

УДК 519.632.4

Особенности процесса накопления погрешностей при решении задач для простейших уравнений математической физики конечно-разностными методами*

В.П. Житников¹, Н.М. Шерыхалина¹, Р.Р. Муксимова²

¹Уфимский государственный авиационный технический университет, ул. К. Маркса, 12, Уфа, 450000

²Санкт-Петербургский государственный университет гражданской авиации, ул. Пилотов, 38, Санкт-Петербург, 196210
E-mails: zhitnik@mail.ru (Житников В.П.), n_sher@mail.ru (Шерыхалина Н.М.), rose.r.mux@gmail.com (Муксимова Р.Р.)

Житников В.П., Шерыхалина Н.М., Муксимова Р.Р. Особенности процесса накопления погрешностей при решении задач для простейших уравнений математической физики конечно-разностными методами // Сиб. журн. вычисл. математики / РАН. Сиб. отд.-ние. — Новосибирск, 2016. — Т. 19, № 2. — С. 139–152.

Рассматривается смешанная задача для одномерного уравнения теплопроводности с несколькими вариантами начальных и краевых условий. Для решения применяются явная и неявная схемы. Для неявной схемы при решении системы уравнений используются методы прогонки и итераций. Для анализа погрешностей метода и округления применяется численная фильтрация конечной последовательности результатов, полученной для различных сеток с возрастающим числом узловых точек. Кроме того, для исследования погрешности округления сравниваются результаты, полученные при нескольких длинах мантииссы машинного слова. Аналогичными методами исследуется численное решение смешанной задачи для волнового уравнения.

Обнаружено явление возникновения детерминированных зависимостей погрешности численного метода и округления от пространственной координаты, времени и числа узлов. На основе анализа результатов вычислительного эксперимента для разных вариантов условий задач построены модели источников для описания поведения погрешностей во времени. В соответствии с этим моделями, подтвержденными экспериментом, погрешности с течением времени могут в зависимости от условий возрастать, уменьшаться или стабилизироваться аналогично изменению энергии или массы.

DOI: 10.15372/SJNM20160202

Ключевые слова: уравнение теплопроводности, волновое уравнение, явная и неявная схемы, число Куранта, модели погрешности, численная фильтрация.

Zhitnikov V.P., Sherykhalina N.M., Muksimova R.R. The peculiarities of error accumulation in solving problems for simple equations of mathematical physics by finite difference methods // Siberian J. Num. Math. / Sib. Branch of Russ. Acad. of Sci. — Novosibirsk, 2016. — Vol. 19, № 2. — P. 139–152.

A mixed problem for a one-dimensional heat equation with several versions of initial and boundary conditions is considered. Explicit and implicit schemes are applied for the solution. The sweep method and the iteration methods are used for the implicit scheme for solving the implicit system of equations. The numerical filtration of a finite sequence of results obtained for different grids with an increasing number of nodal points is used to analyze errors of the method and rounding. In addition, to investigate the rounding errors, the results obtained with several lengths of the machine word mantissa are compared. The numerical solution of the mixed problem for the wave equation is studied by similar methods.

The occurrence of deterministic dependencies of the error in the numerical method and the rounding on spatial coordinates, time and the number of nodes is revealed. The source models to describe the behavior of errors in terms of time are based on the analysis of the results of numerical experiments for different versions

*Работа выполнена при финансовой поддержке Минобрнауки России в рамках базовой части госзадания образовательным организациям высшего образования. Код программы 2229.

of conditions of problems. In accord with such models, which were verified by the experiment, the errors can increase, decrease or stabilize depending on conditions over time similar to changing the energy or mass.

Keywords: *heat equation, explicit and implicit schemes, the Courant number, model error, numerical filtration.*

Введение

Ранее было проведено исследование составляющих погрешностей различных численных методов с помощью фильтрации [1–3].

Фильтрация проводилась на основе априорной модели зависимости результата вычисления искомой величины z от параметра дискретизации n . Для многих вычислительных методов численного дифференцирования, интегрирования и т. д. эта зависимость представляется в виде суммы нескольких слагаемых (компонент):

$$z_n - z = c_1 n^{-k_1} + c_2 n^{-k_2} + \dots + c_L n^{-k_L} + \Delta(n), \quad (1)$$

где z_n — результат вычисления, полученный при значении параметра дискретизации (числа узлов сетки, числа слагаемых в сумме и т. п.), равном n ; k_1, \dots, k_L — произвольные известные действительные числа ($k_1 < k_2 < \dots < k_L$); c_j — неизвестные коэффициенты.

Составляющая $\Delta(n)$ может состоять из не вошедших в сумму слагаемых, остаточного члена, погрешности округления и других составляющих, порожденных несовершенством численного алгоритма и его программной реализации. Существенно то, что величина $\Delta(n)$ не имеет априорной оценки, и предполагается возможным возрастание этой величины при возрастании n .

Для некоторых численных методов компоненты зависимости погрешности от n имеют показательный вид [3, 4].

Исследование погрешностей различных численных методов: численного дифференцирования, интегрирования, решение различных задач для обыкновенных дифференциальных уравнений [1, 2] показало наличие нескольких компонент погрешности степенного вида (1). С помощью фильтрации удалось подавить эти регулярные компоненты и уменьшить погрешность искомых параметров. Подавление этих компонент позволило исследовать поведение составляющей $\Delta(n)$, главная часть которой состояла из погрешности округления, вызванной ограниченностью длины мантиссы машинного слова. В переносимых методах зависимость $\Delta(n)$ имела хаотичный характер и моделировалась случайной величиной с нулевым математическим ожиданием и среднеквадратичным отклонением, растущим как cn^k с ростом n (где c — величина, близкая к 10^{-M} , M — длина мантиссы машинного слова, $k > 0$ — показатель, зависящий от рассматриваемого метода).

При подробном исследовании результатов решения задач для уравнения теплопроводности выяснилось, что $\Delta(n)$ имеет не случайный характер, а возникают зависимости от параметров n , x и t , имеющие детерминированный вид. Это будет показано ниже.

1. Постановка задачи

Рассматривается смешанная задача для одномерного уравнения теплопроводности (диффузии)

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0, \quad 0 < x < 1, \quad 0 < t \leq T. \quad (2)$$

Начальное условие

$$u(x, 0) = f(x), \quad 0 < x < 1. \quad (3)$$

Краевые условия:

$$u(0, t) = \varphi_1(t), \quad u(1, t) = \varphi_2(t), \quad t \geq 0. \quad (4)$$

Рассматривается 3 варианта краевых и начальных условий:

- I. $f(x) = \sin(\pi x)$, $\varphi_1(t) = \varphi_2(t) = 0$,
 II. $f(x) = 0$, $\varphi_1(t) = 1$, $\varphi_2(t) = 0$,
 III. $f(x) = 1 - x$, $\frac{\partial u}{\partial x}(0, t) = \frac{\partial u}{\partial x}(1, t) = 0$.

