



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

Препринты ИПМ • Препринт № 43 за 2025 г.



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

М.А. Бочев

О сходимости релаксации
формы волны для
нелинейных систем
обыкновенных
дифференциальных
уравнений

Статья доступна по лицензии
[Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/)



Рекомендуемая форма библиографической ссылки: Бочев М.А. О сходимости релаксации формы волны для нелинейных систем обыкновенных дифференциальных уравнений // Препринты ИПМ им. М.В.Келдыша. 2025. № 43. 36 с. EDN: [XGMKGE](https://library.keldysh.ru/preprint.asp?id=2025-43)
<https://library.keldysh.ru/preprint.asp?id=2025-43>

**ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
имени М.В. КЕЛДЫША
Российской академии наук**

М.А. Бочев

**О сходимости релаксации формы волны
для нелинейных систем обыкновенных
дифференциальных уравнений**

Москва — 2025

Бочев М.А.

О сходимости релаксации формы волны для нелинейных систем обыкновенных дифференциальных уравнений

Для интегрирования по времени больших систем нелинейных дифференциальных уравнений рассматривается вариант нелинейных итераций релаксации формы волны (также известных как динамические итерации или итерации Пикара–Линделёфа), где на каждой итерации требуется решить линейную неоднородную систему дифференциальных уравнений. Это делается специальным экспоненциальным блочным крыловским методом (ЭБК). Таким образом мы получаем двухуровневый итерационный процесс, где итерационные приближения определяются на определённом временном интервале, а не на слоях по времени, как в пошаговых методах. Такой подход недавно хорошо показал себя в рамках параллельного по времени метода PARAEXP. В данной работе сходимость двухуровневого метода оценивается теоретически и практически. Мы тестируем эффективность метода при решении нелинейных уравнений Бюргерса, трёхмерного уравнения Лиувилля–Брату–Гельфанда и трёхмерного нелинейного уравнения теплопроводности в сравнении с обычными неявными пошаговыми схемами интегрирования по времени.

Ключевые слова: релаксация формы волны, подпространство Крылова, экспоненциальные интеграторы, параллельность по времени, уравнение Бюргерса, уравнение Лиувилля–Брату–Гельфанда, нелинейное уравнение теплопроводности

Mikhail A. Botchev

On convergence of waveform relaxation for nonlinear systems of ordinary differential equations

To integrate large systems of nonlinear differential equations in time, we consider a variant of nonlinear waveform relaxation (also known as dynamic iteration or Picard–Lindelöf iteration), where at each iteration a linear inhomogeneous system of differential equations has to be solved. This is done by the exponential block Krylov subspace (EBK) method. Thus, we have an inner-outer iterative method, where iterative approximations are determined over a certain time interval, with no time stepping involved. This approach has recently been shown to be efficient as a time-parallel integrator within the PARAEXP framework. In this paper, convergence behavior of this method is assessed theoretically and practically. We examine efficiency of the method by testing it on nonlinear Burgers, three-dimensional Liouville–Bratu–Gelfand, and three-dimensional nonlinear heat conduction equations and comparing its performance with that of conventional time-stepping integrators.

Key words: waveform relaxation, Krylov subspaces, exponential time integrators, time-parallel methods, Burgers equation, Liouville–Bratu–Gelfand equation, nonlinear heat equation

Оглавление

1	Введение	3
2	Нелинейная релаксация формы волны и её сходимость	4
3	Численные эксперименты	15
4	Выводы	30
	Список литературы	32

1. Введение

Большие системы обыкновенных дифференциальных уравнений по времени возникают в различных приложениях и во многих случаях должны интегрироваться по времени неявными методами, см., например, [17; 27]. В последние десятилетия ниша неявных методов постепенно заполняется экспоненциальными методами [25]. Для неявных и экспоненциальных методов главный вопрос состоит в том, как эффективно решать возникающие нелинейные системы или как вычислять соответствующие матричные функции. Чтобы достичь эффективности, существуют и применяются различные подходы, такие как неточные методы Ньютона в сочетании с итерационными предобусловленными линейными решателями [8; 9; 46], методы расщепления и Розенброка [1; 10; 11; 27; 53] и приближённые неявные схемы, которые также могут рассматриваться как стабилизированные явные схемы [5; 7; 33; 35; 44; 45; 47; 51; 54].

Другой важный подход для достижения эффективности в неявных и экспоненциальных методах интегрирования по времени основан на релаксации формы волны [34; 41; 49], где итерационные приближения — это зависящие от времени функции, а не значения численного решения на шагах по времени. Эти методы также известны как динамические итерации или итерации Пикара–Линделёфа [38]. Они используются с восьмидесятих годов [28; 36; 37; 41; 52] и недавно получили развитие в связи с параллельным по времени экспоненциальным методом PARAEXP [20] и его обобщением на нелинейные задачи [30; 31].

Обычно действия матричных функций в экспоненциальных интеграторах вычисляются специальными итерационными процедурами линейной алгебры. Как правило, это методы подпространства Крылова или чебышёвские полиномиальные методы [15; 16; 19; 26; 42; 45; 48]. Привлекательное свойство итераций типа релаксации форма волны — это то, что они позволяют применять этот аппарат линейной алгебры сразу на определённом временном интервале, а не на каждом шаге по времени. Таким образом, вычислительные затраты распределяются по времени, что позволяет достичь как вычислительной эффективности, так и параллелизма по времени [49].

Один из успешных экспоненциальных методов этого класса — нелинейный экспоненциальный блочный крыловский (ЭБК) метод [30; 31]. Он использует проекции на блочные подпространства Крылова [2] в сочетании с релаксацией формы волны для того, чтобы работать с нелинейностью. Нелинейные ЭБК методы обладают параллелизмом по времени, и их параллельная реализация может рассматриваться как эффективный вариант параллельного по времени метода PARAEXP [30; 31].

Сходимость релаксации формы волны изучалась в различных ситуациях, например, для линейных задач Коши, для нелинейных методов Гаусса–Зейделя и Якоби (часто используемых в классических методах релаксации формы волны), а также для дискретных по времени формулировок (см. обзоры в работах [38; 49]). Тем не менее результатов по сходимости релаксации формы волны в общих нелинейных постановках не так много, так как обычно предполагалось, что *«...изучение линейного случая в деталях — это то, что действительно может требоваться пользователям»* [38]. Кроме книги [49], которая содержит некоторые базовые результаты по нелинейной сходимости, одна из немногих статей, где изучается сходимость релаксации формы волны, — статья [40]. Ниже (см. замечание 2) мы комментируем результаты, представленные там. Целью данной статьи является обобщение результатов по сходимости релаксации формы волны на нелинейную постановку, используемую в методе ЭБК. Результаты, представленные здесь, также позволяют получить лучшее понимание сходимости метода на практике.

Данная статья организована следующим образом. В разделе 2.1 формулируются постановка задачи и необходимые предположения. Основные результаты по сходимости представлены в разделе 2.2. Поскольку эти результаты сформулированы для функции ошибки, которая обычно неизвестна, в разделе 2.3 приводится оценка ошибки с помощью вычислимой нелинейной невязки. На каждой итерации релаксации формы волны должна быть решена линейная задача Коши, что в данной постановке выполняется итерационным блочным крыловским методом (линейным методом ЭБК). Поэтому в разделе 2.4 мы показываем, как неточное решение линейных задач Коши влияет на сходимость нелинейных итераций. В разделе 2.5 кратко обсуждаются вопросы реализации. Численные эксперименты представлены в разделе 3.1 (для одномерного уравнения Бюргерса), разделе 3.2 (для трёхмерного уравнения Лиувилля–Брату–Гельфанда) и разделе 3.3 (для трёхмерного нелинейного уравнения теплопроводности). Статья заканчивается выводами, сформулированными в разделе 4.

2. Нелинейная релаксация формы волны и её сходимость

2.1. Постановка задачи, предположения и обозначения. Рассматривается задача Коши для нелинейной системы обыкновенных дифференциальных урав-

нений

$$y'(t) = F(t, y(t)), \quad y(0) = v, \quad t \in [0, T], \quad (1)$$

где $F : \mathbb{R} \times \mathbb{R}^N \rightarrow \mathbb{R}^N$, $v \in \mathbb{R}^N$ и $T > 0$ заданы. Не нарушая общности, предполагается, что зависимость функции правой части F от времени имеет вид

$$F(t, y) = \bar{F}(y) + g(t), \quad \bar{F} : \mathbb{R}^N \rightarrow \mathbb{R}^N, \quad g : \mathbb{R} \rightarrow \mathbb{R}^N. \quad (2)$$

Если это не так, мы всегда можем преобразовать (1) к эквивалентному автономному виду (что обычно делается увеличением вектор-функции $y(t)$ на одну компоненту $y_{N+1}(t) = t$ и добавлением уравнения $y'_{N+1}(t) = 1$ к системе).

Итерационные методы, рассматриваемые в этой работе, основаны на семействе расщеплений

$$F(t, y) = -A_k y(t) + f_k(y(t)) + g(t), \quad \forall y \in \mathbb{R}^N, \quad t \geq 0, \quad (3)$$

где k — номер итерации, то есть матрица $A_k \in \mathbb{R}^{N \times N}$ и отображение $f_k : \mathbb{R}^N \rightarrow \mathbb{R}^N$ могут меняться от итерации к итерации. Предполагается, что все отображения f_k в (3) липшицевы с одной и той же константой Липшица L , т.е.

$$\forall k : \quad \|f_k(u) - f_k(v)\| \leq L \|u - v\|, \quad \forall u, v \in \mathbb{R}^N. \quad (4)$$

Здесь и повсюду в статье $\|\cdot\|$ обозначает векторную норму в $\mathbb{C}^N$ или соответствующую индуцированную матричную норму. Кроме того, мы предполагаем, что для всех матриц A_k из (3) существуют такие константы $C > 0$ и $\omega \geq 0$, что

$$\|\exp(-tA_k)\| \leq Ce^{-\omega t}, \quad t \geq 0. \quad (5)$$

Такая оценка выполняется для некоторой матрицы $A = A_k$ в евклидовой норме, например, если числовое поле матрицы A находится в левой комплексной полуплоскости $\{z = x + iy \mid x \geq 0, y \in \mathbb{R}, i^2 = -1\}$. В этом случае $C = 1$, а ω — наименьшее собственное число симметричной части $\frac{1}{2}(A + A^T)$ матрицы A . Условие (5) близко связано с понятием логарифмической матричной нормы [14; 27]. В частности, если $\mu(-A)$ — логарифмическая норма матрицы $-A$, то условие (5) выполняется для $C = 1$ и $t \geq 0$ тогда и только тогда, когда [27, теорема I.2.4]

$$\mu(-A) \leq -\omega.$$

Мы также используем функции φ_j [25, формула (2.10)], определяемых для $j = 0, 1, 2, \dots$ так:

$$\varphi_{j+1}(z) = \frac{\varphi_j(z) - \varphi_j(0)}{z}, \quad \varphi_0(z) = e^z. \quad (6)$$

Здесь полагается $\varphi_j(0) = 1/k!$, так что φ_j — целые функции. В этой работе существенно используется формула вариации произвольных постоянных (см., например, [27, глава I.2.3], [25, формула (1.5)])

$$y(t) = \exp(-tA_k)v + \int_0^t \exp(-(t-s)A_k) [f_k(y(s)) + g(s)] ds, \quad t \in [0, T], \quad (7)$$

где $y(t)$ — это решение задачи Коши (1),(3).

2.2. Нелинейная релаксация формы волны. Пусть $y_0(t)$ — некоторое приближение неизвестной функции $y(t)$ для $t \in [0, T]$, такое что $y_0(0) = v$. Обычно в качестве $y_0(t)$ можно взять $y_0(t) \equiv v$ для $t \in [0, T]$. Для решения задачи (1) в этой работе рассматриваются нелинейные итерации релаксации формы волны, где последовательно для $k = 0, 1, 2, \dots$ решается следующая задача Коши для линейной неоднородной системы уравнений:

$$y'_{k+1}(t) = -A_k y_{k+1}(t) + f_k(y_k(t)) + g(t), \quad y_{k+1}(0) = v, \quad t \in [0, T]. \quad (8)$$

Здесь матрицы $A_k \in \mathbb{R}^{N \times N}$ и отображения $f_k : \mathbb{R}^N \rightarrow \mathbb{R}^N$ определяют расщепление (3) и удовлетворяют предположениям (4),(5) для всех k . Подчеркнём, что приближение $y_{k+1}(t)$ на каждой итерации вычисляется и сохраняется для всего временного интервала $t \in [0, T]$. В разделе 2.5 мы кратко обсуждаем, как делать это эффективно.

Следующее утверждение даёт достаточное условие для сходимости итераций (8).

