}} = 1 - \mathbf{e}_r^T \tilde{\mathbf{A}}^{-1} \tilde{\mathbf{c}} = 0, \quad (3.6)$$

где

$$\tilde{\mathbf{A}} = \begin{bmatrix} \gamma & 0 & \cdots & 0 & 0 \\ a_{32} & \gamma & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{s-1,2} & a_{s-1,3} & \cdots & \gamma & 0 \\ a_{s2} & a_{s3} & \cdots & a_{s,s-1} & \gamma \end{bmatrix}, \quad \tilde{\mathbf{b}} = \begin{bmatrix} a_{s2} \\ a_{s3} \\ \vdots \\ a_{s,s-1} \\ \gamma \end{bmatrix}, \quad \tilde{\mathbf{c}} = \begin{bmatrix} c_2 \\ c_3 \\ \vdots \\ c_{s-1} \\ c_s \end{bmatrix}, \quad \mathbf{e}_r = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

При $p = s - 1$ выполнение этого условия обеспечивается, если γ – корень многочлена $D_p(\gamma)$, а при $p = s - 2$ условие (3.6) приводится к виду

$$a_{s,s-1}a_{s-1,s-2} \cdots a_{32}c_2 = \gamma D_p(\gamma). \tag{3.7}$$

Точность решения ОДУ методом порядка p зависит, прежде всего, от размера шага и от коэффициентов погрешности $e(T_{p+1,i})$. Используя свободные коэффициенты метода как оптимизируемые параметры, можно повысить точность метода, минимизировав коэффициенты погрешности. Такой подход является обычным при построении методов Рунге–Кутты.

При решении жестких ОДУ реальный порядок метода может быть ниже классического порядка, что приводит к заметному снижению точности решения. Для исследования этого явления, которое получило известность как феномен снижения порядка, Протеро и Робинсон (см. [24]) исследовали уравнение

$$y' = \lambda(y - \varphi(t)) + \varphi'(t), \quad y(t_0) = \varphi(t_0), \tag{3.8}$$

с точным решением $y(t) = \varphi(t)$. Используя разложение $\varphi(t)$ в ряд Тейлора, получаем локальную ошибку метода стадийного порядка q в виде

$$\delta_1 = \varphi(t_0 + h) - y_1 = \sum_{i=q+1}^{\infty} e_i(h\lambda) \frac{d^i \varphi(t_0) h^i}{dt^i i!},$$

где $e_i(z) = z\mathbf{b}^T(\mathbf{I} - z\mathbf{A})^{-1}(\mathbf{c}^i - i\mathbf{A}\mathbf{c}^{i-1}) + (1 - i\mathbf{b}^T\mathbf{c}^{i-1})$ – предложенные в [7] функции погрешности. Глобальная ошибка выражается формулой

$$\varphi(t_{n+1}) - y_{n+1} = R(h\lambda)(\varphi(t_n) - y_n) + \delta_{n+1},$$

где δ_{n+1} – локальная ошибка на $(n + 1)$ -м шаге. Для жестко точных методов с явной 1-й стадией функции погрешности можно представить в виде

$$e_i(z) = \mathbf{e}_r^T(\mathbf{I} - z\tilde{\mathbf{A}})^{-1}(\tilde{\mathbf{c}}^i - i\tilde{\mathbf{A}}\tilde{\mathbf{c}}^{i-1}).$$

В [8], [9], [11], [14] были рассмотрены также функции погрешности $e_j(z)$, полученные при анализе ошибок решения простейших модельных уравнений, отличных от уравнения Протеро–Робинсона. Но при этом для всех j имеем $e_{q+1,j}(z) \equiv e_{q+1}(z)$. Таким образом, функция $e_{q+1}(z)$ задает главный член погрешности при решении жестких модельных уравнений, поэтому далее будем рассматривать только функцию $e_3(z)$.

Анализ ошибки численного решения уравнения (3.8) в зависимости от величины $z = h\lambda$ показал важность понятий жесткой точности и стадийного порядка для эффективного решения жестких задач. Жесткая точность обеспечивает малую ошибку при больших по модулю значениях z , а высокий стадийный порядок ограничивает ошибку при умеренных z . Снижение точности и порядка тем заметнее, чем больше разность между классическим порядком p и стадийным порядком q . Можно ограничить снижение точности, если минимизировать функцию $e_3(z)$. Методы ESDIRK с минимизированной функцией погрешности предлагались в [7]–[9], [14]. В [9], [14], [27] были рассмотрены также методы, имеющие $e_3(z) \equiv 0$, но они требуют выполнения двух дополнительных стадий.

Обсудим теперь точность решения ДАУ методами ESDIRK. Системы ДАУ индекса 1 имеют вид

$$\mathbf{y}' = \mathbf{f}(\mathbf{y}, \mathbf{z}), \quad \mathbf{0} = \mathbf{g}(\mathbf{y}, \mathbf{z}),$$

Таблица 3. Необходимые условия для порядков сходимости компонент ДАУ

Условие	ДАУ индекса 2		ДАУ индекса 3		
	$p_y = 4$	$p_z = 3$	$p_y = 3$	$p_z = 3$	$p_u = 2$
(3.4)	+	+	+	+	–
(3.5)	+	–	–	–	–
(3.9)	+	+	+	+	+
(3.10)	–	–	+	+	+
(3.11)	–	–	+	+	–
(3.12)	+	–	–	–	–

где матрица $\mathbf{g}_z = \partial \mathbf{g}(\mathbf{y}, \mathbf{z}) / \partial \mathbf{z}$ обратима в окрестности решения. Жестко точные методы, к которым относятся ESDIRK, обеспечивают точное выполнение алгебраического соотношения, поэтому порядки сходимости дифференциальных и алгебраических компонент совпадают с порядком метода: $p_y = p_z = p$.

Систему ДАУ индекса 2 можно привести к виду

$$\mathbf{y}' = \mathbf{f}(\mathbf{y}, \mathbf{z}), \quad \mathbf{0} = \mathbf{g}(\mathbf{y}),$$

где матрица $\mathbf{g}_y \mathbf{f}_z$ обратима в окрестности решения. Как следствие теоремы 5.2 из [28], порядки сходимости соответствующих компонент при решении таких задач методами ESDIRK (при $p \geq 3$, $q = 2$) следующие: $p_y = \min(p, q + 1) = 3$, $p_z = q = 2$.

Систему ДАУ индекса 3 можно представить в виде

$$\mathbf{y}' = \mathbf{f}(\mathbf{y}, \mathbf{z}), \quad \mathbf{z}' = \mathbf{k}(\mathbf{y}, \mathbf{z}, \mathbf{u}), \quad \mathbf{0} = \mathbf{g}(\mathbf{y}),$$

где матрица $\mathbf{g}_y \mathbf{f}_z \mathbf{k}_u$ обратима в окрестности решения. В [29] получены порядки сходимости компонент \mathbf{y} , \mathbf{z} , \mathbf{u} при решении таких задач методами с обратимой матрицей \mathbf{A} . Для методов с явной 1-й стадией аналогичных результатов мы не нашли, но численные эксперименты с методами ESDIRK давали оценки порядков $\tilde{p}_y = \tilde{p}_z = 2$, $\tilde{p}_u = 1$.

В [10], [11], [14] для исследования сходимости решения ДАУ индексов 2 и 3 использовались простые модельные уравнения, позволившие получить дополнительные условия, необходимые для повышения порядков сходимости различных компонент ДАУ. Для методов ESDIRK при $R(\infty) = 0$, $q = 2$ эти условия имеют вид

$$\mathbf{e}_r^T \tilde{\mathbf{A}}^{-1} \tilde{\mathbf{c}}^3 = 3; \quad (3.9)$$

$$\mathbf{e}_r^T \tilde{\mathbf{A}}^{-2} \tilde{\mathbf{c}}^3 = 6; \quad (3.10)$$

$$\tilde{\mathbf{b}}^T (\tilde{\mathbf{c}} \cdot (\tilde{\mathbf{A}}^{-2} \tilde{\mathbf{c}}^3)) = 2, \quad \tilde{\mathbf{b}}^T (\tilde{\mathbf{A}}^{-2} \tilde{\mathbf{c}}^3)^2 = 12; \quad (3.11)$$

$$\tilde{\mathbf{b}}^T (\tilde{\mathbf{c}} \cdot (\tilde{\mathbf{A}}^{-1} \tilde{\mathbf{c}}^3)) = 3/4. \quad (3.12)$$

В табл. 3 приведены необходимые условия сходимости с заданным порядком для компонент ДАУ, где необходимое условие для каждой компоненты отмечено знаком +. Эти условия являются необходимыми для рассмотренных в [10], [11], [14] модельных уравнений, а значит, и для уравнений более общего вида. На ряде тестовых задач мы убедились, что выполнение этих условий действительно обеспечивает указанные порядки, но у нас нет доказательства, что эти условия являются также и достаточными.

