



ИПМ им.М.В.Келдыша РАН • Электронная библиотека

Препринты ИПМ • Препринт № 227 за 2018 г.



ISSN 2071-2898 (Print)  
ISSN 2071-2901 (Online)

**Жолковский Е.К., Белов А.А.,  
Калиткин Н.Н.**

Решение жестких задач  
Коши явными схемами с  
геометрически-адаптивным  
выбором шага

**Рекомендуемая форма библиографической ссылки:** Жолковский Е.К., Белов А.А., Калиткин Н.Н. Решение жестких задач Коши явными схемами с геометрически-адаптивным выбором шага // Препринты ИПМ им. М.В.Келдыша. 2018. № 227. 20 с. doi:[10.20948/prepr-2018-227](https://doi.org/10.20948/prepr-2018-227)  
URL: <http://library.keldysh.ru/preprint.asp?id=2018-227>

РОССИЙСКАЯ АКАДЕМИЯ НАУК  
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ  
ИМ. М.В. КЕЛДЫША

Е. К. Жолковский, А. А. Белов, Н. Н. Калиткин

РЕШЕНИЕ ЖЕСТКИХ ЗАДАЧ КОШИ  
ЯВНЫМИ СХЕМАМИ  
С ГЕОМЕТРИЧЕСКИ-АДАПТИВНЫМ  
ВЫБОРОМ ШАГА

Москва, 2018

Е. К. Жолковский, А. А. Белов, Н. Н. Калиткин. Решение жестких задач Коши явными схемами с геометрически-адаптивным выбором шага.

Предложен явный метод, позволяющий решать жесткие задачи Коши. Метод использует явные схемы и алгоритм выбора шага, основанный на кривизне интегральной кривой. Для нахождения кривизны построены явные формулы. Для схем Рунге-Кутты с числом стадий до 4 приведены наборы соответствующих коэффициентов схем. Проведена верификация метода на тестовой задаче с известным точным решением. Показано, что метод по точности и надежности не уступает неявным методам, но кардинально превосходит их по экономичности.

*Ключевые слова:* жесткие задачи Коши, явные схемы, автоматический выбор шага, геометрически-адаптивные сетки

E. K. Zholkovskii, A. A. Belov, N. N. Kalitkin. Solution of stiff Cauchy problems with explicit schemes with geometrical-adaptive step selection.

We propose an explicit numerical method for solution of stiff Cauchy problems. The method implies explicit schemes and step selection procedure based on curvature of the integral curve. We propose explicit formulae for the curvature. For the Runge-Kutta schemes with up to 4 stages, the sets of the scheme coefficients are provided. Verification of the method is performed on a test problem with a known exact solution. We show that the method possesses the same accuracy and robustness as implicit methods and sufficiently excels them in efficiency.

*Keywords:* stiff Cauchy problems, explicit schemes, automatic step selection, geometrical-adaptive meshes

Работа поддержана грантом РФФИ №18-01-00175 и программой РУДН «5-100».

# 1 Введение

Жесткие задачи Коши для обыкновенных дифференциальных уравнений (ОДУ) возникают во многих прикладных областях. Это задачи химической кинетики, динамические процессы в широкополосной радиоаппаратуре и микросхемах, электрический пробой и подобные процессы с сингулярностями и др. В таких задачах производные отдельных компонент решения (скорости процессов) различаются на много порядков. Расчет таких задач экономичными явными методами требовал неприемлемо малых шагов. В 1952 году Далквист доказал теорему о том, что явные методы не могут быть  $A$ -устойчивыми [1]. С тех пор все расчеты задач высокой жесткости велись по надежным, но довольно трудоемким неявным схемам. Много таких схем описано в [2].

Однако не прекращались попытки использовать явные схемы, компенсируя отсутствие  $A$ -устойчивости некоторой стратегией выбора шага: сильно измельчать шаг тогда, когда существенны наиболее быстрые процессы, и увеличивать его при преобладании медленных процессов. Такие стратегии были основаны на использовании вложенных схем или локального сгущения шага. Хорошим примером является потребительная программа DOPRI5 [3]. Однако такие подходы не позволяют построить программы, обеспечивающие заданную точность. Нередко точность удовлетворительна, однако известно много случаев срыва расчета в таких программах или формального завершения расчета, но с грубо ошибочными результатами.

Мы предложили другую стратегию выбора шага, содержащую три важнейших элемента. 1) Выбирается новый аргумент: вместо традиционного времени  $t$  в качестве аргумента выбирается длина дуги  $l$  интегральной кривой в многомерном пространстве. Этот прием был предложен давно [4], но в настоящее время мало используется. Он крайне существенен для жестких задач, но и для мягких задач дает заметное улучшение.

2) В качестве критерия выбора шага используется кривизна интегральной кривой: чем больше кривизна, тем меньше должен быть шаг [5], [6]. Такой критерий в сочетании с длиной дуги оказывается более эффективным, чем ранее предлагавшиеся.

3) Расчет не ограничивается одной сеткой. Производится многократное сгущение сетки и сравнение разностных решений на этих сетках. Это позволяет применить метод Ричардсона и получить гарантированную оценку погрешности. Ранее мы применяли этот подход к неявным схемам. Это объясняется тем, что для аккуратного определения кривизны требуется вычислять матрицу Якоби от правых частей, что является очень трудоемкой процедурой. Однако для неявных схем матрица Якоби все равно вычисляется, и вычисление кривизны не увеличивает трудоемкость расчета.

В данной работе этот подход перенесен на явные схемы. Для этого построены явные методы вычисления кривизны, не использующие матрицы Якоби. Эти методы малотрудоемки и просты. При их применении с явными схемами Рунге-Кутты результаты сравнимы по надежности с лучшим неявным схемам, а по экономичности кардинально их превосходят.

## 2 Построение метода

### 2.1 Длина дуги

Рассмотрим задачу Коши для системы ОДУ

$$\frac{du_k}{dt} = f_k(\mathbf{u}, t), \quad 1 \leq k \leq K, \quad \mathbf{u}(0) = \mathbf{u}^0; \quad (1)$$

Если система жесткая, то ее решение содержит участки быстрого изменения решения – начальный и внутренние пограничные слои (последние называют также контрастными структурами). Численное решение таких задач существенно облегчается, если аргумент  $t$  считать дополнительной компонентой решения  $u_0(t) \equiv t$ ,  $f_0 \equiv 1$ , а в качестве нового аргумента выбрать длину дуги интегральной кривой

$$dl = \rho dt, \quad \rho = \left(1 + \sum f_k^2\right)^{1/2}. \quad (2)$$

Тогда система (1) принимает вид

$$\frac{du_k}{dl} = F_k(\mathbf{u}), \quad F_k = f_k \rho^{-1}, \quad 0 \leq k \leq K. \quad (3)$$

Система (3) автономна, а ее порядок на 1 больше, чем у (1).