Утверждение 1. Пусть задача Коши (1) решается итерациями (8), где A_k и f_k удовлетворяют предположениям (4),(5) для $k = 0, 1, \dots$. Тогда ошибка $\epsilon_{k+1}(t) \equiv y(t) - y_{k+1}(t)$ итерационного приближения $y_{k+1}(t)$ удовлетворяет, для $k = 0, 1, \dots$,

$$\|\epsilon_{k+1}(t)\| \leq CLt\varphi_1(-\omega t) \max_{s \in [0, t]} \|\epsilon_k(s)\|, \quad \forall t \in [0, T], \quad (9)$$

а нелинейная релаксация формы волны (8) сходится для $t \in [0, T]$ к решению $y(t)$ задачи Коши (1),(3) при условии

$$CLT\varphi_1(-\omega T) < 1. \quad (10)$$

Доказательство. Вычитая итерационную формулу (8) из дифференциального уравнения $y'(t) = -Ay(t) + f(y(t)) + g(t)$ и принимая во внимание, что $y(0) = y_{k+1}(0) = v$, мы видим, что ошибка $\epsilon_{k+1}(t)$ удовлетворяет задаче Коши

$$\epsilon'_{k+1}(t) = -A_k \epsilon_{k+1}(t) + f_k(y(t)) - f_k(y_k(t)), \quad \epsilon_{k+1}(0) = 0. \quad (11)$$

Тогда, применяя формулу вариации произвольных постоянных (7) к задаче Коши (11), получаем

$$\epsilon_{k+1}(t) = \int_0^t \exp(-(t-s)A_k) [f_k(y(s)) - f_k(y_k(s))] ds.$$

Поэтому, используя (4) и (5), мы можем оценить

$$\begin{aligned} \|\epsilon_{k+1}(t)\| &\leq \int_0^t \|\exp(-(t-s)A_k)\| \|f_k(y(s)) - f_k(y_k(s))\| ds \\ &\leq \int_0^t \|\exp(-(t-s)A_k)\| L \|y(s) - y_k(s)\| ds \\ &\leq L \max_{s \in [0, t]} \|\epsilon_k(s)\| \int_0^t \|\exp(-(t-s)A_k)\| ds \\ &\leq CL \max_{s \in [0, t]} \|\epsilon_k(s)\| \int_0^t e^{-\omega(t-s)} ds = CLt\varphi_1(-\omega t) \max_{s \in [0, t]} \|\epsilon_k(s)\|. \end{aligned}$$

Таким образом, соотношение (9) доказано. Учитывая, что $t\varphi_1(-\omega t)$ — монотонно возрастающая функция от t , мы получаем (10). $\square$

Замечание 1. Утверждение 1 показывает, что мы можем выбрать длину T интервала по времени $[0, T]$ так, чтобы релаксация формы волны сходилась. Чем больше липшицева константа L , тем меньшая величина T может быть выбрана.

Замечание 2. В работе [40] задача Коши для автономной системы дифференциальных уравнений

$$C(y)y'(t) - \bar{F}(y(t)) = 0$$

решается релаксацией формы волны

$$y'_{k+1}(t) - \bar{F}(y_{k+1}(t)) = \hat{g}(y'_k(t)), \quad (12)$$

где $C(y)$ — матрица, такая что $C(y)y' = y' - \hat{g}(y')$. Анализ сходимости итераций (12) проводится в предположении, что $\hat{g}$ — липшицево отображение с константой Липшица, меньшей единицы. Кроме того, в [40] рассмотрен частный случай (12), где $\bar{F}(y) = -Ay + f(y)$, т.е.

$$y'_{k+1}(t) + Ay_{k+1}(t) - f(y_{k+1}(t)) = \hat{g}(y'_k(t)).$$

Таким образом, представленные здесь результаты не пересекаются с результатами [40].

В случае, когда отображение $f_k(y)$ линейное, релаксация формы волны (8) имеет привлекательное свойство суперлинейной сходимости для конечных T , см., например, [3, теорема 5.1] или, для неточной релаксации формы волны, [4, теорема 2]. Как показывает следующий результат, свойство суперлинейной сходимости распространяется и на нелинейный случай итераций (8).

Утверждение 2. Пусть задача Коши (1) решается итерациями (8), где A_k и f_k удовлетворяют предположениям (4), (5) для $k = 0, 1, \dots$. Тогда для ошибки $\epsilon_k(t)$ итерационного предположения $y_k(t)$ выполняется, для $k = 1, 2, \dots$,

$$\begin{aligned} \|\epsilon_k(t)\| &\leq (CL)^k t^k e^{-\omega t} \varphi_k(\omega t) \max_{s \in [0, t]} \|\epsilon_0(s)\| \\ &\leq (CL)^k \frac{t^k}{k!} \max_{s \in [0, t]} \|\epsilon_0(s)\|, \quad \forall t \in [0, T]. \end{aligned} \quad (13)$$

Доказательство. Доказательство очень близко к доказательству теоремы 5.1 из [3], где рассмотрен линейный аналог итераций (1), (8). Покажем справедливость оценки (13) по индукции. Заметим, что $t\varphi_1(-\omega t) = te^{-\omega t}\varphi_1(\omega t)$. Тогда оценка (13) для $k = 1$ совпадает с (9) для $k = 0$ и, следовательно, выполняется. Пусть (13) справедливо для некоторого k . Из доказательства утверждения 1 видим, что

$$\|\epsilon_{k+1}(t)\| \leq CL \int_0^t e^{-(t-s)\omega} \|\epsilon_k(s)\| ds,$$

так что по предположению индукции имеем

$$\begin{aligned} &\leq CL \int_0^t e^{-(t-s)\omega} (CL)^k s^k e^{-\omega s} \varphi_k(\omega s) \max_{\tilde{s} \in [0, s]} \|\epsilon_0(\tilde{s})\| ds \\ &= (CL)^{k+1} e^{-\omega t} \int_0^t s^k \varphi_k(\omega s) \max_{\tilde{s} \in [0, s]} \|\epsilon_0(\tilde{s})\| ds \\ &\leq (CL)^{k+1} e^{-\omega t} \max_{s \in [0, t]} \|\epsilon_0(s)\| \int_0^t s^k \varphi_k(\omega s) ds \\ &= (CL)^{k+1} t^{k+1} e^{-\omega t} \varphi_{k+1}(\omega t) \max_{s \in [0, t]} \|\epsilon_0(s)\|, \end{aligned}$$

где используется соотношение $\int_0^t s^k \varphi_k(\omega s) ds = t^{k+1} \varphi_{k+1}(\omega t)$. Шаг индукции выполнен. Наконец, используя определение функций φ_k и неравенство $\omega \geq 0$, нетрудно проверить, что

$$t^k e^{-\omega t} \varphi_k(\omega t) \leq \frac{t^k}{k!}, \quad t \geq 0.$$

Это доказывает вторую оценку в (13). □

2.3. Невязка нелинейной задачи. Чтобы контролировать на практике сходимость релаксации формы волны (8), естественно рассмотреть невязку $r_k(t)$ итерационного приближения $y_k(t)$ из соотношения (8). Поскольку $y_k(t)$ определяется формулами (8),(3) для $k - 1$, удобно определить невязку по отношению к задаче Коши (1) и расщепления (3) также для $k - 1$, т.е. для задачи Коши

$$y'(t) = -A_{k-1}y(t) + f_{k-1}(y(t)) + g(t), \quad y(0) = v, \quad t \in [0, T]. \quad (14)$$

Тогда невязка определяется соотношением

$$r_k(t) \equiv -A_{k-1}y_k(t) + f_{k-1}(y_k(t)) + g(t) - y'_k(t), \quad t \in [0, T], \quad (15)$$

и нетрудно увидеть, что

$$\begin{aligned} r_k(t) &= -A_{k-1}y_k(t) + f_{k-1}(y_k(t)) + g(t) - y'_k(t) \pm f_{k-1}(y_{k-1}(t)) \\ &= \underbrace{-A_{k-1}y_k(t) + f_{k-1}(y_{k-1}(t)) + g(t) - y'_k(t)}_{=0} + \\ &\quad f_{k-1}(y_k(t)) - f_{k-1}(y_{k-1}(t)) \\ &= f_{k-1}(y_k(t)) - f_{k-1}(y_{k-1}(t)), \quad t \in [0, T], \quad k = 1, 2, \dots \end{aligned} \quad (16)$$

Заметим, что определённая таким образом невязка $r_k(t)$ обладает свойством, требуемым в обратном анализе ошибок: итерационное приближение $y_k(t)$ является точным решением возмущённой задачи Коши (14), где возмущением является невязка $r_k(t)$:

$$y'_k(t) = -A_{k-1}y_k(t) + f_{k-1}(y_k(t)) + g(t) - r_k(t), \quad y_k(0) = v \quad t \in [0, T]. \quad (17)$$

Вычитая уравнение этой задачи Коши из уравнения задачи Коши (14), получим задачу Коши для ошибки $\epsilon_k(t)$:

$$\epsilon'_k(t) = -A_{k-1}\epsilon_k(t) + f_{k-1}(y(t)) - f_{k-1}(y_k(t)) + r_k(t), \quad \epsilon_k(0) = 0. \quad (18)$$

Эта задача эквивалентна задаче Коши (11). Следующее утверждение показывает, что невязку можно рассматривать как верхнюю оценку ошибки.

Утверждение 3. Пусть задача Коши (1) решается итерациями (8), где A_k и f_k удовлетворяют предположениям (4),(5) для $k = 0, 1, \dots$. Если $T > 0$ выбрано так, что достаточное условие сходимости (10) итераций (8) выполнено, т.е.

$$CLT\varphi_1(-\omega T) \leq \delta < 1 \quad (19)$$

для некоторой константы $\delta \in (0, 1)$, то для всех $t \in [0, T]$ справедлива оценка

$$\max_{s \in [0, t]} \|\epsilon_k(s)\| \leq \frac{Ct\varphi_1(-\omega t)}{1 - CLt\varphi_1(-\omega t)} \max_{s \in [0, t]} \|r_k(s)\| \leq \frac{\delta}{(1 - \delta)L} \max_{s \in [0, t]} \|r_k(s)\|. \quad (20)$$

Заметим, что $T\varphi_1(-\omega T)$ — монотонно возрастающая функция от T .

Доказательство. Применяя формулу вариации произвольных постоянных к задаче (18), можно оценить

$$\begin{aligned}
 \|\epsilon_k(t)\| &\leq \int_0^t \|\exp(-(t-s)A_{k-1}) [f_{k-1}(y(s)) - f_{k-1}(y_k(s)) + r_k(s)]\| ds \\
 &\leq \max_{s \in [0,t]} \|f_{k-1}(y(s)) - f_{k-1}(y_k(s)) + r_k(s)\| \int_0^t \|\exp(-(t-s)A_{k-1})\| ds \\
 &\leq \max_{s \in [0,t]} \|f_{k-1}(y(s)) - f_{k-1}(y_k(s)) + r_k(s)\| Ct\varphi_1(-\omega t) \\
 &\leq \left(L \max_{s \in [0,t]} \|\epsilon_k(s)\| + \max_{s \in [0,t]} \|r_k(s)\| \right) Ct\varphi_1(-\omega t) \\
 &\leq CLt\varphi_1(-\omega t) \max_{s \in [0,t]} \|\epsilon_k(s)\| + Ct\varphi_1(-\omega t) \max_{s \in [0,t]} \|r_k(s)\|.
 \end{aligned}$$

Учитывая, что $CLt\varphi_1(-\omega t) \leq CLT\varphi_1(-\omega T) \leq \delta < 1$ для $t \in [0, T]$, получаем

$$\begin{aligned}
 (1 - CLt\varphi_1(-\omega t)) \max_{s \in [0,t]} \|\epsilon_k(s)\| &\leq Ct\varphi_1(-\omega t) \max_{s \in [0,t]} \|r_k(s)\| \\
 \max_{s \in [0,t]} \|\epsilon_k(s)\| &\leq \frac{Ct\varphi_1(-\omega t)}{1 - CLt\varphi_1(-\omega t)} \max_{s \in [0,t]} \|r_k(s)\|,
 \end{aligned}$$

откуда следует (20). □

Как показывает утверждение, приводимое ниже, оценку сверху на нелинейную невязку можно получить с помощью ошибки на предыдущей итерации.

Утверждение 4. Пусть задача Коши (1) решается итерациями (8), где A_k и f_k удовлетворяют предположениям (4),(5) для $k = 0, 1, \dots$. Тогда для $k = 1, 2, \dots$

$$\|r_k(t)\| \leq (1 + CLt\varphi_1(-\omega t))L \max_{s \in [0,t]} \|\epsilon_{k-1}(s)\|, \quad \forall t \in [0, T]. \quad (21)$$

Кроме того, если выполняется условие сходимости (19), то справедлива оценка

$$\|r_k(t)\| \leq (1 + \delta)L \max_{s \in [0,t]} \|\epsilon_{k-1}(s)\| < 2L \max_{s \in [0,t]} \|\epsilon_{k-1}(s)\|, \quad \forall t \in [0, T]. \quad (22)$$

Доказательство. Используя соотношение (16), для $t \in [0, T]$ получаем оценку

$$\begin{aligned}
 r_k(t) &= f_{k-1}(y_k(t)) - f_{k-1}(y(t)) + f_{k-1}(y(t)) - f_{k-1}(y_{k-1}(t)), \\
 r_k(t) &\leq \|f_{k-1}(y_k(t)) - f_{k-1}(y(t))\| + \|f_{k-1}(y(t)) - f_{k-1}(y_{k-1}(t))\| \\
 &\leq L\|y_k(t) - y(t)\| + L\|y(t) - y_{k-1}(t)\| = L(\|e_k(t)\| + \|e_{k-1}(t)\|) \\
 &\leq L \left(CLt\varphi_1(-\omega t) \max_{s \in [0,t]} \|e_{k-1}(s)\| + \|e_{k-1}(t)\| \right),
 \end{aligned}$$

откуда следуют неравенства (21),(22). □

2.4. Внутренние линейные итерации и их невязка. На практике линейную задачу Коши (8), возникающую на каждой итерации релаксации формы волны, можно решить любым удобным численным методом с определённой точностью. В рамках параллельной по времени схемы PARAEXP рассматриваемая нелинейная итерация формы волны показала себя эффективной для задач гидродинамики [30; 31]. При этом линейная задача Коши (8) решалась экспоненциальным блочным крыловским методом (ЭБК) [2]. В этом случае мы имеем двухуровневый итерационный процесс, где на каждой внешней нелинейной итерации релаксации формы волны (8) запускаются внутренние блочно-крыловские итерации $\tilde{y}_{k+1,\ell}(t) \rightarrow y_{k+1}(t)$, где ℓ — номер внутренних итераций. Для простоты обозначений опустим зависимость от ℓ в $\tilde{y}_{k+1,\ell}(t)$ и будем писать $\tilde{y}_{k+1}(t)$. Сходимость внутренних итераций метода ЭБК контролируется по норме «внутренней» линейной невязки, которая определяется для $\tilde{y}_{k+1}(t) \approx y_{k+1}(t)$ так:

$$r_{k+1}^{\text{lin}}(t) \equiv -A_k \tilde{y}_{k+1}(t) + f_k(\tilde{y}_k(t)) + g(t) - \tilde{y}'_{k+1}(t), \quad t \in [0, T]. \quad (23)$$

Здесь мы, предполагая, что предыдущее итерационное приближение $y_k(t)$ также вычислено неточно, заменили $y_k(t)$ на $\tilde{y}_k(t)$. Внутренние итерации останавливаются, как только норма линейной невязки $r_{k+1}^{\text{lin}}(t)$, определённой соотношением (23), достаточно мала. Линейная невязка $r_{k+1}^{\text{lin}}(t)$ легко вычислима [2] и может быть использована для контроля ошибки в линейной задаче Коши (8), см. [2, формула (16)].