4. ПЯТИСТАДИЙНЫЕ МЕТОДЫ

Чтобы различать рассматриваемые методы, условимся обозначать их в виде ESDIRK $_{sp}(\gamma)$, где s – число стадий, p – порядок, γ – диагональный элемент.

Пятистадийный метод 4-го порядка должен удовлетворять условиям (3.3)–(3.5), из которых получаем

$$c_2 = 2\gamma, \quad a_{32} = \frac{c_3^2 - 2\gamma c_3}{4\gamma}, \quad \begin{bmatrix} b_2 \\ b_3 \\ b_4 \end{bmatrix} = \begin{bmatrix} a_{52} \\ a_{53} \\ a_{54} \end{bmatrix} = \begin{bmatrix} c_2 & c_3 & c_4 \\ c_2^2 & c_3^2 & c_4^2 \\ c_2^3 & c_3^3 & c_4^3 \end{bmatrix}^{-1} \begin{bmatrix} 1/2 - \gamma \\ 1/3 - \gamma \\ 1/4 - \gamma \end{bmatrix}, \quad (4.1)$$

$$a_{43} = \frac{1 - 12\gamma + 36\gamma^2 - 24\gamma^3}{24b_4 a_{32} c_2}, \quad a_{42} = \frac{c_4^2 - 2(a_{43}c_3 + \gamma c_4)}{4\gamma},$$

после чего коэффициенты a_{i1} находим из (3.2). Такой метод будет иметь $R(\infty) = 0$, если γ – корень многочлена $D_4(\gamma)$. Мы выбрали значение

$$\gamma = 0.22042841025921, \quad (4.2)$$

которое считаем наиболее подходящим (тогда метод $L(89.55^\circ)$ -устойчив и имеет малые значения двух из девяти коэффициентов погрешности 5-го порядка: $e(T_{58}) = e(T_{59}) = 0.135$). Альтернативное значение $\gamma = 0.5728\dots$ обеспечивает L -устойчивость, но тогда $e(T_{58}) = e(T_{59}) = -3.27$ и $c_2 > 1$.

У нас остались два свободных параметра – c_3 и c_4 . Их подбором можно минимизировать остальные коэффициенты погрешности, но это приводит к большим значениям коэффициентов метода, что нежелательно. Более заметный эффект получим, минимизировав функцию погрешности. При $p = r = s - 1$ функция $e_3(z)$ метода ESDIRK зависит только от γ и c_3 и имеет вид

$$e_3(z) = \frac{z^{p-2}}{(1-\gamma z)^p} 2\gamma D_{p-1}(\gamma) [c_3 - (c_3^* + \gamma) - \gamma z(c_3 - c_3^*)], \quad c_3^* = 4\gamma + \gamma^2 \frac{D_{p-2}(\gamma)}{D_{p-1}(\gamma)}.$$

В [9], [14] было показано, что неравенство $c_3^* \leq c_3 \leq c_3^* + \gamma$ задает множество всех (Парето-оптимальных) значений c_3 , изменяя которые невозможно уменьшить функцию $|e_3(z)|$ сразу во всех точках левой полуплоскости. Отметим также, что значение $c_3 = c_3^*$ обеспечивает выполнение равенства (3.9).

При заданном значении $\gamma(4.2)$ интервал оптимальных значений c_3 получаем в виде $0.701 \leq c_3 \leq 0.921$. Мы выбрали

$$c_3 = (2 + \sqrt{2})\gamma \approx 0.753, \quad c_4 = \frac{a + \gamma\sqrt{b}}{48\gamma^3 - 72\gamma^2 + 24\gamma - 2} \approx 0.601,$$

$$a = (6\gamma^2 - 6\gamma + 1)c_3 + 96\gamma^3 - 100\gamma^2 + 27\gamma - 2, \quad (4.3)$$

$$b = (19872\gamma^3 - 17808\gamma^2 + 4160\gamma - 264)c_3 - 26784\gamma^3 + 24128\gamma^2 - 5656\gamma + 361.$$

Эти значения обеспечивают попадание c_3 в оптимальный интервал, а также L -устойчивость 3-й и 4-й стадий, тогда

$$a_{i1} = a_{i2}, \quad i = 2, \dots, 5. \quad (4.4)$$

Метод, задаваемый формулами (4.1)–(4.4), был предложен в [7], численные значения его коэффициентов приведены в [7], [9], [14]. Обозначим его через ESDIRK54(0.220).

Рассмотрим теперь построение пятистадийных методов, обладающих повышенной точностью при решении ДАУ индексов 2 и 3. Методы 4-го порядка не подходят для этого вследствие недостаточного числа свободных коэффициентов. Поэтому снизим порядок до 3-го и используем условия (3.2)–(3.4), (3.7), (3.9)–(3.11). В этом случае число алгебраических условий совпадает с числом коэффициентов метода. Мы нашли 12 методов, которые удовлетворяют этим условиям. Они разбиваются на 3 группы в зависимости от значений b_2 и b_3 : 1) $b_2 = 0, b_3 \neq 0$ – 6 методов; 2) $b_2 = 0, b_3 = 0$ – 3 метода; 3) $b_2 \neq 0, b_3 \neq 0$ – 3 метода. Для всех методов c_3 задается формулой

$$c_3 = \frac{(1 - 6\gamma + 2\gamma^2)(3\gamma - 6\gamma^2)}{1 - 9\gamma + 18\gamma^2 - 6\gamma^3}. \quad (4.5)$$

Из всех 12 методов только два удовлетворяют нашим требованиям. Первый из них принадлежит 1-й группе, в которой γ – один из корней многочлена $1 - 24\gamma + 186\gamma^2 - 600\gamma^3 + 828\gamma^4 - 432\gamma^5 + 72\gamma^6$. Его коэффициенты находим по формулам (3.2), (3.3), (4.5),

$$\gamma = 0.18157222316139, \quad c_4 = \frac{1 - 20\gamma + 135\gamma^2 - 366\gamma^3 + 378\gamma^4 - 72\gamma^5}{1 - 15\gamma + 81\gamma^2 - 180\gamma^3 + 162\gamma^4 - 36\gamma^5}, \quad b_2 = 0,$$

$$b_3 = \frac{(3 - 6\gamma)c_4 - 2 + 6\gamma}{c_3(c_4 - c_3)}, \quad b_4 = \frac{(3 - 6\gamma)c_3 - 2 + 6\gamma}{c_4(c_3 - c_4)}, \quad a_{43} = \frac{\gamma(1 - 9\gamma + 18\gamma^2 - 6\gamma^3)}{6b_4a_{32}c_2}.$$

Обозначим этот метод через ESDIRK53(0.182).

Второй метод принадлежит 2-й группе, в которой γ – один из трех вещественных корней многочлена $2 - 36\gamma + 201\gamma^2 - 432\gamma^3 + 360\gamma^4 - 72\gamma^5$ (два других корня – комплексно-сопряженные). Его коэффициенты находим по формулам (3.2), (3.3), (4.5),

$$\gamma = 0.21646827973787, \quad c_4 = \frac{2(1 - 3\gamma)}{3(1 - 2\gamma)}, \quad b_2 = b_3 = 0,$$

$$b_4 = \frac{3(1 - 2\gamma)^2}{4(1 - 3\gamma)}, \quad a_{43} = \frac{\gamma(1 - 9\gamma + 18\gamma^2 - 6\gamma^3)}{6b_4a_{32}c_2}.$$

Обозначим этот метод через ESDIRK53(0.216).