### 2.2 Выбор шага

Выбор шага по кривизне интегральной кривой был предложен в [5], [6]. Здесь мы воспользуемся конкретной формулой

$$h = \frac{h_0}{(1 + L^{0.5/\nu}(\varkappa, \varkappa)^{0.25/\nu})^\nu}. \quad (4)$$

Здесь  $L$  – полный промежуток интегрирования по  $l$ ,  $\varkappa$  – кривизна,  $h_0$  – шаг на почти прямолинейном участке интегральной кривой. Напомним, что размерность кривизны  $\varkappa$  равна размерности  $L^{-1}$ . Поэтому в знаменателе стоит безразмерная комбинация величин. Численные расчеты показали, что для сверхжестких задач целесообразно выбирать  $\nu = 2$ , а для задач меньшей жесткости –  $\nu = 1$ .

Выбирая некоторое значение  $h_0$ , строим сетку одновременно с вычислением решения. Тем самым сетка будет адаптирована к данному решению. Эти сетки необходимо сгущать. Процесс сгущения мы опишем ниже.

### 2.3 Вычисление кривизны

Новый вектор правых частей имеет единичную длину

$$|\mathbf{F}|^2 = \sum F_k^2 = 1. \quad (5)$$

Таким образом, правые части (3) образуют единичный вектор касательной к интегральной кривой. Как известно, вектор кривизны в многомерном пространстве есть производная единичного вектора направления по длине дуги

$$\varkappa = \frac{d\mathbf{F}}{dl}. \quad (6)$$

Рассмотрим способ вычисления этого вектора.

Вычисления проводятся на некоторой сетке узлов. Чтобы не вводить номер узла сетки, будем обозначать решение на текущем шаге через  $\mathbf{u}$ , на новом –  $\hat{\mathbf{u}}$ , на предыдущем –  $\check{\mathbf{u}}$ . Тогда простейшая аппроксимация для кривизны вычисляется с использованием предыдущего шага

$$\varkappa = \check{h}^{-1}(\mathbf{F} - \check{\mathbf{F}}), \quad \check{h} = l - \check{l}. \quad (7)$$

Погрешность такой аппроксимации составляет всего  $O(h)$ , так как величина  $\varkappa$  фактически относится к середине предыдущего шага. Однако (7) можно использовать с явными схемами любого типа.

Кроме того, кривизна используется лишь для расчета длины очередного шага. Поэтому она должна слабо влиять на порядок точности самого решения  $\mathbf{u}$ . Проведенные расчеты подтверждают это.

### 2.4 Схемы Рунге-Кутты (РК)

Для конкретных схем можно построить более точные аппроксимации  $\varkappa$ . Например, для явных схем РК можно включить в расчет промежуточные стадии нового шага и не использовать старый шаг. Мы провели такие построения для схем РК с числом стадий до 4. Приведем полученные результаты.

Явные S-стадийные схемы РК имеют следующий вид:

$$\hat{\mathbf{u}} = \mathbf{u} + h \sum_{s=1}^S b_s \mathbf{w}_s, \quad \mathbf{w}_s = \mathbf{F} \left( \mathbf{u} + h \sum_{q=1}^{s-1} a_{sq} \mathbf{w}_q \right). \quad (8)$$

Ограничимся схемами с  $S \leq 4$ . Тогда матрицу Бутчера  $\{a_{sq}\}$  и столбец  $\{b_s\}$  можно подобрать так, чтобы  $S$ -стадийная схема имела аппроксимацию  $O(h^S)$ . Будем использовать только такие схемы.

На текущем шаге рассчитаем величину  $\hat{\boldsymbol{\varkappa}}$ , которая будет использоваться для нахождения следующего шага. Для этого используем стадии  $\mathbf{w}_s$  и вдобавок величину  $\mathbf{F}(\hat{\mathbf{u}})$ , поскольку она все равно будет вычисляться на следующем шаге. Ищем аппроксимацию в виде

$$\hat{\boldsymbol{\varkappa}} = h^{-1} \sum_{q=1}^{S+1} c_q \mathbf{w}_q, \quad \mathbf{w}_{S+1} = \mathbf{F}(\hat{\mathbf{u}}). \quad (9)$$

Коэффициенты  $c_q$ ,  $b_s$ ,  $a_{sq}$  нужно подбирать так, чтобы решение (8) имело аппроксимацию  $O(h^S)$ , а аппроксимация кривизны (9) была как можно лучше. Рекомендуемые наборы коэффициентов приведены в табл. 1 – 4. При  $S = 1, 2$  аппроксимация кривизны есть  $O(h)$ , а при  $S = 3, 4$  –  $O(h^2)$ .

Таблица 1: Коэффициенты схемы (8) и кривизны (9) для  $S = 1$ .

$c_q$	$b_s$	$a_{sq}$
-1	1	-
1	-	-

Таблица 2: Коэффициенты схемы (8) и кривизны (9) для  $S = 2$ .

$c_q$	$b_s$	$a_{sq}$
0	0	0 0
2	1	1/2 0
2	-	- -

Таблица 3: Коэффициенты схемы (8) и кривизны (9) для  $S = 3$ .

$c_q$	$b_s$	$a_{sq}$
2/3	2/9	0 0 0
-2	1/3	1/2 0 0
-8/3	4/9	0 3/4 -
4	-	- - -

Таблица 4: Коэффициенты схемы (8) и кривизны (9) для  $S = 4$ .

$c_q$	$b_s$	$a_{sq}$			
1	1/6	0	0	0	0
-2	1/3	1/2	0	0	0
-2	1/3	0	1/2	0	0
0	1/6	0	0	1	0
3	-	-	-	-	-

### 3 Расчеты с гарантированной точностью

Для определения фактической погрешности нужно проводить расчеты на многократно сгущающихся сетках и применять метод Ричардсона. Он позволяет найти не мажорантную оценку, а асимптотически точное значение погрешности [7]. Оно относится к четным узлам новой сетки (которые должны совпадать с узлами предыдущей сетки). Процедура сгущения хорошо разработана для равномерных и априорно построенных квазиравномерных сеток. Однако для адаптивных сеток, которые строятся в процессе расчета, требуется ее модификация.

Сгущение сеток будем производить в два этапа. На первом этапе шаг сетки выбирается по кривизне интегральной кривой. Сгущение сетки осуществляется за счет уменьшения управляющего параметра  $h_0$  в (4) вдвое. При таком сгущении четные узлы новой сетки стремятся к узлам предыдущей, но точного совпадения нет. Чтобы его обеспечить, на втором этапе сгущения будем дробить шаги предыдущей сетки по некоторому правилу, обеспечивающему квазиравномерность новой сетки. Тогда можно воспользоваться методом Ричардсона и найти апостериорное асимптотически точное значение погрешности.

Здесь возникает два основных вопроса: 1) в какой момент следует переходить от первого («автоматического») этапа ко второму («ручному») и 2) как следует проводить «ручное» сгущение сеток.

#### 3.1 Критерий установления сетки

Для перехода от первого этапа сгущения ко второму этапу необходимо, чтобы две последних сетки были достаточно близки. Понятие близости сеток не вполне строгое, и его можно реализовать разными способами.

Пусть  $l_n$  – узлы старой сетки с числом узлов  $N$ , а  $\hat{l}_{2n}$  – четные узлы новой сетки с числом узлов  $\hat{N}$ . Из-за того что шаг сетки выбирается автоматически, отношение  $\hat{N}/N \approx 2$  может оказаться не равным в точности двум.