Следующий результат показывает, что сходимость нелинейной релаксации формы волны сохраняется, если задача Коши (8) на каждой внешней итерации решается приближённо, так что невязка (23) достаточно мала по норме. Для формулировки утверждения, приводимого ниже, удобно предположить, что норма невязки не превосходит нормы ошибки, достигнутой на предыдущей нелинейной итерации, помноженной на определённую константу, см. соотношение (24).

Утверждение 5. Пусть задача Коши (1) решается итерациями (8), где A_k и f_k удовлетворяют предположениям (4), (5) и выполняется условие сходимости (19). Пусть также задача Коши (8) на каждой нелинейной итерации $k = 0, 1, \dots$ решается приближённо ℓ внутренними итерациями, так что для невязки (23) выполняется одна из двух следующих оценок:

$$\begin{aligned} \text{(a)} \quad & \|r_{k+1}^{\text{lin}}(t)\| \leq \eta \max_{s \in [0, t]} \|f_k(\tilde{y}_{k+1}(s)) - f_k(\tilde{y}_k(s))\|, \quad t \in [0, T], \\ \text{(b)} \quad & \|r_{k+1}^{\text{lin}}(t)\| \leq \eta L \max_{s \in [0, t]} \|\tilde{y}_{k+1}(s) - \tilde{y}_k(s)\|, \quad t \in [0, T], \end{aligned} \quad (24)$$

с константой $\eta \in (0, 1)$. Тогда для ошибки $\tilde{e}_{k+1}(t) \equiv y(t) - \tilde{y}_{k+1}(t)$ приближённого решения $\tilde{y}_{k+1}(t)$ двухуровневых итераций (8), (23) выполняется, для

$k = 0, 1, \dots,$

$$\|\tilde{\epsilon}_{k+1}(t)\| \leq \frac{(1 + \eta)CLt\varphi_1(-\omega t)}{1 - \eta CLt\varphi_1(-\omega t)} \max_{s \in [0, t]} \|\tilde{\epsilon}_k(s)\|, \quad \forall t \in [0, T]. \quad (25)$$

Кроме того, двухуровневый итерационный процесс (8),(23) сходится для $t \in [0, T]$ к решению $y(t)$ задачи Коши (1),(3) при условии

$$\eta < \frac{1 - \delta}{2\delta}. \quad (26)$$

Доказательство. Заметим, что в силу (4)

$$\|f_k(\tilde{y}_{k+1}(t)) - f_k(\tilde{y}_k(t))\| \leq L\|\tilde{y}_{k+1}(t) - \tilde{y}_k(t)\|,$$

и, следовательно, условие (24),(b) следует из условия (24),(a). Поэтому, не теряя общности, можем предположить, что выполняется условие (24),(a). Из соотношения (23) следует, что $\tilde{y}_{k+1}(t)$ является решением возмущённой системы дифференциальных уравнений

$$\tilde{y}'_{k+1}(t) = -A_k \tilde{y}_{k+1}(t) + f_k(\tilde{y}_k(t)) + g(t) - r_{k+1}^{\text{lin}}(t). \quad (27)$$

Вычитая это соотношение из исходной нелинейной системы $y'(t) = -A_k y(t) + f_k(y(t)) + g(t)$ и учитывая, что $\tilde{y}_{k+1}(0) = y(0) = v$, мы получаем задачу Коши для ошибки $\tilde{\epsilon}_{k+1}(t) \equiv y(t) - \tilde{y}_{k+1}(t)$ (см. соотношение (11)):

$$\tilde{\epsilon}'_{k+1}(t) = -A_k \tilde{\epsilon}_{k+1}(t) + f_k(y(t)) - f_k(\tilde{y}_k(t)) + r_{k+1}^{\text{lin}}(t), \quad \tilde{\epsilon}_{k+1}(0) = 0.$$

Применение формулы вариации произвольных постоянных (7) к этой задаче Коши приводит к соотношению

$$\tilde{\epsilon}_{k+1}(t) = \int_0^t \exp(-(t-s)A_k) [f_k(y(s)) - f_k(\tilde{y}_k(s)) + r_{k+1}^{\text{lin}}(s)] ds.$$

Тогда, используя (4),(5),(24), для $t \in [0, T]$ получаем

$$\begin{aligned} \|\tilde{\epsilon}_{k+1}(t)\| &\leq \int_0^t \|\exp(-(t-s)A_k)\| \|f_k(y(s)) - f_k(\tilde{y}_k(s)) + r_{k+1}^{\text{lin}}(s)\| ds \\ &\leq Ct\varphi_1(-\omega t) \left[L \max_{s \in [0, t]} \|\tilde{\epsilon}_k(s)\| + \max_{s \in [0, t]} \|r_{k+1}^{\text{lin}}(s)\| \right] \\ &\leq Ct\varphi_1(-\omega t) \left[L \max_{s \in [0, t]} \|\tilde{\epsilon}_k(s)\| + \eta L \max_{s \in [0, t]} \|\tilde{y}_{k+1}(s) - \tilde{y}_k(s)\| \right]. \end{aligned}$$

Далее, оценим

$$\max_{s \in [0, t]} \|\tilde{y}_{k+1}(s) - \tilde{y}_k(s) \pm y(s)\| \leq \max_{s \in [0, t]} \|\tilde{\epsilon}_{k+1}(s)\| + \max_{s \in [0, t]} \|\tilde{\epsilon}_k(s)\|,$$

так что для $t \in [0, T]$ выполняется

$$\begin{aligned} \|\tilde{\epsilon}_{k+1}(t)\| &\leq CLt\varphi_1(-\omega t) \left[(1 + \eta) \max_{s \in [0, t]} \|\tilde{\epsilon}_k(s)\| + \eta \max_{s \in [0, t]} \|\tilde{\epsilon}_{k+1}(s)\| \right], \\ (1 - CLt\varphi_1(-\omega t)\eta) \max_{s \in [0, t]} \|\tilde{\epsilon}_{k+1}(s)\| &\leq CLt\varphi_1(-\omega t)(1 + \eta) \max_{s \in [0, t]} \|\tilde{\epsilon}_k(s)\|. \end{aligned}$$

Принимая во внимание, что $CLt\varphi_1(-\omega t) \leq CLT\varphi_1(-\omega T) \leq \delta < 1$ и $0 < \eta < 1$, получаем

$$\max_{s \in [0, t]} \|\tilde{\epsilon}_{k+1}(s)\| \leq \frac{CLt\varphi_1(-\omega t)(1 + \eta)}{1 - CLt\varphi_1(-\omega t)\eta} \max_{s \in [0, t]} \|\tilde{\epsilon}_k(s)\| \leq \frac{\delta(1 + \eta)}{1 - \delta\eta} \max_{s \in [0, t]} \|\tilde{\epsilon}_k(s)\|.$$

Соотношение (25) доказано. Условие (26) эквивалентно требованию $\frac{\delta(1+\eta)}{1-\delta\eta} < 1$. $\square$

Замечание 3. Заметим, что при $\eta \rightarrow 0$ полученное условие сходимости (26) преобразуется в (19). Вообще говоря, условие (24), (b), вероятно, легче проверить на практике, чем (24), (a). Однако в приложениях вполне могут встречаться задачи Коши, где условие (24), (a) легко вычислимо. Поскольку $\tilde{y}_k(t)$ и $\tilde{y}_{k+1}(t)$ являются решениями задач Коши с одинаковым начальным вектором v , можно ожидать, что $\max_{s \in [0, t]} \|\tilde{y}_{k+1}(s) - \tilde{y}_k(s)\| \approx \|\tilde{y}_{k+1}(t) - \tilde{y}_k(t)\|$.

Тем не менее, применение условия (24) на практике не представляется необходимым: как показывают численные тесты, представленные ниже, простые остановочные критерии (см. (33), (40)), основанные на малости $\|r_{k+1}^{\text{lin}}(t)\|$, хорошо работают на практике и дают эффективный двухуровневый итерационный метод.

2.5. Реализация нелинейной релаксации формы волны. Зададим функцию начального приближения $y_0(t)$ равной для всех $t \in [0, T]$ вектору начального значения: $y_0(t) \equiv v$. Затем на каждой итерации (8) мы полагаем $\tilde{g}(t) := f_k(y_k(t)) + g(t)$ и решаем линейную задачу Коши

$$y'_{k+1}(t) = -A_k y_{k+1}(t) + \tilde{g}(t), \quad y_{k+1}(0) = v, \quad t \in [0, T]. \quad (28)$$

Поскольку эта задача решается методом ЭБК [2], решение $y_{k+1}(t)$ на каждой итерации вычисляется как линейная комбинация базисных векторов подпространства Крылова, хранящихся как столбцы матрицы $V_{\ell \cdot m} \in \mathbb{R}^{N \times \ell m}$, где ℓ — число внутренних итераций. Поэтому получаем

$$y_{k+1}(t) = V_{\ell \cdot m} u_{k+1}(t), \quad t \in [0, T], \quad u_{k+1} : \mathbb{R} \rightarrow \mathbb{R}^{\ell \cdot m},$$

где $u_{k+1}(t)$ — решение проектированной задачи Коши малого размера (см. подробности в [2]). Это позволяет хранить $y_{k+1}(t)$ в компактном виде для всех $t \in [0, T]$. На каждой внешней итерации k внутренние итерации метода ЭБК выполняются до тех пор, пока невязка (23) не станет достаточно малой по норме. Норму невязки легко вычислима и контролируется для нескольких значений $t \in [0, T]$.

Начиная со второй итерации $k = 1$, прежде чем решается задача (28), вектор-функция $\tilde{g}(y_k(t)) : \mathbb{R} \rightarrow \mathbb{R}^N$ вычисляется в n_s точках $t_1, \dots, t_{n_s}$ покрывающих весь временной интервал $[0, T]$. Обычно разумно выбрать $t_1 = 0, t_{n_s} = T$, а остальные t_j — корнями полинома Чебышёва

$$t_j = \frac{T}{2} \left(1 - \cos \frac{\pi(j - 3/2)}{s - 2} \right), \quad j = 2, \dots, s - 1.$$

Число образцов n_s следует выбирать так, чтобы функция $\tilde{g}(y_k(t))$ достаточно точно аппроксимировалась своей линейной интерполяцией в $t_1, \dots, t_{n_s}$. Затем вычисленные значения $\tilde{g}(t_j) \in \mathbb{R}^N, j = 1, \dots, n_s$ сохраняются как столбцы матрицы $\tilde{G} \in \mathbb{R}^{N \times s}$, и вычисляется «тощее» сингулярное разложение (thin SVD) матрицы $\tilde{G}$, дающее малоранговое представление

$$\tilde{g}(t) \approx Up(t), \quad U \in \mathbb{R}^{N \times m}, \quad t \in [0, T], \quad (29)$$

где обычно $m < 10$. Подробности получения этого малорангового представления можно найти в [2] или в [30]. Как правило, вычислительные затраты для получения представления $\tilde{g}(t) \approx Up(t)$ в (29) пренебрежимо малы по сравнению с другими вычислительными затратами в методе ЭБК. Внешние итерации (8) останавливаются, как только нелинейная невязка (16) мала по норме. Как показывает наш (ограниченный) опыт, на практике достаточно проверить норму невязки в конечный момент времени $t = T$. Получаемый таким образом алгоритм представлен ниже.

Алгоритм 1

Дано: задача Коши (1), вектор начального значения v , допустимая точность tol , размер блока m , число образцов n_s .

Результат: численное решение $y(T)$ задачи Коши (1).

Инициализация нелинейных итераций:

$y_0(t) := v$ (для всех t)

for $k = 0, 1, 2, \dots$ (главный итерационный цикл)

 if ($k = 0$) then

$\text{resnorm} := \|r_0(T)\| = \|\Phi(T, v)\|$

 else

$\text{resnorm} := \|r_k(T)\| = \|f_{k-1}(y_k(T)) - f_{k-1}(y_{k-1}(T))\|$

 endif

 if resnorm достаточно мала (см. (32),(39)) then

 остановить итерации

 endif

 определить $f_k(y)$, A_k , $\tilde{g}(t) := f_k(y_k(t)) + g(t)$

 задать допустимую точность внутренних итераций tol_{lin} (см. (33),(40))

 вычислить малоранговое приближение (29), $\tilde{g}(t) \approx Up(t)$

 решить задачу Коши (28) внутренними итерациями линейного метода ЭБК

$\Rightarrow$ получить $y_{k+1}(t) = V_{\ell,m} u_{k+1}(t)$, $t \in [0, T]$

endfor (главный итерационный цикл)

3. Численные эксперименты

3.1. Одномерное уравнение Бюргерса. Этот тест мы взяли из книги [22, глава IV.10]. Ищется решение $u(x, t)$ начально-краевой задачи для одномерного уравнения Бюргерса

$$u_t = \nu u_{xx} - uu_x, \quad 0 \leq x \leq 1, \quad 0 \leq t \leq T,$$

где ν — вязкость, значения которой выбирались как $\nu = 3 \cdot 10^{-4}$ и $\nu = 3 \cdot 10^{-5}$. Использовались начальные и краевые условия

$$u(x, 0) = \frac{3}{2}x(1-x)^2, \quad u(0, t) = u(1, t) = 0,$$

а конечное время T бралось равным 0.5, 1.0 и 1.5. Уравнение дискретизировалось конечными разностями на регулярной сетке с N внутренними узлами $x_j = j\Delta x$, $j = 1, \dots, N$, $\Delta x = 1/(N+1)$. Чтобы избежать численной диссипации и гарантировать, что дискретизация адвективного члена uu_x даёт кососимметричную матрицу [32], использовались обычные центральные конечные разности второго

порядка для адвективного члена, представленного в виде

$$uu_x = \frac{1}{3}uu_x + \frac{2}{3}\left(\frac{u^2}{2}\right)_x \approx \frac{1}{3}u_i \frac{u_{i+1} - u_{i-1}}{2\Delta x} + \frac{2}{3} \frac{u_{i+1}^2 - u_{i-1}^2}{4\Delta x}.$$

Диффузионный член νu_{xx} дискретизировался обычными конечными разностями второго порядка. Применение такой конечно-разностной аппроксимации приводит к задаче Коши

$$y'(t) = -A_{\text{symm}}y(t) - A_{\text{skew}}(y(t))y(t), \quad y(0) = v, \quad (30)$$

где

$$A_{\text{symm}}y \approx -\nu u_{xx}, \quad A_{\text{skew}}(y)y \approx uu_x,$$

матрица $A_{\text{symm}} = A_{\text{symm}}^T \in \mathbb{R}^{N \times N}$ — симметричная положительно определённая, а матрица $(A_{\text{skew}}(y))^T = -A_{\text{skew}}(y) \in \mathbb{R}^{N \times N}$ — кососимметричная для всех $y \in \mathbb{R}^N$. Заметим, что важно дискретизировать конвективный член так, чтобы получалась кососимметричная матрица [32; 50]. В противном случае получаем вклады конвективного члена в симметричную часть матрицы Якоби функции правой части, которые влияют на правильную диссипацию энергии в системе дифференциальных уравнений и могут нарушить положительную определённость симметричной части. В контексте данной работы это означало бы, что свойство $\omega \geq 0$ может не выполняться.

