Глава 3

ЧИСЛЕННЫЕ МЕТОДЫ РЕШЕНИЯ ЖЕСТКИХ СИСТЕМ

1. Жесткие системы. Простейшие неявные методы. Рассмотрим применение многошаговых методов (2.2) для решения линейной системы (1.120), (1.121) с матрицей простой структуры. Точное решение этой системы дается формулой (1.123):

$$y(x) = \sum_{j=1}^{M} C_j e^{\lambda_j x} \mathbf{e}_j. \tag{1}$$

Разностное уравнение в данном случае примет вид

$$\sum_{i=0}^{k} \alpha_i y_{n+i} = h \sum_{i=0}^{k} \beta_i A y_{n+i}. \tag{2}$$

Разложим y_n по собственным векторам $\mathbf{e}_i$ матрицы A:

$$y_n = \sum_{j=1}^{M} C_n^j \mathbf{e}_j. \tag{3}$$

Подставляя (3) в (2), имеем

$$\sum_{i=0}^{k} \alpha_i \sum_{j=1}^{M} C_{n+i}^j \mathbf{e}_j - h \sum_{i=0}^{k} \beta_i A \sum_{j=1}^{M} C_{n+i}^j \mathbf{e}_j = \sum_{j=1}^{M} \Big(\sum_{i=0}^{k} \alpha_i C_{n+i}^j - h \sum_{i=0}^{k} \beta_i \lambda_j C_{n+i}^j \Big) \mathbf{e}_j = 0.$$

В силу линейной независимости векторов $\mathbf{e}_i$ отсюда следует, что

$$\sum_{i=0}^{k} \alpha_i C_{n+i}^j = h \sum_{i=0}^{k} \beta_i \lambda_j C_{n+i}^j, \qquad j = 1, \ldots, M. \tag{4}$$

Соотношение (4) совпадает с разностным уравнением (2.2), примененным для решения дифференциального уравнения (1.112) с $\lambda = \lambda_j$:

$$y' = \lambda_j y. \tag{5}$$

Оно показывает, как преобразуются коэффициенты C_n^j при численном интегрировании линейной системы (1.120), (1.121) многошаговым методом. Из (3) и (4) видно, что каждая составляющая

$$C_j e^{\lambda_j x} \mathbf{e}_j \tag{6}$$

решения (1), пропорциональная одному из собственных векторов, интегрируется независимо от остальных. Составляющая

$$C_n^j \mathbf{e}_j \tag{7}$$

решения y_n (3) соответствует составляющей (6) решения дифференциального уравнения (1.120), причем формулы преобразования коэффициентов C_n^j совпадают с разностными формулами интегрирования уравнения (5). Поэтому, если уравнение (5) проинтегрировано с достаточной точностью при всех $j = 1, \ldots, M$, то с достаточной точностью будет найдена составляющая (7), соответствующая собственному значению λ_j и, следовательно, обеспечена достаточная точность решения всей системы (1.120), (1.121) многошаговым методом.

Рассмотрим случай действительных отрицательных λ_j. В этом случае уравнение (5) является асимптотически устойчивым и любое его решение стремится к нулю при $x \to\propto$. Обсудим, какому условию должен удовлетворять шаг интегрирования h для того, чтобы решение разностного уравнения (4) хотя бы качественно отражало указанное свойство решений дифференциального уравнения (5). Другими словами, найдем при каких h решение уравнения (4) стремится к нулю при $n \to\propto$.

Для того чтобы всякое решение разностного уравнения стремилось к нулю, необходимо, чтобы все корни его характеристического уравнения были по модулю меньше единицы. Для явных методов Адамса (2.53) было получено необходимое для этого условие (2.94). Для методов Адамса в ординатной форме (2.41) это условие записывается в виде

$$h|\lambda| < \frac{2}{\sum\limits_{i=0}^{k} |B_i|}$$

и приводит к таким ограничениям на шаг интегрирования: для метода Адамса первого порядка аппроксимации

$$h|\lambda| < 2,$$

второго поряда аппроксимации

$$h|\lambda| < 1,$$

третьего порядка аппроксимации

$$h|\lambda| < \frac{2}{44/12} \cong 0.545,$$

четвертого порядка аппроксимации

$$h|\lambda| < \frac{2}{160/24} = 0.3,$$

пятого порядка аппроксимации

$$h|\lambda| < \frac{2}{8816/720} \cong 0.163.$$

Заметим, что для методов Рунге Кутта условие, обеспечивающее стремление к нулю решения соответствующего разностного уравнения (1.118), имеет вид: для метода второго порядка (1.14)

$$h|\lambda| < 2,$$

для метода четвертого порядка (2.22)

$$h|\lambda| < 2.785.$$

Следовательно, для явного метода Адамса необходимое условие асимптотической устойчивости разностного уравнения принимает вид

$$|h| < \frac{\text{const}}{\max\limits_{1 \le j \le M} |\lambda_j|}. \tag{8}$$

Если матрица системы дифференциальных уравнений (1.120) имеет большие по модулю отрицательные собственные значения, т.е.

$$\max_{1 \le j \le M} |\lambda_j| \gg 1, \tag{9}$$

то ограничение (8) на шаг h является на больших интервалах интегрирования слишком обременительным, так как оно требует использования очень малого шага на протяжении всего интервала интегрирования. Это обстоятельство вступает в противоречие с характером поведения точного решения системы. Если матрица системы имеет большой разброс собственных значений, например, в случае двух уравнений $|\lambda_1| \gg |\lambda_2|$, то первая составляющая точного решения

$$y(x) = C_1 e^{\lambda_1 x}\mathbf{e}_1 + C_2 e^{\lambda_2 x}\mathbf{e}_2$$

очень быстро затухает на отрезке, длина которого имеет порядок $\sim \tau = \dfrac{1}{|\lambda_1|}$, а затем становится ничтожно малой. Именно на этом отрезке она вносит свой вклад в решение $y(x)$. Вторая составляющая заметно изменяется на отрезке длины $\sim \dfrac{1}{|\lambda_2|}$, причем длина второго промежутка значительно превосходит длину первого. На втором отрезке именно вторая составляющая определяет поведение всего решения.

На первом отрезке скорость изменения решения определяется скоростью изменения первой составляющей и имеет большую величину. На втором отрезке скорость изменения решения определяется скоростью изменения второй составляющей и имеет относительно малую величину.

Таким образом, на интервале интегрирования выделяются два промежутка с разным характером поведения решения. На первом промежутке, называемом *пограничным слоем*, для воспроизведения быстро изменяющегося решения с приемлемой точностью необходим шаг интегрирования, удовлетворяющий условию $h \ll \dfrac{1}{|\lambda_1|}$. Казалось бы, что на втором промежутке, т.е. после прохождения пограничного слоя, где решение характеризуется малой скоростью изменения, шаг интегрирования мог бы быть увеличен. Однако, как показывает неравенство (8), этого сделать нельзя, так как при значениях шага, нарушающих условие (8), соответствующая составляющая решения y_n разностного уравнения не будет затухать.

Таким образом, на всем интервале определения решения необходимо применять малый шаг интегрирования, что приводит к огромному числу шагов на больших промежутках интегрирования и чрезмерному возрастанию времени решения задачи на ЭВМ.

Описанная ситуация встречается при интегрировании *жестких* систем уравнений. Жесткие системы характеризуются тем, что среди собственных чисел матрицы Якоби $\dfrac{\partial f}{\partial y}$ имеются большие по абсолютной величине, которые обязательно обладают большой по модулю отрицательной действительной частью, а собственные числа с положительной вещественной частью имеют малую величину. Как уже отмечалось, для жестких задач применяются специально сконструированные численные методы, один из которых — неявный метод Эйлера — был приведен в гл. 1.

Как уже было показано, применение неявного метода Эйлера (1.126) (формула которого получается из (2.70) при $k = 0$) для задачи вида (5) приводит к разностному уравнению первого порядка

$$y_{n+1} = F(\lambda h) y_n, \tag{10}$$

где

$$F(\lambda h) = \frac{1}{1 - \lambda h}. \tag{11}$$

Из (10) и (11) видно, что при произвольном $h > 0$ и отрицательном λ решение уравнения (10) стремится к нулю при $n \to \infty$. Тем самым обеспечивается стремление к нулю всех тех составляющих (7) решения (3) разностного уравнения

$$y_{n+1} = y_n + hAy_{n+1}, \tag{12}$$

которые соответствуют отрицательным собственным значениям матрицы A.

Для комплексных

$$\lambda = a + bi$$

с отрицательной реальной частью $a < 0$ также выполняется неравенство

$$|F(h\lambda)| < 1.$$

Поэтому, если среди собственных чисел матрицы A имеются комплексные, то будут стремиться к нулю также и те составляющие решения разностного уравнения, которые соответствуют комплексным собственным значениям с отрицательной реальной частью.

Так как точное решение устойчивого дифференциального уравнения вне пограничного слоя изменяется с малой скоростью и ведет себя плавно, то шаг интегрирования может быть увеличен без нарушения численной устойчивости приближенного решения. Таким образом, применение неявного метода Эйлера (1.126) к численному решению устойчивой линейной системы (1.120), (1.121) позволяет выбирать шаг интегрирования, ограничивая его единственным требованием достижения заданной точности приближенного решения.

Рассмотрим теперь неявный метод Адамса второго порядка аппроксимации. Формула этого метода получается из (1.70) при $k = 1$:

$$y_{n+1} = y_n + \frac{h}{2}\left(f(x_n, y_n) + f(x_{n+1}, y_{n+1})\right). \tag{13}$$

Формула (13) называется также *неявной формулой трапеций*. Применив (13) к решению (5), получим разностное уравнение

$$y_{n+1} = y_n + \frac{h}{2}\,\lambda\,(y_n + y_{n+1}), \tag{14}$$

откуда приходим к (10), где

$$F(\lambda h) = \frac{1 + \lambda\dfrac{h}{2}}{1 - \lambda\dfrac{h}{2}}.$$

Если $\lambda = a + bi$, то

$$|F(\lambda h)| = \left|\frac{1 + \dfrac{h}{2}\,a + \dfrac{h}{2}\,bi}{1 - \dfrac{h}{2}\,a - \dfrac{h}{2}\,bi}\right| = \frac{\sqrt{\left(1 + \dfrac{h}{2}\,a\right)^2 + \left(\dfrac{h}{2}\,b\right)^2}}{\sqrt{\left(1 - \dfrac{h}{2}\,a\right)^2 + \left(\dfrac{h}{2}\,b\right)^2}}.$$

При $a < 0$ имеем $|F(\lambda h)| < 1$, и любое решение уравнения (14) стремится к нулю при $n \to \infty$. Следовательно, обеспечивается стремление к нулю всех составляющих (7) решения (3) разностного уравнения

$$y_{n+1} = y_n + \frac{h}{2}\,(Ay_n + Ay_{n+1}), \tag{15}$$

которые соответствуют отрицательным собственным значениям матрицы A. Если среди собственных чисел матрицы A имеются комплексные, то стремятся к нулю и те составляющие решения разностного уравнения (15), которые соответствуют комплексным собственным значениям с отрицательной реальной частью. Поэтому применение метода трапеций, как и неявного метода Эйлера, для численного решения устойчивой линейной системы позволяет увеличивать длину шага интегрирования вне пограничного слоя без нарушения численной устойчивости приближенного решения. Единственным ограничением при этом является требование достижения заданной точности.

2. Методы дифференцирования назад. Для интегрирования жестких дифференциальных уравнений наиболее распространенными являются методы численного интегрирования, основанные на формулах дифференцирования назад. Эти формулы имеют вид (2.2) и могут быть построены, если производную $f(x_{n+k}, y(x_{n+k}))$ решения в точке $x = x_{n+k}$ аппроксимировать с помощью односторонних формул численного дифференцирования. Односторонние формулы численного дифференцирования можно получить разными способами.

Один из них заключается в следующем. По известным значениям функции $y(x)$ в узлах $x = x_{n+i},\ i = 0, \ldots, k$, строится интерполяционный многочлен Лангранжа $L_{k,n+k}(x)$. Тогда

$$y(x) = L_{k,n+k}(x) + r_{k,n+k}(x), \tag{16}$$

где

$$r_{k,n+k}(x) = \frac{y^{(k+1)}(\xi)}{(k+1)!} \prod_{i=0}^{k} (x - x_{n+i}),$$

ξ — промежуточная точка между x_{n+k} и x_n. Дифференцируя (16), получаем

$$y'(x) = L'_{k,n+k}(x) + r'_{k,n+k}(x).$$

Подставляя $x = x_{n+k}$, имеем

$$y'(x_{n+k}) = \frac{1}{h} \sum_{i=0}^{k} C_i y(x_{n+i}) + r'_{k,n+k}(x_{n+k}), \tag{17}$$

где C_i — вполне определенные числа,

$$r'_{k,n+k}(x_{n+k}) = \frac{y^{(k+1)}(\xi)}{k+1}\, h^k.$$

Умножим (17) на h:

$$hy'(x_{n+k}) = \sum_{i=0}^{k} C_i y(x_{n+i}) + hr'_{k,n+k}(x_{n+k}). \tag{18}$$

Чтобы получить формулу вида (2.2) с условием нормировки (2.16), поделим (18) на C_k:

$$\sum_{i=0}^{k} \alpha_i y(x_{n+i}) - h\beta_k f\left(x_{n+k}, y(x_{n+k})\right) = \rho_{n+k}, \tag{19}$$

где

$$\alpha_i = \frac{C_i}{C_k}, \quad \alpha_k = 1, \quad \beta_k = \frac{1}{C_k},$$

$$\rho_{n+k} = -hr'_{k,n+k}(x_{n+k})\beta_k = -\frac{\beta_k}{k+1}\, h^{k+1} y^{(k+1)}(\xi). \tag{20}$$

Отбрасывая в (19) остаточный член, получаем конечно-разностное уравнение

$$\sum_{i=0}^{k} \alpha_i y_{n+i} - h\beta_k f(x_{n+k}, y_{n+k}) = 0, \qquad \alpha_k = 1, \tag{21}$$

аппроксимирующее дифференциальное уравнение (1.1) с порядком $O(h^k)$. Локальная погрешность (20) формулы (21) может быть представлена в виде

$$\rho_{n+k} = -\frac{\beta_k}{k+1}\, h^{k+1} y^{(k+1)}(x_n) + O(h^{h+2}). \tag{22}$$

Несколько формул вида (21) различной степени вместе с указанием порядка разностного уравнения и коэффициента при главном члене погрешности приведены в табл. 1.

Т а б л и ц а 1.

