

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ**

**ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ**

**САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ**



**ПОБЕДИТЕЛЬ КОНКУРСА ИННОВАЦИОННЫХ ОБРАЗОВАТЕЛЬНЫХ ПРОГРАММ ВУЗОВ**

**Г.П. Мирошниченко, А.Г. Петрашень**

# **ЧИСЛЕННЫЕ МЕТОДЫ**

**Учебное пособие**



**Санкт-Петербург**

**2007**

Авторы:

Г.П. Мирошниченко, А.Г. Петрашень

Численные методы. Учебное пособие.-СПб: СПбГУИТМО,2007.-120 с.

Предлагаемое учебное пособие представляет собой описание лабораторных работ по различным разделам высшей математики. Пособие предназначено для студентов 2-3 курсов, обучающихся по специальности 010500 “Прикладная математика и информатика”. Работа может выполняться в любом математическом пакете, но для конкретности материал методически адаптирован к пакету MATHCAD. Этот пакет выбран не случайно, так как обладает очень удобным пользовательским интерфейсом и на его основе можно реализовать методическую идею ”живого конспекта”. В тексте описания каждой работы содержится компьютерная программа. Лабораторная работа заканчивается списком заданий как теоретического, так и вычислительного свойства, которые рекомендуется выполнить для освоения данного раздела.

Рекомендовано к печати Ученым советом естественнонаучного факультета.

Протокол № 7 от 20.03.2007г.



В 2007 году СПбГУ ИТМО стал победителем конкурса инновационных образовательных программ вузов России на 2007–2008 годы. Реализация инновационной образовательной программы «Инновационная система подготовки специалистов нового поколения в области информационных и оптических технологий» позволит выйти на качественно новый уровень подготовки выпускников и удовлетворить возрастающий спрос на специалистов в информационной, оптической и других высокотехнологичных отраслях экономики.

© Санкт-Петербургский государственный университет информационных технологий, механики и оптики, 2007

© Г.П. Мирошниченко, А.Г. Петрашень, 2007

## Содержание

Введение.....	4
1. Задача интерполирования функции с помощью полинома Лагранжа .....	6
2. Задача интерполирования функции кубическими сплайнами .....	13
3. Задача интерполирования функции дискретным рядом Фурье .....	19
4. Быстрое дискретное преобразование Фурье .....	24
5. Сходимость рядов Фурье .....	30
6. Формула Грина – Римана, теорема Стокса .....	40
6.1. Первая и вторая формулы Грина-Римана.....	40
6.2. Интегральная теорема Стокса.....	43
7. Теорема Остроградского – Гаусса.....	49
8. Численное решение систем обыкновенных дифференциальных уравнений и исследование особых решений .....	58
8.1. Схема Рунге – Кутта .....	58
8.2. Устойчивость точек покоя автономной системы двух дифференциальных уравнений первого порядка.....	60
8.3. Устойчивость предельного цикла .....	63
9. Метод Фурье и сеточный метод решения уравнений в частных производных .....	70
9.1. Метод разделения переменных (метод Фурье) для уравнения теплопроводности .....	70
9.2. Конечно-разностные методы .....	76
10. Интегральное уравнение Фредгольма II рода.....	82
11. Конечные разности .....	90
11.1. Свойства конечных разностей.....	90
11.2. Аппроксимация функций с помощью конечных разностей.....	92
11.3. Исчисление конечных сумм.....	93
12. Разностные уравнения .....	100
12.1. Линейные разностные уравнения.....	100
12.2. Z преобразование Лапласа .....	104
12.3. Дискретный WKВ метод .....	111
Литература .....	117
Кафедра высшей математики .....	118

## Введение

Пособие предназначено для студентов 2-3 курсов, обучающихся по специальности “Прикладная математика и информатика”. Численные методы имеют особое значение в программе обучения студентов, специализирующихся на приложениях высшей математики в физике, технике, экономике, социологии и других областях. Осваивающие данную специальность овладевают приемами математического компьютерного моделирования, методиками постановки и проведения численного эксперимента в выбранной области приложений. Чтобы успешно применить математику в конкретном приложении необходимо на первом этапе освоить язык, бегло ориентироваться в информационной базе этого приложения. На втором этапе необходимо выделить конкретную математическую задачу, адекватную изучаемой ситуации, освоить требуемый для ее решения математический аппарат, окончательно сформулировать задачу на языке математики. На третьем этапе необходимо компьютерное моделирование, численный эксперимент, в процессе которого находится оптимальное решение задачи, выраженное в виде оптимизированной функциональной зависимости, графических иллюстраций или просто числа. На наш взгляд, основная цель прикладной математики – это доведение конкретной математической задачи до числа. В данном пособии отражены элементы всех трех этапов, но особое внимание уделено второму и третьему этапам математического моделирования. Каждая глава представляет собой описание лабораторной работы по изучению некоторого раздела высшей математики. Работа может выполняться в любом математическом пакете, но для конкретности материал методически адаптирован к пакету MATHCAD, поэтому читатель должен иметь минимальные представления о возможностях данного пакета. Этот пакет выбран не случайно, так как обладает очень удобным пользовательским интерфейсом и на его основе можно реализовать методическую идею ”живого конспекта”. Целями каждой работы являются:

1. проведение вычислений по проверке справедливости соответствующих математических теорем изучаемого раздела
2. изучение приближенных аналитических методов решения задач и их сравнение с точными (возможно численными) методами
3. изучение различных собственно численных методов решения задач и проверка теорем о точности вычислений данным методом
4. проведение численного, возможно стохастического, эксперимента, демонстрирующего возможности численных методов.

Описание каждой работы, по возможности, содержит необходимые сведения о соответствующем разделе математики. Теоремы, как правило, формулируются, но не доказываются. В описании работы содержится компьютерная программа данной лабораторной работы. Каждая работа

заканчивается списком заданий как теоретического, так и вычислительного свойства, которые рекомендуется выполнить для освоения данного раздела. Методическая новизна предлагаемого упорядочения материала, на наш взгляд, состоит в следующем. В каждом описании содержатся формулы, имеющие разное назначение. Формулы, в которые входит знак “=” (например,  $a = b$ ) играют вспомогательную роль для освоения раздела. Такая формула в пакете MATHCAD играет роль комментария и не выполняется, а сам знак “=” имеет смысл логического равенства. Такие формулы центрируются в тексте слева, и, если необходимо, нумеруются справа. Формулы, соединенные знаком “:=” (в MATHCADe это операция присвоения, например,  $a := b$ ) являются собственно программой, они центрируются в тексте справа и, если надо, нумеруются слева. При внимательном чтении текста необходимо рабочие формулы в тех же общематематических обозначениях (это и есть главное удобство пакета) переписать на экран, соблюдая последовательность (слева направо, сверху вниз) и полученный на экране конспект лекции “оживает”, с его помощью можно производить соответствующие вычисления. Выполняющему лабораторную работу необходимо ввести исходные данные (самостоятельно, или по заданию преподавателя). Место в программе, куда требуется ввести данные, отмечено специальным знаком, например  $a := \blacklozenge$ . В тексте используются обозначения специальных операторов и функций пакета MATHCAD. Это:  $\text{floor}(x)$  - минимальная целая часть числа,  $\text{runif}(a,b,c)$  - равномерно распределенная случайная переменная,  $\text{linterp}(a,b,c)$  - линейный сплайн,  $\text{root}(a,b,c)$  - корень функции,  $\text{CreateMesh}(a,b,c,d,e,f,g)$  - генерация сетки для построения линий уровня функции двух переменных,  $\text{eigenvals}(A)$ ,  $\text{eigenvecs}(A)$  - поиск собственных векторов и собственных чисел матрицы,  $\text{combin}(n,k)$  - биномиальный коэффициент,  $\text{rnorm}(N,a,b)$  - гауссовская случайная переменная,  $J_n(N,z)$  - функция Бесселя,  $\text{if}$  - условный оператор,  $\text{for}$  - оператор цикла,  $|$  - вертикальная черта. Подробное описание данных операций можно найти в help-е MATHCADa.

# 1. Задача интерполирования функции с помощью полинома Лагранжа

Интерполирование – это один из способов приближения одной функции с помощью другой, вообще говоря, более простой [1], [6], [7], [8]. Необходимость интерполирования возникает, например, тогда, когда заданная функция имеет сложную структуру и для некоторых целей (например, для ее интегрирования) ее рационально упростить. В другом случае, например, при получении экспериментальных данных в виде таблицы, требуется интерполировать табличные значения плавной (гладкой) функцией, совпадающей с табличными значениями и дающей некоторое представление об изучаемой зависимости в промежутках между измерениями. Общая постановка задачи интерполяции состоит в следующем. Задаем интерполируемую функцию

$$f(x) := \diamond .$$

Необходимо задать количество узлов интерполяции  $n$

$$n := \diamond ,$$

задать интервал интерполяции

$$a := \diamond \quad b := \diamond ,$$

заполнить вектор размерности  $n$  положениями узлов (генерация сетки). При этом крайние узлы могут слегка отличаться от крайних точек интервала. В зависимости от конкретных ситуаций, узлы сетки могут располагаться по-разному. Зададим две сетки узлов. Первая – это равномерная сетка  $x1_k$  с шагом  $\Delta x1$

$$\begin{aligned} \Delta a &:= \diamond \quad \Delta b := \diamond \\ k &:= 0 \dots n - 1 \\ m &:= 0 \dots n - 1 \\ \Delta x1 &:= \frac{b - a - \Delta b - \Delta a}{n - 1} \\ x1_k &:= a + \Delta a + \Delta x \cdot k \end{aligned} .$$

Здесь  $\Delta a$  и  $\Delta b$  - отступы от краев интервала интерполяции (могут быть равными нулю). Узел  $x_0$  расположен в точке  $a + \Delta a$ , узел  $x_{n-1}$  расположен в точке  $b - \Delta b$ . Вторая – неравномерная сетка узлов Чебышева  $x2_k$

$$(1.1) \quad x2_k := \frac{a+b}{2} + \frac{b-a}{2} \cdot \cos \left[ \pi \cdot \frac{2 \cdot k + 1}{2 \cdot n} \right].$$

Свойства неравномерной сетки изучаются ниже. Зададим вектор значений интерполируемой функции  $f(x)$  на узлах  $x1_k$  и  $x2_k$

$$f1_k := f(x1_k)$$

$$f2_k := f(x2_k)$$

Выбираем какую либо “простую” *интерполирующую* функцию, содержащую  $n$  произвольных параметров

$$y = g(x; a_0, a_1, \dots, a_{n-1})$$

и подбираем набор параметров  $\{a_0, a_1, \dots, a_{n-1}\}$  так, чтобы удовлетворились  $n$  уравнений (принцип интерполирования, равенство интерполирующей и интерполируемой функций на узлах)

$$f(x_k) = g(x_k; a_0, a_1, \dots, a_{n-1})$$

$$k = 0, 1, \dots, n-1$$
(1.2)

Система является основной в любой задаче интерполирования. Рассмотрим *задачу линейного интерполирования*. Для этого функцию  $g(x; a_0, a_1, \dots, a_{n-1})$  выберем в виде линейной комбинации некоторого линейно независимого набора функций  $\varphi(k, x)$

$$g(x; a_0, a_1, \dots, a_{n-1}) = \sum_{k=0}^{n-1} a_k \cdot \varphi(k, x).$$

Конкретно, по Лагранжу, выбираем степенные функции  $\varphi(k, x)$

$$\varphi(k, x) = x^k$$

$$k = 0, 1, \dots, n-1$$

Полученный интерполяционный полином степени  $n-1$  называется полиномом Лагранжа

$$L(x) = \sum_{k=0}^{n-1} a_k \cdot x^k$$
(1.3)

Тогда линейная система уравнений для параметров  $\{a_0, a_1, \dots, a_{n-1}\}$  имеет вид

$$f(x_m) = \sum_{k=0}^{n-1} a_k \cdot (x_m)^k$$

$$m = 0, 1, \dots, n-1$$
(1.4)

Матрица этой системы на узлах  $x1_m$  и  $x2_m$  имеет вид

$$(1.5) \quad M1_{m,k} := (x1_m)^k$$

$$M2_{m,k} := (x2_m)^k$$

Решаем систему (1.4) с матрицами (1.5). Детерминант таких матриц не равен нулю, поэтому обратная матрица существует

$$a1 := (M1)^{-1} \cdot f1$$

$$a2 := (M2)^{-1} \cdot f2$$

Соответствующие полиномы Лагранжа для двух сеток равны

$$L1(x) := \sum_{k=0}^{n-1} a1_k \cdot x^k$$

$$L2(x) := \sum_{k=0}^{n-1} a2_k \cdot x^k$$

Изучим вопрос оценки погрешности интерполирования. Тот факт, что погрешность интерполирования сложной функции интерполирующим полиномом может быть малой, следует из теоремы Вейерштрасса:

• **Теорема:**

*Если функция  $f(x)$  непрерывна на отрезке  $[a, b]$ , то для любого  $\varepsilon > 0$  существует многочлен  $P_m(x)$  степени  $m = m(\varepsilon)$ , абсолютное отклонение которого от функции  $f(x)$  на  $[a, b]$  меньше  $\varepsilon$ :*

$$\max_{[a,b]} |f(x) - P_m(x)| < \varepsilon. \quad \bullet$$

Допустим, интерполируемая функция конечна в узлах. Тогда погрешность интерполирования с помощью полинома Лагранжа (1.3) в точке  $x$  имеет вид

$$R(x) = f(x) - L(x).$$

Очевидно на узлах  $x_m$  погрешность  $R(x_m) = 0$ ,  $m = 0, 1, \dots, n-1$ . Величина погрешности между узлами определяется свойствами функции  $f(x)$ , числом и расположением узлов интерполирования. Допустим, интерполируемая функция  $n$  раз дифференцируема. Определим вспомогательный полином степени  $n$ , имеющий нули на узлах сетки и единичный коэффициент при старшей степени

$$\Omega(x) = \prod_{k=0}^{n-1} (x - x_k). \quad (1.6)$$

Составим вспомогательную функцию

$$\Phi(x) = f(x) - L(x) - K \cdot \Omega(x).$$

Здесь  $a \leq x \leq b$  - произвольная точка, где нас интересует величина погрешности. Подберем параметр  $K$  так, чтобы функция  $\Phi(x)$  обратилась в ноль в выбранной точке  $x$ .

$$K = \frac{f(x) - L(x)}{\Omega(x)}. \quad (1.7)$$

При таком выборе  $K$  функция  $\Phi(x)$  обращается в ноль в  $n+1$  точке в интервале  $[a, b]$ . Тогда ее первая производная обратится в ноль  $n$  раз, а  $n$ -ая производная обратится в ноль один раз в некоторой точке  $a \leq \zeta \leq b$ .



Найдем  $n$ -ую производную функции  $\Phi(x)$  в точке  $\zeta$  ( $n$ -ая производная полинома  $L(x)$  равна нулю)

$$(\Phi(\zeta))^{(n)} = 0 = (f(\zeta))^{(n)} - 0 - K \cdot n!,$$

и выразим параметр  $K$  через эту точку  $\zeta$ .

$$K = \frac{(f(\zeta))^{(n)}}{n!}. \quad (1.8)$$

Приравнявая (1.8) и (1.7), получаем формулу для погрешности интерполирования в точке  $x$

$$\|R(x)\| = \max_{[a,b]} |f(x) - L(x)| \leq \frac{\|(f(x))^{(n)}\| \cdot \|\Omega(x)\|}{n!}. \quad (1.9)$$

Итак, погрешность интерполирования можно изменять для выбранной функции  $f(x)$  и заданном числе узлов интерполяции  $n$  с помощью изменения величины полинома  $\|\Omega(x)\|$ , то есть с помощью *оптимального расположения узлов интерполяции*. Теория показывает, что полином  $\Omega(x)$  (1.6) (степени  $n$ , имеющий единичный коэффициент при старшей степени), наименее уклоняющийся от оси  $x$  на интервале  $[a, b]$ , существует. Это вытекает из свойств множества *полиномов Чебышева* (степени  $m$ )  $T_m(x)$ ,  $m = 0, 1, \dots, \infty$ , определенных на интервале  $-1 \leq x \leq 1$ . Система полиномов Чебышева на интервале  $[-1, 1]$  определяется следующим рекуррентным соотношением

$$T_0(x) = 1, \quad T_1(x) = x, \\ T_{m+1}(x) = 2 \cdot x \cdot T_m(x) - T_{m-1}(x).$$

Решение этого соотношения имеет вид

$$T_m(x) = \frac{1}{2} \cdot \left( \left( x + \sqrt{x^2 - 1} \right)^m + \left( x - \sqrt{x^2 - 1} \right)^m \right). \quad (1.10)$$

$$n = 0, 1, \dots, \quad -1 \leq x \leq 1$$

Существует другая форма записи полиномов Чебышева

$$T_m(x) = \cos(m \cdot \arccos(x)), \quad (1.11)$$

что легко проверить, воспользовавшись тригонометрической формулой

$$\cos((m+1)\theta) = 2 \cdot \cos(\theta) \cdot \cos(m \cdot \theta) - \cos((m-1)\theta).$$

Из (1.11) следует, что  $|T_m(x)| \leq 1$  при  $|x| \leq 1$ . Полиномы  $T_m(x)$  имеют нули в точках

$$x_k^{(0)} = \cos\left(\frac{\pi \cdot (2 \cdot k + 1)}{2 \cdot m}\right),$$

$$k = 0, 1, \dots, m-1$$

экстремумы в точках

$$x_k^{(ext)} = \cos\left(\frac{\pi \cdot k}{m}\right),$$

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

Экстремальные значения в точках  $x_k^{(ext)}$  равны  $(-1)^k$ . Коэффициент перед старшей степенью полинома  $T_m(x)$  равен  $2^{m-1}$ . Рассмотрим систему полиномов  $\bar{T}_m(x)$ , определяемую с помощью  $T_m(x)$  по формуле

$$\bar{T}_m(x) = 2^{1-m} \cdot T_m(x).$$

Эта система полиномов имеет дополнительное свойство: коэффициент при старшей степени у полинома  $\bar{T}_m(x)$  равен единице. Эти полиномы обладают еще одним замечательным экстремальным свойством, которое сформулируем в виде теоремы:

• **Теорема:**

для любого полинома  $P_m(x)$  степени  $m$  и с коэффициентом перед старшей степенью, равным 1, на интервале  $-1 \leq x \leq 1$  выполнено неравенство  $\max_{[-1,1]} |P_m(x)| \geq \max_{[-1,1]} |\bar{T}_m(x)| = 2^{1-m}$ . •

То есть, полином  $\bar{T}_m(x)$  наименее уклоняется от нуля на  $-1 \leq x \leq 1$  по сравнению с любыми полиномами  $P_m(x)$ . Получим систему полиномов, наименее уклоняющихся от нуля, на произвольном интервале  $[a, b]$

$$\bar{T}_m^{(a,b)}(x) = (b-a)^m \cdot 2^{1-2m} \cdot T_m\left(\frac{2 \cdot x - b - a}{b-a}\right).$$

Эти полиномы обладают следующими свойствами:

1. степень полинома  $m$
2. коэффициент перед старшей степенью равен 1
3. на интервале  $[a, b]$  уклоняется от нуля меньше чем все другие полиномы такого типа
4. нули этого полинома (сетка Чебышева) даются формулой (1.1)
5. максимальное уклонение от нуля

$$\max_{[a,b]} |\bar{T}_m^{(a,b)}(x)| = (b-a)^m \cdot 2^{1-2m}.$$

Запрограммируем полиномы  $\Omega(x)$  (1.6), входящие в формулу оценки погрешности интерполирования (1.9). Для равномерной сетки

$$(1.12) \quad \Omega_1(x) := \prod_{k=0}^{n-1} (x - x_{1k}).$$

Для сетки Чебышева

$$(1.13) \quad \Omega_2(x) := \prod_{k=0}^{n-1} (x - x_{2k}).$$

Сравним численно погрешности интерполирования на двух сетках. Так как  $n$ -ую производную интерполируемой функции (входящую в определение погрешности (1.9)) оценить трудно, то для проверки формулы (1.9) можно взять отношение максимальных погрешностей, полученных численно на двух сетках. Это отношение можно сравнить с отношением соответствующих теоретических погрешностей. С помощью полинома Лагранжа можно построить интерполяционные формулы для производной и первообразной интерполируемой функции, соответственно продифференцировав и проинтегрировав полином. Меняя число узлов интерполирования, можно изучить зависимость погрешности от числа узлов и оценить устойчивость процедуры интерполирования.

### Задания к работе

1. Выбрать исходные данные: интерполируемую функцию  $f(x)$ , границы интервала интерполяции  $a$  и  $b$ , число узлов интерполяции  $n$ . Составить программу, используя формулы, записанные в описании справа.
2. Показать, что матрицы (1.5) не особенные.
3. Получить два интерполяционных полинома Лагранжа и графически убедиться, что основная задача интерполяции на двух сетках решена.
4. Доказать, что формулы (1.10), (1.11) действительно дают одинаковые серии полиномов Чебышева и получить несколько первых полиномов.
5. Составить таблицу погрешностей интерполяции на двух сетках в зависимости от числа узлов  $n$ . Величины погрешностей находятся по графику численно. Убедиться, что, начиная с некоторого  $n$ , погрешность резко возрастает (граница неустойчивости метода). Почему это происходит?
6. При одном и том же  $n$  найти отношение численных погрешностей для двух сеток. Сравнить это отношение с соответствующим отношением теоретических погрешностей. Отношение теоретических погрешностей рассчитать с помощью максимальных значений полиномов (1.12), (1.13). Сделать и обосновать вывод о применимости теоретической формулы для погрешности интерполирования (1.9).

7. Найти интерполяционные полиномы Лагранжа для производной и какой-либо первообразной интерполируемой функции  $f(x)$ . Численно найти погрешности их интерполирования.
8. Представить следующий графический материал: график интерполируемой функции и ее интерполирующего полинома для равномерной сетки на интервале интерполяции; то же для неравномерной сетки; график абсолютной погрешности интерполирования для равномерной сетки; то же для неравномерной сетки; графики для соответственно первой производной и первообразной от изучаемой функции и их интерполяционных полиномов; графики соответствующих абсолютных погрешностей; графики нескольких первых многочленов Чебышева.

## 2. Задача интерполирования функции кубическими сплайнами

К достоинствам интерполяционных полиномов Лагранжа следует отнести их высокую степень гладкости (бесконечная дифференцируемость) и простоту процедуры построения. К существенным недостаткам следует отнести необходимость большой степени полинома, если интерполируется функция на большом числе узлов, в результате чего происходит потеря точности при вычислениях. Если высокая степень гладкости интерполирующей функции не требуется, то в качестве таковой можно использовать так называемый сплайн. Так, например, если достаточна непрерывность второй производной (и не выше), то можно ограничиться *кубическим сплайном* [1], [11], [12]. Процедура построения кубического сплайна более трудоемка, чем процедура построения полинома Лагранжа, но кубический сплайн обладает существенно большей устойчивостью к увеличению числа узлов интерполяции. А, следовательно, позволяет добиться высокой точности интерполирования.

Согласно общей постановке задачи интерполяции, задаем интерполируемую функцию

$$f(x) := \diamond .$$

Необходимо задать  $n$  узлов интерполяции

$$n := \diamond ,$$

задать интервал интерполяции

$$a := \diamond \quad b := \diamond ,$$

заполнить вектор размерности  $n$  положениями узлов (генерация сетки). При этом крайние узлы совместим с крайними точками интервала

$$(2.1) \quad \begin{aligned} k &:= 0 \dots n-1, \\ \Delta x &:= \frac{b-a}{n-1}, \quad x_k := a + \Delta x \cdot k. \end{aligned}$$

Зададим вектор значений функции на узлах

$$F_k := f(x_k).$$

В (2.1) выбрана равномерная сетка узлов, в общем случае это не обязательно. На оси  $x$  получаем  $n-1$  интервалов интерполяции. Присвоим каждому интервалу номер, совпадающий с номером его левого узла, и запишем в вектор  $h$  длин этих интервалов

$$\begin{aligned} m &:= 0 \dots n-2 \\ h_m &:= x_{m+1} - x_m \end{aligned}$$

Найдем интерполирующую функцию, совпадающую с  $f(x)$  на узлах и имеющую непрерывные первую и вторую производные в каждой точке интервала  $[a, b]$ . Для этого на каждом интервале  $k$  зададим кубичный полином, определяемый четырьмя коэффициентами  $b_{j,k}$

$$P_3(k, x) = b_{1,k} + b_{2,k} \cdot x + b_{3,k} \cdot x^2 + b_{4,k} \cdot x^3,$$

$$k = 0, 1 \dots n - 2.$$

Полную интерполирующую функцию (сплайн) составим из отдельных полиномов, определенных на каждом интервале. Для обеспечения гладкости (непрерывность второй производной) на каждой границе интервалов интерполирования полиномы “сшиваются” с помощью соответствующего выбора неизвестных коэффициентов. Всего неизвестно  $4 \cdot (n - 1)$  коэффициентов, для их нахождения требуется соответствующее число уравнений. Запишем следующие  $4 \cdot n - 6$  уравнения

$$1. P_3(k, x_{k+1}) = F_{k+1},$$

$$2. P_3(0, x_0) = F_0,$$

$$3. P_3(n - 2, x_{n-1}) = F_{n-1},$$

$$4. P_3(k, x_{k+1}) = P_3(k + 1, x_{k+1}), \tag{2.2}$$

$$5. P_3'(k, x_{k+1}) = P_3'(k + 1, x_{k+1}),$$

$$6. P_3''(k, x_{k+1}) = P_3''(k + 1, x_{k+1}),$$

$$k = 0, 1 \dots n - 3.$$

В (2.2) три уравнения 1, 2, 3, решают основную задачу интерполирования – равенство значений интерполируемой и интерполирующей функций на  $n$  узлах ( $n$  уравнений). Уравнения 4, 5, 6 ( $3 \cdot (n - 2)$  уравнения) обеспечивают требуемую гладкость: непрерывность сплайна, его первой и второй производных на  $n - 2$  (внутренних) узлах. Осталось сформулировать два последних уравнения. Эти два условия выбираются из соображений удобства, их можно формулировать по-разному, в зависимости от известных свойств интерполируемой функции. От их выбора зависит погрешность интерполирования функции. Рассмотрим два способа выбора последних двух условий. *Способ первый*: предполагается известной вторая производная функции на краях интервала

$$P_3''(0, a) = f''(a), \tag{2.3}$$

$$P_3''(n - 2, b) = f''(b).$$

*Способ второй*: “естественный”, применяется тогда, когда дополнительных сведений об интерполируемой функции нет

$$\begin{aligned}
P_3''(0, a) &= 0, \\
P_3''(n-2, b) &= 0.
\end{aligned}
\tag{2.4}$$

Можно показать, что существует *сокращенный набор параметров*  $M_k$ ,  $k = 1, 2, \dots, n-2$  в количестве  $n-2$ , достаточный для построения кубического сплайна. В качестве этого набора можно выбрать *величины вторых производных на внутренних* ( $k = 1, 2, \dots, n-2$ ) узлах

$$\begin{aligned}
M_k &= P_3''(k, x_k), \\
k &= 1, 2, \dots, n-2.
\end{aligned}
\tag{2.5}$$

Компоненты  $M_0$  и  $M_{n-1}$  (вторые производные на граничных узлах) определяются для двух способов выбора согласно (2.3), (2.4). Для реализации этой идеи сокращенного числа параметров необходимо кубичный полином выбрать специальным образом, выразив его через параметры  $M_k$

$$\begin{aligned}
P_3(k, y) &= M_k \cdot \frac{(x_{k+1} - y)^3}{6 \cdot h_k} + M_{k+1} \cdot \frac{(y - x_k)^3}{6 \cdot h_k} + \\
&+ \left( F_k - \frac{M_k \cdot (h_k)^2}{6} \right) \cdot \frac{x_{k+1} - y}{h_k} + \left( F_{k+1} - \frac{M_{k+1} \cdot (h_k)^2}{6} \right) \cdot \frac{y - x_k}{h_k}, \\
k &= 0, 1, \dots, n-2.
\end{aligned}
\tag{2.6}$$

Нетрудно показать, что набора параметров (2.5) достаточно, чтобы с помощью полинома (2.6) решить задачу построения сплайна (2.2). Уравнения 1 – 4, 6 с помощью (2.6) *удовлетворяются автоматически*. Уравнения 5 (всего  $n-2$  уравнения) необходимо использовать для нахождения неизвестных  $n-2$  параметров  $M_k$ . Это линейные уравнения и легко решаются средствами любого математического пакета. Удобно вместо вектора  $M_k$  ввести вспомогательный вектор  $Q_k$ , связанный с  $M_k$  соотношением (сдвиг нумерации на единицу)

$$M_{k+1} = Q_k, \quad k = 0, 1, \dots, n-3.$$

Система уравнений 5 (2.2), выраженная через  $Q_k$ , имеет вид

$$\begin{aligned}
Q_{k+1} \cdot \frac{h_{k+1}}{6} + Q_{k-1} \cdot \frac{h_k}{6} + Q_k \cdot \frac{h_k + h_{k+1}}{3} &= \\
= \frac{F_k - F_{k+1}}{h_k} + \frac{F_{k+2} - F_{k+1}}{h_{k+1}}.
\end{aligned}
\tag{2.7}$$

Заполним матрицу  $W$  размерности  $(n-2) \otimes (n-2)$  системы (2.7)

$$k := 0 \dots n - 4 \quad m := 0 \dots n - 3$$

$$W_{k,k+1} := \frac{h_{k+1}}{6} \quad W_{k+1,k} := \frac{h_{k+1}}{6} \quad ,$$

$$W_{m,m} := \frac{h_{m+1} + h_m}{3}$$

и вектор  $B$  размерности  $n - 2$  правой части уравнения (2.7)

$$B_m := \frac{F_m - F_{m+1}}{h_m} + \frac{F_{m+2} - F_{m+1}}{h_{m+1}}.$$

Решим систему  $W \cdot Q = B$

$$Q := W^{-1} \cdot B.$$

Заполним вектор  $M$

$$M_{m+1} := Q_m.$$

Выберем один из двух способов интерполяции: (2.3) или (2.4)

(2.8)

$$M_0 := \blacklozenge$$

$$M_{n-1} := \blacklozenge$$

Определим кубичный полином  $P(k, y)$  на  $k$ -ом интервале

$$P(k, y) := M_k \cdot \frac{(x_{k+1} - y)^3}{6 \cdot h_k} + M_{k+1} \cdot \frac{(y - x_k)^3}{6 \cdot h_k} + \left( F_k - \frac{M_k \cdot (h_k)^2}{6} \right) \cdot \frac{x_{k+1} - y}{h_k} + \left( F_{k+1} - \frac{M_{k+1} \cdot (h_k)^2}{6} \right) \cdot \frac{y - x_k}{h_k}.$$

Из полиномов  $P(k, y)$  составим единую функцию – сплайн  $S(y)$ . Для этого для каждого аргумента  $y$  найдем номер интервала  $K(y)$  (по номеру его левого края), куда попадает этот  $y$ , по формуле

$$K(y) := \text{floor} \left( \frac{y - a}{\Delta x} \right).$$

Здесь  $\text{floor}(x)$  - ближайшее к  $x$  целое число слева. Тогда сплайн  $S(y)$  запишем в виде

(2.9)

$$S(y) := P(K(y), y).$$

Исследуем возможности этой интерполяционной формулы для разных функций, сеток, области определения. Погрешность интерполирования  $\Delta S$  оценим по формуле

$$\Delta S = \max_{a \leq x \leq b} |f(x) - S(x)|.$$

Составим таблицу, показывающую зависимость погрешности  $\Delta S$  в зависимости от числа узлов интерполяции  $n$  для двух способов



определения сплайнов (2.3), (2.4), (2.8) . По аналогии составим таблицу интерполяции для первой и второй производных функции  $f(x)$  с помощью соответственно первой и второй производных сплайна  $S(y)$ . Как показывает теория, погрешность интерполяции функции и ее производных для первого способа задания сплайна (2.3) можно оценить по формуле

$$\max_{a \leq x \leq b} |f^{(r)}(x) - S^{(r)}(x)| \leq h^{4-r} \cdot \max_{a \leq x \leq b} |f^{(4)}(x)|, \quad (2.10)$$

$$r = 0, 1, 2$$

Формула применима, по крайней мере, для четырехкратно непрерывно дифференцируемых функций  $f(x)$ . Здесь символом  $f^{(r)}(x)$  обозначена  $r$ -ая производная. Погрешность интерполирования для сплайна с граничными условиями (2.4) больше и оценивается, как  $O(h^2)$ . Увеличение погрешности следует ожидать для функций, у которых вторые производные на краях интервала не равны нулю. При этом погрешность локализуется на краях интервала интерполяции. Величина погрешности уменьшается при удалении от краев. Это свойство локализации погрешности очень интересно и его можно проверить, например, следующим способом. Внесем искусственно небольшую погрешность в значение интерполируемой функции на каком-либо узле (добавим к какой-нибудь компоненте вектора  $F_k$  небольшую добавку  $\delta$ ). Изучим графически, как погрешность “распространяется” на соседние узлы. “Естественный” кубический сплайн с граничным условием (2.4) обладает важным экстремальным свойством, о котором следующая

• **Теорема:**

Пусть на сетке узлов  $a = x_0 < x_1 < \dots < x_{n-1} = b$  заданы числа  $\{y_i\}_{i=0}^{n-1}$  (значения интерполируемой функции). Среди всех дважды непрерывно дифференцируемых функций, таких, что  $f(x_i) = y_i$ ,  $f''(a) = f''(b) = 0$ , на “естественном” сплайне достигается минимум функционала

$$F(f) = \int_a^b (f''(x))^2 dx$$

(свойство минимальной кривизны). •

Эта теорема может быть проверена численно.

