ГЛАВА 6

ЧИСЛЕННЫЕ МЕТОДЫ РЕШЕНИЯ ЗАДАЧИ КОШИ ДЛЯ ОБЫКНОВЕННЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

§ 1. Исходная задача и примеры численных методов ее решения

1. Постановка исходной задачи. Будем рассматривать задачу Коши для системы обыкновенных дифференциальных уравнений

$$\frac{du(t)}{dt} = f(t, u), \quad t > 0, \quad u(0) = u^{(0)} \tag{1}$$

или, подробнее,

$$\frac{du_i(t)}{dt} = f_i(t, u_1, u_2, \ldots, u_m), \quad t > 0, \; i = 1, 2, \ldots, m, \tag{2}$$

$$u_i(0) = u_i^{(0)}, \quad i = 1, 2, \ldots, m. \tag{3}$$

Хорошо известны условия, гарантирующие существование и единственность решения задачи Коши (см. [39, с. 49]). Предположим, что функции f_i, $i = 1, 2, \ldots, m$, непрерывны по всем аргументам в замкнутой области

$$D = \{|t| \leqslant a, \quad |u_i - u_i^{(0)}| \leqslant b, \quad i = 1, 2, \ldots, m\}.$$

Из непрерывности функций f_i следует их ограниченность, т. е. существование константы $M > 0$ такой, что всюду в D выполняются неравенства $|f_i| \leqslant M$, $i = 1, 2, \ldots, m$.

Предположим, кроме того, что в D функции f_i удовлетворяют условию Липшица по аргументам $u_1, u_2, \ldots, u_m$, т. е.

$$|f_i(t, u_1', u_2', \ldots, u_m') - f_i(t, u_1'', u_2'', \ldots, u_m'')| \leqslant$$
$$\leqslant L\{|u_1' - u_1''| + |u_2' - u_2''| + \ldots + |u_m' - u_m''|\}$$

для любых точек $(t, u_1', \ldots, u_m')$ и $(t, u_1'', u_2'', \ldots, u_m'')$ области D.

Если выполнены сформулированные выше предположения, то существует единственное решение

$$u_1 = u_1(t), \; u_2 = u_2(t), \ldots, u_m = u_m(t)$$

системы (2), определенное при $|t| \leqslant t_0 = \min(a, b/M)$ и принимающее при $t = 0$ заданные начальные значения (3).

При исследовании численных методов для задачи Коши будем заранее предполагать, что ее решение существует, единственно и обладает необходимыми свойствами гладкости.

2. Примеры численных методов. Существуют две группы численных методов решения задачи Коши: многошаговые разностные методы и методы Рунге — Кутта. Приведем примеры и поясним основные понятия, возникающие при использовании численных методов. Для простоты изложения будем рассматривать сейчас одно уравнение

$$\frac{du}{dt}=f(t, u), \quad t>0, \quad u(0)=u_0. \tag{4}$$

Введем по переменному t *равномерную сетку с шагом* $\tau>0$, т. е. рассмотрим множество точек

$$\omega_\tau=\{t_n=n\tau,\ n=0,\ 1,\ 2, \ldots\}.$$

Будем обозначать через $u(t)$ точное решение задачи (4), а через $y_n=y(t_n)$ — приближенное решение. Заметим, что приближенное решение является *сеточной функцией*, т. е. определено только в точках сетки ω_τ.

Пример 1. *Метод Эйлера.* Уравнение (4) заменяется разностным уравнением

$$\frac{y_{n+1}-y_n}{\tau}-f(t_n, y_n)=0, \quad n=0,\ 1,\ 2, \ \ldots, \quad y_0=u_0. \tag{5}$$

Решение этого уравнения находится явным образом по рекуррентной формуле

$$y_{n+1}=y_n+\tau f(t_n, y_n), \quad n=0,\ 1, \ldots, y_0=u_0.$$

При использовании приближенных методов основным является вопрос о сходимости. Понятие сходимости приближенного метода можно сформулировать по-разному. Применительно к разностным методам, к которым относится и метод Эйлера (5), наибольшее распространение получило понятие *сходимости при* $\tau\to 0$. Оно означает следующее. Фиксируем точку t и построим последовательность сеток ω_τ таких, что $\tau\to 0$ и $t_n=n\tau=t$ (тогда необходимо $n\to\infty$). Говорят, что *метод* (5) *сходится в точке* t, если $|y_n-u(t_n)|\to 0$ при $\tau\to 0$, $t_n=t$.

Метод *сходится на* отрезке $(0, T]$, если он сходится в каждой точке $t\in(0, T]$.

Говорят, что *метод имеет p-й порядок точности*, если существует число $p>0$ такое, что $|y_n-u(t_n)|=O(\tau^p)$ при $\tau\to 0$.

Получим уравнение, которому удовлетворяет *погрешность метода* $z_n=y_n-u(t_n)$. Подставляя $y_n=z_n+u_n$ в (5), получим

$$\frac{z_{n+1}-z_n}{\tau}=f(t_n, u_n+z_n)-\frac{u_{n+1}-u_n}{\tau}. \tag{6}$$

Правую часть уравнения (6) можно представить в виде суммы

$$\psi_n^{(1)}+\psi_n^{(2)},$$

где

$$\psi_n^{(1)}=-\frac{u_{n+1}-u_n}{\tau}+f(t_n, u_n),$$

$$\psi_n^{(2)}=f(t_n, u_n+z_n)-f(t_n, u_n).$$

Функция $\psi_n^{(1)}$ называется *невязкой* или *погрешностью аппроксимации разностного уравнения* (5) *на решении исходного уравнения* (4). Видно, что невязка представляет собой результат подстановки точного решения $u=u(t)$ в левую часть разностного уравнения (5). Если бы приближенное решение y_n совпадало с точным $u(t_n)$, то невязка равнялась бы нулю. Говорят, что *разностный метод аппроксимирует исходное дифференциальное уравнение*, если $\psi_n^{(1)}\to 0$ при $\tau\to 0$. Разностный метод имеет *p-й порядок аппроксимации*, если $\psi_n^{(1)}=O(\tau^p)$. В дальнейшем будет показано, что при очень общих предположениях порядок точности разностного метода совпадает с порядком аппроксимации.

Функция

$$\psi_n^{(2)}=f(t_n, u_n+z_n)-f(t_n, u_n)$$

обращается в нуль, если правая часть f не зависит от решения u. В общем случае $\psi_n^{(2)}$ пропорциональна погрешности z_n, так как по формуле конечных приращений имеем

$$\psi_n^{(2)}=\frac{\partial f}{\partial u}(t_n, u_n+\theta z_n)\, z_n, \qquad |\theta|\leqslant 1.$$

Порядок аппроксимации метода Эйлера (5) нетрудно найти, используя разложение по формуле Тейлора. Поскольку

$$\frac{u_{n+1}-u_n}{\tau}=u'(t_n)+O(\tau),$$

то в силу уравнения (4)

$$\psi_n^{(1)}=-u'(t_n)+f(t_n, u_n)+O(\tau)=O(\tau),$$

т. е. метод Эйлера имеет первый порядок аппроксимации. При выводе предполагалась ограниченность $u''(t)$.

Пример 2. *Симметричная схема.* Уравнение (4) заменяется разностным уравнением

$$\frac{y_{n+1}-y_n}{\tau}-\frac{1}{2}\left(f(t_n, y_n)+f(t_{n+1}, y_{n+1})\right)=0, \quad n=0, 1, \ldots, \; y_0=u_0. \tag{7}$$

Данный метод более сложен в реализации, чем метод Эйлера (5), так как новое значение y_{n+1} определяется по найденному ранее y_n путем решения уравнения

$$y_{n+1}-0{,}5\tau f(t_{n+1}, y_{n+1})=F_n,$$

где $F_n=y_n+0{,}5\tau f(t_n, y_n)$. По этой причине метод называется *неявным*. Преимуществом метода (7) по сравнению с (5) является более высокий порядок точности.

Для невязки

$$\psi_n^{(1)}=-\frac{u_{n+1}-u_n}{\tau}+\frac{1}{2}\left(f(t_n, y_n)+f(t_{n+1}, y_{n+1})\right)$$

справедливо разложение

$$\psi_n^{(1)}=-u_n'-\frac{\tau}{2}u_n''+O(\tau^2)+\frac{1}{2}(u_n'+u_{n+1}')=$$

$$=-u_n'-\frac{\tau}{2}u_n''+\frac{1}{2}(u_n'+u_n'+\tau u_n''+O(\tau^2)),$$

т. е. $\psi_n^{(1)} = O(\tau^2)$. Таким образом, метод (7) имеет второй порядок аппроксимации. Из результатов § 3 будет следовать, что он имеет и второй порядок точности.