5. ШЕСТИСТАДИЙНЫЕ МЕТОДЫ

При построении шестистадиийных методов 4-го порядка часто используют значение $\gamma = 1/4$, которое обеспечивает L -устойчивость и малые коэффициенты погрешности. В [13] при таком значении γ свободные параметры метода выбирались из условий L -устойчивости внутренних стадий и минимизации коэффициентов погрешности. Построенный метод ESDIRK4(3)6L [2]SA (см. [13], табл. 16) “is recommended as a good default method for solving stiff problems at moderate error tolerances”. В наших обозначениях это метод ESDIRK64(1/4).

В [10], [14] рассматривались шестистадиийные методы ESDIRK 4-го порядка, удовлетворяющие условиям (3.9)–(3.12). При $\gamma = 1/6$ такие методы образуют однопараметрическое семейство со свободным параметром c_4 . Из этого семейства мы выбрали метод, имеющий $c_4 = 2/3$ и таблицу Бутчера

0	0					
1/3	1/6	1/6				
8/15	31/150	4/25	1/6			
2/3	23/88	8/99	125/792	1/6		
1/2	61/384	13/72	125/1152	-11/96	1/6	
1	1/6	0	0	0	2/3	1/6
b_i	1/6	0	0	0	2/3	1/6

Обозначим этот метод через ESDIRK64(1/6).

Рассмотрим также шестистадиийные методы 3-го порядка, удовлетворяющие условиям (3.9)–(3.11). С уменьшением порядка появились дополнительные свободные параметры, которые позволили обеспечить удобную и эффективную реализацию методов с контролем ошибки. Приведем два таких метода.

Таблица 4. Характеристики методов ESDIRK

Метод $sp(\gamma)$	α	$\ e(T_{p+1})\ $	$\ e(T_{p+2})\ $	$\ e_3(z)\ _R$	$\ e_3(z)\ _C$
ESDIRK53(0.182)	79.35°	0.180	0.78	0.0027	0.0083
ESDIRK53(0.216)	88.81°	0.084	0.44	0.0058	0.0179
ESDIRK63(1/6)	88.91°	0.026	0.05	0.0040	0.0138
ESDIRK63(1/5)	90°	0.015	0.28	0.0013	0.0052
ESDIRK54(0.220)	89.55°	0.25	1.64	0.0020	0.0116
ESDIRK64(1/6)	89.95°	0.13	0.72	0.0016	0.0086
ESDIRK64(1/4)	90°	0.16	1.11	0.0114	0.0326

Первый метод является $L_2(88.91^\circ)$ -устойчивым (см. строку 6 в табл. 1) и имеет таблицу коэффициентов

$$\begin{array}{c|cccccc}
 0 & 0 & & & & & \\
 1/3 & 1/6 & 1/6 & & & & \\
 2/3 & 1/6 & 1/3 & 1/6 & & & \\
 1 & 1/3 & 0 & 1/2 & 1/6 & & \\
 1 & 7/16 & 0 & 3/16 & 5/24 & 1/6 & \\
 1 & 1/8 & 3/8 & 3/8 & 1/360 & -2/45 & 1/6 \\
 \hline
 b_i & 1/8 & 3/8 & 3/8 & 1/360 & -2/45 & 1/6
 \end{array} \tag{5.1}$$

Обозначим его через ESDIRK63(1/6).

Второй метод является L -устойчивым (см. строку 7 в табл. 1). При его построении мы старались минимизировать функцию $e_3(z)$. Полученный метод ESDIRK63(1/5) имеет таблицу коэффициентов

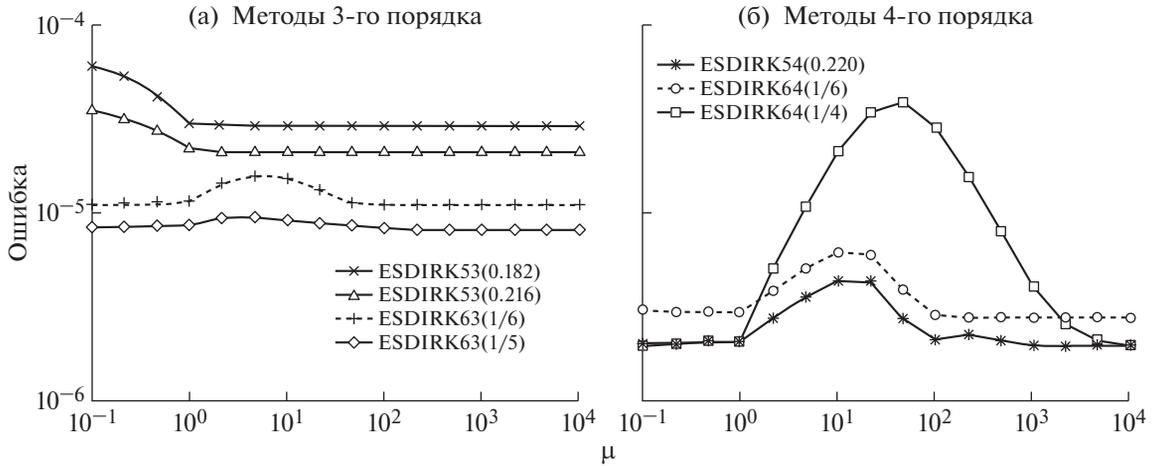
$$\begin{array}{c|cccccc}
 0 & 0 & & & & & \\
 2/5 & 1/5 & & 1/5 & & & \\
 4/5 & 1/5 & & 2/5 & 1/5 & & \\
 0 & -877/8040 & -731/4020 & 731/8040 & 1/5 & & \\
 1 & 257423/2807040 & 59/1920 & 1381/3840 & 7437/23392 & 1/5 & \\
 1 & 5047/29240 & 8/15 & 29/120 & -4489/109650 & -8/75 & 1/5 \\
 \hline
 b_i & 5047/29240 & 8/15 & 29/120 & -4489/109650 & -8/75 & 1/5
 \end{array}$$

Основные характеристики рассмотренных методов приведены в табл. 4, где

$$\|e(T_i)\| = \left(\sum_{j=1}^{v_i} e(T_{ij})^2 \right)^{1/2}, \quad \|e_3(z)\|_R = \max(|e_3(x)|, x \leq 0), \\
 \|e_3(z)\|_C = \max(|e_3(z)|, \operatorname{Re} z \leq 0) = \max(|e_3(iy)|, y \geq 0).$$

6. РЕЗУЛЬТАТЫ РАСЧЕТОВ С ПОСТОЯННЫМ РАЗМЕРОМ ШАГА

Чтобы обеспечить одинаковые условия для всех методов, задаем размер шага таким, чтобы на всем интервале было выполнено заданное число неявных стадий Nr . В этом случае будет выполнено $N = Nr/r$ шагов размером $h = T/N$. На каждой неявной стадии выполняем достаточно большое число итераций, заведомо обеспечивающее сходимость.



Фиг. 2.

Посмотрим сначала, как влияет жесткость задачи на точность решения. Для этого используем задачу

$$\begin{bmatrix} y_1' \\ y_2' \end{bmatrix} = \begin{bmatrix} a & b \\ b & a \end{bmatrix} \begin{bmatrix} y_1 - \sin t \\ y_2 - \cos t \end{bmatrix} + \begin{bmatrix} \cos t \\ -\sin t \end{bmatrix}, \quad \mathbf{y}(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

$$a = -(\mu + 1)/2, \quad b = (\mu - 1)/2, \quad 0 \leq t \leq 2\pi,$$

решение которой $y_1(t) = \sin t$, $y_2(t) = \cos t$, а собственные числа матрицы Якоби равны -1 и $-\mu$. Задаем $Nr = 120$ и вычисляем максимальное значение евклидовой нормы ошибки на всем интервале интегрирования. Полученные результаты (зависимости ошибки от μ) приведены на фиг. 2. Большой “горб” кривой для ESDIRK64(1/4) объясняется поведением функции погрешности $e_3(z)$, которая у жестко точных методов достигает максимального значения при умеренных значениях z . У других методов такой горб проявляется значительно меньше или совсем отсутствует, что объясняется меньшими значениями функции погрешности и меньшей разностью $p - q$ у методов 3-го порядка.