Точные решения для вариантов I–III:

$$u(x, t) = e^{-\pi^2 t} \sin(\pi x), \quad (5)$$

$$u(x, t) = -\frac{2}{\pi} \sum_{m=1}^{\infty} \frac{1}{m} e^{-m^2 \pi^2 t} \sin m\pi x + 1 - x, \quad (6)$$

$$u(x, t) = \frac{1}{2} + \frac{4}{\pi^2} \sum_{m=1}^{\infty} \frac{1}{(2m-1)^2} e^{-(2m-1)^2 \pi^2 t} \cos(2m-1)\pi x. \quad (7)$$

Задачи с вариантами условий II и III моделируют эксперименты, проводимые с образцами пород, добытых при бурении скважин [5].

Задачи решаются методом конечных разностей. На плоскости (x, t) строится сетка с шагом h по переменной x ($x_i = ih$, $i = 0, \dots, n$, $h = 1/n$) и с шагом τ по переменной t ($t_j = j\tau$, $j = 0, \dots, m$, $\tau = T/m$). Вводятся обозначения $u(x_i, t_j) = u_{i,j}$.

Обозначим $\lambda = \tau/h^2$ (число Куранта). Тогда разностные уравнения для неявной схемы запишутся в виде:

$$\begin{aligned} -\lambda u_{i-1,j} + (1 + 2\lambda) u_{i,j} - \lambda u_{i+1,j} &= u_{i,j-1}, \quad i = 1, \dots, n-1; \quad j = 1, \dots, m, \\ u_{i,0} &= f(x_i), \quad u_{0,j} = \varphi_1(t_j), \quad u_{n,j} = \varphi_2(t_j). \end{aligned} \quad (8)$$

Для явной схемы —

$$u_{i,j+1} = \lambda u_{i-1,j} + (1 - 2\lambda) u_{i,j} + \lambda u_{i+1,j}, \quad i = 1, \dots, n-1; \quad j = 0, \dots, m-1. \quad (9)$$

Исследование устойчивости [6], проводимое путем возмущения правой части (2) некоторой величиной $\xi(x, t)$ и анализа отклика $v(x, t)$ при нулевых начальных условиях

$$\frac{\partial v}{\partial t} - \frac{\partial^2 v}{\partial x^2} = \xi(x, t), \quad \xi(x, 0) = 0, \quad (10)$$

показывает, что справедливо неравенство

$$\max_{0 \leq j \leq m} \max_{0 \leq i \leq n} |v_{i,j}| \leq T \max_{0 \leq j \leq m} \max_{0 \leq i \leq n} |\xi_{i,j}|, \quad (11)$$

причем для явной схемы это справедливо при $\lambda \leq \frac{1}{2}$. Отметим, что оценка (11) не учитывает погрешности округления.

Задачей настоящей работы является исследование как убывающих регулярных компонент погрешности, так и составляющей $\Delta(n)$ при различных значениях независимых переменных x , t , n и m (или x , t , n и λ).

2. Результаты исследований

Для исследования величина n изменялась от значения $n = 5$ до $n = 5120$ последовательным удвоением текущего значения.

На рис. 1 в логарифмической шкале представлены результаты сравнения с точным решением данных, полученных при численном решении задачи с условиями вида I по явной схеме (9) для $\lambda = 1/8$; $x = 0.2$; $T = 0.1$. По оси ординат отложены десятичные логарифмы (со знаком минус) относительных погрешностей, т. е. точность, выраженная в количестве точных десятичных значащих цифр. По оси абсцисс отложены десятичные логарифмы n . При таком выборе шкалы каждая компонента зависимости (1) отображается прямой линией.

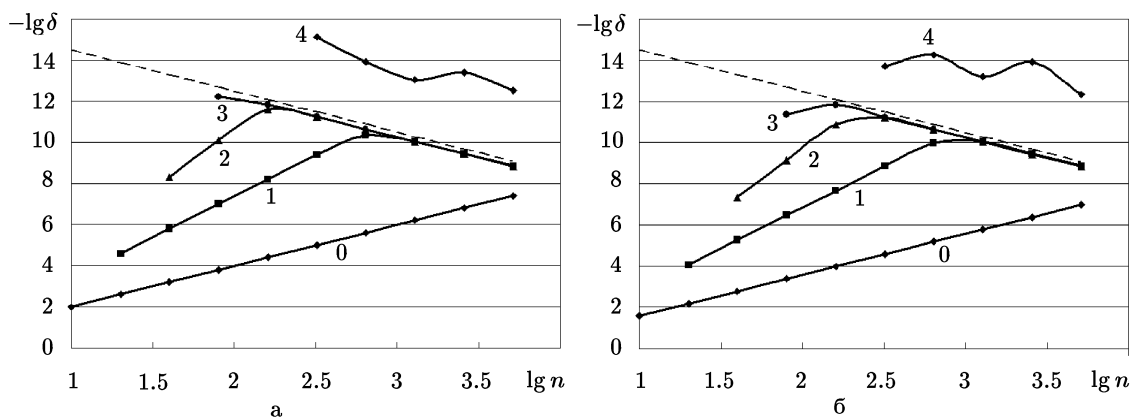


Рис. 1. Оценки погрешностей для явной схемы при $\lambda = 1/8$: а — искомых величин $u(x, t)$; б — производных $\frac{\partial u}{\partial x}(x, t)$. Пунктирная прямая $y = 16.5 - 2 \lg n$

Линии, обозначенные цифрой 0 на рис. 1, соответствуют точности вычисленных непосредственно значений разностной производной; цифрами 1, 2 и т. д. — точности результатов первой, второй и т. д. фильтрации.

В результате каждой фильтрации меняется тангенс угла наклона линий (равный k_1, k_2, \dots для $j = 0, 1, \dots$) и происходит сдвиг вверх, что свидетельствует о повышении точности. Такое поведение линий подтверждает наличие и результат устранения конкретных компонент зависимости.

На рис. 1 линии 1–3 имеют два четко выделенных участка. Первый участок имеет наклон, соответствующий показателю степенной функции. На втором участке угловой коэффициент приблизительно равен -2 . Это связано с преобладанием составляющей погрешности $\Delta(n)$, которую можно аппроксимировать прямой $y = 16.5 - 2 \lg n$.

Анализ этих графиков позволяет установить, что составляющие погрешности метода $c_l n^{-k_l}$ имеют такие же значения показателей как вторая разностная частная производная $\frac{\partial^2 u}{\partial x^2}(x, t) \approx \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2}$, т. е. $k_l = 2, 4, 6, \dots$. Отсутствие нечетных слагаемых, которые имеют место в разложении первой частной производной по времени, объясняется тем, что для сохранения значения λ при удвоении n величину m приходится увеличивать в 4 раза, т. е. $\lambda m = n^2$.