Формулы дифференцирования назад для решения жестких уравнений

	Формула	Порядок k	Степень s	Локальная погрешность $C_{k+1}\, h^{k+1} y^{(k+1)}(x_n)$
1	2	3	4	5
1	$y_{n+2} = \frac{4}{3}\, y_{n+1} - \frac{1}{3}\, y_n + \frac{2}{3}\, h\, f_{n+2}$	2	2	$-\frac{2}{9}\, h^3 y^{(3)}$
2	$y_{n+3} = \frac{18}{11}\, y_{n+2} - \frac{9}{11}\, y_{n+1} + \frac{2}{11}\, y_n + \frac{6}{11}\, h\, f_{n+3}$	3	3	$-\frac{3}{22}\, h^4 y^{(4)}$
3	$y_{n+4} = \frac{48}{25}\, y_{n+3} - \frac{36}{25}\, y_{n+2} + \frac{16}{25}\, y_{n+1} - \frac{3}{25}\, y_n + \frac{12}{25}\, h\, f_{n+4}$	4	4	$-\frac{12}{125}\, h^5 y^{(5)}$
4	$y_{n+5} = \frac{300}{137}\, y_{n+4} - \frac{300}{137}\, y_{n+3} + \frac{200}{137}\, y_{n+2} - \frac{75}{137}\, y_{n+1} + \frac{12}{137}\, y_n + \frac{60}{137}\, h\, f_{n+5}$	5	5	$-\frac{10}{137}\, h^6 y^{(6)}$
5	$y_{n+6} = \frac{360}{147}\, y_{n+5} - \frac{450}{147}\, y_{n+4} + \frac{400}{147}\, y_{n+3} - \frac{225}{147}\, y_{n+2} + \frac{72}{147}\, y_{n+1} - \frac{10}{147}\, y_n + \frac{60}{147}\, h\, f_{n+6}$	6	6	$-\frac{20}{343}\, h^7 y^{(7)}$

Заметим, что неявный метод Эйлера (1.126) является методом дифференцирования назад первого порядка.

Рассмотрим применение метода второго порядка для решения линейной системы (1.120) с начальным условием (1.121), матрица A коэффициентов которой имеет отрицательные собственные значение. Применим формулу (21) при $k = 2$

$$y_{n+2} - \frac{4}{3}\, y_{n+1} + \frac{1}{3}\, y_n - \frac{2}{3}\, h f_{n+2} = 0 \tag{23}$$

к уравнению (5). Тогда получим разностное уравнение

$$y_{n+2} - \frac{4}{3}\, y_{n+1} + \frac{1}{3}\, y_n - \frac{2}{3}\, h\lambda y_{n+2} = 0,$$

или

$$\left(1 - \frac{2}{3}\, h\lambda\right) y_{n+2} - \frac{4}{3}\, y_{n+1} + \frac{1}{3}\, y_n = 0. \tag{24}$$

Общее решение (24) имеет вид

$$y_n = C_1 z_1^n + C_2 z_2^n, \tag{25}$$

где z_1, z_2 — корни характеристического уравнения

$$\left(1 - \frac{2}{3}\, h\lambda\right) z^2 - \frac{4}{3}\, z + \frac{1}{3} = 0. \tag{26}$$

Непосредственные вычисления дают

$$z_{1,2} = \frac{\frac{4}{3} \pm \frac{2}{3}\sqrt{1 + 2h\lambda}}{2\left(1 - \frac{2}{3} h\lambda\right)} = \frac{2 \pm \sqrt{1 + 2h\lambda}}{3\left(1 - \frac{2}{3} h\lambda\right)}.$$

Пусть $\lambda < 0$ и $h > 0$. Если $1 + 2h\lambda > 0$, то $|z_{1,2}| < 1$. Если $1 + 2h\lambda < 0$, то

$$\sqrt{1+2h\lambda} = i\sqrt{2h|\lambda| - 1}\,.$$

Тогда

$$|z_{1,2}|^2 = \frac{4 + (2h|\lambda| - 1)}{9\left(1 + \dfrac{2}{3}\,h|\lambda|\right)^2} = \frac{3 + 2h|\lambda|}{(3 + 2h|\lambda|)^2} = \frac{1}{3 + 2h|\lambda|} < \frac{1}{3} < 1.$$

Таким образом, $|z_{1,2}| < 1$ при любом $\lambda < 0$. Из (25) следует, что при $\lambda < 0$ любое решение разностного уравнение (24) стремится к нулю при $n \to\infty$. Если формула (23) применяется к решению задачи (1.120), (1.121), то все составляющие (7) решения (3) разностного уравнения

$$y_{n+2} - \frac{4}{3}\,y_{n+1} + \frac{1}{3}\,y_n - \frac{2}{3}\,hAy_{n+2} = 0$$

стремятся к нулю при $n \to\infty$. Значит, применение метода дифференцирования назад, основанного на формуле (23), позволяет увеличивать шаг интегрирования после прохождения пограничного слоя. Таким же свойством обладают все методы, приведенные в табл. 1.

Показано, что и для комплексных $\lambda = a + bi$ с отрицательной реальной частью $a < 0$ корни уравнения (26) по модулю меньше единицы. Следовательно, и для устойчивых систем, имеющих кроме действительных еще и комплексные собственные значения с отрицательными реальными частями, применение (23) позволяет увеличивать шаг интегрирования, ограничивая его единственным требованием достижения заданной точности.

Заметим, что константа в локальной погрешности неявного метода трапеций меньше соответствующей константы формулы (23).

Теперь перейдем к рассмотрению других методов, основанных на формулах дифференцирования назад. Методы дифференцирования назад (21), примененные к уравнению вида (5), приводят к разностному уравнению

$$\sum_{i=0}^{k} \alpha_i y_{n+i} - h\beta_k \lambda y_{n+k} = 0. \tag{27}$$

Соответствующее характеристическое уравнение имеет вид

$$\sum_{i=0}^{k} \alpha_i z^i - h\beta_k \lambda z^k = 0.$$

Доказано, что при $k = 3, 4, 5, 6$ корни этого уравнения также удовлетворяют условию $|z| < 1$ для всех $h\lambda = h(a+bi)$, принадлежащих некоторой области комплексной плоскости, содержащей бесконечный клин

$$\pi - \alpha < \arg\lambda < \pi + \alpha, \tag{28}$$

где

$$0 < \alpha < \frac{\pi}{2}.$$

Кроме указанного клина эта область содержит полуплоскость, задаваемую неравенством

$$ha < D,$$

где D — некоторое отрицательное число. Следовательно, все составляющие (7) решения (3) разностного уравнения

$$\sum_{i=0}^{k} \alpha_i y_{n+i} - h\beta_k A y_{n+k} = 0$$

будут затухать, если для всех собственных чисел λ_j значения $h\lambda_j$ принадлежат указанной области.

3. Реализация неявных методов. Для решения разностных уравнений (2.2) в гл. 2 предлагался метод простой интерации (2.96), необходимым условием сходимости которого является условие (2.97)

$$hL|\beta_k| < 1.$$

Однако в случае жестких систем уравнений выполняется условие

$$\max_{1 \le j \le M} |\lambda_j| \gg 1.$$

Следовательно, в силу соотношения

$$\max_{1 \le j \le M} |\lambda_j| \le L$$

константа Липшица для таких систем велика:

$$L \gg 1.$$

Поэтому условие (2.97) сходимости метода простой интерации приводит к сильному ограничению на шаг интегрирования, которое мы старались устранить. Вот почему при интегрировании жестких систем от использования метода простой итерации следует отказаться.

Перепишем формулы неявных методов Эйлера (1.126), трапеции (13) и дифференцирования назад (21) в следующем виде:

$$y_{n+1} - y_n - hf(x_{n+1}, y_{n+1}) = 0, \tag{29}$$

$$y_{n+1} - y_n - \frac{h}{2} f(x_n, y_n) - \frac{h}{2} f(x_{n+1}, y_{n+1}) = 0, \tag{30}$$

$$y_{n+k} + \sum_{i=0}^{k-1} \alpha_i y_{n+i} - h\beta_k f(x_{n+k}, y_{n+k}) = 0 \tag{31}$$

и применим для решения этих уравнений метод Ньютона. Тогда итерационный процесс для уравнения (31) может быть записан так: для $\nu = 0, 1, \dots$

$$y_{n+k}^{(\nu+1)} = y_{n+k}^{(\nu)} - \left(E - h\beta_k \frac{\partial f(x_{n+k}, y_{n+k}^{(\nu)})}{\partial y}\right)^{-1} \times \left(y_{n+k}^{(\nu)} + \sum_{i=0}^{k-1} \alpha_i y_{n+i} - h\beta_k f(x_{n+k}, y_{n+k}^{(\nu)})\right). \tag{32}$$

Формула (32) может быть реализована либо вычислением обратной для

$$G = E - h\beta_k \frac{\partial f(x_{n+k}, y_{n+k}^{(\nu)})}{\partial y}$$

матрицы и непосредственным нахождением $y_{n+k}^{(\nu+1)}$ по формуле

$$y_{n+k}^{(\nu+1)} = y_{n+k}^{(\nu)} - G^{-1}(y_{n+k}^{(\nu)} + \upsilon - h\beta_k f(x_{n+k}, y_{n+k}^{(\nu)})), \tag{32'}$$

где

$$\upsilon = \sum_{i=0}^{k-1} \alpha_i y_{n+i},$$

либо решением линейной системы уравнений

$$G(y_{n+k}^{(\nu+1)} - y_{n+k}^{(\nu)}) = -y_{n+k}^{(\nu)} - \upsilon + h\beta_k f(x_{n+k}, y_{n+k}^{(\nu)}). \tag{33}$$

Для решения системы (33) может быть применено LU–разложение матрицы G, где L — нижняя треугольная матрица, U — верхняя треугольная матрица, и последовательное решение двух линейных систем: сначала решение системы

$$Lz = -y_{n+k}^{(\nu)} - \upsilon + h\beta_k f(x_{n+k}, y_{n+k}^{(\nu)}) \tag{34}$$

с нижней треугольной матрицей, затем решение системы

$$Uw = z \tag{35}$$

с верхней треугольной матрицей. После этого решение (33) находится по формуле

$$y_{n+k}^{(\nu+1)} = y_{n+k}^{(\nu)} + w.$$

На каждой итерации требуется вычисление матрицы Якоби, обращение матрицы G или решение системы (33). Часто применяется модифицированный метод Ньютона, который заключается в том, что вычисление матрицы Якоби и обращение матрицы G или ее LU–разложение производятся только один раз и все итерации для $\nu = 1, 2, \ldots$ выполняются с одной и той же обратной матрицей G^{-1} или с одними и теми же матрицами L и U:

$$y_{n+k}^{(\nu+1)} = y_{n+k}^{(\nu)} - \left(E - h\beta_k \frac{\partial f(x_{n+k}, y_{n+k}^{(0)})}{\partial y}\right)^{-1} \times \left(y_{n+k}^{(\nu)} + \upsilon - h\beta_k f(x_{n+k}, y_{n+k}^{(\nu)})\right).$$

Далее, если матрица Якоби мало изменяется от точки (x_{n+k}, y_{n+k}) к точке $(x_{n+k+1},\, y_{n+k+1})$, то одна и та же матрица

$$G^{-1} = \left(E - h\beta_k \frac{\partial f(x_{n+k}, y_{n+k}^{(0)})}{\partial y}\right)^{-1}$$

(или одна и та же LU–факторизация) может быть использована для вычисления решения в нескольких точках $x_{n+k},\, x_{n+k+1},\, x_{n+k+2}, \ldots$.

Если же при нахождении решения в некоторой точке x_{n+k+m} сходимость не достигается за максимально допустимое число итераций, то вновь вычисляется матрица Якоби и вновь находится обратная матрица

$$G^{-1} = \left(E - h\beta_k \frac{\partial f(x_{n+k+m}, y_{n+k+m}^{(0)})}{\partial y}\right)^{-1}$$

или снова производится LU–разложение матрицы G. Аналогичный итерационный процесс производится и при решении уравнений (29), (30).

Теперь обсудим вопрос о выборе начального приближения. Выбор хорошего начального приближения имеет существенное значение, так как от него зависит не только сходимость, но и количество итераций, необходимых для достижения заданной точности, а значит, и число вычислений и обращений (или факторизаций) итерационной матрицы G. Для вычисления начального приближения используются различные способы. Например, начальное приближение может быть найдено с помощью явной формулы вида

$$y_{n+k} = -\sum_{i=0}^{k-1} \alpha_i y_{n+i} + h\beta_{k-1} f(x_{n+k-1}, y_{n+k-1}), \tag{36}$$

которую можно интерпретировать как интерполяционную формулу Эрмита, построенную по значениям функции $y(x)$ в точках $x = x_j,\ j = n, \ldots, n+k-1$, и ее производной в точке x_{n+k-1}. При этом локальная погрешность (36) имеет порядок $O(h^{k+1})$.

Несколько таких формул разных степеней приведены ниже. Формула второй степени:

$$y_{n+2} = y_n + 2hf(x_{n+1}, y_{n+1}).$$

Формула третьей степени:

$$y_{n+3} = -\frac{3}{2}\,y_{n+2} + 3y_{n+1} - \frac{1}{2}\,y_n + 3hf(x_{n+2}, y_{n+2}).$$

Формула четвертой степени:

$$y_{n+4} = -\frac{10}{3}\,y_{n+3} + 6y_{n+2} - 2y_{n+1} + \frac{1}{3}\,y_n + 4hf(x_{n+3}, y_{n+3}).$$

Формула пятой степени:

$$y_{n+5} = -\frac{65}{12}\,y_{n+4} + 10y_{n+3} - 5y_{n+2} + \frac{5}{3}\,y_{n+1} - \frac{1}{4}\,y_n + 5hf(x_{n+4}, y_{n+4}).$$

Формула шестой степени:

$$y_{n+6} = -\frac{77}{10}\,y_{n+5} + 15y_{n+4} - 10y_{n+3} + 5y_{n+2} - \frac{3}{2}\,y_{n+1} + \frac{1}{5}\,y_n + 6hf(x_{n+5}, y_{n+5}).$$

Приведенные формулы могут быть построены также методом неопределенных коэффициентов. Например, для того чтобы получить формулу k-й степени (36), коэффициенты α_i и β_{k-1} подбираются таким образом, чтобы формула (36) была точна для всех алгебраических многочленов степени не выше k.

Начальное приближение может быть найдено также по экстраполяционной формуле

$$y_{n+k} = \sum_{i=0}^{k} A_i y_{n-1+i},$$

построенной по значениям функции $y(x)$ в узлах $x = x_j,\ j = n-1, n, \ldots, n+k-1$, при этом локальная погрешность предсказанного значения также имеет порядок $O(h^{k+1})$. Несколько таких формул разных степеней приведены ниже. Формула первой степени:

$$y_{n+2} = 2y_{n+1} - y_n.$$

Формула второй степени:

$$y_{n+3} = 3y_{n+2} - 3y_{n+1} + y_n.$$

Формула третьей степени:

$$y_{n+4} = 4y_{n+3} - 6y_{n+2} + 4y_{n+1} - y_n.$$

Формула четвертой степени:

$$y_{n+5} = 5y_{n+4} - 10y_{n+3} + 10y_{n+2} - 5y_{n+1} + y_n.$$

Формула пятой степени:

$$y_{n+6} = 6y_{n+5} - 15y_{n+4} + 20y_{n+3} - 15y_{n+2} + 6y_{n+1} - y_n.$$

Рассмотрим теперь вопрос о вычислении матрицы Якоби $\dfrac{\partial f}{\partial y}$. При вычислении решения жесткой системы наряду с формулами для вычисления значений правой части $f(x, y)$ используются формулы для частных производных. Однако бывает так, что этими формулами не всегда удобно воспользоваться. Поэтому для вычисления частных производных применяются формулы численного дифференцирования, например, формулы первого порядка по Δy^j

$$\frac{\partial f^i}{\partial y^j} = \frac{f^i(x,\, y^1, \dots, y^{j-1}, y^j + \Delta y^j, y^{j+1}, \dots, y^M) - f^i(x,\, y^1, \dots, y^{j-1}, y^j, \dots, y^M)}{\Delta y^j} \tag{37}$$

и второго порядка по Δy^i

$$\frac{\partial f^j}{\partial y^j} = \frac{f^i(x, y^1, \dots, y^{j-1}, y^j + \Delta y^j, y^{j+1}, \dots, y^M) - f^i(x, y^1, \dots, y^{j-1}, y^j - \Delta y^j, y^{j+1}, \dots, y^M)}{2\Delta y^j}. \tag{38}$$

При вычислении матрицы Якоби с помощью (37) требуется $M+1$ вычислений правой части $f(x,y) = (f^1(x,y), \dots, f^M(x,y))$, а с помощью (38) — $2M$ вычислений.