Приведенные примеры представляют собой простейшие случаи *разностных методов*, или, как их еще называют, *разностных схем. Методы Рунге — Кутта* отличаются от разностных методов тем, что в них допускается вычисление правых частей $f(t, u)$ не только в точках сетки, но и в некоторых промежуточных точках.

Пример 3. *Метод Рунге — Кутта второго порядка точности.* Предположим, что приближенное значение y_n решения исходной задачи в момент $t=t_n$ уже известно. Для нахождения $y_{n+1}=y(t_{n+1})$ поступим следующим образом. Сначала, используя схему Эйлера

$$\frac{y_{n+1/2}-y_n}{0,5\tau}=f(t_n,\ y_n), \tag{8}$$

вычислим промежуточное значение $y_{n+1/2}$, а затем воспользуемся разностным уравнением

$$\frac{y_{n+1}-y_n}{\tau}=f(t_n+0,5\tau,\ y_{n+1/2}), \tag{9}$$

из которого явным образом найдем искомое значение y_{n+1}.

Для исследования невязки подставим промежуточное значение $y_{n+1/2}=y_n+0,5\tau f_n$, где $f_n=f(t_n,\ y_n)$, в уравнение (9). Тогда получим разностное уравнение

$$\frac{y_{n+1}-y_n}{\tau}-f(t_n+0,5\tau,\ y_n+0,5\tau f_n)=0, \tag{10}$$

невязка которого равна

$$\psi_n^{(1)}=-\frac{u_{n+1}-u_n}{\tau}+f(t_n+0,5\tau,\ u_n+0,5\tau f(t_n,\ u_n)). \tag{11}$$

Имеем

$$\frac{u_{n+1}-u_n}{\tau}=u'_n+0,5\tau u''_n+O(\tau^2),$$

$$f(t_n+0,5\tau,\ u_n+0,5\tau f(t_n,\ u_n))=f(t_n,\ u_n)+$$
$$+0,5\tau\left(\frac{\partial f(t_n,\ u_n)}{\partial t}+0,5\tau f(t_n,\ u_n)\frac{\partial f(t_n,\ u_n)}{\partial u}\right)=f(t_n,\ u_n)+0,5\tau u''_n,$$

так как в силу (4) справедливо равенство

$$\frac{d^2u}{dt^2}=\frac{\partial f}{\partial t}+f\frac{\partial f}{\partial u}.$$

Таким образом, метод (10) имеет второй порядок погрешности аппроксимации, $\psi_n^{(1)} = O(\tau^2)$, и в отличие от (7) является явным.

Реализация метода (10) в виде двух этапов (8), (9) называется *методом предиктор — корректор* (предсказывающе-исправляющим), поскольку на первом этапе (8) приближенное значение предсказывается с невысокой точностью $O(\tau)$, а на втором этапе (9) это предсказанное значение исправляется, так что результирующая погрешность имеет второй порядок по τ.

Тот же самый метод (10) можно реализовать несколько иначе. А именно, сначала вычислим последовательно функции

$$k_1=f(t_n, y_n), \quad k_2=f(t_n+0{,}5\tau, y_n+0{,}5\tau k_1),$$

а затем найдем y_{n+1} из уравнения $(y_{n+1}-y_n)/\tau=k_2$.

Такая форма реализации метода (10) называется *методом Рунге — Кутта*. Поскольку требуется вычислить две промежуточные функции, k_1 и k_2, данный метод относится к *двухэтапным методам*. В следующем параграфе будут рассмотрены более общие m-этапные методы Рунге — Кутта, позволяющие получить бо́льшую точность.

§ 2. Методы Рунге — Кутта

1. Общая формулировка методов. Семейство методов второго порядка. По-прежнему рассматриваем задачу Коши для одного уравнения

$$\frac{du}{dt}=f(t, u), \quad t>0, \qquad u(0)=u_0. \tag{1}$$

Явный m-этапный метод Рунге — Кутта состоит в следующем. Пусть решение $y_n=y(t_n)$ уже известно. Задаются числовые коэффициенты a_i, b_{ij}, $i=2, 3, \ldots, m$, $j=1, 2, \ldots, m-1$, σ_i, $i=1, 2, \ldots$ $\ldots, m$, и последовательно вычисляются функции

$$\begin{aligned}
&k_1=f(t_n, y_n),\ k_2=f(t_n+a_2\tau,\ y_n+b_{21}\tau k_1),\\
&k_3=f(t_n+a_3\tau,\ y_n+b_{31}\tau k_1+b_{32}\tau k_2),\ \ldots,\\
&k_m=f(t_n+a_m\tau,\ y_n+b_{m1}\tau k_1+b_{m2}\tau k_2+\ldots+b_{m,m-1}\tau k_{m-1}).
\end{aligned}$$

Затем из формулы

$$\frac{y_{n+1}-y_n}{\tau}=\sum_{i=1}^{m}\sigma_i k_i \tag{2}$$

находится новое значение $y_{n+1}=y(t_{n+1})$.

Коэффициенты a_i, b_{ij}, σ_i выбираются из соображений точности. Например, для того чтобы уравнение (2) аппроксимировало исходное уравнение (1), необходимо потребовать $\sum\limits_{i=1}^{m}\sigma_i=1$. Отметим, что методы Рунге — Кутта при $m>5$ не используются.

Остановимся более подробно на отдельных методах. При $m=$ $=1$ получаем схему Эйлера, рассмотренную в примере 1 из предыдущего параграфа. При $m=2$ получаем семейство методов

$$\begin{gathered}
k_1=f(t_n, y_n), \quad k_2=f(t_n+a_2\tau,\ y_n+b_{21}\tau k_1),\\
y_{n+1}=y_n+\tau(\sigma_1 k_1+\sigma_2 k_2).
\end{gathered} \tag{3}$$

Исследуем погрешность аппроксимации методов (3) в зависимости от выбора параметров. Исключая из последнего уравнения функции k_1 и k_2, получаем

$$\frac{y_{n+1}-y_n}{\tau}=\sigma_1 f(t_n, y_n)+\sigma_2 f(t_n+a_2\tau, y_n+b_{21}\tau f(t_n, y_n)). \qquad (4)$$

По определению погрешностью аппроксимации или невязкой метода (3) называется выражение

$$\psi_n^{(1)}=-\frac{u_{n+1}-u_n}{\tau}+\sigma_1 f(t_n, u_n)+\sigma_2 f(t_n+a_2\tau, u_n+b_{21} f(t_n, u_n)), \qquad (5)$$

полученное заменой в (4) приближенного решения y_n точным решением $u_n=u(t_n)$.

Найдем порядок погрешности аппроксимации в предположении достаточной гладкости решения $u(t)$ и функции $f(t, u)$. Для этого разложим все величины, входящие в выражение (5), по формуле Тейлора в точке t_n. Имеем

$$\frac{u_{n+1}-u_n}{\tau}=u'(t_n)+\frac{\tau}{2}u''(t_n)+O(\tau^2),$$

$$f(t_n+a_2\tau, u_n+b_{21}\tau f_n)=f_n+a_2\tau\frac{\partial f_n}{\partial t}+b_{21}\tau f_n\frac{\partial f_n}{\partial u}+O(\tau^2),$$

где $f_n=f(t_n, u_n)$, $\frac{\partial f_n}{\partial u}=\frac{\partial f}{\partial u}(t_n, u_n)$. Далее, согласно уравнению (1), получим

$$u''=\frac{\partial f}{\partial t}+\frac{\partial f}{\partial u}u'=\frac{\partial f}{\partial t}+\frac{\partial f}{\partial u}f.$$

Поэтому

$$\psi_n^{(1)}=-u_n'+(\sigma_1+\sigma_2)f_n+$$
$$+\tau\left[(\sigma_2 b_{21}-0{,}5)f_n\frac{\partial f_n}{\partial u}+(\sigma_2 a_2-0{,}5)\frac{\partial f_n}{\partial t}\right]+O(\tau^2). \qquad (6)$$

Отсюда видно, что методы (3) имеют первый порядок аппроксимации, если $\sigma_1+\sigma_2=1$.

Если же дополнительно потребовать $\sigma_2 a_2=\sigma_2 b_{21}=0{,}5$, то получим методы второго порядка аппроксимации. Таким образом, имеется однопараметрическое семейство двухэтапных методов Рунге — Кутта второго порядка аппроксимации. Это семейство методов можно записать в виде

$$\frac{y_{n+1}-y_n}{\tau}=(1-\sigma)f(t_n, y_n)+\sigma f(t_n+a\tau, y_n+a\tau f(t_n, y_n)), \qquad (7)$$

где $\sigma a=0{,}5$.