Однако графики показывают два “нетипичных” свойства накапливающейся погрешности округления. Первое обнаруживается при анализе ее зависимости от n . Зависимость $y = 16.5 - 2 \lg n$ на рис. 1 является, в отличие от численного дифференцирования и интегрирования, не хаотичной, а близкой к закономерной. Эта компонента cn^2 может быть

отфильтрована, и результат фильтрации, подтверждающий “закономерность” этой компоненты, показан на рис. 1 линиями с номером 4.

Второе “нетипичное” свойство обнаруживается при сравнении линий на рис. 1а, 1б. При численном определении первой производной погрешность округления отсчитывается от исходного уровня нерегулярной погрешности Δ_0 как $\sim \frac{\Delta_0}{h}$, при этом $\frac{\Delta_0}{h} \approx \frac{cn^2}{h} = cn^3$. Однако вычислительный эксперимент обнаруживает погрешность на уровне cn^2 . Тем самым, можно сделать вывод о том, что погрешность округления, имеющая порядок cn^2 , накопленная при решении задачи с помощью разностной схемы (9), имеет регулярный характер в зависимости также и от переменной x .

На вопрос, является ли источником данной регулярной компоненты $\Delta(n)$ погрешность округления, помогает ответить применение попарного суммирования при вычислении по явной схеме. Из рис. 2а следует, что в этом случае погрешность округления при увеличении n ограничена на уровне 10^{-15} – 10^{-16} , т. е. отмеченный эффект полностью исчезает.

Результаты решения задачи с помощью неявной схемы с условиями II (6) при $m = 2n$ при применении метода итераций приведены на рис. 2б. В отличие от рассмотренных выше результатов решения с помощью явной схемы, имеют место как четные, так и нечетные составляющие погрешности метода cn^{-k_l} . Это объясняется линейным ростом m (в связи с чем $\lambda = 0.5n^2/m \rightarrow \infty$) и первым порядком аппроксимации первой частной производной по t . Что касается погрешности округления, то она имеет такие же свойства, как и в явной схеме (квадратично растет с увеличением n). Разница заключается в том, что погрешность решения системы уравнений итерационным методом ограничивалась значением $\varepsilon = 10^{-9}$. Это значение, а не длина мантииссы машинного слова (около 16-ти десятичных разрядов), определяло коэффициент зависимости cn^2 (хотя c совпадает с ε только приближенно, но имеет место приближенно прямо-пропорциональная зависимость от ε).

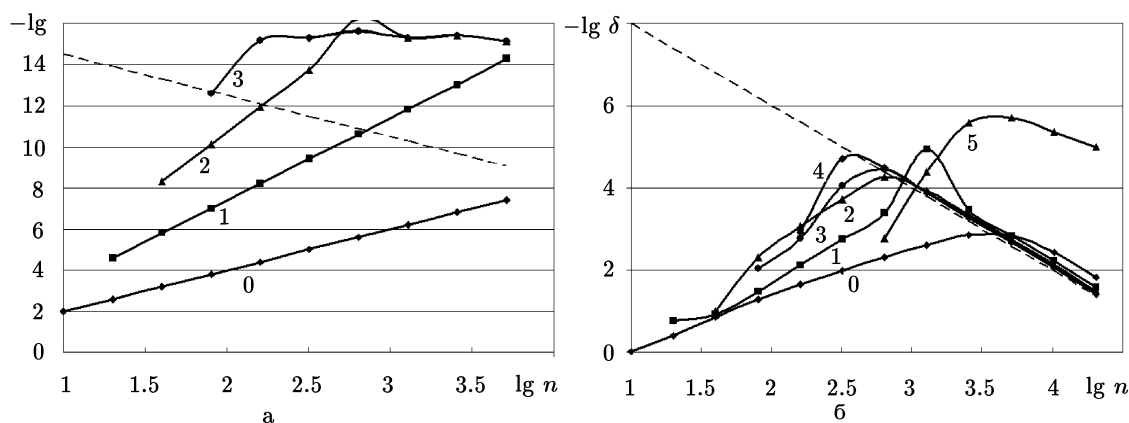


Рис. 2. Оценки погрешностей: а — для явной схемы с попарным суммированием; б — для неявной схемы с итерационным решением системы уравнений (8). Пунктирные прямые $y = 16.5 - 2 \lg n$ и $y = 10 - 2 \lg n$

В связи с приведенными результатами представляет интерес исследование зависимости составляющей погрешности $\Delta(n)$ также и от x . На рис. 3 приведены зависимости погрешности округления (отнесенной к точному решению и к n^2) от x при разных n для неявной схемы: на рис. 3а — для условий вида I, на рис. 3б — для условий вида II (линии 1–4 соответствуют $n = 1280; 2560; 5120; 10240$).

Из рис. 3 видно, что относительная погрешность округления почти не зависит от x . Это само по себе не вызывает вопросов, так как погрешность округления определяется порядком числа в машинном представлении (в данном случае это $u_{i,j}$). Вызывает вопросы регулярность этой зависимости, причем как по x , так и по n .

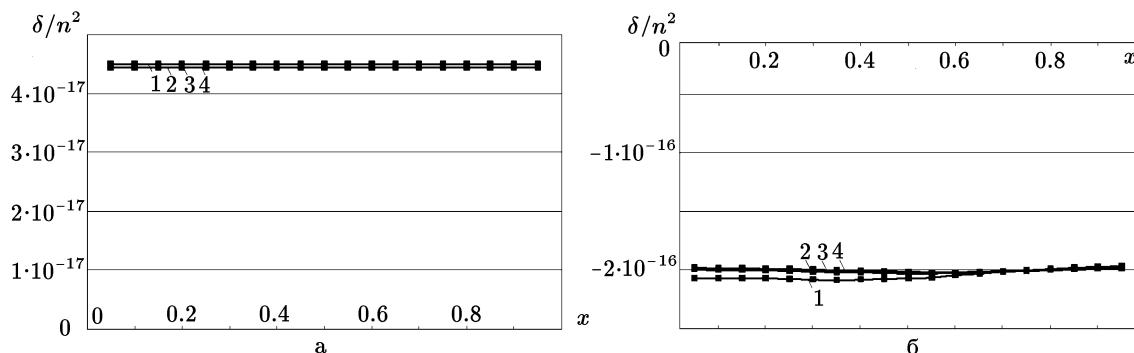


Рис. 3. Зависимости относительной погрешности округления от x при разных n для неявной схемы: а — с условиями вида I; б — с условиями вида II

Отметим, что при зависимости погрешности вида $\delta \approx cn^2$ отношение δ/n^2 приближенно определяет константу c . Представляет интерес исследовать зависимость $c(\lambda)$. На рис. 4 представлены зависимости относительной погрешности округления от λ для неявной схемы с условиями вида I. Графики построены с разными шагами определения параметра λ : на рис. 4б шаг в 4 раза меньше, чем на рис. 4а. Видна “фрактальность” этой зависимости. Итак, по параметру λ зависимость имеет квазислучайный характер с ограничением $|c(\lambda)| \leq 10^{-16}$. Для других видов условий это свойство также имеет место.