4. Метод Гира. *4.1. Вывод формул численного интегрирования.* Перейдем к изложению еще одного метода решения жестких систем. Этот метод строится на основе формул дифференцирования назад (21).

Предположим, что начальное приближение для решения уравнения (21) находится по формуле (36). Обозначим $m = n + k$ и перепишем формулу (36) в следующем виде:

$$y_m = -\sum_{i=1}^{k} \alpha_{k-i} y_{m-i} + h\beta_{k-1} f(x_{m-1}, y_{m-1}).$$

Обозначим $A_i = -\alpha_{k-i},\ B_1 = \beta_{k-1}$ и заменим m на n. Тогда

$$y_n = \sum_{i=1}^{k} A_i y_{n-i} + hB_1 f(x_{n-1}, y_{n-1}).$$

Аналогично преобразуем формулу (21):

$$y_m = -\sum_{i=1}^{k} \alpha_{k-i} y_{m-i} + h\beta_k f(x_m, y_m).$$

Обозначим $a_i = -\alpha_{k-i},\ b_0 = \beta_k$ и заменим m на n. Тогда

$$y_n = \sum_{i=1}^{k} a_i y_{n-i} + hb_0 f(x_n, y_n).$$

Рассмотрим следующий итерационный процесс:

$$y_n^{(0)} = \sum_{i=1}^{k} A_i y_{n-i} + hB_1 f_{n-1}, \tag{39}$$

$$y_n^{(\nu+1)} = \sum_{i=1}^{k} a_i y_{n-i} + hb_0 f(x_n, y_n^{(\nu)}), \qquad \nu = 0, 1, 2, \dots. \tag{40}$$

Из (40) получаем

$$y_n^{(\nu+1)} = y_n^{(\nu)} + b_0(hf(x_n, y_n^{(\nu)}) - hf(x_n, y_n^{(\nu-1)})), \qquad \nu = 1, 2, \dots.$$

Определим ν-е приближение для производной $hy'(x_n)$ по формуле

$$hy_n'^{(\nu)} = hf(x_n, y_n^{(\nu-1)}), \qquad \nu = 1, 2, \dots; \tag{41}$$

погрешность этого приближения для $hy'(x_n)$ имеет порядок $O(h^{k+2})$. Тогда

$$y_n^{(\nu+1)} = y_n^{(\nu)} + b_0(hf(x_n, y_n^{(\nu)}) - hy_n'^{(\nu)}), \qquad \nu = 1, 2 \dots. \tag{42}$$

Вычтем (39) из (40) при $\nu = 0$:

$$\begin{aligned} y_n^{(1)} &= y_n^{(0)} + \sum_{i=1}^{k}(a_i - A_i)y_{n-i} + hb_0 f(x_n, y_n^{(0)}) - hB_1 f_{n-1} = \\ &= y_n^{(0)} + b_0\left(hf(x_n, y_n^{(0)}) - \sum_{i=1}^{k}\frac{A_i - a_i}{b_0}\,y_{n-i} - h\,\frac{B_1}{b_0}\,f_{n-1}\right). \end{aligned} \tag{43}$$

Обозначим $\gamma_i = \dfrac{A_i - a_i}{b_0}$, $\delta_1 = \dfrac{B_1}{b_0}$ и определим начальное приближение для производной $hy'(x_n)$ по формуле

$$hy_n^{'(0)} = \sum_{i=1}^{k}\gamma_i y_{n-i} + h\delta_1 f_{n-1}. \tag{44}$$

Заметим, что погрешность этого приближения к $hy'(x_n)$ имеет порядок $O(h^{k+1})$. Тогда (43) примет вид

$$y_n^{(1)} = y_n^{(0)} + b_0(hf(x_n, y_n^{(0)}) - hy_n^{'(0)}). \tag{45}$$

Из (45) следует, что формула (42) справедлива также для $\nu = 0$. При этом $hy_n^{'(0)}$ определяется с помощью (44).

Введем в рассмотрение векторы

$$\mathbf{Y}_n = (y_n, hy_n', y_{n-1}, \ldots, y_{n-k+1})^T$$

и

$$\mathbf{Y}(x_n) = (y(x_n), hy'(x_n), y(x_{n-1}), \ldots, y(x_{n-k+1}))^T.$$

Первая компонента y_n вектора $\mathbf{Y}_n$ аппроксимирует $y(x_n)$, вторая компонента hy_n' вектора $\mathbf{Y}_n$ аппроксимирует $hy'(x_n)$. Остальные компоненты y_{n-l} аппроксимируют значения точного решения $y(x_{n-l})$, $l = 1, 2, \ldots, k-1$. Определим также вектор

$$\mathbf{Y}_n^{(\nu)} = (y_n^{(\nu)}, hy_n^{'(\nu)}, y_{n-1}, \ldots, y_{n-k+1})^T$$

и матрицу

$$D = \begin{pmatrix} A_1 & B_1 & A_2 & \ldots & A_{k-1} & A_k \\ \gamma_1 & \delta_1 & \gamma_2 & \ldots & \gamma_{k-1} & \gamma_k \\ 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 & 0 \end{pmatrix}.$$

Тогда предсказывающие формулы (39) и (44) для решения $y(x_n)$ и производной $hy'(x_n)$ могут быть представлены в виде следующего разностного векторного соотношения:

$$\mathbf{Y}_n^{(0)} = D\mathbf{Y}_{n-1}. \tag{46}$$

Погрешность первой и второй компоненты вектора $\mathbf{Y}_n^{(0)}$ имеет порядок $O(h^{k+1})$, а для остальных равна нулю, если $y_{n-i} = y(x_{n-i})$, $i = 1, \ldots, k-1$, и $O(h^{k+2})$, если $y(x_{n-i}) - y_{n-i} = O(h^{k+2})$. Следовательно,

$$\mathbf{Y}(x_n) - \mathbf{Y}_n^{(0)} = O(h^{k+1}). \tag{47}$$

Из (41) следует, что

$$hy_n^{'(\nu+1)} = hy_n^{'(\nu)} + hf(x_n, y_n^{(\nu)}) - hy_n^{'(\nu)}. \tag{48}$$

Введем вектор $\mathbf{c} = (b_0, 1, 0, \ldots, 0)^T$ и функцию невязки

$$F(\mathbf{Y}_n^{(\nu)}) = hf(x_n, y_n^{(\nu)}) - hy_n'^{(\nu)}. \tag{49}$$

Тогда исправляющие формулы (42) и (48) для решения $y(x_n)$ и производной $hy'(x_n)$ могут быть представлены в виде

$$\mathbf{Y}_n^{(\nu+1)} = \mathbf{Y}_n^{(\nu)} + \mathbf{c}F(\mathbf{Y}_n^{(\nu)}). \tag{50}$$

Погрешность первой компоненты вектора $\mathbf{Y}_n^{(\nu+1)}$ имеет порядок $O(h^{k+1})$, погрешность второй компоненты — $O(h^{k+2})$, а для остальных компонент погрешность такая же, как для $\mathbf{Y}_n^{(0)}$. Следовательно,

$$\mathbf{Y}(x_n) - \mathbf{Y}_n^{(\nu+1)} = O(h^{k+1}). \tag{51}$$

Итерационный процесс (39), (40) сходится к y_n , если h достаточно мало. Следовательно, будет сходиться итерационный процесс (46), (50):

$$\mathbf{Y}_n^{(\nu)} \longrightarrow \mathbf{Y}_n$$

и

$$\mathbf{Y}(x_n) - \mathbf{Y}_n = O(h^{k+1}). \tag{52}$$

Определим линейное преобразование Q, переводящее вектор $\mathbf{Y}(x_n)$ в вектор $\mathbf{Z}(x_n) + O(h^{k+1})$, где

$$\mathbf{Z}(x_n) = \left(y(x_n), hy'(x_n), \frac{h^2y''(x_n)}{2}, \ldots, \frac{h^ky^{(k)}(x_n)}{k!}\right)^T,$$

т.е.

$$Q\mathbf{Y}(x_n) = \mathbf{Z}(x_n) + O(h^{k+1}), \tag{53}$$

или, записывая покомпонентно,

$$\begin{pmatrix} y(x_n) \\ hy'(x_n) \\ y(x_{n-1}) \\ y(x_{n-2}) \\ \cdots \\ y(x_{n-k+1}) \end{pmatrix} \xrightarrow{Q} \begin{pmatrix} y(x_n) \\ hy'(x_n) \\ \dfrac{h^2y''(x_n)}{2} \\ \dfrac{h^3y'''(x_n)}{6} \\ \cdots \\ \dfrac{h^ky^{(k)}(x_n)}{k!} \end{pmatrix} + O(h^{k+1}). \tag{53$'$}$$

Вектор $\mathbf{Z}(x_n)$ называется *вектором Нордсика.*

Матрицу Q такого преобразования можно найти следующим образом. Так как преобразование Q не меняет первые две компоненты вектора $\mathbf{Y}(x_n)$, то первые две строки матрицы Q имеют вид

$$\begin{array}{ccccc} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0. \end{array}$$

Чтобы найти l-ю строку матрицы Q

$$q_{l1}, q_{l2}, \ldots, q_{l,k+1} \qquad (l \geq 3),$$

надо разложить все слагаемые в левой части равенства

$$q_{l1}y(x_n) + q_{l2}hy'(x_n) + q_{l3}y(x_{n-1}) + \ldots + q_{l,k+1}y(x_{n-k+1}) = \frac{h^{l-1}y^{(l-1)}(x_n)}{(l-1)!} + O(h^{k+1})$$

по степеням h, а затем приравнять коэффициенты при одинаковых степенях h справа и слева. Получим $k+1$ уравнений с $k+1$ неизвестными. Решая эту систему, найдем элементы l-й строки. Таким способом находятся строки третья, четвертая, ..., $(k+1)$-я. Например, если $k=3$, то матрица преобразования

$$\begin{pmatrix} y(x_n) \\ hy'(x_n) \\ y(x_{n-1}) \\ y(x_{n-2}) \end{pmatrix} \xrightarrow{Q} \begin{pmatrix} y(x_n) \\ hy'(x_n) \\ \dfrac{h^2 y''(x_n)}{2} \\ \dfrac{h^3 y'''(x_n)}{6} \end{pmatrix} + O(h^4)$$

равна

$$Q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -\dfrac{7}{4} & \dfrac{3}{2} & 2 & -\dfrac{1}{4} \\ -\dfrac{3}{4} & \dfrac{1}{2} & 1 & -\dfrac{1}{4} \end{pmatrix}.$$

Элементы третьей строки являются решением системы уравнений

$$\begin{cases} q_{31} + q_{33} + q_{34} = 0, \\ q_{32} - q_{33} - 2q_{34} = 0, \\ \dfrac{q_{33}}{2} + 2q_{34} = \dfrac{1}{2}, \\ -\dfrac{q_{33}}{6} - \dfrac{8}{6}\,q_{34} = 0, \end{cases}$$

а элементы четвертой строки — решением системы уравнений

$$\begin{cases} q_{41} + q_{43} + q_{44} = 0, \\ q_{42} - q_{43} - 2q_{44} = 0, \\ \dfrac{q_{43}}{2} + 2q_{44} = 0, \\ -\dfrac{q_{43}}{6} - \dfrac{8}{6}\,q_{44} = \dfrac{1}{6}. \end{cases}$$

Заметим, что матрица Q не зависит от h. При $k=4$ матрица преобразования

$$\begin{pmatrix} y(x_n) \\ hy'(x_n) \\ y(x_{n-1}) \\ y(x_{n-2}) \\ y(x_{n-3}) \end{pmatrix} \xrightarrow{Q} \begin{pmatrix} y(x_n) \\ hy'(x_n) \\ \dfrac{h^2 y''(x_n)}{2} \\ \dfrac{h^3 y'''(x_n)}{3!} \\ \dfrac{h^4 y^4(x_n)}{4!} \end{pmatrix} + O(h^5)$$

равна

$$Q = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ -\dfrac{85}{36} & \dfrac{11}{6} & 3 & -\dfrac{3}{4} & \dfrac{1}{9} \\ \dfrac{5}{3} & -1 & -\dfrac{5}{2} & 1 & -\dfrac{1}{6} \\ -\dfrac{11}{36} & \dfrac{1}{6} & \dfrac{1}{2} & -\dfrac{1}{4} & \dfrac{1}{18} \end{pmatrix}.$$

Элементы l-й строки, $l = 3, 4, 5$, являются решением следующей системы уравнений:

$$\begin{cases} q_{l1} + q_{l3} + q_{l4} + q_{l5} & = 0 & \text{— коэффициент при} & h^0, \\ q_{l2} - q_{l3} - 2q_{l4} - 3q_{l5} & = 0 & \text{— коэффициент при} & h, \\ \dfrac{q_{l3}}{2} + 2q_{l4} + \dfrac{9}{2}\, q_{l5} & = \dfrac{\delta_{l3}}{2} & \text{— коэффициент при} & h^2, \\ -\dfrac{q_{l3}}{6} - \dfrac{8}{6} q_{l4} - \dfrac{27}{6} q_{l5} & = \dfrac{\delta_{l4}}{6} & \text{— коэффициент при} & h^3, \\ \dfrac{q_{l3}}{24} + \dfrac{16}{24} q_{l4} + \dfrac{81}{24} q_{l5} & = \dfrac{\delta_{l5}}{24} & \text{— коэффициент при} & h^4. \end{cases}$$

Здесь $\delta_{l3},\ \delta_{l4},\ \delta_{l5}$ — символ Кронекера.