Было рассмотрено несколько количественных критериев близости сеток. Опишем их, начиная с наиболее адекватного критерия.

1) Сетки тем лучше совпадают, чем меньше разности  $l_n - \hat{l}_{2n}$ . Из этих разностей можно составить сеточные нормы  $s$  и  $l_2$ . При этом разности целесообразно учитывать с разумно подобранными весами.

Удобный вес можно построить на основе обратной величины шага. Очевидно, если ошибка в положении узла превосходит величину шага, то ее следует считать большой. В качестве таких шагов целесообразно использовать шаги подробной сетки, примыкающие к узлу  $2n$ . Это приводит к следующим критериям.

$$\delta_c = \max_n |l_n - \hat{l}_{2n}| (\hat{h}_{2n}^{-1} + \hat{h}_{2n+1}^{-1}), \quad (10)$$

$$\delta_2 = \sqrt{\frac{1}{N-1} \sum_{n=1}^{N-1} (l_n - \hat{l}_{2n})^2 (\hat{h}_{2n}^{-1} + \hat{h}_{2n+1}^{-1})^2}. \quad (11)$$

Среднеквадратичная ошибка рассчитана на один узел. Обе погрешности (10), (11) безразмерны.

2) Этот критерий учитывает отношение шагов двух соседних сеток, точнее – расстояние между соответствующими узлами. Рассмотрим соседние узлы грубой сетки  $l_n, l_{n+1}$  и шаг между ними  $h_n = l_{n+1} - l_n$ . На подробной сетке им соответствуют узлы  $\hat{l}_{2n}, \hat{l}_{2n+2}$  и расстояние между ними  $\hat{l}_{2n+2} - \hat{l}_{2n} = \hat{h}_{2n} + \hat{h}_{2n+1}$ . Запишем следующий безразмерный критерий, использующий отношение соответствующих шагов:

$$\delta = \left[ \frac{1}{N-1} \sum_{n=1}^{N-1} (\sqrt{\xi_n} - 1/\sqrt{\xi_n})^2 \right]^{1/2}, \quad \xi = \frac{\hat{h}_{2n} + \hat{h}_{2n+1}}{h_n}. \quad (12)$$

По существу это уклонение соответствующих шагов подробной сетки от шага грубой сетки в расчете на один узел.

3) Этот критерий очень близок ко второму, если отношение шагов невелико. Но при большом отношении шагов он дает меньшее значение нормы, поэтому на задачах высокой жесткости он оказывается не столь эффективным.

$$\delta = \left[ \frac{1}{N-1} \sum_{n=1}^{N-1} \ln^2(\xi_n) \right]^{1/2}, \quad \xi = \frac{\hat{h}_{2n} + \hat{h}_{2n+1}}{h_n}. \quad (13)$$

4) Этот критерий представляет собой корень из среднего квадрата расстояний между соответствующими узлами сеток, нормированный на полную длину дуги  $L$ :

$$\delta = \frac{1}{\hat{l}_{2N}} \sqrt{\frac{1}{N} \sum_{n=1}^N (\hat{l}_{2n} - l_n)^2}. \quad (14)$$

Недостатком данного критерия является то, что в него вносит непропорционально малый вклад расстояние между узлами в начале сетки, так как ошибка с увеличением  $L$  и, соответственно, времени возрастает.

5) В этом критерии для каждой пары узлов отдельно производится нормировка на усредненную длину интегральной кривой в текущем узле:

$$\delta = \sqrt{\frac{1}{N} \sum_{n=1}^N \left( \frac{\hat{l}_{2n} - l_n}{(\hat{l}_{2n} + l_n)/2} \right)^2}. \quad (15)$$

Это позволило исправить указанный недостаток критерия (14).

Пусть в процессе сгущения последняя пара сеток оказалась достаточно близкой в смысле выбранного критерия. Тогда можно считать, что общая неравномерность расположения узлов установилась. Последнюю сетку будем считать квазиравномерной и перейдем ко второму этапу сгущения сеток.

### 3.2 Квазиравномерное сгущение

На втором этапе мы продолжаем сгущать сетку, полученную на первом этапе. Чтобы оценить погрешность с помощью метода Ричардсона, требуется точное совпадение четных узлов новой сетки с соответствующими узлами предыдущей. Будем последовательно квазиравномерно сгущать имеющуюся сетку, используя следующие формулы:

$$\hat{h}_0 = h_0 \frac{\sqrt{h_0}}{\sqrt{h_0} + \sqrt{h_1}}, \quad \hat{h}_1 = h_0 \frac{\sqrt{h_1}}{\sqrt{h_0} + \sqrt{h_1}}, \quad (16)$$

$$\hat{h}_{2n-1} = h_n \frac{\sqrt[4]{h_{n-1}}}{\sqrt[4]{h_{n-1}} + \sqrt[4]{h_{n+1}}}, \quad \hat{h}_{2n} = h_n \frac{\sqrt[4]{h_{n+1}}}{\sqrt[4]{h_{n-1}} + \sqrt[4]{h_{n+1}}}, \quad (17)$$

$$\hat{h}_{2N-1} = h_N \frac{\sqrt{h_{N-1}}}{\sqrt{h_{N-1}} + \sqrt{h_N}}, \quad \hat{h}_{2N} = h_N \frac{\sqrt{h_N}}{\sqrt{h_{N-1}} + \sqrt{h_N}}. \quad (18)$$

Для расчетов, полученных на новой сетке, можно применять метод оценки погрешности Ричардсона. Оценка погрешности вычисляется поточечно (в каждом четном узле новой сетки) для каждой компоненты

$$(\varepsilon_k)_{2n} = \frac{(\hat{u}_k)_{2n} - (u_k)_n}{2^p - 1}, \quad n = 1, \dots, N, \quad k = 0, \dots, K. \quad (19)$$

Здесь  $p$  – порядок точности схемы. Эта погрешность относится к расчетным значениям длины дуги.

Однако длина дуги не имеет непосредственного физического смысла и неудобна для пользователя. Пользователю обычно интереснее поведение

погрешности в зависимости от физического времени. Пересчет погрешности компонент  $u_1, \dots, u_K$  от аргумента  $l$  к аргументу  $t$  производится по следующей формуле:

$$(\epsilon_k)_{2n} = (\epsilon_k)_{2n} - (\epsilon_0)_{2n} f_k(\mathbf{u}, u_0), \quad k = 1, \dots, K. \quad (20)$$

### 3.3 Погрешность на первом этапе

Строго говоря, применять метод Рундсона к первому этапу нельзя. Во-первых, он требует сравнения решений на двух сетках при одинаковых значениях аргумента, а узлы  $l_n$  и  $\hat{l}_{2n}$  не совпадают. Во-вторых, схема состоит из разнородных элементов. Вычисление решения может проводиться по формулам высокого порядка точности. В то же время вычисление шага производится по кривизне, которая рассчитывается по формуле более низкого порядка точности.