Наиболее простой способ применить нелинейные итерации (8) к задаче (30) — в расщеплении (3) положить $A_k := A_{\text{symm}}$, $f_k(y) := -A_{\text{skew}}(y)y$. Однако, чтобы уменьшить липшицеву константу функции $f_k(y)$ (см. условие (10)), мы добавляем к члену $-A_k y$ в (3) ещё и линейную часть члена $-A_{\text{skew}}(y)y$. Таким образом, итерации (8) применяются к решению уравнения Бюргерса (30) в виде

$$y'_{k+1}(t) = -\underbrace{(A_{\text{symm}} + A_{\text{skew}}(\bar{y}_k))}_{A_k} y_{k+1}(t) + \underbrace{[A_{\text{skew}}(\bar{y}_k) - A_{\text{skew}}(y_k(t))]}_{f_k(y_k(t))} y_k(t), \quad (31)$$

$k = 0, 1, 2, \dots,$

где $\bar{y}_k = y_k(T)$. При этом получаем, что нелинейная невязка (16) определяется, для $k = 0, 1, \dots$, соотношениями

$$\begin{aligned} r_{k+1}(t) &= f_k(y_{k+1}(t)) - f_k(y_k(t)) \\ &= [A_{\text{skew}}(\bar{y}_k) - A_{\text{skew}}(y_{k+1}(t))] y_{k+1}(t) - [A_{\text{skew}}(\bar{y}_k) - A_{\text{skew}}(y_k(t))] y_k(t), \\ r_{k+1}(T) &= [A_{\text{skew}}(\bar{y}_k) - A_{\text{skew}}(y_{k+1}(T))] y_{k+1}(T). \end{aligned}$$

Поскольку симметричная часть матрицы A_k — это $\frac{1}{2}(A_k + A_k^T) = A_{\text{symm}}$, параметр ω — это наименьшее собственное значение матрицы A_{symm} . Низкочастотные собственные моды должны хорошо аппроксимироваться даже на самой грубой из используемых сеток. Действительно, определяя ω численно (что для этой одномерной задачи нетрудно), получаем примерно один и тот же результат на всех сетках: $\omega \approx 2.96\text{e-}03$ при $\nu = 3 \cdot 10^{-4}$ и $\omega \approx 2.96\text{e-}04$ при $\nu = 3 \cdot 10^{-5}$. Нелинейный итерационный процесс (8) останавливался, как только

$$\|r_k(T)\| \leq \text{tol}, \quad (32)$$

где tol — требуемая точность. Поскольку мы прежде всего нацелены на решение задач Коши, полученных при дискретизации по пространству уравнений в частных производных, следует учитывать, что ошибка дискретизации по пространству значительна, и, следовательно, разумно брать умеренные значения точности по времени. Поэтому в этом тесте берётся $\text{tol} = 10^{-3}$. Для получения малорангового представления (29) на каждой итерации k брались $n_s = 100$ точек временного интервала t_j , $j = 1, \dots, n_s$, а число сингулярных векторов бралось равным $m = 7$. При этом наибольшее «отсечённое» сингулярное значение σ_{m+1} , служащее оценкой ошибки малорангового представления (29), имело порядок 10^{-6} для $T = 0.5$ и 10^{-3} для $T = 1.5$.

Во внутренних линейных итерациях ЭБК, используемых для решения (8), допустимая точность на норму внутренней невязки выбиралась как $\text{tol}_{\text{lin}} = \text{tol}$, т.е. внутренние итерации останавливались, как только линейная невязка удовлетворяла условию

$$\|r_{k+1}^{\text{lin}}(T)\| \leq \text{tol}_{\text{lin}}. \quad (33)$$

Размерность блочного подпространства Крылова бралась равной 10. Таким образом, поскольку блочный размер $m = 7$, затраты памяти в методе ЭБК равнялись $m(10 + 1) = 77$ векторам длины N . Поскольку задача на мелких сетках становится жёсткой, метод ЭБК применялся в режиме «сдвиг-обращение» (СО) [18; 39], т.е. блочное подпространство Крылова вычислялось для матрицы СО $(I + \gamma A_k)^{-1}$, $\gamma = T/10$, где матрица A_k определена в (31). Для этого разреженное (в этом тесте — ленточное) LU-разложение матрицы $I + \gamma A_k$ вычисляется один раз на каждой нелинейной итерации и применяется каждый раз, когда требуется вычислить действие $(I + \gamma A_k)^{-1}$.

Чтобы сравнить наш нелинейный решатель ЭБК с другими методами, мы также решаем эту тестовую задачу методом `ode15s`, доступным в пакете Матлаб и предназначенным для решения жёстких систем дифференциальных уравнений. Это неявный многошаговый метод с автоматически выбираемыми длиной шага по времени и порядком [43]. Программа применялась с абсолютными и относительными точностями $\text{AbsTol} = \text{tol}$ и $\text{RelTol} = 10^4 \cdot \text{tol}$, соответственно, при

значении $\text{tol} = 10^{-9}$. При этих значениях допустимых точностей `ode15s` давал результаты, близкие по точности к результатам наших нелинейных итераций ЭБК. Значения относительной ошибки, которые приводятся ниже, вычислялись как

$$\frac{\|y - y_{\text{ref}}\|}{\|y_{\text{ref}}\|}, \quad (34)$$

где y — численное решение в конечный момент времени T , а y_{ref} — референтное решение, вычисленное в конечный момент времени методом `ode15s` со строгими значениями допустимой точности $\text{AbsTol} = 10^{-12}$ и $\text{RelTol} = 10^{-8}$. Поскольку y_{ref} вычислялось на той же сетке по пространству, что и y , ошибка, вычисленная таким образом, может считаться правильным показателем точности по времени [44].

Полученные значения ошибки и соответствующая требуемая вычислительная работа для обоих методов показаны в таблице 1 (для вязкости $\nu = 3 \cdot 10^{-4}$) и таблице 2 (для $\nu = 3 \cdot 10^{-5}$). Для нашего нелинейного решателя ЭБК вычислительная работа измеряется как число нелинейных итераций, с одним LU-разложением и одним матрично-векторным умножением на итерацию, и общее число решений линейных систем, равное числу крыловских итераций, помноженному на размер блока m . Приведённое ниже число применений LU-разложений необязательно является кратным размеру блока $m = 7$, потому что на первой нелинейной итерации приближённое решение постоянно по времени и, следовательно, в (29) задаётся $m = 1$. Вычислительная работа для решателя `ode15s` измеряется как число шагов по времени, число LU-разложений и число вычислений функции правой части решаемой системы дифференциальных уравнений (fevals). В таблицах 1 и 2 также приведены значения нормы матрицы $A_{\text{symm}} + A_{\text{skew}}(y(T))$ — их можно рассматривать как меру жёсткости решаемой системы дифференциальных уравнений.

Как видно из таблиц 1 и 2, наш решатель требует меньшего числа LU-разложений, чем решатель `ode15s`. Нелинейный ЭБК требует большего числа применений LU-разложений, но они относительно дешёвы и, кроме того, в ЭБК они применяются одновременно к блокам из m векторов правых частей. Для обоих решателей на рис. 1 мы приводим график зависимости числа LU-разложений (представленных в таблицах) от размера сетки. Число нелинейных итераций в решателе ЭБК (с одним LU-разложением на итерацию) остаётся на сгущающихся сетках практически постоянным, тогда как число LU-разложений в `ode15s` увеличивается. Кроме того, мы видим, что решателю `ode15s` при сгущении сетки требуется больше шагов по времени и значительно больше вычислений функции правой части (fevals). Значение $T = 1.5$ — это примерно максимально допустимая длина временного интервала, при котором наш нелинейный решатель ЭБК сходится в этом тесте. Из рис. 1 мы видим, что более высокая

Таблица 1. Одномерный тест Бюргерса, вязкость $\nu = 3 \cdot 10^{-4}$. Полученная ошибка и вычислительная работа для нашего нелинейного метода ЭБК и решателя ode15s. Для метода ЭБК работа показана как число нелинейных итераций, число LU-разложений, число их применений и число матрично-векторных умножений (matvecs). Для решателя ode15s работа показана как число шагов по времени, число LU-разложений, число их применений и число вычислений функций правой части (fevals).

T	метод, точность	итер./ шаги	LU (примен. LU), matvecs/fevals	полученная ошибка
сетка $N = 500$, $\ A_{\text{symm}} + A_{\text{skew}}(y(T))\ _1 \approx 300$				
0.5	нелин.ЭБК($m = 7$), $1e-03$	5	5 (141), 5	$5.17e-06$
1.0	нелин.ЭБК($m = 7$), $1e-03$	7	7 (220), 7	$2.03e-05$
1.5	нелин.ЭБК($m = 7$), $1e-03$	10	10 (340), 10	$5.31e-05$
0.5	ode15s, $1e-09$	59	14 (73), 575	$1.93e-06$
1.0	ode15s, $1e-09$	70	16 (86), 588	$8.38e-06$
1.5	ode15s, $1e-09$	83	19 (118), 620	$2.21e-05$
сетка $N = 1000$, $\ A_{\text{symm}} + A_{\text{skew}}(y(T))\ _1 \approx 1\,200$				
0.5	нелин.ЭБК($m = 7$), $1e-03$	5	5 (170), 5	$5.06e-06$
1.0	нелин.ЭБК($m = 7$), $1e-03$	7	7 (256), 7	$2.00e-05$
1.5	нелин.ЭБК($m = 7$), $1e-03$	10	10 (389), 10	$5.30e-05$
0.5	ode15s, $1e-09$	67	15 (83), 1085	$1.76e-06$
1.0	ode15s, $1e-09$	78	17 (104), 1106	$1.13e-05$
1.5	ode15s, $1e-09$	91	19 (134), 1136	$2.48e-05$
сетка $N = 2000$, $\ A_{\text{symm}} + A_{\text{skew}}(y(T))\ _1 \approx 4\,800$				
0.5	нелин.ЭБК($m = 7$), $1e-03$	5	5 (177), 5	$5.07e-06$
1.0	нелин.ЭБК($m = 7$), $1e-03$	7	7 (277), 7	$2.00e-05$
1.5	нелин.ЭБК($m = 7$), $1e-03$	11	11 (452), 11	$4.38e-05$
0.5	ode15s, $1e-09$	73	16 (91), 2093	$2.29e-06$
1.0	ode15s, $1e-09$	85	19 (107), 2109	$7.96e-06$
1.5	ode15s, $1e-09$	98	22 (139), 2141	$2.22e-05$
сетка $N = 4000$, $\ A_{\text{symm}} + A_{\text{skew}}(y(T))\ _1 \approx 19\,200$				
0.5	нелин.ЭБК($m = 7$), $1e-03$	5	5 (193), 5	$5.06e-06$
1.0	нелин.ЭБК($m = 7$), $1e-03$	8	8 (347), 8	$4.82e-06$
1.5	нелин.ЭБК($m = 7$), $1e-03$	11	11 (501), 11	$4.38e-05$
0.5	ode15s, $1e-09$	79	18 (99), 4101	$1.93e-06$
1.0	ode15s, $1e-09$	90	20 (112), 4114	$8.25e-06$
1.5	ode15s, $1e-09$	103	23 (144), 4146	$2.21e-05$

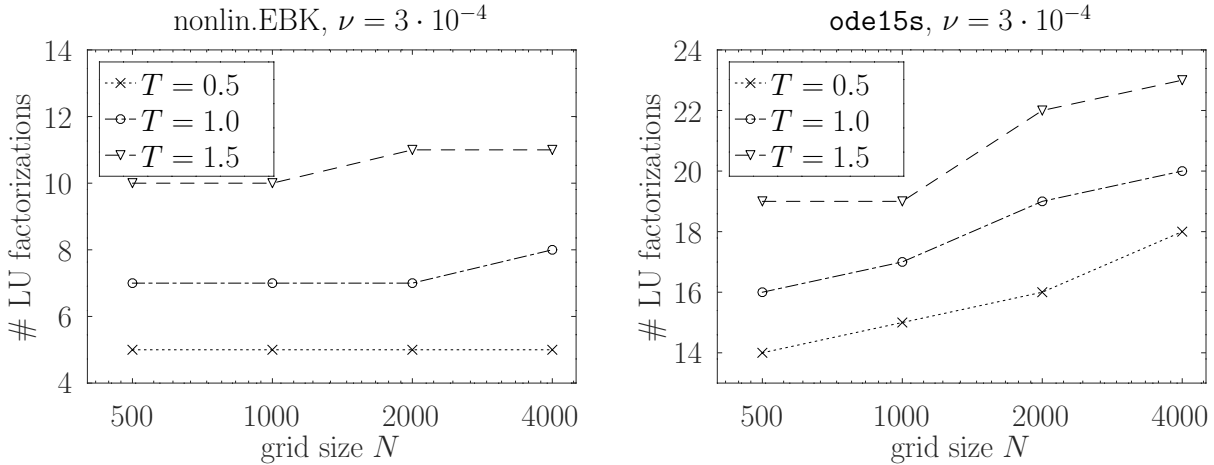


Рис. 1. Тест Бюргерса. Число требуемых LU-разложений в зависимости от размера сетки в нелинейном методе ЭБК (слева) и ode15s (справа) для различных значений конечного времени T и вязкости $\nu = 3 \cdot 10^{-4}$.