Преобразование Q переводит вектор $\mathbf{Y}_n$ в вектор

$$\mathbf{Z}_n = Q\mathbf{Y}_n = Q(\mathbf{Y}(x_n) + O(h^{k+1})) = Q\mathbf{Y}(x_n) + O(h^{k+1}) = \mathbf{Z}(x_n) + O(h^{k+1}). \tag{54}$$

Если записать это покомпонентно, то имеем

$$\begin{pmatrix} y_n \\ hy'_n \\ y_{n-1} \\ y_{n-2} \\ \dots \\ y_{n-k+1} \end{pmatrix} \xrightarrow{Q} \begin{pmatrix} y(x_n) \\ hy'(x_n) \\ \dfrac{h^2 y''(x_n)}{2} \\ \dfrac{h^3 y'''(x_n)}{6} \\ \dots \\ \dfrac{h^k y^{(k)}(x_n)}{k!} \end{pmatrix} + O(h^{k+1}). \tag{54$'$}$$

Теперь применим преобразование Q к (46). Обозначим $\mathbf{Z}_n^{(0)} = Q\mathbf{Y}_n^{(0)}$, $\mathbf{Z}_{n-1} = Q\mathbf{Y}_{n-1}$, и пусть $\mathbf{Y}_{n-1} = \mathbf{Y}(x_{n-1})$. Тогда

$$Q\mathbf{Y}_n^{(0)} = QD\mathbf{Y}_{n-1} = QD(Q^{-1}Q)\mathbf{Y}_{n-1} = (QDQ^{-1})Q\mathbf{Y}_{n-1} = (QDQ^{-1})\mathbf{Z}_{n-1}. \tag{55}$$

Учитывая (47), получаем

$$Q\mathbf{Y}_n^{(0)} = Q(\mathbf{Y}(x_n) + O(h^{k+1})) = Q\mathbf{Y}(x_n) + O(h^{k+1}) = \mathbf{Z}(x_n) + O(h^{k+1})$$

или

$$\mathbf{Z}_n^{(0)} = \mathbf{Z}(x_n) + O(h^{k+1}). \tag{56}$$

Следовательно,

$$(QDQ^{-1})\mathbf{Z}_{n-1} = \mathbf{Z}(x_n) + O(h^{k+1}). \tag{57}$$

Воспользуемся (54), заменяя n на $n-1$. Тогда из (57) следует

$$\mathbf{Z}(x_n) = (QDQ^{-1})\mathbf{Z}(x_{n-1}) + O(h^{k+1}). \tag{58}$$

Обозначим $P = QDQ^{-1}$ и распишем (58) покомпонентно:

$$\begin{aligned}
y(x_n) &= p_{11}y(x_{n-1}) + p_{12}hy'(x_{n-1}) + p_{13}\,\frac{h^2y''(x_{n-1})}{2} + \ldots + p_{1,k+1}\,\frac{h^ky^{(k)}(x_{n-1})}{k!} + O(h^{k+1}),\\
hy'(x_n) &= p_{21}y(x_{n-1}) + p_{22}hy'(x_{n-1}) + p_{23}\,\frac{h^2y''(x_{n-1})}{2} + \ldots + p_{2,k+1}\,\frac{h^ky^{(k)}(x_{n-1})}{k!} + O(h^{k+1}),\\
\frac{h^2y''(x_n)}{2} &= p_{31}y(x_{n-1}) + p_{32}hy'(x_{n-1}) + p_{33}\,\frac{h^2y''(x_{n-1})}{2} + \ldots + p_{3,k+1}\,\frac{h^ky^{(k)}(x_{n-1})}{k!} + O(h^{k+1}),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\frac{h^ky^{(k)}(x_n)}{k!} &= p_{k+1,1}y(x_{n-1}) + p_{k+1,2}hy'(x_{n-1}) + p_{k+1,3}\,\frac{h^2y''(x_{n-1})}{2} + \ldots + p_{k+1,k+1}\,\frac{h^ky^{(k)}(x_{n-1})}{k!} +\\
&\quad + O(h^{k+1}).
\end{aligned}$$

Раскладывая левые части этих равенств по формуле Тейлора и приравнивая коэффициенты при одинаковых степенях h слева и справа, получаем, что матрица P является треугольной матрицей Паскаля:

$$P = \begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & \ldots & 1\\
 & 1 & 2 & 3 & 4 & 5 & 6 & \ldots & k\\
 & & 1 & 3 & 6 & 10 & 15 & \ldots & \dfrac{k(k-1)}{2}\\
 & & & 1 & 4 & 10 & 20 & \ldots & .\\
 & & & & 1 & 5 & 15 & \ldots & .\\
 & & & & & 1 & 6 & \ldots & .\\
 & & & & & & 1 & \ldots & .\\
 & 0 & & & & & & \ldots & .\\
 & & & & & & & \ldots & \dfrac{k(k-1)}{2}\\
 & & & & & & & . & k\\
 & & & & & & & & 1
\end{pmatrix}. \tag{59}$$

Обозначим $\mathbf{Z}_n^{(\nu)} = Q\mathbf{Y}_n^{(\nu)}$. Из (51) следует

$$\mathbf{Z}_n^{(\nu)} = Q\mathbf{Y}_n^{(\nu)} = Q(\mathbf{Y}(x_n) + O(h^{k+1})) = Q\mathbf{Y}(x_n) + O(h^{k+1}) = \mathbf{Z}(x_n) + O(h^{k+1}). \tag{60}$$

Применим преобразование Q к (50). Тогда

$$Q\mathbf{Y}_n^{(\nu+1)} = Q\mathbf{Y}_n^{(\nu)} + Q\mathbf{c}F(\mathbf{Y}_n^{(\nu)}) = \mathbf{Z}_n^{(\nu)} + \mathbf{I}\cdot F(\mathbf{Y}_n^{(\nu)}),$$

где $\mathbf{I} = Q\mathbf{c} = (l_0, l_1, \ldots, l_k)^T$. Так как функция невязки $F(\mathbf{Y}_n^{(\nu)})$ зависит только от первых двух компонент вектора $\mathbf{Y}_n^{(\nu)}$, а преобразование Q первые две компоненты не меняет, то эти компоненты вектора $\mathbf{Y}_n^{(\nu)}$ совпадают с первыми двумя компонентами вектора $\mathbf{Z}_n^{(\nu)}$. Поэтому

$$\mathbf{Z}_n^{(\nu+1)} = \mathbf{Z}_n^{(\nu)} + \mathbf{I}\cdot F(\mathbf{Z}_n^{(\nu)}).$$

Из сходимости $\mathbf{Y}_n^{(\nu)}$ к $\mathbf{Y}_n$ следует сходимость $\mathbf{Z}_n^{(\nu)}$ к $\mathbf{Z}_n$. Таким образом, мы приходим к следующему сходящемуся итерационному процессу:

$$\mathbf{Z}_n^{(0)} = P\mathbf{Z}_{n-1}, \tag{61}$$

$$\mathbf{Z}_n^{(\nu+1)} = \mathbf{Z}_n^{(\nu)} + \mathbf{I}\cdot F(\mathbf{Z}_n^{(\nu)}), \qquad \nu = 0, 1, \ldots\,. \tag{62}$$

Из (62) следует, что

$$\mathbf{Z}_n^{(\nu+1)} = \mathbf{Z}_n^{(0)} + \mathbf{I}\Big(F(\mathbf{Z}_n^{(0)}) + F(\mathbf{Z}_n^{(1)}) + \ldots + F(\mathbf{Z}_n^{(\nu)})\Big).$$

Из сходимости (62) вытекает, что $\mathbf{Z}_n^{(\nu+1)}$ сходится к

$$\mathbf{Z}_n = \mathbf{Z}_n^{(0)} + \mathbf{I}\omega, \tag{63}$$

где $\omega = \lim\limits_{\nu\to\infty}\sum\limits_{j=0}^{\nu} F(\mathbf{Z}_n^j)$, причем, как следует из (62),

$$F(\mathbf{Z}_n) = 0. \tag{64}$$

Подставляя (63) в (64), получаем уравнение

$$F(\mathbf{Z}_n^{(0)} + \mathbf{I}\omega) = 0. \tag{65}$$

Решаем уравнение (65) относительно ω методом Ньютона:

$$\omega^{(\nu+1)} = \omega^{(\nu)} - \left(\frac{\partial F(\mathbf{Z}_n^{(0)} + \mathbf{I}\omega^{(\nu)})}{\partial \mathbf{Z}}\cdot\mathbf{I}\right)^{-1} F(\mathbf{Z}_n^{(0)} + \mathbf{I}\omega^{(\nu)}).$$

Умножим обе части равенства на $\mathbf{I}$:

$$\mathbf{I}\omega^{(\nu+1)} = \mathbf{I}\omega^{(\nu)} - \mathbf{I}\left(\frac{\partial F(\mathbf{Z}_n^{(0)} + \mathbf{I}\omega^{(\nu)})}{\partial \mathbf{Z}}\cdot\mathbf{I}\right)^{-1} F(\mathbf{Z}_n^{(0)} + \mathbf{I}\omega^{(\nu)}).$$

Прибавим к обеим частям равенства $\mathbf{Z}_n^{(0)}$ и обозначим $\mathbf{Z}_n^{(\nu)} = \mathbf{Z}_n^{(0)} + \mathbf{I}\omega^{(\nu)}$. Тогда имеем следующий итерационный процесс:

$$\mathbf{Z}_n^{(\nu+1)} = \mathbf{Z}_n^{(\nu)} - \mathbf{I}\left(\frac{\partial F(\mathbf{Z}_n^{(\nu)})}{\partial \mathbf{Z}}\cdot\mathbf{I}\right)^{-1} F(\mathbf{Z}_n^{(\nu)}).$$

Начальное приближение для этого процесса определяется с помощью (61). Из (49) находим

$$\frac{\partial F}{\partial \mathbf{Z}} = \left(\frac{\partial F}{\partial \mathbf{Z}^1}, \frac{\partial F}{\partial \mathbf{Z}^2}, \ldots, \frac{\partial F}{\partial \mathbf{Z}^{k+1}}\right) = \left(h\,\frac{\partial f}{\partial y}, -1,\, 0,\, \ldots, 0\right).$$

Обозначим

$$W = \left(\frac{\partial F(\mathbf{Z}_n^{(\nu)})}{\partial \mathbf{Z}}\cdot\mathbf{I}\right)^{-1} = \left(hl_0\,\frac{\partial f(x_n, y_n^{(\nu)})}{\partial y} - l_1\right)^{-1}.$$

Тогда

$$\mathbf{Z}_n^{(\nu+1)} = \mathbf{Z}_n^{(\nu)} - \mathbf{I}WF(\mathbf{Z}_n^{(\nu)}), \tag{66}$$

или

$$\mathbf{Z}_n^{(\nu+1)} = \mathbf{Z}_n^{(\nu)} - \mathbf{I}\left(hl_0\,\frac{\partial f(x_n, y_n^{(\nu)})}{\partial y} - l_1\right)^{-1}\Big(hf(x_n, y_n^{(\nu)}) - hy_n^{\prime(\nu)}\Big).$$

Таблица 2

Коэффициенты метода Гира $\mathbf{I} = (l_0, l_1, \ldots, l_k)^T$

Коэффициенты	Порядок метода k				
	2	3	4	5	6
l_0	$\frac{2}{3}$	$\frac{6}{11}$	$\frac{12}{25}$	$\frac{60}{137}$	$\frac{20}{49}$
l_1	1	1	1	1	1
l_2	$\frac{1}{3}$	$\frac{6}{11}$	$\frac{7}{10}$	$\frac{225}{274}$	$\frac{58}{63}$
l_3		$\frac{1}{11}$	$\frac{1}{5}$	$\frac{85}{274}$	$\frac{5}{12}$
l_4			$\frac{1}{50}$	$\frac{15}{274}$	$\frac{25}{252}$
l_5				$\frac{1}{274}$	$\frac{1}{84}$
l_6					$\frac{1}{1764}$

Итерационный процесс (66) отличается от итерационного процесса (62) тем, что сложению вектора $\mathbf{I} \cdot F$ с $\mathbf{Z}_n^{(\nu)}$ в (62) предшествует умножение F на матрицу W в (66) и только после этого производится коррекция значения $\mathbf{Z}_n^{(\nu)}$.

Заметим, что если правая часть дифференциального уравнения $f(x, y)$ линейна по y, то итерационный процесс (66) сойдется за одну итерацию.

Компоненты вектора $\mathbf{I} = (l_0, l_1, \ldots, l_k)^T$ для методов разного порядка аппроксимации приведены в табл. 2.

4.2. Оценка погрешности. Автоматический выбор шага и порядка. Первые две компоненты вектора $\mathbf{Y}_n$ равны первым двум компонентам вектора $\mathbf{Z}_n$, так как преобразование Q не меняет первые две компоненты. Следовательно, для оценки погрешности первой компоненты y_n решения $\mathbf{Z}_n$ уравнения (64) можно воспользоваться оценкой (22) локальной погрешности формулы (21)

$$\rho_n = C_{k+1} h^{k+1} y^{(k+1)}(x_n) + O(h^{k+2}), \tag{67}$$

где

$$C_{k+1} = -\frac{\beta_k}{k+1}.$$

Обсудим, как практически воспользоваться оценкой (67). Неизвестной величиной в (67) является $y^{(k+1)}(x_n)$. Если заменим производную $hy^{(k+1)}(x_n)$ разностью назад $\bigtriangledown y^{(k)}(x_n)$, то при этом допустим погрешность порядка $O(h^2)$. Если заменим $h^{k+1}y^{(k+1)}(x_n)$ на разность $\bigtriangledown(h^k y^{(k)}(x_n))$, то допустим ошибку порядка $O(h^{k+2})$:

$$h^{k+1} y^{(k+1)}(x_n) = \nabla(h^k y^{(k)}(x_n)) + O(h^{k+2}).$$

Конечную разность $\nabla(h^k y^{(k)}(x_n))$ можно получить, если из последней компоненты $Z_{n,k+1}$ вектора $\mathbf{Z}_n$ вычесть последнюю компоненту $Z_{n-1,k+1}$ вектора $\mathbf{Z}_{n-1}$ и полученную разность умножить на $k!$. Следовательно,

$$h^{k+1} y^{(k+1)}(x_n) \cong \nabla Z_{n,k+1} \cdot k!$$

В результате имеем

$$\rho_n \cong C_{k+1} \nabla Z_{n,k+1} \cdot k! \tag{68}$$

или, расписывая покомпонентно,

$$\rho_n^i \cong C_{k+1} \nabla Z_{n,k+1}^i \cdot k!, \quad i = 1, 2, \ldots, M.$$

Заметим, что последняя компонента вектора $\mathbf{Z}_{n-1}$ совпадает с последней компонентой вектора $\mathbf{Z}_n^{(0)}$. Величина ρ_n^i — абсолютная погрешность y_n^i. Введем также другие характеристики точности: относительную погрешность

$$\frac{\rho_n^i}{|y_n^i|}, \quad \text{если } y_n^i \neq 0,$$

так называемую *стандартную погрешность*

$$\frac{\rho_n^i}{\max\limits_{0 \le j \le n} |y_j^i|},$$

меру погрешности

$$\mu = \begin{cases} \dfrac{\rho_n^i}{|y_n^i|}, & \text{если} \quad |y_n^i| > p_i, \\ \rho_n^i, & \text{если} \quad |y_n^i| \le p_i. \end{cases}$$

Все перечисленные характеристики точности можно записать в общем виде

$$\frac{\rho_n^i}{P_n^i},$$

где

$$P_n^i = \begin{cases} 1 & \text{для абсолютной погрешности,} \\ |y_n^i| & \text{для относительной погрешности,} \\ \max\limits_{0 \le j \le n} |y_j^i| & \text{для стандартной погрешности,} \\ \textbf{если } |y_n^i| \le p_i, \ \textbf{то } 1, \ \textbf{иначе } |y_n^i| & \text{для меры погрешности.} \end{cases}$$

Будем вести контроль точности по норме

$$\sum_{i=1}^{M} \left(\frac{C_{k+1} \nabla Z_{n,k+1}^i \cdot k!}{P_n^i} \right)^2 = (C_{k+1} \cdot k!)^2 \sum_{i=1}^{M} \left(\frac{\nabla Z_{n,k+1}^i}{P_n^i} \right)^2. \tag{69}$$

Обозначим

$$E = \left(\frac{\varepsilon}{C_{k+1} k!} \right)^2,$$

где ε — заданная точность вычисления приближенного решения, и

$$V = \sum_{i=1}^{M} \left(\frac{\nabla Z_{n,k+1}^i}{P_n^i} \right)^2.$$

Тогда, если

$$V > E, \tag{70}$$

то считается, что на данном шаге метод не достигает требуемой точности и вычисленное значение $\mathbf{Z}_n$ вместе с точкой x_n исключаются из рассмотрения. Выбирается новый размер шага с помощью соотношения (2.154)

$$h^* = \xi h,$$

где величина ξ выбирается так, чтобы на шаге h^* достигалась требуемая точность. Это приводит к следующему значению для ξ:

$$\xi = \left(\frac{E}{V}\right)^{\frac{1}{2(k+1)}}. \tag{71}$$

Формула (71) аналогична формуле (1.97) для одношаговых методов.