В частности, при $\sigma=1$, $a=0{,}5$ получаем метод, рассмотренный в примере 3 предыдущего параграфа. При $\sigma=0{,}5$, $a=1$ получаем другой метод второго порядка:

$$k_1=f(t_n, y_n), \quad k_2=f(t_n+\tau, y_n+\tau k_1),$$
$$y_{n+1}=y_n+0{,}5\tau(k_1+k_2).$$

Двухэтапных методов третьего порядка аппроксимации не существует. Чтобы убедиться в этом, достаточно рассмотреть уравнение $u'=u$. Для него двухэтапный метод Рунге — Кутта (7) принимает вид

$$\frac{y_{n+1}-y_n}{\tau}=(1+\tau\sigma a)\,y_n$$

и погрешность аппроксимации равна

$$\psi_n^{(1)}=-\frac{u_{n+1}-u_n}{\tau}+(1+\tau\sigma a)\,u_n.$$

Разлагая $\psi_n^{(1)}$ по формуле Тейлора и учитывая, что $u'''=u''=u'=$ $=u$, получим

$$\psi_n^{(1)}=\tau(\sigma a-0{,}5)\,u-\frac{\tau^2}{6}u(t_n+\theta\tau),\qquad 0\leqslant\theta\leqslant 1.$$

Отсюда видно, что наивысший достижимый порядок аппроксимации равен двум.

Приведем примеры методов Рунге — Кутта более высокого порядка аппроксимации.

Метод третьего порядка:

$$k_1=f(t_n,y_n),\quad k_2=f\left(t_n+\frac{\tau}{2},\ y_n+\frac{\tau}{2}k_1\right),$$

$$k_3=f(t_n+\tau,\ y_n-\tau k_1+2\tau k_2),\ \frac{y_{n+1}-y_n}{\tau}=\frac{1}{6}(k_1+4k_2+k_3).$$

Метод третьего порядка:

$$k_1=f(t_n,y_n),\quad k_2=f\left(t_n+\frac{\tau}{3},\ y_n+\frac{\tau k_1}{3}\right),$$

$$k_3=f\left(t_n+\frac{2\tau}{3},\ y_n+\frac{2\tau k_2}{3}\right),\ \frac{y_{n+1}-y_n}{\tau}=\frac{1}{4}(k_1+3k_3).$$

Метод четвертого порядка:

$$k_1=f(t_n,y_n),\quad k_2=f\left(t_n+\frac{\tau}{2},\ y_n+\frac{\tau k_1}{2}\right),$$

$$k_3=f\left(t_n+\frac{\tau}{2},\ y_n+\frac{\tau k_2}{2}\right),\ k_4=f(t_n+\tau,\ y_n+\tau k_3),$$

$$\frac{y_{n+1}-y_n}{\tau}=\frac{1}{6}(k_1+2k_2+2k_3+k_4).$$

Метод четвертого порядка:

$$k_1=f(t_n,y_n),\quad k_2=f\left(t_n+\frac{\tau}{4},\ y_n+\frac{\tau k_1}{4}\right),$$

$$k_3=f\left(t_n+\frac{\tau}{2},\ y_n+\frac{\tau k_2}{2}\right),$$

$$k_4=f(t_n+\tau,\ y_n+\tau k_1-2\tau k_2+2\tau k_3),$$

$$\frac{y_{n+1}-y_n}{\tau}=\frac{1}{6}(k_1+4k_3+k_4)$$

Приведенные здесь методы являются частными случаями методов Рунге — Кутта третьего и четвертого порядков, рассмотренных подробнее в пп. 3 и 4.

2. Доказательство сходимости. Докажем, что методы Рунге — Кутта сходятся и порядок их точности совпадает с порядком аппроксимации.

Выпишем уравнение, которому удовлетворяет погрешность $z_n = y_n - u(t_n)$. Основное уравнение метода Рунге — Кутта имеет вид

$$\frac{y_{n+1}-y_n}{\tau} = \sum_{i=1}^{m} \sigma_i k_i(y), \tag{8}$$

где

$$k_i(y) = f\left(t_n + a_i\tau,\ y_n + \sum_{j=1}^{i-1} \tau b_{ij} k_j(y)\right), \quad i = 2, 3, \ldots, m, \tag{9}$$

$$k_1(y) = f(t_n, y_n).$$

Подставим в левую часть уравнения (8) вместо y_l выражения $u_l + z_l$ при $l = n$, $n+1$, а в правой части этого уравнения добавим и вычтем сумму

$$\sum_{i=1}^{m} \sigma_i k_i(u),$$

где

$$k_i(u) = f\left(t_n + a_i\tau,\ u_n + \sum_{j=1}^{i-1} \tau b_{ij} k_j(u)\right), \quad i = 2, 3, \ldots, m,$$

$$k_1(u) = f(t_n, u_n). \tag{10}$$

Тогда уравнение (8) примет вид

$$\frac{z_{n+1}-z_n}{\tau} = \psi_n^{(1)} + \psi_n^{(2)}, \tag{11}$$

где

$$\psi_n^{(1)} = -\frac{u_{n+1}-u_n}{\tau} + \sum_{i=1}^{m} \sigma_i k_i(u) \tag{12}$$

есть по определению погрешность аппроксимации метода (8), (9) на решении исходной задачи (1) (невязка) и

$$\psi_n^{(2)} = \sum_{i=1}^{m} \sigma_i (k_i(y) - k_i(u)). \tag{13}$$

Будем рассматривать (11) как уравнение для погрешности метода. Оно выполняется для $n = 0, 1, \ldots$ Поскольку начальные значения y_0 задаются точно, $y_0 = u(0)$, величина z_0 равна нулю. Будем считать, что задача (1) решается на ограниченном отрезке времени $0 < t \leqslant T$, и, следовательно, при любых n и τ выполняется неравенство $t_n = n\tau \leqslant T$.

Предположим, что в рассматриваемой области изменения переменных (t, u) функция $f(t, u)$ удовлетворяет условию Липшица по u с константой L, не зависящей от t. При этих предположениях оценим сначала функцию $\psi_n^{(2)}$, а затем и решение z_{n+1} уравнения (11).

Из выражений (9), (10), используя условие Липшица, получим

$$|k_i(y) - k_i(u)| \leqslant L\left(|y_n - u_n| + \sum_{j=1}^{i-1} \tau |b_{ij}|\, |k_j(y) - k_j(u)|\right),$$

$$i=2, 3, \ldots, m, \quad |k_1(y) - k_1(u)| \leqslant L|y_n - u_n|.$$

Обозначим

$$r_i = |k_i(y) - k_i(u)|, \quad i = 1, 2, \ldots, m,$$
$$b = \max_{\substack{2 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant i-1}} |b_{ij}|, \quad g = L|y_n - u_n|. \tag{14}$$

Тогда согласно предыдущему неравенству будем иметь

$$r_i \leqslant Lb \sum_{j=1}^{i-1} \tau r_j + g, \quad i = 2, 3, \ldots, m, \quad r_1 \leqslant g,$$

или, что то же самое,

$$r_{i+1} \leqslant Lb \sum_{j=0}^{i} \tau r_j + g, \quad i = 0, 1, \ldots, m-1, \quad r_0 = 0. \tag{15}$$

Лемма 1. *Из неравенств* (15) *при* $Lb\tau > 0$ *следуют оценки*

$$r_i \leqslant \rho^{i-1} g, \quad i = 1, 2, \ldots, m, \tag{16}$$

где $\rho = 1 + Lb\tau$.

Доказательство. Оценка (16) при $i=1$ совпадает с оценкой (15) для $i=0$. Пусть неравенство (16) выполнено для $i = 1, 2, \ldots, k$. Покажем, что оно выполнено и для $i=k+1$. Из (15) при $i=k$ получим

$$r_{k+1} \leqslant Lb \sum_{j=1}^{k} \tau r_j + g.$$

Согласно предположению индукции имеем

$$r_j \leqslant \rho^{j-1} g, \quad j = 1, 2, \ldots, k,$$

следовательно,

$$r_{k+1} \leqslant \left(Lb\tau \sum_{j=1}^{k} \rho^{j-1} + 1\right) g = \left(Lb\tau \frac{\rho^k - 1}{\rho - 1} + 1\right) g = \rho^k g,$$

что и требовалось.