**Задания к работе**

1. Выбрать исходные данные: интерполируемую функцию  $f(x)$ , границы интервала интерполяции  $a$  и  $b$ , число узлов интерполяции  $n$ . Составить программу, используя формулы, записанные в описании справа.

2. Показать, что полином (2.6) автоматически решает систему уравнений 1, 2, 3, 4, 6 из системы (2.2).
3. Проверить смысл параметров  $M_k$  для полиномов (2.6).
4. Проверить, что система уравнений (2.7) решает уравнение 5 системы (2.2).
5. Убедиться в справедливости формулы (2.9).
6. Графически проверить, что  $S(y)$  решает поставленную задачу интерполяции и составить таблицу погрешности интерполяции в зависимости от числа узлов.
7. Графически убедиться, что параметры  $M_k$  имеют нужный геометрический смысл.
8. Графически убедиться, что сплайн имеет требуемую гладкость.
9. Численно исследовать погрешность интерполирования производной функции  $f(x)$  в зависимости от числа узлов.
10. Численно проверить справедливость теоретической оценки погрешности интерполяции сплайном (2.10) и справедливость свойства минимальной кривизны для “естественного” сплайна.

### 3. Задача интерполирования функции дискретным рядом Фурье

В линейной задаче интерполирования интерполирующая функция выбирается в виде линейной комбинации линейно независимого набора функций. Так, например, в методе Лагранжа используются степенные функции. В качестве такого набора можно использовать тригонометрические функции. Для краткости записи синусы и косинусы удобно комбинировать в эйлеровские экспоненты с мнимым показателем степени. Интерполирующая функция, построенная на конечном комплексном экспоненциальном базисе, обычно называется дискретным рядом Фурье интерполируемой функции [1], [6]. Комплексные коэффициенты линейной комбинации экспонент называют дискретным Фурье-образом изучаемой функции (или коэффициентами дискретного Фурье-преобразования).

Определим границы интервала интерполирования

$$a := \diamond \quad b := \diamond .$$

Определим период  $T$  интерполирующей функции

$$T := b - a .$$

Зададим интерполируемую функцию

$$f(t) := \diamond ,$$

которая, в силу периодичности, для уменьшения погрешности интерполирования, должна иметь *равные значения* на краях интервала. Зададим число  $M$  (выберем его четным, для удобства)

$$M := \diamond ,$$

определяющее полное число  $M + 1$  (внутренних) узлов интерполяции. Определим шаг равномерной сетки узлов интерполяции

$$\Delta t := \frac{T}{M + 1} .$$

Заполним вектор  $\tau$  размерности  $M + 1$  положениями (внутренних) узлов (генерация сетки), края интервала  $[a, b]$  отстоят от нулевого и  $M$ -ого узлов на половину шага сетки  $\Delta/2$

(3.1)

$$k := 0..M$$
$$\tau_k := a + \frac{\Delta t}{2} + \Delta t \cdot k .$$

Заполним вектор  $F$  размерности  $M + 1$  значениями интерполируемой функции на узлах

(3.2)

$$F_k := f(\tau_k) .$$

Определим шаг сетки частот  $\Delta\omega$

$$\Delta\omega := \frac{2 \cdot \pi}{T}.$$

Введем равномерную сетку частот  $\omega$

$$(3.3) \quad \omega_k := \Delta\omega \cdot \left( k - \frac{M}{2} \right).$$

Существенно, что частоты расположены симметрично относительно нуля  $\omega_{M/2} := 0$  и принимают положительные и отрицательные значения ( $M/2$  положительных и  $M/2$  отрицательных значений). Введем базисный набор функций  $\phi(k, t)$  - экспонент, из которых далее будет скомбинирована интерполирующая функция

$$(3.4) \quad \phi(k, t) := \exp(i \cdot \omega_k \cdot t).$$

Зададим функции базисного набора (3.4) на узлах (3.1). Получим матрицу  $W$ , где в каждый столбик записаны значения соответствующей базисной функции на узлах

$$(3.5) \quad \begin{aligned} m &:= 0..M \\ W_{k,m} &:= \phi(m, \tau_k) \end{aligned}$$

Интерполирующую функцию строим как линейную комбинацию базисных функций  $\phi(k, t)$  по формуле

$$P(t) = \sum_{m=0}^M F\Omega_m \cdot \phi(m, t). \quad (3.6)$$

Система линейных уравнений (для неизвестных коэффициентов  $F\Omega_m$ ), решающая задачу интерполяции, имеет вид

$$F_k = \sum_{m=0}^M F\Omega_m \cdot W_{k,m}. \quad (3.7)$$

Здесь использованы формулы (3.2), (3.5), (3.6). Перепишем (3.7) в векторных обозначениях

$$F = \sum_{m=0}^M F\Omega_m \cdot W^{(m)}. \quad (3.8)$$

В (3.8) использован вектор  $F$  (3.2) и обозначен вектор  $W^{(m)}$  -  $m$ -ый столбик матрицы  $W$  (3.5) (значения  $m$ -ой базисной функции на узлах (3.1)). Векторы  $\{W^{(m)}, m=0,1..M\}$  обладают замечательным свойством: они образуют ортонормированный базис во множестве  $M+1$ -мерных векторов, имеющих, вообще говоря, комплексные координаты. Ортонормированность можно проверить аналитически или численно, определив скалярное произведение  $S(A, B)$  комплексных векторов  $A$  и  $B$  по формуле (черта сверху обозначает комплексное сопряжение)

$$(3.9) \quad S(A, B) := \frac{1}{M+1} \cdot \sum_{m=0}^M A_m \cdot \bar{B}_m.$$

Проверим ортонормированность базиса численно. Для этого заполним матрицу  $DS$  взаимных скалярных произведений базисных векторов  $W^{(m)}$ , и если свойство имеет место, то  $DS$  будет единичной

$$DS_{m,k} := S(W^{(m)}, W^{(k)}).$$

Запись (3.8) представляет вектор  $F$  в виде разложения по ортонормированному базису  $W^{(m)}$  с коэффициентами разложения  $F\Omega_m$ . Отсюда следует, что  $F\Omega_m$  являются координатами вектора  $F$  и находятся единственным образом. Для нахождения  $F\Omega_m$  умножим скалярно *справа* вектор  $W^{(m)}$  на  $F$

$$(3.10) \quad F\Omega_m := S(F, W^{(m)}).$$

Координаты  $F\Omega_m$  найдены, тогда интерполирующая функция имеет вид

$$P(t) := \sum_{m=0}^M F\Omega_m \cdot \phi(m, t).$$

С помощью построенной по методу дискретного Фурье преобразования интерполяционной функции можно получить интерполяционные формулы для производной и первообразной интерполируемой функции. Меняя число узлов интерполирования, можно изучить зависимость погрешности от числа узлов и оценить устойчивость процедуры интерполирования.

Аппарат дискретного преобразования Фурье широко применяется для анализа временных сигналов. В реальной ситуации полезный сигнал, поступающий на анализирующую аппаратуру, зашумлен помехами, которые, например, можно описать белым шумом. Белый шум, в отличие от полезного сигнала, имеет широкий спектральный состав. Этим свойством шума можно воспользоваться, чтобы отделить полезный сигнал от помех. Для этого можно использовать так называемые полосовые фильтры, пропускающие, в идеале, без искажений, спектральные составляющие сигнала в определенной полосе частот, и задерживающие остальные составляющие. Если высокочастотные составляющие в сигнале связаны с шумом, то фильтрация (пропускание) низкочастотных составляющих и срезание высокочастотных уменьшает уровень шума в сигнале, выделяя полезный сигнал. Это можно сделать с помощью фильтра низких частот. Промоделируем это свойство численно, добавив к полезному сигналу  $f(t)$  помеху, сгенерированную с помощью равномерно распределенной случайной величины. Для этого, используя средства математического пакета, сгенерируем отрезок случайной, равномерно распределенной (например, на интервале  $[-0.1, 0.1]$ ), последовательности, длительностью в  $M+1$  элемент

$$VY := \text{runif}(M + 1, -0.1, 0.1).$$

Тогда непрерывный отрезок случайной помехи на интервале  $[a, b]$  получим с помощью линейного сплайна

$$WN(t) := \text{linterp}(\tau, VY, t).$$

Программа  $\text{linterp}(,,)$  генерирует линейный сплайн на узлах сетки  $\tau$  (3.1). Добавим помеху к  $f(t)$ , получим зашумленный сигнал

$$(3.11) \quad \varepsilon := \diamond, \quad fN(t) := f(t) + \varepsilon \cdot WN(t).$$

Здесь параметр  $\varepsilon$  определяет вклад шума в полезный сигнал. Введем модель идеального фильтра низких частот, пропускающего без искажений составляющие спектра с номерами  $mg \leq n \leq M - mg$  (нулевая частота отвечает составляющей спектра с номером  $M/2$ ) (3.3)

$$FIL(mg, n) := \begin{cases} s \leftarrow 0 \\ s \leftarrow 1 \text{ if } mg \leq n \leq M - mg. \\ s \end{cases}$$

Зададим зашумленный сигнал на сетке

$$FN_k := fN(\tau_k).$$

Найдем его коэффициенты Фурье

$$FN\Omega_m := S(FN, W^{(m)}).$$

“Срежем” фильтром высокочастотные составляющие в зашумленном сигнале, выбрав границу срезания  $mg$

$$mg := \diamond, \quad FNF\Omega_m := FN\Omega_m \cdot FIL(mg, m)$$

Восстановим “очищенный” сигнал

$$fNf(t) := \sum_{m=0}^M FNF\Omega_m \cdot \phi(m, t).$$

Сравним графически чистый сигнал  $f(t)$ , зашумленный сигнал  $fN(t)$ , очищенный сигнал  $fNf(t)$  для разных  $\varepsilon$ ,  $mg$ ,  $M$  и сделаем вывод о возможности и точности отделения шума с помощью фильтра низкой частоты.

### Задания к работе

1. Выбрать исходные данные: интерполируемую функцию  $f(t)$ , границы интервала интерполяции  $a$  и  $b$ , число узлов интерполяции

- $M$  и составить программу, используя формулы, записанные в тексте справа.
2. Почему интервал частот Фурье  $\omega_k$  должен быть симметричен относительно нуля?
  3. Доказать аналитически, что векторы  $\{W^{(m)}, m = 0, 1 \dots M\}$  образуют ортонормированный набор. Для какой цели в определении скалярного произведения введен множитель  $\frac{1}{M+1}$ ?
  4. Для какой цели при определении скалярного произведения (3.9) компоненты правого вектора берутся комплексно сопряженными? Почему при определении коэффициента Фурье (3.10) вектор  $W^{(m)}$  умножается на  $F$  справа?
  5. Графически проверить, что  $P(t)$  решает поставленную задачу интерполяции и составить таблицу погрешности интерполяции в зависимости от числа узлов. Исследовать два типа функций: с равными и не равными значениями на границах интервала интерполяции. В чем причина отличий?
  6. Исследовать погрешность интерполирования для производной и первообразной заданной функции в зависимости от числа узлов.
  7. Изучить возможности отделения шума от полезного сигнала с помощью фильтра низкой частоты в зависимости от уровня шума и граничной частоты отсечки фильтра.

## 4. Быстрое дискретное преобразование Фурье

Процедуру построения дискретного ряда Фурье для выбранной функции можно оптимизировать, существенно сократив время вычислений [1], [6]. Эта возможность следует из следующего наблюдения. Допустим, нам задана функция  $f(t)$  на  $M$  узлах сетки

$$f(t) := \diamond .$$

Функция задана в интервале

$$a := \diamond \quad b := \diamond .$$

Определим период  $T$  дискретного ряда Фурье

$$T := b - a .$$

Зададим число  $M$  (выберем его четным, для удобства)

$$M := \diamond ,$$

определяющее полное число  $M$  (внутренних) узлов равномерной сетки. Определим шаг сетки

$$\Delta t := \frac{T}{M} .$$

Заполним вектор  $\tau$  размерности  $M$  положениями (внутренних) узлов (генерация сетки), края интервала  $[a, b]$  отстоят от нулевого и  $M - 1$  - ого узлов на половину шага сетки  $\Delta t/2$

(4.1)

$$k := 0..M - 1$$
$$\tau_k := a + \frac{\Delta t}{2} + \Delta t \cdot k .$$

Заполним вектор  $F$  размерности  $M$  значениями функции на узлах

(4.2)

$$F_k := f(\tau_k) .$$

Определим интервал частот  $\Delta \omega$

$$\Delta \omega := \frac{2 \cdot \pi}{T} .$$

Введем равномерную сетку частот  $\omega$

$$\omega_k := \Delta \omega \cdot k .$$

Затем для нахождения каждого коэффициента  $\{F\Omega_m ; m = 0, 1 \dots M - 1\}$  дискретного Фурье (полное число коэффициентов Фурье равно  $M$ ) необходимо проделать  $M$  суммирований

$$F\Omega_m = \frac{1}{M} \cdot \sum_{n=0}^{M-1} F_n \cdot \exp(-i \cdot \omega_m \cdot \tau_n) . \quad (4.3)$$



Полное число операций для вычисления всех компонент  $F\Omega_m$  оказывается порядка  $M^2$ . Это число операций можно сократить. Предположим, что число  $M$  разбивается на два сомножителя

$$M = p_1 \cdot p_2.$$

Перепишем формулу (4.3) для коэффициентов дискретного Фурье  $F\Omega_m$

$$F\Omega_m = \frac{1}{M} \cdot \exp(-i \cdot \theta \cdot m) \cdot \sum_{n=0}^{M-1} F_n \cdot \exp\left(-2 \cdot \pi \cdot i \cdot \frac{m \cdot n}{M}\right), \quad (4.4)$$

и введем новое правило суммирования. Здесь фаза

$$\theta = \Delta\omega \cdot (a + \Delta t/2).$$

Зададим целые числа  $\{m; m = 0, 1..M-1\}$  и  $\{n; n = 0, 1..M-1\}$  в виде

$$\begin{cases} m = m_1 + p_1 \cdot m_2 \\ n = n_2 + p_2 \cdot n_1 \end{cases}, \quad 0 \leq n_1, m_1 \leq p_1 - 1, \quad 0 \leq n_2, m_2 \leq p_2 - 1. \quad (4.5)$$

Здесь введены новые индексы суммирования  $n_1$  и  $n_2$  (вместо  $n$ ) и новые индексы  $m_1$  и  $m_2$  (вместо  $m$ ), нумерующие коэффициенты Фурье. Тогда сумма по  $n$  разобьется на две суммы

$$F\Omega_m = \exp(-i \cdot \theta \cdot m) \cdot \left\{ \frac{1}{p_2} \cdot \sum_{n_2=0}^{p_2-1} \exp\left(-2 \cdot \pi \cdot i \cdot \frac{m_1 + p_1 \cdot m_2}{p_1 \cdot p_2} \cdot n_2\right) \times \right. \\ \left. \times \frac{1}{p_1} \cdot \sum_{n_1=0}^{p_1-1} F_{n_2 + p_2 \cdot n_1} \cdot \exp\left(-2 \cdot \pi \cdot i \cdot \frac{m_1}{p_1} \cdot n_1\right) \right\}. \quad (4.6)$$

В формуле (4.6) в первой сумме по значку  $n_1$

$$\tilde{F}(m_1; n_2) = \frac{1}{p_1} \cdot \sum_{n_1=0}^{p_1-1} F_{n_2 + p_2 \cdot n_1} \cdot \exp\left(-2 \cdot \pi \cdot i \cdot \frac{m_1}{p_1} \cdot n_1\right)$$

заполняется  $M$  - мерный (промежуточный) вектор  $\tilde{F}(m_1; n_2)$ . Для его заполнения требуется порядка  $(p_1)^2 \cdot p_2$  операций. Во второй сумме по значку  $n_2$  заполняется искомый вектор коэффициентов Фурье  $F\Omega_m$ . Для этого требуется порядка  $(p_2)^2 \cdot p_1$  операций. Если числа  $p_1$  и  $p_2$  одного порядка, то полное число операций для получения вектора  $F\Omega_m$  имеет порядок  $M^{3/2}$  (вместо  $M^2$ ). Налицо ускорение процесса. Для проведения быстрого Фурье – преобразования выбирают число узлов  $M$  в виде

$$M := 2^r,$$

где  $r$  - целое число. Тогда требуемое число операций для получения вектора  $F\Omega_m$  имеет порядок  $M \cdot \log_2(M)$ . Имеем существенное ускорение процесса. Для реализации этой возможности требуется новая организация суммирования в формуле (4.4). Для этой цели вместо индексов  $m$  и  $n$ ,

введем  $2 \cdot r$  новых индексов  $m_1, m_2, \dots, m_r$  и  $n_1, n_2, \dots, n_r$ , принимающих два значения 0 и 1:

$$m = m_1 + 2 \cdot m_2 + \dots + 2^{r-1} \cdot m_r, \quad (4.7)$$

$$n = n_r + 2 \cdot n_{r-1} + \dots + 2^{r-1} \cdot n_1. \quad (4.8)$$

С помощью (4.7), (4.8) преобразуем произведение  $\frac{m \cdot n}{M}$ , входящее в показатель экспоненты в формуле (4.4)

$$\begin{aligned} \frac{m \cdot n}{M} &= \frac{m \cdot n}{2^r} = \frac{m_1 + 2 \cdot m_2 + \dots + 2^{r-1} \cdot m_r}{2^r} \cdot n_r + \\ &+ \frac{m_1 + 2 \cdot m_2 + \dots + 2^{r-1} \cdot m_r}{2^{r-1}} \cdot n_{r-1} + \dots + \frac{m_1 + 2 \cdot m_2 + \dots + 2^{r-1} \cdot m_r}{2} \cdot n_1 \end{aligned} \quad (4.9)$$

Выделим *целое число* из выражения (4.9). В силу периодичности экспоненты  $\exp\left(-2 \cdot \pi \cdot i \cdot \frac{m \cdot n}{M}\right)$  это целое число не существенно.

Соотношение (4.9) принимает вид

$$\begin{aligned} \frac{m \cdot n}{2^r} &= (\text{целое число}) + \frac{m_1 + 2 \cdot m_2 + \dots + 2^{r-1} \cdot m_r}{2^r} \cdot n_r + \\ &+ \frac{m_1 + 2 \cdot m_2 + \dots + 2^{r-2} \cdot m_{r-1}}{2^{r-1}} \cdot n_{r-1} + \dots + \frac{m_1 + 2 \cdot m_2}{2^2} \cdot n_2 + \frac{m_1}{2} \cdot n_1 \end{aligned} \quad (4.10)$$

При подстановке (4.10) экспонента  $\exp\left(-2 \cdot \pi \cdot i \cdot \frac{m \cdot n}{M}\right)$  разбивается на произведение экспонент, где каждый сомножитель зависит только от одной переменной суммирования. Запись числа  $m$  в виде (4.7) эквивалентна переходу в двоичную систему счисления. В двоичной системе число  $m$  задается вектором размерности  $r$  из нулей и единиц  $(m_1, m_2, \dots, m_r)$ . Компоненты вектора имеют нумерацию слева направо: от  $m_1$  до  $m_r$ . Число  $n$  в формуле (4.8) также представлено в двоичной форме. Вектор размерности  $r$ , для этого числа, также записываем в виде  $(n_1, n_2, \dots, n_r)$ , но порядок записи множителей, кратных двойке, перед этими числами, обратный к соответствующему порядку для числа  $m$ . Такую форму записи будем называть *транспонированной* по отношению к *основной* форме записи числа в двоичной форме для  $m$ . С учетом вышесказанного, формула (4.4) принимает вид

$$\begin{aligned} F\Omega_m &= \exp(-i\theta m) \cdot \left\{ \frac{1}{2} \sum_{n_r=0}^1 \exp\left(-2\pi i \frac{m_1 + 2m_2 + \dots + 2^{r-1}m_r}{2^r} n_r\right) \times \dots \times \right. \\ &\left. \times \frac{1}{2} \sum_{n_2=0}^1 \exp\left(-2\pi i \frac{m_1 + 2m_2}{2^2} n_2\right) \cdot \frac{1}{2} \sum_{n_1=0}^1 F_{n_r + 2 \cdot n_{r-1} + \dots + 2^{r-1} \cdot n_1} \cdot \exp\left(-2\pi i \frac{m_1}{2} n_1\right) \right\} \end{aligned} \quad (4.11)$$

Итак, компоненты Фурье  $F\Omega_m$  удобно нумеровать числами в двоичной системе, записанными в *основной* форме, а компоненты вектора  $F_n$  удобно нумеровать числами, записанными в *транспонированной* двоичной форме. Надо научиться каждому целому числу, имеющему порядок  $r$  в двоичной системе, сопоставлять его транспонированную форму, так, чтобы, задавая двоичный вектор  $(n_1, n_2, \dots, n_r)$ , получать транспонированное целое число  $n$ . Для этого для каждого числа  $0 \leq n \leq M = 2^r$  получим соответствующую двоичную форму, записав двоичный вектор в  $n$ -ый столбик матрицы  $nb$

$$(4.12) \quad nb_{s,n} := \frac{\left(1 - (-1)^{\text{floor}(n \cdot 2^{-s})}\right)}{2} \quad \begin{matrix} s := 0..r-1 & n := 0..M-1 \end{matrix}$$

Затем по числу  $n$  получим его транспонированный вариант  $nt_n$

$$(4.13) \quad nt_n := \sum_{s=0}^{r-1} 2^{r-1-s} \cdot nb_{s,n}$$

Например, для числа  $n=7$  при  $r=5$  получаем двоичный вектор  $nb^{(7)} = (1,1,1,0,0)$ . Тогда по формуле (4.13) транспонированное число  $nt_7 = 28$ . Обозначим выражение в фигурных скобках в (4.11) через  $G_m$

$$(4.14) \quad G_m = \frac{1}{2} \sum_{n_r=0}^1 \exp\left(-2\pi i \frac{m_1 + 2m_2 + \dots + 2^{r-1}m_r}{2^r} n_r\right) \times \dots \times \frac{1}{2} \sum_{n_2=0}^1 \exp\left(-2\pi i \frac{m_1 + 2m_2}{2^2} n_2\right) \cdot \frac{1}{2} \sum_{n_1=0}^1 F_{n_r + 2 \cdot n_{r-1} + \dots + 2^{r-1} \cdot n_1} \cdot \exp\left(-2\pi i \frac{m_1}{2} n_1\right)$$

В этой формуле имеется  $r$  промежуточных суммирований, для выполнения которых можно предложить следующее рекуррентное соотношение, позволяющее переходить от  $s$ -ого промежуточного суммирования к  $s+1$ -ому ( $0 \leq s \leq r-1$ )

$$(4.15) \quad \begin{aligned} Q(m_1, m_2, \dots, m_{s-1}; m_s; n_{s+1}, n_{s+2}, \dots, n_r) &= \\ &= \frac{1}{2} \sum_{n_s=0}^1 Q(m_1, m_2, \dots, m_{s-1}; n_s; n_{s+1}, \dots, n_r) \cdot \exp\left(-i2\pi \left(m_1 + 2m_2 + \dots + 2^{s-1}m_s\right) \frac{n_s}{2^s}\right) \end{aligned}$$

Здесь  $M$ -мерный вектор  $Q$ , компоненты которого нумеруются индексами  $(m_1, m_2, \dots, m_{s-1}; n_s; n_{s+1}, \dots, n_r)$ , где  $1 \leq s \leq r$ , равен вектору  $F$  при  $s=1$  и равен вектору  $G$  при  $s=r$

$$\begin{aligned} Q(m_1, m_2, \dots, m_r) &= G_m, \\ Q(n_1, n_2, \dots, n_r) &= F_n \end{aligned}$$

Некоторое неудобство для программирования рекуррентного соотношения (4.15) представляет то обстоятельство, что в этом выражении используется смешанная система нумерации компонент вектора  $Q$ . Индексы

$m_1, m_2 \dots m_{s-1}$  принадлежат основной форме двоичного числа, а индексы  $n_s; n_{s+1} \dots n_r$  принадлежат транспонированной форме. Удобно выбрать единую систему нумерации, например, основную. Для этого следует заполнить компоненты вектора  $F$  согласно следующему правилу: в  $n$ -ую компоненту следует записать компоненту, имеющую номер  $nt_n$ , где  $nt_n$  - число, транспонированное по отношению  $n$  (4.13). Переписанный таким образом вектор обозначим  $FT$

$$FT_n := F_{nt_n}.$$

Сумма (4.14) может быть записана в виде

$$G(FT) := \left| \begin{array}{l} D \leftarrow FT \\ \text{for } s \in 1..r \\ \quad \text{for } k \in 0..2^{s-1} - 1 \\ \quad \quad \text{for } k1 \in 0..2^{r-s} - 1 \\ \quad \quad \quad \left| \begin{array}{l} p0 \leftarrow k + 0 \cdot 2^{s-1} + k1 \cdot 2^s \\ p1 \leftarrow k + 1 \cdot 2^{s-1} + k1 \cdot 2^s \\ D0 \leftarrow \frac{1}{2} \cdot \left( D_{p0} + D_{p1} \cdot \exp\left(-2 \cdot \pi \cdot i \cdot \frac{k}{2^s}\right) \right) \\ D1 \leftarrow \frac{1}{2} \cdot \left( D_{p0} + D_{p1} \cdot \exp\left(-2 \cdot \pi \cdot i \cdot \frac{k + 2^{s-1}}{2^s}\right) \right) \\ D_{p0} \leftarrow D0 \\ D_{p1} \leftarrow D1 \end{array} \right. \\ \quad \quad \quad D \end{array} \right.$$

Окончательно получим коэффициенты  $F\Omega_n$  быстрого Фурье - преобразования

$$G := G(FT)$$

$$F\Omega_n := G_n \cdot \exp(-i \cdot \theta \cdot n).$$

Программу для коэффициентов  $F\Omega_n$  обычного (не быстрого) преобразования Фурье запишем в виде

$$F\Omega_1(F) := \left| \begin{array}{l} D_{M-1} \leftarrow 0 \\ \text{for } n \in 0..M-1 \\ \quad \text{for } k \in 0..M-1 \\ \quad \quad D_n \leftarrow \frac{F_k}{M} \cdot \exp(-\Delta\omega \cdot i \cdot n \cdot \tau_k) + D_n \\ \quad \quad \quad D \end{array} \right. .$$

Вектор коэффициентов обычного Фурье  $F\Omega_n$  равен

$$F\Omega := F\Omega(F).$$

### Задания к работе

1. Выбрать исходные данные: функцию  $f(x)$ , границы интервала  $a$  и  $b$ , число узлов сетки  $M$ .
2. Составить программу, используя формулы, записанные справа.
3. На примере (4.6) объяснить принцип ускорения вычислений коэффициентов дискретного Фурье – преобразования.
4. Какой вид имеет *целое число* в формуле (4.10)?
5. Почему формула (4.11) дает ускорение вычислений и во сколько раз?
6. Объяснить правило *транспонирования* числа в двоичной форме и работу формул (4.12), (4.13).
7. Объяснить особенности работы программы для вычисления  $G(FT)$ .
8. В чем отличие программы для вычисления коэффициентов обычного (не быстрого) Фурье  $F\Omega$  этой работы и соответствующей программы предыдущей работы?
9. Составить программы и получить дискретные ряды Фурье для выбранной функции двумя способами.
10. Сравнить численно коэффициенты Фурье для одной и той же функции, полученные в данной работе двумя способами. С помощью секундомера оценить скорости работы обоих программ и проверить теоретическую формулу для величины относительного ускорения счета коэффициентов Фурье.