Поэтому вычисление погрешности по разности решений на двух соседних сетках является не вполне адекватным. Зависимость такой погрешности от числа узлов сетки нередко оказывается достаточно сложной. Графики погрешности в двойном логарифмическом масштабе  $\lg \epsilon$  от  $\lg N$  могут заметно отличаться от прямых линий с некоторым наклоном. Если же в них отмечаются прямолинейные участки, то возможно аномальное появление высокого эффективного порядка точности на грубых сетках и малого на очень подробных.

Поэтому апостериорные оценки погрешности на первой стадии расчета недостаточно надежны, и алгоритм их нахождения трудно автоматизировать.

## 4 Верификация метода

### 4.1 Тестовая задача

Рассмотрим тестовую задачу

$$\frac{du}{dt} = f(u, t), \quad f(u, t) = -\lambda \cos t \frac{(u^2 - a^2)^2}{u^2 + a^2}, \quad \lambda, a = \text{const}. \quad (21)$$

Для нее нетрудно построить точное решение

$$u(t) = -\frac{2\Lambda(t)a^2}{1 + \sqrt{1 + 4a^2\Lambda(t)^2}}, \quad \Lambda(t) = \frac{u^0}{(u^0)^2 - a^2} + \lambda \sin t. \quad (22)$$

Оно представляет собой функцию с контрастными структурами, решение которой попеременно притягивается к верхнему и нижнему стационарам.

Пограничные слои имеют ширину порядка  $1/\lambda$ , на них производная имеет очень большое значение. Чем больше параметр  $\lambda$ , тем больше жесткость задачи.

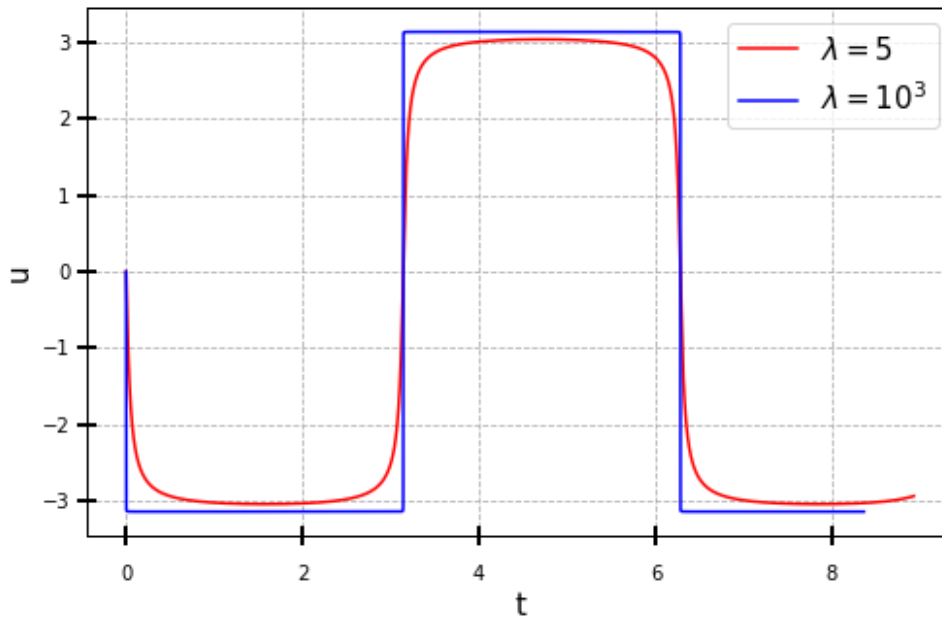


Рис. 1: Решение задачи (21) при  $\lambda = 5$  и  $\lambda = 10^3$ .

На рис. 1 представлена зависимость решения  $u$  от времени  $t$  для разных значений параметра  $\lambda$ .

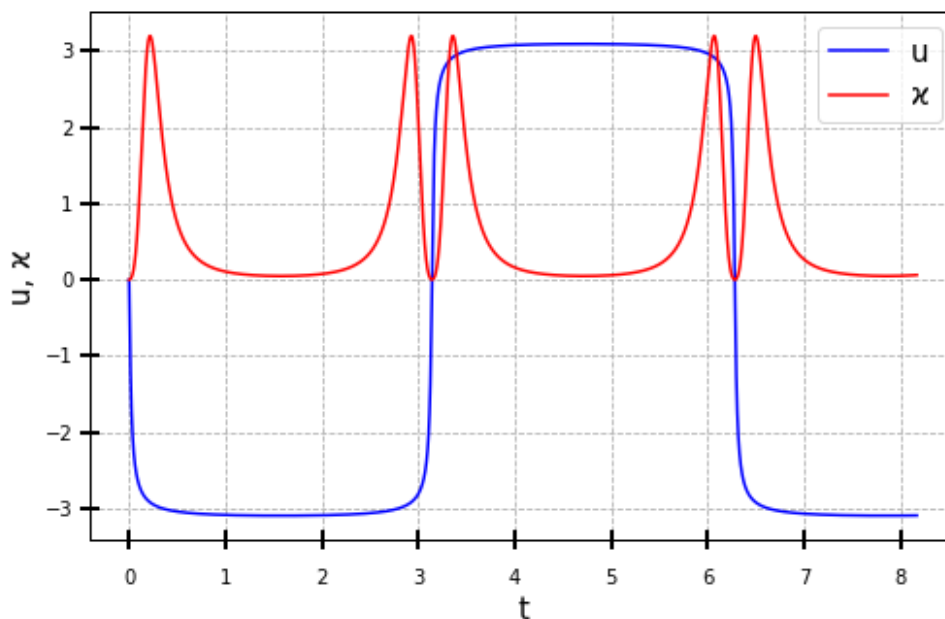


Рис. 2: Решение и кривизна в задаче (21), аргумент  $t$ ,  $\lambda = 10$ .

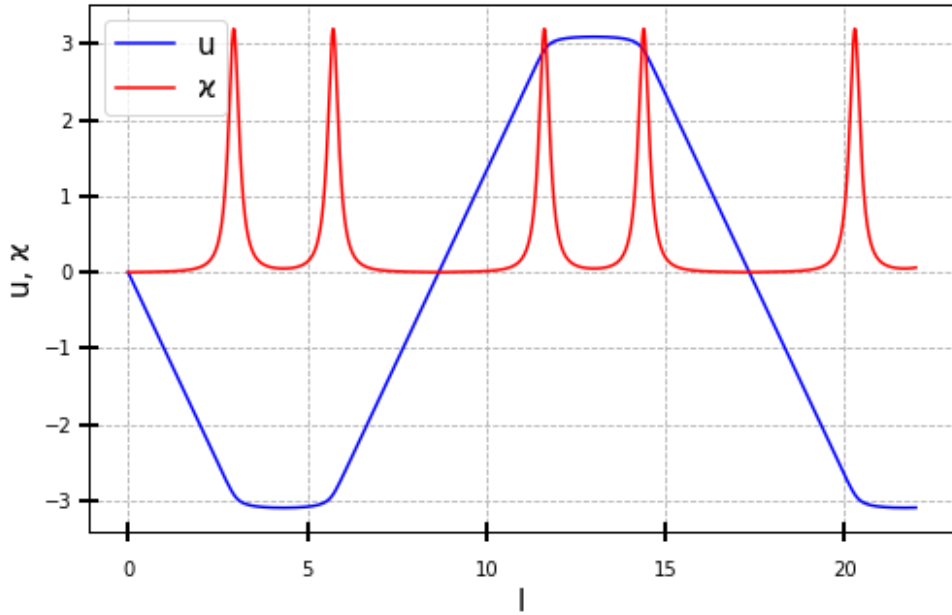


Рис. 3: Решение и кривизна в задаче (21), аргумент  $l$ ,  $\lambda = 10$ .