эффективность в методе ЭБК достигается при больших значениях T (т.е. при значениях T , близких к максимально возможному). Действительно, на самой мелкой сетке $N = 4000$ требуется $5 \times 2 = 10$ LU-разложений в единицу времени при $T = 0.5$, 8 LU-разложений — при $T = 1$ и $11/1.5 \approx 7$ разложений — при $T = 1.5$.

Сравнивая таблицы 1 и 2, мы видим, что производительность решателя ode15s повышается для меньшего значения вязкости $\nu = 3 \cdot 10^{-5}$. Это ожидаемо, поскольку жёсткость дискретизированного по пространству уравнения Бюргерса снижается с ν . Действительно, на мелких сетках, где вклад A_{symm} (пропорциональный $(\Delta x)^{-2}$) преобладает над вкладом $A_{\text{skew}}(t)$ (пропорциональным $(\Delta x)^{-1}$), получаем $\|A_{\text{symm}} + A_{\text{skew}}(t)\| \approx \mathcal{O}(\nu)$, что подтверждается значениями $\|A_{\text{symm}} + A_{\text{skew}}(T)\|$, представленными в таблицах.

На рис. 2 представлены референтное решение $y_{\text{ref}}(T)$ и первые три итерационных приближения $y_0(T) = v$, $y_1(T)$ и $y_2(T)$. Графики сходимости норм невязки и ошибки показаны на рис. 3. Представленная здесь норма невязки вычислена в момент времени $t = T$ согласно (16), а норма ошибки посчитана согласно (34). Как видим, невязка оказывается надёжным индикатором ошибки. Кроме того, вязкость ν не влияет на характер сходимости, что соответствует теории: ν не меняет нелинейный член f_k в (8) и его постоянную Липшица.

3.2. Трёхмерное уравнение Брату. В этом тесте рассматривается задача Коши для трёхмерного уравнения Лиувилля–Брату–Гельфанда [29]: найти такую

Таблица 2. Одномерный тест Бюргерса, вязкость $\nu = 3 \cdot 10^{-5}$. Полученная ошибка и вычислительная работа для нашего нелинейного метода ЭБК и решателя ode15s. Для метода ЭБК работа показана как число нелинейных итераций, число LU-разложений, число их применений и число матрично-векторных умножений (matvecs). Для решателя ode15s работа показана как число шагов по времени, число LU-разложений, число их применений и число вычислений функций правой части (fevals).

T	метод, точность	итер./ шаги	LU (примен. LU), matvecs/fevals	полученная ошибка
сетка $N = 500$, $\ A_{\text{symm}} + A_{\text{skew}}(y(T))\ _1 \approx 90$				
0.5	нелин.ЭБК($m = 7$), $1e-03$	5	5 (69), 5	$1.82e-05$
1.0	нелин.ЭБК($m = 7$), $1e-03$	7	7 (139), 7	$2.26e-05$
1.5	нелин.ЭБК($m = 7$), $1e-03$	13	13 (414), 13	$1.10e-04$
0.5	ode15s, $1e-09$	28	7 (38), 540	$3.71e-06$
1.0	ode15s, $1e-09$	38	9 (53), 555	$1.23e-05$
1.5	ode15s, $1e-09$	52	13 (89), 591	$3.38e-05$
сетка $N = 1000$, $\ A_{\text{symm}} + A_{\text{skew}}(y(T))\ _1 \approx 210$				
0.5	нелин.ЭБК($m = 7$), $1e-03$	5	5 (90), 5	$6.20e-06$
1.0	нелин.ЭБК($m = 7$), $1e-03$	7	7 (176), 7	$2.25e-05$
1.5	нелин.ЭБК($m = 7$), $1e-03$	12	12 (430), 12	$1.07e-04$
0.5	ode15s, $1e-09$	35	9 (44), 1046	$2.44e-06$
1.0	ode15s, $1e-09$	45	10 (60), 1062	$1.64e-05$
1.5	ode15s, $1e-09$	59	15 (99), 2102	$3.83e-05$
сетка $N = 2000$, $\ A_{\text{symm}} + A_{\text{skew}}(y(T))\ _1 \approx 540$				
0.5	нелин.ЭБК($m = 7$), $1e-03$	5	5 (120), 5	$5.29e-06$
1.0	нелин.ЭБК($m = 7$), $1e-03$	7	7 (190), 7	$2.22e-05$
1.5	нелин.ЭБК($m = 7$), $1e-03$	12	12 (494), 12	$1.06e-04$
0.5	ode15s, $1e-09$	42	10 (55), 2057	$3.91e-06$
1.0	ode15s, $1e-09$	52	12 (70), 2072	$1.32e-05$
1.5	ode15s, $1e-09$	66	17 (108), 4111	$3.32e-05$
сетка $N = 4000$, $\ A_0 + A_{\text{skew}}(y(T))\ _1 \approx 1920$				
0.5	нелин.ЭБК($m = 7$), $1e-03$	5	5 (149), 5	$5.24e-06$
1.0	нелин.ЭБК($m = 7$), $1e-03$	8	8 (276), 8	$5.52e-06$
1.5	нелин.ЭБК($m = 7$), $1e-03$	12	12 (578), 12	$1.07e-04$
0.5	ode15s, $1e-09$	47	11 (61), 4063	$3.86e-06$
1.0	ode15s, $1e-09$	58	15 (80), 4082	$1.22e-05$
1.5	ode15s, $1e-09$	71	18 (112), 8115	$3.36e-05$

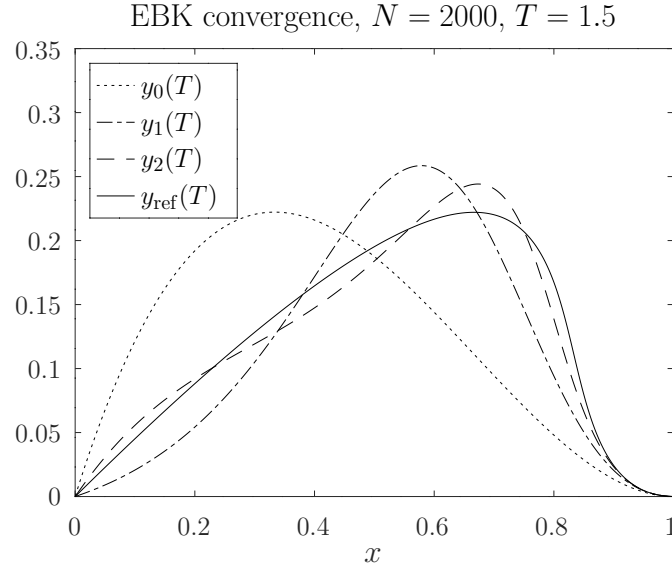


Рис. 2. Тест Бюргерса. Первые три итерационные приближения $y_0(T) = v$, $y_1(T)$, $y_2(T)$ и референтное решение $y_{\text{ref}}(T)$, сетка $N = 2000$, конечное время $T = 1.5$, вязкость $\nu = 3 \cdot 10^{-4}$

функцию $u = u(x, y, z, t)$, что

$$\begin{aligned} u_t &= 10^4 u_{xx} + 10^2 u_{yy} + u_{zz} + C e^u + g^{\text{src}}(x, y, z, t), \\ (x, y, z) &\in \Omega = [0, 1]^3, \quad 0 \leq t \leq T, \quad C = 3 \cdot 10^4, \\ u(x, y, z, 0) &= e^{-100((x-0.2)^2 + (y-0.4)^2 + (z-0.5)^2)}, \quad u|_{\partial\Omega} = 0. \end{aligned} \quad (35)$$

Здесь функция источника g^{src} задана как

$$\begin{aligned} g^{\text{src}}(x, y, z, t) &= \begin{cases} e^{-100((x-x_0(t))^2 + (y-y_0(t))^2 + (z-0.5)^2)} + C u(x, y, z, 0), & t \leq 5 \cdot 10^{-5}, \\ e^{-100((x-x_0(t))^2 + (y-y_0(t))^2 + (z-0.5)^2)}, & t > 5 \cdot 10^{-5}, \end{cases} \\ x_0(t) &= 0.5 + 0.3 \cos(2000\pi t), \quad y_0(t) = 0.5 + 0.3 \sin(2000\pi t), \end{aligned}$$

а конечное время T выбирается как $T = 5 \cdot 10^{-5}$ или $T = 1 \cdot 10^{-4}$. Решение задачи (35) показано на рис. 4.

Применяя к (35) обычную конечно-разностную пространственную аппроксимацию второго порядка на регулярной сетке, получаем задачу Коши

$$y'(t) = -Ay(t) + \hat{f}(y(t)) + g(t), \quad y(0) = v, \quad (36)$$

где элементы вектор-функций $y(t)$ и $g(t)$ содержат значения, соответственно, численного решения и функции $g^{\text{src}}(x, y, z, t)$ в узлах сетки, а

$$Ay(t) \approx -(10^4 u_{xx} + 10^2 u_{yy} + u_{zz}). \quad (37)$$

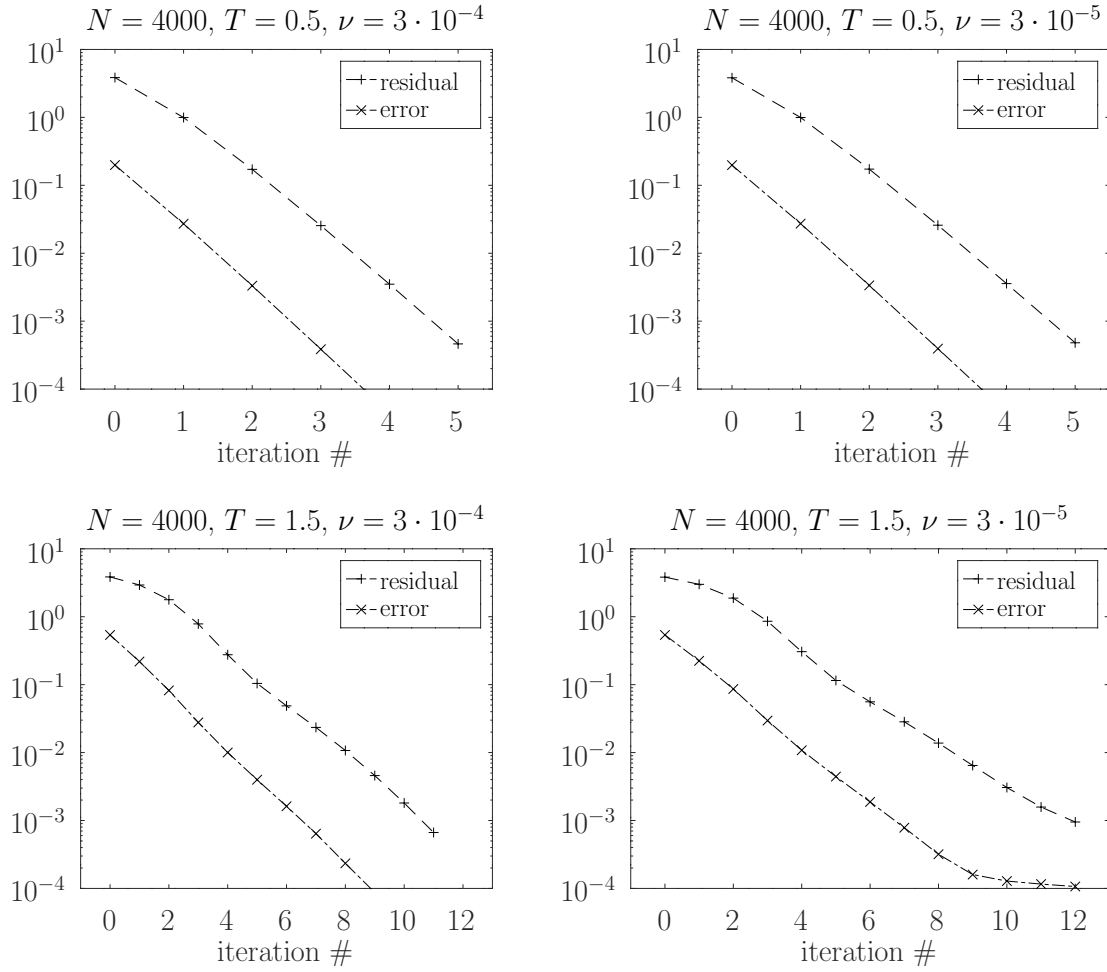


Рис. 3. Тест Бюргерса. Норма невязки (16) и относительная норма ошибки (34) в зависимости от номера итерации для $T = 0.5$ (верхние графики) и $T = 1.5$ (нижние графики), $\nu = 3 \cdot 10^{-4}$ (левые графики) и $\nu = 3 \cdot 10^{-5}$ (правые графики). Сетка $N = 4000$. Стагнации ошибки на нижнем правом графике можно избежать увеличением размера блока m .