## 5. Сходимость рядов Фурье

В задаче линейного интерполирования интерполирующая функция выбирается в виде линейной комбинации линейно независимого набора функций. Количество таких функций должно равняться количеству узлов сетки интерполирования. Погрешность интерполирования (для достаточно гладких функций) убывает с ростом числа узлов на интервале интерполяции. Возникает естественное предположение увеличить (в пределе до бесконечности) число узлов интерполяции и в результате получить “идеальную” интерполирующую функцию. При таком предельном переходе бесконечное, но счетное множество узлов сетки интерполирования разместится на конечном отрезке  $[a, b]$  вещественной прямой. Эти узлы займут не все точки отрезка, но заполнят его *всюду плотно*, так, что в окрестности любой точки отрезка расположится бесконечное число узлов. В математическом анализе доказывается, что скалярное произведение, определяемое для дискретного ряда Фурье формулой

$$\frac{1}{M} \sum_{m=0}^{M-1} f_m \cdot \bar{g}_m \xrightarrow{M \rightarrow \infty} \int_a^b f(t) \bar{g}(t) dt,$$

где векторы  $f_m$  и  $g_m$  есть значения функций  $f(t)$  и  $g(t)$  на узлах сетки  $\{a \leq t_m \leq b; m = 0, 1 \dots M-1\}$ , в пределе  $M \rightarrow \infty$  будет пропорционально интегралу от этих функций по интервалу  $[a, b]$ . Это верно, если соответствующие интегралы сходятся, функции квадратично интегрируемы. В этом пределе интерполирующая функция, построенная, например, с помощью коэффициентов дискретного Фурье - преобразования на базисе тригонометрических функций, переходит в тригонометрический (непрерывный, в отличие от дискретного) ряд Фурье [3], [4]

$$\sum_{k=-(M-1)/2}^{(M-1)/2} F\Omega_k \cdot \exp(i \cdot \omega_k \cdot t) \xrightarrow{M \rightarrow \infty} \sum_{k=-\infty}^{\infty} C(k) \cdot \exp(i \cdot \omega_k \cdot t). \quad (5.1)$$

Здесь  $F\Omega_k$ ,  $C(k)$  - коэффициенты соответственно дискретного и непрерывного Фурье – преобразования.

Эти наводящие соображения можно использовать, чтобы развить теорию *аппроксимации* заданной функции на выбранном отрезке с помощью (бесконечного) ряда Фурье. В отличие от задачи интерполяции задача аппроксимации ставится так. Рассматривается множество, например, квадратично интегрируемых на отрезке  $[a, b]$ , вообще говоря, комплекснозначных функций  $f(t)$  (пространство  $L_2$ )

$$\int_a^b |f(t)|^2 dt < \infty. \quad (5.2)$$

В этом множестве естественным образом вводят линейные операции умножения элементов множества (функций) на числа и их сложение. Определяют операцию скалярного произведения функций (векторов) множества, как интеграл по отрезку

$$(f(t), g(t)) = \int_a^b f(t) \cdot \bar{g}(t) dt. \quad (5.3)$$

Здесь черта сверху обозначает комплексное сопряжение. Вводят норму вектора

$$N(f) = \sqrt{\int_a^b |f(t)|^2 dt}$$

и определяют расстояние между векторами - метрику

$$\rho(f, g) = \sqrt{\int_a^b |f(t) - g(t)|^2 dt}. \quad (5.4)$$

Полное по норме линейное нормированное множество со скалярным произведением (5.3) становится гильбертовым пространством. Далее будем рассматривать гильбертово пространство  $L_2$  (5.2) на  $[a, b]$ . В этом пространстве выберем полный ортонормированный базисный набор функций  $\{\varphi(n, t), n = 1, 2, \dots\}$ , такой, что

$$(\varphi(n, t), \varphi(m, t)) = \delta_{n, m}, \quad n, m = 1, 2, 3, \dots,$$

где  $\delta_{n, m}$  - символ Кронекера. Выберем произвольную квадратично интегрируемую функцию  $f(t)$  и разложим ее по выбранному базису. Для этого найдем ее “координаты” в этом базисе – коэффициенты Фурье

$$C(n) = (f(t), \varphi(n, t)), \quad n = 1, 2, \dots. \quad (5.5)$$

Составим частичную сумму обобщенного ряда Фурье

$$F(k, t) = \sum_{n=1}^k C(n) \cdot \varphi(n, t). \quad (5.6)$$

Ряд Фурье выбранной функции  $f(t)$  - это предел частичных сумм

$$F(t) = \lim_{k \rightarrow \infty} \sum_{n=1}^k C(n) \cdot \varphi(n, t). \quad (5.7)$$

Необходимо исследовать вопрос о сходимости данного ряда. Убедимся, что частичная сумма (5.6) обладает свойством *наилучшего приближения в среднем* функции  $f(t)$ . Это свойство означает, что из всех функций, принадлежащих линейной оболочке, построенной на элементах базиса с

номера от 1 до  $k$ , именно частичная сумма  $F(k, t)$  наименее удалена от функции  $f(t)$  в смысле расстояния (5.4). Для доказательства выберем произвольную линейную комбинацию с коэффициентами  $\{\alpha_n, n = 1, 2, \dots\}$

$$G(t) = \sum_{n=1}^k \alpha_n \cdot \varphi(n, t)$$

и найдем квадрат расстояния от  $G(t)$  до  $f(t)$  по формуле (5.4)

$$\rho(f, G)^2 = \int_a^b \left| f(t) - \sum_{n=1}^k \alpha_n \cdot \varphi(n, t) \right|^2 dt = R(\alpha_1, \dots, \alpha_k).$$

Эта величина  $R(\alpha_1, \dots, \alpha_k)$  положительна и зависит от всех коэффициентов линейной комбинации. Воспользовавшись свойством ортонормированности базиса и определением (5.5), запишем  $R(\alpha_1, \dots, \alpha_k)$  в виде

$$R(\alpha_1, \dots, \alpha_k) = N(f)^2 - 2 \cdot \sum_{n=1}^k \alpha_n \cdot C(n) + \sum_{n=1}^k \alpha_n^2. \quad (5.8)$$

Здесь для упрощения полагаем, что коэффициенты  $\alpha_n$  и  $C(n)$  вещественны. Из этого соотношения следует, что минимальное значение расстояния получается при равенстве

$$\alpha_n = C(n), \quad n = 1, 2, \dots, k. \quad (5.9)$$

Подставим (5.9) в (5.8), получим в пределе

$$0 \leq R(\alpha_1, \dots, \alpha_k) = N(f)^2 - \sum_{n=1}^k C(n)^2 \xrightarrow{k \rightarrow \infty} N(f)^2 - \sum_{n=1}^{\infty} C(n)^2.$$

Или в другой записи имеем неравенство Бесселя для квадратов модулей коэффициентов Фурье

$$N(f)^2 \geq \sum_{n=1}^{\infty} |C(n)|^2.$$

Если выбранный базис *полный*, то неравенство Бесселя переходит в равенство Парсеваля

$$N(f)^2 = \sum_{n=1}^{\infty} |C(n)|^2. \quad (5.10)$$

Отсюда следует, что числовой ряд для квадратов модулей коэффициентов Фурье для функции из  $L_2$  сходится. Можно показать, используя свойство полноты ортонормированного базиса и свойство *полноты гильбертова пространства*  $L_2$ , что частичная сумма ряда (5.6) сходится к исходной функции  $f(t)$  *в среднем*, то есть



$$\lim_{k \rightarrow \infty} \rho(f(t), F(k, t)) = \lim_{k \rightarrow \infty} \sqrt{\int_a^b |f(t) - F(k, t)|^2 dt} = \sqrt{\int_a^b |f(t) - F(t)|^2 dt} = 0.$$

Это равенство означает, что ряд Фурье (5.7) совпадает с исходной функцией  $f(t)$  “почти везде”.

Проверим высказанные предположения численно. Для этого введем полный ортонормированный базис  $\varphi(m, t)$  на отрезке  $[a, b]$

$$a := \blacklozenge \quad b := \blacklozenge$$

$$T := b - a \quad \omega := \frac{2 \cdot \pi}{T}$$

$$\varphi(m, t) := \frac{1}{\sqrt{T}} \cdot \exp(i \cdot \omega \cdot m \cdot t).$$

Зададим четное число  $M$ , определяющее общее число  $M + 1$  учитываемых базисных элементов

$$M := \blacklozenge .$$

Проверим ортонормированность набора. Для этого заполним матрицу  $Q$  взаимными скалярными произведениями векторов базиса

$$n := 0..M \quad m := 0..M$$

$$Q_{n,m} := \int_a^b \varphi\left(n - \frac{M}{2}, t\right) \cdot \bar{\varphi}\left(m - \frac{M}{2}, t\right) dt.$$

Матрица  $Q$  должна быть близка к единичной. Введем изучаемую функцию  $f(t)$

$$f(t) := \blacklozenge .$$

Найдем ее норму

$$(5.11) \quad Nf := \sqrt{\int_a^b |f(t)|^2 dt}.$$

Интеграл должен сходиться. Заполним вспомогательный вектор  $c$  коэффициентами (в количестве  $M + 1$ )

$$c_m := \int_a^b f(t) \cdot \bar{\varphi}\left(m - \frac{M}{2}, t\right) dt.$$

Определим коэффициенты Фурье

$$C(n) := c_{n + \frac{M}{2}} .$$

Тогда частичная сумма ряда Фурье, построенная на  $2 \cdot p + 1$   $\left( p \leq \frac{M}{2} \right)$  базисном элементе равна

$$(5.12) \quad F(p, t) := \sum_{n=-p}^p C(n) \cdot \varphi(n, t).$$

Очевидно, с помощью компьютера невозможно воспроизвести весь бесконечный ряд Фурье (5.1) выбранной функции. Поэтому ряд приходится обрывать, то есть использовать частичную сумму ряда (5.12), вместо полной суммы, в результате чего возникает *погрешность аппроксимации* функции. Сравним качественно графики функции  $f(t)$  и частичной суммы ее ряда Фурье в зависимости от числа учтенных слагаемых  $p$ . Найдем разность  $\Delta Nf_p$  квадрата нормы (5.11) функции  $f(t)$  и квадрата нормы частичной суммы ряда (5.12)

$$(5.13) \quad p := 0.. \frac{M}{2} \quad \Delta Nf_p := Nf^2 - \sum_{k=-p}^p |C(k)|^2.$$

Численно убедимся, что с ростом числа учтенных слагаемых неравенство Бесселя приближается к равенству Парсеваля. Как показывает теория, числовой ряд из квадратов модулей коэффициентов Фурье для функций из  $L_2$  сходится. Численно оценим скорость убывания коэффициентов, умножив  $|C(n)|^2$  на  $n^\varepsilon$

$$(5.14) \quad \Delta C_n := |C(n)|^2 \cdot n^\varepsilon$$

и оценим показатель  $\varepsilon$ , при котором получается убывание.

Сравнить функцию  $f(t)$  и ее ряд Фурье  $F(t)$  (5.7) в каждой точке на интервале  $[a, b]$  можно с помощью понятия *поточечной сходимости*. Для этого близость двух функций будем оценивать с помощью расстояния, вычисляемого по формуле

$$\rho_c(f, g) = \max_{a \leq t \leq b} |f(t) - g(t)|. \quad (5.15)$$

Сформулируем известную теорему Дирихле о поточечной сходимости ряда Фурье по тригонометрической системе функций. Говорят, что функция удовлетворяет условиям Дирихле на интервале  $[a, b]$ , если она или непрерывна на этом интервале или имеет на  $[a, b]$  конечное число разрывов *первого рода*. Для функций, удовлетворяющих условиям Дирихле справедлива:

• **Теорема Дирихле.** Если  $f(t)$ , заданная на  $[a,b]$ , удовлетворяет условиям Дирихле на этом интервале, то ряд Фурье этой функции сходится на всем интервале  $[a,b]$  и сумма этого ряда:

1) равна  $f(t)$  во всех точках непрерывности  $f(t)$ , лежащих внутри интервала;

2) равна

$$\frac{f(t+0) + f(t-0)}{2}$$

во всех точках разрыва непрерывности;

3) равна

$$\frac{f(a+0) + f(b-0)}{2}$$

на концах интервала, то есть в точках  $t = a$  и  $t = b$ . •

Обрыв ряда Фурье приводит к появлению дополнительной погрешности аппроксимации функции. Сформулируем известные теоремы о связи степени гладкости функции и скорости сходимости ее тригонометрического ряда Фурье, с помощью которых можно оценить погрешность аппроксимации в каждой точке на интервале  $[a,b]$ . Вначале дадим определение понятию: *гладкость функции порядка  $q$* . Это:

• **Определение.** Функцию  $f(t)$  назовем гладкой порядка  $q$  на интервале  $[a,b]$ , если сама функция  $f(t)$  и ее производные до порядка  $q$  включительно  $f^{(1)}(t), \dots, f^{(q)}(t)$  непрерывны на интервале  $[a,b]$  и имеют равные значения на краях интервала

$$f(a+0) = f(b-0), f^{(1)}(a+0) = f^{(1)}(b-0), \dots, f^{(q)}(a+0) = f^{(q)}(b-0). \bullet$$

Для функций, имеющих гладкость порядка  $q$  справедлива:

• **Теорема.** Если функция  $f(t)$  на интервале  $[a,b]$  имеет гладкость порядка  $q$ , а ее  $q+1$  производная  $f^{(q+1)}(t)$  на  $[a,b]$  кусочно-непрерывна (то есть имеет конечное число разрывов первого рода на  $[a,b]$ ), то для коэффициентов Фурье функции по тригонометрической системе справедлива оценка

$$|C(k)| = o\left(\frac{1}{k^{q+1}}\right) \quad (5.16)$$

и ряд

$$\sum_{k=1}^{\infty} k^{\nu} \cdot |C(k)|, \quad \nu = 0, 1, \dots, q$$

сходится.

•

Соотношение (5.16) означает, что

$$|C(k)| \cdot k^{q+1} \xrightarrow{k \rightarrow \infty} 0.$$

Из последней теоремы следует оценка погрешности, которая возникает при замене суммы тригонометрического ряда Фурье его частичной суммой. При выполнении условия гладкости функции порядка  $q$  на интервале  $[a, b]$  получаем оценку модуля остатка (хвоста) ряда Фурье, то есть *оценку погрешности замены ряда на частичную сумму*, содержащую  $2 \cdot p + 1$  слагаемое, в зависимости от степени гладкости

$$\Delta F(q, p) = \rho_c(f(t), F(p, t)) = \max_{a \leq t \leq b} |F(p, t) - f(t)| = O\left(\frac{1}{p^{q+0.5}}\right). \quad (5.17)$$

Это соотношение означает, что  $\lim_{p \rightarrow \infty} \Delta F(q, p) \cdot p^{q+0.5}$  конечен при всех  $q$ .

Проверим высказанные положения численно. Для этой цели зададим две функции  $f_0(t)$  и  $f_1(t)$ , имеющие степень гладкости  $q=0$  и  $q=1$  соответственно. Для получения функции  $f_0(t)$  вначале выберем произвольную непрерывную функцию  $\varphi(t)$

$$\varphi(t) := \blacklozenge$$

а затем “исправим” ее, добавив линейное по аргументу  $t$  слагаемое с произвольным параметром  $\alpha$ . Подберем  $\alpha$  так, чтобы сумма  $\varphi(t) + \alpha \cdot t$  оказалась бы функцией гладкости  $q=0$

$$(5.18) \quad f_0(t) := \varphi(t) + \alpha \cdot t.$$

Для получения функции гладкости  $q=1$  выберем произвольную функцию

$$\chi(t) := \blacklozenge$$

и “исправим” ее, добавив линейное и квадратичное по аргументу  $t$  слагаемое с произвольными параметрами  $\alpha$  и  $\beta$ . Тогда

$$(5.19) \quad f_1(t) := \varphi(t) + \alpha \cdot t + \beta \cdot t^2.$$

Подберем параметры  $\alpha$  и  $\beta$  так, чтобы  $f_1(t)$  оказалась бы функцией гладкости  $q=1$ . Найдем коэффициенты Фурье для  $f_0(t)$  и  $f_1(t)$

$$c_{0_m} := \int_a^b f_0(t) \cdot \bar{\varphi}\left(m - \frac{M}{2}, t\right) dt,$$

$$C_0(n) := c_{0_{n+\frac{M}{2}}},$$

$$c1_m := \int_a^b f1(t) \cdot \bar{\varphi}\left(m - \frac{M}{2}, t\right) dt,$$

$$C1(n) := c1_{n + \frac{M}{2}}.$$

Получим соответствующие частичные суммы

$$F0(p, t) := \sum_{n=-p}^p C0(n) \cdot \varphi(n, t),$$

$$F1(p, t) := \sum_{n=-p}^p C1(n) \cdot \varphi(n, t).$$

По аналогии с (5.14) найдем численно скорости убывания модулей коэффициентов  $C0(n)$  и  $C1(n)$  и сравним полученные результаты с выводами теоремы (5.16). Оценим *численно*, с помощью графика, погрешность поточечной аппроксимации (5.17) функций  $f0(t)$  и  $f1(t)$  их частичными суммами  $F0(p, t)$  и  $F1(p, t)$  для двух значений числа  $p$ , отличающихся, например, в два раза

$$\Delta F(q, p) := \diamond$$

$$\Delta F(q, 2 \cdot p) := \diamond$$

Тогда, согласно (5.17)

$$\frac{\Delta F(q, p)}{\Delta F(q, 2 \cdot p)} \approx 2^{q+0.5}.$$

В справедливости этой оценки для разных гладкостей  $q$  можно убедиться.

На идее представления негладкой функции в виде суммы гладкой функции и простой (линейной) функции (5.18) основан метод ускорения сходимости ряда Фурье, предложенный А.Н.Крыловым. Ряд Фурье для негладкой функции медленно сходится. Возникает вопрос, нельзя ли из медленно сходящегося тригонометрического ряда выделить такой медленно сходящийся тригонометрический ряд Фурье, сумма которого известна в замкнутой форме, чтобы оставшийся ряд сходился бы достаточно быстро. Такой подход называется *улучшением сходимости ряда Фурье*. Для реализации этой идеи введем кусочно-непрерывную функцию  $U(t)$ , имеющую разные значения на краях и имеющую разрыв в некоторой точке. Для этой цели зададим точку, где происходит разрыв функции

$$h := \diamond$$

Зададим, условно, левую  $ua(t)$  и правую  $ub(t)$  ветви искомой функции  $U(t)$

$$ua(t) := \blacklozenge \quad ub(t) := \blacklozenge .$$

Используя условные операторы, составим из этих ветвей единую функцию  $U(t)$

$$U(t) := \begin{cases} c \leftarrow 0 \\ c \leftarrow ua(t) \text{ if } a \leq t \leq h \\ c \leftarrow ub(t) \text{ if } h < t \leq b \\ c \end{cases} .$$

Выделим из разрывной функции  $U(t)$  гладкую функцию  $V(t)$ , линейную функцию  $\alpha \cdot t$  и функцию “скачка”  $\Delta \cdot \sigma(t-h)$

$$V(t) = U(t) + \alpha \cdot t + \Delta \cdot \sigma(t-h).$$

Здесь  $\alpha$  и  $\Delta$  - дополнительные параметры,  $\sigma(t)$  - функция единичного скачка в точке  $t=0$  (функция Хевисайда). Подберем параметры  $\alpha$  и  $\Delta$  так, чтобы функция  $V(t)$  получила гладкость  $q=0$

$$\alpha := \blacklozenge \quad \Delta := \blacklozenge .$$

Получим частичную сумму ряда Фурье функции  $V(t)$

$$cv_m := \int_a^b V(t) \cdot \bar{\varphi}\left(m - \frac{M}{2}, t\right) dt,$$

$$CV(n) := cv_{n + \frac{M}{2}},$$

$$FV(p, t) := \sum_{n=-p}^p CV(n) \cdot \varphi(n, t).$$

Получим частичную сумму ряда Фурье для функции  $U(t)$

$$cu_m := \int_a^b U(t) \cdot \bar{\varphi}\left(m - \frac{M}{2}, t\right) dt,$$

$$CU(n) := cu_{n + \frac{M}{2}},$$

$$FU(p, t) := \sum_{n=-p}^p CU(n) \cdot \varphi(n, t).$$

Будем аппроксимировать одну и ту же функцию  $U(t)$  двумя способами: с помощью частичной суммы ее ряда Фурье  $FU(p, t)$  и с помощью функции  $UA(p, t)$ , построенной с помощью идеи улучшения сходимости

$$UA(p, t) := FV(p, t) - \alpha \cdot t - \Delta \cdot \sigma(t-h).$$

Найдем численно с помощью соответствующих графиков, погрешности (5.17) аппроксимации функции  $U(t)$  для двух способов аппроксимации. Составим таблицу погрешностей для разных  $p$  и убедимся в эффективности метода улучшения сходимости ряда Фурье.

### Задания к работе

1. Доказать формулу (5.8) и равенство (5.9).
2. Выбрать исходные данные: интерполируемую функцию  $f(t)$ , границы интервала  $a$  и  $b$ , число (четное) базисных элементов  $M$ , убедиться в ортонормированности тригонометрического базиса. Составить программу, используя формулы, записанные в тексте справа.
3. Найти коэффициенты Фурье функции  $f(t)$  и изучить скорость изменения квадратов модулей коэффициентов Фурье и скорость сходимости числового ряда из квадратов модулей коэффициентов. Оценить точность, с которой выполняется равенство Парсеваля.
4. Найти частичную сумму ряда Фурье для функции  $f(t)$  и графически проверить выводы теоремы Дирихле.
5. Получить формулы для коэффициентов  $\alpha$  и  $\beta$  и построить две функции гладкости  $q = 0$  и  $q = 1$  (5.18), (5.19).
6. Получить частичные суммы Фурье для функций гладкости  $q = 0$  и  $q = 1$  (5.18) и (5.19) и численно проверить утверждения теоремы о связи гладкости функции и скорости сходимости ее тригонометрического ряда Фурье.
7. Проверить численно возможность *улучшения сходимости* тригонометрического ряда Фурье для кусочно-непрерывной функции.

## 6. Формула Грина – Римана, теорема Стокса

### 6.1. Первая и вторая формулы Грина-Римана

Рассмотрим криволинейный интеграл по кривой  $y = \psi(\gamma, x)$  на плоскости  $xu$ , проходящей через две заданные точки  $A(xa, ya)$  и  $B(xb, yb)$

$$xa := \blacklozenge \quad ya := \blacklozenge \quad xb := \blacklozenge \quad yb := \blacklozenge .$$

Форму кривой можно изменять при помощи параметра  $\gamma$ , при этом точки  $A$  и  $B$  на плоскости не должны зависеть от параметра  $\gamma$ . Получить такую функцию  $\psi(\gamma, x)$  можно, например, так. Выберем произвольную функцию  $f(\gamma, x)$  и добавим к ней линейные слагаемые  $\alpha(\gamma) \cdot x + \beta(\gamma)$ . Подберем функции  $\alpha(\gamma)$  и  $\beta(\gamma)$  так, чтобы выполнялись два условия неподвижности точек  $A$  и  $B$  на плоскости при изменении параметра  $\gamma$

$$f(\gamma, x) := \blacklozenge ,$$

$$\alpha(\gamma) := \blacklozenge \quad \beta(\gamma) := \blacklozenge ,$$

$$\psi(\gamma, x) := f(\gamma, x) + \alpha(\gamma) \cdot x + \beta(\gamma) .$$

Проверим графически свойство функции  $\psi(\gamma, x)$ , построив два графика при двух значениях параметра

$$\gamma 1 := \blacklozenge \quad \gamma 2 := \blacklozenge .$$

Получим замкнутый контур  $\Gamma$  на плоскости  $xu$ , охватывающий область  $Dxu$ . Зададим на плоскости силовое поле

$$P(x, y) := \blacklozenge \quad Q(x, y) := \blacklozenge ,$$

$$F(x, y) := \begin{pmatrix} P(x, y) \\ Q(x, y) \end{pmatrix} .$$

Как известно, для односвязной области, в которой функции  $P(x, y)$  и  $Q(x, y)$  не имеют особенностей, необходимое и достаточное условие независимости криволинейного интеграла  $I = \int_{AB} P(x, y) dx + Q(x, y) dy$  по

кривой  $AB$  от пути является выполнение равенства Римана [3]

$$\frac{\partial}{\partial y} P(x, y) = \frac{\partial}{\partial x} Q(x, y) . \quad (6.1)$$

Проверим это утверждение численно. Для этого определим вектор касательной к кривой



$$\tau(\gamma, x) := \begin{pmatrix} 1 \\ \frac{d}{dx} \psi(\gamma, x) \end{pmatrix},$$

и перепишем криволинейный интеграл через определенный интеграл по параметру, в качестве которого выберем координату  $x$  точки на кривой

$$I(\gamma) := \int_{xa}^{xb} F(x, \psi(\gamma, x)) \cdot \tau(\gamma, x) dx.$$

Построим график  $I(\gamma)$  для двух случаев: условие (6.1) выполнено, и условие (6.1) не выполнено. По виду графика сделаем вывод о справедливости теоремы о независимости криволинейного интеграла от формы кривой. Составим криволинейный интеграл по замкнутому контуру  $\Gamma$  и рассчитаем криволинейный интеграл для двух случаев. Составим ротор (для плоского случая) силового поля

$$R(x, y) := \frac{d}{dx} Q(x, y) - \frac{d}{dy} P(x, y).$$

Тогда согласно формуле Грина выполнено равенство

$$\iint_{D_{xy}} \left( \frac{d}{dx} Q(x, y) - \frac{d}{dy} P(x, y) \right) dx dy = \int_{\Gamma} P(x, y) dx + Q(x, y) dy. \quad (6.2)$$

Двойной интеграл в (6.2) вычислим по “прямоугольникам”, заменив интегрирование на двойную интегральную сумму. Для этого поместим область интегрирования в прямоугольник, размеры которого можно определить по рисунку области  $D_{xy}$

$$\begin{aligned} X_{min} &:= xa & X_{max} &:= xb \\ Y_{min} &:= \blacklozenge & Y_{max} &:= \blacklozenge \end{aligned}$$

Введем характеристическую функцию области  $D_{xy}$

$$\chi_{xy}(x, y) := \begin{cases} a \leftarrow 0 \\ a \leftarrow 1 \text{ if } X_{min} \leq x \leq X_{max} \wedge \psi(\gamma_2, x) \leq y \leq \psi(\gamma_1, x), \\ a \end{cases}$$

предполагая, что кривая  $\psi(\gamma_2, x)$  лежит ниже кривой  $\psi(\gamma_1, x)$ . Зададим равномерную сетку узлов на выбранном прямоугольнике

$$\begin{aligned} N_x &:= \blacklozenge & N_y &:= \blacklozenge \\ k &:= 0..N_x & n &:= 0..N_y \\ \Delta x &:= \frac{X_{max} - X_{min} + 0.04}{N_x} \end{aligned}$$

$$\Delta y := \frac{Y_{max} - Y_{min} + 0.04}{N_y}$$

$$x_k := \Delta x \cdot k + X_{min} - 0.02 .$$

$$y_n := \Delta y \cdot n + Y_{min} - 0.02$$

Заполним матрицу  $M$  значениями подынтегральной функции (6.2) на сетке

$$M_{k,n} := R(x_k, y_n) \cdot \chi_{xy}(x_k, y_n).$$

Двойной интеграл приближенно равен

$$AD := \sum_{k=0}^{N_x} \sum_{n=0}^{N_y} M_{k,n} \cdot \Delta x \cdot \Delta y .$$

Сравним это значение с величиной криволинейного интеграла по замкнутому контуру и сделаем вывод о применимости первой формулы Грина-Римана (6.2).

Аналогично проверим применимость второй формулы Грина-Римана. Для этого введем две дважды дифференцируемые функции  $u(x, y)$  и  $v(x, y)$ , не имеющие особенностей на области  $D_{xy}$

$$u(x, y) := \blacklozenge \quad v(x, y) := \blacklozenge .$$

Определим вектор силы  $F(x, y)$

$$P(x, y) := v(x, y) \cdot \frac{d}{dy} u(x, y) - u(x, y) \cdot \frac{d}{dy} v(x, y)$$

$$Q(x, y) := u(x, y) \cdot \frac{d}{dx} v(x, y) - v(x, y) \cdot \frac{d}{dx} u(x, y)$$

$$F(x, y) := \begin{pmatrix} P(x, y) \\ Q(x, y) \end{pmatrix} .$$

Поддействуем оператором Лапласа на функции  $u(x, y)$  и  $v(x, y)$

$$\Delta u(x, y) := \frac{d^2}{dx^2} u(x, y) + \frac{d^2}{dy^2} u(x, y)$$

$$\Delta v(x, y) := \frac{d^2}{dx^2} v(x, y) + \frac{d^2}{dy^2} v(x, y)$$

Вторая формула имеет вид

$$\iint_{D_{xy}} (u(x, y) \cdot \Delta v(x, y) - v(x, y) \cdot \Delta u(x, y)) dx dy = \int_{\Gamma} P(x, y) dx + Q(x, y) dy .$$

Эта формула проверяется вычислением.