Оценим теперь функцию $\psi_n^{(2)}$, определенную согласно (13). Из (14), (16) следует неравенство

$$|\psi_n^{(2)}| < \sum_{i=1}^{m} |\sigma_i| \, |r_i| \leqslant \sigma g \sum_{i=1}^{m} \rho^{i-1} \leqslant \sigma g m \rho^{m-1},$$

где $\sigma = \max\limits_{1\leqslant i\leqslant m} |\sigma_i|$, $\rho = 1 + Lb\tau$, $g = L|z_n|$.

Итак, окончательно имеем следующую оценку для $\psi_n^{(2)}$:

$$|\psi_n^{(2)}| \leqslant \sigma L m (1 + Lb\tau)^{m-1} |z_n|. \tag{17}$$

Таким образом, при возрастании погрешности $|z_n|$ величина $|\psi_n^{(2)}|$ растет не быстрее первой степени погрешности.

Теперь уже несложно оценить погрешность $z_n = y_n - u(t_n)$. Из уравнения (11) имеем

$$z_{n+1} = z_n + \tau\psi_n^{(2)} + \tau\psi_n^{(1)},$$

откуда, учитывая (17), получаем неравенство

$$|z_{n+1}| \leqslant (1 + \alpha\tau)|z_n| + \tau|\psi_n^{(1)}|, \quad n = 0, 1, \ldots, \tag{18}$$

где

$$\alpha = \alpha(\tau) = \sigma L m (1 + Lb\tau)^{m-1}. \tag{19}$$

Заметим, что $\alpha(\tau) \to \sigma L m$ при $\tau \to 0$. Если $\tau \leqslant \tau_0$, то $\alpha(\tau) \leqslant \sigma L m e^{Lb(m-1)\tau_0}$, т. е. $\alpha(\tau)$ ограничена равномерно по τ. В качестве τ_0 с большим загрублением можно взять T.

Из неравенства (18) следует оценка

$$|z_{n+1}| \leqslant (1 + \alpha\tau)^{n+1} |z_0| + \sum_{j=0}^{n} \tau (1 + \alpha\tau)^{n-j} |\psi_j^{(1)}|, \tag{20}$$

которую легко доказать по индукции.

Загрубляя оценку (20) и учитывая, что $z_0 = 0$, получим

$$|z_{n+1}| \leqslant (n+1)\tau(1 + \alpha\tau)^n \max_{0\leqslant j\leqslant n} |\psi_j^{(1)}| \leqslant t_{n+1} e^{\alpha t_n} \max_{0\leqslant j\leqslant n} |\psi_j^{(1)}|,$$

где $t_n = n\tau \leqslant T$.

Таким образом, доказана

Теорема 1. *Пусть правая часть уравнения* (1) $f(t, u)$ *удовлетворяет условию Липшица по второму аргументу с константой* L. *Пусть* $\psi_j^{(1)}$ — *невязка метода Рунге — Кутта* (2), *определенная согласно* (12). *Тогда для погрешности метода при* $n\tau \leqslant T$ *справедлива оценка*

$$|y_n - u(t_n)| \leqslant T e^{\alpha T} \max_{0\leqslant j\leqslant n-1} |\psi_j^{(1)}|, \tag{21}$$

где

$$\alpha = \sigma L m (1 + Lb\tau)^{m-1},$$

$$\sigma = \max_{1\leqslant i\leqslant m} |\sigma_i|, \quad b = \max_{\substack{2\leqslant i\leqslant m\\ 1\leqslant j\leqslant i-1}} |b_{ij}|.$$

Следствие. *Если метод Рунге — Кутта аппроксимирует исходное уравнение, то он сходится при $\tau \to 0$, причем порядок точности совпадает с порядком аппроксимации.*

Доказательство этого утверждения сразу следует из оценки (21) и замечания о равномерной ограниченности $\alpha(\tau)$.

3. Методы третьего порядка точности. При решении обыкновенных дифференциальных уравнений часто используются методы третьего и четвертого порядка точности. Приведем вывод таких методов. Сначала рассмотрим трехэтапный метод

$$\begin{gathered} k_1 = f(t_n, y_n), \quad k_2 = f(t_n + a_2\tau, y_n + b_{21}\tau k_1), \\ k_3 = f(t_n + a_3\tau, y_n + b_{31}\tau k_1 + b_{32}\tau k_2), \\ y_{n+1} = y_n + \tau(\sigma_1 k_1 + \sigma_2 k_2 + \sigma_3 k_3). \end{gathered} \tag{22}$$

Выясним, каким условиям должны удовлетворять параметры a_i, b_{ij}, σ_i для того, чтобы данный метод имел третий порядок аппроксимации. Погрешность аппроксимации метода (22) дается выражением

$$\psi_n^{(1)} = -\frac{u_{n+1} - u_n}{\tau} + \sigma_1 k_1 + \sigma_2 k_2 + \sigma_3 k_3, \tag{23}$$

где

$$\begin{gathered} k_1 = f(t_n, u_n), \ k_2 = f(t_n + a_2\tau, u_n + b_{21}\tau k_1), \\ k_3 = f(t_n + a_3\tau, u_n + b_{31}\tau k_1 + b_{32}\tau k_2) \end{gathered} \tag{24}$$

и $u_n = u(t_n)$ — решение исходного уравнения (1).

Применяя разложение по формуле Тейлора для функции двух переменных и учитывая, что $k_1 = f(t_n, u_n)$, получим

$$k_2 = f + a_2\tau f_t + b_{21}\tau f_t f_u + \frac{\tau^2}{2}[a_2^2 f_{tt} + 2a_2 b_{21} f f_{tu} + b_{21}^2 f^2 f_{uu}] + O(\tau^3). \tag{25}$$

Здесь значения функции $f(t, u)$ и ее частных производных берутся при $t = t_n$, $u = u_n$. Точно так же

$$k_3 = f + a_3\tau f_t + (b_{31}k_1 + b_{32}k_2)\tau f_u + \\ + \frac{\tau^2}{2}[a_3^2 f_{tt} + 2a_3(b_{31}k_1 + b_{32}k_2) f_{tu} + (b_{31}k_1 + b_{32}k_2)^2 f_{uu}] + O(\tau^3).$$

Подставляя сюда

$$k_1 = f, \quad k_2 = f + a_2\tau f_t + b_{21}\tau f f_u + O(\tau^2),$$

получим

$$k_3 = f + \tau[a_3 f_t + (b_{31} + b_{32}) f f_u] + \frac{\tau^2}{2}[a_3^2 f_{tt} + 2a_3(b_{31} + b_{32}) f f_{tu} + \\ + (b_{31} + b_{32})^2 f^2 f_{uu} + 2b_{32}a_2 f_t f_u + 2b_{32}b_{21} f (f_u)^2] + O(\tau^3). \tag{26}$$

Далее, из разложений (24)—(26) следует

$$\sigma_1 k_1 + \sigma_2 k_2 + \sigma_3 k_3 = (\sigma_1 + \sigma_2 + \sigma_3) f + \\ + \tau[(\sigma_2 a_2 + \sigma_3 a_3) f_t + (\sigma_2 b_{21} + \sigma_3(b_{31} + b_{32})) f f_u] + \\ + \frac{\tau^2}{2}[(\sigma_2 a_2^2 + \sigma_3 a_3^2) f_{tt} + 2(\sigma_2 a_2 b_{21} + \sigma_3 a_3(b_{31} + b_{32})) f f_{tu} + \\ + ((\sigma_2 b_{21}^2 + \sigma_3(b_{31} + b_{32})^2) f^2 f_{uu} + 2\sigma_3 b_{32} a_2 f_t f_u + 2\sigma_3 b_{32} b_{21} f (f_u)^2] + O(\tau^3). \tag{27}$$

Получим теперь разложение по степеням τ разностного отношения $(u_{n+1}-u_n)/\tau$, входящего в выражение для погрешности аппроксимации (23). При этом учтем, что в силу уравнения (1) справедливы следующие соотношения:

$$u'=f, \quad u''=f_t+ff_u,$$
$$u'''=f_{tt}+2ff_{tu}+f^2f_{uu}+f_uf_t+f(f_u)^2. \tag{28}$$

Тогда будем иметь

$$\frac{u_{n+1}-u_n}{\tau}=u_n'+\frac{\tau}{2}u_n''+\frac{\tau^2}{6}u_n'''+O(\tau^3)=$$
$$=f+\frac{\tau}{2}(f_t+ff_u)+\frac{\tau^2}{6}[f_{tt}+2ff_{tu}+f^2f_{uu}+f_uf_t+f(f_u)^2]+O(\tau^3).$$