Вторая задача – тест PLATE (ее описание дано в [2]). Она содержит 80 переменных и имеет комплексный спектр матрицы Якоби. Интегрирование выполняем на интервале $[0, 7]$ с размерами шага $h_1 = 7/N$ и $h_2 = 0.1h_1$, где $N = 280/r$. Точность решения оцениваем величиной

$$scd = -\lg \left(\max_i \left| \frac{y_i - \tilde{y}_i}{y_i} \right| \right), \tag{6.1}$$

где y_i – точное, а \tilde{y}_i – численное решение по i -й компоненте в конце интервала интегрирования. Величина scd (significant correct digits) показывает число правильных значащих цифр численного решения. Результаты приведены в табл. 5. Видно, что порядок метода практически не влияет на точность решения этой задачи, и лучшие результаты показывают методы, имеющие малые нормы функции погрешности (см. табл. 4). Такие результаты характерны для умеренно жестких задач, а также для задач с распределенным спектром матрицы Якоби, к которым относится PLATE.

Таблица 5. Значения scd при решении задачи PLATE

h	Метод ESDIRK						
	53(0.182)	53(0.216)	63(1/6)	63(1/5)	54(0.220)	64(1/6)	64(1/4)
h_1	3.68	3.51	3.43	3.91	3.77	3.67	2.78
$0.1h_1$	6.32	5.95	5.87	6.33	6.29	6.38	5.49

Таблица 6. Результаты решения задачи (6.2) при $h = \text{const}$

Метод	e_y	e_z	\tilde{p}_y	\tilde{p}_z
ESDIRK53(0.182)	1.55×10^{-5}	7.55×10^{-5}	3.05	3.04
ESDIRK53(0.216)	7.96×10^{-6}	7.93×10^{-5}	3.04	3.00
ESDIRK63(1/6)	1.26×10^{-5}	2.02×10^{-4}	3.06	2.99
ESDIRK63(1/5)	1.13×10^{-5}	4.92×10^{-4}	3.01	2.99
ESDIRK54(0.220)	4.61×10^{-6}	3.31×10^{-4}	3.08	2.02
ESDIRK64(1/6)	1.15×10^{-6}	1.20×10^{-4}	3.98	3.01
ESDIRK64(1/4)	1.65×10^{-5}	2.67×10^{-3}	3.00	2.00

Таблица 7. Результаты решения задачи (6.3) при $h = \text{const}$

Метод	e_y	e_z	e_u	\tilde{p}_y	\tilde{p}_z	\tilde{p}_u
ESDIRK53(0.182)	6.88×10^{-6}	5.95×10^{-6}	8.26×10^{-4}	3.01	3.01	2.00
ESDIRK53(0.216)	3.70×10^{-6}	2.26×10^{-6}	4.30×10^{-4}	3.00	3.00	2.00
ESDIRK63(1/6)	3.04×10^{-6}	2.18×10^{-6}	4.66×10^{-4}	3.03	3.04	2.00
ESDIRK63(1/5)	1.43×10^{-6}	4.35×10^{-6}	1.52×10^{-3}	3.03	3.00	2.00
ESDIRK54(0.220)	5.50×10^{-5}	5.56×10^{-5}	8.57×10^{-3}	2.00	2.01	1.00
ESDIRK64(1/6)	4.74×10^{-6}	3.31×10^{-6}	1.88×10^{-3}	2.98	3.00	2.00
ESDIRK64(1/4)	2.18×10^{-4}	2.60×10^{-4}	7.61×10^{-3}	2.00	2.01	1.06

Еще две задачи – ДАУ индексов 2 и 3. Задача индекса 2 задается уравнениями

$$\begin{aligned}
 y_1' &= y_2 z, & y_2' &= y_1 (z - 2 \cos t), & 2y_1 y_2 &= \sin(2 \sin t), \\
 y_1(0) &= 0, & y_2(0) &= z(0) = 1, & 0 \leq t \leq 2\pi
 \end{aligned}
 \tag{6.2}$$

и имеет решение $y_1(t) = \sin(\sin t)$, $y_2(t) = \cos(\sin t)$, $z(t) = \cos t$. Задаем $Nr = 200$ и вычисляем e_y и e_z – максимальные ошибки соответствующих компонент на всем интервале (для вектора $\mathbf{y} = (y_1, y_2)^T$ используем евклидову норму ошибки). Вычисляем также оценки порядков сходимости \tilde{p}_y и \tilde{p}_z по соответствующим компонентам (для этого используем ошибки, полученные при размерах шага h и $h/2$). Результаты приведены в табл. 6.

Задача индекса 3 задается уравнениями

$$\begin{aligned}
 y_1' &= z_1, & y_2' &= z_2, & z_1' &= -y_1 u - y_2 \sin t, & z_2' &= -y_2 u + y_1 \sin t, \\
 & & & & y_1^2 + y_2^2 &= 1, \\
 y_1(0) &= z_2(0) = 0, & y_2(0) &= z_1(0) = u(0) = 1, & 0 \leq t \leq 2\pi
 \end{aligned}
 \tag{6.3}$$

и имеет решение $y_1(t) = \sin(\sin t)$, $y_2(t) = \cos(\sin t)$, $z_1(t) = \cos(\sin t) \cdot \cos t$, $z_2(t) = -\sin(\sin t) \cdot \cos t$, $u(t) = \cos^2 t$. Задаем $Nr = 1000$, и аналогично предыдущей задаче вычисляем ошибки и оценки порядков сходимости. Результаты приведены в табл. 7.

Из приведенных результатов следует, что методы ESDIRK53(0.220) и ESDIRK64(1/4), которые не удовлетворяют условиям (3.9)–(3.11), уступают остальным методам, для которых эти условия выполняются. Аналогичные результаты были получены и при решении других задач индексов 2 и 3.

7. РЕАЛИЗАЦИЯ МЕТОДОВ С КОНТРОЛЕМ ОШИБКИ

Для реализации с переменным размером шага были выбраны методы 3-го порядка ESDIRK63(1/6) и ESDIRK63(1/5) (как более удобные для реализации) и все три метода 4-го порядка. Используем экономичную схему реализации с двухшаговым прогнозом (см. [14], [30], [31]). Рассмотрим эту схему применительно к системе ДАУ

$$\mathbf{y}' = \mathbf{f}(t, \mathbf{y}, \mathbf{z}), \quad \mathbf{0} = \mathbf{g}(t, \mathbf{y}, \mathbf{z}), \quad \mathbf{y}(t_0) = \mathbf{y}_0, \quad \mathbf{z}(t_0) = \mathbf{z}_0. \quad (7.1)$$

Формулы $(n + 1)$ -го шага решения этой системы методом ESDIRK запишутся в виде

$$\left. \begin{aligned} \mathbf{F}_i &= \mathbf{f}_n, \\ \mathbf{Y}_i &= \mathbf{y}_n + h \sum_{j=1}^{i-1} a_{ij} \mathbf{F}_j + h\gamma \mathbf{F}_i, \quad \mathbf{F}_i = \mathbf{f}(t_n + c_i h, \mathbf{Y}_i, \mathbf{Z}_i), \\ \mathbf{g}(t_n + c_i h, \mathbf{Y}_i, \mathbf{Z}_i) &= \mathbf{0} \\ \mathbf{y}_{n+1} &= \mathbf{Y}_s, \quad \mathbf{z}_{n+1} = \mathbf{Z}_s, \quad \mathbf{f}_{n+1} = \mathbf{F}_s, \end{aligned} \right\} \quad i = 2, \dots, s, \quad (7.2)$$

где выполнение неявных стадий сводится к решению нелинейных алгебраических уравнений.