## 6.2. Интегральная теорема Стокса

Обобщением формулы Грина-Римана на трехмерный случай является формула Стокса. Математическая теорема Стокса широко применяется в физике и технике для изучения свойств силовых полей. Рассмотрим замкнутый контур  $L$  в трехмерном пространстве. Пусть в каждой точке контура на пробное тело действует сила  $F(x, y, z)$ . Выделим ориентированный вдоль вектора касательной элемент контура  $d\tau(x, y, z)$ , расположенный в точке  $r(x, y, z)$  на контуре. Медленно сместим пробное тело вдоль элемента  $d\tau(x, y, z)$ . Предположим, что сила достаточно гладко зависит от координат, а элемент контура достаточно мал, так, что можно полагать, что сила постоянна во всех точках элемента. Тогда элемент работы, совершенной полем над пробным телом  $dA(x, y, z)$  рассчитывается по формуле

$$dA(x, y, z) = (F(x, y, z), d\tau(x, y, z)).$$

Разобьем замкнутый контур  $L$  на элементы, найдем работу на каждом смещении, просуммируем вдоль контура, получим работу над пробным телом, которое совершило обход по замкнутому контуру. Очевидно, величина работы (циркуляции по замкнутому контуру) выражается криволинейным интегралом

$$A = \oint_L (F(x, y, z), d\tau(x, y, z)). \quad (6.3)$$

Интерес представляют силовые поля, для которых интеграл (6.3) не равен нулю, так как в этом случае возможен циклический двигатель, использующий энергию поля. В поле, для которого существует потенциальная энергия  $U(x, y, z)$ , силовое поле записывается в виде

$$F(x, y, z) = \text{grad}(U(x, y, z)). \quad (6.4)$$

В этом случае интеграл по контуру между двумя точками не зависит от контура, и, следовательно, работа по замкнутому контуру равна 0. Поэтому ожидать ненулевого значения для  $A$  можно только для вихревых полей, то есть для полей, не выражаемых в виде (6.4). Вихревые поля, в отличие от потенциальных, имеют не нулевой вектор ротора  $\text{rot}(F(x, y, z)) \neq 0$ . Как известно, если в данной точке  $r(x, y, z)$  на бесконечно малом ориентированном по нормали  $n(x, y, z)$  элементе поверхности  $d\Sigma(x, y, z)$  ротор силового поля не равен нулю, то работа при обходе элемента поверхности в положительном направлении (нормаль от ног к голове, поверхность остается слева) равна

$dA(x, y, z) = (\text{rot}(F(x, y, z)), d\Sigma(x, y, z))$ . Разобьем поверхность  $\Sigma$ , границей которой является контур  $L$  из (6.3), на элементы  $d\Sigma(x, y, z)$ ,

найдем элементы работы, просуммируем по поверхности, получим полную работу  $A\Sigma$ . Очевидно  $A\Sigma$  выражается через поверхностный интеграл и равна потоку вектора ротора

$$A\Sigma = \iint_{\Sigma} (\text{rot}(F(x, y, z)), d\Sigma(x, y, z)).$$

Теорема Стокса утверждает, что для согласованных направлений  $A = A\Sigma$ , или

$$\oint_L (F(x, y, z), d\tau(x, y, z)) = \iint_{\Sigma} (\text{rot}(F(x, y, z)), d\Sigma(x, y, z)). \quad (6.5)$$

Проверим соотношение численно. Для этого на плоскости  $xu$  зададим контур, состоящий из двух пересекающихся в двух точках, кривых. Ближняя кривая, по отношению к положительному направлению оси  $x$ , задается уравнением  $y = y_n(x)$ , дальняя соответственно задается уравнением  $y = y_f(x)$

$$y_n(x) := \blacklozenge \quad y_f(x) := \blacklozenge .$$

На плоскости  $xu$  эти кривые ограничивают область  $Dxu$ , по которой производится интегрирование при вычислении потока ротора. Поместим  $Dxu$  в прямоугольник и найдем его размеры. По оси  $x$  это  $x_{min}$  и  $x_{max}$  - точки пересечения кривых  $y_n(x)$  и  $y_f(x)$

$$x_{min} := \blacklozenge \quad x_{max} := \blacklozenge .$$

По оси  $y$  это максимум на кривой  $y_f(x)$  и минимум на кривой  $y_n(x)$

$$y_{min} := \blacklozenge \quad y_{max} := \blacklozenge .$$

Зададим сетку интегрирования по области  $Dxu$ . Для этого задаем количество узлов сетки  $Nx$  и  $Ny$  вдоль оси  $x$  и  $y$

$$Nx := \blacklozenge \quad Ny := \blacklozenge .$$

Определяем шаги равномерной сетки на плоскости  $xu$

$$\Delta x := \frac{x_{max} - x_{min} + 0.02}{Nx}$$

$$\Delta y := \frac{y_{max} - y_{min} + 0.02}{Ny}$$

$$i := 0..Nx \quad j := 0..Ny$$

$$x_i := x_{min} - 0.01 + \Delta x \cdot i .$$

$$y_j := y_{min} - 0.01 + \Delta y \cdot j$$

Сетка может немного заходить за границы прямоугольника. Изобразим поверхность  $\Sigma$  и контур  $L$  на ней графически, с помощью 3D графики математического пакета. Это можно сделать разными способами.

Наиболее удобен параметрический метод задания поверхности. Для построения поверхности выберем в качестве параметров координаты точки  $x$  и  $y$  на плоскости  $xу$  и зададим три функции от двух аргументов  $x$  и  $y$

$$\begin{aligned}x\sigma(x, y) &:= x \\y\sigma(x, y) &:= y \\z\sigma(x, y) &:= \diamond\end{aligned} .$$

Заполним три матрицы значениями координат точек на поверхности

$$\begin{aligned}X_{i,j} &:= x\sigma(x_i, y_j) \\Y_{i,j} &:= y\sigma(x_i, y_j) \\Z_{i,j} &:= z\sigma(x_i, y_j)\end{aligned} .$$

Для получения контура  $L$  на  $\Sigma$  зададим три однопараметрические функции, выбрав в качестве параметра аргумент  $x$ . Дальняя часть контура задается уравнениями

$$\begin{aligned}x\sigma f(x) &:= x \\y\sigma f(x) &:= yf(x) \\z\sigma f(x) &:= z\sigma(x, yf(x))\end{aligned} .$$

Ближняя часть контура задается уравнениями

$$\begin{aligned}x\sigma n(x) &:= x \\y\sigma n(x) &:= yn(x) \\z\sigma n(x) &:= z\sigma(x, yn(x))\end{aligned} .$$

Заполним три соответствующих вектора на сетке значений параметра  $x$

$$\begin{aligned}XF_i &:= x\sigma f(x_i) \\YF_i &:= y\sigma f(x_i) \\ZF_i &:= z\sigma f(x_i) \\XN_i &:= x\sigma n(x_i) \\YN_i &:= y\sigma n(x_i) \\ZN_i &:= z\sigma n(x_i)\end{aligned} .$$

Заносим заполненные матрицы в окно 3D графопостроителя и получаем график исследуемой поверхности. Зададим характеристическую функцию области  $Dху$

$$\chi_{xy}(x, y) := \begin{cases} a \leftarrow 0 \\ a \leftarrow 1 \text{ if } y_n(x) \leq y \leq y_f(x) \\ a \end{cases}$$

Зададим вектор силы

$$F(x, y, z) := \begin{pmatrix} \blacklozenge \\ \blacklozenge \\ \blacklozenge \end{pmatrix}.$$

Найдем вектор ротора силы

$$(6.6) \quad \text{rot}(x, y, z) := \begin{pmatrix} \frac{d}{dy} F(x, y, z)_2 - \frac{d}{dz} F(x, y, z)_1 \\ \frac{d}{dz} F(x, y, z)_0 - \frac{d}{dx} F(x, y, z)_2 \\ \frac{d}{dx} F(x, y, z)_1 - \frac{d}{dy} F(x, y, z)_0 \end{pmatrix}.$$

Элемент касательной  $d\tau(x, y, z)$  в точке  $r(x, y, z)$  определим по формуле

$$d\tau f(x) = \left( i \cdot \frac{d}{dx} x\sigma f(x) + j \cdot \frac{d}{dx} y\sigma f(x) + k \cdot \frac{d}{dx} z\sigma f(x) \right) \cdot dx$$

$$d\tau n(x) = \left( i \cdot \frac{d}{dx} x\sigma n(x) + j \cdot \frac{d}{dx} y\sigma n(x) + k \cdot \frac{d}{dx} z\sigma n(x) \right) \cdot dx,$$

или вектор касательной к двум участкам контура получим в виде

$$\tau f(x) := \begin{pmatrix} \frac{d}{dx} x\sigma f(x) \\ \frac{d}{dx} y\sigma f(x) \\ \frac{d}{dx} z\sigma f(x) \end{pmatrix},$$

$$\tau n(x) := \begin{pmatrix} \frac{d}{dx} x\sigma n(x) \\ \frac{d}{dx} y\sigma n(x) \\ \frac{d}{dx} z\sigma n(x) \end{pmatrix}.$$

Криволинейный интеграл по некоторой кривой  $\Gamma$ , заданной параметрическим уравнением  $x = x(t)$ ,  $y = y(t)$ ,  $z = z(t)$ , где параметр  $t$  изменяется в интервале  $a \leq t \leq b$ , переписывается через определенный интеграл по параметру  $t$

$$\int_{\Gamma} (F(x, y, z), d\tau(x, y, z)) = \int_a^b (F(x(t), y(t), z(t)), \tau(x(t), y(t), z(t))) dt .$$

Кривая  $L$  разбита на два участка. Найдем интеграл по ближней кривой в положительном направлении

$$AN := \int_{x_{min}}^{x_{max}} F(x\sigma n(x), y\sigma n(x), z\sigma n(x)) \cdot \tau n(x) \cdot dx .$$

Найдем интеграл по дальней кривой в положительном направлении

$$AF := \int_{x_{max}}^{x_{min}} F(x\sigma f(x), y\sigma f(x), z\sigma f(x)) \cdot \tau f(x) \cdot dx .$$

Циркуляция вектора  $F(x, y, z)$  по кривой  $L$  равна

$$A := AF + AN .$$

Для вычисления потока вектора ротора от силы  $F(x, y, z)$  (6.6) зададим ненормированную нормаль к поверхности  $\Sigma$ , согласованную с направлением обхода кривой  $L$

$$NN\sigma(x, y) := \begin{pmatrix} -\frac{d}{dx} z\sigma(x, y) \\ -\frac{d}{dy} z\sigma(x, y) \\ 1 \end{pmatrix} .$$

Нормируем нормаль

$$N\sigma(x, y) := \frac{NN\sigma(x, y)}{|NN\sigma(x, y)|} .$$

Интеграл по поверхности  $\Sigma$  от ротора силы переписывается через двойной интеграл по области  $Dxy$

$$\iint_{\Sigma} (\text{rot}(F(x, y, z)), d\Sigma(x, y, z)) = \iint_{Dxy} RN(x, y, z\sigma(x, y)) \cdot \frac{dx \cdot dy}{|N\sigma(x, y)_2|} . \quad (6.7)$$

Здесь  $RN(x, y, z)$  - нормальная компонента вектора ротора силы

$$RN(x, y, z) := \text{rot}(x, y, z) \cdot N\sigma(x, y) ,$$

соотношение  $\frac{dx \cdot dy}{|N\sigma(x, y)_2|}$  - есть проекция элемента поверхности  $d\Sigma(x, y, z)$

на координатную плоскость  $xy$ . Двойной интеграл в (6.7) найдем численно по “прямоугольникам”, заменив интеграл интегральной суммой

$$A\Sigma := \sum_{i=0}^{Nx} \sum_{j=0}^{Ny} RN(x_i, y_j, z\sigma(x_i, y_j)) \cdot \chi_{xy}(x_i, y_j) \cdot \frac{\Delta x \cdot \Delta y}{|N\sigma(x_i, y_j)_2|} .$$

$$A =$$

$$A\Sigma =$$

### Задания к работе

1. Выбрать исходные данные для проверки формул Грина-Римана и построить пучок кривых на плоскости  $xu$ , проходящих через две заданные точки.
2. Численно проверить теорему о независимости криволинейного интеграла от формы кривой, первую и вторую формулы Грина-Римана. В каких случаях эти теоремы могут нарушаться?
3. Привести примеры использования формулы Стокса в задачах оптики, в применении к электромагнитным явлениям.
4. Выбрать исходные данные: уравнения границ  $u_n(x)$  и  $u_f(x)$  на плоскости  $xu$ , уравнение  $z\sigma(x, y)$  поверхности. Изобразить область интегрирования  $Dxu$  на плоскости  $xu$  и вписать ее в прямоугольник. Создать сетку узлов на плоскости  $xu$ . Записать имена соответствующих матриц и векторов в графический редактор, построить с помощью 3D графопостроителя поверхность с контуром и научиться управлять графиком с помощью окна графопостроителя.
5. Задать нормаль к поверхностям. Объяснить формулы и дать комментарий к выбору знаков.
6. Задать касательные к контурам, вектор силы и его ротор. Как можно изменять вектор силы, чтобы величина работы не изменялась? Проверить численно.
7. Для какой цели введена характеристическая функция области  $Dxu$ ?
8. Рассчитать циркуляцию и поток вектора ротора силы и оценить погрешность вычисления.



## 7. Теорема Остроградского – Гаусса

Математическая теорема Остроградского – Гаусса широко применяется в физике и технике для формулировки различных законов сохранения. Рассмотрим стационарный поток жидкости и охарактеризуем его вектором плотности потока массы  $S(x, y, z)$ , имеющем размерность  $g/(сек \cdot см^2)$ . Плотность потока задана в каждой точке  $r(x, y, z)$  в момент времени  $t$ . Выделим внутри потока объем  $V$  с поверхностью  $\Sigma$ , ориентированной внешней единичной нормалью к поверхности  $n(x, y, z)$ . Тогда количество жидкости (массы)  $d\Phi(x, y, z)$  (размерность  $g/сек$ ), протекающее через ориентированный элемент поверхности  $d\Sigma(x, y, z)$  за секунду (поток через  $d\Sigma(x, y, z)$ ) равен

$$d\Phi(x, y, z) = (S(x, y, z), d\Sigma(x, y, z)).$$

Элемент поверхности  $d\Sigma(x, y, z)$  ориентирован по внешней нормали  $n(x, y, z)$  и расположен в точке  $r(x, y, z)$  на поверхности  $\Sigma$ . Знак элемента потока  $d\Phi(x, y, z)$  имеет смысл: положительный – в данной точке жидкость вытекает из элемента поверхности, отрицательный – в данной точке жидкость втекает внутрь  $d\Sigma(x, y, z)$ . Разобьем всю замкнутую поверхность  $\Sigma$  на малые элементы, просуммируем все элементарные потоки по поверхности, получим полный поток  $\Phi$  через замкнутую поверхность. Смысл потока – полная масса жидкости, вытекающей и втекающей за секунду через всю замкнутую поверхность (с учетом знака). Очевидно,  $\Phi$  вычисляется с помощью поверхностного интеграла

$$\Phi = \iint_{\Sigma} (S(x, y, z), d\Sigma(x, y, z)). \quad (7.1)$$

Вектор  $S(x, y, z)$  направлен по касательной в каждой точке к линии тока. Если жидкость создается источником, который расположен вне рассматриваемого замкнутого объема, а внутри объема отсутствуют стоки (точки, где жидкость исчезает) и источники (точки, где жидкость может возникнуть), то поток  $\Phi$  такой жидкости через  $\Sigma$  равен нулю. Количество втекающей в объем и вытекающей из объема жидкости за секунду одинаково. Если же внутри выделенного объема имеются источники жидкости или ее стоки, то поток будет не ноль. Количество жидкости, возникающей (или исчезающей, в зависимости от знака) в единице объема в окрестности точки  $r(x, y, z)$  определяется дивергенцией вектора плотности потока

$$\operatorname{div}(\mathbf{S}(x, y, z)) = \frac{\partial(S)_x}{\partial x} + \frac{\partial(S)_y}{\partial y} + \frac{\partial(S)_z}{\partial z}.$$

Размерность дивергенции  $z/(\text{сек} \cdot \text{см}^3)$ . Исходя из смысла дивергенции, количество жидкости  $\Phi V$ , исчезающее (или возникающее, в зависимости от знака) во всем объеме за секунду, определяется объемным интегралом

$$\Phi V = \iiint_V \operatorname{div}(\mathbf{S}(x, y, z)) \cdot dV.$$

Очевидно, если все причины исчезновения или рождения жидкости учтены, то в стационарном потоке жидкости выполнено равенство  $\Phi = \Phi V$ , или имеет место теорема Остроградского – Гаусса [3]

$$\iint_{\Sigma} (\mathbf{S}(x, y, z), d\Sigma(x, y, z)) = \iiint_V \operatorname{div}(\mathbf{S}(x, y, z)) \cdot dV. \quad (7.2)$$

Похожие рассуждения можно провести для потока энергии, потока силовых линий электрического поля и так далее.

Проверим формулу (7.2) численными методами. Для этого сконструируем тело объема  $V$ , для простоты, с цилиндрической боковой поверхностью с образующей, параллельной оси  $z$ . Обозначим:  $SF_y$  - дальняя боковая поверхность (по отношению к положительному направлению оси  $y$ ) и  $SN_y$  - ближняя боковая поверхность. Спроектируем тело на координатную плоскость  $xu$  и зададим границы полученной области  $Dxu$  с помощью двух кривых  $yn(x)$  и  $yf(x)$ , пересекающихся в двух крайних, если смотреть вдоль оси  $x$ , точках  $xmin$ ,  $xmax$

$$yn(x) := \blacklozenge \quad yf(x) := \blacklozenge.$$

По области  $Dxu$  далее будет производиться интегрирование для вычисления потока через нижнюю  $SBz$  (по отношению к положительному направлению оси  $z$ ) и верхнюю  $STz$  поверхности выделенного объема  $V$ . Назовем  $yf(x)$  - дальней границей,  $yn(x)$  - ближней границей области на плоскости  $xu$ . Впишем область  $Dxu$  в прямоугольник с границами  $xmin$ ,  $xmax$  вдоль оси  $x$  и с границами  $ymin$ ,  $ymax$  вдоль оси  $y$ . Очевидно  $xmin$ ,  $xmax$  - это точки пересечения границ  $yf(x)$  и  $yn(x)$  и их можно найти, например, с помощью программы `root(,,)` математического пакета

$$xa := \blacklozenge$$

$$xmin := \text{root}(yf(xa) - yn(xa), xa)'$$

$$xb := \blacklozenge$$

$$xmax := \text{root}(yf(xb) - yn(xb), xb)'$$

Здесь  $x_a$ ,  $x_b$  - стартовые значения аргумента. Точка  $y_{max}$  является максимумом функции  $y_f(x)$ , а точка  $y_{min}$  - это точка минимума функции  $y_n(x)$ . Все эти параметры необходимо найти с высокой точностью

$$y_{min} := \blacklozenge \quad y_{max} := \blacklozenge \quad .$$

Для удобства интегрирования введем характеристическую функцию области  $D_{xy}$ . Эта функция  $\chi_{xy}(x, y)$  равна 1, если точка лежит на  $D_{xy}$  и равна 0 в противоположном случае. С помощью условных операторов это условие можно, например, записать так

$$\chi_{xy}(x, y) := \begin{cases} a \leftarrow 0 \\ a \leftarrow 1 \text{ if } y_n(x) \leq y \leq y_f(x) . \\ a \end{cases}$$

Зададим сетку интегрирования на плоскости  $xy$ . Для этого задаем количество узлов сетки  $N_x$  и  $N_y$  вдоль оси  $x$  и  $y$

$$N_x := \blacklozenge \quad N_y := \blacklozenge \quad .$$

Определяем шаги равномерной сетки на плоскости  $xy$

$$\Delta x := \frac{x_{max} - x_{min} + 0.02}{N_x}$$

$$\Delta y := \frac{y_{max} - y_{min} + 0.02}{N_y}$$

$$i := 0..N_x \quad j := 0..N_y$$

$$x_i := x_{min} - 0.01 + \Delta x \cdot i \quad .$$

$$y_j := y_{min} - 0.01 + \Delta y \cdot j$$

Сетка может немного заходить за границы прямоугольника. Изобразим тело графически, с помощью 3D графики математического пакета. Это можно сделать разными способами. Наиболее удобен параметрический метод задания поверхности. Для построения верхней и нижней поверхностей выберем в качестве параметров координаты точки  $x$  и  $y$  на плоскости  $xy$  и зададим три функции от двух аргументов  $x$  и  $y$ . Для нижней поверхности

$$x_b(x, y) := x$$

$$y_b(x, y) := y \quad .$$

$$z_b(x, y) := \blacklozenge$$

Для верхней поверхности

$$\begin{aligned} xt(x, y) &:= x \\ yt(x, y) &:= y \\ zt(x, y) &:= \diamond \end{aligned} .$$

Заполним три матрицы значениями координат для нижней поверхности

$$\begin{aligned} XB_{i,j} &:= xb(x_i, y_j) \\ YB_{i,j} &:= yb(x_i, y_j) , \\ ZB_{i,j} &:= zb(x_i, y_j) \end{aligned}$$

И для верхней поверхности

$$\begin{aligned} XT_{i,j} &:= xt(x_i, y_j) \\ YT_{i,j} &:= yt(x_i, y_j) . \\ ZT_{i,j} &:= zt(x_i, y_j) \end{aligned}$$

Удобно по границам между верхней и нижней поверхностями и боковой поверхностью провести контуры. Для этого задаются три однопараметрические функции. Граница между дальней боковой поверхностью и верхней поверхностью

$$\begin{aligned} xtf(x) &:= x \\ ytf(x) &:= yf(x) \\ ztf(x) &:= zt(x, yf(x)) \end{aligned} .$$

Граница между ближней боковой поверхностью и верхней поверхностью

$$\begin{aligned} xtn(x) &:= x \\ ytn(x) &:= yn(x) \\ ztn(x) &:= zt(x, yn(x)) \end{aligned} .$$

Граница между дальней боковой поверхностью и нижней поверхностью

$$\begin{aligned} xbf(x) &:= x \\ ybf(x) &:= yf(x) \\ zbf(x) &:= zb(x, yf(x)) \end{aligned} .$$

Граница между ближней боковой поверхностью и нижней поверхностью

$$\begin{aligned} xbn(x) &:= x \\ ybn(x) &:= yn(x) \\ zbn(x) &:= zb(x, yn(x)) \end{aligned} .$$

Заполним три соответствующих вектора на сетке значений параметра (в данном случае параметр – это координата  $x$  точки)

$$\begin{aligned}
XTF_i &:= xtf(x_i) \\
YTF_i &:= ytf(x_i) \quad , \\
ZTF_i &:= ztf(x_i) \\
XTN_i &:= xtn(x_i) \\
YTN_i &:= ytn(x_i) \quad , \\
ZTN_i &:= ztn(x_i) \\
XBF_i &:= xbf(x_i) \\
YBF_i &:= ybf(x_i) \quad , \\
ZBF_i &:= zbf(x_i) \\
XBN_i &:= xbn(x_i) \\
YBN_i &:= ybn(x_i) \quad . \\
ZBN_i &:= zbn(x_i)
\end{aligned}$$

Заносим заполненные матрицы в окно 3D графопостроителя и получаем график исследуемого тела. Для вычисления потоков через боковые поверхности спроектируем дальнюю и ближнюю поверхности на плоскость  $xz$ . Изобразим эти области  $DNxz$  и  $DFxz$  графически. Для этого изобразим проекции: верхней дальней боковой границы (параметрическое уравнение  $z = ztf(x)$ ,  $x = xtf(x)$ ), верхней ближней боковой (параметрическое уравнение  $z = ztn(x)$ ,  $x = xtn(x)$ ), нижней дальней боковой (параметрическое уравнение  $z = zbf(x)$ ,  $x = xbf(x)$ ), нижней ближней боковой (параметрическое уравнение  $z = zbn(x)$ ,  $x = xbn(x)$ ). Найдем максимум на кривой  $z = ztf(x)$  и минимум на кривой  $z = zbf(x)$

$$z_{fmax} := \blacklozenge \quad z_{fmin} := \blacklozenge \quad .$$

Найдем максимум на кривой  $z = ztn(x)$  и минимум на кривой  $z = zbn(x)$

$$z_{nmax} := \blacklozenge \quad z_{nmin} := \blacklozenge \quad .$$

Получим сетки интегрирования  $zn_k$  и  $zf_k$  по оси  $z$  по областям  $DNxz$  и  $DFxz$

$$\begin{aligned}
Nzn &:= \blacklozenge \quad Nzf := \blacklozenge \\
\Delta zn &:= \frac{z_{nmax} - z_{nmin} + 0.02}{Nzn} \\
\Delta zf &:= \frac{z_{fmax} - z_{fmin} + 0.02}{Nzf} \\
kn &:= 0..Nzn \quad kf := 0..Nzf
\end{aligned}$$

$$zn_{kn} := znmin + \Delta zn \cdot kn - 0.01$$

$$zf_{kf} := zfmin + \Delta zf \cdot kf - 0.01$$

Для удобства интегрирования введем характеристические функции  $\chi_{nxz}(x, z)$  и  $\chi_{fxz}(x, z)$  областей  $DNxz$  и  $DFxz$

$$\chi_{nxz}(x, z) := \begin{cases} a \leftarrow 0 \\ a \leftarrow 1 \text{ if } zbn(x) \leq z \leq ztn(x), \\ a \end{cases}$$

$$\chi_{fxz}(x, z) := \begin{cases} a \leftarrow 0 \\ a \leftarrow 1 \text{ if } zbf(x) \leq z \leq ztf(x). \\ a \end{cases}$$

Для расчета потоков необходимо ориентировать поверхности внешней нормалью. Зададим внешние нормали. Ненормированная внешняя нормаль к верхней поверхности

$$NNt(x, y) := \begin{pmatrix} -\frac{d}{dx}zt(x, y) \\ -\frac{d}{dy}zt(x, y) \\ 1 \end{pmatrix},$$

ненормированная внешняя нормаль к нижней поверхности

$$NNb(x, y) := \begin{pmatrix} \frac{d}{dx}zb(x, y) \\ \frac{d}{dy}zb(x, y) \\ -1 \end{pmatrix},$$

ненормированная внешняя нормаль к дальней боковой поверхности

$$NNf(x) := \begin{pmatrix} -\frac{d}{dx}yf(x) \\ 1 \\ 0 \end{pmatrix},$$

ненормированная внешняя нормаль к ближней боковой поверхности

$$\text{NNn}(x) := \begin{pmatrix} \frac{d}{dx} y_n(x) \\ -1 \\ 0 \end{pmatrix}.$$

Соответствующие нормированные (единичные) нормали

$$\text{Nt}(x, y) := \frac{\text{NNt}(x, y)}{|\text{NNt}(x, y)|},$$

$$\text{Nb}(x, y) := \frac{\text{NNb}(x, y)}{|\text{NNb}(x, y)|},$$

$$\text{Nf}(x) := \frac{\text{NNf}(x)}{|\text{NNf}(x)|},$$

$$\text{Nn}(x) := \frac{\text{NNn}(x)}{|\text{NNn}(x)|}.$$

Зададим вектор плотности потока

$$\mathbf{S}(x, y, z) := \begin{pmatrix} \blacklozenge \\ \blacklozenge \\ \blacklozenge \end{pmatrix}.$$

Найдем нормальные компоненты вектора плотности потока, скалярно умножив соответствующую единичную нормаль  $\text{Nt}(x, y)$ ,  $\text{Nb}(x, y)$ ,  $\text{Nf}(x)$ ,  $\text{Nn}(x)$  на вектор  $\mathbf{S}(x, y, z)$

$$WT(x, y, z) := \mathbf{S}(x, y, z) \cdot \text{Nt}(x, y)$$

$$WB(x, y, z) := \mathbf{S}(x, y, z) \cdot \text{Nb}(x, y)$$

$$WN(x, y, z) := \mathbf{S}(x, y, z) \cdot \text{Nn}(x)$$

$$WF(x, y, z) := \mathbf{S}(x, y, z) \cdot \text{Nf}(x)$$

Поверхностный интеграл (7.1) вычислим, выразив его через двойной. Поток через верхнюю поверхность  $\Phi T$

$$\Phi T = \iint_{STz} (\mathbf{S}(x, y, z), d\Sigma(x, y, z)) = \iint_{Dxy} WT(x, y, zt(x, y)) \cdot \frac{dx \cdot dy}{|\text{Nt}(x, y)_2|},$$

поток через нижнюю поверхность

$$\Phi B = \iint_{SBz} (\mathbf{S}(x, y, z), d\Sigma(x, y, z)) = \iint_{Dxy} WB(x, y, zb(x, y)) \cdot \frac{dx \cdot dy}{|\text{Nb}(x, y)_2|},$$

поток через дальнюю боковую поверхность

$$\Phi F = \iint_{SFy} (\mathbf{S}(x, y, z), d\Sigma(x, y, z)) = \iint_{DFxz} WF(x, yf(x), z) \cdot \frac{dx \cdot dz}{|\mathbf{Nf}(x)_1|},$$

поток через ближнюю боковую поверхность

$$\Phi N = \iint_{SNy} (\mathbf{S}(x, y, z), d\Sigma(x, y, z)) = \iint_{DNxz} WB(x, yn(x), z) \cdot \frac{dx \cdot dz}{|\mathbf{Nn}(x)_1|}$$

Двойной интеграл вычисляем по соответствующим областям  $Dxy$ ,  $DNxz$ ,  $DFxz$  по “прямоугольникам”, заменив интеграл двойной суммой

$$\Phi T := \sum_{i=0}^{Nx} \sum_{j=0}^{Ny} WT(x_i, y_j, zt(x_i, y_j)) \cdot \chi_{xy}(x_i, y_j) \cdot \frac{\Delta x \cdot \Delta y}{|\mathbf{Nt}(x_i, y_j)_2|},$$

$$\Phi B := \sum_{i=0}^{Nx} \sum_{j=0}^{Ny} WB(x_i, y_j, zb(x_i, y_j)) \cdot \chi_{xy}(x_i, y_j) \cdot \frac{\Delta x \cdot \Delta y}{|\mathbf{Nb}(x_i, y_j)_2|},$$

$$\Phi F := \sum_{i=0}^{Nx} \sum_{kf=0}^{Nzf} WF(x_i, yf(x_i), zf_{kf}) \cdot \chi_{fz}(x_i, zf_{kf}) \cdot \frac{\Delta x \cdot \Delta z f}{|\mathbf{Nf}(x_i)_1|},$$

$$\Phi N := \sum_{i=0}^{Nx} \sum_{kn=0}^{Nzn} WN(x_i, yn(x_i), zn_{kn}) \cdot \chi_{nz}(x_i, zn_{kn}) \cdot \frac{\Delta x \cdot \Delta z n}{|\mathbf{Nn}(x_i)_1|}.$$

Полный поток

$$\Phi := \Phi B + \Phi T + \Phi N + \Phi F.$$

Рассчитаем правую часть формулы (7.2). Найдем дивергенцию вектора плотности потока

$$\operatorname{div}(x, y, z) := \frac{d}{dx} S(x, y, z)_0 + \frac{d}{dy} S(x, y, z)_1 + \frac{d}{dz} S(x, y, z)_2.$$

Зададим характеристическую функцию  $\chi(x, y, z)$  объема

$$\chi(x, y, z) := \begin{cases} a \leftarrow 0 \\ a \leftarrow 1 \text{ if } zb(x, y) \leq z \leq zt(x, y). \\ a \end{cases}$$

Найдем минимальное  $zmin$  и максимальное  $zmax$  значение координаты  $z$  по телу с помощью средств математического пакета

$$zmax := \max(ZT)$$

$$zmin := \min(ZB)$$

Зададим сетку интегрирования вдоль оси  $z$  для всего тела  $V$

$$Nz := \blacklozenge$$

$$\Delta z := \frac{zmax - zmin + 0.02}{Nz}.$$

$$k := 0..Nz$$

$$z_k := zmin + \Delta z \cdot k - 0.01$$



Мощность всех источников в объеме  $V$  найдем численно по “прямоугольникам”, заменив тройной интеграл (7.2) тройной суммой по сетке

$$\Phi V := \sum_{i=0}^{N_x} \sum_{j=0}^{N_y} \sum_{k=0}^{N_z} \operatorname{div}(x_i, y_j, z_k) \cdot \chi(x_i, y_j, z_k) \cdot \chi_{xy}(x_i, y_j) \cdot \Delta x \cdot \Delta y \cdot \Delta z .$$

Сравним  $\Phi$  и  $\Phi V$  и сделаем вывод о справедливости теоремы Остроградского – Гаусса.