Отсюда и из (27) получаем следующее разложение выражения для погрешности аппроксимации (23):

$$\psi_n^{(1)}=(\sigma_1+\sigma_2+\sigma_3-1)f+$$
$$+\frac{\tau}{2}\{[2(\sigma_2a_2+\sigma_3a_3)-1]f_t+[2(\sigma_2b_{21}+\sigma_3(b_{31}+b_{32}))-1]ff_u\}+$$
$$+\frac{\tau^2}{6}\{[3(\sigma_2a_2^2+\sigma_3a_3^2)-1]f_{tt}+[6(\sigma_2a_2b_{21}+\sigma_3a_3(b_{31}+b_{32}))-2]ff_{tu}+$$
$$+[3(\sigma_2b_{21}^2+\sigma_3(b_{31}+b_{32})^2)-1]f^2f_{uu}+$$
$$+(6\sigma_3b_{32}a_2-1)f_uf_t+(6\sigma_3b_{32}b_{21}-1)f(f_u)^2\}+O(\tau^3).$$

Приравнивая нулю коэффициенты при τ^j, $j=0, 1, 2$, получаем условия третьего порядка аппроксимации:

$$\sigma_1+\sigma_2+\sigma_3=1,$$
$$\sigma_2a_2+\sigma_3a_3=\sigma_2b_{21}+\sigma_3(b_{31}+b_{32})=0{,}5,$$
$$\sigma_2a_2^2+\sigma_3a_3^2=\sigma_2a_2b_{21}+\sigma_3a_3(b_{31}+b_{32})=\sigma_2b_{21}^2+\sigma_3(b_{31}+b_{32})^2=\frac{1}{3},$$
$$\sigma_3b_{32}a_2=\sigma_3b_{32}b_{21}=\frac{1}{6}.$$

После проведения эквивалентных преобразований эту систему уравнений можно записать в более простом виде:

$$\sigma_2a_2+\sigma_3a_3=\frac{1}{2},\quad \sigma_2a_2^2+\sigma_3a_3^2=\frac{1}{3},$$
$$a_2=b_{21},\quad a_3=b_{31}+b_{32},\quad \sigma_3b_{32}a_2=\frac{1}{6}, \tag{29}$$
$$\sigma_1=1-\sigma_2-\sigma_3,\quad \sigma_3\neq 0,\quad a_2\neq 0.$$

Исключим с помощью (29) из выражений (22) коэффициенты b_{ij}. Тогда получим метод

$$k_1=f(t_n, y_n),\quad k_2=f(t_n+a_2\tau, y_n+a_2\tau k_1),$$
$$k_3=f\left(t_n+a_3\tau, y_n+a_3\tau k_1+\frac{\tau(k_2-k_1)}{6\sigma_3a_2}\right), \tag{30}$$
$$\frac{y_{n+1}-y_n}{\tau}=\sigma_1k_1+\sigma_2k_2+\sigma_3k_3,\qquad \sigma_1=1-\sigma_2-\sigma_3,$$

который имеет третий порядок аппроксимации при условиях

$$\sigma_2 a_2 + \sigma_3 a_3 = \frac{1}{2}, \quad \sigma_2 a_2^2 + \sigma_3 a_3^2 = \frac{1}{3}, \quad \sigma_3 \neq 0, \quad a_2 \neq 0. \tag{31}$$

Таким образом, в общем случае существует двухпараметрическое семейство трехэтапных методов Рунге — Кутта, имеющих третий порядок аппроксимации. Задавая a_2 и a_3 в качестве свободных параметров, получим из (31)

$$\sigma_2 = \frac{\frac{1}{2} a_3 - \frac{1}{3}}{a_2 (a_3 - a_2)}, \quad \sigma_3 = \frac{\frac{1}{3} - \frac{1}{2} a_2}{a_3 (a_3 - a_2)}. \tag{32}$$

Кроме того, система (31) имеет два однопараметрических семейства решений, определяемых условиями

$$a_2 = a_3 = \frac{2}{3}, \quad \sigma_2 + \sigma_3 = \frac{3}{4}, \quad \sigma_3 \neq 0, \tag{33}$$

$$a_2 = \frac{2}{3}, \quad a_3 = 0, \quad \sigma_2 = \frac{3}{4}, \quad \sigma_3 \neq 0 \quad \text{любое.} \tag{34}$$

Например, полагая $a_2 = \frac{1}{2}$, $a_3 = 1$, получим из (30), (32) следующий метод третьего порядка аппроксимации:

$$\begin{gathered}
k_1 = f(t_n, y_n), \quad k_2 = f\left(t_n + \frac{\tau}{2}, \, y_n + \frac{\tau}{2} k_1\right), \\
k_3 = f(t_n + \tau, \, y_n - \tau k_1 + 2\tau k_2), \\
\frac{y_{n+1} - y_n}{\tau} = \frac{1}{6}(k_1 + 4k_2 + k_3).
\end{gathered} \tag{35}$$

4. **Методы четвертого порядка точности.** Рассмотрим теперь четырехэтапный метод

$$\begin{gathered}
k_1 = f(t_n, y_n), \quad k_2 = f(t_n + a_2\tau, \, y_n + b_{21}\tau k_1), \\
k_3 = f(t_n + a_3\tau, \, y_n + b_{31}\tau k_1 + b_{32}\tau k_2), \\
k_4 = f(t_n + a_4\tau, \, y_n + b_{41}\tau k_1 + b_{42}\tau k_2 + b_{43}\tau k_3), \\
y_{n+1} = y_n + \tau(\sigma_1 k_1 + \sigma_2 k_2 + \sigma_3 k_3 + \sigma_4 k_4).
\end{gathered} \tag{36}$$

Погрешность аппроксимации метода (36) равна по определению

$$\psi_n^{(1)} = -\frac{u_{n+1} - u_n}{\tau} + \sigma_1 k_1 + \sigma_2 k_2 + \sigma_3 k_3 + \sigma_4 k_4, \tag{37}$$

где функции k_l, $l = 1, 2, 3, 4$, получаются из (36) путем замены y_n на точное решение $u_n = u(t_n)$.

Чтобы построить схемы четвертого порядка аппроксимации, необходимо разложить функции, входящие в (37), по формуле Тейлора до величин третьего порядка по τ включительно и приравнять нулю коэффициенты при степенях τ^n, $n = 0, 1, 2, 3$. Необходимо при этом учесть соотношения (28) и аналогичное выражение для u^{IV}. Опуская выкладки, приведем систему уравнений, которой должны удовлетворять коэффициенты метода (36) для того, чтобы данный метод имел четвертый порядок аппроксимации:

$$\sigma_1+\sigma_2+\sigma_3+\sigma_4=1,$$

$$b_{21}=a_2,\quad b_{31}+b_{32}=a_3,\ b_{41}+b_{42}+b_{43}=a_4; \tag{38}$$

$$\sigma_2a_2+\sigma_3a_3+\sigma_4a_4=\frac{1}{2},$$

$$\sigma_2a_2^2+\sigma_3a_3^2+\sigma_4a_4^2=\frac{1}{3}, \tag{39}$$

$$\sigma_2a_2^3+\sigma_3a_3^3+\sigma_4a_4^3=\frac{1}{4};$$

$$(\sigma_3b_{32}+\sigma_4b_{42})\,a_2+\sigma_4b_{43}a_3=\frac{1}{6},$$

$$(\sigma_3b_{32}+\sigma_4b_{42})\,a_2^2+\sigma_4b_{43}a_3^2=\frac{1}{12}, \tag{40}$$

$$\sigma_3b_{32}a_2a_3+\sigma_4b_{42}a_2a_4+\sigma_4b_{43}a_3a_4=\frac{1}{8};$$

$$\sigma_4b_{43}b_{32}a_2=\frac{1}{24}. \tag{41}$$

Система (38)—(41) состоит из одиннадцати уравнений и содержит тринадцать неизвестных. Выберем в качестве независимых параметров неизвестные a_2 и a_3 и выразим остальные величины через эти неизвестные.

Для этого сначала разрешим группу уравнений (39) относительно переменных σ_2, σ_3, σ_4. Определитель δ этой системы

$$\delta=a_2a_3a_4(a_3-a_2)(a_4-a_2)(a_4-a_3). \tag{42}$$

Это означает, что мы не рассматриваем значения параметров a_2, a_3, a_4, удовлетворяющие хотя бы одному из условий

$$a_2=0,\quad a_3=0,\quad a_4=0,\quad a_2=a_3,\quad a_2=a_4,\quad a_3=a_4.$$

Заметим сразу же, что $a_2\neq0$, $\sigma_4\neq0$ согласно (41).

Предположим, что $\delta\neq0$. Случай $\delta=0$ будет рассмотрен позже.