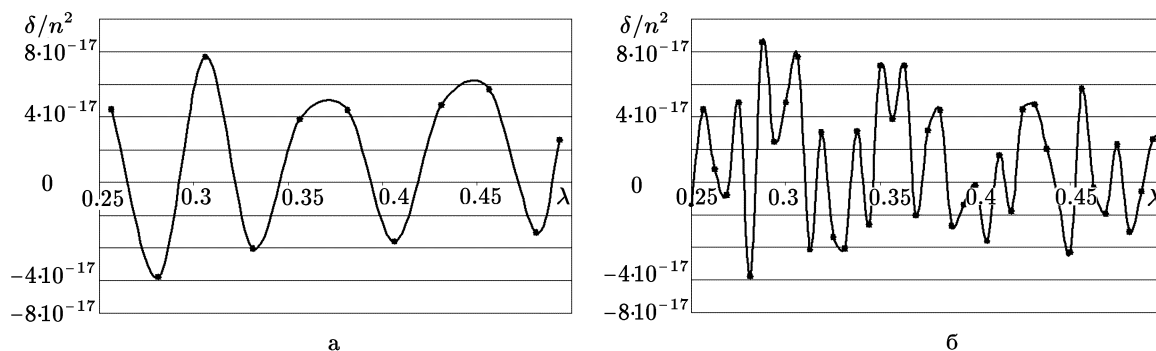


Рис. 4. Графики зависимости относительной погрешности округления от числа Куранта λ для неявной схемы, построенные с разными шагами определения параметра: а — по 10 точкам; б — по 40 точкам

Перейдем к исследованию зависимости погрешности округления от времени. На рис. 5а приведены зависимости относительной погрешности округления (отнесенной к точному решению (5)) от времени для неявной схемы при $\lambda = 0.5$ с шагом по безразмерному времени $\tau = 0.05$ (в диапазоне $0 \leq t \leq 1$). Погрешность округления найдена путем вычитания результатов, вычисленных с машинными словами с разной длиной мантиссы (16 и 19 десятичных разрядов).

Эксперимент показывает наличие закономерной зависимости как от x , так и от t . Тем не менее, исследование первых шагов работы алгоритма (8) (результаты, полученные

алгоритмом (9) имеют аналогичный вид) показывает, что закономерные зависимости складываются из квазислучайных элементов.

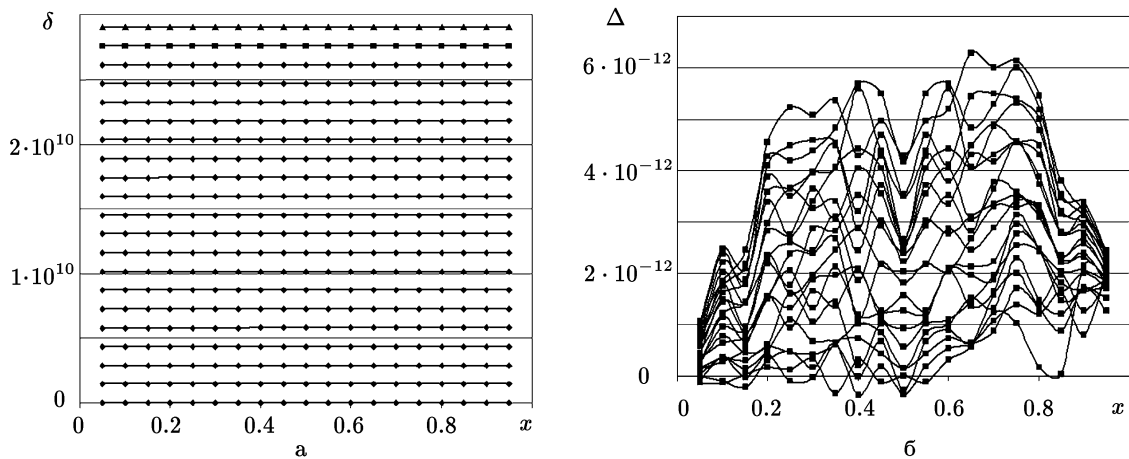


Рис. 5. Зависимости погрешности округления от времени при $\lambda = 0.5$: а — с шагом $\tau = 0.05$; б — первые шаги при $n = 2560$

Для большей наглядности на рис. 6 приведены графики зависимостей погрешности округления от времени для $x = 0.05$ (линии 1) и $x = 0.5$ (линии 2). Видно, что абсолютная погрешность (рис. 6а) сначала возрастает, а затем начинает убывать и приближается к нулю. При этом зависимость относительной погрешности от времени близка к линейной (рис. 6б).

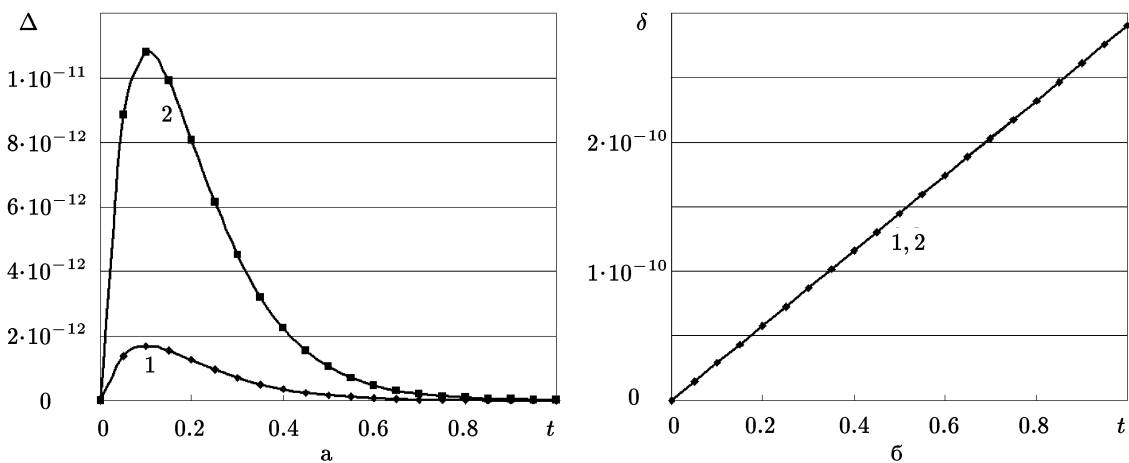


Рис. 6. Зависимости погрешности округления от времени при $\lambda = 0.5$: а — абсолютной; б — относительной

Такой же вид имеют зависимости от времени погрешности численного метода, только при этом порядок величин отличается, и зависимость от n имеет противоположный вид (погрешность округления растет, а погрешность метода уменьшается при увеличении n).

Заметим, что теоретическая оценка сверху (11) показывает монотонно растущую с увеличением t зависимость, что в данном случае не имеет места. Этот факт опровергает наличие чисто накопительного механизма образования погрешности, лежащего в основе оценки (11).