### Задания к работе

1. Привести примеры использования формулы Остроградского – Гаусса в задачах оптики, в применении к электромагнитным явлениям.
2. Выбрать исходные данные: уравнения границ  $yn(x)$  и  $yf(x)$  на плоскости  $xu$ , уравнения верхней  $zt(x,y)$  и нижней  $zb(x,y)$  поверхности. Изобразить область интегрирования  $Dxu$  на плоскости  $xu$  и вписать ее в прямоугольник. Создать сетку узлов на плоскости  $xu$ . Записать имена соответствующих матриц и векторов в графический редактор, построить с помощью 3D графопостроителя тело и научиться управлять графиком с помощью окна графопостроителя математического пакета.
3. Изобразить области интегрирования  $DNxz$  и  $DFxz$  на плоскости  $xz$ , вписать их в соответствующие прямоугольники и задать сетку узлов по оси  $z$  для этих областей.
4. Задать внешние нормали к поверхностям. Объяснить формулы и дать комментарий к выбору знаков.
5. Задать вектор плотности потока и его дивергенцию. Как можно изменить вектор плотности потока, чтобы величина потока не изменилась? Проверить численно.
6. Для какой цели введены характеристические функции областей?
7. Рассчитать потоки и мощность источников и оценить погрешность вычисления.

## 8. Численное решение систем обыкновенных дифференциальных уравнений и исследование особых решений

### 8.1. Схема Рунге – Кутта

Численные методы построения решения систем обыкновенных дифференциальных уравнений на сетке узлов весьма разнообразны. Изучим возможности наиболее распространенного и вместе с тем простого метода Рунге – Кутта [1], [2], [5], [7], [9], [10]. Общая идея построения приближенного одношагового метода состоит в следующем. Рассмотрим задачу Коши для системы дифференциальных уравнений первого порядка

$$\begin{aligned} \frac{d}{dt}y(t) &= F(t, y(t)) \\ y(t)|_{t=t_0} &= Y_0 \end{aligned} \quad (8.1)$$

Здесь  $y(t) = \{y(t)_0, y(t)_1, \dots, y(t)_{n-1}\}$  -  $n$  – мерный вектор, зависящий от вещественной переменной  $t$ . Правая часть  $F(t, y(t))$  -  $n$  – мерный вектор, зависящий от  $n+1$  переменной  $(t, x_0, x_1, \dots, x_{n-1})$ . Если правая часть не зависит явно от аргумента  $t$ , то такая система называется автономной. Зададим сетку узлов аргумента  $t$  на интервале  $t_0 \leq t \leq t_1$ , на котором необходимо решение и для которого выполнены условия существования и единственности решения системы (8.1). Для простоты сетка выбирается равномерной с шагом  $h$ , число узлов сетки  $N$

$$\begin{aligned} (8.2) \quad t_0 &:= \blacklozenge \quad t_1 := \blacklozenge \\ N &:= \blacklozenge \\ h &:= \frac{t_1 - t_0}{N - 1} \\ k &:= 0..N - 1 \\ t_k &:= t_0 + h \cdot k \end{aligned}$$

Рассмотрим интервал изменения аргумента шириной в один шаг сетки, начинающийся в произвольной точке  $t$  и заканчивающийся в точке  $t + h$ . Проинтегрируем левую часть уравнения (8.1) по этому интервалу, получим соотношение

$$y(t + h) = y(t) + \int_0^h \frac{d}{d\tau} y(\tau + t) \cdot d\tau. \quad (8.3)$$

Для построения приближенного метода необходимо оценить правую часть (8.3). Возьмем интеграл приближенно, методом “прямоугольников” (без центральной точки). Получаем соотношение

$$y(t+h) = y(t) + h \cdot \frac{d}{dt} y(t) + O(h^2). \quad (8.4)$$

Подставим (8.1) в (8.4), получим

$$y(t+h) = y(t) + h \cdot F(t, y(t)) + O(h^2).$$

Отбрасывая слагаемые порядка  $h^2$ , получаем приближенный метод Эйлера решения системы уравнений. Обозначим  $y_k$  - значение функции  $y(t)$  на  $k$  узле

$$y_k = y(t_k).$$

Получаем рекуррентное соотношение метода Эйлера

$$y_{k+1} = y_k + h \cdot F(t_k, y_k) \quad k = 0, 1 \dots N-1.$$

Уточним метод, взяв интеграл (8.3) по “трапециям”. Получаем

$$y(t+h) = y(t) + \frac{h}{2} \cdot (F(t, y(t)) + F(t+h, y(t+h))) + O(h^3).$$

Неявная формула метода (Адамар) имеет вид

$$y_{k+1} = y_k + \frac{h}{2} \cdot (F(x_k, y_k) + F(x_{k+1}, y_{k+1})) \quad k = 0, 1 \dots N-1. \quad (8.5)$$

Для того чтобы сделать схему явной, заменим правую часть формулы (8.5) приближенным выражением  $y_{k+1} \approx y_{k+1}^1$ , которое, тем не менее, обеспечило бы требуемую точность  $O(h^3)$  на каждом шаге вычислений.

Очевидно,  $y_{k+1}^1$  достаточно найти с точностью  $O(h^2)$ , тогда требуемая точность будет обеспечена. Такую точность как раз обеспечивает метод Эйлера. Получаем метод Рунге – Кутта с одной итерацией (один из семейства методов Рунге – Кутта). Реализуем метод

$$(8.6) \quad RK(t_0, Y_0, F, N, h, t) := \begin{array}{l} y_0 \leftarrow Y_0 \\ \text{for } n \in 0..N-2 \\ \quad \left| \begin{array}{l} y_{n+1}^1 \leftarrow y_n + h \cdot F(t_n, y_n) \\ y_{n+1} \leftarrow y_n + \frac{h}{2} \cdot (F(t_n, y_n) + F(t_{n+1}, y_{n+1}^1)) \end{array} \right. \\ y \end{array} .$$

В качестве аргументов  $y$  функции  $RK(t_0, Y_0, F, N, h, t)$  выбраны все необходимые для запуска программы величины. Точность приближенного решения, полученного по схеме Рунге – Кутта с одной итерацией на всем интервале  $O(h^2)$ .

Проверим формулу для точности интегрирования по схеме (8.6). Для этого, например, для системы двух дифференциальных уравнений первого порядка введем вектор правой части такой, чтобы задача Коши имела бы точное решение

$$\begin{aligned} P(t,y) &:= \diamond & Q(t,y) &:= \diamond \\ F(t,y) &:= \begin{pmatrix} P(t,y) \\ Q(t,y) \end{pmatrix} \end{aligned} .$$

Зададим начальные данные задачи Коши

$$Y_0 := \begin{pmatrix} \diamond \\ \diamond \end{pmatrix} .$$

Решим задачу Коши точно и численно, по схеме (8.6) на сетке (8.2). Построим графики обоих решений и численно проверим справедливость формулы для оценки погрешности метода (например, изменив шаг сетки в два раза).

## 8.2. Устойчивость точек покоя автономной системы двух дифференциальных уравнений первого порядка

Точки покоя являются особыми решениями систем дифференциальных уравнений. Рассмотрим систему двух *автономных* уравнений первого порядка

$$\begin{cases} \frac{d}{dt} p(t) = P(p,q) \\ \frac{d}{dt} q(t) = Q(p,q) \end{cases} . \quad (8.7)$$

Зададим правые части системы

$$(8.8) \quad P(p,q) := \diamond \quad Q(p,q) := \diamond .$$

Точки покоя удовлетворяют системе двух уравнений

$$\begin{cases} P(p,q) = 0 \\ Q(p,q) = 0 \end{cases} . \quad (8.9)$$

Найдем решения этой системы двух уравнений численно. Это решение можно получить приближенно, например, графически, построив с помощью графопостроителя математического пакета линии уровня двух функций. Для надежности работы графопостроителя надо сгенерировать сетку на плоскости  $pq$ , задав размеры прямоугольника и число узлов по осям

$$\begin{aligned}
pmin &:= \blacklozenge & qmin &:= \blacklozenge \\
pmax &:= \blacklozenge & qmax &:= \blacklozenge \\
np &:= \blacklozenge & nq &:= \blacklozenge
\end{aligned}$$

Значения функций на сетке генерируем с помощью программы

$$FQ := CreateMesh(Q, pmin, pmax, qmin, qmax, np, nq)$$

$$FP := CreateMesh(P, pmin, pmax, qmin, qmax, np, nq)$$

Найдем точки пересечения линий, на которых функции равны нулю. Выберем одну точку приближенно, по графику. Приближенные координаты точки покоя

$$ps := \blacklozenge \quad qs := \blacklozenge$$

Запишем эти координаты в вектор  $ys$

$$ys := \begin{pmatrix} ps \\ qs \end{pmatrix}$$

Линеаризуем систему уравнений (8.9) в окрестности выбранной точки. Для этого разложим правые части (8.7) в ряд Тейлора и ограничимся линейными по смещению из  $ys$  слагаемыми

$$P(p, q) \approx B_0 + M(ps, qs)_{0,0} \cdot \Delta p + M(ps, qs)_{0,1} \cdot \Delta q$$

$$Q(p, q) \approx B_1 + M(ps, qs)_{1,0} \cdot \Delta p + M(ps, qs)_{1,1} \cdot \Delta q$$

Здесь введены элементы матрицы

$$M(p, q) = \begin{pmatrix} \frac{\partial}{\partial p} P(p, q) & \frac{\partial}{\partial q} P(p, q) \\ \frac{\partial}{\partial p} Q(p, q) & \frac{\partial}{\partial q} Q(p, q) \end{pmatrix}$$

и вектора

$$B := \begin{pmatrix} P(ys_0, ys_1) \\ Q(ys_0, ys_1) \end{pmatrix}$$

Возьмем производные “вручную” и заполним матрицу  $M(p, q)$

$$M(p, q) := \blacklozenge$$

Уточненное значение точки покоя (решение системы уравнений (8.9)) получим с помощью соотношения

$$ys := ys - M^{-1}(ys_0, ys_1) \cdot B$$

Совершим несколько итераций, так, чтобы установился четвертый знак после запятой. Окончательно особое решение системы имеет вид

$$\begin{cases} p(t) = y_{s_0} \\ q(t) = y_{s_1} \end{cases}$$

Исследуем тип точки покоя с помощью *метода линеаризации*. Найдем собственные числа и собственные векторы матрицы  $Ms = M(y_{s_0}, y_{s_1})$  средствами математического пакета

$$Ms := M(y_{s_0}, y_{s_1})$$

$$\lambda := \text{eigenvals}(Ms)$$

$$U := \text{eigenvecs}(Ms)$$

По виду собственных чисел делаем вывод о типе точки покоя: устойчивый узел, неустойчивый узел, седло, устойчивый фокус, неустойчивый фокус, центр и так далее. Построим линеаризованное решение задачи Коши для выбранной точки покоя для выбранных начальных данных. Представим вектор решения  $Y(t)$  в виде суммы

$$(8.10) \quad \begin{aligned} \Delta Y0 &:= \diamond \\ Y0 &:= y_s + \Delta Y0 \\ \alpha &:= U^{-1} \cdot \Delta Y0 \end{aligned}$$

$$(8.11) \quad y(t) := Y0 + \alpha_0 \cdot U^{(0)} \cdot \exp(\lambda_0 \cdot t) + \alpha_1 \cdot U^{(1)} \cdot \exp(\lambda_1 \cdot t).$$

Здесь  $\Delta Y0$  - вектор смещения из точки покоя,  $Y0$  - вектор начальных данных,  $\alpha$  - вектор смещения в базисе  $U^{(0)}$ ,  $U^{(1)}$ . Начальный момент времени выбран равным нулю, это не ограничивает общности из-за автономности системы. Получим решение поставленной задачи Коши численно. Для этого создадим схему Рунге – Кутты в упрощенном варианте, для автономной системы

$$(8.12) \quad RKa(Y0, F, N, h) := \begin{array}{l} \left| \begin{array}{l} y_0 \leftarrow Y0 \\ \text{for } n \in 0..N-2 \\ \left| \begin{array}{l} y1_{n+1} \leftarrow y_n + h \cdot F(y_n) \\ y_{n+1} \leftarrow y_n + \frac{h}{2} \cdot (F(y_n) + F(y1_{n+1})) \end{array} \right. \\ \left. \right. \end{array} \right. \\ y \end{array}$$

Для запуска схемы введем правую часть

$$Fa(Y) := \begin{pmatrix} P(Y_0, Y_1) \\ Q(Y_0, Y_1) \end{pmatrix},$$

шаг и количество шагов

$$h := \diamond \quad N := \diamond$$

Получим точное (численное) решение  $y_e$

$$y_e := RKa(Y0, Fa, N, h).$$

Введем сетку узлов с шагом  $h$  и построим график на фазовой плоскости точного решения  $y_e$  и приближенного (8.11), полученного методом линеаризации. Удобно в одном графическом окне построить несколько графиков для разных начальных данных. Сравнение позволит сделать вывод о правильности определения типа точки покоя.

### 8.3. Устойчивость предельного цикла

В приложениях широко используется задача о колебательном движении одной степени свободы системы, которая описывается двумя автономными дифференциальными уравнениями первого порядка [13]. Здесь простейшим является приближение гармонических колебаний, применимое тогда, когда возвращающая сила линейна по смещению из положения равновесия. В специально выбранных единицах измерения система уравнений имеет вид

$$\begin{aligned} \frac{d}{dt}y &= Fl(y) \\ y &= \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} \end{aligned} \quad (8.13)$$

Здесь компонента  $y_0$  имеет смысл смещения из положения равновесия, компонента  $y_1$  определяет скорость изменения смещения. Вектор правой части имеет вид

$$Fl(y) = \begin{pmatrix} y_1 \\ -\omega^2 \cdot y_0 \end{pmatrix}.$$

Здесь компонента  $Fl(y)_1 = -\omega^2 \cdot y_0$  равна возвращающей силе, а  $\omega$  равна частоте колебаний (в выбранных единицах). Общее решение системы (8.13) известно

$$y(t) = \begin{pmatrix} A \cdot \cos(\omega \cdot t - \varphi) \\ -\omega \cdot A \cdot \sin(\omega \cdot t - \varphi) \end{pmatrix}. \quad (8.14)$$

Здесь  $A, \varphi$  - произвольные, не зависящие от времени амплитуда и начальная фаза гармонических колебаний. Для решения задачи Коши можно задать начальную амплитуду  $A = A0$  и фазу  $\varphi = \varphi0$ . Тогда начальный (в момент  $t0 = 0$ ) вектор  $y0$  равен

$$(8.15) \quad y0(A0, \varphi0) := \begin{pmatrix} A0 \cdot \cos(\varphi0) \\ \omega \cdot A0 \cdot \sin(\varphi0) \end{pmatrix}.$$

Для начальных данных (8.15) решение (8.14) имеет вид

$$y(t) = \begin{pmatrix} A_0 \cdot \cos(\omega \cdot t - \varphi_0) \\ -\omega \cdot A_0 \cdot \sin(\omega \cdot t - \varphi_0) \end{pmatrix}.$$

Благодаря автономности системы (8.13), имеется интеграл движения – полная энергия системы

$$E = \omega^2 \cdot y(t)_0^2 + y(t)_1^2 = \omega^2 \cdot A_0^2.$$

Как следствие, траектории в фазовом пространстве замкнуты, имеют форму эллипса с полуосями  $\{A_0, \omega \cdot A_0\}$ . Добавим к возвращающей силе нелинейное слагаемое  $\mu \cdot v(y)$ , величина которого управляется (малым) параметром  $\mu$

$$\begin{aligned} \mu &:= \diamond, \\ v(y) &:= \diamond. \end{aligned}$$

Тогда правая часть системы (8.13) с учетом нелинейности запишется

$$(8.16) \quad FN(y) := \begin{pmatrix} y_1 \\ -\omega^2 \cdot y_0 + \mu \cdot v(y) \end{pmatrix}.$$

Уравнение нелинейных колебаний имеет вид

$$\frac{d}{dt} y = FN(y). \quad (8.17)$$

Добавка нелинейности может существенно изменить характер колебательного процесса. У системы уравнений могут появиться особые решения – устойчивые или не устойчивые *предельные циклы*. Малостью параметра нелинейности  $\mu < 1$  можно воспользоваться, чтобы найти приближенные решения системы. Воспользуемся методом Ван-дер-Поля (метод усреднений Боголюбова-Митропольского) для учета нелинейности. Для этого введем новые неизвестные функции  $A(t)$  и  $\varphi(t)$ , связанные со “старыми”  $y(t)_0$  и  $y(t)_1$  с помощью соотношений

$$\begin{aligned} y(t)_0 &= A(t) \cdot \cos(\omega \cdot t - \varphi(t)) \\ y(t)_1 &= -\omega \cdot A(t) \cdot \sin(\omega \cdot t - \varphi(t)) \end{aligned} \quad (8.18)$$

Подставим (8.18) в (8.17), получим систему уравнений для новых неизвестных  $A(t)$  и  $\varphi(t)$

$$\begin{aligned} \frac{d}{dt} A &= -\frac{\mu}{\omega} \cdot v(y(A, \psi)) \cdot \sin(\psi), \\ \frac{d}{dt} \varphi &= \frac{\mu}{\omega} \cdot \frac{1}{A} \cdot v(y(A, \psi)) \cdot \cos(\psi). \end{aligned} \quad (8.19)$$

В (8.19) для краткости введены обозначения  $A = A(t)$ ,  $\psi = \psi(t, \varphi)$



$$(8.20) \quad \psi(t, \varphi) := \omega \cdot t - \varphi,$$

$$(8.21) \quad y(A, \psi) := \begin{pmatrix} A \cdot \cos(\psi) \\ -\omega \cdot A \cdot \sin(\psi) \end{pmatrix}.$$

Метод Ван-дер-Поля применяется к системам, для которых можно ввести иерархию характерных времен развития. В исследуемой системе имеются “быстрые” движения, которые происходят с периодом

$$T = \frac{2 \cdot \pi}{\omega}.$$

Включение малой нелинейности приводит к появлению у системы “медленных” движений, имеющих большие характерные времена. Согласно (8.19), медленно изменяются во времени новые функции – амплитуда  $A(t)$  и фаза  $\varphi(t)$ . Это следует из того факта, что их производные – правые части (8.19) – пропорциональны малому параметру  $\mu$ . Движения с разными характерными временами можно приближенно разделить с помощью оператора усреднений Боголюбова - Митропольского. Оператор усреднений имеет вид

$$\bar{f}(t) = \frac{1}{T} \cdot \int_0^T f(t, \tau) d\tau. \quad (8.22)$$

Здесь усредняется функция  $f(t)$ , где, для удобства введены два временных аргумента  $t$  и  $\tau$ . Такое разделение весьма условно. Аргумент  $t$  связан с медленным развитием функции  $f(t)$ , аргумент  $\tau$  связан с быстрыми изменениями этой же функции. Другими словами, при интегрировании надо зафиксировать медленно изменяющиеся параметры под интегралом и проинтегрировать по “быстрому” времени. В (8.22) чертой помечена усредненная функция. Для выделения медленных движений, согласно идее метода Боголюбова - Митропольского, применим операцию усреднения к уравнениям (8.19). Применение операции к левой части первого уравнения дает результат

$$\frac{1}{T} \cdot \int_0^T \left( \frac{\partial}{\partial t} A(t, \tau) + \frac{\partial}{\partial \tau} A(t, \tau) \right) d\tau = \frac{d}{dt} \bar{A}(t).$$

Здесь под интегралом записана полная производная по времени

$$\frac{d}{dt} A(t) = \left( \frac{\partial}{\partial t} A(t, \tau) + \frac{\partial}{\partial \tau} A(t, \tau) \right) \Big|_{t=\tau}.$$

Применим операцию усреднения к правой части первого уравнения (8.19)

$$\begin{aligned} & \frac{1}{T} \cdot \int_0^T \left( -\frac{\mu}{\omega} \cdot v(y(A(t, \tau), \omega \cdot \tau - \varphi(t, \tau))) \cdot \sin(\omega \cdot \tau - \varphi(t, \tau)) \right) d\tau \approx \\ & \approx -\frac{\mu}{\omega} \cdot \frac{1}{2 \cdot \pi} \cdot \int_0^{2\pi} v(y(\bar{A}, \psi)) \cdot \sin(\psi) d\psi = \Phi(\bar{A}) \end{aligned} \quad (8.23)$$

Здесь под знаком усреднения можно, сохраняя первый порядок точности по параметру  $\mu$ , заменить амплитуду и фазу на усредненные значения. Аналогично, усредняя второе уравнение, получаем систему уравнений в первом порядке метода усреднений (укороченные уравнения Ван-дер-Поля)

$$\begin{cases} \frac{d}{dt} \bar{A} = \Phi(\bar{A}) \\ \frac{d}{dt} \bar{\varphi} = \Psi(\bar{A}) \end{cases} \quad (8.24)$$

Здесь

$$\Psi(\bar{A}) = \frac{\mu}{\omega} \cdot \frac{1}{\bar{A}} \cdot \frac{1}{2 \cdot \pi} \cdot \int_0^{2\pi} v(y(\bar{A}, \psi)) \cdot \cos(\psi) d\psi.$$

Часто выполнено  $\Psi(\bar{A}) = 0$ . Тогда  $\bar{\varphi} = \varphi_0$ , где  $\varphi_0$  - начальная фаза (8.15).

Далее будем полагать

$$\varphi_0 := 0.$$

Далее усредненное значение амплитуды будем обозначать той же буквой  $A$ , без черты сверху. Для ускорения работы программы удобно интеграл в формуле (8.23) взять явно

$$\Phi(A) := \diamond. \quad (8.25)$$

Исследуем первое уравнение системы (8.24). Построим график правой части в зависимости от положительных значений  $A$  и найдем приближенно, по графику, точки покоя этого уравнения

$$\Phi(A) = 0.$$

С помощью программы *root(,,)* уточним значения этих точек. Получим

$$A_1 := \diamond \quad A_2 := \diamond \quad A_3 := \diamond.$$

Не равные нулю амплитуды  $A_1, A_2, A_3, \dots$  называются амплитудами предельных циклов. Устойчивость предельного цикла определяется соотношением знаков функции  $\Phi(A)$ : если при переходе через 0 знак “плюс” меняется на знак “минус”, то этот предельный цикл асимптотически устойчив. В противном случае цикл не устойчив. В первом уравнении (8.24) делятся переменные, и его решение имеет вид

$$(8.26) \quad t(A, A_0) := \int_{A_0}^A \frac{1}{\Phi(A)} dA.$$

Будем рассматривать положительные значения амплитуды:  $0 \leq A(t) < \infty$ . Эта область значений разбивается точками  $A1, A2, A3, \dots$  на несколько подобластей. В зависимости от типа точек покоя, которые являются границами выбранной подобласти, график амплитуды  $A(t)$  будет монотонно убывать или возрастать. При  $t \rightarrow \infty$  график стремится к границе (точке покоя). Построим графики изменения амплитуды от времени с помощью обратной функции  $t(A, A0)$  (8.26). Для правильного отображения графика необходимо, в зависимости от свойства подобласти, правильно выбрать начальное

$$A1a := \blacklozenge \quad A2a := \blacklozenge \quad A3a := \blacklozenge$$

и конечное

$$A1b := \blacklozenge \quad A2b := \blacklozenge \quad A3b := \blacklozenge$$

значение амплитуд на каждой подобласти. Вводим *свою сетку узлов значений амплитуды на каждом интервале*. Для этого определяем число узлов и шаг сетки на каждом интервале

$$\begin{aligned} n1 &:= \blacklozenge \quad n2 := \blacklozenge \quad n3 := \blacklozenge \quad , \\ \Delta 1 &:= \frac{A1b - A1a}{n1} \quad \Delta 2 := \frac{A2b - A2a}{n2} \quad \Delta 3 := \frac{A3b - A3a}{n3} \quad , \\ k1 &:= 0..n1 \quad a1_{k1} := A1a + \Delta 1 \cdot k1 \\ k2 &:= 0..n2 \quad a2_{k2} := A2a + \Delta 2 \cdot k2 . \\ k3 &:= 0..n3 \quad a3_{k3} := A3a + \Delta 3 \cdot k3 \end{aligned}$$

По сетке амплитудных узлов введем соответствующую *сетку временных узлов* по формуле (8.26)

$$\begin{aligned} t1_{k1} &:= t(a1_{k1}, A1a) \\ t2_{k2} &:= t(a2_{k2}, A2a) \quad . \\ t3_{k3} &:= t(a3_{k3}, A3a) \end{aligned}$$

В одном окне графопостроителя строим графики изменения амплитуды для всех подобластей. Построим фазовые траектории с помощью формул (8.21), (8.20), отложив по горизонтальной оси значения  $y(A, \psi)_0$  и по вертикальной оси значения  $y(A, \psi)_1$  на соответствующей сетке амплитудных и временных узлов

$$\begin{aligned} Y10_{k1} &:= y(a1_{k1}, \psi(t1_{k1}, \varphi 0))_0 \quad Y11_{k1} := y(a1_{k1}, \psi(t1_{k1}, \varphi 0))_1 \\ Y20_{k2} &:= y(a2_{k2}, \psi(t2_{k2}, \varphi 0))_0 \quad Y21_{k2} := y(a2_{k2}, \psi(t2_{k2}, \varphi 0))_1 . \\ Y30_{k3} &:= y(a3_{k3}, \psi(t3_{k3}, \varphi 0))_0 \quad Y31_{k3} := y(a3_{k3}, \psi(t3_{k3}, \varphi 0))_1 \end{aligned}$$

В одном окне построим фазовые траектории для всех подобластей по формулам (8.21), (8.20). Это приближенные траектории, найденные с

помощью метода усреднений. Решим систему уравнений с правой частью  $FN(y)$  (8.16) численно, с помощью схемы Рунге – Кутта (8.12). Для этого определим минимальное и максимальное значение времени на каждом интервале

$$\begin{aligned} t_{min1} &:= 0 & t_{max1} &:= t1_{n1} \\ t_{min2} &:= 0 & t_{max2} &:= t2_{n2} \\ t_{min3} &:= 0 & t_{max3} &:= t3_{n3} \end{aligned} .$$

Введем число шагов на каждой подобласти

$$N1 := \blacklozenge \quad N2 := \blacklozenge \quad N3 := \blacklozenge .$$

Определим шаг равномерной сетки на каждой подобласти

$$\delta1 := \frac{t_{max1} - t_{min1}}{N1} \quad \delta2 := \frac{t_{max2} - t_{min2}}{N2} \quad \delta3 := \frac{t_{max3} - t_{min3}}{N3} .$$

Определим начальные данные для каждой траектории, соответствующие начальным амплитудам  $A1a, A2a, A3a, \dots$  по формуле (8.15)

$$\begin{aligned} Y01 &:= y0(A1a, \varphi0) \\ Y02 &:= y0(A2a, \varphi0) \\ Y03 &:= y0(A3a, \varphi0) \end{aligned} .$$

Запустим схему Рунге – Кутта (8.12)  $RKa(Y0, F, N, h)$  несколько раз для разных начальных данных, разного числа шагов и разной величины шага и построим численные фазовые траектории в одном окне графопостроителя. Сравним численный расчет с приближенным расчетом по методу усреднений и оценим возможности приближенного метода.