Компоненты $[\hat{f}(y)]_i$ вектор-функции $\hat{f}(y)$ определяются как

$$[\hat{f}(y)]_i = Ce^{y_i}, \quad i = 1, \dots, N.$$

Чтобы применить нелинейную релаксацию формы волны (8) к решению (36), можно было бы положить $A_k := A$ и $f_k(y) := \hat{f}(y)$. Однако, чтобы уменьшить липшицеву константу L в $f_k(y)$, дополним матрицу A_k линейной частью $\hat{f}$:

$$y'_{k+1}(t) = -(\underbrace{A - J(\bar{y}_k)}_{A_k})y_{k+1}(t) + \underbrace{[\hat{f}(y_k(t)) - J(\bar{y}_k)y_k(t)]}_{f_k(y_k(t))} + g(t), \quad (38)$$

$$k = 0, 1, 2, \dots,$$

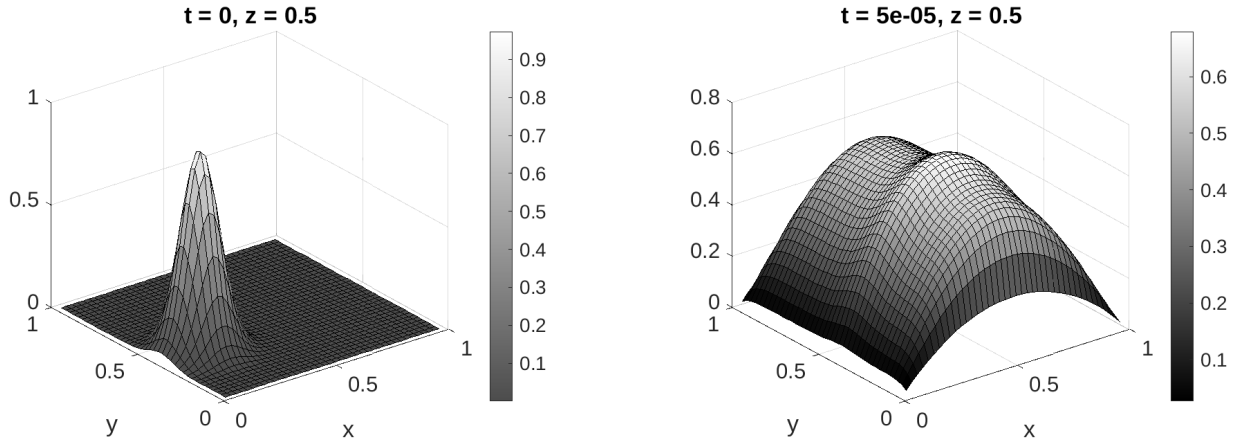


Рис. 4. Трёхмерный тест Брату. Левый график: функция начального условия $u(x, y, z, 0)$ на регулярной сетке $40 \times 40 \times 40$ для $z = 0.5$. Правый график: численное решение $u(x, y, z, t)$ на той же сетке для $z = 0.5$ и $t = 5 \cdot 10^{-5}$.

где $J(\bar{y}_k)$ — якобиан $\hat{f}(y)$ в $\bar{y}_k = y_k(T)$ и используется приближение

$$\hat{f}(y_{k+1}(t)) \approx \hat{f}(y_k(t)) + J(\bar{y}_k)(y_{k+1}(t) - y_k(t)).$$

Тогда нелинейная невязка (16) принимает вид, для $k = 0, 1, \dots$,

$$\begin{aligned} r_{k+1}(t) &= f_k(y_{k+1}(t)) - f_k(y_k(t)) \\ &= \left[\hat{f}(y_{k+1}(t)) - J(\bar{y}_k)y_{k+1}(t) \right] - \left[\hat{f}(y_k(t)) - J(\bar{y}_k)y_k(t) \right] \\ &= \hat{f}(y_{k+1}(t)) - \hat{f}(y_k(t)) - J(\bar{y}_k)(y_{k+1}(t) - y_k(t)). \end{aligned}$$

Поскольку решаемая задача Коши — жёсткая и анизотропная, нелинейная невязка может достигать по норме величин $\mathcal{O}(10^8)$. Поэтому мы переходим от абсолютного остановочного критерия $\|r_k(T)\| \leq \text{tol}$, использованного в предыдущем тесте, к относительному остановочному критерию

$$\|r_k(T)\| \leq \text{tol} \|r_0(T)\|, \quad (39)$$

со значениями допустимой точности tol , варьируемыми в этом тесте от 10^{-2} до 10^{-4} (конкретные значения указываются в таблице 3, обсуждаемой ниже). Заметим, что, как оговорено в разделе 2.5, берётся $y_0(t) \equiv v$.

Малоранговое представление (29) на каждой итерации k вычисляется в $n_s = 100$ точках t_j , $j = 1, \dots, s$ для $m = 4$ или $m = 5$ членов. Линейный решатель ЭБК, применяемый для решения задачи (8), останавливается, как только

$$\|r_{k+1}^{\text{lin}}(T)\| \leq \text{tol}_{\text{lin}} = \|f_k(y_k(0)) + g(0)\| \text{tol}/10. \quad (40)$$

Это означает, что линейная невязка мала по отношению к остальным членам возмущённого дифференциального уравнения (27). Таким образом, поскольку в

этой жёсткой анизотропной задаче член $f_k(y_k(t)) + g(t)$ может быть очень велик по норме, мы переходим от абсолютного остановочного критерия внутренних итераций (33) к критерию (40).

Размерность подпространства Крылова полагалась равной 10. Поэтому при размере блока $m = 5$, решатель ЭБК требует хранения $5 \cdot (10 + 1) = 55$ векторов. Поскольку задача жёсткая, так же, как и в предыдущем тесте, ЭБК применяется в режиме СО: разреженное LU-разложение $I + \gamma A_k$ вычисляется и используется на каждой нелинейной итерации. Представленные значения ошибки вычислены согласно (34).

Мы сравниваем наш нелинейный метод ЭБК с двухстадийным методом Розенброка ROS2, см. [1] или [27, глава IV.5]. Для задачи Коши (1) его можно записать в виде

$$\begin{aligned} y^{\ell+1} &= y^\ell + \frac{3}{2}\tau k_1 + \frac{1}{2}\tau k_2, \\ (I - \gamma\tau\hat{A})k_1 &= F(t_\ell, y^\ell), \\ (I - \gamma\tau\hat{A})k_2 &= F(t_{\ell+1}, y^\ell + \tau k_1) - 2k_1. \end{aligned} \tag{41}$$

Здесь $y^\ell \approx y(\ell\tau)$ — численное решение на шаге $\ell = 0, 1, 2, \dots$ ($y^0 = v$), полученное с длиной шага $\tau > 0$. Эта схема имеет второй порядок точности для любой матрицы $\hat{A} \in \mathbb{R}^{N \times N}$. Здесь в качестве матрицы $\hat{A}$ мы берём якобиан Φ , вычисленный в $t_\ell = \ell\tau$ и y_ℓ . На каждом шаге ℓ можно вычислить разреженное LU-разложение $I - \gamma\tau\hat{A}$ и решить им две линейные системы для $k_{1,2}$. Численные эксперименты, представленные здесь, выполнены в Матлабе, и, как оказывается, на использованных сетках существенно быстрее вычислять действия обратной матрицы $(I - \gamma\tau\hat{A})^{-1}$ матлабовским оператором $\backslash$ (а не вычислять на каждом шаге по времени сначала LU-разложение, а затем применять его дважды*). Вероятной причиной этого являются дополнительные затраты Матлаба для вызова процедур разреженного LU-разложения (пакета UMFPACK [12; 13]) и хранения вычисленных LU-множителей в виде матлабовских переменных. Заметим, что как оператор $\backslash$, так и LU-разложение для больших разреженных матриц в Матлабе используют UMFPACK.

Результаты сравнения нашего нелинейного метода ЭБК и метода ROS2 представлены в таблице 3. Результаты ясно показывают, что для этой тестовой задачи нелинейный решатель ЭБК намного эффективнее метода ROS2 как по числу LU-разложений (и их применений), так и по расчётному времени. Мы также видим, что число нелинейных итераций релаксации формы волны не зависит от размера сетки и конечного времени T . В то время как независимость

*Чтобы решить линейную систему $Ax=b$ с квадратной невырожденной матрицей A в Матлабе, можно применить оператор $\backslash$ так: $x=A \backslash b$. При этом оператор вычисляет и применяет LU-разложение.

Таблица 3. Трёхмерный тест Брату. Полученная ошибка и вычислительная работа для нашего нелинейного метода ЭБК и схемы ROS2. Для метода ЭБК работа показана как число нелинейных итераций, число LU-разложений, число их применений и число матрично-векторных умножений (matvecs). Для схемы ROS2 работа показана как число шагов по времени, число LU-разложений, число их применений и число вычислений функций правой части (fevals).

метод, точность / τ	итер./ шаги	LU (примен. LU), matvecs/fevals	полученная ошибка	расчётное время, с
сетка $40 \times 40 \times 40$, $T = 5 \cdot 10^{-5}$, $\ A(y(T))\ _1 \approx 6.79e+07$				
нелин.ЭБК($m = 4$), $1e-02$	4	4 (131), 4	$6.21e-05$	41
нелин.ЭБК($m = 5$), $1e-02$	4	4 (161), 4	$1.22e-05$	42
ROS2, $\tau = T/320$	320	320 (640), 640	$2.36e-04$	225
ROS2, $\tau = T/640$	640	640 (1280), 1280	$5.73e-07$	442
сетка $40 \times 40 \times 40$, $T = 1 \cdot 10^{-4}$, $\ A(y(T))\ _1 \approx 6.79e+07$				
нелин.ЭБК($m = 4$), $1e-02$	4	4 (131), 4	$2.19e-05$	40
нелин.ЭБК($m = 5$), $1e-02$	4	4 (161), 4	$1.52e-05$	43
ROS2, $\tau = T/320$	320	320 (640), 640	$1.51e-05$	201
ROS2, $\tau = T/640$	320	640 (1280), 1280	$7.51e-06$	401
сетка $60 \times 60 \times 60$, $T = 5 \cdot 10^{-5}$, $\ A(y(T))\ _1 \approx 1.50e+08$				
нелин.ЭБК($m = 4$), $1e-02$	4	4 (131), 4	$6.16e-05$	288
нелин.ЭБК($m = 5$), $1e-02$	4	4 (161), 4	$1.23e-05$	308
ROS2, $\tau = T/320$	320	320 (640), 640	$2.35e-04$	2177
ROS2, $\tau = T/640$	640	640 (1280), 1280	$1.66e-05$	4790
сетка $60 \times 60 \times 60$, $T = 1 \cdot 10^{-4}$, $\ A(y(T))\ _1 \approx 1.50e+08$				
нелин.ЭБК($m = 4$), $1e-02$	4	4 (131), 4	$2.16e-05$	288
нелин.ЭБК($m = 5$), $1e-02$	4	4 (161), 4	$1.53e-05$	310
ROS2, $\tau = T/320$	320	320 (640), 640	$2.11e-05$	2013
ROS2, $\tau = T/640$	640	640 (1280), 1280	$1.66e-05$	4368

сходимости от размера сетки ожидаема (поскольку размер сетки не влияет на нелинейный член Ce^u), слабая зависимость сходимости от T , вероятно, является свойством конкретной тестовой задачи. Из графиков на рис. 5 мы видим, что характер сходимости от T всё же зависит.

Расчёты нелинейным методом ЭБК, результаты которых представлены в таблице 3, получены для размеров блока $m = 4$ и $m = 5$. Значение $m = 4$ едва приемлемо по точности, но удовлетворительные результаты всё же даёт. По-

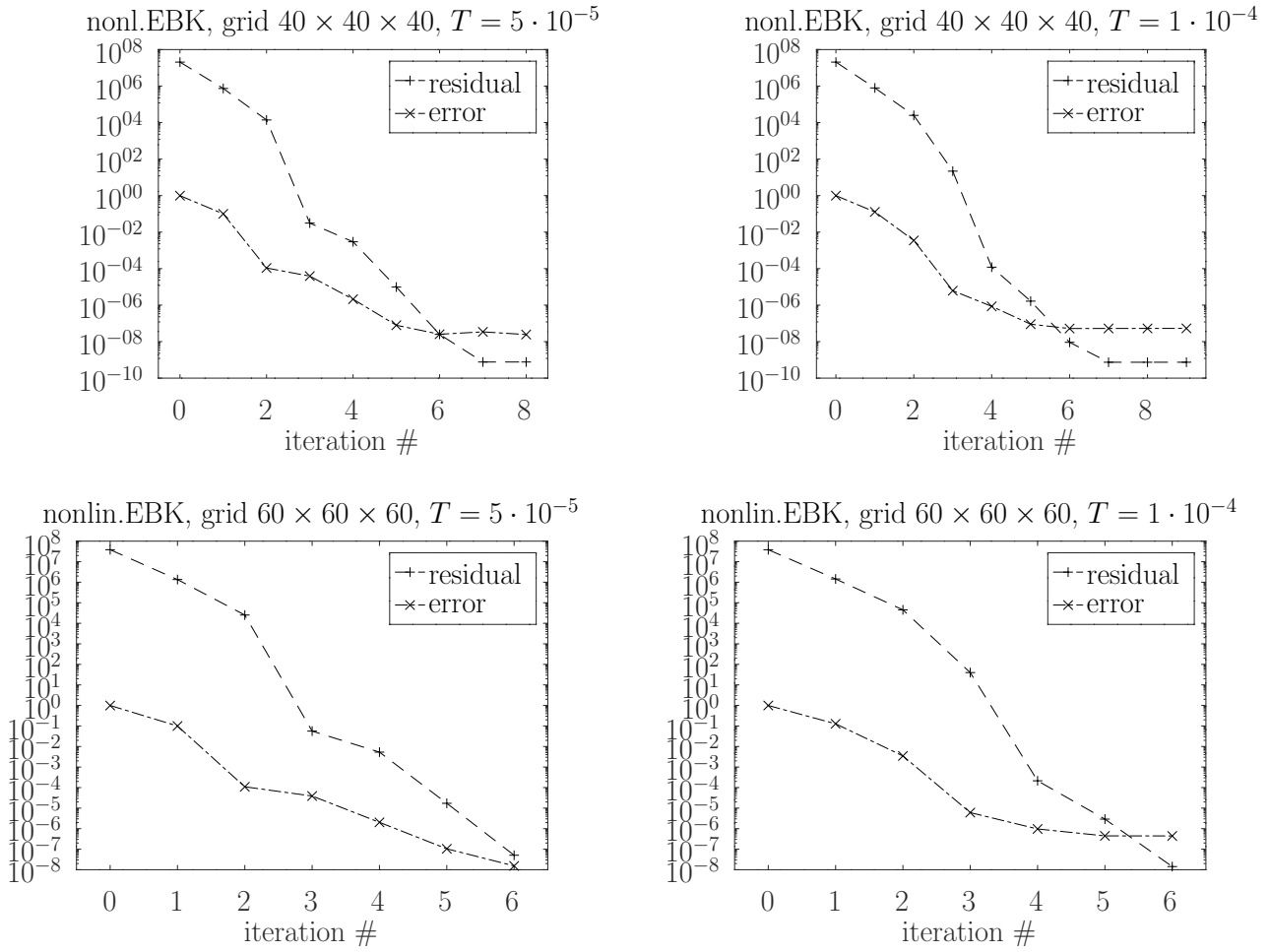


Рис. 5. Трёхмерный тест Брату. Нормы невязки и ошибки в зависимости от номера итераций в нелинейном методе ЭБК на сетке $40 \times 40 \times 40$ (верхние графики) и $60 \times 60 \times 60$ (нижние графики), $T = 5 \cdot 10^{-5}$ (левые графики) и $T = 1 \cdot 10^{-4}$ (правые графики). Графики сделаны для увеличенного размера блока $m = 12$.