При $\delta\neq0$ система (39) имеет следующее решение:

$$\sigma_2=\frac{6a_3a_4-4a_3-4a_4+3}{12a_2(a_3-a_2)(a_4-a_2)}, \tag{43}$$

$$\sigma_3=-\frac{6a_2a_4-4a_2-4a_4+3}{12a_3(a_3-a_2)(a_4-a_3)}, \tag{44}$$

$$\sigma_4=\frac{6a_2a_3-4a_2-4a_3+3}{12a_4(a_4-a_2)(a_4-a_3)}. \tag{45}$$

Точно так же, решая систему (40), получим

$$\sigma_3b_{32}=\frac{4a_4-3}{24a_2(a_4-a_3)}, \tag{46}$$

$$\sigma_4b_{42}=-\frac{2(1-2a_2)(a_4-a_3)-(3-4a_3)(a_3-a_2)}{24(a_3-a_2)(a_4-a_3)a_2}, \tag{47}$$

$$\sigma_4b_{43}=\frac{1-2a_2}{12(a_3-a_2)a_3}. \tag{48}$$

Учтем теперь соотношение (41). Прежде всего заметим, что $\sigma_3 \neq 0$. Действительно, при $\sigma_3=0$ из (44) и (46) получаем

$$a_4 = \frac{3}{4}, \quad a_2 = 0,$$

но в силу (41) имеем $a_2 \neq 0$. Таким образом, уравнение (41) эквивалентно уравнению

$$(\sigma_3 b_{32})(\sigma_4 b_{43}) a_2 = \frac{\sigma_3}{24}.$$

Подставляя сюда выражения для $\sigma_3 b_{32}$, $\sigma_4 b_{43}$, σ_3 из (44), (46), (48) и приводя подобные члены, получим уравнение $a_2(1-a_4)=0$, из которого следует, что $a_4=1$.

Таким образом, при $\delta \neq 0$ система (38)—(41) имеет следующее двухпараметрическое семейство решений:

$$a_4 = 1, \quad \sigma_2 = \frac{2a_3 - 1}{12a_2(a_3 - a_2)(1 - a_2)},$$

$$\sigma_3 = -\frac{2a_2 - 1}{12a_3(a_3 - a_2)(1 - a_3)},$$

$$\sigma_4 = \frac{6a_2a_3 - 4a_2 - 4a_3 + 3}{12(1 - a_2)(1 - a_3)},$$

$$b_{42} = -\frac{4a_3^2 - a_2 - 5a_3 + 2}{24\sigma_4 a_2(a_3 - a_2)(1 - a_3)},$$

$$b_{43} = \frac{1 - 2a_2}{12\sigma_4 a_3(a_3 - a_2)},$$

$$b_{41} = 1 - b_{42} - b_{43},$$

$$b_{31} = a_2 - b_{32},$$

$$b_{21} = a_2,$$

$$\sigma_1 = 1 - \sigma_2 - \sigma_3 - \sigma_4.$$

Здесь, как уже отмечалось, $\sigma_3 \neq 0$, $\sigma_4 \neq 0$, т. е. $a_2 \neq 0{,}5$, $6a_2a_3 - 4a_2 - 4a_3 + 3 \neq 0$. Приведенное выше решение справедливо при $\delta \neq 0$, т. е. когда параметры a_2, a_3, a_4 удовлетворяют условиям

$$a_i \neq 0, \quad i=2,\ 3,\ 4, \quad a_2 \neq a_3, \quad a_2 \neq a_4, \quad a_3 \neq a_4.$$

Рассмотрим систему (38)—(41) при тех значениях параметров a_2, a_3, a_4, когда $\delta=0$. При $a_2=0$ система не имеет решения вследствие (41). При $a_3=0$ система (40) принимает вид

$$(\sigma_3 b_{32} + \sigma_4 b_{42}) a_2 = \frac{1}{6}, \quad (\sigma_3 b_{32} + \sigma_4 b_{42}) a_2^2 = \frac{1}{12}, \tag{49}$$

$$\sigma_4 a_2 a_4 b_{42} = \frac{1}{8},$$

откуда следует, что $a_2 = \frac{1}{2}$ и

$$\sigma_3 b_{32} + \sigma_4 b_{42} = \frac{1}{3}. \tag{50}$$

Далее, система (39) при $a_2 = \frac{1}{2}$, $a_3=0$ принимает вид

$$\sigma_2 + 2\sigma_4 a_4 = 1, \quad \sigma_2 + 4\sigma_4 a_4^2 = \frac{4}{3}, \quad \sigma_2 + 8\sigma_4 a_4^3 = 2$$

и имеет единственное решение

$$a_4 = 1, \quad \sigma_2 = \frac{2}{3}, \quad \sigma_4 = \frac{1}{6}.$$

Подставляя эти значения a_4, σ_2, σ_4 и $a_2 = \frac{1}{2}$ в уравнения (49), (50), получаем

$$b_{42} = \frac{3}{2}, \quad b_{32} = \frac{1}{12\sigma_3}.$$

Кроме того, из (41) имеем

$$b_{43} = \frac{1}{24\sigma_4 b_{32} a_2} = 6\sigma_3.$$

Таким образом, система (38) имеет следующее семейство решений, зависящее от параметра $\sigma_3 \neq 0$:

$$a_2 = \frac{1}{2}, \quad a_3 = 0, \quad a_4 = 1,$$
$$\sigma_2 = \frac{2}{3}, \quad \sigma_4 = \frac{1}{6},$$
$$b_{32} = \frac{1}{12\sigma_3}, \quad b_{42} = \frac{3}{2}, \quad b_{43} = 6\sigma_3,$$
$$b_{41} = -\frac{1}{2} - 6\sigma_3, \quad b_{31} = -\frac{1}{12\sigma_3}, \quad b_{21} = \frac{1}{2}, \quad \sigma_1 = \frac{1}{6} - \sigma_3.$$

Точно так же при условии $a_2=a_3$ система (38)—(41) имеет решение

$$a_2 = a_3 = \frac{1}{2}, \quad a_4 = 1,$$
$$\sigma_2 = \frac{2}{3} - \sigma_3, \quad \sigma_4 = \frac{1}{6},$$
$$b_{32} = \frac{1}{6\sigma_3}, \quad b_{42} = 1 - 3\sigma_3, \quad b_{43} = 3\sigma_3,$$
$$b_{41} = 0, \quad b_{31} = \frac{1}{2} - \frac{1}{6\sigma_3}, \quad b_{21} = \frac{1}{2}, \quad \sigma_1 = \frac{1}{6},$$

зависящее от параметра $\sigma_3 \neq 0$.

При $a_2=a_4$ имеем решение

$$a_2 = a_4 = 1, \quad a_3 = \frac{1}{2}, \quad \sigma_2 = \frac{1}{6} - \sigma_4, \quad \sigma_3 = \frac{2}{3},$$
$$b_{32} = \frac{1}{8}, \quad b_{42} = -\frac{1}{6\sigma_4}, \quad b_{43} = \frac{1}{3\sigma_4},$$
$$b_{41} = 1 - \frac{1}{6\sigma_4}, \quad b_{31} = \frac{3}{8}, \quad b_{21} = 1,$$
$$\sigma_1 = \frac{1}{6},$$

зависящее от параметра $\sigma_4 \neq 0$.

Определитель δ обращается в нуль еще в двух случаях: при $a_4=0$ и $a_3=a_4$. Оказывается, что в этих случаях система (38)—(41) не имеет решения. Пусть, например, $a_4=0$. Тогда из первых двух уравнений системы (39) получим

$$\sigma_2=\frac{3a_2-2}{6a_2(a_3-a_2)}, \quad \sigma_3=\frac{2-3a_2}{6a_3(a_3-a_2)}.$$

При этом последнее уравнение системы (39) приводит к условию

$$6a_2a_3-4a_2-4a_3+3=0. \tag{51}$$

Аналогично находим, что система (40), (41) разрешима относительно b_{32}, b_{42}, b_{43} только при условии

$$6a_2a_3-6a_2-4a_3+3=0. \tag{52}$$

Из (51), (52) находим $a_2=0$, что невозможно в силу (41). Точно так же доказывается, что не существует решений с $a_3=a_4$.

§ 3. Многошаговые разностные методы

1. Формулировка методов. Для решения задачи Коши

$$\frac{du}{dt}=f(t,u), \quad t>0, \qquad u(0)=u_0 \tag{1}$$

введем сетку

$$\omega_\tau=\{t_n=n\tau, \quad n=0, 1, \ldots\}$$

с постоянным шагом $\tau>0$. Обозначим через $y_n=y(t_n)$, $f_n=f(t_n, y_n)$ функции, определенные на сетке ω_τ. *Линейным m-шаговым разностным методом* называется система разностных уравнений

$$\frac{a_0y_n+a_1y_{n-1}+\ldots+a_my_{n-m}}{\tau}=b_0f_n+b_1f_{n-1}+\ldots+b_mf_{n-m}, \tag{2}$$
$$n=m, m+1, \ldots,$$

где a_k, b_k — числовые коэффициенты, не зависящие от n, $k=0, 1, \ldots$ $\ldots, m$, причем $a_0\neq0$.