На рис. 2 представлены вид решения  $u$  и кривизны  $\kappa$  задачи с параметром  $\lambda = 10$  в аргументе «время»  $t$ , на рис. 3 – в аргументе «длина дуги»  $l$ .

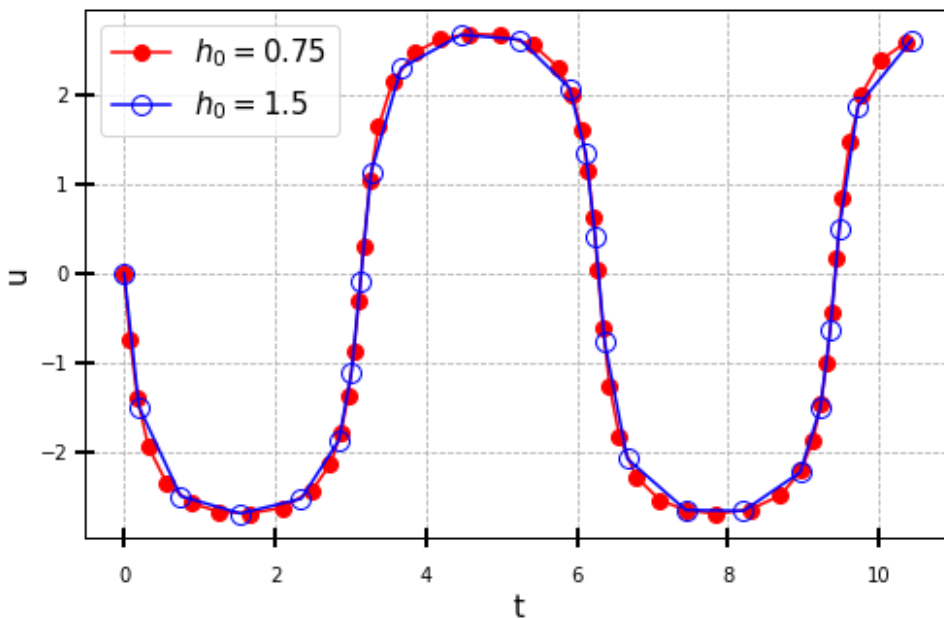


Рис. 4: Ступение сеток на первом этапе.

Видно, что пограничному слою с резким изменением решения в аргументе времени  $t$  соответствует почти прямолинейный участок с единичным наклоном в аргументе длины дуги  $l$ . Кривизна интегральной кривой  $\kappa$  оказывается большой только в переходных зонах, расположенных между

пограничными слоями и регулярными участками. Это означает, что мельчить шаг по длине дуги имеет смысл только в переходных зонах.

Будем выбирать шаг сетки по формуле (4), где кривизна интегральной кривой  $\kappa$  вычисляется по формуле (7), имеющей первый порядок точности. В дальнейшем будет показано, что такое вычисление кривизны не влияет на точность расчета.

На рис. 4 показаны два численных решения задачи с параметром  $\lambda = 1$  и различающимися в два раза длинами начальных шагов. Видно, что соответствующие узлы двух соседних сеток близки, но точного совпадения нет. Поэтому будем последовательно уменьшать начальный шаг в два раза, пока не выполнится условие  $\delta < \Delta$  в смысле какого-то из критериев разд. 3.1.

## 4.2 Результаты расчетов

Все последующие расчеты самого численного решения выполнены по схеме РК4 точности  $O(h^4)$ . Расчеты кривизны выполнялись почти всюду с точностью  $O(h)$ ; однако для контроля проводились расчеты, при которых кривизна вычислялась с точностью  $O(h^2)$ , и даже с точными значениями кривизны.

Кажется естественным, что при сгущении сетки критерий качества сетки как функция числа шагов  $N$  должен монотонно убывать. Соответствующая кривая в двойном логарифмическом масштабе должна выходить на прямую, наклон которой соответствует порядку точности формулы расчета кривизны.

Для кривой погрешности решения также естественно ожидать монотонного убывания, однако априори неясно, каков будет наклон асимптотической прямой: будет он соответствовать порядку точности основной схемы, по которой вычисляется решение, или порядку точности формулы, по которой вычисляется кривизна.

Однако численные расчеты показали, что описанная картина справедлива лишь при достаточно подробных сетках. На грубых сетках расположен нерегулярный участок кривой, не соответствующий теоретическим представлениям. Опишем результаты численных расчетов. Оправдывается только второе предположение о выходе на прямую с определенным наклоном. Первое предположение не подтверждается: начало кривой при  $N < 200$  оказалось возрастающей функцией. Это означает, что на грубых сетках поведение кривой нерегулярно.

**Критерий качества сетки.** На рис. 5 показана зависимость величины (12) от числа узлов. Каждая линия соответствует определенной жесткости задачи  $\lambda$ . На задачах малой жесткости ( $\lambda = 1$ ) зависимость критерия качества от  $N$  полностью соответствует теоретическим представлениям.

Расчетная кривая близка к прямой с наклоном  $-1$ . Первое звено удовлетворяет требованию монотонности, но его наклон несколько отличается от теоретического. Это является зарождением нерегулярного участка кривой.

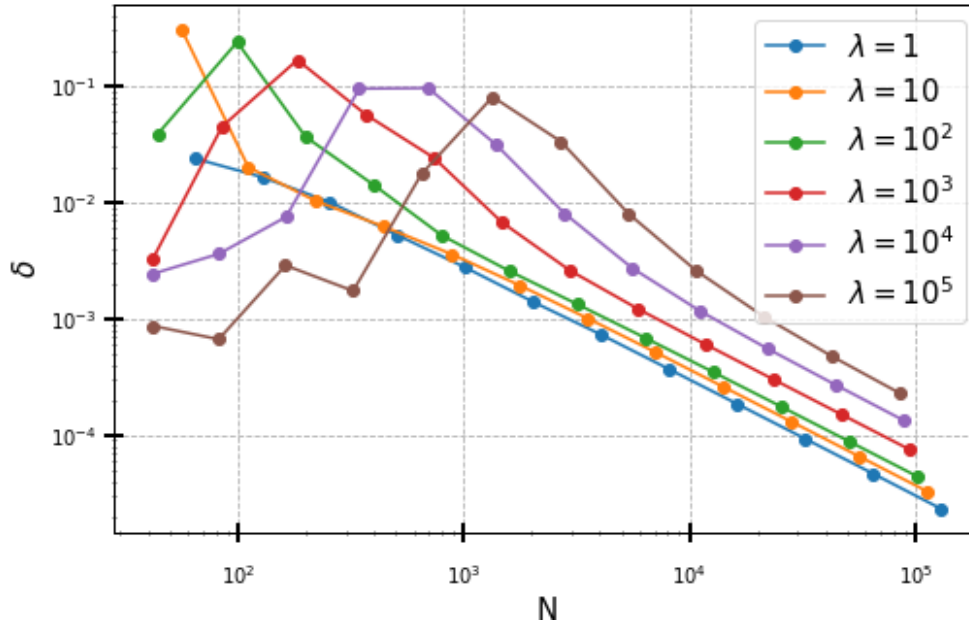


Рис. 5: Зависимость критерия качества (12) от числа интервалов сетки. Около каждой кривой указана величина жесткости  $\lg \lambda$ .