Обозначим через \mathbf{f}_y , \mathbf{f}_z , \mathbf{g}_y и \mathbf{g}_z соответствующие матрицы частных производных, вычисленные в некоторой точке численного решения (предполагается, что эти матрицы не изменяются в течение нескольких шагов). Итерации метода Ньютона при реализации i -й стадии запишутся в виде

$$\begin{aligned} \begin{bmatrix} \Delta \mathbf{Y}_i^k \\ \Delta \mathbf{Z}_i^k \end{bmatrix} &= \begin{bmatrix} \Delta \mathbf{Y}_i^{k-1} \\ \Delta \mathbf{Z}_i^{k-1} \end{bmatrix} + \begin{bmatrix} \mathbf{I} - h\gamma \mathbf{f}_y & -h\gamma \mathbf{f}_z \\ -\mathbf{g}_y & -\mathbf{g}_z \end{bmatrix}^{-1} \begin{bmatrix} h \sum_{j=1}^{i-1} a_{ij} \mathbf{F}_j + h\gamma \mathbf{F}_i^{k-1} - \Delta \mathbf{Y}_i^{k-1} \\ \mathbf{G}_i^{k-1} \end{bmatrix}, \\ \mathbf{Y}_i^k &= \mathbf{y}_n + \Delta \mathbf{Y}_i^k, \quad \mathbf{Z}_i^k = \mathbf{z}_n + \Delta \mathbf{Z}_i^k, \quad k = 1, \dots, m; \\ \mathbf{F}_i^k &= \mathbf{f}(t_n + c_i h, \mathbf{Y}_i^k, \mathbf{Z}_i^k), \quad \mathbf{G}_i^k = \mathbf{g}(t_n + c_i h, \mathbf{Y}_i^k, \mathbf{Z}_i^k), \quad k = 1, \dots, m-1; \\ \mathbf{Y}_i &= \mathbf{Y}_i^m, \quad \mathbf{Z}_i = \mathbf{Z}_i^m, \end{aligned}$$

где m – число итераций.

Перед началом итераций нужно задать начальные значения $\Delta \mathbf{Y}_i^0 = \mathbf{Y}_i^0 - \mathbf{y}_n$, $\Delta \mathbf{Z}_i^0 = \mathbf{Z}_i^0 - \mathbf{z}_n$, \mathbf{F}_i^0 , \mathbf{G}_i^0 . Для уменьшения числа итераций применяем двухшаговый прогноз, задаваемый в виде

$$\mathbf{Y}_i^0 = \sum_{j=1}^{s-1} \alpha_{ij} \bar{\mathbf{Y}}_j + \sum_{j=1}^{i-1} \beta_{ij} \mathbf{Y}_j, \quad \mathbf{Z}_i^0 = \sum_{j=1}^{s-1} \alpha_{ij} \bar{\mathbf{Z}}_j + \sum_{j=1}^{i-1} \beta_{ij} \mathbf{Z}_j,$$

где $\bar{\mathbf{Y}}_j$, $\bar{\mathbf{Z}}_j$ – стадийные значения предыдущего шага (на 1-м шаге принимаем $\bar{\mathbf{Y}}_j = \mathbf{y}_0$, $\bar{\mathbf{Z}}_j = \mathbf{z}_0$).

Значения \mathbf{F}_i^0 и \mathbf{G}_i^0 обычно вычисляют как правые части в (7.1) при $\mathbf{y} = \mathbf{Y}_i^0$, $\mathbf{z} = \mathbf{Z}_i^0$. В экономичной схеме вместо этого используем такой же прогноз, как для переменных, в результате получаем

$$\mathbf{F}_i^0 = \sum_{j=1}^{s-1} \alpha_{ij} \bar{\mathbf{F}}_j + \sum_{j=1}^{i-1} \beta_{ij} \mathbf{F}_j, \quad \mathbf{G}_i^0 = \mathbf{0}. \quad (7.3)$$

После выполнения итераций следует вычислить значение \mathbf{F}_i , которое будет использовано на последующих стадиях. Определяем его из формулы вычисления \mathbf{Y}_i в (7.2), откуда

$$\mathbf{F}_i = \frac{1}{\gamma} \left(\frac{\Delta \mathbf{Y}_i}{h} - \sum_{j=1}^{i-1} a_{ij} \mathbf{F}_j \right). \quad (7.4)$$

Использование (7.3), (7.4) позволяет сэкономить одно вычисление правой части на каждой неявной стадии, а для жестких задач (как показали эксперименты) дает более точный результат по сравнению со стандартной схемой.

Остановимся на формулах прогноза. Для методов ESDIRK прогноз 2-го порядка формируем как значение интерполяционного многочлена, построенного по уже вычисленным трем стадийным значениям. Пусть для формирования прогноза используются два стадийных значения предыдущего шага: $\bar{\mathbf{Y}}_i$ и $\bar{\mathbf{Y}}_j$ (мы задаем $i = 1$ и j – наибольшее значение, при котором $0.5 \leq c_j < 1$).

Принимаем также $w = h/\bar{h}$, где \bar{h} – размер предыдущего шага. Формулы прогноза на стадиях 2, 3, 4 запишутся в виде

$$\begin{aligned} \mathbf{Y}_2^0 &= \alpha_{2i}\bar{\mathbf{Y}}_i + \alpha_{2j}\bar{\mathbf{Y}}_j + \beta_{21}\mathbf{Y}_1, \\ \alpha_{2i} &= \frac{(wc_2 - c_j + 1)wc_2}{(c_i - c_j)(c_i - 1)}, \quad \alpha_{2j} = \frac{(wc_2 - c_i + 1)wc_2}{(c_j - c_i)(c_j - 1)}, \quad \beta_{21} = 1 - \alpha_{2i} - \alpha_{2j}; \\ \mathbf{Y}_3^0 &= \alpha_{3j}\bar{\mathbf{Y}}_j + \beta_{31}\mathbf{Y}_1 + \beta_{32}\mathbf{Y}_2, \\ \beta_{31} &= \frac{c_3 - c_2}{c_2} \left(\frac{wc_3}{c_j - 1} - 1 \right), \quad \beta_{32} = \frac{c_3(wc_3 - c_j + 1)}{c_2(wc_2 - c_j + 1)}, \quad \alpha_{3j} = 1 - \beta_{31} - \beta_{32}; \\ \mathbf{Y}_4^0 &= \beta_{41}\mathbf{Y}_1 + \beta_{42}\mathbf{Y}_2 + \beta_{43}\mathbf{Y}_3, \\ \beta_{42} &= \frac{c_4(c_4 - c_3)}{c_2(c_2 - c_3)}, \quad \beta_{43} = \frac{c_4(c_4 - c_2)}{c_3(c_3 - c_2)}, \quad \beta_{41} = 1 - \beta_{42} - \beta_{43}. \end{aligned}$$

На 5-й стадии имеет смысл формировать прогноз 3-го порядка в виде

$$\mathbf{Y}_5^0 = \beta_{51}\mathbf{Y}_1 + \beta_{52}\mathbf{Y}_2 + \beta_{53}\mathbf{Y}_3 + \beta_{54}\mathbf{Y}_4, \quad \beta_{51} = 1 - \beta_{52} - \beta_{53} - \beta_{54},$$

где коэффициенты $\beta_{52}, \beta_{53}, \beta_{54}$ находим из уравнений

$$\begin{aligned} \beta_{52}c_2 + \beta_{53}c_3 + \beta_{54}c_4 &= c_5, \\ \beta_{52}c_2^2 + \beta_{53}c_3^2 + \beta_{54}c_4^2 &= c_5^2, \\ \beta_{53}a_{32}c_2^2 + \beta_{54}(a_{42}c_2^2 + a_{43}c_3^2) &= a_{52}c_2^2 + a_{53}c_3^2 + a_{54}c_4^2 \end{aligned}$$

(под порядком прогноза понимают порядок аппроксимации стадийного значения формулой прогноза).