Уравнение (2) следует рассматривать как рекуррентное соотношение, выражающее новое значение $y_n=y(t_n)$ через найденные ранее значения y_{n-1}, y_{n-2}, $\ldots$, y_{n-m}.

Расчет начинается с $n=m$, т. е. с уравнения

$$\frac{a_0y_m+a_1y_{m-1}+\ldots+a_my_0}{\tau}=b_0f_m+b_1f_{m-1}+\ldots+b_mf_0.$$

Отсюда видно, что для начала расчета необходимо задать m начальных значений y_0, y_1, $\ldots$, y_{m-1}. Значение y_0 определяется исходной задачей (1), а именно полагают $y_0=u_0$. Величины y_1, y_2, $\ldots$ $\ldots$, y_{m-1} можно вычислить, например, с помощью метода Рунге — Кутта. В дальнейшем будем предполагать, что начальные значения y_0, y_1, $\ldots$, y_{m-1} заданы.

Из уравнения (2) видно, что в отличие от методов Рунге — Кутта многошаговые разностные методы допускают вычисление правых частей только в точках основной сетки ω_τ.

Метод (2) называется *явным*, если $b_0=0$, и, следовательно, искомое значение y_n выражается явным образом через предыдущие значения $y_{n-1}, y_{n-2}, \ldots, y_{n-m}$. В противном случае (т. е. когда $b_0 \neq 0$) метод называется *неявным*. Тогда для нахождения y_n приходится решать нелинейное уравнение

$$\frac{a_0}{\tau} y_n - b_0 f(t_n, y_n) = F[y_{n-1}, y_{n-2}, \ldots, y_{n-m}],$$

где

$$F[y_{n-1}, y_{n-2}, \ldots, y_{n-m}] = \sum_{k=1}^{m} \left(b_k f_{n-k} - \frac{a_k}{\tau} y_{n-k}\right).$$

Обычно это уравнение решают методом Ньютона, выбирая начальное приближение $y_n^{(0)}$ равным y_{n-1}.

Заметим, что коэффициенты уравнения (2) определены с точностью до множителя. Чтобы устранить этот произвол, будем считать, что выполнено условие

$$\sum_{k=0}^{m} b_k = 1, \tag{3}$$

означающее, что правая часть разностного уравнения (2) аппроксимирует правую часть дифференциального уравнения (1).

В практике вычислений наибольшее распространение получили методы Адамса, которые представляют собой частный случай многошаговых методов (2), когда производная $u'(t)$ аппроксимируется только по двум точкам, t_n и t_{n-1}, т. е.

$$a_0 = -a_1 = 1, \quad a_k = 0, \quad k = 2, 3, \ldots, m.$$

Таким образом, *методы Адамса* имеют вид

$$\frac{y_n - y_{n-1}}{\tau} = \sum_{k=0}^{m} b_k f_{n-k}. \tag{4}$$

В случае $b_0=0$ методы Адамса называются *явными*, в случае $b_0 \neq 0$ — *неявными*.

При изучении разностных методов (2) мы рассмотрим прежде всего, как влияет выбор коэффициентов a_k, b_k на погрешность аппроксимации, а затем исследуем тесно связанные между собой вопросы устойчивости и сходимости.

2. Погрешность аппроксимации многошаговых методов. *Погрешностью аппроксимации на решении* или *невязкой разностного метода* (2) называется функция

$$\psi_n = -\sum_{k=0}^{m} \frac{a_k}{\tau} u_{n-k} + \sum_{k=0}^{m} b_k f(t_{n-k}, u_{n-k}), \tag{5}$$

получающаяся в результате подстановки точного решения $u(t)$ дифференциальной задачи (1) в разностное уравнение (2).

Выясним вопрос о порядке погрешности аппроксимации при $\tau\to 0$ в зависимости от выбора коэффициентов a_k, b_k, $k=0, 1, \ldots, m$. Будем предполагать при этом, что все рассматриваемые функции обладают необходимой гладкостью.

Разлагая функции $u_{n-k}=u(t_n-k\tau)$ в точке $t=t_n$ по формуле Тейлора, получим

$$u_{n-k}=\sum_{l=0}^{p}\frac{(-k\tau)^l u^{(l)}(t_n)}{l!}+O(\tau^{p+1}),$$

$$f(t_{n-k}, u_{n-k})=u'(t_n-k\tau)=$$
$$=\sum_{l=0}^{p-1}\frac{(-k\tau)^l u^{(l+1)}(t_n)}{l!}+O(\tau^p),\quad k=1, 2, \ldots, m.$$

Подставляя эти разложения в выражение (5) для погрешности аппроксимации, будем иметь

$$\psi_n=-\sum_{k=0}^{m}\frac{a_k}{\tau}\left(\sum_{l=0}^{p}\frac{(-k\tau)^l u^{(l)}(t_n)}{l!}\right)+$$
$$+\sum_{k=0}^{m}b_k\left(\sum_{l=0}^{p-1}\frac{(-k\tau)^l u^{(l+1)}(t_n)}{l!}\right)+O(\tau^p)=$$
$$=-\sum_{l=0}^{p}\left(\sum_{k=0}^{m}\frac{a_k}{\tau}\cdot\frac{(-k\tau)^l u^{(l)}(t_n)}{l!}\right)+$$
$$+\sum_{l=1}^{p}\left(\sum_{k=0}^{m}b_k\frac{(-k\tau)^{l-1}u^{(l)}(t_n)}{(l-1)!}\right)+O(\tau^p).$$

После очевидных преобразований приходим к разложению

$$\psi_n=-\left(\sum_{k=0}^{m}\frac{a_k}{\tau}\right)u(t_n)+$$
$$+\sum_{l=1}^{p}\left(\sum_{k=0}^{m}(-k\tau)^{l-1}\left(a_k\frac{k}{l}+b_k\right)\right)\frac{u^{(l)}(t_n)}{(l-1)!}+O(\tau^p).\tag{6}$$

Отсюда видно, что погрешность аппроксимации имеет порядок p, если выполнены условия

$$\sum_{k=0}^{m}a_k=0,\tag{7}$$

$$\sum_{k=0}^{m}k^{l-1}(ka_k+lb_k)=0,\quad l=1, 2, \ldots, p.\tag{8}$$

Вместе с условием нормировки (3) уравнения (7), (8) образуют систему из $p+2$ линейных алгебраических уравнений относительно $2(m+1)$ неизвестных

$$a_0, a_1, \ldots, a_m, b_0, b_1, \ldots, b_m.$$

Можно несколько упростить эту систему. А именно, рассмотрим уравнение (8) при $l=1$,

$$\sum_{k=0}^{m} ka_k + \sum_{k=0}^{m} b_k = 0$$

и учтем условие нормировки (3). Тогда получим уравнение

$$\sum_{k=0}^{m} ka_k = -1.$$

Окончательно получаем систему уравнений

$$\sum_{k=1}^{m} ka_k = -1, \tag{9}$$

$$\sum_{k=1}^{m} k^{l-1}(ka_k + lb_k) = 0, \quad l = 2, 3, \ldots, p,$$

которая содержит p уравнений и $2m$ неизвестных $a_1, a_2, \ldots, a_m$, $b_1, b_2, \ldots, b_m$. Коэффициенты a_0 и b_0 вычисляются по формулам

$$a_0 = -\sum_{k=1}^{m} a_k, \quad b_0 = 1 - \sum_{k=1}^{m} b_k. \tag{10}$$

Для того чтобы система (9) не была переопределена, необходимо потребовать, чтобы $p \leqslant 2m$. Это требование означает, что *порядок аппроксимации линейных m-шаговых разностных методов не может превосходить $2m$.*

Итак, наивысший достижимый порядок аппроксимации неявных m-шаговых методов равен $2m$, а явных — $2m-1$.

Заметим, что если в системе (8) отбросить последние n уравнений, $n=1, 2, \ldots, p-1$, то получим условия, обеспечивающие порядок аппроксимации $p-n$.