3. Модели образования и изменения погрешности округления

Экспериментальный факт линейной зависимости относительной погрешности округления от времени позволяет предложить следующую модель образования и изменения погрешности округления. Поскольку величина погрешности округления определяется порядком операндов формул (8) и (9), рассмотрим задачу, аналогичную (10):

$$\frac{\partial v}{\partial t} - \frac{\partial^2 v}{\partial x^2} = c_1 u(x, t), \quad v(0, t) = v(1, t) = 0, \quad (12)$$

или для условий вида III: $\frac{\partial v}{\partial x}(0, t) = \frac{\partial v}{\partial x}(1, t) = 0$, где $c_1 u(x, t)$ — модель возникающей при выполнении арифметических действий погрешности округления (возмущения); c_1 — некоторое число, не зависящее от x и t , которое связано с коэффициентом обнаруженной выше зависимости: $c_1 = cn^2$ ($|c| \leq 10^{-16}$); $u(x, t)$ — решение задачи (2); v — отклик на возмущение правой части (12).

Для задачи с условиями вида I: $u(x, t) = e^{-\pi^2 t} \sin(\pi x)$, тогда частное решение (12) $v(x, t) = tc_1 e^{-\pi^2 t} \sin(\pi x) = tc_1 u(x, t)$, что и показывает рис. 6б.

Для задачи с условиями вида II точное решение дается формулой (6). Для простоты рассмотрим предельное при $t \rightarrow \infty$ решение $u(x, \infty) = 1 - x$. Частным решением (12) с данной правой частью является $v(x, t) = c_1 [1 - x - (1 - x)^3]/6$.

На рис. 7а приведены зависимости погрешности округления, отнесенной к этому предельному решению. Видно, что при увеличении времени устанавливается величина, слабо зависящая от x .

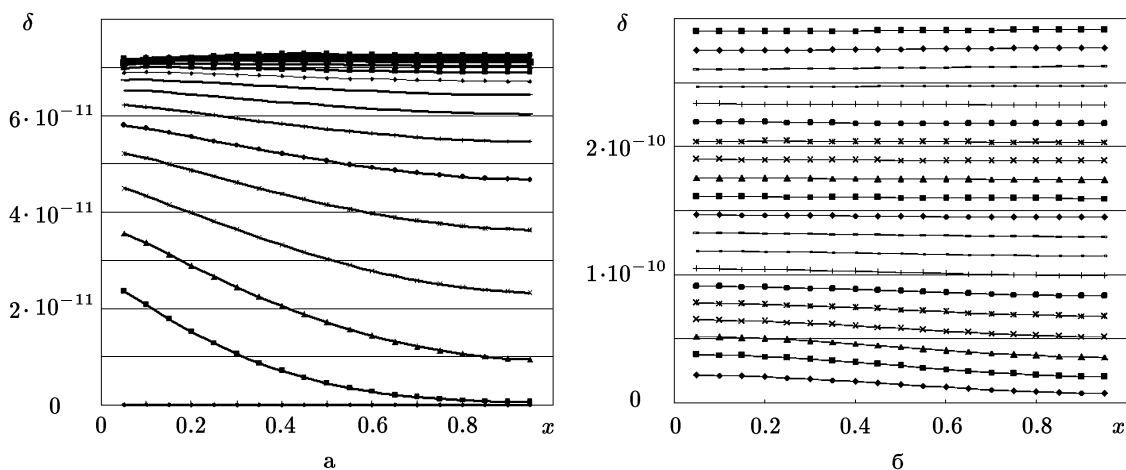


Рис. 7. Зависимости погрешности округления от времени при $\lambda = 0.5$ с шагом $\tau = 0.05$: а — для условий типа II; б — для условий типа III

Аналогично для задачи с условиями вида III точное решение дается формулой (7) с предельным решением $u(x, \infty) = 1/2$. Частным решением (12) с данной правой частью является $v(x, t) = c_1 t/2$. Зависимости погрешности округления, приведенные на рис. 7б, аналогичны рис. 7а.

На рис. 8 даны графики зависимостей относительной погрешности округления от времени для $x = 0.05$ (линии 1) и $x = 0.5$ (линии 2). Видно, что при условиях II погрешность возрастает и устанавливается на уровне $\sim 7 \cdot 10^{-11}$ (рис. 6а), а при условиях III линейно

растет, что согласуется с предложенной моделью (12). В последнем случае абсолютная и относительная погрешности отличаются только коэффициентом.

Следует отметить, что условия III (второго рода) соответствуют изолированным концам, когда поток (жидкости, тепла) через границы отсутствует. При этом погрешность округления накапливается линейно, как предсказывает оценка (11). В остальных случаях устанавливается баланс, т. е. стационарное решение, когда количество принятой от источника сущности (т. е. жидкости, тепла или погрешности) равно количеству ушедшей через границы. Интересно, что погрешность округления ведет себя так же, как жидкость или тепло, т. е. подчиняется законам физики, например закону сохранения.

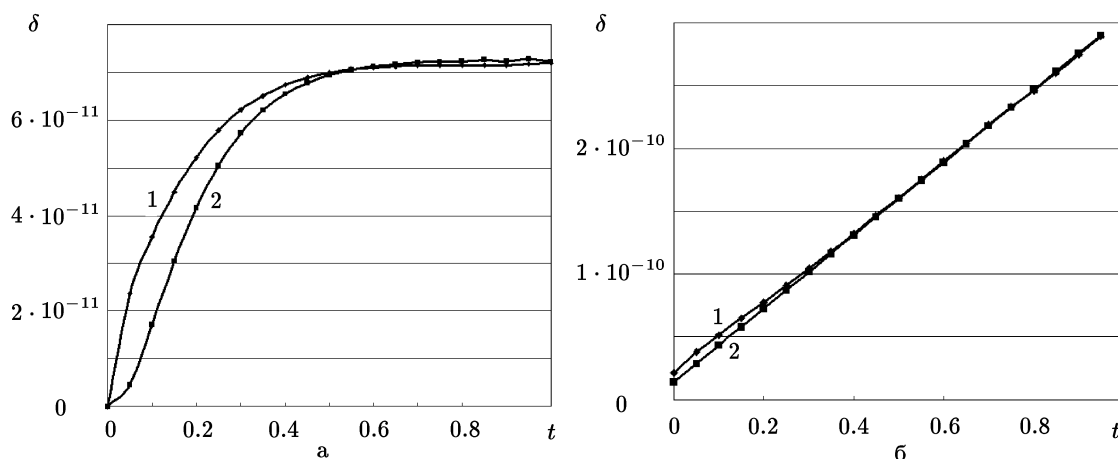


Рис. 8. Зависимости относительной погрешности округления от времени при $\lambda = 0.5$, $n = 2560$: а — для условий типа II; б — для условий типа III

На рис. 9 определяется коэффициент $c = c_1/n^2$ для вариантов условий I и III для разных n (кривые 1–7 соответствуют $n = 20; 40; 80; 160; 320; 640; 1280$). Видно, что для $n \geq 160$ кривые близки к прямым и практически сливаются.