На 6-й стадии к условиям порядка можно добавить условие L -сходимости прогноза. Для методов ESDIRK64(1/6) и ESDIRK64(1/4) такой прогноз 3-го порядка принимаем в виде

$$\mathbf{Y}_6^0 = \beta_{61}\mathbf{Y}_1 + \dots + \beta_{65}\mathbf{Y}_5, \quad \beta_{61} = 1 - \beta_{62} - \dots - \beta_{65},$$

где $\beta_{62}, \dots, \beta_{65}$ находим из системы линейных уравнений

$$\begin{aligned} \hat{\beta}^T \hat{c} = 1, \quad \hat{\beta}^T \hat{c}^2 = 1, \quad \hat{\beta}^T \hat{A} \hat{c}^2 = 1/3, \quad \hat{\beta}^T \hat{A}^{-1} \hat{c} = 1, \\ \hat{\beta} = \begin{bmatrix} \beta_{62} \\ \beta_{63} \\ \beta_{64} \\ \beta_{65} \end{bmatrix}, \quad \hat{c} = \begin{bmatrix} c_2 \\ c_3 \\ c_4 \\ c_5 \end{bmatrix}, \quad \hat{A} = \begin{bmatrix} \gamma & 0 & 0 & 0 \\ a_{32} & \gamma & 0 & 0 \\ a_{42} & a_{43} & \gamma & 0 \\ a_{52} & a_{53} & a_{54} & \gamma \end{bmatrix}. \end{aligned}$$

Для метода ESDIRK63(1/6) принимаем $\mathbf{Y}_6^0 = (\mathbf{Y}_4 + 2\mathbf{Y}_5)/3$, что обеспечивает L -сходимость прогноза, а для ESDIRK63(1/5) принимаем $\mathbf{Y}_6^0 = \mathbf{Y}_5$.

Формулу прогноза последней стадии используем также и для формирования нормированной оценки ошибки, которую принимаем в виде

$$err = \max_i \left(\frac{|\delta y_i|}{Rtol \times \max(|y_{n,i}|, |y_{n+1,i}|) + Atol} \right), \quad \delta y = K(\mathbf{y}_{n+1} - \mathbf{Y}_s^0),$$

где $Rtol$ – допустимая относительная ошибка, $Atol$ – допустимая абсолютная ошибка. Такая оценка соответствует применению вложенной формулы, записанной в виде $\hat{\mathbf{y}}_{n+1} = \mathbf{y}_{n+1} - \delta \mathbf{y}$. Принимаем следующие значения коэффициента K : ESDIRK54(0.220) – $K = 0.5$, ESDIRK64(1/6) – $K = 0.125$, ESDIRK64(1/4) – $K = 0.0712123$ (это значение эквивалентно вложенной формуле в [13], табл. 16), ESDIRK63(1/6) и ESDIRK63(1/5) – $K = 0.25$. Шаг считаем успешным, если $err \leq 1$.

Размер следующего шага принимаем в виде $h_{new} = fac \times err^{-1/p} h$, где $fac = 0.7$ для методов 3-го порядка и $fac = 0.75$ для методов 4-го порядка.

Сходимость итерационных схем решения задачи Коши исследовалась в [32], [33]. При определенных условиях каждая итерация с неточной матрицей Якоби повышает порядок схемы на 1 до тех пор, пока не будет достигнут порядок метода. Численные эксперименты показали, что при использовании экономичной схемы и приведенных формул прогноза для решения жестких ОДУ достаточно выполнить две итерации, первая из которых не требует вычисления правой части. На последней стадии добавляем 3-ю итерацию, которая позволяет получить оценку сходимости итераций в виде $\theta = \delta_3/\delta_2$, где δ_i – норма приращения стадийных значений на i -й итерации. Перерасчет матрицы Якоби выполняем, если $\theta > 0.1$ и при этом δ_3 не очень мало – это позволяет исключить перерасчет при незначительном изменении матрицы.

При решении ДАУ высших индексов сходимость итераций более медленная, поэтому следует предусмотреть возможность выполнения большего числа итераций. В этом случае усложняется также контроль ошибки для переменных высших индексов. Исследуем точность этих компонент на примере системы

$$\begin{aligned} 0 &= y - \varphi(t), & y' &= z, & z' &= u, \\ y_0 &= \varphi(t_0), & z_0 &= \varphi'(t_0), & u_0 &= \varphi''(t_0), \end{aligned}$$

где переменная z имеет индекс 2, а переменная u – индекс 3. Глобальные ошибки решения этих уравнений запишутся в виде

$$\begin{aligned} \varphi_{n+1} - y_{n+1} &= \alpha_0 (\varphi_n - y_n) + \delta y_{n+1}, \\ \varphi'_{n+1} - z_{n+1} &= \alpha_0 (\varphi'_n - z_n) + h^{-1} \alpha_1 (\varphi_n - y_n) + \delta z_{n+1}, \\ \varphi''_{n+1} - u_{n+1} &= \alpha_0 (\varphi''_n - u_n) + h^{-1} \alpha_1 (\varphi'_n - z_n) + h^{-2} \alpha_2 (\varphi_n - y_n) + \delta u_{n+1}, \end{aligned} \tag{7.5}$$

где $\alpha_0 = R(\infty)$, $\alpha_1 = \lim_{z \rightarrow \infty} z(R(z) - \alpha_0)$, $\alpha_2 = \lim_{z \rightarrow \infty} z[z(R(z) - \alpha_0) - \alpha_1]$, δy_{n+1} , δz_{n+1} , δu_{n+1} – локальные ошибки соответствующих компонент. Для жестко точных методов $y_n = \varphi_n$ и $\delta y_{n+1} = 0$.

Аналогичные выражения получаем для вложенной формулы (заменяем α_i , δy_{n+1} , δz_{n+1} , δu_{n+1} соответствующими значениями $\hat{\alpha}_i$, $\hat{\delta} y_{n+1}$, $\hat{\delta} z_{n+1}$, $\hat{\delta} u_{n+1}$ вложенного метода). Из приведенных выражений следует, что в оценке ошибки появляются составляющие, пропорциональные h^{-1} , а также составляющие, пропорциональные только глобальной ошибке. В результате размер шага может уменьшаться до нуля, что приводит к аварийной остановке численного решения. Чтобы этого не происходило, в [34] предлагалось специальным образом масштабировать либо игнорировать оценки ошибки для переменных высших индексов.

Заметим, что если $\hat{\alpha}_0 = \alpha_0$, $\hat{\alpha}_1 = \alpha_1$ и при этом $\alpha_0 = \alpha_1 = 0$, то глобальные ошибки будут равны локальным ошибкам. В этом случае можно применять обычный контроль ошибки для всех переменных, не опасаясь аварийной остановки. Вложенная пара такого типа была построена на основе метода ESDIRK63(1/6), в который добавлена еще одна (6-я) стадия специально для оценивания ошибки. Эта же стадия используется как прогноз $Y_7^0 = Y_6$ для заключительной стадии. Полученный метод имеет таблицу Бутчера

0	0						
1/3	1/6	1/6					
2/3	1/6	1/3	1/6				
1	1/3	0	1/2	1/6			
1	7/16	0	3/16	5/24	1/6		
1	7/48	17/48	17/48	1/80	-1/30	1/6	
1	1/8	3/8	3/8	1/360	-2/45	0	1/6
b_i	1/8	3/8	3/8	1/360	-2/45	0	1/6
\hat{b}_i	7/48	17/48	17/48	1/80	-1/30	1/6	0

Здесь основной шестистадийный метод (5.1) и вложенная формула образуют семистадийный метод, который обозначим через ESDIRK73(1/6).