Вместо (71) можно взять несколько меньшее значение

$$\xi = \frac{1}{1.2}\left(\frac{E}{V}\right)^{\frac{1}{2(k+1)}}. \tag{72}$$

Теперь вычисление можно повторить по формулам (61), (66), исходя из точки x_{n-1}. Для этого необходимо пересчитать вектор $\mathbf{Z}_{n-1}$. Новые значения компонент этого вектора $\mathbf{Z}^*_{n-1}$ вычисляются по простым формулам

$$Z^*_{n-1,j} = \xi^{j-1} Z_{n-1,j}, \qquad j = 1, \ldots, k+1.$$

Если (70) не выполняется, то считается, что полученное приближение y_n удовлетворяет требуемой точности и значение x_n принимается в качестве текущего узла интегрирования. Дальнейшее интегрирование можно вести, исходя из точки x_n с шагом h^*, который выбирается по формуле (72).

Гир вместе с автоматическим выбором шага решает вопрос об автоматическом выборе порядка точности метода. Предположим, что решение в точке x_n было получено не методом порядка k, а методом порядка на единицу меньше $k-1$. Тогда локальная погрешность решения равна

$$\rho_n \cong C_k h^k y^{(k)}(x_n) + O(h^{k+1}),$$

или

$$\rho_n \cong C_k Z_{n,k+1} \cdot k!$$

Аналогично (69) норма для погрешности запишется в следующем виде:

$$\sum_{i=1}^{M}\left(\frac{C_k Z^i_{n,k+1} \cdot k!}{P^i_n}\right)^2 = (C_k \cdot k!)^2 \sum_{i=1}^{M}\left(\frac{Z^i_{n,k+1}}{P^i_n}\right)^2.$$

Обозначим

$$\breve{E} = \left(\frac{\varepsilon}{C_k \cdot k!}\right)^2, \quad \breve{V} = \sum_{i=1}^{M}\left(\frac{Z^i_{n,k+1}}{P^i_n}\right)^2.$$

Если провести все рассуждения аналогично тому, как это было сделано для метода k-го порядка, то мы получим аналогичное выражение для константы $\breve{\xi}$ изменения шага в методе порядка $k-1$:

$$\breve{\xi} = \left(\frac{\breve{E}}{\breve{V}}\right)^{\frac{1}{2k}}. \tag{73}$$

Вместо (73) можно взять несколько меньшее значение

$$\breve{\xi} = \frac{1}{1.3}\left(\frac{\breve{E}}{\breve{V}}\right)^{\frac{1}{2k}}. \tag{74}$$

Теперь рассмотрим, как следовало бы изменить шаг интегрирования, если бы решение в точке x_n было получено методом порядка на единицу больше. В этом случае погрешность решения была бы равна

$$\rho_n = C_{k+2} h^{k+2} y^{(k+2)}(x_n) + O(h^{k+3}).$$

Выразим $h^{k+2} y^{(k+2)}(x_n)$ через разность

$$h^{k+2} y^{(k+2)}(x_n) = \nabla^2 (h^k y^{(k)}(x_n)) + O(h^{k+3}).$$

Далее,
$$\nabla^2(h^k y^{(k)}(x_n)) \cong \nabla^2(Z_{n,k+1}\cdot k!) = \nabla(\nabla Z_{n,k+1}\cdot k!) = k!\nabla(\nabla Z_{n,k+1}).$$
В результате имеем
$$\rho_n \cong C_{k+2}\nabla(\nabla Z_{n,k+1})k!. \tag{75}$$
Аналогично (69) норма погрешности запишется в виде
$$\sum_{i=1}^{M}\left(\frac{C_{k+2}\nabla(\nabla Z^i_{n,k+1})k!}{P^i_n}\right)^2 = (C_{k+2}\cdot k!)^2\sum_{i=1}^{M}\left(\frac{\nabla(\nabla Z^i_{n,k+1})}{P^i_n}\right)^2.$$
Обозначим
$$\hat{E} = \Big(\frac{\varepsilon}{C_{k+2}\cdot k!}\Big)^2 \qquad \hat{V} = \sum_{i=1}^{M}\left(\frac{\nabla(\nabla Z^i_{n,k+1}}{P^i_n}\right)^2.$$
Аналогично предыдущему приходим к значению константы изменения шага интегрирования в методе порядка $k+1$
$$\hat{\xi} = \Big(\frac{\hat{E}}{\hat{V}}\Big)^{\frac{1}{2(k+2)}}. \tag{76}$$
Вместо (76) можно взять несколько меньшее значение
$$\hat{\xi} = \frac{1}{1.4}\Big(\frac{\hat{E}}{\hat{V}}\Big)^{\frac{1}{2(k+2)}}. \tag{77}$$
Выбором дополнительных множителей в (74), (77) отдается предпочтение методу порядка k или $k-1$.

Для эффективного использования оценки (77) следует при переходе от точки x_{m-1} к точке x_m сохранять значение $\nabla Z_{m-1,k+1}$, которое используется в (75) и $\hat{V}$.

После вычисления ξ, $\breve{\xi}$, $\hat{\xi}$ выбирается тот метод, для которого соответствующая константа изменения шага максимальна.

Преимущество метода Гира по сравнению с другими многошаговыми методами численного интегрирования, рассмотренными в п. 2 данной главы, состоит в легкости изменения шага интегрирования.

Описанная процедура выбора шага и порядка в действительности несколько видоизменяется. Во-первых, если новая предполагаемая длина h^* шага интегрирования увеличивается менее, чем в 1.1 раза по сравнению со старым значением h, то шаг не изменяется. Во-вторых, изменение шага и порядка не допускается в течение $k+1$ шагов после последнего изменения, если только используемое значение шага обеспечивает заданную точность. Причина этого заключается в том, что более частые изменения шага интегрирования могут вызвать дополнительный рост погрешности. В-третьих, если после $k+1$ шагов увеличение шага невозможно, то дальнейшая проверка возможности увеличения шага откладывается до десяти шагов, чтобы сократить накладные расходы, связанные с вычислением констант изменения шага.

Вычисление итерационной матрицы W в (66) включает нахождение матрицы Якоби и обращение матрицы $(\partial F(\mathbf{Z}_n^{(\nu)})/\partial\mathbf{Z})\cdot\mathbf{I}$. Каждая такая операция достаточно трудоемка. Если матрица $\partial f/\partial y$ мало изменяется, то и итерационная матрица мало изменяется от итерации к итерации и даже от одного узла интегрирования к другому. Поэтому применяется модификация метода Ньютона, при которой итерационная матрица W определяется один раз и не перевычисляется ни во время итераций, ни при переходе от точки к точке до тех пор, пока не произойдет изменение шага интегрирования или порядка метода или пока не будет обнаружено, что матрица Якоби существенно изменилась. Об изменении матрицы Якоби судят по количеству выполняемых итераций. Если за три итерации поправка $WF(\mathbf{Z}_n^{(\nu)}) = WF(\mathbf{Z}_n^{(3)})$ не станет достаточно малой, то считается, что матрица Якоби существенно изменилась и итерационная матрица W

вычисляется вновь. Если опять за три итерации указанная точность коррекции не достигается, то шаг h сокращается до значения $h/4$ и вычисление повторяется из точки x_{n-1} с новым шагом.

Если в процессе автоматического выбора порядка метода последний оказался слишком большим, то при известном законе распространения допущенных в ходе численного интегрирования ошибок в y_m на конечные разности функции f_n $(m \leq n)$ старшие производные, входящие в $\mathbf{Z}_n$, становятся лишенными математического смысла. Это приводит к тому, что в одной и той же точке отвергается подряд несколько шагов с последовательно уменьшающейся длиной. К такому же плохому поведению разностей может привести разрыв в какой-нибудь производной решения. В этом случае рекомендуется уменьшить порядок метода настолько, чтобы конечные разности вели себя гладко. В методе Гира принято, что если в одной и той же точке подряд отвергаются три шага, то порядок метода полагается равным единице и вычисления проводятся заново из этой точки.

Метод Гира относится к методам *переменного порядка*. Самый первый шаг выполняется методом первого порядка. Для этого требуется знать вектор $\mathbf{Z}_0 = (y_0, hy_0')^T$. Так как y_0 известно, то может быть вычеслено и $hy_0' = hf(x_0, y_0)$.

Уменьшение порядка на единицу приводит к отбрасыванию последней компоненты $Z_{n,k+1}$ вектора $\mathbf{Z}(x_n)$, равной $h^k y^{(k)}(x_n)/k!$. Увеличение порядка на единицу требует присоединения еще одной компоненты к вектору $\mathbf{Z}(x_n)$, равной $h^{k+1}y^{(k+1)}(x_n)/(k+1)!$. В качестве $h^{k+1}y^{(k+1)}(x_n)$ можно взять конечную разность от $h^k y^{(k)}(x_n)$:

$$h^{k+1}y^{(k+1)}(x_n) \cong \nabla h^k y^{(k)}(x_n).$$

При этом будет допущена ошибка порядка $O(h^{k+2})$. Тогда недостающую компоненту $Z_{n,k+2}$ вектора $\mathbf{Z}_n$ можно найти по формуле

$$Z_{n,k+2} = \frac{\nabla Z_{n,k+1}}{k+1}.$$

5. Экспоненциальный метод. В этом пункте излагаются методы численного интегрирования линейных систем обыкновенных дифференциальных уравнений специального вида, основанные на точном представлении решения в аналитической форме и вычислении матричной экспоненты. Предлагаемые алгоритмы особенно эффективны для решения важного класса систем с большой константой Липшица, в частности для жестких систем уравнений.

5.1. Интегрирование линейных однородных систем с постоянными коэффициентами. Если формулу Рунге-Кутта четвертого порядка точности (1.22) применить к интегрированию линейной однородной системы обыкновенных дифференциальных уравнений первого порядка (1.120), (1.121), то получим следующее выражение для приближенного значения решения на одном шаге интегрирования:

$$y_1 = \left(E + hA + \frac{(hA)^2}{2} + \frac{(hA)^3}{6} + \frac{(hA)^4}{24}\right) y_0. \tag{78}$$

Сравнивая (78) с точным решением задачи (1.120), (1.121)

$$y(x_0 + h) = e^{Ah} y_0, \tag{79}$$

где

$$e^{Ah} = \sum_{k=0}^{\infty} \frac{(hA)^k}{k!}, \tag{80}$$

видим, что приближенное значение решения, получаемое по методу Рунге-Кутта, учитывает только частичную сумму

$$F = E + hA + \frac{(hA)^2}{2} + \frac{(hA)^3}{6} + \frac{(hA)^4}{24} \tag{81}$$

из первых пяти членов ряда (80). Если матрица A имеет большие по модулю собственные числа, то аппроксимация (81) для матричной экспоненты (80) будет грубой для больших значений шага h. Следовательно, счет по формуле Рунге-Кутта приведет к численной неустойчивости. Если метод Рунге-Кутта применять при малых h, то он потребует огромных затрат машинного времени на больших промежутках интегрирования, причем с увеличением длительности счета вследствие большого числа округлений накапливается вычислительная погрешность.

В этом случае для определения решения задачи (1.120), (1.121) в заданной точке $x_n = x_0+nh$ целесообразно организовать вычисления следующим образом. Сначала аппроксимировать при малом h матричную экспоненту (80) выражением (81); затем, используя алгоритм быстрого умножения, вычислить решение

$$y_n = \left(E + hA + \frac{(hA)^2}{2} + \frac{(hA)^3}{6} + \frac{(hA)^4}{24}\right)^n y_0 = F^n y_0$$

в заданной точке x_n. Затраты машинного времени будут порядка $log_2 n$, что существенно меньше времени, затрачиваемого методом Рунге-Кутта, которое пропорционально n.

Рассмотрим аппроксимацию матричной экспоненты более общего вида и определим

$$y_n = F^n y_0 = \left(\sum_{k=0}^{S} \frac{(Ah)^k}{k!}\right)^n y_0. \tag{82}$$

Покажем, что можно выбрать S таким, чтобы получить решение системы (1.120), (1.121) на одном шаге h с заданной точностью ε.

Положим

$$e^{Ah} = F + R$$

и рассмотрим соотношения

$$y = e^{Ah} y_0 \quad \text{и} \quad y_F = F y_0.$$

Введем относительную ошибку δ для вектора y_F:

$$\delta = \frac{\|y - y_F\|}{\|y_F\|} = \frac{\|y_R\|}{\|y_F\|}. \tag{83}$$

S выбираем так, чтобы выполнялось условие $\delta \leq \varepsilon$, которое означает, что на фиксированной сетке узлов мы применяем такой метод (82) численного интегрирования, для которого относительная погрешность δ приближенного решения на одном шаге не превосходит наперед заданной величины ε.

В дальнейшем под нормой матрицы понимаем норму, согласованную с нормой вектора. Для $\|R\|$ имеем следующую оценку:

$$\|R\| = \|e^{Ah} - F\| = \left\|\sum_{k=S+1}^{\infty} \frac{(Ah)^k}{k!}\right\| \leq \frac{(\|A\|h)^{S+1}}{(S+1)!} \sum_{k=0}^{\infty} \frac{(S+1)!(S+2)^k p^k}{(k+S+1)!}, \tag{84}$$

где

$$p = \frac{\|A\|h}{S+2}.$$

Пусть S_0 — наименьшее значение S такое, что

$$p < 1, \tag{85}$$

и положим $S \geq S_0$. Тогда сумма в неравенстве (84) ограничена величиной

$$\sum_{k=0}^{\infty} p^k = \frac{1}{1-p},$$

так что
$$\|R\| \le \frac{(\|A\|h)^{S+1}}{(S+1)!} \cdot \frac{1}{1-p}\,. \tag{86}$$
Далее,
$$F = e^{Ah} - R = e^{Ah}(E - e^{-Ah}R).$$
Предположим, что
$$e^{\|A\|h}\,\|R\| < 1. \tag{87}$$
Тогда для
$$F^{-1} = (E - e^{-Ah}R)^{-1}e^{-Ah} = \sum_{k=0}^{\infty}(e^{-Ah}R)^k e^{-Ah}$$
справедлива следующая оценка:
$$\|F^{-1}\| \le \sum_{k=0}^{\infty}(e^{\|A\|h}\|R\|)^k\, e^{\|A\|h} = \frac{e^{\|A\|h}}{1 - \|R\|e^{\|A\|h}}. \tag{88}$$

Используя неравенства
$$\|y_R\| \le \|R\|\,\|y_0\|$$
и
$$\|y_F\| \ge \|F^{-1}\|^{-1}\|y_0\|,$$
получаем из (83)
$$\delta \le \|R\|\,\|F^{-1}\|. \tag{89}$$
Подставляя (88) в (89), имеем
$$\delta \le \frac{\|R\|e^{\|A\|h}}{1 - \|R\|e^{\|A\|h}}.$$
Так как в правой части неравенства (86) стоит убывающая к нулю функция от S, то всегда найдется такое значение $S \ge S_0$, что будут иметь место неравенство
$$q = e^{\|A\|h}\,\frac{(\|A\|h)^{S+1}}{(S+1)!} \cdot \frac{1}{1-p} < 1$$
и оценка
$$\delta < \frac{q}{1-q} < \varepsilon. \tag{90}$$

Покажем теперь, что если выполняется условие
$$\|F\|\,(1+\varepsilon) < 1,$$
то
$$\lim_{n\to\infty} y_n = \lim_{n\to\infty} y(x_n) = 0. \tag{91}$$
В самом деле,
$$\|y(x_n) - y_n\| \le \|(e^{Ah})^n - F^n\| \cdot \|y_0\| \le ((\|F\| + \|R\|)^n - \|F\|^n)\,\|y_0\|.$$
Далее,
$$(\|F\| + \|R\|)^n \le \left(\|F\| + \frac{\varepsilon}{\|F^{-1}\|}\right)^n \le (\|F\|\,(1+\varepsilon))^n.$$
Из последнего неравенства и неравенства
$$\|y_n\| \le \|F\|^n\|y_0\|$$

вытекает справедливость нашего утверждения. Отсюда следует устойчивость алгоритма (82) для решения задачи (1.120), (1.121) при постоянном h и $n \to \infty$.