этому графики рис. 5, представленные, чтобы получить понимание сходимости ошибки, сделаны с блоками большего размера $m = 32$ и для $n_s = 300$ точек. На обоих графиках ясно видна стагнация сходимости. Затрудняясь дать строгое объяснение этой стагнации, отметим, что стагнация исчезает при отключении сильной анизотропии в (37), если в качестве A берётся лапласиан. Уровень точности, на котором происходит стагнация, оказывается связанным с числом обусловленности матрицы Якоби решаемой системы обыкновенных дифференциальных уравнений. Понятно, что эта матрица задействована нелинейным решателем ЭБК в регулярно выполняемых малоранговых приближениях, необходимых для внутренних итераций на блочных крыловских подпространствах. Другой возможный источник снижения точности — решение линейных систем в этих внутренних итерациях, где линейные системы решаются для матриц $I + \gamma A_k$ при значениях $\gamma > 0$ на порядки больших, чем обычные шаги по времени в

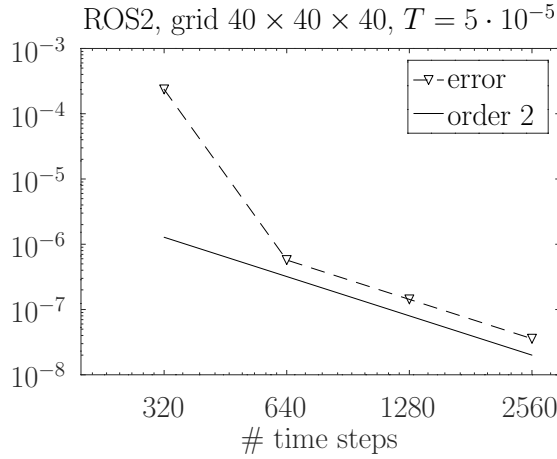


Рис. 6. Трёхмерный тест Брату. Относительная ошибка метода ROS2 в момент времени $t = T$, вычисленная согласно (34), в зависимости от числа шагов по времени.

неявных схемах. Тем не менее, стагнация сходимости наблюдается на уровне точности $\approx 10^{-8}$, что кажется вполне приемлемым при численном решении уравнений в частных производных.

В заключение коротко обсудим сходимость метода ROS2. Поскольку порядок его сходимости не может быть отслежен по значениям ошибки в таблице 3, мы даём график сходимости метода на рис. 6. Здесь значения относительной ошибки, вычисленной согласно (34), показаны в зависимости от числа шагов по времени. Ясно виден второй порядок сходимости.

3.3. Трёхмерное нелинейное уравнение теплопроводности. В рассмотренных выше тестах функция правой части решаемой системы дифференциальных уравнений имела выделенную линейную часть (а именно, $-A_{\text{symm}}y(t)$ в первом тесте и $-Ay(t)$ — во втором, см. соотношения (30) и (36) соответственно). Чтобы проверить, применим ли наш нелинейный решатель ЭБК к задачам, где такой выделенной линейной части нет, рассмотрим следующую нелинейную задачу теплопроводности, описанную в [55]: найти такую функцию $u = u(x, y, z, t)$, что

$$\begin{aligned}
 u_t &= (k^{(x)}(u)u_x)_x + (k^{(y)}(u)u_y)_y + (k^{(z)}(u)u_z)_z, \\
 (x, y, z) &\in [0, 1]^3, \quad 0 \leq t \leq T_{\text{final}}, \\
 u|_{x=0} &= u|_{x=1}, \quad u|_{y=0} = 900, \quad u|_{y=1} = 300, \quad u_z|_{z=0} = u_z|_{z=1} = 0, \\
 u(x, y, z, 0) &= 1800 e^{-60((x-0.5)^2 + (y-0.5)^2 + (z-0.5)^2)}, \\
 k^{(x)}(u) &= u/300, \quad k^{(y)}(u) = k^{(z)}(u) = k^{(x)}(u)/10.
 \end{aligned} \tag{42}$$

Для дискретизации этой начально-краевой задачи по пространству используем равномерную сетку из $n_x \times n_y \times n_z$ узлов (x_i, y_j, z_k) , $x_i = ih_x$, $i = 1, \dots, n_x$,

$y_j = jh_y$, $j = 1, \dots, n_y$, $z_k = (k - 1/2)h_z$, $k = 1, \dots, n_z$, с шагами сетки $h_{x,y} = 1/(n_{x,y} + 1)$ и $h_z = 1/n_z$. Здесь по направлению z используется сдвинутая на полшага сетка, чтобы легче аппроксимировать краевые условия Неймана. Стандартную конечно-разностную аппроксимацию второго порядка по направлению x запишем в виде

$$(k^{(x)}(u)u_x)_x|_{i,j,k} \approx \frac{k^{(x)}(u_{i+1/2,j,k})(u_{i+1,j,k} - u_{i,j,k}) - k^{(x)}(u_{i-1/2,j,k})(u_{i,j,k} - u_{i-1,j,k})}{h_x^2},$$

где $u_{i\pm 1/2,j,k} = (u_{i\pm 1,j,k} + u_{i,j,k})/2$. По направлениям y и z конечные разности применяются аналогично, при этом следует учитывать сдвиг узлов по направлению z . Такая дискретизация по пространству задачи (42) приводит к задаче Коши

$$y'(t) = -A(y(t))y(t) + g, \quad y(0) = v, \quad (43)$$

где элементы вектор-функции $y(t)$ — это значения искомого численного решения на сетке по пространству, а постоянный по времени вектор g содержит вклады неоднородных краевых условий.

Зададим конечное время $T_{\text{final}} = 0.1$ и будем применять схему ЭБК последовательно на n_T временных интервалах $[0, T]$, $[T, 2T]$, $\dots$ с $T = T_{\text{final}}/n_T$. Ещё раз подчеркнём принципиальное отличие применения схемы ЭБК последовательно на временных интервалах длины T от обычных схем интегрирования по времени, применяемых с шагом по времени τ : точность схемы ЭБК не зависит от T , что обычно позволяет брать $T \gg \tau$. Итерации релаксации формы волны в этом тесте будем использовать с расщеплением (3), где

$$A_k = A(\bar{y}_k), \quad f_k(y) = A(\bar{y}_k)y - A(y)y, \quad (44)$$

а $\bar{y}_k$ — значение k -го итерационного приближения в конце текущего временного интервала, т.е. $\bar{y}_k = y_k(iT)$ на интервале по времени i . В соотношении (44) нетрудно узнать расщепление, использованное в первом тесте для решения уравнения Бюргерса — для этого в расщеплении из теста Бюргерса следует выбрать $A_{\text{symm}} = 0$ и $A_{\text{skew}} = A$.

Вложенные (внутренние и внешние) итерации ЭБК применим точно так же, как и в предыдущем тесте, с теми же остановочными критериями (см. (39), (40)) и таким же числом точек n_s . Размерность подпространства Крылова для внутренних итераций задавалась равной 10, что при использованном размере блока $m = 6$ или $m = 10$ означает хранение 66 или 110 векторов соответственно. Как и в предыдущем тесте, мы сравниваем решатель ЭБК с интегратором ROS2. Результаты тестовых запусков представлены в таблице 4. Как видим, на сетке по пространству $40 \times 40 \times 40$ эти две схемы интегрирования затрачивают примерно одинаковое процессорное время. При этом решатель ЭБК требует намного

Таблица 4. Трёхмерное нелинейное уравнение теплопроводности. Полученная ошибка и вычислительная работа для нашего нелинейного метода ЭБК и схемы ROS2. Для метода ЭБК работа показана как число нелинейных итераций, число LU-разложений, число их применений и число матрично-векторных умножений (matvecs). Для схемы ROS2 работа показана как число шагов по времени, число LU-разложений, число их применений и число вычислений функций правой части (fevals). Метод ЭБК применялся последовательно на n_T интервалах по времени длиной $T = T_{\text{final}}/n_T$.

метод, временные интервалы n_T точность / τ	итер./ шаги	LU, (примен. LU) matvecs/fevals	полученная ошибка	расчётное время, с
сетка $40 \times 40 \times 40$, $T_{\text{final}} = 0.1$, $T_{\text{final}} \ A(y(T_{\text{final}}))\ _1 \approx 2.19\text{e}+03$				
нелин.ЭБК($m = 6$), $n_T = 10$, $1\text{e}-02$	52	52 (1223), 52	$1.30\text{e}-04$	387
нелин.ЭБК($m = 10$), $n_T = 10$, $1\text{e}-02$	52	52 (1936), 52	$5.91\text{e}-05$	448
ROS2, $\tau = T_{\text{final}}/250$	250	250 (500)	$7.26\text{e}-05$	347
ROS2, $\tau = T_{\text{final}}/500$	500	500 (1000)	$2.02\text{e}-05$	821
сетка $60 \times 60 \times 60$, $T_{\text{final}} = 0.1$, $T_{\text{final}} \ A(y(T_{\text{final}}))\ _1 \approx 4.96\text{e}+04$				
нелин.ЭБК($m = 6$), $n_T = 20$, $1\text{e}-02$	94	94 (2068), 94	$4.41\text{e}-05$	8229
ROS2, $\tau = T_{\text{final}}/500$	500	500 (1000)	$3.99\text{e}-05$	19578

меньше LU-разложений. На более мелкой сетке $60 \times 60 \times 60$ решатель ЭБК явно превосходит ROS2 как по процессорному времени, так и по числу требуемых LU-разложений. При переходе на более мелкую сетку для сохранения сходимости в схеме ЭБК нам потребовалось уменьшить длину временного интервала T с $T = 0.01$ до $T = 0.005$. Таким образом, в этом тесте мы наблюдаем ухудшение сходимости метода ЭБК при измельчении сетки. Также наблюдается и небольшое увеличение числа нелинейных итераций: при $T = 0.005$ ЭБК требует всего 73 итерации на сетке $40 \times 40 \times 40$ и 94 итерации на сетке $60 \times 60 \times 60$.

4. Выводы

В этой работе теоретически и практически изучаются нелинейные итерации релаксации формы волны и их реализация на основе блочных подпространств

Крылова (нелинейный метод ЭБК). Теоретически мы показали, что сходимость рассматриваемой нелинейной релаксации формы волны определяется постоянной Липшица L нелинейного члена f_k в (3), логарифмической нормой ω матрицы линейных членов A_k и длиной временного интервала T , см. соотношение (10). Нами также установлена суперлинейная сходимость рассматриваемой нелинейной релаксации формы волны (утверждение 2), известная в линейном случае. Кроме того, установлена связь между невязкой и ошибкой нелинейных итераций релаксации формы волны, см. оценку (20). Также показано, что неточное решение линейных задач Коши, возникающих на каждой нелинейной итерации, приводит к процессу, сходимость которого эквивалентна сходимости точной нелинейной релаксации формы волны с увеличенной константой Липшица $L + l$, где l — параметр точности решения линейных задач Коши.

Практическая эффективность нелинейных итераций релаксации формы волны в значительной степени зависит от возможности быстро решать возникающую на каждой итерации линейную задачу Коши и экономно хранить итерационные решения во времени. Как показано в этой работе, обе эти цели (решение линейных задач Коши и хранение полученных решений) могут быть успешно достигнуты с помощью экспоненциального блочного крыловского метода ЭБК [2]. В представленных экспериментах метод ЭБК применялся в режиме СО («сдвига-обращение»), т.е. нами предполагалось, что линейные системы СО (аналогичные возникающим в неявных схемах интегрирования по времени) могут быть эффективно решены.

Сравнение с неявными методами интегрирования по времени в численных тестах продемонстрировали превосходство и потенциал наших нелинейных итераций релаксации формы волны на основе ЭБК-СО. Показано, что представленный остановочный критерий на основе нелинейной невязки является надёжным способом контроля сходимости и оценки точности. Более того, как показывают численные тесты, сеточная независимость сходимости, характерная для линейных методов СО подпространства Крылова [18; 21; 39], наследуется нашими нелинейными итерациями релаксации формы волны.

Тем не менее для успешной работы представленного подхода требуется определённая настройка параметров (например, правильный выбор размера блока m и длины T интервала по времени). Предполагается рассмотреть эти вопросы в будущих исследованиях.