Таблица 8. Результаты решения задачи VDPOLE

Метод	<i>Tol</i>	<i>scd</i>	<i>mescd</i>	<i>Nf</i>	<i>NJ</i>
ESDIRK54(0.220)	10 ⁻³	3.23	3.56	1256	67
	10 ⁻⁴	4.05	4.38	1676	70
ESDIRK64(1/6)	10 ⁻³	3.11	3.43	1213	55
	10 ⁻⁴	3.89	4.21	1477	70
ESDIRK64(1/4)	10 ⁻³	2.26	2.59	1123	41
	10 ⁻⁴	3.68	3.89	1465	44
ESDIRK73(1/6)	10 ⁻³	3.84	4.16	1531	39
	10 ⁻⁴	3.92	4.24	2611	55
ESDIRK73(1/5)	10 ⁻³	2.63	2.96	1669	36
	10 ⁻⁴	4.45	4.78	2683	53
RADAU5	10 ⁻⁴	4.96	5.28	2253	162

Аналогичным образом на основе метода ESDIRK63(1/5) построен метод ESDIRK73(1/5), 6-я стадия которого является вложенной формулой для оценивания ошибки и имеет коэффициенты

$$\hat{b}_i = (2065/11008, 1019/1920, 869/3840, -5293/103200, -7/75, 1/5)$$

(но в этом случае $\hat{\alpha}_1 = \alpha_1 \neq 0$). Численные эксперименты показали, что методы (вложенные пары) ESDIRK73(1/6) и ESDIRK73(1/5) более эффективны, чем соответствующие методы (пары) ESDIRK63(1/6) и ESDIRK63(1/5), поэтому далее приводим результаты методов ESDIRK73(1/6) и ESDIRK73(1/5).

8. РЕЗУЛЬТАТЫ РАСЧЕТОВ С ПЕРЕМЕННЫМ РАЗМЕРОМ ШАГА

При реализации всех методов используем одинаковое число итераций и одинаковые условия обновления матрицы Якоби. При решении ОДУ выполняем две итерации экономичной схемы на предварительных стадиях и три итерации на последней стадии. Для решения ДАУ индекса 3 такого числа итераций оказалось недостаточно, в этом случае выполняем три итерации на каждой стадии и вычисляем матрицу Якоби на каждом шаге. Вложенные формулы (6-е стадии) в методах ESDIRK73(1/6) и ESDIRK73(1/5) реализуем, делая одну итерацию (без вычисления правой части) при решении ОДУ и две итерации при решении ДАУ.

Для решения с переменным шагом были выбраны четыре жестких теста: VDPOLE, HIRES, PLATE и BEAM, подробные описания которых приведены в [2], [35]. Интервалы интегрирования и размеры начального шага берем, как в [35], [36]. Для всех задач принимаем $Rtol = Atol = Tol$, где $Rtol$ – относительная допустимая ошибка, $Atol$ – абсолютная допустимая ошибка. Как и в [35], [36] точность решения оцениваем значениями scd (6.1) и

$$mescd = -\lg \left(\max_i \left(\frac{|y_i - \tilde{y}_i|}{Atol/Rtol + |y_i|} \right) \right),$$

где y_i – точное, а \tilde{y}_i – численное решение по i -й компоненте в конце интервала интегрирования. При решении задачи BEAM значение scd вычисляем для первых 40 (из 80) переменных, поскольку “they refer to the physically important quantities” [35]. Каждую задачу решаем при двух значениях Tol : 10⁻³ и 10⁻⁴, но если при $Tol = 10^{-3}$ точность решения низкая ($scd < 1$), то приводим результаты при $Tol = 10^{-4}$, 10⁻⁵. Вычислительные затраты оцениваем числом вычислений правой части Nf и числом вычислений матрицы Якоби NJ . Результаты расчетов приведены в табл. 8–11. Приводим также некоторые результаты решателя RADAU5, взятые из [36]. Результаты решения этих задач другими известными методами приведены в [35], [36].

Приведенные результаты подтверждают высокую эффективность методов ESDIRK при решении жестких задач с умеренной точностью. Задача BEAM имеет чисто мнимый спектр матрицы Якоби. По этой причине многие известные методы требуют большого объема вычислений

Таблица 9. Результаты решения задачи HIREС

Метод	<i>Tol</i>	<i>scd</i>	<i>mescd</i>	<i>Nf</i>	<i>NJ</i>
ESDIRK54(0.220)	10^{-3}	1.95	4.16	136	12
	10^{-4}	3.07	5.42	176	12
ESDIRK64(1/6)	10^{-4}	1.62	3.83	175	10
	10^{-5}	2.28	4.49	235	12
ESDIRK64(1/4)	10^{-4}	0.92	3.13	163	9
	10^{-5}	2.49	5.03	229	11
ESDIRK73(1/6)	10^{-4}	1.91	4.11	295	10
	10^{-5}	1.96	4.17	415	9
ESDIRK73(1/5)	10^{-4}	1.79	4.00	217	15
	10^{-5}	3.14	5.35	421	18
RADAU5	10^{-5}	1.35	3.55	381	23

Таблица 10. Результаты решения задачи PLATE

Метод	<i>Tol</i>	<i>scd</i>	<i>mescd</i>	<i>Nf</i>	<i>NJ</i>
ESDIRK54(0.220)	10^{-3}	2.83	4.53	101	1
	10^{-4}	3.72	5.61	216	1
ESDIRK64(1/6)	10^{-3}	1.78	3.92	97	1
	10^{-4}	3.18	5.13	211	1
ESDIRK64(1/4)	10^{-4}	1.87	3.81	121	1
	10^{-5}	2.39	4.40	253	1
ESDIRK73(1/6)	10^{-3}	1.18	3.32	79	1
	10^{-4}	2.96	4.90	175	1
ESDIRK73(1/5)	10^{-4}	2.46	4.60	133	1
	10^{-5}	4.38	6.49	307	1
RADAU5	10^{-4}	1.62	3.77	74	4

Таблица 11. Результаты решения задачи BEAM

Метод	<i>Tol</i>	<i>scd</i>	<i>mescd</i>	<i>Nf</i>	<i>NJ</i>
ESDIRK54(0.220)	10^{-3}	2.25	2.54	296	5
	10^{-4}	3.53	3.10	626	5
ESDIRK64(1/6)	10^{-3}	1.69	2.51	151	4
	10^{-4}	2.96	2.50	321	4
ESDIRK64(1/4)	10^{-3}	1.45	2.27	116	5
	10^{-4}	2.04	2.42	211	6
ESDIRK73(1/6)	10^{-3}	1.86	2.35	259	4
	10^{-4}	3.00	2.99	841	3
ESDIRK73(1/5)	10^{-3}	1.84	2.59	193	7
	10^{-4}	2.65	2.58	469	4
RADAU5	10^{-4}	2.49	3.57	406	43

Таблица 12. Результаты решения задачи (6.3)

Метод	Tol	e_y	e_z	e_u	N_f	N_J
ESDIRK54(0.220)	10^{-5}	3.55×10^{-4}	3.67×10^{-4}	1.90×10^{-2}	801	99
	10^{-6}	7.55×10^{-5}	7.64×10^{-5}	8.79×10^{-3}	1673	208
ESDIRK64(1/6)	10^{-3}	2.76×10^{-4}	2.54×10^{-4}	2.02×10^{-2}	711	71
	10^{-4}	3.89×10^{-6}	8.48×10^{-6}	3.78×10^{-3}	2271	226
ESDIRK64(1/4)	10^{-4}	4.58×10^{-4}	5.95×10^{-4}	1.91×10^{-2}	1351	133
	10^{-5}	5.14×10^{-5}	6.15×10^{-5}	4.08×10^{-3}	4361	436
ESDIRK73(1/6)	10^{-3}	1.13×10^{-4}	1.10×10^{-4}	6.67×10^{-3}	749	68
	10^{-4}	4.70×10^{-6}	3.70×10^{-6}	1.34×10^{-3}	1970	179
ESDIRK73(1/5)	10^{-3}	5.25×10^{-4}	7.16×10^{-4}	2.94×10^{-2}	551	50
	10^{-4}	9.16×10^{-6}	2.11×10^{-5}	3.59×10^{-3}	1706	154

при решении этой задачи даже с невысокой точностью (см. [35], [36]). Рассмотренные методы ESDIRK (в том числе и $L(\alpha)$ -устойчивые) успешно решают эту задачу с малыми вычислительными затратами. Этот факт показывает, что строгое соблюдение условия L -устойчивости не является необходимым при построении эффективных методов, достаточно, чтобы метод был $L(\alpha)$ -устойчивым при α , близком к 90° .