Заметим, что матричная экспонента (80) является значением при $x = h$ фундаментальной матрицы решений e^{Ax} системы (1.120), нормированной в нуле. Поэтому ее нахождение можно свести к численному интегрированию системы (1.120) с начальными условиями

$$y^i(0) = 1,$$
$$y^j(0) = 0, \quad j \neq i, \quad i = 1, 2, \ldots, M,$$

на отрезке $[0, h]$. Данный метод вычисления матрицы e^{Ah} состоит, следовательно, в M-кратном решении задачи Коши для системы (1.120) на малом отрезке интегрирования. Поэтому все эти задачи Коши могут быть решены быстро с требуемой точностью каким-нибудь методом численного интегрирования из рассмотренных в гл. 1, 2.

5.2. Интегрирование линейных однородных систем с переменными коэффициентами. Для линейной однородной системы с переменными коэффициентами

$$\begin{cases} y'(x) = A(x)y(x), \\ y(x_0) = y_0 \end{cases} \tag{92}$$

аппроксимируем матрицу $A(x)$ кусочно-постоянной матрицей. Для этого возьмем некоторое разбиение интервала интегрирования

$$x_0 < x_1 < x_2 < \ldots < x_l = x_0 + X$$

и на каждом сегменте $[x_i, x_{i+1}]$ будем решать задачу Коши для линейной системы с постоянными коэффициентами

$$\frac{dy^{(i)}(x)}{dx} = \overline{A}^{(i)} y^{(i)}(x), \tag{93}$$

$$y^{(i)}(x_i) = y^{(i-1)}(x_i), \tag{94}$$

где $\overline{A}^{(i)} = A(\tau),\ \tau \in [x_i, x_{i+1}],\ i = 0, 1, \ldots, l-1,\ y^{(0)}(x_0) = y_0$. Функции $y^{(i)}(x)$ примем в качестве приближенного решения на $[x_i, x_{i+1}]$ исходной задачи (92).

Найдем уклонение $y^{(0)}(x)$ от решения исходной задачи $y(x)$ в узле $x_1 = x_0 + T$. Положим

$$z(x) = y(x) - y^{(0)}(x), \quad x_0 \leq x \leq x_1.$$

Функция $z(x)$ удовлетворяет дифференциальному уравнению

$$z'(x) = A(x)z(x) + ((A(x) - A(\tau))\, y^{(0)}(x) \tag{95}$$

и начальному условию

$$z(x_0) = 0. \tag{96}$$

Решение (95), (96) можно представить в виде

$$z(x) = \int\limits_{x_0}^{x} \Omega(\xi, x)(A(\xi) - A(\tau))\, y^{(0)}(\xi)\, d\xi, \qquad x_0 \leq x \leq x_1, \tag{97}$$

где $\Omega(\xi, x)$ — матрициант матрицы $A(x)$. Предположим, что матрица $A(x)$ трижды непрерывно дифференцируема. Тогда выражение (97) для $z(x)$ можно преобразовать так:

$$\begin{aligned} z(x) &= \int\limits_{x_0}^{x} \Omega(\xi, x)\left(A'(\tau)(\xi - \tau) + A''(\tau)\,\frac{(\xi - \tau)^2}{2}\right) y^{(0)}(\xi)\, d\xi + O(T^4) = \\ &= v(x) + O(T^4), \end{aligned} \tag{98}$$

где $v(x)$ — решение следующей задачи Коши:

$$\begin{cases} v'(x) = A(x)v(x) + \Big(A'(\tau)(x-\tau) + A''(\tau)\,\dfrac{(x-\tau)^2}{2}\Big)\, y^{(0)}(x), \quad x_0 \le x \le x_1, & (99) \\ v(x_0) = 0. & (100) \end{cases}$$

Разложив $v(x_1)$ по формуле Тейлора, имеем

$$z(x_1) = Tv'(x_0) + \frac{T^2}{2}\, v''(x_0) + \frac{T^3}{6}\, v'''(x_0) + O(T^4). \tag{101}$$

Далее, последовательно дифференцируя (99) и учитывая (100), находим

$$\begin{aligned} v'(x_0) &= Dy_0, \\ v''(x_0) &= A(x_0)\, Dy_0 + D'y_0 + DA(\tau)y_0, \\ v'''(x_0) &= 2A'(x_0)\, Dy_0 + A^2(x_0)\, Dy_0 + A(x_0)\, D'y_0 + A(x_0)\, DA(\tau)y_0 + \\ &\quad + A''(\tau)y_0 + 2D'A(\tau)y_0 + DA^2(\tau)y_0, \end{aligned}$$

где

$$\begin{aligned} D &= A'(\tau)(x_0-\tau) + A''(\tau)\,\frac{(x_0-\tau)^2}{2}, \\ D' &= A'(\tau) + A''(\tau)(x_0-\tau). \end{aligned}$$

Подставляя найденные производные в (101) и группируя члены одного порядка малости, получаем

$$\begin{aligned} z(x_1) = A'(\tau)\Big((x_0-\tau)T + \frac{T^2}{2}\Big)\, y_0 + A''(\tau)\Big(\frac{(x_0-\tau)^2}{2}T + \frac{T^3}{6} + (x_0-\tau)\frac{T^2}{2}\Big)\, y_0 + \\ + A(x_0)A'(\tau)\Big((x_0-\tau)\frac{T^2}{2} + \frac{T^3}{6}\Big)\, y_0 + A'(\tau)A(\tau)\Big((x_0-\tau)\frac{T^2}{2} + 2\cdot\frac{T^3}{6}\Big)\, y_0 + O(T^4). \end{aligned}$$

Полагая $\tau = x_0 + \dfrac{T}{2}$, приходим к следующему выражению для $z(x_1)$:

$$z(x_1) = \frac{1}{24}\Big(A''(\tau) - 2(A(x_0)A'(\tau) - A'(\tau)A(\tau))\Big)\, T^3 y_0 + O(T^4),$$

которое можно записать также в виде

$$y(x_1) - y^{(0)}(x_1) = \frac{1}{24}\Big(A''(\tau) - 2(A(\tau)A'(\tau) - A'(\tau)A(\tau))\Big)\, T^3 y_0 + O(T^4). \tag{102}$$

Отметим, что задачи Коши (93), (94) с постоянными матрицами можно решать описанными в п. 5.1 методами.

5.3. Интегрирование линейных неоднородных систем со специальной правой частью. Решение задачи Коши для линейной неоднородной системы с постоянными коэффициентами

$$\begin{cases} y' = Ay + f(x), \\ y(x_0) = y_0 \end{cases} \tag{103}$$

может быть представлено следующим образом:

$$y(x) = e^{A(x-x_0)}y_0 + \int\limits_{x_0}^{x} e^{A(x-\tau)} f(\tau)\, d\tau. \tag{104}$$

Преобразуем это представление в предположении, что свободный член $f(\tau)$ задан в виде

$$f(\tau) = (\alpha_1 e^{\beta_1 \tau}, \alpha_2 e^{\beta_2 \tau}, \dots, \alpha_M e^{\beta_M \tau})^T = \sum_{i=1}^{M} \alpha_i e^{\beta_i \tau} \mathbf{e}_i, \tag{105}$$

вектор $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0)^T$ имеет все компоненты, равные 0, кроме i-й, которая равна 1 (i-й орт). Подставив (105) в (104) и воспользовавшись соотношением

$$E e^{\beta_i \tau} = e^{\beta_i E \tau},$$

имеем

$$y(x) = e^{A(x-x_0)} y_0 + \sum_{i=1}^{M} \int_{x_0}^{x} e^{A(x-\tau)} \alpha_i e^{\beta_i \tau} \, d\tau \cdot \mathbf{e}_i =$$
$$= e^{A(x-x_0)} y_0 + \sum_{i=1}^{M} \alpha_i \int_{x_0}^{x} e^{A(x-\tau) + \beta_i E \tau} \, d\tau \cdot \mathbf{e}_i.$$

Положим $T = x - x_0$, тогда

$$y(x) = y(x_0 + T) = e^{AT} y_0 + \sum_{i=1}^{M} \alpha_i \int_{0}^{T} e^{A(T-t) + \beta_i E t + \beta_i E x_0} \, dt \cdot \mathbf{e}_i =$$
$$= e^{AT} y_0 + \sum_{i=1}^{M} \alpha_i e^{\beta_i (x_0 + T)} \int_{0}^{T} e^{(A - \beta_i E)(T-t)} \, dt \cdot \mathbf{e}_i.$$

Вводя обозначения

$$\Phi_i(T) = \int_{0}^{T} e^{\tilde{A}(T-t)} \, dt = \int_{0}^{T} e^{(A - \beta_i E)(T-t)} \, dt$$

и

$$g_i(T) = \Phi_i(T) \, \mathbf{e}_i,$$

получаем следующее выражение для $y(x)$:

$$y(x) = y(x_0 + T) = e^{AT} y_0 + \sum_{i=1}^{M} \alpha_i e^{\beta_i (x_0 + T)} \, g_i(T). \tag{106}$$

Матрицы $e^{AT}, \Phi_i(T)$ и векторы $g_i(T)$ удовлетворяют рекуррентным соотношениям

$$e^{AT} = e^{AT/2} \, e^{AT/2}, \tag{107}$$

$$\Phi_i(T) = \Phi_i(T/2)(E + e^{\tilde{A}T/2}), \tag{108}$$

$$g_i(T) = (E + e^{\tilde{A}T/2}) \, g_i(T/2), \tag{109}$$

где

$$e^{\tilde{A}T/2} = e^{AT/2} \, e^{-\beta_i T/2}$$

и, кроме того,

$$e^{AT} = \sum_{k=0}^{\infty} \frac{(AT)^k}{k!}, \tag{110}$$

$$\Phi_i(T) = \sum_{k=1}^{\infty} \frac{\tilde{A}^{k-1} T^k}{k!}. \tag{111}$$

Покажем, например, справедливость формулы (108):

$$\Phi_i(T) = \int\limits_0^T e^{\tilde{A}(T-t)}\,dt = \int\limits_0^{T/2} e^{\tilde{A}(T-t)}\,dt + \int\limits_{T/2}^{T} e^{\tilde{A}(T-t)}\,dt =$$
$$= \int\limits_0^{T/2} e^{\tilde{A}(T-t)}\,dt + \int\limits_0^{T/2} e^{\tilde{A}(T/2-\tau)}\,d\tau = \int\limits_0^{T/2} e^{\tilde{A}(T/2-t)+\tilde{A}T/2}\,dt + \int\limits_0^{T/2} e^{\tilde{A}(T/2-t)}\,dt =$$
$$= \int\limits_0^{T/2} e^{\tilde{A}(T/2-t)}\,dt \cdot (e^{\tilde{A}T/2} + E) = \Phi_i(T/2)(E + e^{\tilde{A}T/2}).$$

Из (108) и перестановочности матриц $\Phi_i(T/2)$ и $E + e^{\tilde{A}T/2}$ при умножении вытекает соотношение (109):

$$g_i(T) = \Phi_i(T)\,\mathbf{e}_i = \Phi_i(T/2)(E + e^{\tilde{A}T/2})\,\mathbf{e}_i = (E + e^{\tilde{A}T/2})\,\Phi_i(T/2)\,\mathbf{e}_i =$$
$$= (E + e^{\tilde{A}T/2})\,g_i(T/2).$$

Решение задачи (103) может быть найдено, если известны матрица e^{AT} и векторы $g_i(T)$. А эти матрица и векторы могут быть вычислены на основе последовательного применения рекуррентных соотношений (107), (109) и вычисления e^{Ah}, $\quad g_i(h)$ при достаточно малом $h = T/2^N$ с помощью частичных сумм рядов (110), (111):

$$e^{Ah} = \sum_{k=0}^{r} \frac{(Ah)^k}{k!}, \quad \Phi_i(h) = \sum_{k=1}^{s} \frac{\tilde{A}^{k-1}h^k}{k!} = h\sum_{k=1}^{s} \frac{\tilde{A}^{k-1}h^{k-1}}{k!} = h\sum_{i=0}^{s-1} \frac{\tilde{A}^i h^i}{(i+1)!}$$

и

$$g_i(h) = \Phi_i(h)\,\mathbf{e}_i.$$

Далее решение задачи (103) вычисляется по формуле (106), которую можно представить в виде

$$y(x_{n+1}) = y(x_n + T) = e^{AT}y(x_n) + \sum_{i=1}^{M} \alpha_i e^{\beta_i(x_n+T)}\,g_i(T), \quad n = 0, 1, \ldots. \tag{112}$$

Приведем еще один алгоритм вычисления решения задачи (103), (105). Предполагая, что обратные матрицы существуют и могут быть эффективно определены, интеграл $\Phi_i(T)$ можно представить следующим образом:

$$\Phi_i(T) = -(A - \beta_i E)^{-1} e^{(A-\beta_i E)(T-t)}\Big|_0^T = (A - \beta_i E)^{-1}\,(e^{(A-\beta_i E)T} - E).$$

Тогда формулу (112) можно записать так:

$$\begin{aligned} y(x_{n+1}) = y(x_n + T) &= e^{AT}y(x_n) + \sum_{i=1}^{M} \alpha_i e^{\beta_i(x_n+T)}(A - \beta_i E)^{-1}(e^{(A-\beta_i E)T} - E)\,\mathbf{e}_i = \\ &= e^{AT}y(x_n) + \sum_{i=1}^{M} \alpha_i (A - \beta_i E)^{-1} e^{\beta_i x_n}(e^{AT} - e^{\beta_i T}E)\,\mathbf{e}_i. \end{aligned} \tag{113}$$

Если каждая компонента вектор-функции $f(\tau)$ является суммой экспонент:

$$f(\tau) = \sum_{i=1}^{M}\sum_{j=1}^{p} \alpha_{ij} e^{\beta_{ij}\tau}\mathbf{e}_i, \tag{114}$$

то в (112) должна быть двойная сумма и формула (113) принимает вид

$$y(x_{n+1}) = y(x_n + T) = e^{AT} y(x_n) + \sum_{i=1}^{M} \sum_{j=1}^{p} \alpha_{ij} (A - \beta_{ij} E)^{-1} \times$$
$$\times\, e^{\beta_{ij} x_n} (e^{AT} - e^{\beta_{ij} T} E)\, \mathbf{e}_i.$$

Аналогично записывается формула (112).