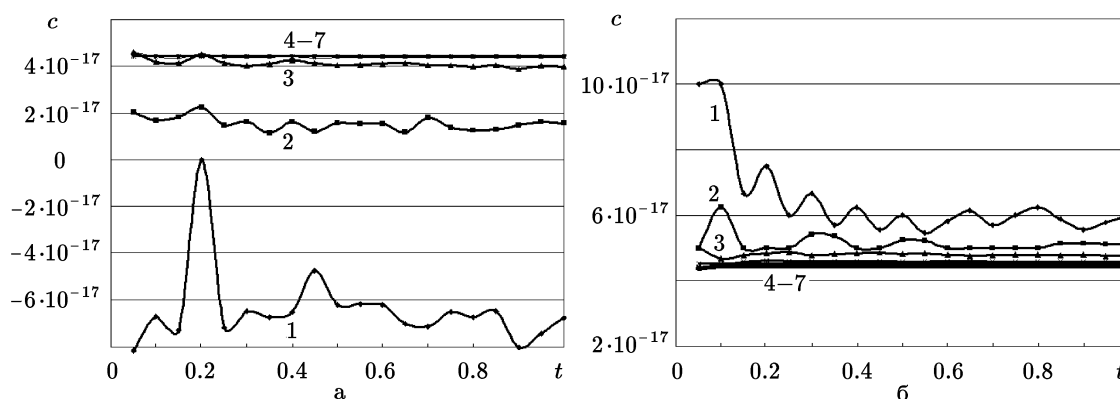


Рис. 9. Зависимости относительной погрешности округления от времени при $\lambda = 0.5$: а — для условий типа I; б — для условий типа III

Отметим, что частные решения задачи (12), приведенные выше, являются аттракторами, так как общие решения (при различных начальных условиях) содержат, кроме частного решения, экспоненциально убывающие компоненты, что и показывает вычислительный эксперимент.

Результаты, представленные на рис. 3–9, получены с помощью неявной схемы при решении систем линейных алгебраических уравнений (СЛАУ) (8) методом прогонки.

Рассмотрим результаты решения задачи (2) по неявной схеме с итерационным методом решения СЛАУ. В этом случае источником погрешности является не округление при выполнении арифметических операций, а погрешность решения СЛАУ, регулируемая сравнением максимальной по модулю разности последовательных итераций с заданным числом ε (см. рис. 2б). В связи с этим правую часть дифференциального уравнения, моделирующего погрешность, необходимо изменить

$$\frac{\partial v}{\partial t} - \frac{\partial^2 v}{\partial x^2} = c_2 g(x), \quad (13)$$

где $g(x)$ — множитель, зависящий от x , соответствующий наиболее медленной из убывающих компонент разложений (5)–(7). Для задачи с условиями I и II: $g(x) = \sin(\pi x)$, с условиями III: $g(x) = \cos(\pi x)$. Для величины c_2 верхней оценкой является значение ε , т. е. $|c_2| \leq \varepsilon$. Это можно объяснить тем, что при шаге по времени при достаточно больших t разность между значениями искомого $u(x, t) - u(x, t + \tau) \approx c_3 e^{-\pi^2 t} (1 - e^{-\pi^2 \tau}) g(x)$, что является начальной погрешностью в итерационном процессе. В процессе решения СЛАУ невязки уменьшаются, но распределение погрешности по x остается приблизительно пропорциональным $g(x)$. Итерационный процесс заканчивается, когда абсолютная величина разности, т. е. коэффициента перед $g(x)$, станет меньше ε . Частным решением такой задачи является $v(x, t) = \pi^{-2} c_2 g(x)$, что и подтверждается вычислительным экспериментом.

4. Особенности образования погрешности округления для явной схемы

Для анализа погрешности округления явная схема имеет преимущество по сравнению с неявной в связи с простотой проводимых арифметических операций. Для уменьшения числа операций умножения, приводящих к дополнительной погрешности, видоизменим схему (9):

$$u_{i,j+1} = u_{i,j} + \lambda (u_{i-1,j} - 2u_{i,j} + u_{i+1,j}). \quad (14)$$

Кроме того, вместо формулы $\lambda = \tau/h^2$ зададим непосредственно $\lambda = \frac{1}{k}$, $k = 2, 3, 4, \dots$. При $T = 0.1$, $n \geq 10$, $h = 1/n$, $\tau = T/m$ число $m = Tn^2/\lambda = kn^2/10$ является целым.

После таких видоизменений алгоритма эффект установления закономерной зависимости погрешности округления от n исчезает (рис. 10а, кривые 1–5 соответствуют $n = 320$; 640; 1280; 2560; 5120). При этом закономерность по x сохраняется. Однако при малом изменении λ на величину $\Delta_\lambda = -1 \cdot 10^{-16}$ эффект снова появляется при $c \approx 1 \cdot 10^{-17}$ (рис. 10б).

При изменении знака приращения $\Delta_\lambda = 1 \cdot 10^{-16}$ меняется знак c при сохранении модуля. Более того, при увеличении Δ_λ до величины порядка $1000 \cdot 10^{-16}$ значение $c \approx -1 \cdot 10^{-17}$ практически не меняется. При дальнейшем увеличении Δ_λ зависимость погрешности от n начинает “распадаться”.

Тем самым, зависимость $c(\Delta_\lambda)$ представляет собой функцию типа $c_0 \operatorname{sign}(\Delta_\lambda)$ при $c_0 \approx 1 \cdot 10^{-17}$. Аналогичное поведение имеет зависимость $c(\Delta_\lambda)$ при отсчете от других $\lambda = \frac{1}{k}$ при четных k (расчеты проводились до $k = 20$). Более того, значение $c_0 \approx -1 \cdot 10^{-17}$ сохраняется для всех четных k .

Но для нечетных k эффект установления квадратичной зависимости погрешности округления от n не возникает даже при больших Δ_λ .

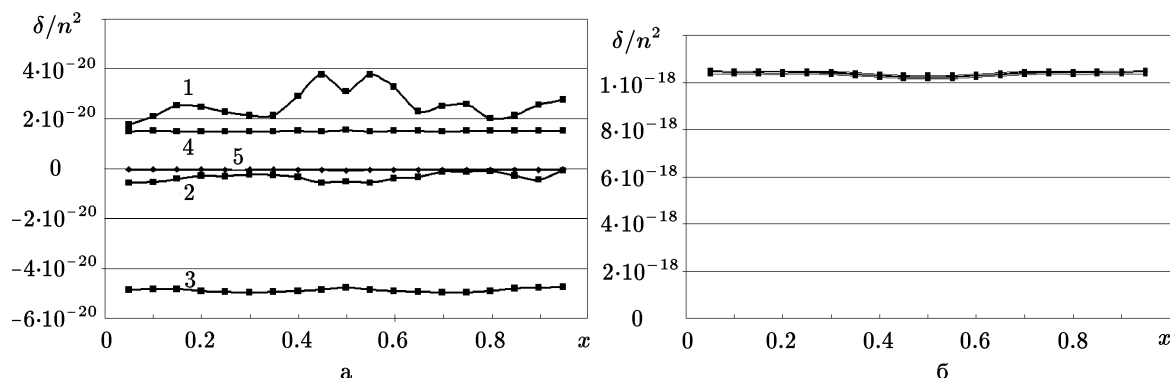


Рис. 10. Зависимости относительной погрешности округления от x при разных n для явной схемы с условиями вида I для $\lambda = 1/2$: а — для $\Delta_\lambda = 0$; б — для $\Delta_\lambda = -1 \cdot 10^{-16}$

Все это справедливо для всех рассмотренных видов краевых условий. При этом меняется только c_0 .