### Задания к работе

1. Доказать, что точность схемы Эйлера на каждом шаге оценивается как  $O(h^2)$ , а схемы Рунге – Кутта с одной итерацией – как  $O(h^3)$ .  
Запрограммировать неавтономную схему Эйлера и схему Рунге – Кутта с одной итерацией. Рассмотреть точно решаемую неавтономную систему уравнений и проверить численно, что точность схемы Эйлера на интервале  $t0 \leq t \leq t1$  оценивается как  $O(h)$ , а схемы Рунге – Кутта – как  $O(h^2)$ .
2. Задать правые части  $P(p, q)$  и  $Q(p, q)$  автономной системы уравнений и графически найти приближенные координаты точки покоя системы на фазовой плоскости. Уточнить местоположение точки с заданной точностью с помощью итерационного процесса.
3. Найти тип выбранной точки покоя и построить несколько приближенных решений задачи Коши в окрестности точки покоя с помощью метода линеаризации для разных начальных данных. Построить численное решение для тех же начальных данных и

сделать вывод о применимости метода линеаризации и правильности определения типа точки покоя.

4. Вывести систему уравнений (8.19) для неизвестных амплитуды  $A(t)$  и фазы  $\varphi(t)$ .
5. С помощью метода усреднений получить укороченную систему уравнений Ван-дер-Поля.
6. Ввести нелинейную добавку к силе  $\mu \cdot v(y)$  и найти явное выражение правой части первого укороченного уравнения  $\Phi(A)$  (8.25).
7. Определить амплитуды предельных циклов и их тип. Построить графики зависимости амплитуды от времени, решив укороченное уравнение.
8. Построить графики траекторий на фазовой плоскости для решений, найденных методом усреднений и сравнить соответствующие траектории с траекториями, полученными численным методом по схеме Рунге – Кутта. Сделать вывод о применимости метода усреднений.

## 9. Метод Фурье и сеточный метод решения уравнений в частных производных

### 9.1. Метод разделения переменных (метод Фурье) для уравнения теплопроводности

Изучим уравнение теплопроводности, определяющее распределение температуры по объему тела, находящегося в тепловом контакте с внешней средой и нагреваемого источниками тепла “внутри” тела [1], [2], [5], [6], [7]. Обозначим  $u(x, y, z, t)$  - температуру тела в точке  $(x, y, z)$  в момент  $t$ . Введем характеристику теплового потока – вектор плотности потока тепла  $S(x, y, z, t)$  (размерность  $эрг/(см^2сек)$ ). По проверенному в экспериментах закону (закон Фурье) этот вектор связан с температурой формулой

$$S(x, y, z, t) = -\kappa \cdot \text{grad}(u(x, y, z, t)).$$

Здесь  $\kappa$  - коэффициент теплопроводности. Для неоднородных и нестационарных сред  $\kappa$  зависит от точки и времени. Рассмотрим тело  $D$ , ограниченное поверхностью  $\Sigma$ . Обозначим  $f(x, y, z, t)$  - объемную плотность источников тепла (размерность  $эрг/(см^3сек)$ ) в точке  $(x, y, z)$  в момент  $t$ . Количество тепла  $\Delta Q_S$ , созданное источниками в объеме  $D$  за время  $\Delta t$  равно

$$\Delta Q_S = \iiint_D f(x, y, z, t) \cdot dV \cdot \Delta t.$$

Расход тепла за счет потока через поверхность  $\Sigma$  за время  $\Delta t$  равно

$$\Delta Q_F = \iint_{\Sigma} (S(x, y, z, t), n(x, y, z)) \cdot d\Sigma \cdot \Delta t.$$

Здесь  $n(x, y, z)$  - внешняя нормаль к поверхности в точке  $(x, y, z)$  на поверхности. Отмеченные выше две причины приводят к изменению количества тепла в теле за время  $\Delta t$ , и это приведет к изменению температуры в каждой точке тела. Изменение количества тепла в теле  $\Delta Q$  связано с изменением температуры по формуле

$$\Delta Q = \iiint_D C \cdot \rho(x, y, z, t) \cdot u_t(x, y, z, t) \cdot dV \cdot \Delta t.$$

Здесь:  $C$  - коэффициент теплоемкости вещества (при постоянном объеме),  $\rho(x, y, z, t)$  - объемная плотность массы,  $u_t(x, y, z, t) = \frac{\partial}{\partial t} u(x, y, z, t)$

Применим теорему Остроградского – Гаусса. Закон сохранения тепла

$$\Delta Q = \Delta Q_S - \Delta Q_F \tag{9.1}$$

запишем в виде

$$\begin{aligned} & \iiint_D C \cdot \rho(x, y, z, t) \cdot u_t(x, y, z, t) \cdot dV = \\ & = \iiint_D \left( \operatorname{div}(\kappa \cdot \operatorname{grad}(u(x, y, z, t))) + f(x, y, z, t) \right) \cdot dV \end{aligned}$$

В силу произвольности объема  $D$  получаем уравнение теплопроводности

$$\operatorname{div}(\kappa \cdot \operatorname{grad}(u(x, y, z, t))) + f(x, y, z, t) = C \cdot \rho(x, y, z, t) \cdot u_t(x, y, z, t). \quad (9.2)$$

Для однородной стационарной среды уравнение упрощается

$$K \cdot \widehat{\Delta} u(x, y, z, t) + g(x, y, z, t) = u_t(x, y, z, t). \quad (9.3)$$

Здесь коэффициент температуропроводности  $K = \frac{\kappa}{C \cdot \rho}$ ,

$$g(x, y, z, t) = \frac{f(x, y, z, t)}{C \cdot \rho}, \quad \widehat{\Delta} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad - \quad \text{оператор Лапласа.}$$

Сформулируем задачу Коши для уравнения (9.3). Будем искать решение  $u(x, y, z, t)$  при  $t \geq 0$  в объеме  $D$ , удовлетворяющее дополнительным условиям:

а) начальным

$$u(x, y, z, t)|_{t=0} = \Phi(x, y, z), \quad (x, y, z) \in D \quad (9.4)$$

б) краевым (граничным)

$$\begin{aligned} & \left( \gamma_1(x, y, z) \cdot (\mathbf{n}(x, y, z), \operatorname{grad}(u(x, y, z, t))) + \gamma_2(x, y, z) \cdot u(x, y, z, t) \right) \Big|_{\Sigma} = \\ & = \beta(x, y, z, t) \end{aligned}$$

Здесь  $\beta(x, y, z, t)$  - температура внешней среды в точках поверхности тела,  $\mathbf{n}(x, y, z)$  - внешняя единичная нормаль к поверхности. Граничное условие есть закон теплообмена Ньютона на контакте двух сред: поток тепла вдоль нормали к поверхности контакта пропорционален разности температур контактирующих сред (в области контакта). Далее рассмотрим стационарные однородные граничные условия

$$\left( \gamma_1 \cdot (\mathbf{n}(x, y, z), \operatorname{grad}(u(x, y, z, t))) + \gamma_2 \cdot u(x, y, z, t) \right) \Big|_{\Sigma} = 0. \quad (9.5)$$

Здесь не зависящие от времени и координат коэффициенты  $\gamma_1$  и  $\gamma_2$  определяют тип теплового контакта тела с внешней средой (через границу). Температура внешней среды считается равной нулю во все моменты времени и во всех точках.

Задачу (9.3), (9.4), (9.5) можно решить с помощью метода Фурье. Для проведения данного метода необходимо, чтобы в *однородном уравнении*

$$K \cdot \widehat{\Delta} u(x, y, z, t) = u_t(x, y, z, t), \quad (9.6)$$

которое получается из неоднородного (9.3) обращением в нуль неоднородности  $g(x, y, z, t)$ , переменные разделились. А именно пространственные переменные отделились от времени, то есть у уравнения (9.6) существовали бы частные решения вида  $u(x, y, z, t) = X(x, y, z) \cdot T(t)$ . В разделенном виде однородное уравнение (9.6) имеет вид системы двух уравнений: *однородного пространственного и временного* уравнения. Пространственное уравнение для функции  $X(x, y, z)$  имеет вид

$$K \cdot \widehat{\Delta} X(x, y, z) = \lambda \cdot X(x, y, z). \quad (9.7)$$

Решение  $X(x, y, z)$  удовлетворяет граничным условиям

$$\left( \gamma_1 \cdot (n(x, y, z), \text{grad}(X(x, y, z))) + \gamma_2 \cdot X(x, y, z) \right) \Big|_{\Sigma} = 0. \quad (9.8)$$

Однородное временное уравнение для функции  $T(t)$  имеет вид

$$T_t(t) = \lambda \cdot T(t).$$

Здесь  $\lambda$  - константа деления переменных. Как показывает теория, краевая задача (9.7), (9.8) на *собственные векторы и собственные числа* имеет счетное число решений – собственных векторов  $R(k, x, y, z)$ , для счетного набора неположительных собственных чисел  $\lambda_k$ . Это множество решений образует *ортонормированный базис Фурье*. Базисные функции удовлетворяют уравнениям

$$\begin{aligned} K \cdot \widehat{\Delta} R(k, x, y, z) &= \lambda_k \cdot R(k, x, y, z), \quad k = 0, 1, 2, \dots \\ \left( \gamma_1 \cdot (n(x, y, z), \text{grad}(R(k, x, y, z))) + \gamma_2 \cdot R(k, x, y, z) \right) \Big|_{\Sigma} &= 0. \end{aligned} \quad (9.9)$$

Базисные элементы ортонормированны в смысле скалярного произведения

$$\iiint_D R(k, x, y, z) \cdot R(m, x, y, z) dV = \delta_{k,m},$$

где  $\delta_{k,m}$  - символ Кронекера. Если удобный базис найден, то решение *неоднородного* уравнения (9.3), удовлетворяющее граничным условиям и начальным данным, следует искать в виде ряда Фурье по этому базису

$$u(x, y, z, t) = \sum_{k=0}^{\infty} T(k, t) \cdot R(k, x, y, z).$$

Функция  $u(x, y, z, t)$  автоматически удовлетворяет граничным условиям. Для удовлетворения начальным данным (9.4) временные функции  $T(k, t)$  для каждого  $k$  должны удовлетворять *неоднородному* (если функция  $g(x, y, z, t)$  не равна нулю) *временному уравнению*

$$\begin{aligned} T_t(k, t) &= \lambda_k \cdot T(k, t) + G(k, t) \\ T(k, t) \Big|_{t=0} &= \varphi_k, \quad k = 0, 1, 2, \dots \end{aligned} \quad (9.10)$$



Здесь введены обозначения компонент Фурье для  $g(x, y, z, t)$  (9.3) и  $\Phi(x, y, z)$  (9.4)

$$G(k, t) = \iiint_D g(x, y, z, t) \cdot R(k, x, y, z) dV,$$

$$\varphi_k = \iiint_D \Phi(x, y, z) \cdot R(k, x, y, z) dV.$$

Проверим метод Фурье с помощью вычислений для упрощенного одномерного случая. Предположим, что температура среды постоянна во всех точках плоскости  $yz$  и меняется вдоль оси  $x$ . Такая ситуация возможна для слоистых сред, когда поперечные размеры слоя существенно превышают его толщину. Тогда задача становится одномерной по пространственным переменным и уравнения приобретают вид:

1) уравнение теплопроводности

$$K \cdot u_{xx}(x, t) + g(x, t) = u_t(x, t), \quad (9.11)$$

2) начальные данные

$$u(x, t)|_{t=0} = \Phi(x) \quad 0 \leq x \leq L, \quad (9.12)$$

3) граничные условия на левом  $x = 0$  (нормаль направлена против оси  $x$ ) и правом  $x = L$  краю (нормаль направлена по оси  $x$ )

$$\begin{aligned} (s1 \cdot u_x(x, t) - h1 \cdot u(x, t))|_{x=0} &= 0 \\ (s2 \cdot u_x(x, t) + h2 \cdot u(x, t))|_{x=L} &= 0 \end{aligned} \quad (9.13)$$

Здесь  $u_x(x, t) = \frac{\partial}{\partial x} u(x, t)$ ,  $u_{xx}(x, t) = \frac{\partial^2}{\partial x^2} u(x, t)$ . Зададим:  $L$  - толщину слоя

$$L := \blacklozenge,$$

коэффициент теплопроводности

$$K := \blacklozenge,$$

параметры теплового контакта с внешней средой (положительные числа)

$$s1 := \blacklozenge \quad h1 := \blacklozenge \quad s2 := \blacklozenge \quad h2 := \blacklozenge,$$

источники тепла

$$g(x, t) := \blacklozenge,$$

начальное распределение температуры

$$\Phi(x) := \blacklozenge.$$

В одномерном случае краевая задача (9.9) имеет вид (краевая задача Штурма – Лиувилля)

$$\begin{aligned}
K \cdot R_{xx}(x) &= \lambda \cdot R(x) \\
(s1 \cdot R_x(x) - h1 \cdot R(x)) \Big|_{x=0} &= 0 \quad . \\
(s2 \cdot R_x(x) + h2 \cdot R(x)) \Big|_{x=L} &= 0
\end{aligned} \tag{9.14}$$

Общее решение уравнения (9.14) имеет вид

$$R(x) = A \cdot \cos(Z(\lambda) \cdot x) + B \cdot \sin(Z(\lambda) \cdot x). \tag{9.15}$$

Здесь

$$Z(\lambda) := \sqrt{-\lambda/K}.$$

Подставим (9.15) в граничные условия (9.14), получим систему однородных уравнений для неизвестных  $A, B, \lambda$

$$\begin{aligned}
-h1 \cdot A + Z(\lambda) \cdot s1 \cdot B &= 0 \\
(-Z(\lambda) \cdot s2 \cdot \sin(Z(\lambda) \cdot L) + h2 \cdot \cos(Z(\lambda) \cdot L)) \cdot A + \\
+(Z(\lambda) \cdot s2 \cdot \cos(Z(\lambda) \cdot L) + h2 \cdot \sin(Z(\lambda) \cdot L)) \cdot B &= 0
\end{aligned} \tag{9.16}$$

Система (9.16) имеет нетривиальные решения, при условии, что ее детерминант равен нулю

$$\begin{aligned}
det(\lambda) := -Z(\lambda) \cdot (h1 \cdot s2 + h2 \cdot s1) \cdot \cos(Z(\lambda) \cdot L) + \\
+(s1 \cdot s2 \cdot Z(\lambda)^2 - h1 \cdot h2) \cdot \sin(Z(\lambda) \cdot L) \quad .
\end{aligned}$$

Получаем трансцендентное уравнение

$$det(\lambda) = 0. \tag{9.17}$$

Корни уравнения (9.17) определяют собственные числа задачи Штурма – Лиувилля. Найдем значения  $m$  отрицательных младших корней приближенно, построив график функции  $det(\lambda)$ . Затем, с помощью программы `root(,,)`, найдем положение корней с заданной точностью. Найденные собственные числа упорядочим в вектор  $\lambda$

$$\begin{aligned}
m &:= \blacklozenge \quad , \\
\lambda &:= \blacklozenge \quad .
\end{aligned}$$

Для каждого собственного числа найдем соответствующий собственный вектор. Если  $h1 \neq 0$ , то положим

$$B := 1 \quad A(n) := \frac{Z(\lambda_n) \cdot s1}{h1}.$$

Тогда ненормированный собственный вектор имеет вид

$$(9.18) \quad r(n, x) := A(n) \cdot \cos(Z(\lambda_n) \cdot x) + \sin(Z(\lambda_n) \cdot x).$$

Введем скалярное произведение и определим нормы векторов (9.18)

$$k := 0..m-1$$

$$N_k := \sqrt{\int_0^L r(k, x)^2 dx} \quad .$$

Тогда ортонормированный базис Фурье имеет вид

$$R(n, x) := \frac{r(n, x)}{N_n} \quad .$$

Проверим ортонормировку численно, заполнив матрицу

$$p := 0..m-1$$

$$M_{k,p} := \int_0^L R(k, x) \cdot R(p, x) dx \quad .$$

Матрица  $M$  должна быть близка к единичной. Точность ортонормировки существенно зависит от точности вычисления собственных чисел  $\lambda_k$ . Приближенное решение задачи (9.11), (9.12), (9.13) ищем в виде частичной суммы ряда Фурье

$$u(x, t) = \sum_{n=0}^m T(n, t) \cdot R(n, x).$$

Погрешность, вносимая обрывом ряда, зависит от скорости убывания коэффициентов Фурье при увеличении их номера. Ряд Фурье будет быстро сходиться, если исходная, разлагаемая в ряд, функция будет удовлетворять граничным условиям (9.14) во все моменты времени. Отсюда следует, что быструю сходимость ряда следует ожидать тогда, когда начальное условие  $\Phi(x)$  (9.12) и функция  $g(x, t)$  (9.11) будут удовлетворять граничным условиям (9.13). Сконструируем функцию  $g(x, t)$ , удовлетворяющую граничным условиям (9.13). Для этого выберем произвольную гладкую функцию  $H(x, t)$

$$H(x, t) := \blacklozenge \quad .$$

“Исправим” эту функцию, добавив к ней линейное по  $x$  слагаемое  $\chi(t) \cdot x + \theta(t)$ . Для каждого момента  $t$  функции  $\theta(t)$ ,  $\chi(t)$  подберем так, чтобы граничные условия (9.13) оказались выполненными. Решим полученную систему двух уравнений, получим

$$\theta(t) := \blacklozenge \quad \chi(t) := \blacklozenge \quad .$$

Тогда функция  $g(x, t)$  запишется

$$(9.19) \quad g(x, t) := H(x, t) + \chi(t) \cdot x + \theta(t).$$

По аналогии получим начальное условие  $\Phi(x)$ , выбрав произвольную гладкую функцию  $\Psi(x)$  и исправив ее добавкой  $a \cdot x + b$

$$\Psi(x) := \blacklozenge \quad .$$

Подберем  $a$ ,  $b$  так, чтобы  $\Phi(x)$  удовлетворяла (9.13), получаем

$$(9.20) \quad \begin{aligned} a &:= \blacklozenge & b &:= \blacklozenge \\ \Phi(x) &:= \Psi(x) + a \cdot x + b \end{aligned}$$

Решим временное уравнение (9.10). Для этого найдем компоненты Фурье функций  $g(x,t)$  и  $\Phi(x)$

$$\begin{aligned} G(n,t) &:= \int_0^L g(x,t) \cdot R(n,x) dx, \\ \varphi_p &:= \int_0^L \Phi(x) \cdot R(p,x) dx. \end{aligned}$$

Решение задачи Коши (9.10) имеет вид

$$T(n,t) := \exp(\lambda_n \cdot t) \left( \varphi_n + \int_0^t G(n,t) \cdot \exp(-\lambda_n \cdot t) dt \right).$$

Окончательно (приближенное, из-за обрыва ряда) решение (9.11) - (9.13) имеет вид

$$u(x,t) := \sum_{n=0}^m T(n,t) \cdot R(n,x).$$

Средства математических пакетов позволяют анимировать картину остывания слоя, записать анимацию в видеофайл и детально изучить закономерности остывания в зависимости от типа граничных условий и особенностей нагрева источниками.

## 9.2. Конечно-разностные методы

При решении уравнений в частных производных сеточными методами получают приближенные значения искомой функции на сетке пространственных и временных узлов (сеточная функция). В случае необходимости можно построить интерполяционную формулу для приближенного представления решения между узлами. Дифференциальное уравнение на сетке заменяется на разностное. При этом возникают следующие проблемы:

- 1) сходится ли точное решение разностной задачи к решению дифференциальной (вопрос *сходимости*)
- 2) насколько сильно изменяется решение разностной задачи, если при вычислениях допускаются погрешности и насколько сильно решение зависит от изменения начальных данных (вопрос *устойчивости*)
- 3) какова *погрешность аппроксимации*, то есть погрешность, вносимая при замене дифференциальной операции на разностную (вопрос *порядка аппроксимации*).

В случае линейных дифференциальных уравнений разностное уравнение представляет собой систему линейных алгебраических уравнений. Чтобы получить эту систему необходимо дифференциальным операциям (производным по времени и по координатам) сопоставить соответствующую конечно-разностную схему. На конечно-разностные уравнения необходимо заменить само уравнение и граничные условия. Порядок аппроксимации рассчитывается с помощью разложения функции и ее производных в ряды Тейлора в точках сетки узлов с использованием шаблона. Так, если остаточное слагаемое оценивается как  $O(h^k)$ , где  $h$  - шаг равномерной сетки  $x_m$  по координате  $x$ , то порядок аппроксимации равен  $k$ . Найдем конечно-разностное выражение для второй производной произвольной функции  $f(x)$  с порядком аппроксимации  $k=2$  шаблоне  $\begin{matrix} \circ & - & \bullet & - & \circ \\ m-1 & & m & & m+1 \end{matrix}$ . Здесь под черным кружком отмечен номер узла, в котором надо вычислить производную, и эта производная должна быть выражена через значения функции на остальных узлах шаблона. Другими словами, для данного шаблона надо найти коэффициенты  $a_0, a_1, a_2$  в выражении

$$f''(x_m) = a_0 \cdot f(x_{m-1}) + a_1 \cdot f(x_m) + a_2 \cdot f(x_{m+1}) + O(h^2).$$

Для этого используем разложение Тейлора до третьего порядка по  $h$

$$f(x_{m+1}) = f(x_m) + f'(x_m) \cdot h + f''(x_m) \cdot \frac{h^2}{2} + f'''(x_m) \cdot \frac{h^3}{6} + O(h^4)$$

$$f(x_{m-1}) = f(x_m) - f'(x_m) \cdot h + f''(x_m) \cdot \frac{h^2}{2} - f'''(x_m) \cdot \frac{h^3}{6} + O(h^4)$$

Сложим эти два уравнения, получим требуемое разложение

$$f''(x_m) = \frac{f(x_{m+1}) + f(x_{m-1}) - 2 \cdot f(x_m)}{h^2} + O(h^2). \quad (9.21)$$

Используем шаблон  $\begin{matrix} \bullet & - & \circ & - & \circ \\ 0 & & 1 & & 2 \end{matrix}$  для аппроксимации первой производной произвольной функции  $f(x)$  с точностью  $O(h^2)$  в левом граничном условии (9.14). Получаем

$$f_x(x_0) = \frac{4 \cdot f(x_1) - f(x_2) - 3 \cdot f(x_0)}{2 \cdot h} + O(h^2). \quad (9.22)$$

Обозначим  $n$  - номер последнего узла координатной сетки. Используем шаблон  $\begin{matrix} \circ & - & \circ & - & \bullet \\ n-2 & & n-1 & & n \end{matrix}$  для аппроксимации первой производной произвольной функции  $f(x)$  с точностью  $O(h^2)$  в правом граничном условии (9.14).

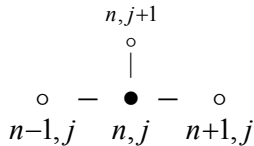
Получаем

$$f_x(x_n) = \frac{3 \cdot f(x_n) + f(x_{n-2}) - 4 \cdot f(x_{n-1})}{2 \cdot h} + O(h^2). \quad (9.23)$$

Введем обозначение  $\tau$  - шаг равномерной сетки  $t_j$  по времени, используем шаблон  $\bullet - \circ$  для аппроксимации первой производной по времени с порядком  $k = 1$

$$\varphi_t(t_j) = \frac{\varphi(t_{j+1}) - \varphi(t_j)}{\tau} + O(\tau). \quad (9.24)$$

Тогда задача (9.11) - (9.13) на шаблоне



имеет порядок аппроксимации  $O(h^2 + \tau)$ . Используя формулы (9.21) - (9.24), получаем *явное* конечно-разностное соотношение

$$u_{m, j+1} - u_{m, j} = \mu \cdot (u_{m+1, j} + u_{m-1, j} - 2 \cdot u_{m, j}) + \tau \cdot g_{m, j}, \quad (9.25)$$

$$s1 \cdot \frac{4 \cdot u_{1, j} - u_{2, j} - 3 \cdot u_{0, j}}{2 \cdot h} - h1 \cdot u_{0, j} = 0, \quad (9.26)$$

$$s2 \cdot \frac{3 \cdot u_{n, j} + u_{n-2, j} - 4 \cdot u_{n-1, j}}{2 \cdot h} + h2 \cdot u_{n, j} = 0, \quad (9.27)$$

$$u_{m, 0} = \Phi(x_m). \quad (9.28)$$

Здесь  $u_{m, j} = u(x_m, t_j)$  - решение на пространственной и временной сетках. Существует *теорема*, связывающая между собой понятия аппроксимации, устойчивости, сходимости.

• **Теорема:**

*Если решение исходной дифференциальной задачи существует, а разностная схема устойчива и аппроксимирует решаемую задачу на данном решении с порядком  $k$ , то разностное решение сходится к точному со скоростью  $O(h^k)$ .* •

Введем параметры задачи (9.25) - (9.28). Количество узлов координатной сетки

$$n := \blacklozenge,$$

шаг координатной сетки

$$h := \frac{L}{n},$$

координатная сетка

$$m := 0..n$$

$$x_m := h \cdot m,$$

шаг временной сетки

$$\tau := \diamond ,$$

количество шагов временной сетки

$$N := \diamond ,$$

временная сетка

$$j := 0..N$$

$$t_j := \tau \cdot j$$

источники тепла на пространственной и временной сетках

$$g_{m,j} := g(x_m, t_j),$$

начальные данные на пространственной сетке

$$\Phi_m := \Phi(x_m).$$

Теория показывает, что *явная* схема является *условно устойчивой*. Это означает, что устойчивость разностной схемы может быть выполнена при определенном ограничении на параметр

$$\mu := \frac{K \cdot \tau}{h^2}.$$

Это ограничение имеет вид

$$\mu < 0.5. \tag{9.29}$$

Соотношение (9.29) можно проверить с помощью *спектрального признака устойчивости*, для применения которого необходимо найти спектр (набор собственных чисел) разностного оператора (9.25). Разделим дискретные переменные в однородном разностном уравнении

$$u_{m,j+1} - u_{m,j} = \mu \cdot (u_{m+1,j} + u_{m-1,j} - 2 \cdot u_{m,j}). \tag{9.30}$$

Для этого найдем набор частных решений, имеющих вид

$$u_{m,j} = v_m \cdot w_j. \tag{9.31}$$

Подстановка (9.31) в (9.30) дает систему двух разностных уравнений, связанных положительной константой разделения переменных  $\lambda$

$$\begin{cases} -\lambda \cdot v_m = v_{m+1} - v_m \\ -\lambda \cdot w_j = \mu \cdot (w_{j+1} + w_{j-1} - 2 \cdot w_j) \end{cases}, \quad j = 0, 1, \dots, n, \quad m = 0, 1, \dots, N. \tag{9.32}$$

Найдем положительные собственные числа  $\lambda$  из второго разностного уравнения (9.32). Решение ищем в виде

$$w_j = \gamma^j.$$

Подставим в (9.32), получим уравнение для параметра  $\gamma$

$$\gamma^2 - \left( \frac{\lambda}{\mu} + 2 \right) \cdot \gamma + 1 = 0. \tag{9.33}$$

Решение будет устойчивым, если параметры  $\gamma$  по модулю будут равны 1. Будем искать  $\gamma$  в виде

$$\gamma = \exp(i \cdot \chi). \quad (9.34)$$

Тогда из (9.33) получаем дисперсионное соотношение для положительных  $\lambda$  в виде

$$\lambda = 2 \cdot \mu \cdot (1 - \cos \chi).$$

Отсюда следует, что

$$0 \leq \lambda \leq 4 \cdot \mu. \quad (9.35)$$

Решение первого уравнения системы (9.32) ищем в виде  $v_m = \delta^m$ . Подставим это решение в (9.32), получим соотношение  $\delta = 1 - \lambda$ . Решение (9.32) будет устойчивым при увеличении  $m$ , если  $|1 - \lambda| < 1$ , или

$$0 \leq \lambda < 2. \quad (9.36)$$

При выполнении (9.34) все собственные числа  $\lambda$  (9.35) будут удовлетворять (9.36), если выполнено соотношение (9.29). Запишем рекуррентное соотношение (9.25) в матричной форме. Для этого введем матрицу  $D$  второй производной

$$\begin{aligned} D_{m,m} &:= -2 \\ m1 &:= 0..n-1 \\ D_{m1,m1+1} &:= 1 \quad D_{m1+1,m1} := 1 \end{aligned} .$$

Решение на сетке, удовлетворяющее начальным и граничным условиям, имеет вид

$$U := \begin{cases} u^{(0)} \leftarrow \Phi \\ \text{for } j \in 0..N \\ \left| \begin{array}{l} u^{(j+1)} \leftarrow u^{(j)} + \mu \cdot D \cdot u^{(j)} + \tau \cdot g^{(j)} \\ u_{0,j+1} \leftarrow \frac{s1}{3 \cdot s1 + 2 \cdot h1 \cdot h} \cdot (4 \cdot u_{1,j+1} - u_{2,j+1}) \\ u_{n,j+1} \leftarrow \frac{s2}{3 \cdot s2 + 2 \cdot h2 \cdot h} \cdot (4 \cdot u_{n-1,j+1} - u_{n-2,j+1}) \end{array} \right. \\ u \end{cases} .$$

Сравним графически решения, полученные методом Фурье и решения на сетке. Проверим формулы для погрешности вычислений.