При еще большей жесткости  $\lambda = 10^2$  нерегулярный участок захватывает уже 3 первых звена. При этом первое звено кривой является нарастающим, то есть на нерегулярном участке нарушается монотонность критерия качества. При дальнейшем увеличении жесткости увеличивается как длина нерегулярного участка, так и отрезок нарушения монотонности. При этом на начальном участке возникает небольшая пилообразность.

Поясним причину возникновения возрастающего нерегулярного участка. В пределах расчетного интервала находятся 3 пограничных слоя. Для мягкой задачи сетка сразу же хорошо строится во всех этих трех слоях. Однако если жесткость велика, то картина выглядит иначе.

Погрешность численного решения увеличивается к концу промежутка интегрирования. Поэтому во втором пограничном слое погрешность гораздо больше, чем в первом (начальном). Накопление погрешности в решении приводит к накоплению ошибки при вычислении шага. Поэтому сетка хорошо строится в начальном пограничном слое, ощутимо хуже – во втором слое, а в третий слой узлы сетки могут и вовсе не попадать. Поэтому на грубых сетках разностная погрешность в третьем слое вообще не вычисляется.

По мере сгущения сетки узлы начинают попадать в третий слой и в них возникает значительная погрешность. Поэтому при первых сгущени-

ях погрешность увеличивается с увеличением числа интервалов сетки. Эти сетки образуют нерегулярный участок. При дальнейшем сгущении, когда в каждом пограничном слое становится достаточно много узлов, погрешность начинает уменьшаться в соответствии с теорией.

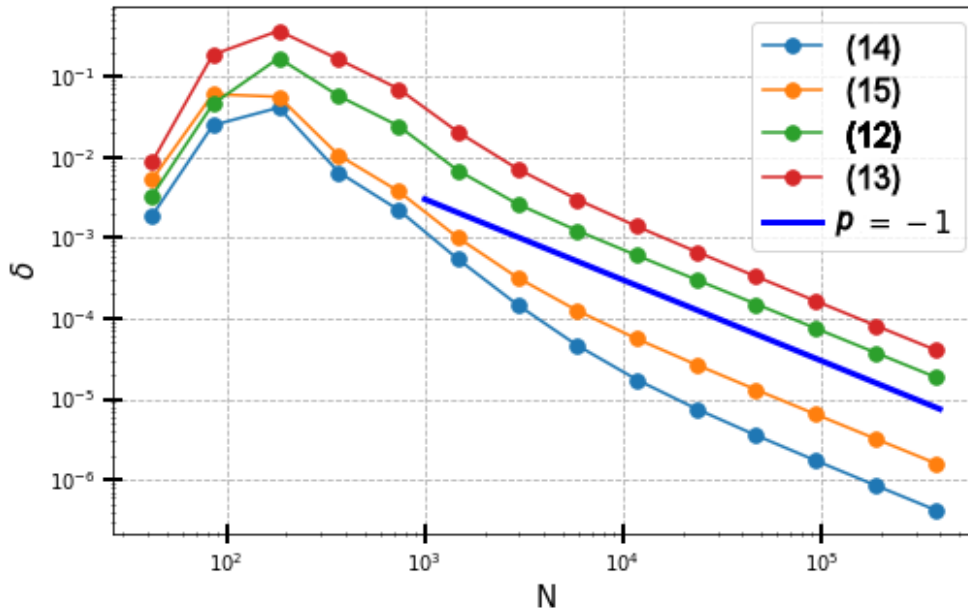


Рис. 6: Сравнение критериев качества (12), (13), (14), (15) при  $\lambda = 10^3$ . Сплошная линия – прямая с наклоном  $-1$ .

Для всех жесткостей переход на регулярный участок происходит при одновременном выполнении двух условий: 1) кривая монотонно убывающая (чтобы исключить начальную нерегулярность) и 2) величина критерия качества принимает значение  $\delta \leq 10^{-2}$  для любой жесткости. Это означает, что соответственные шаги соседних сеток в среднем отличаются на  $\sim 1\%$ , однако максимальное отличие шагов друг от друга может быть существенно больше.

Для других критериев качества поведение кривых имеет сходный вид (см. рис. 6), однако количественное значение перехода на регулярный участок будет своим для каждого критерия.

**Квазиравномерное сгущение.** Будем квазиравномерно сгущать последнюю сетку, полученную на первом этапе, используя формулы (16) – (18). Теперь четные узлы новой сетки точно совпадают с соответствующими узлами предыдущей сетки, и можно применять метод Ричардсона для оценки погрешности.

Проиллюстрируем второй этап на задаче с жесткостью  $\lambda = 10^3$ . На рис. 7 приведена зависимость от числа узлов оценки ошибки по методу Ричардсона, вычисленной по формулам (19) и (20), и фактической ошибки,



вычисленной непосредственным сравнением численного решения с точным решением (22). Вычислялись две нормы погрешности: нормы  $C$  и  $L_2$ .

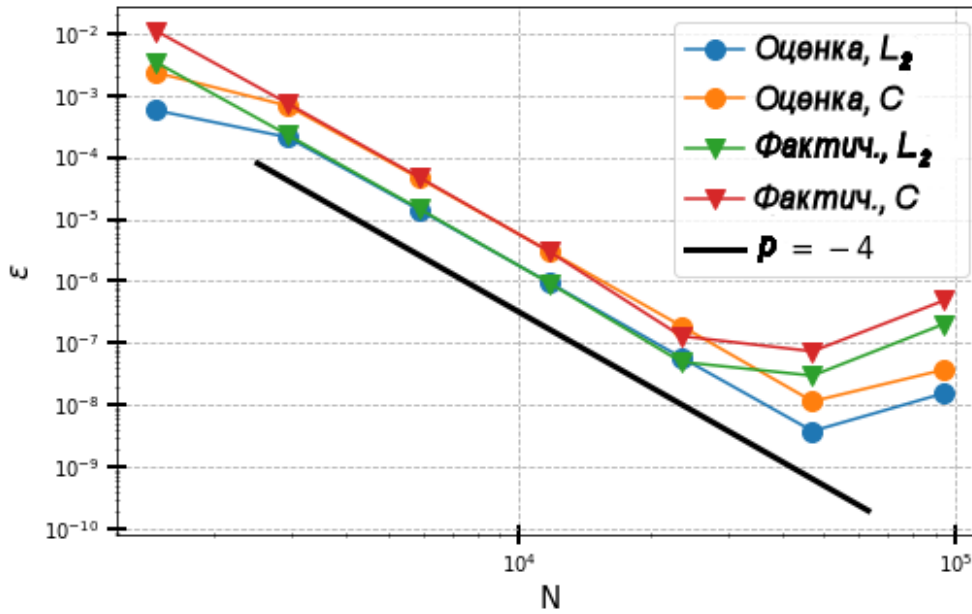


Рис. 7: Нормы погрешности для  $\lambda=10^3$ .