Для методов Адамса (4) условия p-го порядка аппроксимации (9) принимают вид

$$l\sum_{k=1}^{m} k^{l-1}b_k = 1, \quad l = 2, 3, \ldots, p, \quad b_0 = 1 - \sum_{k=1}^{m} b_k. \tag{11}$$

Отсюда видно, что *наивысший порядок аппроксимации m-шагового метода Адамса равен $m+1$, а наивысший порядок аппроксимации явного метода Адамса ($b_0=0$) равен m.*

3. Устойчивость и сходимость разностных методов. Оказывается, что методы наивысшего порядка аппроксимации практически непригодны для расчетов, так как они неустойчивы. Подробно вопросы устойчивости и сходимости разностных методов будут рассмотрены в следующем параграфе, а сейчас ограничимся изложением самых необходимых сведений.

Рассмотрим наряду с (2) однородное разностное уравнение

$$a_0 v_n + a_1 v_{n-1} + \ldots + a_m v_{n-m} = 0, \qquad n = m,\ m+1,\ \ldots, \tag{12}$$

и будем искать решения уравнения (12), имеющие вид $v_n = q^n$, где q — число, подлежащее определению. Тогда для нахождения q получаем уравнение

$$a_0 q^m + a_1 q^{m-1} + \ldots + a_{m-1} q + a_m = 0, \tag{13}$$

которое называется *характеристическим уравнением разностного метода* (2).

Говорят, что метод (2) удовлетворяет *условию корней*, если все корни $q_1, q_2, \ldots, q_m$ характеристического уравнения (13) лежат внутри или на границе единичного круга комплексной плоскости, причем на границе единичного круга нет кратных корней. Разностный метод (2), удовлетворяющий условию корней, называют *устойчивым методом*. Существует определенное ограничение на порядок аппроксимации устойчивого метода. Приведем без доказательства следующее утверждение.

Пусть метод (2) *удовлетворяет условию корней и имеет порядок аппроксимации p. Тогда $p \leqslant m+1$ при m нечетном и $p \leqslant m+2$ при m четном. Для явных m-шаговых устойчивых методов порядок аппроксимации не превосходит m.*

В § 4 будет доказана следующая теорема о связи между устойчивостью и сходимостью разностного метода (2) (см. теорему 2 из § 4).

Пусть метод (2) *удовлетворяет условию корней и $|f_u(t, u)| \leqslant L$ при $0 \leqslant t \leqslant T$. Тогда при $m\tau \leqslant t_n = n\tau \leqslant T$, $n \geqslant m$ и всех достаточно малых τ выполнена оценка*

$$|y_n - u(t_n)| \leqslant M\left(\max_{0 \leqslant j \leqslant m-1} |y_j - u(t_j)| + \max_{0 \leqslant k \leqslant n-m} |\psi_k|\right), \tag{14}$$

где ψ_k — погрешность аппроксимации, $y_j - u(t_j)$, $j = 0, 1, \ldots, m-1$ — погрешности в задании начальных условий и M — константа, зависящая от L, T и не зависящая от n.

Из оценки (14) следует, что если начальные погрешности $y_j - u(t_j)$, $j = 0, 1, \ldots, m-1$, и погрешность аппроксимации ψ_k, $k = 0, 1, \ldots, n-m$, являются величинами $O(\tau^p)$, $p > 0$, то и $y_n - u(t_n) = O(\tau^p)$ при $n \geqslant m$, т. е. метод сходится и имеет p-й порядок точности.

Таким образом, исследование сходимости метода (2) сводится к анализу погрешности аппроксимации и проверке условия корней.

Заметим, что методы Адамса

$$\frac{y_n - y_{n-1}}{\tau} = \sum_{k=0}^{m} b_k f(t_{n-k}, y_{n-k})$$

всегда удовлетворяют условию корней, так как для них $a_0 = -a_1 = 1$, т. е. $q = q_1 = 1$.

Простым примером метода, не удовлетворяющего условию корней, является явный двухшаговый метод

$$\frac{y_n + 4y_{n-1} - 5y_{n-2}}{6\tau} = \frac{2f_{n-1} + f_{n-2}}{3},$$

имеющий третий порядок аппроксимации.

4. Примеры многошаговых разностных методов. Наивысший порядок аппроксимации явных m-шаговых методов Адамса

$$\frac{y_n - y_{n-1}}{\tau} = b_1 f_{n-1} + b_2 f_{n-2} + \ldots + b_m f_{n-m} \qquad (15)$$

равен m. Согласно (11) условия m-го порядка аппроксимации имеют вид

$$\sum_{k=1}^{m} k^{l-1} b_k = \frac{1}{l}, \qquad l = 1, 2, \ldots, m. \qquad (16)$$

Решая систему (16), можно найти коэффициенты метода наивысшего порядка (15), (16) при каждом конкретном m. Так, при $m=1$ получаем метод Эйлера

$$\frac{y_n - y_{n-1}}{\tau} = f_{n-1}.$$

При $m=2, 3, 4, 5$ получаем соответственно следующие методы m-го порядка аппроксимации:

$$\frac{y_n - y_{n-1}}{\tau} = \frac{3}{2} f_{n-1} - \frac{1}{2} f_{n-2}, \qquad m = 2,$$

$$\frac{y_n - y_{n-1}}{\tau} = \frac{1}{12}(23f_{n-1} - 16f_{n-2} + 5f_{n-3}), \qquad m = 3,$$

$$\frac{y_n - y_{n-1}}{\tau} = \frac{1}{24}(55f_{n-1} - 59f_{n-2} + 37f_{n-3} - 9f_{n-4}), \qquad m = 4,$$

$$\frac{y_n - y_{n-1}}{\tau} = \frac{1}{720}(1901f_{n-1} - 2774f_{n-2} + 2616f_{n-3} - 1274f_{n-4} + 251f_{n-5}), \qquad m = 5.$$

Для неявных m-шаговых методов Адамса

$$\frac{y_n - y_{n-1}}{\tau} = b_0 f_n + b_1 f_{n-1} + \ldots + b_m f_{n-m} \qquad (17)$$

наивысший порядок аппроксимации равен $m+1$. Коэффициенты метода (17) наивысшего порядка находятся из системы (11) с $p=m+1$. При $m=1$ получаем метод второго порядка аппроксимации

$$\frac{y_n - y_{n-1}}{\tau} = \frac{1}{2}(f_n + f_{n-1}), \qquad p = 2,$$

называемый методом трапеций. При $m=2, 3, 4$ получаем соответственно следующие методы $(m+1)$-го порядка аппроксимации:

$$\frac{y_n-y_{n-1}}{\tau}=\frac{1}{12}(5f_n+8f_{n-1}-f_{n-2}), \quad p=3,$$

$$\frac{y_n-y_{n-1}}{\tau}=\frac{1}{24}(9f_n+19f_{n-1}-5f_{n-2}+f_{n-3}), \quad p=4,$$

$$\frac{y_n-y_{n-1}}{\tau}=\frac{1}{720}(251f_n+646f_{n-1}-264f_{n-2}+106f_{n-3}-19f_{n-4}), \quad p=5.$$

Выписанные выше неявные методы содержат искомое значение нелинейно, поэтому для их реализации необходимо применять итерационные методы. Например, для неявного метода Адамса четвертого порядка используется итерационный метод

$$\frac{y_n^{(s+1)}-y_{n-1}}{\tau}=\frac{1}{24}(9f(t_n, y_n^{(s)})+ \\ +19f(t_{n-1}, y_{n-1})-5f(t_{n-2}, y_{n-2})+f(t_{n-3}, y_{n-3})), \quad (18)$$

где s — номер итерации, $s=0, 1, \ldots$ В качестве начального значения $y_n^{(0)}$ можно взять решение, полученное с помощью явного метода Адамса третьего порядка, т. е. метода

$$\frac{y_n^{(0)}-y_{n-1}}{\tau}=\frac{1}{12}(23f(t_{n-1}, y_{n-1})-16f(t_{n-2}, y_{n-2})+5f(t_{n-3}, y_{n-3})). \quad (19)$$

Записывая (18) в виде

$$y_n^{(s+1)}=\frac{3\tau}{8}f(t_n, y_n^{(s)})+F,$$

получаем, что если $\left|\frac{\partial f}{\partial y}\right|\leqslant M$, то итерационный метод сходится при условии $\frac{3\tau M}{8}<1$, которое выполнено при достаточно малом τ.

Если в (18) ограничиться только одной итерацией $s=0$, то получим метод, называемый методом предиктор — корректор (предсказывающе-исправляющий).

§ 4. Сходимость и оценка погрешности многошагового разностного метода *)

1. Уравнение для погрешности. Для задачи Коши

$$\frac{du}{dt}=f(t, u), \quad t>0, \quad u(0)=u_0 \quad (1)$$

*) При первом чтении этот параграф можно опустить.