Другой подход с использованием экспоненциальных схем интегрирования по времени на крыловских подпространствах предложен для решения нелинейных уравнений теплопроводности в [6]. В этой задаче быстрое изменение решения, похоже, исключает применение представленной здесь релаксации формы волны с блочными подпространствами Крылова. Дальнейшие исследования могут показать, как выбирать на практике между подходами, представленными

здесь и в [6]. Наконец, было бы полезно получить понимание того, как эти два подхода соотносятся в теории и практике с другими экспоненциальными интеграторами нелинейных задач по времени [23—25].

Благодарности. При выполнении работы использовались ресурсы гибридного суперкомпьютера К-100, установленного в Центре коллективного пользования ИПМ им. М.В. Келдыша РАН.

Список литературы

1. A second order Rosenbrock method applied to photochemical dispersion problems / J. G. Verwer [et al.] // SIAM J. Sci. Comput. — 1999. — Vol. 20. — P. 456–480.
2. Botchev M. A. A block Krylov subspace time-exact solution method for linear ordinary differential equation systems // Numer. Linear Algebra Appl. — 2013. — Vol. 20, no. 4. — P. 557–574. — <http://doi.org/10.1002/nla.1865>.
3. Botchev M. A., Grimm V., Hochbruck M. Residual, Restarting and Richardson Iteration for the Matrix Exponential // SIAM J. Sci. Comput. — 2013. — Vol. 35, no. 3. — A1376–A1397. — <http://doi.org/10.1137/110820191>.
4. Botchev M. A., Oseledets I. V., Tyrtysnikov E. E. Iterative across-time solution of linear differential equations: Krylov subspace versus waveform relaxation // Computers & Mathematics with Applications. — 2014. — Vol. 67, no. 12. — P. 2088–2098. — ISSN 0898-1221. — <http://doi.org/10.1016/j.camwa.2014.03.002>.
5. Botchev M. A., Vorst H. A. van der. A parallel nearly implicit scheme // Journal of Computational and Applied Mathematics. — 2001. — Vol. 137. — P. 229–243. — [https://doi.org/10.1016/S0377-0427\(01\)00358-2](https://doi.org/10.1016/S0377-0427(01)00358-2).
6. Botchev M. A., Zhukov V. T. Exponential Euler and backward Euler methods for nonlinear heat conduction problems // Lobachevskii J. Math. — 2023. — Vol. 44, no. 1. — P. 10–19. — <https://doi.org/10.1134/S1995080223010067>.
7. Botchev M. A., Sleijpen G. L. G., Vorst H. A. van der. Stability control for approximate implicit time stepping schemes with minimum residual iterations // Appl. Numer. Math. — 1999. — Vol. 31, no. 3. — P. 239–253. — ISSN 0168-9274. — [https://doi.org/10.1016/S0168-9274\(98\)00138-X](https://doi.org/10.1016/S0168-9274(98)00138-X).
8. Brown P. N., Saad Y. Convergence Theory of Nonlinear Newton–Krylov Algorithms // SIAM J. Optimization. — 1994. — Vol. 4, no. 2. — P. 297–330.
9. Choquet R., Erhel J. Newton–GMRES Algorithm Applied to Compressible Flows // Int. J. Numer. Meth. Fluids. — 1996. — Vol. 23. — P. 177–190.

10. *Csomós P., Faragó I., Havasi Á.* Weighted sequential splittings and their analysis // *Comput. Math. Appl.* — 2005. — Vol. 50, no. 7. — P. 1017–1031.
11. *D'yakonov E. G.* Difference systems of second order accuracy with a divided operator for parabolic equations without mixed derivatives // *USSR Comput. Math. Math. Phys.* — 1964. — Vol. 4, no. 5. — P. 206–216.
12. *Davis T. A.* A column pre-ordering strategy for the unsymmetric-pattern multifrontal method // *ACM Trans. Math. Software.* — 2004. — Vol. 30, no. 2. — P. 167–195.
13. *Davis T. A.* Algorithm 832: UMFPACK V4.3—an unsymmetric-pattern multifrontal method // *ACM Trans. Math. Software.* — 2004. — Vol. 30, no. 2. — P. 196–199.
14. *Dekker K., Verwer J. G.* Stability of Runge–Kutta methods for stiff non-linear differential equations. — North-Holland Elsevier Science Publishers, 1984.
15. *Druskin V. L., Knizhnerman L. A.* Krylov subspace approximations of eigenpairs and matrix functions in exact and computer arithmetic // *Numer. Lin. Alg. Appl.* — 1995. — Vol. 2. — P. 205–217.
16. *Druskin V. L., Knizhnerman L. A.* Two polynomial methods of calculating functions of symmetric matrices // *U.S.S.R. Comput. Maths. Math. Phys.* — 1989. — Vol. 29, no. 6. — P. 112–121.
17. *Elman H. C., Silvester D. J., Wathen A. J.* Finite Elements and Fast Iterative Solvers with Applications in Incompressible Fluid Dynamics. — Oxford : Oxford University Press, 2005.
18. *Eshof J. van den, Hochbruck M.* Preconditioning Lanczos approximations to the matrix exponential // *SIAM J. Sci. Comput.* — 2006. — Vol. 27, no. 4. — P. 1438–1457.
19. *Gallopoulos E., Saad Y.* Efficient solution of parabolic equations by Krylov approximation methods // *SIAM J. Sci. Statist. Comput.* — 1992. — Vol. 13, no. 5. — P. 1236–1264. — ISSN 0196-5204. — <http://doi.org/10.1137/0913071>.
20. *Gander M. J., Güttel S.* PARAEXP: A parallel integrator for linear initial-value problems // *SIAM Journal on Scientific Computing.* — 2013. — Vol. 35, no. 2. — P. C123–C142.
21. *Grimm V.* Resolvent Krylov subspace approximation to operator functions // *BIT.* — 2012. — Vol. 52, no. 3. — P. 639–659.

22. *Hairer E., Wanner G.* Solving Ordinary Differential Equations II. Stiff and Differential–Algebraic Problems. — 2nd ed. — Springer–Verlag, 1996. — (Springer Series in Computational Mathematics 14).
23. *Hansen E., Ostermann A.* High-order splitting schemes for semilinear evolution equations // BIT Numerical Mathematics. — 2016. — Vol. 56, no. 4. — P. 1303–1316.
24. *Hipp D., Hochbruck M., Ostermann A.* An exponential integrator for non-autonomous parabolic problems // ETNA. — 2014. — Vol. 41. — P. 497–511. — <http://etna.mcs.kent.edu/volumes/2011-2020/vol41/>.
25. *Hochbruck M., Ostermann A.* Exponential integrators // Acta Numer. — 2010. — Vol. 19. — P. 209–286. — ISSN 0962-4929.
26. *Hochbruck M., Lubich C.* On Krylov Subspace Approximations to the Matrix Exponential Operator // SIAM J. Numer. Anal. — 1997. — Oct. — Vol. 34, no. 5. — P. 1911–1925.
27. *Hundsdorfer W., Verwer J. G.* Numerical Solution of Time-Dependent Advection–Diffusion–Reaction Equations. — Springer Verlag, 2003.
28. *Janssen J., Vandewalle S.* Multigrid Waveform Relaxation of Spatial Finite Element Meshes: The Continuous-Time Case // SIAM J. Numer. Anal. — 1996. — Vol. 33, no. 2. — P. 456–474.
29. *Jon J., Klaus S.* The Liouville–Bratu–Gelfand Problem for Radial Operators // Journal of Differential Equations. — 2002. — Vol. 184. — P. 283–298. — <https://doi.org/10.1006/jdeq.2001.4151>.
30. *Kooij G. L., Botchev M. A., Geurts B. J.* A block Krylov subspace implementation of the time-parallel Paraexp method and its extension for nonlinear partial differential equations // Journal of Computational and Applied Mathematics. — 2017. — Vol. 316, Supplement C. — P. 229–246. — <https://doi.org/10.1016/j.cam.2016.09.036>.
31. *Kooij G., Botchev M. A., Geurts B. J.* An exponential time integrator for the incompressible Navier–Stokes equation // SIAM J. Sci. Comput. — 2018. — Vol. 40, no. 3. — B684–B705. — <https://doi.org/10.1137/17M1121950>.
32. *Krukier L. A.* Implicit difference schemes and an iterative method for solving them for a certain class of systems of quasi-linear equations // Sov. Math. — 1979. — Vol. 23, no. 7. — P. 43–55. — Translation from Izv. Vyssh. Uchebn. Zaved., Mat. 1979, No. 7(206), 41–52 (1979).

33. *Lebedev V. I.* Explicit difference schemes for solving stiff systems of ODEs and PDEs with complex spectrum // Russian J. Numer. Anal. Math. Modelling. — 1998. — Vol. 13, no. 2. — P. 107–116. — <https://doi.org/10.1515/rnam.1998.13.2.107>.
34. *Lelarasmee E., Ruehli A. E., Sangiovanni-Vincentelli A. L.* The Waveform Relaxation Method for Time-Domain Analysis of Large Scale Integrated Circuits // IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. — 1982. — Vol. 1, no. 3. — P. 131–145.
35. *Lokutsievskii V. O., Lokutsievskii O. V.* On numerical solution of boundary value problems for equations of parabolic type // Soviet Math. Dokl. — 1987. — Vol. 34, no. 3. — P. 512–516. — <https://zbmath.org/?q=an:0637.65098>.
36. *Lubich C., Ostermann A.* Multi-grid dynamic iteration for parabolic equations // BIT Numerical Mathematics. — 1987. — Vol. 27, issue 2. — P. 216–234. — ISSN 0006-3835. — URL: <http://doi.org/10.1007/BF01934186> ; 10.1007/BF01934186.
37. *Miekkala U., Nevanlinna O.* Convergence of Dynamic Iteration Methods for Initial Value Problems // SIAM Journal on Scientific and Statistical Computing. — 1987. — Vol. 8, no. 4. — P. 459–482. — <https://doi.org/10.1137/0908046>.
38. *Miekkala U., Nevanlinna O.* Iterative solution of systems of linear differential equations // Acta Numerica. — 1996. — Vol. 5. — P. 259–307. — <https://doi.org/10.1017/S096249290000266X>.
39. *Moret I., Novati P.* RD rational approximations of the matrix exponential // BIT. — 2004. — Vol. 44. — P. 595–615.
40. *Nevanlinna O., Odeh F.* Remarks on the convergence of waveform relaxation method // Numerical functional analysis and optimization. — 1987. — Vol. 9, no. 3/4. — P. 435–445. — <https://doi.org/10.1080/01630568708816241>.
41. *Newton A. R., Sangiovanni-Vincentelli A. L.* Relaxation-based electrical simulation // IEEE Transactions on Electron Devices. — 1983. — Vol. 30, no. 9. — P. 1184–1207.
42. *Park T. J., Light J. C.* Unitary quantum time evolution by iterative Lanczos reduction // J. Chem. Phys. — 1986. — Vol. 85. — P. 5870–5876.
43. *Shampine L. F., Reichelt M. W.* The MATLAB ODE suite // SIAM J. Sci. Comput. — 1997. — Vol. 18, no. 1. — P. 1–22. — Available at www.mathworks.com.

44. *Sommeijer B. P., Shampine L. F., Verwer J. G.* RKC: An Explicit Solver for Parabolic PDEs // J. Comput. Appl. Math. — 1998. — Vol. 88. — P. 315–326. — [https://doi.org/10.1016/S0377-0427\(97\)00219-7](https://doi.org/10.1016/S0377-0427(97)00219-7).
45. *Tal-Ezer H.* Spectral Methods in Time for Parabolic Problems // SIAM J. Numer. Anal. — 1989. — Vol. 26, no. 1. — P. 1–11.
46. *Tromeur-Dervout D., Vassilevski Y.* Choice of initial guess in iterative solution of series of systems arising in fluid flow simulations // Journal of Computational Physics. — 2006. — Vol. 219, no. 1. — P. 210–227. — ISSN 0021-9991. — URL: <https://www.sciencedirect.com/science/article/pii/S0021999106001483>.
47. *van der Houwen P. J., Sommeijer B. P.* On the internal stability of explicit m -stage Runge–Kutta methods for large values of m . // Z. Angew. Math. Mech. — 1980. — Vol. 60. — P. 479–485.
48. *van der Vorst H. A.* An Iterative Solution Method for Solving $f(A)x = b$, using Krylov Subspace Information Obtained for the Symmetric Positive Definite Matrix A // J. Comput. Appl. Math. — 1987. — Vol. 18. — P. 249–263.
49. *Vandewalle S.* Parallel Multigrid Waveform Relaxation for Parabolic Problems. — Stuttgart : Teubner, 1993.
50. *Verstappen R. W. C. P., Veldman A. E. P.* Symmetry-preserving discretization of turbulent flow // J. Comput. Phys. — 2003. — Vol. 187, no. 1. — P. 343–368. — [http://doi.org/10.1016/S0021-9991\(03\)00126-8](http://doi.org/10.1016/S0021-9991(03)00126-8).
51. *Verwer J. G., Sommeijer B. P., Hundsdorfer W.* RKC time-stepping for advection–diffusion–reaction problems // J. Comput. Phys. — 2004. — Vol. 201, no. 1. — P. 61–79. — ISSN 0021-9991. — URL: <https://www.sciencedirect.com/science/article/pii/S0021999104001925>.
52. Waveform Relaxation: Theory and Practice : tech. rep. / J. White [et al.] ; EECS Department, University of California, Berkeley. — 1985. — UCB/ERL M85/65. — www.eecs.berkeley.edu/Pubs/TechRpts/1985/543.html.
53. *Yanenko N. N.* The method of fractional steps. The solution of problems of mathematical physics in several variables. — New York : Springer-Verlag, 1971. — P. viii+160.
54. *Zhukov V. T.* Explicit methods of numerical integration for parabolic equations // Math. Models Comput. Simul. — 2011. — Vol. 3, no. 3. — P. 311–332. — <https://doi.org/10.1134/S2070048211030136>.
55. *Zhukov V. T., Novikova N., Feodoritova O. B.* On direct solving conjugate heat transfer of gas mixture and solid body // KIAM Preprint. — 2023. — Vol. 2023, no. 12. — <https://doi.org/10.20948/prepr-2023-12>.