Приведенные в данной главе методы применялись в НИВЦ МГУ для решения ряда задач вычислительного эксперимента в физике и технике.

Так, алгоритмы п. 5.1, 5.2 были использованы в физической кинетике для численного исследования динамики заселения энергетических уровней атома и иона гелия. Математическая модель соответствующих переходных процессов сводится к линейной системе с большим числом уравнений и с большими по модулю отрицательными собственными числами матрицы системы. Попытка решать такие задачи традиционными численными методами Рунге-Кутта и Адамса приводит к необходимости выбора шага интегрирования, много меньшего продолжительности интервала наблюдения решения, и значительному возрастанию машинного времени. Использование методов, основанных на вычислении матричной экспоненты, позволяет эффективно решать эти задачи, причем весь процесс решения укладывается в отладочное время.

Успешным оказался этот метод и для численного исследования надежности некоторых систем управления, моделируемых с помощью марковских случайных процессов с дискретными состояниями и непрерывным временем и описываемых дифференциальными уравнениями Колмогорова для вероятностей состояний.

Рассмотренный в п. 5.3 метод применялся в электротехнике при расчете наведенных токов в электрических цепях нагрузки кабельных линий связи при воздействии на них одиночных импульсов грозового разряда. Соответствующая неоднородная система дифференциальных уравнений характеризуется большим разбросом и наличием комплексных собственных значений и имеет быстро затухающие и сильно осциллирующие компоненты решения. Возмущающая функция имеет вид (114) и каждая ее ненулевая составляющая является суммой нескольких экспонент с различными коэффициентами α_{ij} и показателями степеней β_{ij}. Решение этой задачи вычислялось с использованием алгоритма с обратными матрицами с удвоенным числом значащих цифр на ЭВМ БЭСМ-6. Решение также было проведено за отладочное время.

6. Неявные методы Рунге—Кутта. Важным классом одношаговых методов решения жестких задач являются неявные методы Рунге—Кутта, имеющие следующий вид:

$$y_1 = y_0 + p_1 k_1(h) + p_2 k_2(h) + \ldots + p_q k_q(h), \tag{115}$$

где

$$\begin{aligned}
k_1(h) &= hf(x_0 + \alpha_1 h, \quad y_0 + \beta_{11} k_1(h) + \beta_{12} k_2(h) + \ldots + \beta_{1q} k_q(h)),\\
k_2(h) &= hf(x_0 + \alpha_2 h, \quad y_0 + \beta_{21} k_1(h) + \beta_{22} k_2(h) + \ldots + \beta_{2q} k_q(h)),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
k_q(h) &= hf(x_0 + \alpha_q h, \quad y_0 + \beta_{q1} k_1(h) + \beta_{q2} k_2(h) + \ldots + \beta_{qq} k_q(h)).
\end{aligned} \tag{116}$$

Числа α_i, β_{ij}, p_i выбираются так, чтобы разложение выражения (115) по степеням h совпадало с разложением (1.6) до некоторой степени h^s включительно. Проилллюстрируем вывод неявных формул Рунге—Кутта на примере двучленной формулы

$$\begin{aligned}
y_1 &= y_0 + p_1 k_1(h) + p_2 k_2(h),\\
k_1(h) &= hf(x_0 + \alpha_1 h, y_0 + \beta_{11} k_1(h) + \beta_{12} k_2(h)),\\
k_2(h) &= hf(x_0 + \alpha_2 h, y_0 + \beta_{21} k_1(h) + \beta_{22} k_2(h)).
\end{aligned} \tag{117}$$

Находим первую производную

$$k_1'(h) = f(x_0 + \alpha_1 h, y_0 + \beta_{11} k_1 + \beta_{12} k_2) + h(f_x' \alpha_1 + f_y'(\beta_{11} k_1' + \beta_{12} k_2')).$$

Находим вторую производную

$$k_1''(h) = 2(f_x' \alpha_1 + f_y'(\beta_{11} k_1' + \beta_{12} k_2')) + h\psi(h),$$

где

$$\psi(h) = f_{xx}'' \alpha_1^2 + 2 f_{xy}''(\beta_{11} k_1' + \beta_{12} k_2')\,\alpha_1 + f_{yy}''(\beta_{11} k_1' + \beta_{12} k_2')^2 + f_y'(\beta_{11} k_1'' + \beta_{12} k_2'').$$

Находим третью производную

$$k_1'''(h) = 3\psi(h) + h\psi'(h).$$

Аналогично вычисляются производные $k_2'(h), k_2''(h), k_2'''(h)$. Подставляем $h = 0$ и получаем

$$\begin{aligned}
k_1(0) &= 0, \quad k_2(0) = 0,\\
k_1'(0) &= f(x_0, y_0) = f_0, \quad k_2'(0) = f_0,\\
k_1''(0) &= 2(f_x'\alpha_1 + \beta_{11} f_y' f_0 + \beta_{12} f_y' f_0) = 2(f_x'\alpha_1 + f_y' f_0(\beta_{11} + \beta_{12})),\\
k_2''(0) &= 2(f_x'\alpha_2 + f_y' f_0(\beta_{21} + \beta_{22})),\\
k_1'''(0) &= 3(f_{xx}''\alpha_1^2 + 2 f_{xy}'' f_0(\beta_{11} + \beta_{12})\alpha_1 + f_{yy}'' f_0^2(\beta_{11} + \beta_{12})^2 + 2 f_y' f_x' \beta_{11}\alpha_1 +\\
&\quad + 2 f_y'^2 f_0 \beta_{11}(\beta_{11} + \beta_{12}) + 2 f_y' f_x' \alpha_2 \beta_{12} + 2 f_y'^2 f_0 \beta_{12}(\beta_{21} + \beta_{22})),\\
k_2'''(0) &= 3(f_{xx}''\alpha_2^2 + 2 f_{xy}'' f_0(\beta_{21} + \beta_{22})\alpha_2 + f_{yy}'' f_0^2(\beta_{21} + \beta_{22})^2 + 2 f_y' f_x' \beta_{21}\alpha_1 +\\
&\quad + 2 f_y'^2 f_0 \beta_{21}(\beta_{11} + \beta_{12}) + 2 f_y' f_x' \alpha_2 \beta_{22} + 2 f_y'^2 f_0 \beta_{22}(\beta_{21} + \beta_{22})).
\end{aligned}$$

Найденные производные подставляются в разложение

$$\begin{aligned}
y_1 = y_0 + p_1 k_1(h) + p_2 k_2(h) = y_0 + (p_1 k_1'(0) + p_2 k_2'(0))\,h +\\
+ (p_1 k_1''(0) + p_2 k_2''(0))\,\frac{h^2}{2} + (p_1 k_1'''(0) + p_2 k_2'''(0))\,\frac{h^3}{6} + \ldots .
\end{aligned} \tag{118}$$

Коэффициент при h в разложении (118) имеет вид $(p_1 + p_2) f_0$. Приравнивая его к коэффициенту при h в разложении (1.6), получаем уравнение

$$p_1 + p_2 = 1. \tag{119}$$

Приравнивая коэффициенты при подобных членах, содержащих h^2 в разложениях (118) и (1.6), получаем уравнения

$$p_1\alpha_1 + p_2\alpha_2 = \frac{1}{2}, \tag{120}$$

$$p_1\,(\beta_{11} + \beta_{12}) + p_2\,(\beta_{21} + \beta_{22}) = \frac{1}{2}. \tag{121}$$

Приравнивая коэффициенты при подобных членах, содержащих h^3 в разложениях (118) и (1.6), получаем уравнения

$$3 p_1 \alpha_1^2 + 3 p_2 \alpha_2^2 = 1, \tag{122}$$

$$3 p_1 \alpha_1 (\beta_{11} + \beta_{12}) + 3 p_2 \alpha_2 (\beta_{21} + \beta_{22}) = 1, \tag{123}$$

$$3 p_1 (\beta_{11} + \beta_{12})^2 + 3 p_2 (\beta_{21} + \beta_{22})^2 = 1, \tag{124}$$

$$6 p_1 (\alpha_1 \beta_{11} + \alpha_2 \beta_{12}) + 6 p_2 (\alpha_1 \beta_{21} + \alpha_2 \beta_{22}) = 1, \tag{125}$$

$$6p_1(\beta_{11}(\beta_{11}+\beta_{12})+\beta_{12}(\beta_{21}+\beta_{22}))+6p_2(\beta_{21}(\beta_{11}+\beta_{12})+\beta_{22}(\beta_{21}+\beta_{22}))=1. \tag{126}$$

Восемь уравнений (119) — (126) обеспечивают локальную погрешность формулы (117) порядка $O(h^4)$.

Положим

$$\begin{aligned} \alpha_1 &= \beta_{11}+\beta_{12}, \\ \alpha_2 &= \beta_{21}+\beta_{22}. \end{aligned} \tag{127}$$

Тогда уравнение (121) совпадает с уравнением (120), уравнения (124) и (123) совпадут с уравнением (122), а уравнение (126) — с уравнением (125). Вместо восьми уравнений будем рассматривать четыре уравнения (119), (120), (122), (125). Из уравнения (119) имеем

$$p_1 = 1 - p_2.$$

Подставляем это выражение в (120):

$$\alpha_1 - \alpha_1 p_2 + p_2\alpha_2 = \frac{1}{2},$$

или

$$p_2(\alpha_2-\alpha_1) = \frac{1}{2} - \alpha_1. \tag{128}$$

Уравнение (122) запишется так:

$$(1-p_2)\alpha_1^2 + p_2\alpha_2^2 = \frac{1}{3},$$

или

$$p_2(\alpha_2^2-\alpha_1^2) = \frac{1}{3} - \alpha_1^2. \tag{129}$$

Поделив (129) на (128), имеем

$$\alpha_2 + \alpha_1 = \frac{\frac{1}{3}-\alpha_1^2}{\frac{1}{2}-\alpha_1},$$

откуда

$$\alpha_2 = -\alpha_1 + \frac{\frac{1}{3}-\alpha_1^2}{\frac{1}{2}-\alpha_1}.$$

Найдем решение уравнений (119), (120), (122), (125) в нескольких частных случаях. Пусть $\beta_{11}=\beta_{22}$, $\beta_{12}=0$. Тогда $\beta_{11}=\alpha_1$. Положим $\alpha_1 = \dfrac{3+\sqrt{3}}{6}$. Тогда $\alpha_2 = \dfrac{3-\sqrt{3}}{6}$, $p_1=\dfrac{1}{2}$, $p_2=\dfrac{1}{2}$. В результате приходим к неявной формуле

$$\begin{aligned} y_1 &= y_0 + \frac{1}{2}\,k_1 + \frac{1}{2}\,k_2, \\ k_1 &= hf\biggl(x_0 + \frac{3+\sqrt{3}}{6}\,h,\ y_0 + \frac{3+\sqrt{3}}{6}\,k_1\biggr), \\ k_2 &= hf\biggl(x_0 + \frac{3-\sqrt{3}}{6}\,h,\ y_0 - \frac{\sqrt{3}}{3}\,k_1 + \frac{3+\sqrt{3}}{6}\,k_2\biggr). \end{aligned} \tag{130}$$

Полученная формула имеет локальную погрешность $O(h^4)$. Следовательно, это метод третьего порядка точности.

Применим формулу (130) к линейной системе (1.120), (1.121). Формула для k_1 принимает вид

$$k_1 = hA(y_0+\alpha_1 k_1) = hAy_0 + \alpha_1 hAk_1.$$

Отсюда находим
$$k_1 = (E - \alpha_1 hA)^{-1} hAy_0.$$
Формула для k_2 принимает вид
$$k_2 = hA\Big(y_0 + (\alpha_2 - \alpha_1)(E - \alpha_1 hA)^{-1} hAy_0\Big) + \alpha_1 hAk_2.$$
Отсюда находим
$$k_2 = (E - \alpha_1 hA)^{-1} hA\Big(E + (\alpha_2 - \alpha_1)(E - \alpha_1 hA)^{-1} hA\Big)y_0.$$
Подставляем найденные выражения для k_1 и k_2 в (130):
$$\begin{aligned} y_1 &= \Big(E + \frac{1}{2}(E - \alpha_1 hA)^{-1} hA + \frac{1}{2}(E - \alpha_1 hA)^{-1} hA\Big(E + (\alpha_2 - \alpha_1)(E - \alpha_1 hA)^{-1} hA\Big)\Big)y_0 = \\ &= \Big(E + (E - \alpha_1 hA)^{-1} hA + \frac{1}{2}(\alpha_2 - \alpha_1)((E - \alpha_1 hA)^{-1})^2 h^2 A^2\Big)y_0 = \\ &= ((E - \alpha_1 hA)^{-1})^2\Big((E - \alpha_1 hA)^2 + (E - \alpha_1 hA)hA + \frac{1}{2}(\alpha_2 - \alpha_1)h^2 A^2\Big)y_0 = \\ &= (E - 2\alpha_1 hA + \alpha_1^2 h^2 A^2)^{-1}\Big(E + (1 - 2\alpha_1)hA + (\alpha_1^2 - \frac{3}{2}\,\alpha_1 + \frac{1}{2}\,\alpha_2)h^2 A^2\Big)y_0. \end{aligned}$$
Приходим к соотношению
$$y_1 = F(Ah)y_0, \tag{131}$$
где
$$F(Ah) = \left(E - \frac{3+\sqrt{3}}{3}\,Ah + \frac{2+\sqrt{3}}{6}\,A^2h^2\right)^{-1}\left(E - \frac{\sqrt{3}}{3}\,Ah - \frac{1+\sqrt{3}}{6}\,A^2h^2\right). \tag{132}$$
Из соотношения (131) получаем
$$y_{n+1} = F(Ah)y_n. \tag{133}$$

Рассмотрим неявную формулу Рунге—Кутта (117) второго порядка. В этом случае достаточно удовлетворить уравнения (119) и (120). Пусть $\beta_{11} = \beta_{12} = 0$. Тогда $\alpha_1 = 0$. Положим $\beta_{22} = \beta$ и рассмотрим неявную формулу вида
$$\begin{aligned} y_1 &= y_0 + p_1 k_1 + p_2 k_2, \\ k_1 &= hf(x_0, y_0), \\ k_2 &= hf(x_0 + \alpha_2 h, y_0 + (\alpha_2 - \beta)\, k_1 + \beta k_2). \end{aligned} \tag{134}$$
Применим формулу (134) к линейной системе (1.120), (1.121). Формула для k_1 принимает вид
$$k_1 = Ahy_0.$$
Формула для k_2 принимает вид
$$k_2 = Ahy_0 + (\alpha_2 - \beta)A^2h^2y_0 + \beta Ahk_2.$$
Отсюда находим
$$k_2 = (E - \beta Ah)^{-1}(Ah + (\alpha_2 - \beta)A^2h^2)\, y_0.$$
Подставляем найденные выражения для k_1 и k_2 в (134):
$$\begin{aligned} y_1 &= \Big(E + p_1 Ah + p_2(E - \beta Ah)^{-1}(Ah + (\alpha_2 - \beta)A^2h^2)\Big)\, y_0 = \\ &= (E - \beta Ah)^{-1}\Big((E - \beta Ah)(E + p_1 Ah) + p_2(Ah + (\alpha_2 - \beta)\, A^2h^2)\Big)\, y_0 = \\ &= (E - \beta Ah)^{-1}(E + (p_1 - \beta)\, Ah - \beta p_1 A^2h^2 + p_2 Ah + p_2(\alpha_2 - \beta)\, A^2h^2)\, y_0 = \\ &= (E - \beta Ah)^{-1}\Big(E + (p_1 + p_2 - \beta)\, Ah + (p_2\alpha_2 - \beta(p_1 + p_2))\, A^2h^2\Big)\, y_0. \end{aligned}$$