Проводились также исследования зависимости $c_0(\lambda)$ в ограниченных диапазонах с фиксированным шагом, при этом эффект то появлялся, то исчезал, значение $c_0(\lambda)$ иногда совпадало, иногда менялось, т. е. характер зависимости нельзя назвать ни чисто случайным, ни закономерным.

Следует отметить, что представленная выборка иллюстраций является лишь малой долей из множества проведенных экспериментов с помощью разных трансляторов и с использованием различных размеров машинных слов. На разных трансляторах даже одного языка программирования могут получаться количественно разные результаты с сохранением качественного совпадения. При этом для коротких (4-х байтных) слов при больших n возникает более грубая погрешность отсечения слагаемых, которые становятся по абсолютной величине меньше чем единица последнего разряда второго слагаемого (14). Причина возникновения или невозникновения явления формирования закономерности для явной схемы не объясняется экспериментом, и выяснение ее является задачей будущих исследований.

5. Волновое уравнение

Рассмотрим смешанную задачу для волнового уравнения

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = 0, \quad 0 < x < 1, \quad 0 < t \leq T, \tag{15}$$

с граничными условиями: $u(0, t) = u(1, t) = 0, 0 < t \leq T$, и начальными условиями: $u(x, 0) = f_1(x) = \sin \pi x, \frac{\partial u}{\partial t}(x, 0) = f_2(x) = 0, 0 < x < 1$. Точное решение этой задачи

$$u(x, t) = \sin \pi x \cos \pi t. \tag{16}$$

Рассмотрим явную трехслойную схему:

$$u_{i,j+1} = 2(1 - \lambda)u_{ij} + \lambda(u_{i+1,j} + u_{i-1,j}) - u_{i,j-1}, \quad \lambda = \frac{\tau^2}{h^2}, \tag{17}$$

$$i = 1, \dots, n - 1; \quad j = 0, \dots, m - 1.$$

Начальные условия: $u_{i,0} = f_1(x_i)$, $u_{i,1} = u_{i,0} + \tau f_2(x_i)$, $i = 1, \dots, n$. Краевые условия: $u_{0,j+1} = 0$, $u_{n,j+1} = 0$, $j = 0, \dots, m-1$.

Явная схема является устойчивой при $\lambda = \frac{\tau^2}{h^2} \leq 1$ [6]. Для данной задачи и схемы модель источников (12) примет вид

$$\frac{\partial^2 v}{\partial t^2} - \frac{\partial^2 v}{\partial x^2} = c_1 u(x, t), \quad v(0, t) = v(1, t) = 0. \quad (18)$$

Тогда для точного решения $u(x, t)$ (16) имеет место частное решение уравнения (14):

$$v(x, t) = \frac{t}{2\pi} c_1 \sin(\pi x) \sin(\pi t).$$

Вычислительный эксперимент показывает поведение в общих чертах аналогичное с явной схемой для уравнения теплопроводности. При $\lambda = \frac{1}{k}$, $k = 1, 2, 3, \dots$, регулярной зависимости погрешности округления по n не возникает. При малом отклонении от этих значений (для волнового уравнения как для четных, так и для нечетных k) регулярные зависимости, подтверждающие модель (14), возникают (рис. 11, 12).

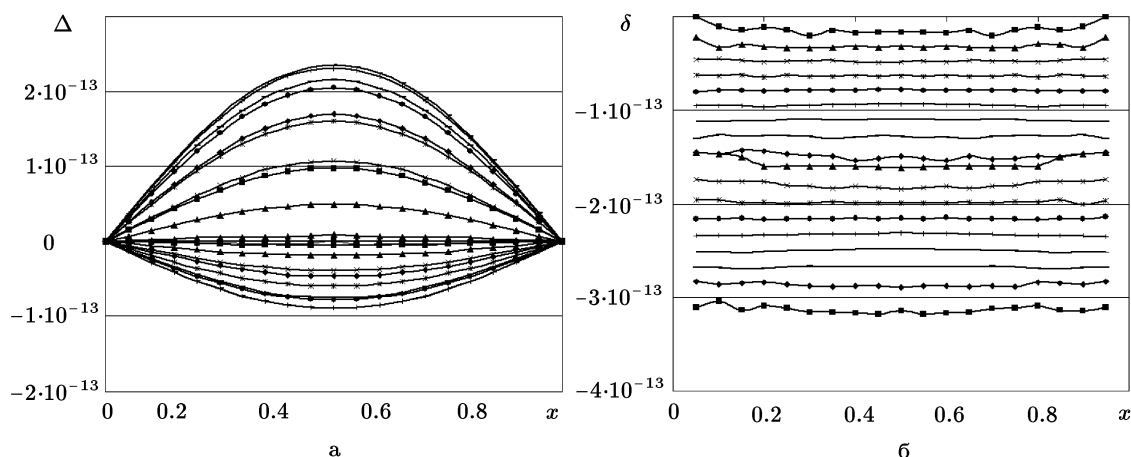


Рис. 11. Зависимости погрешности округления от координаты и времени при $\lambda = 1 \cdot 10^{-13}$ с шагом $\tau = 0.05$: а — абсолютной; б — относительной