### Задания к работе

1. Объяснить расстановку знаков в законе сохранения тепла (9.1) и дать подробный вывод уравнения теплопроводности (9.2). Объяснить физический смысл граничных условий (9.5).



2. Разделить переменные в однородном уравнении теплопроводности (9.5), (9.6).
3. Доказать самосопряженность оператора Лапласа на множестве функций, удовлетворяющих граничным условиям (9.8). Каковы свойства собственных векторов и собственных чисел самосопряженного оператора?
4. Получить и решить неоднородное временное уравнение (9.10), вывести уравнение (9.17). Получить функции  $g(x,t)$  (9.19) и  $\Phi(x)$  (9.20), удовлетворяющие граничным условиям.
5. Используя средства математического пакета, детально изучить закономерности остывания слоя в зависимости от типа граничных условий и особенностей нагрева источниками. Оценить скорость остывания и погрешность метода Фурье.
6. Получить численное решение задачи сеточным методом, подробно исследовать проблему устойчивости разностной схемы и точности аппроксимации и сделать вывод о сходимости численного решения и его погрешности.

## 10. Интегральное уравнение Фредгольма II рода

В различных разделах математической физики широко используется интегральное уравнение Фредгольма II рода [1], [2], [5], [9], [14]

$$y(x) = \lambda \cdot \int_a^b K(x, \tilde{x}) \cdot y(\tilde{x}) d\tilde{x} + f(x). \quad (10.1)$$

Здесь  $y(x)$  - неизвестная функция, удовлетворяющая условиям непрерывности и интегрируемости на интервале  $a \leq x \leq b$ ,  $K(x, \tilde{x})$  - ядро интегрального оператора,  $f(x)$  - заданная функция (правая часть уравнения),  $\lambda$  - параметр уравнения. Обозначим для краткости интегральный оператор символом  $\hat{J}$

$$\hat{J}y(x) = \int_a^b K(x, \tilde{x}) \cdot y(\tilde{x}) d\tilde{x} \quad (10.2)$$

и перепишем (10.1) в виде

$$(\hat{I} - \lambda \cdot \hat{J})y(x) = f(x). \quad (10.3)$$

Здесь  $\hat{I}$  - единичная операция. Для поиска решения используется понятие резольвенты. Решение уравнения (10.3) запишем в виде

$$y(x) = \hat{R}ez(\lambda) \cdot f(x). \quad (10.4)$$

Здесь обозначен (интегральный) оператор резольвенты

$$\hat{R}ez(\lambda) = (\hat{I} - \lambda \cdot \hat{J})^{-1},$$

имеющий ядро  $Rez(\lambda, x, \tilde{x})$  и действующий на функцию  $f(x)$  (10.4) согласно правилу

$$y(x) = \int_a^b Rez(\lambda, x, \tilde{x}) \cdot f(\tilde{x}) d\tilde{x}. \quad (10.5)$$

Найдем ядро  $Rez(\lambda, x, \tilde{x})$ . Для этого предположим, что интегральный оператор  $\hat{J}$  обладает дискретным спектром характеристических чисел  $\mu_n$ ,  $n = 0, 1, 2, \dots$  и счетным набором собственных функций  $\chi_n(x)$

$$\mu_m \cdot \hat{J}\chi_m(x) = \chi_m(x). \quad (10.6)$$

Собственные функции нормированы условием

$$\int_a^b \chi_n^*(x) \cdot \chi_m(x) dx = \delta_{n,m}. \quad (10.7)$$

Здесь  $\delta_{n,m}$  - символ Кронекера,  $*$  - комплексное сопряжение. Ядро  $K(x, \tilde{x})$  можно разложить по (бесконечному) базису (10.6), (10.7)

$$K(x, \tilde{x}) = \sum_{\substack{n=0 \\ m=0}}^{\infty} \chi_m(x) \cdot K_{m,n} \cdot \chi_n^*(\tilde{x}). \quad (10.8)$$

Матрица ядра  $K_{m,n}$  находится при подстановке разложения (10.8) в (10.6)

$$K_{m,n} = \frac{1}{\mu_n} \delta_{n,m}. \quad (10.9)$$

Ядро  $\text{Rez}(\lambda, x, \tilde{x})$  будем искать также в виде разложения по собственному базису

$$\text{Rez}(\lambda, x, \tilde{x}) = \sum_{\substack{n=0 \\ m=0}}^{\infty} \chi_m(x) \cdot \text{REZ}(\lambda)_{m,n} \cdot \chi_n^*(\tilde{x}). \quad (10.10)$$

Здесь  $\text{REZ}(\lambda)$  - матрица оператора резольвенты с матричными элементами  $\text{REZ}(\lambda)_{m,n}$ . Для ее нахождения разложим функцию  $y(x)$  и правую часть  $f(x)$  по базису  $\chi_m(x)$

$$y(x) = \sum_{m=0}^{\infty} Y_m \cdot \chi_m(x), \quad Y_m = \int_a^b \chi_m^*(x) \cdot y(x) dx, \quad (10.11)$$

$$f(x) = \sum_{m=0}^{\infty} F_m \cdot \chi_m(x), \quad F_m = \int_a^b \chi_m^*(x) \cdot f(x) dx. \quad (10.12)$$

Подставим (10.8), (10.9), (10.11), (10.12) в уравнение (10.1), используем соотношение (10.7) и запишем интегральное уравнение (10.1) в матричном алгебраическом виде

$$(E - \lambda \cdot K)Y = F. \quad (10.13)$$

Здесь  $E$  - единичная матрица,  $K$  - матрица ядра. Решение алгебраического уравнения (10.13) имеет вид

$$Y = (E - \lambda \cdot K)^{-1} \cdot F. \quad (10.14)$$

Перепишем решение (10.5) в базисе  $\chi_m(x)$

$$Y = \text{REZ}(\lambda) \cdot F. \quad (10.15)$$

Сравнивая (10.15) с (10.14), получим матричные элементы матрицы резольвенты

$$\text{REZ}(\lambda)_{m,n} = (E - \lambda \cdot K)_{m,n}^{-1} = \left(1 - \frac{\lambda}{\mu_n}\right)^{-1} \delta_{n,m}. \quad (10.16)$$

Ядро оператора резольвенты (10.10) имеет вид

$$\text{Rez}(\lambda, x, \tilde{x}) = \sum_{n=0}^{\infty} \chi_n(x) \cdot \left(1 - \frac{\lambda}{\mu_n}\right)^{-1} \cdot \chi_n^*(\tilde{x}). \quad (10.17)$$

Из (10.17) следует, что резольвента существует, если параметр  $\lambda$  не совпадает с каким либо характеристическим числом  $\mu_n$  (10.6). С помощью матрицы  $\text{REZ}(\lambda)_{m,n}$  (10.16) получаем решение алгебраического уравнения (10.13)

$$Y_n = \left(1 - \frac{\lambda}{\mu_n}\right)^{-1} \cdot F_n.$$

Решение интегрального уравнения (10.1) получаем из (10.11) в виде обобщенного ряда Фурье

$$y(x) = \sum_{k=0}^{\infty} \left(1 - \frac{\lambda}{\mu_k}\right)^{-1} \cdot F_k \cdot \chi_k(x). \quad (10.18)$$

Изложенный метод решения требует знания собственного базиса ядра  $K(x, \tilde{x})$ , что не всегда возможно. Бесконечную систему уравнений (10.13) можно оборвать и, используя теоремы о сходимости обобщенных рядов Фурье, оценить точность обрыва. На идее конечности базисного набора основан метод построения точного решения интегральных уравнений (10.1) с *вырожденным ядром*. Рассмотрим данный метод. Выберем  $n$  линейно независимых функций (любая функция из этого набора не должна быть линейной комбинацией каких либо других функций этого же набора)

$$n := \blacklozenge$$

$$C(x) := \begin{pmatrix} \blacklozenge \\ \blacklozenge \\ \text{---} \\ \blacklozenge \end{pmatrix}$$

и рассмотрим функции

$$g(m, x) := C(x)_m.$$

Из этих функций составим вырожденное ядро

$$(10.19) \quad K(x, z) := \sum_{k=0}^n g(k, x) \cdot g(k, z).$$

Определим границы интервала

$$a := \blacklozenge \quad b := \blacklozenge.$$

Определим функцию правой части, зависящую от дополнительного параметра

$$f(x) := \blacklozenge.$$

Подставим ядро (10.19) в уравнение (10.1) и получим общий вид решения интегрального уравнения (10.1) с вырожденным ядром

$$y(x) = \lambda \cdot \sum_{k=0}^n g(k, x) \cdot A_k + f(x). \quad (10.20)$$

Здесь введен пока еще неизвестный вектор с компонентами  $A_k$ , определяемыми из соотношения

$$A_k = \int_a^b g(k, x) \cdot y(x) dx. \quad (10.21)$$

Подставим (10.20) в (10.21) и получим уравнение для неизвестных  $A_k$

$$A_k - \lambda \cdot \sum_{m=0}^n M_{k,m} \cdot A_m = FF_k. \quad (10.22)$$

Здесь введены обозначения

$$(10.23) \quad \begin{aligned} k &:= 0 \dots n & m &:= 0 \dots n \\ M_{k,m} &:= \int_a^b g(k, x) \cdot g(m, x) dx. \\ FF_k &:= \int_a^b g(k, x) \cdot f(x) dx \end{aligned}$$

Определим матрицу резольвенты  $REZg(\lambda)$  в линейной оболочке  $L_g$ , построенной на функциях  $\{g(k, x), k = 0, 1 \dots n\}$

$$(10.24) \quad REZg(\lambda) := (E - \lambda \cdot M)^{-1}.$$

Тогда решение уравнения (10.22) имеет вид

$$A(\lambda) := REZg(\lambda) \cdot FF.$$

Решение интегрального уравнения (10.20) запишется

$$(10.25) \quad y(\lambda, x) := \lambda \cdot \sum_{k=0}^n g(k, x) \cdot A(\lambda)_k + f(x).$$

Здесь дополнительно введена зависимость от параметра  $\lambda$  для удобства. Можно убедиться, что данное выражение есть решение (10.1), например, сравнив графически левую и правую части уравнения (10.1) (после подстановки).

Построим точное решение (10.1) с помощью собственного базиса ядра по формуле (10.18). Для этого найдем собственные векторы ядра, решив задачу (10.6). Собственный вектор  $\chi_m(x)$ , принадлежащий линейной оболочке  $L_g$ , построенной на функциях  $\{g(k, x), k = 0, 1 \dots n\}$ , следует искать в виде (проверяется подстановкой в (10.6))

$$\chi_m(x) = \sqrt{\mu_m} \cdot \sum_{k=0}^n X_{k,m} \cdot g(k, x), \quad m = 0, 1 \dots n. \quad (10.26)$$

Множество собственных векторов, ортогональных  $L_g$ , обозначим  $\{\varphi_p(x), p = n+1, n+2, \dots, \infty\}$ . Подставим (10.26) в (10.6) и получим алгебраическое уравнение для ортонормированных векторов нового базиса  $\vec{X}^{(m)}$  с компонентами  $X_{k,m}$

$$\begin{aligned} \mu_m \cdot M X^{(m)} &= X^{(m)}, \\ (X^{(m)}, X^{(k)}) &= \delta_{m,k}. \end{aligned} \quad (10.27)$$

Матрица  $M$  имеет матричные элементы  $M_{k,m}$  (10.23). Найдем  $\mu_m$  и  $X^{(m)}$  с помощью средств математического пакета

$$(10.28) \quad \begin{aligned} v &:= \text{eigenvals}(M) \\ X &:= \text{eigenvecs}(M) \end{aligned}$$

$$(10.29) \quad \mu_m := \frac{1}{v_m}.$$

В (10.28) получены  $n+1$  характеристическое число и столько же собственных векторов. Векторы  $\chi_m(x)$  (10.26)

$$(10.30) \quad \chi_m(x) := \sqrt{\mu_m} \cdot \sum_{k=0}^n X_{k,m} \cdot g(k, x)$$

при выполнении (10.27), ортонормированны

$$EN_{m,k} := \int_a^b \chi_m(x) \cdot \chi_k(x) dx$$

(проверить,  $EN_{m,k}$  - единичная матрица). Остальные собственные векторы ортогональны подпространству  $L_g$  и их характеристические числа равны  $\infty$ . Это следует из соотношения, где  $K(x, \tilde{x})$  - вырожденное ядро (10.19):

$$\int_a^b K(x, \tilde{x}) \cdot \varphi_p(\tilde{x}) d\tilde{x} = 0, \quad p = n+1, n+2, \dots, \infty.$$

Матрица  $X_{k,m}$  осуществляет преобразование векторов при замене базиса (исходный базис заменяется на собственный  $X^{(m)}$ ). Матрица резольвенты  $REZg(\lambda)$  (10.24) переписывается через матрицу резольвенты в собственном базисе (в подпространстве  $L_g$ ) (10.16) по формуле (запрограммировать)

$$REZg1(\lambda)_{m,k} = \sum_{p=0}^n X_{m,p} \cdot \frac{1}{1 - \lambda/\mu_p} \cdot X_{k,p}. \quad (10.31)$$

Заполним вектор  $F_k$  (10.12)

$$F_m := \int_a^b \chi_m(x) \cdot f(x) dx.$$

Решение уравнения (10.1) запишется в виде (вместо (10.25))

$$(10.32) \quad y_1(\lambda, x) := \lambda \cdot \sum_{k=0}^n \frac{1}{\mu_k - \lambda} \cdot F_k \cdot \chi_k(x) + f(x).$$

Сравним (10.25) и (10.32) графически.

Рассмотрим некоторые способы численного решения уравнения (10.1) и сравним их между собой. Построим приближенные решения для (10.1) с помощью метода сжимающих отображений (метод итераций). Определим нулевую итерацию как

$$y_0(x) = f(x)$$

и выразим через нее остальные итерации (использована сокращенная запись (10.3))

$$(10.33) \quad \begin{aligned} y_{m+1}(x) &= \lambda \cdot \hat{J} y_m(x) + f(x), \\ m &= 0, 1, 2, \dots \end{aligned}$$

В результате решения рекуррентного соотношения (10.33) получаем ряд Неймана для оператора резольвенты (10.4)

$$\hat{R}ez(\lambda) = \hat{I} + \lambda \cdot \hat{R}F(\lambda),$$

$$(10.34) \quad \hat{R}F(\lambda) = \sum_{k=0}^{\infty} \lambda^k \cdot \hat{J}^{k+1}.$$

Здесь введено обозначение для резольвенты Фредгольма  $\hat{R}F(\lambda)$ . Как известно, ряд Неймана сходится по операторной норме, если параметр  $\lambda$  удовлетворяет соотношению

$$\lambda < NN^{-1},$$

где обозначена норма интегрального ядра

$$NN := \sqrt{\int_a^b \int_a^b K^2(x, z) dx dz}.$$

Норма может также быть оценена через минимальное характеристическое число  $\mu_m$  (10.28) (проверить). Для численного решения введем сетку узлов на оси  $x$ . Обозначим количество узлов

$$Nx := \blacklozenge$$

$$p := 0..Nx \quad q := 0..Nx.$$

Шаг сетки

$$\Delta x := \frac{b-a}{Nx}.$$

Сетка

$$x_p := a + \Delta x \cdot p$$

$$z_q := a + \Delta x \cdot q$$

Матрица ядра на сетке

$$Kc_{p,q} := K(x_p, z_q).$$

Правая часть на сетке

$$fc_p := f(x_p).$$

Тогда интегральный оператор  $\hat{J}$  (10.2) превращается в алгебраический (интеграл вычисляем по прямоугольникам)

$$\int_a^b K(x_q, \tilde{x}) \cdot y(\tilde{x}) d\tilde{x} = \sum_{p=0}^{Nx} Kc_{q,p} \cdot y(x_p) \cdot \Delta x. \quad (10.35)$$

Оборванный на  $s$  – ом шаге ряд Неймана для резольвенты Фредгольма (10.34) запишем в виде

$$(10.36) \quad RFc(\lambda, s) := \sum_{j=0}^s \lambda^j \cdot Kc^{j+1} \cdot (\Delta x)^j.$$

Решение с  $s$  итерациями с помощью (10.36) имеет вид

$$yf(\lambda, s) := fc + \lambda \cdot RFc(\lambda, s) \cdot fc \cdot \Delta x.$$

Решим интегральное уравнение напрямую, заменив интеграл суммированием по формуле (10.35) (формула прямоугольников) и записав резольвенту (10.5) на сетке в виде

$$Ec := \diamond$$

$$Rezc(\lambda) := (Ec - \lambda \cdot Kc \cdot \Delta x)^{-1}.$$

Здесь  $Ec$  - единичная матрица размерности  $(Nx+1) \otimes (Nx+1)$ . Прямое решение на сетке имеет вид

$$yc(\lambda) := Rezc(\lambda) \cdot fc.$$

### Задания к работе

1. Выбрать исходные данные: размерность  $n+1$  линейной оболочки  $L_g$ , набор линейно независимых функций  $g(k, x)$ , границы интервала  $a$  и  $b$ , правую часть  $f(x)$ , число узлов сетки  $Nx$ , Составить программу в математическом пакете, используя формулы, записанные справа.
2. Получить точное решение  $y(\lambda, x)$  (10.25) и графически убедиться, что полученная функция удовлетворяет уравнению (10.1).
3. Получить характеристические числа и собственные функции вырожденного ядра (10.29), (10.30) и проверить ортонормированность базиса в  $L_g$ .



4. Запрограммировать матрицу резольвенты  $REZg_1(\lambda)$  в форме (10.31) и сравнить ее с матрицей резольвенты  $REZg(\lambda)$  в форме (10.24).
5. Убедиться графически, что функция  $y_1(\lambda, x)$  является решением (10.1) (полученным вторым способом) и сравнить это решение с решением  $y(\lambda, x)$  (10.25) графически и аналитически.
6. Оценить норму интегрального оператора и сравнить оценку с характеристическими числами ядра  $\mu_m$ .
7. Сравнить точное решение  $y(\lambda, x)$  с решениями, полученными с помощью метода итераций (использовать оборванный ряд Неймана для резольвенты Фредгольма) для различного числа итераций и различных параметрах  $\lambda$ .
8. Сравнить точное решение  $y(\lambda, x)$  с решением, полученным прямым сеточным методом. Проверить численным счетом теоретическую формулу для погрешности решения прямым методом (использовать теоретическую формулу оценки точности интегрирования по прямоугольникам).

## 11. Конечные разности

### 11.1. Свойства конечных разностей

Конечной разностью первого порядка для функции  $f(x)$  называют функцию  $\Delta f(x)$ , зависящую от точки, где разность вычисляется, и от шага сетки  $h$  [1], [15]

$$\Delta f(x) = f(x+h) - f(x).$$

Конечная разность  $n$ -ого порядка определяется естественным образом

$$\Delta^{(n)} f(x) = \Delta(\Delta^{(n-1)} f(x)),$$

$$\Delta^{(0)} f(x) = f(x).$$

Математические свойства конечных разностей во многом похожи на свойства производных функции  $f(x)$ . Приведем некоторые из них:

1.  $\Delta c = 0$  ( $c - const$ )

2. линейность операции конечной разности

$$\Delta(\alpha \cdot f(x) + \beta \cdot g(x)) = \alpha \cdot \Delta f(x) + \beta \cdot \Delta g(x)$$

3. конечная разность произведения

$$\begin{aligned} \Delta(u(x) \cdot v(x)) &= v(x+h) \cdot \Delta u(x) + u(x) \cdot \Delta v(x) = \\ &= u(x+h) \cdot \Delta v(x) + v(x) \cdot \Delta u(x) \end{aligned} \quad (11.1)$$

4. конечная разность частного

$$\Delta\left(\frac{u(x)}{v(x)}\right) = \frac{v(x) \cdot \Delta u(x) - u(x) \cdot \Delta v(x)}{v(x) \cdot v(x+h)}. \quad (11.2)$$

Зададим изучаемую функцию  $f(x)$  и шаг сетки  $h$

$$h := \blacklozenge$$

$$f(x) := \blacklozenge$$

Можно показать, что  $n$ -ая конечная разность функции  $f(x)$  вычисляется по формуле

$$(11.3) \quad \Delta(n, x) := \sum_{k=0}^n (-1)^{n-k} \cdot \text{combin}(n, k) \cdot f(x + k \cdot h).$$

Здесь введено обозначение программы счета биномиальных коэффициентов

$$\text{combin}(n, k) = \frac{n!}{k! \cdot (n-k)!}$$

и введено удобное обозначение  $n$ -ой конечной разности

$$\Delta^{(n)} f(x) \equiv \Delta(n, x).$$

Графически нетрудно убедиться в справедливости предельного перехода

$$\lim_{h \rightarrow 0} \frac{\Delta(n, x)}{h^n} = \frac{d^n}{dx^n} f(x).$$

Для произвольного многочлена  $P_n(x)$  степени  $n$  справедливо равенство

$$\Delta^{(n+1)} P_n(x) = 0.$$

Отсюда следует важное свойство конечных разностей: если функция достаточно гладкая и на данном интервале хорошо аппроксимируется полиномом степени  $n$ , то, начиная с некоторого места, ее конечные разности будут малы. Однако значения функции, вычисленные с плавающей запятой, всегда содержат некоторую ошибку (по крайней мере, ошибку округления). Если ошибка имеет систематический характер, то ее вклад в разностях может в среднем уничтожаться. Если ошибка случайна, носит характер “шума”, то она может “разрастаться” с увеличением номера разности. Изучим закономерности этого “разрастания”. Для этого предположим, что в каждом узле сетки значение функции возмущено гауссовским “белым” шумом с нулевым средним и стандартным отклонением  $\sigma$ . Введем максимальный номер  $N$  конечной разности

$$N := \blacklozenge,$$

сгенерируем  $M$  реализаций гауссовского “белого” шума

$$M := \blacklozenge \quad \sigma := \blacklozenge,$$

$$m := 0..M$$

$$\varepsilon^{(m)} := \text{norm}(N+1, 0, \sigma)$$

Здесь все реализации “шума” записаны в столбики матрицы  $\varepsilon$  размерности  $(N+1) \times (M+1)$ . Так как операция  $\Delta^{(n)} f(x)$  линейна, то для изучения закона зависимости стандартного отклонения  $n$ -ой конечной разности от ее номера  $n$  достаточно подействовать этой разностью на “шум”. Величину  $n$ -ой конечной разности на реализации “шума”  $\delta$  вычислим по формуле

$$(11.4) \quad \Delta \delta(n, \delta) := \sum_{k=0}^n (-1)^{n-k} \cdot \text{combin}(n, k) \cdot \delta_k.$$

Так как математическое ожидание гауссовского “шума” на каждом узле равно нулю, то математическое ожидание конечной разности также равно нулю (в силу линейности). Выборочное стандартное отклонение  $\sigma_s(n)$  “зашумленной”  $n$ -ой конечной разности найдем с помощью математической статистики

$$(11.5) \quad \sigma_s(n) := \sqrt{\frac{\sum_{m=0}^M \Delta\delta(n, \varepsilon^{(m)})^2}{M+1}}.$$

Можно найти теоретическую формулу для зависимости стандартного отклонения  $\sigma_t(n)$   $n$ -ой конечной разности от номера. Для этого возведем  $\Delta\delta(n, \delta)$  (11.4) в квадрат, и, воспользовавшись статистической независимостью “шума”  $\delta$  на разных узлах, получим

$$(11.6) \quad \sigma_t(n) := \sigma \cdot \sqrt{\text{combin}(2n, n)}.$$

Таким образом, “шум” в  $n$ -ой конечной разности увеличивается в  $\sqrt{\text{combin}(2n, n)}$  раз. Правильность полученной формулы можно проверить численно. Можно убедиться, что если “зашумляется” третий десятичный знак изучаемой функции ( $\sigma = 0.001$ ), то четвертые разности оказываются минимальными по величине. То есть функция хорошо аппроксимируется полиномом третьего порядка. Более высокие разности определяются в основном “шумом”. Если уровень шума исходных данных  $\sigma$  неизвестен, то его можно оценить с помощью формул (11.5), (11.6).

## 11.2. Аппроксимация функций с помощью конечных разностей

Конечные разности находят свое применение в задачах аппроксимации функций. Как известно, заданную достаточно гладкую функцию  $f(x)$  в некоторой окрестности точки, например  $x=0$ , можно с заданной точностью аппроксимировать отрезком ряда Тейлора – полиномом  $P(n, x)$  степени  $n$ . Коэффициенты подбираются так, чтобы полином и функция были бы близки. Под близостью полинома и функции будем понимать совпадение в точке  $x=0$  всех их производных до  $n$ -ой включительно. Как известно, тогда коэффициенты полинома определяются по формуле Тейлора

$$(11.7) \quad \begin{aligned} n &:= \diamond \quad x := 0 \\ p &:= 0..n \\ a_p &:= \frac{1}{p!} \cdot \frac{d^p}{dx^p} f(x). \\ P(n, x) &= \sum_{p=0}^n a_p \cdot x^p \end{aligned}$$

Вывод этой формулы существенно опирается на свойство дифференцирования степенной функции

$$\frac{d}{dx} x^n = n \cdot x^{n-1}. \quad (11.8)$$

Для конечных разностей существуют функции с похожим свойством. Это так называемые *факториальные полиномы*. Определим факториальный полином степени  $n$

$$\begin{aligned}\varphi(n, x) &= 1 \cdot x \cdot (x - h) \cdot (x - 2 \cdot h) \cdot \dots \cdot (x - (n - 1) \cdot h), \\ \varphi(0, x) &= 1.\end{aligned}\tag{11.9}$$

Запрограммируем

$$\varphi(n, x) := \begin{cases} s \leftarrow 1 \text{ if } n = 0 \\ s \leftarrow \prod_{p=0}^{n-1} (x - p \cdot h) \text{ if } n \neq 0 \\ s \end{cases} .$$

Свойство этих полиномов для конечной разности аналогично (11.8)

$$\Delta\varphi(n, x) = n \cdot h \cdot \varphi(n - 1, x).\tag{11.10}$$

По аналогии с (11.7) для функции  $f(x)$  найдем аппроксимирующий полином  $PF(n, x)$ , для которого конечные разности до  $n$  - ой включительно в точке  $x = 0$  совпадают с соответствующими разностями для аппроксимируемой функции  $f(x)$ . Этот полином имеет вид

$$\begin{aligned}(11.11) \quad af_p &:= \frac{1}{p! \cdot h^p} \cdot \Delta(p, x) \\ PF(n, x) &:= \sum_{p=0}^n af_p \cdot x^p.\end{aligned}$$

Графически можно сравнить качество аппроксимации функции  $f(x)$  в некоторой окрестности точки  $x = 0$  полиномами  $P(n, x)$  (11.7) и  $PF(n, x)$  (11.11).

### 11.3. Исчисление конечных сумм

Доопределим факториальный полином (11.9) и формулу (11.10) на случай отрицательных целых степеней  $n < 0$ . Очевидно следующее равенство

$$\varphi(n - m, x - m \cdot h) = 1 \cdot (x - m \cdot h) \cdot (x - m \cdot h - h) \cdot \dots \cdot (x - (n - 1) \cdot h).$$

Из этого соотношения следует тождество

$$\varphi(n, x) = \varphi(m, x) \cdot \varphi(n - m, x - m \cdot h).$$

Это тождество справедливо при всех  $n$ , в частности, при  $n = 0$  получаем

$$\varphi(0, x) = 1 = \varphi(m, x) \cdot \varphi(-m, x - m \cdot h).$$

Это соотношение можно переписать в следующем виде

$$\varphi(-m, x) = \frac{1}{\varphi(m, x + m \cdot h)}. \quad (11.12)$$

Формула (11.12) обобщает факториальный полином на случай целых отрицательных степеней  $m < 0$ .