В табл. 12 приведены результаты решения системы ДАУ индекса 3 (6.3). Задаем $Rtol = Tol$, $Atol = 10^{-4} Tol$, $h_0 = Tol$ и вычисляем ошибки по соответствующим компонентам. Из всех методов только ESDIRK73(1/6) обеспечил устойчивое управление размером шага при контроле ошибки по всем компонентам. Поэтому в методах ESDIRK64(1/6), ESDIRK64(1/4) и ESDIRK73(1/5) контроль ошибки выполняем по y - и z -компонентам, а в методе ESDIRK54(0.220) – только по y -компоненте (в отличие от других методов, вложенная формула в ESDIRK54(0.220) имеет $\hat{\alpha}_0 \neq 0$). Чтобы получить приемлемые результаты, пришлось уменьшить значение Tol для метода ESDIRK54(0.220) на 2 порядка и для ESDIRK64(1/4) на порядок. Три других метода оказались более эффективными, при этом для решения задач индекса 3 удобнее всего использовать метод ESDIRK73(1/6), позволяющий осуществлять контроль ошибки по всем переменным.

СПИСОК ЛИТЕРАТУРЫ

1. Alexander R. Diagonally implicit Runge-Kutta methods for stiff O.D.E.'s // SIAM J. Numer. Anal. 1977. V. 14. № 6. P. 1006–1021.
2. Хайпер Э., Ваннер Г. Решение обыкновенных дифференциальных уравнений. Жесткие и дифференциально-алгебраические задачи. М.: Мир, 1999.
3. Cameron F., Palmroth M., Piche R. Quasi stage order conditions for SDIRK methods // Appl. Numer. Math. 2002. V. 42. № 1–3. P. 61–75.
4. Williams R., Burrage K., Cameron I., Kerr M. A four-stage index 2 diagonally implicit Runge-Kutta method // Appl. Numer. Math. 2002. V. 40. № 3. P. 415–432.
5. Alexander R. Design and implementation of DIRK integrators for stiff systems // Appl. Numer. Math. 2003. V. 46. № 1. P. 1–17.
6. Kværnø A. Singly diagonally implicit Runge-Kutta methods with an explicit first stage // BIT. 2004. V. 44. № 3. P. 489–502.
7. Скворцов Л.М. Диагонально неявные FSAL-методы Рунге-Кутты для жестких и дифференциально-алгебраических систем // Матем. моделирование. 2002. Т. 14. № 2. С. 3–17.
8. Скворцов Л.М. Точность методов Рунге-Кутты при решении жестких задач // Ж. вычисл. матем. и матем. физ. 2003. Т. 43. № 9. С. 1374–1384.
9. Скворцов Л.М. Диагонально неявные методы Рунге-Кутты для жестких задач // Ж. вычисл. матем. и матем. физ. 2006. Т. 46. № 12. С. 2209–2222.
10. Скворцов Л.М. Диагонально-неявные методы Рунге-Кутты для дифференциально-алгебраических уравнений индексов 2 и 3 // Ж. вычисл. матем. и матем. физ. 2010. Т. 50. № 6. С. 1047–1059.
11. Скворцов Л.М. Модельные уравнения для исследования точности методов Рунге-Кутты // Матем. моделирование. 2010. Т. 22. № 5. С. 146–160.

12. *Rang J.* An analysis of the Prothero–Robinson example for constructing new adaptive ESDIRK methods of order 3 and 4 // *Appl. Numer. Math.* 2015. V. 94. P. 75–87.
13. *Kennedy C.A., Carpenter M.H.* Diagonally implicit Runge–Kutta methods for ordinary differential equations // *A Rev. NASA report NASA/TM-2016-219173.* 2016.
14. *Скворцов Л.М.* Численное решение обыкновенных дифференциальных и дифференциально-алгебраических уравнений. М: ДМК Пресс, 2018.
15. *Boom P.D., Zingg D.W.* Optimization of high-order diagonally-implicit Runge–Kutta methods // *J. Comput. Phys.* 2018. V. 371. P. 168–191.
16. *Kennedy C.A., Carpenter M.H.* Diagonally implicit Runge–Kutta methods for stiff ODEs // *Appl. Numer. Math.* 2019. V. 146. P. 221–244.
17. *Hosea M.E., Shampine L.F.* Analysis and implementation of TR-BDF2 // *Appl. Numer. Math.* 1996. V. 20. № 1–2. P. 21–37.
18. *Bonaventura L., Marmol M.G.* The TR-BDF method for second order problems in structural mechanics // *Comput. Math. Appl.* 2021. V. 92. P. 13–26.
19. *Брагин М.Д., Rogov B.B.* Метод итерированной приближенной факторизации операторов высокоточной бикомпактной схемы для систем многомерных неоднородных квазилинейных уравнений гиперболического типа // *Ж. вычисл. матем. и матем. физ.* 2018. Т. 58. № 3. С. 313–325.
20. *Rogov B.B., Чикиткин А.В.* О сходимости и точности метода итерированной приближенной факторизации операторов многомерных высокоточных бикомпактных схем // *Матем. моделирование.* 2019. Т. 31. № 12. С. 119–144.
21. *Kennedy C.A., Carpenter M.H.* Additive Runge–Kutta schemes for convection–diffusion–reaction equations // *Appl. Numer. Math.* 2003. V. 44. P. 139–181.
22. *Kennedy C.A., Carpenter M.H.* Higher-order additive Runge–Kutta schemes for ordinary differential equations // *Appl. Numer. Math.* 2019. V. 136. P. 183–205.
23. *Карташов Б.А., Шаббаев Е.А., Козлов О.С., Шекатуров А.М.* Среда динамического моделирования технических систем SimInTech. М: ДМК Пресс, 2017.
24. *Prothero A., Robinson A.* On the stability and accuracy of one-step methods for solving stiff systems of ordinary differential equations // *Math. Comput.* 1974. V. 28. № 125. P. 145–162.
25. *Кочетков К.А., Ширков П.Д.* L-затухающие ROW-методы третьего порядка точности // *Ж. вычисл. матем. и матем. физ.* 1997. Т. 37. № 6. С. 699–710.
26. *Хайрер Э., Нёрсетт С., Ваннер Г.* Решение обыкновенных дифференциальных уравнений. Нежесткие задачи. М.: Мир, 1990.
27. *Скворцов Л.М.* Как избежать снижения точности и порядка методов Рунге–Кутты при решении жестких задач // *Ж. вычисл. матем. и матем. физ.* 2017. Т. 57. № 7. С. 1126–1141.
28. *Jay L.* Convergence of a class of Runge–Kutta methods for differential-algebraic systems of index 2 // *BIT.* 1993. V. 33. № 1. P. 137–150.
29. *Jay L.* Convergence of Runge–Kutta methods for differential-algebraic systems of index 3 // *Appl. Numer. Math.* 1995. V. 17. № 2. P. 97–118.
30. *Скворцов Л.М.* Экономичная схема реализации неявных методов Рунге–Кутты // *Ж. вычисл. матем. и матем. физ.* 2008. Т. 48. № 11. С. 2008–2018.
31. *Скворцов Л.М., Козлов О.С.* Эффективная реализация диагонально-неявных методов Рунге–Кутты // *Матем. моделирование.* 2014. Т. 26. № 1. С. 96–108.
32. *Куликов Г.Ю.* Теоремы сходимости для итерационных методов Рунге–Кутты с постоянным шагом интегрирования // *Ж. вычисл. матем. и матем. физ.* 1996. Т. 36. № 8. С. 73–89.
33. *Jackson K.R., Kværnø A., Nørsett S.P.* An analysis of the order of Runge–Kutta methods that use an iterative scheme to compute their internal stage values // *BIT.* 1996. V. 36. № 4. P. 713–765.
34. *Hairer E., Lubich Ch., Roche M.* The numerical solution of differential-algebraic systems by Runge–Kutta methods. Berlin: Springer-Verlag, 1989.
35. *Mazzia F., Magherini C.* Test set for initial value problem solvers. Release 2.4. 2008. URL: <http://pitagora.dm.uniba.it/~testset/report/testset.pdf>.
36. The codes BiM and BiMD home page. 2014. URL: <http://web.math.unifi.it/users/brugnano/BiM/>.