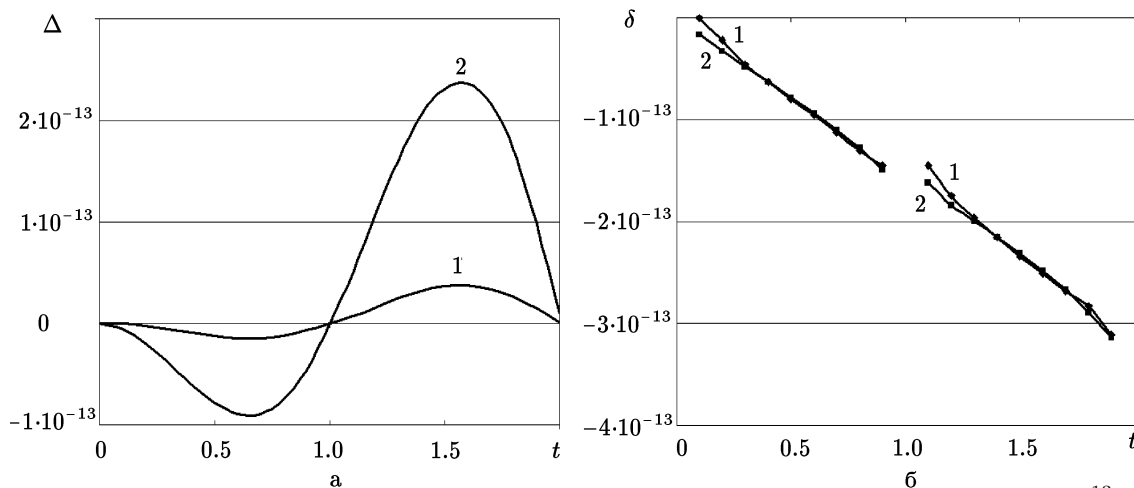


Рис. 12. Зависимости погрешности округления от времени при $T = 2$, $\lambda = 1 \cdot 10^{-13}$: а — абсолютной; б — относительной. Линии 1 — для $x = 0.05$; линии 2 — для $x = 0.5$

Отличие от уравнения теплопроводности заключается в том, что в разложении решения волнового уравнения компоненты имеют не экспоненциальные, а периодические зависимости от t . Возможно, в связи с этим зависимости погрешности округления при возрастании t могут разрушаться.

Были проведены исследования некоторых других задач математической физики: задачи Дирихле для уравнения Лапласа [1, 2] и аналогичной задачи для уравнения с частными производными 4-го порядка.

Эти задачи описывают стационарные процессы, т. е. не содержат времени. Ничего другого, кроме случайного характера роста погрешности округления от числа разбиений n (никакой детерминированной закономерности) как n^2 для уравнения Лапласа, не наблюдается. Для уравнений более высокого порядка возможно увеличение показателя степени (например для уравнения 4-го порядка до n^4). Тем не менее, это зависит не только от порядка частных производных, но и от самой задачи. К примеру, в смешанной задаче для волнового уравнения возрастания погрешности от n , более чем n в первой степени (причем случайного) не обнаруживалось.

Таким образом, погрешность округления при решении задач математической физики обладает сложными свойствами и требует индивидуального исследования для каждой задачи. Надо также учесть, что нерегулярная составляющая может зависеть от языка программирования и транслятора.

Выводы

С помощью вычислительного эксперимента выяснено, что зависимость погрешности численного метода от числа узлов n (1) состоит из тех же компонент, что и погрешность разностных производных. При этом абсолютное значение этой погрешности для вариантов условий I и II (краевые условия 1-го рода) при увеличении времени сначала растет, затем убывает и приближается к нулю. Отнесенная к функции (5) эта погрешность линейно возрастает. Для варианта III (краевые условия 2-го рода) справедливо то же самое, если отнести к функции, аналогичной (5) при замене синуса на косинус. Тем самым, модель источников (12) применима для исследования влияния на результат погрешности разностных производных.

Поведение погрешности округления в рассмотренной задаче намного сложнее, чем в задачах численного дифференцирования, интегрирования функций и решения задачи Коши для обыкновенных дифференциальных уравнений, в которых эта погрешность представлялась случайной величиной. В рассмотренной выше задаче на фоне случайной природы этой погрешности возникали закономерные зависимости от пространственной координаты, времени и числа узлов сетки n . На основе анализа результатов вычислительного эксперимента построена математическая модель возникновения и изменения во времени погрешности округления. В этой модели исходная погрешность, возникающая при выполнении арифметических действий, выступает в роли источника, распределенного в пространстве пропорционально искомой величине. С течением времени изменение погрешности аналогично изменению энергии или массы. Зависимость коэффициента, определяющего общую интенсивность источника от числа Куранта λ , в неявной схеме представляется квазислучайной величиной, ограниченной единицей последнего разряда машинного слова. В явной схеме существует ряд значений λ , при которых эффекта возникновения регулярной зависимости от числа узлов не возникает, но при малых отклонениях от этих чисел она появляется.

Литература

1. **Житников В.П., Шерыхалина Н.М.** Моделирование течений несжимаемой жидкости с применением методов многокомпонентного анализа. — Уфа: Гилем, 2009.
2. **Житников В.П., Шерыхалина Н.М.** Многокомпонентный анализ численных результатов. — Saarbrücken, Germany: LAP LAMBERT Academic Publishing, 2012.
3. **Житников В.П., Шерыхалина Н.М.** Применение многократной фильтрации при численном решении задач методами теории функций комплексного переменного // Вычислительные технологии. — 2013. — Т. 18, № 1. — С. 15–24.
4. **Житников В.П., Муксимова Р.Р.** Решение задачи нестационарной электрохимической обработки плоским электрод-инструментом с ограниченной неровностью // Вестник УГАТУ. — 2011. — Т. 15, № 1. — С. 113–118.
5. **Борщук О.С., Житников В.П.** Нелинейная фильтрация в низкопроницаемых коллекторах. Численная схема, анализ устойчивости и сходимости // Научно-техн. вестник ОАО “НК “Роснефть”. — 2013. — № 2. — С. 13–16.
6. **Волков Е.А.** Численные методы, 2-е изд. испр. и доп. — М.: Наука, 1988.

*Поступила в редакцию 1 июля 2015 г.,
в окончательном варианте 18 августа 2015 г.*

Литература в транслитерации

1. **Zhitnikov V.P., Sherykhalina N.M.** Modelirovanie techeniy vesomoy zhidkosti s primeneniem metodov mnogokomponentnogo analiza. — Ufa: Gilem, 2009.
2. **Zhitnikov V.P., Sherykhalina N.M.** Mnogokomponentnyy analiz chislennykh rezul'tatov. — Saarbrücken, Germany: LAP LAMBERT Academic Publishing, 2012.
3. **Zhitnikov V.P., Sherykhalina N.M.** Primenenie mnogokratnoy fil'tratsii pri chislennom reshenii zadach metodami teorii funktsiy kompleksnogo peremennogo // Vychislitel'nye tekhnologii. — 2013. — Т. 18, № 1. — С. 15–24.
4. **Zhitnikov V.P., Muksimova R.R.** Reshenie zadachi nestatsionarnoy elektrohimicheskoy obrabotki ploskim elektrod-instrumentom s ogranichennoy nerovnost'yu // Vestnik UGATU. — 2011. — Т. 15, № 1. — С. 113–118.
5. **Borshchuk O.S., Zhitnikov V.P.** Nelineynaya fil'tratsiya v nizkopronitsaemykh kollektorah. Chislennaya skhema, analiz ustoychivosti i skhodimosti // Nauchno-tekhn. vestnik ОАО “НК “Rosneft”. — 2013. — № 2. — С. 13–16.
6. **Volkov E.A.** Chislennyye metody, 2-e izd. ispr. i dop. — М.: Nauka, 1988.