Учитывая (119) и (120), получаем

$$y_1 = (E - \beta Ah)^{-1}(E + (1-\beta)\,Ah + (1/2-\beta)\,A^2h^2)\,y_0.$$

Положим $\beta = \dfrac{1}{2}$. Тогда приходим к соотношению (131), в котором

$$F(Ah) = \Big(E - \frac{1}{2}\,Ah\Big)^{-1}\Big(E + \frac{1}{2}\,Ah\Big). \tag{135}$$

Из (119) и (120) имеем

$$p_2 = \frac{1}{2\alpha_2}\,, \qquad p_1 = \frac{2\alpha_2 - 1}{2\alpha_2}.$$

При $\alpha_2 = 1$ формула (134) приобретает вид

$$\begin{aligned} &y_1 = y_0 + \frac{1}{2}\,k_1 + \frac{1}{2}\,k_2,\\ &k_1 = hf(x_0, y_0),\\ &k_2 = hf\Big(x_0 + h,\, y_0 + \frac{1}{2}\,k_1 + \frac{1}{2}\,k_2\Big), \end{aligned} \tag{136}$$

а при $\alpha_2 = \dfrac{1}{2}$

$$\begin{aligned} &y_1 = y_0 + k_2,\\ &k_1 = hf(x_0, y_0),\\ &k_2 = hf\Big(x_0 + \frac{1}{2}\,h,\, y_0 + \frac{1}{2}\,k_2\Big). \end{aligned} \tag{137}$$

Во всех рассмотренных случаях применение неявного метода Рунге—Кутта к линейной системе (1.120), (1.121) приводит к разностному уравнению (133), в котором $F(Ah)$ — функция от матрицы, зависящая от используемой неявной формулы. Раскладывая y_n по собственным векторам матрицы A, имеем

$$y_{n+1} = F(Ah)y_n = \sum_{j=1}^{M} F(Ah)C_n^j\mathbf{e}_j.$$

Учитывая соотношение

$$F(Ah)\,\mathbf{e}_j = F(\lambda_j h)\,\mathbf{e}_j,$$

которое получается из формулы Лагранжа—Сильвестра для матричной функции, выражение для y_{n+1} можно представить следующим образом:

$$y_{n+1} = \sum_{j=1}^{M} F(\lambda_j h)\,C_n^j\mathbf{e}_j.$$

Отсюда получается формула преобразования коэффициентов

$$C_{n+1}^j = F(\lambda_j h)\,C_n^j,$$

которая совпадает с формулой численного интегрирования уравнения (5)

$$y_{n+1} = F(\lambda_j h)y_n. \tag{138}$$

Поэтому о поведении составляющей (7) решения y_n уравнения (133) можно судить по поведению решения уравнения (5).

Для (130) выражение для $F(\lambda_j h)$ имеет вид

$$F(\lambda_j h) = \frac{1 - \dfrac{\sqrt{3}}{3}\,\lambda_j h - \dfrac{1+\sqrt{3}}{6}\,\lambda_j^2 h^2}{1 - \dfrac{3+\sqrt{3}}{3}\,\lambda_j h + \dfrac{2+\sqrt{3}}{6}\,\lambda_j^2 h^2}.$$

Если $\lambda_j < 0$, то $|F(\lambda_j h)| < 1$. Показано, что если $\lambda_j = a + bi$ и $a < 0$, то $|F(\lambda_j h)| < 1$. Для формул (136), (137)

$$F(\lambda_j h) = \frac{1 + \dfrac{1}{2}\,\lambda_j h}{1 - \dfrac{1}{2}\,\lambda_j h}. \tag{139}$$

Если $\lambda_j = a + bi$ и $a < 0$, то $|F(\lambda_j h)| < 1$.

Таким образом, применение указанных формул для интегрирования уравнения (5) обеспечивает убывание решения уравнения (138) при $n \to \infty$ при произвольном h. Следовательно, обеспечивается стремление к нулю всех составляющих (7) решения разностного уравнения (133), которые соответствуют собственным значениям с отрицательной действительной частью. Поэтому для устойчивых систем после прохождения пограничного слоя шаг интегрирования может быть увеличен. При этом сохраняется численная устойчивость приближенного решения.

Мы построили неявные двухчленные формулы Рунге—Кутта второго и третьего порядка, применение которых для жестких систем позволяет увеличивать шаг интегрирования, ограничивая его единственным требованием достижения заданной точности. Батчер доказал, что для каждого q существует единственная неявная формула Рунге—Кутта (115), (116) порядка точности $2q$, причем коэффициенты этой формулы удовлетворяют следующим условиям:

1) коэффициенты $\alpha_i = \sum\limits_{j=1}^{q} \beta_{ij}$, $i = 1, \ldots, q$, являются нулями многочлена Лежандра q-й степени

$$L_q(2\alpha - 1) = \frac{1}{2^q q!} \cdot \left.\frac{d^q (x^2-1)^q}{dx^q}\right|_{x=2\alpha-1};$$

2) коэффициенты p_i удовлетворяют уравнениям

$$\sum_{i=1}^{q} p_i \alpha_i^{k-1} = \frac{1}{k}, \quad k = 1, \ldots, q; \tag{140}$$

3) коэффициенты β_{ij} удовлетворяют уравнениям

$$\sum_{j=1}^{q} \beta_{ij} \alpha_j^{k-1} = \frac{1}{k}\,\alpha_i^k, \quad i, k = 1, \ldots, q. \tag{141}$$

Эти методы называются методами *оптимального порядка*.

Построим неявные методы Рунге-Кутта оптимального порядка для $q = 1, 2, 3$. Многочлен Лежандра первой степени имеет вид

$$L_1(x) = x.$$

Следовательно,

$$L_1(2\alpha - 1) = 2\alpha - 1.$$

Отсюда $\alpha = \dfrac{1}{2}$. Коэффициент p определяется из уравнения (140) явно: $p = 1$. Таким образом, одночленная формула второго порядка имеет вид

$$\begin{aligned} y_{n+1} &= y_n + k, \\ k &= hf\Big(x_n + \frac{1}{2}\,h,\; y_n + \frac{1}{2}\,k\Big). \end{aligned} \tag{142}$$

Формула (142) может быть названа *неявной формулой средних прямоугольников*. Применение формулы (142) для задачи (1.120), (1.121) приводит к следующему выражению для k:

$$k = Ahy_n + \frac{1}{2}\,Ahk.$$

Отсюда

$$k = \Big(E - \frac{1}{2}\,Ah\Big)^{-1} Ahy_n$$

и

$$y_{n+1} = y_n + \Big(E - \frac{1}{2}\,Ah\Big)^{-1} Ahy_n = \Big(E + \Big(E - \frac{1}{2}\,Ah\Big)^{-1} Ah\Big)\, y_n =$$

$$= \Big(E - \frac{1}{2}\,Ah\Big)^{-1}\Big(E - \frac{1}{2}\,Ah + Ah\Big)\, y_n = \Big(E - \frac{1}{2}\,Ah\Big)^{-1}\Big(E + \frac{1}{2}\,Ah\Big)\, y_n.$$

Следовательно, матричная функция $F(Ah)$ для данного метода имеет вид (135), т.е. такой же, как и для двухчленных формул второго порядка. Применение (142) к уравнению (5) также приводит к разностному уравнению (138), в котором $F(\lambda_j h)$ задается с помощью (139).

Многочлен Лежандра второй степени имеет вид

$$L_2(x) = \frac{1}{2}\,(3x^2 - 1).$$

Следовательно,

$$L_2(2\alpha - 1) = \frac{1}{2}\,(3(2\alpha - 1)^2 - 1).$$

Отсюда

$$\alpha_1 = \frac{1}{2} - \frac{\sqrt{3}}{6}, \quad \alpha_2 = \frac{1}{2} + \frac{\sqrt{3}}{6}.$$

Коэффициенты p_1 и p_2 определяются из уравнений (140)

$$\begin{cases} p_1 + p_2 = 1, \\ p_1\alpha_1 + p_2\alpha_2 = \dfrac{1}{2}, \end{cases}$$

решение которых дает следующие значения:

$$p_1 = \frac{\dfrac{1}{2} - \alpha_2}{\alpha_1 - \alpha_2} = \frac{1}{2}, \quad p_2 = \frac{1}{2}.$$

Коэффициенты β_{11} и β_{12} определяются из уравнений (141) при $i = 1$:

$$\begin{cases} \beta_{11} + \beta_{12} = \alpha_1, \\ \beta_{11}\alpha_1 + \beta_{12}\alpha_2 = \dfrac{1}{2}\,\alpha_1^2, \end{cases}$$

решение которых дает следующие значения:

$$\beta_{11} = \frac{\dfrac{1}{2}\,\alpha_1^2 - \alpha_1\alpha_2}{\alpha_1 - \alpha_2} = \frac{1}{4}, \quad \beta_{12} = \frac{1}{4} - \frac{\sqrt{3}}{6}.$$

Коэффициенты β_{21} и β_{22} определяются из уравнений (141) при $i = 2$:

$$\begin{cases} \beta_{21} + \beta_{22} = \alpha_2, \\ \beta_{21}\alpha_1 + \beta_{22}\alpha_2 = \dfrac{1}{2}\,\alpha_2^2, \end{cases}$$

решение которых дает следующие значения:

$$\beta_{21} = \frac{-\frac{1}{2}\alpha_2^2}{\alpha_1 - \alpha_2} = \frac{1}{4} + \frac{\sqrt{3}}{6}, \quad \beta_{22} = \frac{1}{4}.$$

Таким образом, двучленная формула четвертого порядка имеет вид

$$\begin{aligned} y_{n+1} &= y_n + \frac{1}{2}\,k_1 + \frac{1}{2}\,k_2, \\ k_1 &= hf\left(x_n + \left(\frac{1}{2} - \frac{\sqrt{3}}{6}\right)h, \quad y_n + \frac{1}{4}\,k_1 + \left(\frac{1}{4} - \frac{\sqrt{3}}{6}\right)k_2\right), \\ k_2 &= hf\left(x_n + \left(\frac{1}{2} + \frac{\sqrt{3}}{6}\right)h, \quad y_n + \left(\frac{1}{4} + \frac{\sqrt{3}}{6}\right)k_1 + \frac{1}{4}\,k_2\right). \end{aligned} \tag{143}$$

Применение (143) к уравнению (5) приводит к разностному уравнению (138), в котором

$$F(\lambda_j h) = \frac{12 + 6\lambda_j h - \lambda_j^2 h^2}{12 - 6\lambda_j h + \lambda_j^2 h^2}.$$

Показано, что если $\mathrm{Re}(\lambda_j) < 0$, то $|F(\lambda_j h)| < 1$.

Многочлен Лежандра третьей степени имеет вид

$$L_3(x) = \frac{1}{2}\,(5x^3 - 3x).$$

Следовательно,

$$L_3(2\alpha - 1) = \frac{1}{2}\,(5(2\alpha - 1)^3 - 3(2\alpha - 1)).$$

Отсюда

$$\alpha_1 = \frac{1}{2} - \frac{\sqrt{15}}{10}, \quad \alpha_2 = \frac{1}{2}, \quad \alpha_3 = \frac{1}{2} + \frac{\sqrt{15}}{10}.$$

Коэффициенты p_i определяются из системы уравнений

$$\begin{cases} p_1 + p_2 + p_3 = 1, \\ p_1\alpha_1 + p_2\alpha_2 + p_3\alpha_3 = \dfrac{1}{2}, \\ p_1\alpha_1^2 + p_2\alpha_2^2 + p_3\alpha_3^2 = \dfrac{1}{3}. \end{cases}$$

Коэффициенты $\beta_{11}, \beta_{12}, \beta_{13}$ определяются из системы уравнений (141) при $i = 1$:

$$\begin{cases} \beta_{11} + \beta_{12} + \beta_{13} = \alpha_1, \\ \beta_{11}\alpha_1 + \beta_{12}\alpha_2 + \beta_{13}\alpha_3 = \dfrac{1}{2}\,\alpha_1^2, \\ \beta_{11}\alpha_1^2 + \beta_{12}\alpha_2^2 + \beta_{13}\alpha_3^2 = \dfrac{1}{3}\,\alpha_1^3. \end{cases}$$

Коэффициенты β_{2j} и β_{3j}, $j = 1, 2, 3$, определяются из аналогичных уравнений. Определители этих систем являются определителями Вандермонда, поэтому каждая такая система имеет единственное решение.

В результате мы приходим к следующей трехчленной формуле шестого порядка:

$$
\begin{aligned}
&y_{n+1} = y_n + \frac{5}{18}\,k_1 + \frac{4}{9}\,k_2 + \frac{5}{18}\,k_3,\\
&\quad k_1 = hf\left(x_n + \left(\frac{1}{2} - \frac{\sqrt{15}}{10}\right)h,\quad y_n + \frac{5}{36}\,k_1 + \left(\frac{2}{9} - \frac{\sqrt{15}}{15}\right)k_2 + \left(\frac{5}{36} - \frac{\sqrt{15}}{30}\right)k_3\right),\\
&\quad k_2 = hf\left(x_n + \frac{1}{2}\,h,\quad y_n + \left(\frac{5}{36} + \frac{\sqrt{15}}{24}\right)k_1 + \frac{2}{9}\,k_2 + \left(\frac{5}{36} - \frac{\sqrt{15}}{24}\right)k_3\right),\\
&\quad k_3 = hf\left(x_n + \left(\frac{1}{2} + \frac{\sqrt{15}}{10}\right)h,\quad y_n + \left(\frac{5}{36} + \frac{\sqrt{15}}{30}\right)k_1 + \left(\frac{2}{9} + \frac{\sqrt{15}}{15}\right)k_2 + \frac{5}{36}\,k_3\right).
\end{aligned}
\tag{144}
$$

Применение (144) к уравнению (5) приводит к разностному уравнению (138), в котором

$$
F(\lambda_j h) = \frac{120 + 60\lambda_j h + 12\lambda_j^2 h^2 + \lambda_j^3 h^3}{120 - 60\lambda_j h + 12\lambda_j^2 h^2 - \lambda_j^3 h^3}.
$$

Показано, что если $\mathrm{Re}(\lambda_j) < 0$, то $|F(\lambda_j h)| < 1$.

Таким образом, применение неявных формул Рунге—Кутта для интегрирования жестких систем (1.120), (1.121) обеспечивает при $n \to \infty$ при произвольном h затухание всех составляющих (7) решения разностного уравнения (133), которые соответствуют собственным значениям с отрицательной действительной частью. Это обстоятельство позволяет увеличивать шаг интегрирования после прохождения пограничного слоя, сохраняя численную устойчивость приближенного решения.

В заключение приведем список литературы, рекомендуемый к гл.3: [1, 3—5, 7, 16, 18, 19, 22, 24, 32, 34, 41, 42].