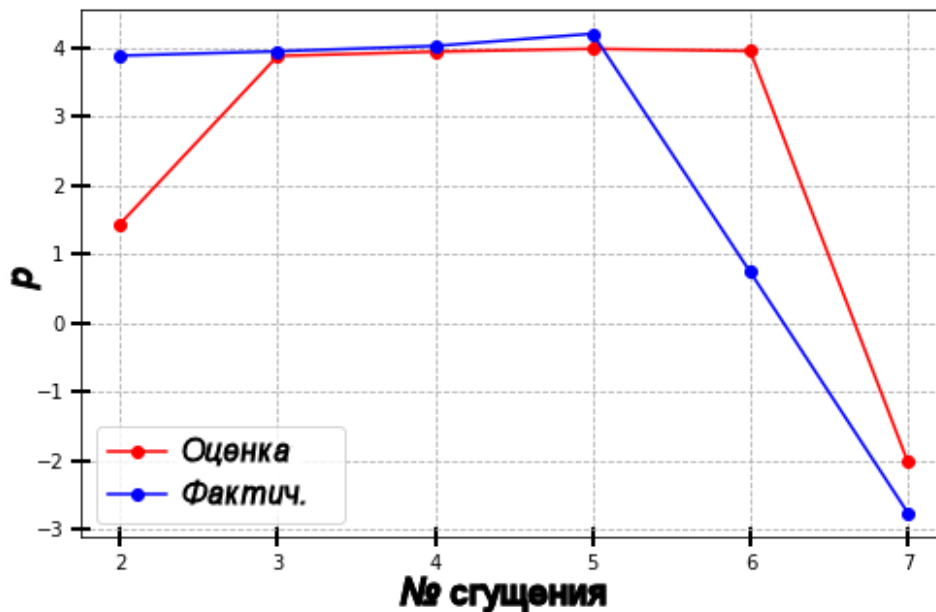


Рис. 8: Эффективные порядки точности для  $\lambda = 10^3$ . Зависимость эффективного порядка точности  $r$ , рассчитанного для реальной и оцененной ошибок, от номера сгущения.

Видно, что оценка ошибки по методу Ричардсона в обеих нормах практически совпадает с фактической погрешностью. Хорошо виден регуляр-

ный участок сходимости, имеющий в двойном логарифмическом масштабе коэффициент наклона  $-4$ , что совпадает с теоретическим порядком точности схемы ERK4. Он невелик, так как расчет быстро выходит на ошибки округления.

Также видно, что ошибка в нормах  $L_2$  и  $C$  отличается не более чем в 3 раза. Это означает, что ошибки в разных узлах сетки достаточно близки друг к другу. В самом деле, если бы во всех узлах величина ошибки была бы одинаковой, то нормы  $C$  и  $L_2$  этой ошибки точно бы совпали. Чем больше различие между минимальной и максимальной ошибкой, тем больше разница между указанными нормами. Это показывает, что исходная сетка для второго этапа хорошо сгущается в нужных местах.

На рис. 8 показаны эффективные порядки точности, соответствующие наклонам звеньев рис. 7. Видно, что эти порядки практически совпадают с теоретическим порядком точности схемы вплоть до момента выхода на ошибки округления.

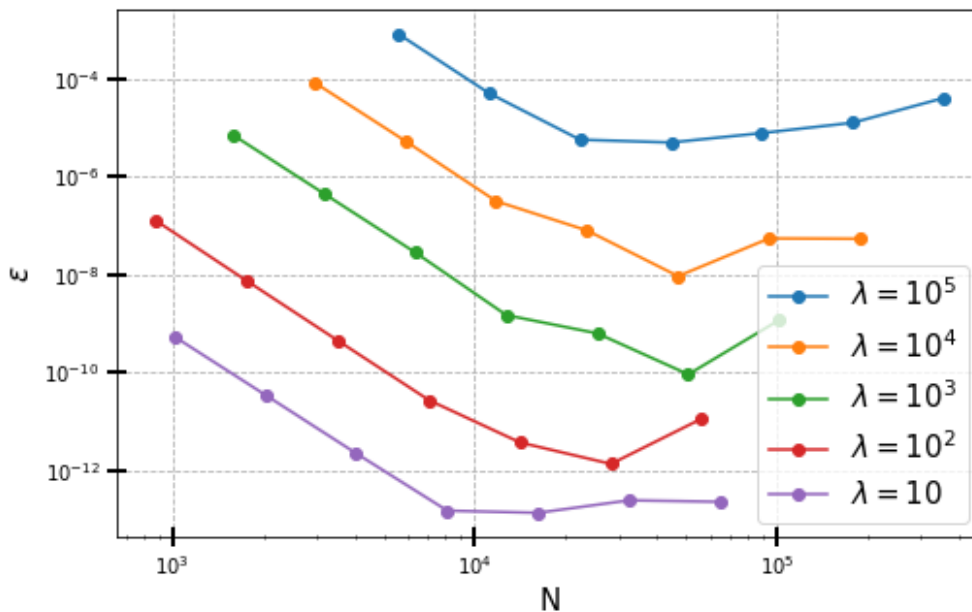


Рис. 9: Нормы погрешности  $L_2$  для разных  $\lambda$ .

На рис. 9 показаны графики погрешности в норме  $L_2$  для разных жесткостей  $\lambda$ . Видно, что чем больше жесткость  $\lambda$ , тем позже происходит переход ко второму этапу с квазиравномерным сгущением сеток. Кроме того, при увеличении жесткости задачи ухудшается количественная точность (но теоретический порядок точности сохраняется) и возрастает фон ошибок округления. Задача с  $\lambda = 10^5$  является очень жесткой, но даже она решается предложенным методом.

**Влияние способа вычисления кривизны.** Рассмотрим три способа вычисления кривизны, которая используется для расчета шага на пер-

вом этапе: 1) по разностной формуле (7) первого порядка точности, 2) по разностной формуле (9) второго порядка точности, 3) непосредственно по кривизне точного решения по следующей формуле:

$$\varkappa = \frac{f(u, t)f_u(u, t) + f_t(u, t)}{(1 + f(u, t)^2)^{3/2}},$$

$$f_u = -2\lambda \cos(\omega t) \frac{u(u^2 - a^2)(u^2 + 3a^2)}{(u^2 + a^2)^2}, \quad f_t = \lambda\omega \sin(\omega t) \frac{(u^2 - a^2)^2}{u^2 + a^2}. \quad (23)$$

Графики нормы  $L_2$  погрешности второго этапа для  $\lambda = 10^3$  в этих случаях представлены на рис. 10.

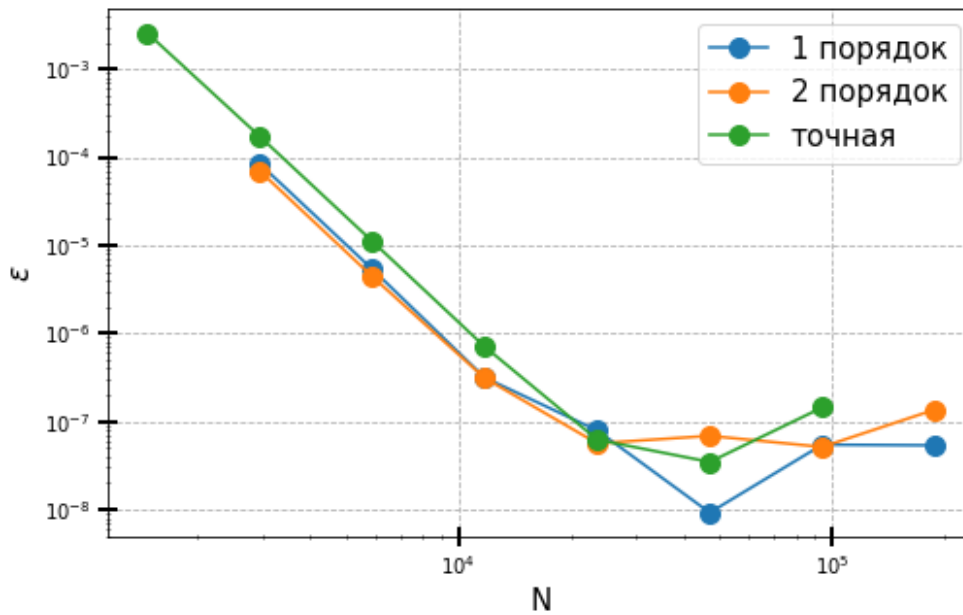


Рис. 10: Погрешности при  $\lambda = 10^3$  и разных способах вычисления кривизны.

Из рис. 10 видно, что все три способа дают примерно одинаковые результаты. Вероятная причина состоит в том, что кривизна является вспомогательной величиной для вычисления шага сетки, так что ее погрешность не сильно влияет на точность самого численного решения. Поэтому для вычисления кривизны вполне достаточно формул первого порядка точности, не говоря уже о втором.

**Сравнение с неявными схемами.** Одной из наиболее надежных схем четвертого порядка точности является чисто неявная оптимальная схема BORK [8], [9] из класса обратных схем Рунге-Кутты. В этой схеме кривизна рассчитывается по матрице Якоби численного решения, то есть она полностью согласована с численным решением. Напомним, что схема BORK требует итерационного уравнивания на новом слое и несравненно более трудоемка, чем явные схемы.

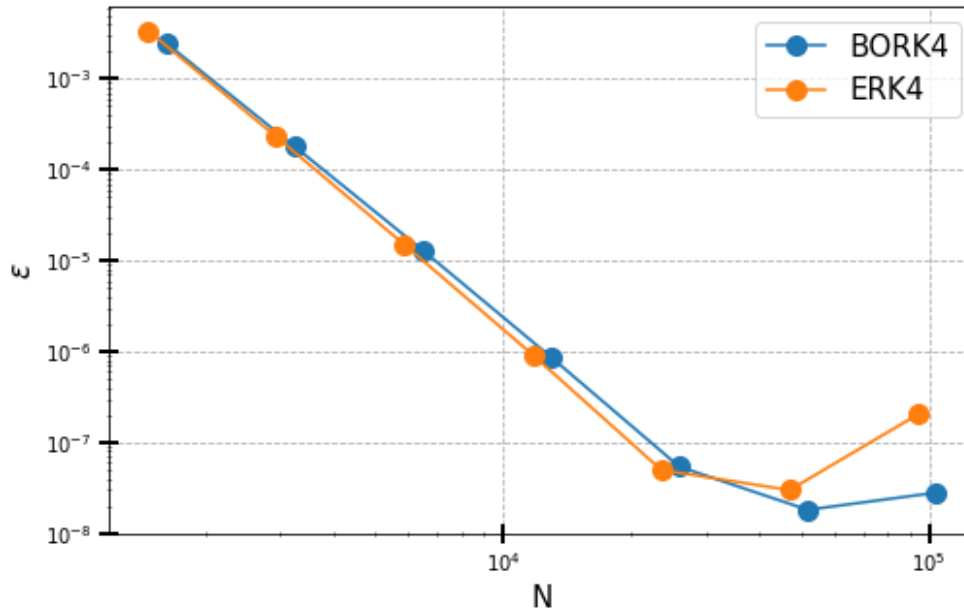


Рис. 11: Погрешности по схемам ERK4 и BORK4 при  $\lambda = 10^3$ .

Сравним расчеты по этой схеме с результатами по явной схеме ERK4, в которой кривизна на первом этапе берется по простейшим формулам первого порядка точности (7). На рис. 11 приведены графики погрешности второго этапа для этих схем в норме  $L_2$  для достаточно жесткой задачи с  $\lambda = 10^3$ . Эти кривые почти совпадают. Это показывает, что даже на задачах высокой жесткости схема ERK4 с выбором шага по кривизне сопоставима по точности и надежности с чисто неявными схемами, хотя ее трудоемкость несравненно меньше.

Конечно, на задачах сверхвысокой жесткости явные схемы даже при описанном выборе шага могут оказаться недостаточно надежными. Однако при сверхвысокой жесткости даже чисто неявные схемы тоже могут оказаться не вполне надежными: итерации в них не всегда сходятся.

**Выводы.** Все приведенные расчеты показывают конкурентоспособность предложенных в данной работе явных методов.

## Список литературы

1. Dahlquist G. Convergence and stability in the numerical integration of ordinary differential equations // Math. Scand. 1956. Vol. 4. P. 33–53.
2. Хайрер Э., Ваннер Г. Решение обыкновенных дифференциальных уравнений. Жесткие и дифференциально-алгебраические задачи. М.: Мир, 1999.

3. Dormand J. R., Prince P. J. A family of embedded Runge-Kutta formulae // J. Comp. Appl. Math. 1980. Vol. 6. P. 223–232.
4. Riks E. The application of Newton's method to the problem of elastic stability. // J. Appl. Mech. 1972. Vol. 39, №4. P. 1060–1065.
5. Белов А. А., Калиткин Н. Н., Пошивайло И. П. Геометрически-адаптивные сетки для жестких задач Коши // Докл. РАН. 2016. Т. 466. №3. С. 276–281.
6. Белов А. А., Калиткин Н. Н. Выбор шага по кривизне для жестких задач Коши // Матем. моделир. 2016. Т. 28. №11. С. 97–112.
7. Калиткин Н. Н., Альшин А. Б., Альшина Е. А., Рогов Б. В. Вычисления на квазиравномерных сетках, М.: Физматлит, 2005.
8. Калиткин Н. Н., Пошивайло И. П. Обратные Ls-устойчивые схемы Рунге-Кутты // Докл. РАН. 2012. Т.442. №2. С.175–180.
9. Калиткин Н. Н., Пошивайло И. П. Вычисления с использованием обратных схем Рунге-Кутты // Матем. моделир. 2013. Т. 25. №10. С.79–96.