Положим шаг сетки равным единице  $h=1$ . Тогда факториальный полином запишется так

$$\begin{aligned} \varphi(n, x) &= 1 \cdot x \cdot (x-1) \cdot (x-2) \cdot \dots \cdot (x-(n-1)), \\ \varphi(0, x) &= 1. \end{aligned} \quad (11.13)$$

Факториальный полином от отрицательной степени

$$\varphi(-m, x) = \frac{1}{\varphi(m, x + m)}. \quad (11.14)$$

Правило исчисления первой конечной разности

$$\Delta \varphi(n, x) = n \cdot \varphi(n-1, x). \quad (11.15)$$

Используем правило взятия первой конечной разности от частного (11.2) и возьмем конечную разность от (11.12). Получаем естественное соотношение

$$\Delta \varphi(-n, x) = -n \cdot \varphi(-n-1, x). \quad (11.16)$$

Факториальный полином степени  $n$  (11.13) это полином степени  $n$ , имеющий  $n$  целочисленных корней  $0, 1, \dots, n-1$  и целочисленные коэффициенты, которые называются *числами Стирлинга первого рода*  $S_{n,k}$

$$\varphi(n, x) = \sum_{k=0}^n S_{n,k} \cdot x^k, \quad \varphi(0, x) = 1. \quad (11.17)$$

Числа  $S_{n,k}$  связаны рекуррентным соотношением

$$\begin{aligned} S_{n+1,k} &= S_{n,k-1} - n \cdot S_{n,k}, \\ n \geq k &= 1, 2, \dots, \quad S_{0,0} = 1, \quad S_{n,0} = 0, \quad S_{n,n} = 1. \end{aligned} \quad (11.18)$$

Это соотношение можно получить с помощью очевидного тождества

$$\varphi(n+1, x) = (x-n) \cdot \varphi(n, x).$$

Числа Стирлинга второго рода  $Q_{n,k}$  являются коэффициентами в разложении степенной функции  $x^n$  по факториальным многочленам  $\varphi(n, x)$

$$x^n = \sum_{k=0}^n Q_{n,k} \cdot \varphi(k, x). \quad (11.19)$$

Числа  $Q_{n,k}$  связаны рекуррентным соотношением

$$Q_{n+1,k} = Q_{n,k-1} + k \cdot Q_{n,k},$$

$$n \geq k = 1, 2, \dots, S_{k,k} = 1, S_{n,n+k} = 0. \quad (11.20)$$

Это соотношение можно получить с помощью очевидного тождества

$$x^{n+1} = x \cdot x^n.$$

Часто необходимо конечную сумму первых  $n$  элементов некоторой последовательности выразить в виде явной функции от  $n$ . Приемы исчисления таких конечных сумм основаны на следующем свойстве конечной разности

$$\sum_{k=1}^n \Delta P(k) = (P(2) - P(1)) + \dots + (P(n+1) - P(n)) = P(n+1) - P(1). \quad (11.21)$$

Из (11.21) следует, что для того, чтобы найти конечную сумму от последовательности  $p(k)$ , достаточно найти для  $p(k)$  ее разностную первообразную  $P(k)$ , то есть представить  $p(k)$  в виде конечной разности

$$p(k) = \Delta P(k).$$

Тогда

$$\sum_{k=1}^n p(k) = \sum_{k=1}^n \Delta P(k) = P(n+1) - P(1).$$

Разностная первообразная известна для факториальных полиномов (11.15), (11.16). Используя свойство линейности, получим первообразную для линейной комбинации факториальных полиномов

$$p(k) = \sum_{m=0}^q C_m \cdot \varphi(m, k),$$

$$P(k) = \sum_{m=0}^q C_m \cdot \frac{\varphi(m+1, k)}{m+1}, \quad (11.22)$$

$$\Delta P(k) = p(k).$$

Например

$$p(k) = \varphi(1, k) = k, \quad P(k) = \frac{\varphi(2, k)}{2} = \frac{k \cdot (k-1)}{2},$$

$$\sum_{k=1}^n p(k) = \sum_{k=1}^n k = \sum_{k=1}^n \varphi(1, k) = \sum_{k=1}^n \Delta \frac{\varphi(2, k)}{2} =$$

$$= \sum_{k=1}^n \Delta \frac{k \cdot (k-1)}{2} = \frac{n \cdot (n+1)}{2}.$$

Итак, конечная сумма от линейной комбинации факториальных полиномов целочисленного аргумента (из небольшого числа слагаемых) легко вычисляется. Так как степенная функция с помощью формулы (11.19) переписывается через комбинацию факториальных полиномов, то

конечная сумма от небольшого полинома целочисленного аргумента вычисляется также просто. Легко вычисляются также конечные суммы от линейной комбинации факториальных полиномов отрицательной степени. Для таких расчетов следует использовать формулы (11.14), (11.16). Можно найти разностные первообразные и для более сложных функций. Например, для показательной функции

$$\Delta a^x = a^x \cdot (a - 1).$$

С помощью этой формулы суммируем прогрессию

$$\sum_{k=1}^n a^k = \sum_{k=1}^n \frac{\Delta a^k}{a-1} = \frac{a^{n+1} - 1}{a-1}.$$

Для тригонометрических функций справедливы соотношения

$$\cos(a \cdot x) = \frac{\Delta \sin(a \cdot (x - 0.5))}{2 \cdot \sin(0.5 \cdot a)},$$

$$\sin(a \cdot x) = \frac{\Delta \cos(a \cdot (x - 0.5))}{2 \cdot \sin(0.5 \cdot a)}.$$

С помощью этих соотношений находим конечные суммы

$$\sum_{k=0}^n \cos(a \cdot k) = \frac{\sin(a \cdot (n + 0.5))}{2 \cdot \sin(0.5 \cdot a)} - 0.5,$$

$$\sum_{k=0}^n \sin(a \cdot k) = \frac{\cos(0.5 \cdot a) - \cos(a \cdot (n + 0.5))}{2 \cdot \sin(0.5 \cdot a)}.$$

Конечные суммы можно вычислять с помощью формулы суммирования по частям, которая напоминает известную формулу интегрирования по частям. Для ее вывода используем формулу (11.1) взятия конечной разности от произведения. Для функций целочисленного аргумента эта формула приобретает вид

$$\Delta(u(k) \cdot v(k)) = u(k) \cdot \Delta v(k) + v(k+1) \cdot \Delta u(k).$$

Возьмем частичную сумму от обеих частей, используем (11.21), получим искомую формулу

$$\sum_{k=1}^n u(k) \cdot \Delta v(k) = u(n+1) \cdot v(n+1) - u(1) \cdot v(1) - \sum_{k=1}^n v(k+1) \cdot \Delta u(k). \quad (11.23)$$

Проиллюстрируем применение этой формулы на примере показательной функции. Так как  $a^k = \Delta a^k / (a - 1)$ , то получаем

$$\begin{aligned} \sum_{k=1}^{n-1} k \cdot a^k &= \frac{1}{a-1} \cdot \sum_{k=1}^{n-1} k \cdot \Delta a^k = \frac{n \cdot a^n - a}{a-1} - \frac{1}{a-1} \cdot \sum_{k=1}^{n-1} \Delta a^{k+1} = \\ &= \frac{n \cdot a^n - a}{a-1} - \frac{a^{n+2} - a^2}{(a-1)^2}. \end{aligned}$$



### Задания к работе

1. Проверить свойства конечных разностей и формулу (11.3) для их вычисления.
2. Доказать теоретическую формулу (11.6) для стандартного отклонения  $n$ -ой конечной разности, “зашумленной” гауссовским “белым” шумом. Проверить эту формулу для разных  $n$ , используя методы математической статистики.
3. Доказать, что для полинома  $PF(n, x)$  (11.11) конечные разности до  $n$ -ой включительно в точке  $x = 0$  совпадают с соответствующими разностями для аппроксимируемой функции  $f(x)$ . Графически изучить аппроксимационные возможности полинома Тейлора  $P(n, x)$  и полинома  $PF(n, x)$  для функции  $f(x)$ .
4. Используя формулы (11.13), (11.15), (11.17) - (11.22), вычислить следующие конечные суммы (результаты проверить численно):

$$\begin{aligned} & 1. \sum_{k=1}^n 8 + 6k + k^2 + 2k^3, 2. \sum_{k=1}^n 2 + 7k + 4k^2 - k^3, 3. \sum_{k=1}^n 6 - k - 8k^2 - 9k^3, \\ & 4. \sum_{k=1}^n 4 - 7k + 7k^2 + 6k^3, 5. \sum_{k=1}^n 3 - 9k + 8k^2 + 9k^3, 6. \sum_{k=1}^n -8 - 9k - 5k^2 - 4k^3, \\ & 7. \sum_{k=1}^n -9 + 5k + 6k^2 - 8k^3, 8. \sum_{k=1}^n -3 - 7k - 9k^2 - 7k^3, 9. \sum_{k=1}^n -1 + k - 7k^2 - k^3, \\ & 10. \sum_{k=1}^n 5 + 5k - 4k^2 + 6k^3, 11. \sum_{k=1}^n 8 - 7k + 9k^2 + 3k^3, 12. \sum_{k=1}^n 7 + 6k - 8k^2 - 7k^3, \\ & 13. \sum_{k=1}^n 1 - 6k + 8k^2 + 6k^3, 14. \sum_{k=1}^n -8 + 2k - 9k^2 - 9k^3, 15. \sum_{k=1}^n -k + 9k^2 - 5k^3, \\ & 16. \sum_{k=1}^n 9 + 9k - 6k^2 - 8k^3, 17. \sum_{k=1}^n -9 + 7k - 7k^2 + 6k^3, 18. \sum_{k=1}^n -8 - 7k + k^2 + 2k^3, \\ & 19. \sum_{k=1}^n 6 - 8k + 4k^2, 20. \sum_{k=1}^n 6 + 7k + 9k^2 + 2k^3, 21. \sum_{k=1}^n 7 + 5k - 4k^3, \\ & 22. \sum_{k=1}^n 4 - 2k - 5k^2 + 7k^3, 23. \sum_{k=1}^n 5 - 4k - 4k^2 - 9k^3, 24. \sum_{k=1}^n 4 - 3k - 6k^2 - 6k^3 \end{aligned}$$

5. Используя формулы (11.14), (11.16), вычислить следующие конечные суммы (результаты проверить численно):

$$\begin{aligned} & 1. \sum_{k=0}^n \frac{1}{(7+k) \cdot (8+k) \cdot (9+k)}, 2. \sum_{k=0}^n \frac{1}{(6+k) \cdot (7+k) \cdot (8+k)}, \\ & 3. \sum_{k=0}^n \frac{1}{(7+k) \cdot (8+k) \cdot (9+k) \cdot (10+k)}, 4. \sum_{k=0}^n \frac{1}{(10+k) \cdot (11+k) \cdot (12+k)}, \end{aligned}$$

$$\begin{aligned}
& 5. \sum_{k=0}^n \frac{1}{(10+k) \cdot (11+k) \cdot (12+k) \cdot (13+k)}, 6. \sum_{k=0}^n \frac{1}{k \cdot (1+k) \cdot (2+k)}, \\
& 7. \sum_{k=0}^n \frac{1}{k \cdot (1+k) \cdot (2+k) \cdot (3+k)}, 8. \sum_{k=0}^n \frac{1}{(8+k) \cdot (9+k) \cdot (10+k)}, \\
& 9. \sum_{k=0}^n \frac{1}{(8+k) \cdot (9+k) \cdot (10+k) \cdot (11+k)}, 10. \sum_{k=0}^n \frac{1}{(3+k) \cdot (4+k) \cdot (5+k)}, \\
& 11. \sum_{k=0}^n \frac{1}{(3+k) \cdot (4+k) \cdot (5+k) \cdot (6+k)}, 12. \sum_{k=0}^n \frac{1}{(2+k) \cdot (3+k) \cdot (4+k)}, \\
& 13. \frac{1}{(2+k) \cdot (3+k) \cdot (4+k) \cdot (5+k)}, 14. \sum_{k=0}^n \frac{1}{(6+k) \cdot (7+k) \cdot (8+k)}, \\
& 15. \sum_{k=0}^n \frac{1}{(6+k) \cdot (7+k) \cdot (8+k) \cdot (9+k)}, 16. \sum_{k=0}^n \frac{1}{(4+k) \cdot (5+k) \cdot (6+k)}, \\
& 17. \sum_{k=0}^n \frac{1}{(4+k) \cdot (5+k) \cdot (6+k) \cdot (7+k)}, 18. \sum_{k=0}^n \frac{1}{(1+k) \cdot (2+k) \cdot (3+k)}, \\
& 19. \sum_{k=0}^n \frac{1}{(1+k) \cdot (2+k) \cdot (3+k) \cdot (4+k)}, 20. \sum_{k=0}^n \frac{1}{(5+k) \cdot (6+k) \cdot (7+k)}, \\
& 21. \sum_{k=0}^n \frac{1}{(5+k) \cdot (6+k) \cdot (7+k) \cdot (8+k)}, 22. \sum_{k=0}^n \frac{1}{(9+k) \cdot (10+k) \cdot (11+k)}, \\
& 23. \sum_{k=0}^n \frac{1}{(9+k) \cdot (10+k) \cdot (11+k) \cdot (12+k)}, 24. \sum_{k=0}^n \frac{1}{(11+k) \cdot (12+k) \cdot (13+k)}.
\end{aligned}$$

6. С помощью формулы суммирования по частям (11.23) вычислить суммы (результаты проверить численно):

$$\begin{aligned}
& 1. \sum_{k=1}^n (8 - 6 \cdot k + 4 \cdot k^2) \cdot \cos(a \cdot k), 2. \sum_{k=1}^n (1 - k + k^2) \cdot \sin(a \cdot k), \\
& 3. \sum_{k=1}^n (8 - 5 \cdot k - 3 \cdot k^2) \cdot \cos(a \cdot k), 4. \sum_{k=1}^n (7 - 2 \cdot k + 9 \cdot k^2) \cdot \sin(a \cdot k), \\
& 5. \sum_{k=1}^n (8 + 6 \cdot k - 2 \cdot k^2) \cdot \cos(a \cdot k), 6. \sum_{k=1}^n (2 + 4 \cdot k + 6 \cdot k^2) \cdot \sin(a \cdot k), \\
& 7. \sum_{k=1}^n (8 + k - 7 \cdot k^2) \cdot \cos(a \cdot k), 8. \sum_{k=1}^n (5 + k^2) \cdot \sin(a \cdot k), \\
& 9. \sum_{k=1}^n (5 - 9 \cdot k + 4 \cdot k^2) \cdot \cos(a \cdot k), 10. \sum_{k=1}^n (7 + 7 \cdot k + 3 \cdot k^2) \cdot \sin(a \cdot k), \\
& 11. \sum_{k=1}^n (8 - 8 \cdot k - 2 \cdot k^2) \cdot \cos(a \cdot k), 12. \sum_{k=1}^n (1 - 3 \cdot k - 8 \cdot k^2) \cdot \sin(a \cdot k),
\end{aligned}$$

$$\begin{aligned}
& 13. \sum_{k=1}^n (8 + 7 \cdot k - 7 \cdot k^2) \cdot \cos(a \cdot k), 14. \sum_{k=1}^n (2 - 6 \cdot k + 2 \cdot k^2) \cdot \sin(a \cdot k), \\
& 15. \sum_{k=1}^n (4 + 2 \cdot k + 8 \cdot k^2) \cdot \cos(a \cdot k), 16. \sum_{k=1}^n (3 - 6 \cdot k + 2 \cdot k^2) \cdot \sin(a \cdot k), \\
& 17. \sum_{k=1}^n (8 - 5 \cdot k + 8 \cdot k^2) \cdot \cos(a \cdot k), 18. \sum_{k=1}^n (8 - 5 \cdot k + 2 \cdot k^2) \cdot \sin(a \cdot k), \\
& 19. \sum_{k=1}^n (3 + 9 \cdot k + 9 \cdot k^2) \cdot \cos(a \cdot k), 20. \sum_{k=1}^n (9 - 5 \cdot k + 2 \cdot k^2) \cdot \sin(a \cdot k), \\
& 21. \sum_{k=1}^n (8 - 2 \cdot k + 5 \cdot k^2) \cdot \cos(a \cdot k), 22. \sum_{k=1}^n (5 + 5 \cdot k - 2 \cdot k^2) \cdot \sin(a \cdot k), \\
& 23. \sum_{k=1}^n (2 + 5 \cdot k - 6 \cdot k^2) \cdot \cos(a \cdot k), 24. \sum_{k=1}^n (1 - 8 \cdot k + 4 \cdot k^2) \cdot \sin(a \cdot k).
\end{aligned}$$

## 12. Разностные уравнения

### 12.1. Линейные разностные уравнения

Теория *линейных разностных уравнений* во многом аналогична теории обыкновенных линейных дифференциальных уравнений. Так, задача Коши для линейных дифференциальных уравнений  $n$ -ого порядка заключается в нахождении решения уравнения, удовлетворяющего начальным условиям

$$a_0(x) \cdot y^{(n)}(x) + a_1(x) \cdot y^{(n-1)}(x) + \dots + a_n(x) \cdot y(x) = f(x)$$
$$y(x_0) = y_0, y^{(1)}(x_0) = y_1, \dots, y^{(n-1)}(x_0) = y_{n-1}$$

Здесь обозначено

$$y^{(n)}(x) = \frac{d^n}{dx^n} y(x).$$

Аналогично задача Коши для линейных разностных уравнений заключается в нахождении решения уравнения, удовлетворяющего начальным условиям

$$c_0(m) \cdot \Delta^{(n)} y(m) + c_1(m) \cdot \Delta^{(n-1)} y(m) + \dots + c_{n+1}(m) \cdot y(m) = f(m) \quad (12.1)$$
$$y(m_0) = y_0, \Delta y(m_0) = y_1, \dots, \Delta^{(n-1)} y(m_0) = y_{n-1}$$

Здесь  $\Delta^{(n)} y(x)$  -  $n$ -ая конечная разность, решение ищется на сетке  $x + m \cdot h$  с шагом  $h$ . Введено краткое обозначение аргументов функций  $f(m) \equiv f(x + m \cdot h)$ .

Если правая часть  $f(m) = 0$ , то говорят об однородном линейном разностном уравнении, в противном случае – о неоднородном. Воспользовавшись определением конечной разности, можно линейное разностное уравнение (12.1) переписать в другой форме, которая называется *рекуррентным соотношением порядка  $n$*

$$\sum_{p=0}^n b_p(m) \cdot y(m+p) = f(m). \quad (12.2)$$

Общим решением линейного разностного уравнения  $n$ -ого порядка называется функция  $y = y(m, C_1, C_2, \dots, C_n)$  - числовая последовательность, зависящая от независимой переменной (номера узла  $m$ ) и от  $n$  произвольных постоянных  $C_1, C_2, \dots, C_n$ , такая, что, во-первых, она является решением соответствующего уравнения, и, во-вторых, произвольные постоянные однозначно определяются по начальным условиям. Как и в случае линейных дифференциальных уравнений, общее решение неоднородного линейного разностного уравнения следует строить в виде

суммы общего решения однородного линейного разностного уравнения и частного решения неоднородного.

Законченный вид имеет теория линейного разностного уравнения с постоянными коэффициентами. Однородное линейное разностное уравнение с постоянными коэффициентами порядка  $n$  записывается так [15]

$$b_n \cdot y(m+n) + b_{n-1} \cdot y(m+n-1) + \dots + b_0 \cdot y(m) = 0. \quad (12.3)$$

Покажем, что любое такое уравнение имеет в качестве частного решения геометрическую прогрессию  $\{q^m, m = 0, \pm 1, \pm 2, \dots\}$ . Подставим  $q^m$  в (12.3) и получим характеристическое уравнение для знаменателя прогрессии  $q$

$$b_n \cdot q^n + b_{n-1} \cdot q^{n-1} + \dots + b_0 = 0. \quad (12.4)$$

Корни характеристического полинома (12.4) дают все возможные знаменатели прогрессий, являющихся частными решениями (12.3). Если корень имеет кратность  $r \geq 1$ , то на каждый такой корень  $q$  приходится  $r$  разных частных решений уравнения (12.3). Этот набор имеет вид

$$q^m, m \cdot q^m, m^2 \cdot q^m, \dots, m^{r-1} \cdot q^m.$$

Общее решение  $y = y(m, c_1, c_2, \dots, c_n)$  (12.3) строится, как линейная комбинация всех различных частных решений  $\{y_{(p)}(m), p = 1, 2, \dots, n\}$

$$y(m, c_1, c_2, \dots, c_n) = c_1 \cdot y_{(1)}(m) + c_2 \cdot y_{(2)}(m) + \dots + c_n \cdot y_{(n)}(m). \quad (12.5)$$

Линейно независимый набор частных решений называется фундаментальным решением однородного уравнения. Зная общее решение однородного уравнения (12.3) можно найти частное решение неоднородного уравнения

$$b_n \cdot y(m+n) + b_{n-1} \cdot y(m+n-1) + \dots + b_0 \cdot y(m) = f(m) \quad (12.6)$$

с произвольной правой частью  $f(m)$ . Для этого можно воспользоваться методом вариации произвольных постоянных, по аналогии с методом Лагранжа в теории линейных дифференциальных уравнений. Для этого в решении  $y(m, c_1, c_2, \dots, c_n)$  (12.5) объявим произвольные коэффициенты линейной комбинации функциями (последовательностями) дискретной переменной  $m$ . Подставим решение  $y(m, c_1(m), c_2(m), \dots, c_n(m))$  в уравнение (12.6). Добавим  $n-1$  дополнительное условие, связывающее  $n$  неизвестных функций  $c_1(m), c_2(m), \dots, c_n(m)$ , получим систему  $n$  уравнений для этих функций. В силу линейной независимости фундаментального решения, полученную систему уравнений можно “расцепить” и по отдельности найти все решения  $c_1(m), c_2(m), \dots, c_n(m)$ .

Рассмотрим неоднородное линейное разностное уравнение второго порядка, широко применяемое в технике и в экономических задачах

$$y_m = \alpha \cdot y_{m-1} + \beta \cdot y_{m-2} + \varepsilon_m. \quad (12.7)$$

Здесь  $\alpha$ ,  $\beta$  - коэффициенты,  $\varepsilon_m$  - “белый” шум. Это уравнение носит название стохастической авторегрессионной модели второго порядка  $AR(2)$  [16]. Соответствующее однородное уравнение имеет вид (трехчленное рекуррентное соотношение)

$$y_m - \alpha \cdot y_{m-1} - \beta \cdot y_{m-2} = 0. \quad (12.8)$$

Характеристическое уравнение (12.4) для этой модели

$$q^2 - \alpha \cdot q - \beta = 0. \quad (12.9)$$

Корни этого уравнения

$$q_{1,2} = \frac{\alpha}{2} \pm \sqrt{\left(\frac{\alpha}{2}\right)^2 + \beta}.$$

Если  $q_1 \neq q_2$ , то фундаментальная система решений имеет вид  $\{(q_1)^m, (q_2)^m\}$ , в случае  $q_1 = q_2 = q$  фундаментальная система решений запишется  $\{q^m, m \cdot q^m\}$ . В зависимости от детерминанта  $D = \alpha^2 + 4 \cdot \beta$  уравнения возможны следующие случаи.

1.  $D > 0$ . Характеристическое уравнение имеет два различных вещественных корня. Если оба корня положительны, то оба частных решения однородного уравнения – монотонные геометрические прогрессии. Если какой либо корень отрицателен, то ему соответствует знакочередующееся решение.

2.  $D = 0$ . Корни совпадают, фундаментальная система имеет вид  $\{q^m, m \cdot q^m\}$ .

3.  $D < 0$ . Характеристическое уравнение имеет пару сопряженных комплексных корней  $q_{1,2} = s_1 \pm i \cdot s_2$ . Фундаментальная система решений запишется в виде  $\{q^m \cdot \cos(m \cdot \omega), q^m \cdot \sin(m \cdot \omega)\}$ , где параметры  $q$  и  $\omega$  естественным образом связаны с  $s_1$  и  $s_2$ . Общее решение носит характер колебаний с возрастающей амплитудой  $q > 1$ , или с убывающей -  $q < 1$ .

Решение однородного уравнения (12.8) называют *равновесным*, если  $y(m)$  не зависит от  $m$ . Очевидно  $y_m = 0$  есть равновесное решение (12.8).

Решение (12.8) называют *устойчивым*, если  $\lim_{m \rightarrow \infty} y_m = 0$ . В противном

случае решение неустойчиво. Очевидно, решение устойчиво, если  $|q_{1,2}| < 1$ .

Условие устойчивости выражается через коэффициенты

$$-1 < \beta < 1 - |\alpha|. \quad (12.10)$$

Зададим параметры  $\alpha$  и  $\beta$  так, чтобы два корня  $q_1 = a, q_2 = b$  характеристического уравнения (12.9) были бы разными, вещественными и по модулю меньшими единицы

(12.11)

$$\alpha := \diamond \quad \beta := \diamond$$

$$a := \frac{\alpha}{2} + \sqrt{\left(\frac{\alpha}{2}\right)^2 + \beta},$$

$$b := \frac{\alpha}{2} - \sqrt{\left(\frac{\alpha}{2}\right)^2 + \beta}.$$

Тогда общее решение однородного уравнения (12.8) запишется

$$y_m = A \cdot a^m + B \cdot b^m. \quad (12.12)$$

Частное решение неоднородного уравнения (12.7) ищем в виде (метод вариации произвольных постоянных)

$$y_p = A_m \cdot a^m + B_m \cdot b^m. \quad (12.13)$$

Подставим (12.13) в уравнение (12.7), добавим дополнительное соотношение между  $A_m$  и  $B_m$ , упрощающее уравнение после подстановки, получаем систему уравнений для неизвестных последовательностей  $A_m$  и  $B_m$

$$\begin{cases} (A_m - A_{m-1}) \cdot a^m + (B_m - B_{m-1}) \cdot b^m = \varepsilon_m \\ (A_m - A_{m-1}) \cdot a^{m-1} + (B_m - B_{m-1}) \cdot b^{m-1} = 0 \end{cases}.$$

Или, расцепляя уравнения, получаем

$$\begin{cases} A_m = A_{m-1} + \frac{\varepsilon_m}{a^{m-1} \cdot (a-b)} \\ B_m = B_{m-1} - \frac{\varepsilon_m}{b^{m-1} \cdot (a-b)} \end{cases}.$$

Получаем линейные разностные уравнения первого порядка, которые легко решаются

$$A_m = A_0 + \frac{1}{a-b} \cdot \sum_{p=1}^m \frac{\varepsilon_p}{a^{p-1}}, \quad (12.14)$$

$$B_m = B_0 - \frac{1}{a-b} \cdot \sum_{p=1}^m \frac{\varepsilon_p}{b^{p-1}}.$$

Подставим (12.14) в (12.13), добавим к результату (12.12), получим общее решение неоднородного линейного разностного уравнения (12.7). Частное решение получим при фиксировании произвольных постоянных  $A$  и  $B$

$$A := \diamond \quad B := \diamond .$$

Зададим общее количество шагов сетки  $N$

$$N := \diamond .$$

Сгенерируем реализацию гауссовского “белого” шума длительностью в  $N + 1$  значений со средним квадратичным отклонением  $\sigma\varepsilon$

$$\sigma\varepsilon := \diamond$$

$$\varepsilon := \text{morm}(N+1, 0, \sigma\varepsilon)$$

Частное решение (12.7) запишем в виде

$$m := 1..N$$

$$y_m := A \cdot a^m + B \cdot b^m + \frac{1}{a-b} \cdot \sum_{p=1}^m \varepsilon_p \cdot (a^{m-p+1} - b^{m-p+1}).$$

$$y_0 := A + B$$

Так как математическое ожидание шума равно нулю, то среднее значение решения на  $m$  шаге  $My_m$  равно

$$My_m = A \cdot a^m + B \cdot b^m.$$

Очевидно в нашем случае (12.11)

$$\lim_{m \rightarrow \infty} My_m = 0.$$

Можно показать, что дисперсия случайного процесса  $y_m$  в пределе равна

$$\lim_{m \rightarrow \infty} Dy_m = \sigma\varepsilon^2 \cdot \frac{1+a \cdot b}{(1-a \cdot b) \cdot (1-a^2) \cdot (1-b^2)}. \quad (12.15)$$

Проверим эти утверждения с помощью методов математической статистики на одной достаточно большой реализации “белого” шума, предположив, что, начиная с достаточно большого  $m$ , распределения случайных величин  $y_m$  перестают зависеть от номера (стационарность случайного процесса  $y_m$ ).

## 12.2. Z преобразование Лапласа

Разностные уравнения (с постоянными коэффициентами) широко применяются для целей цифровой дискретной линейной обработки сигналов [17]. Для этого вводится понятие *линейной дискретной системы* или *дискретного фильтра*. Под этим понимают любую систему обработки дискретного сигнала, обладающую свойствами линейности и стационарности. На вход фильтра подается входной сигнал  $x(k)$ , с выхода снимается выходной сигнал  $y(k)$ . Здесь  $k$  - номер дискретного момента времени. Линейность означает, что отклик фильтра на сумму сигналов равен сумме откликов на каждый сигнал. Стационарность предполагает, что задержка входного сигнала приводит к задержке выходного, не влияя на форму последнего. Любой фильтр обладает частотной характеристикой, его коэффициент передачи различен на разных частотах сигнала. Передаточная характеристика фильтра может зависеть от нескольких предыдущих значений входного дискретного сигнала. То есть фильтр может обладать памятью. В рекурсивных фильтрах могут запоминаться также и несколько предыдущих значений выходного сигнала. В общем



случае выходной сигнал  $y(k)$  в линейном дискретном фильтре в момент времени  $k$  имеет вид линейной комбинации с весовыми коэффициентами некоторого количества входных отсчетов вплоть до  $k$ -ого и некоторого количества выходных отсчетов. Другими словами временной сигнал удовлетворяет разностному уравнению

$$\begin{aligned} y(k) + a_1 \cdot y(k-1) + \dots + a_n \cdot y(k-n) = \\ = \kappa \cdot (x(k) + b_1 \cdot x(k-1) \dots + b_m \cdot x(k-m)) \end{aligned} \quad (12.16)$$

Адекватный математический аппарат, применимый для анализа таких дискретных последовательностей есть  $Z$  – преобразование Лапласа (другое название – дискретное преобразование Лапласа).  $Z$  – преобразование ставит в соответствие дискретной последовательности  $x(k)$  функцию  $X(z)$ , вообще говоря, комплексной переменной  $z$ , по правилу

$$X(z) = \sum_{k=0}^{\infty} x(k) \cdot z^{-k}. \quad (12.17)$$

Естественно, степенной ряд Лорана определен в своем кольце сходимости. Дискретное  $Z$  – преобразование связано с интегральными преобразованиями Лапласа и Фурье. Преобразование Лапласа для функции  $s(t)$  определяется так

$$S(p) = \int_0^{\infty} s(t) \cdot e^{-p \cdot t} dt. \quad (12.18)$$

Сопоставим последовательности временных отсчетов сигнала  $\{x(k), k=0,1,\dots\}$ , полученных с шагом дискретизации  $T$ , временной сигнал в виде

$$s(t) = \sum_{k=0}^{\infty} x(k) \cdot \delta(t - k \cdot T). \quad (12.19)$$

Здесь  $\delta(t)$  - обобщенная  $\delta$  - функция, равная 0 во всех точках, кроме  $t=0$ , где она бесконечна. Для  $\delta(t)$  выполнено:  $\int_{-\Delta}^{\Delta} \delta(t) dt = 1$ . Подставим

(12.19) в (12.18), получим связь интегрального преобразования Лапласа  $S(p)$  и дискретного  $Z$  – преобразования в виде

$$S(p) = X(e^{p \cdot T}), \quad X(z) = S\left(\frac{1}{T} \cdot \ln(z)\right).$$

Интегральное преобразование Фурье для функции  $s(t)$  определяется так

$$\hat{S}(\omega) = \int_{-\infty}^{\infty} s(t) \cdot e^{-i \cdot \omega \cdot t} dt.$$

Интегральное преобразование Фурье  $\dot{S}(\omega)$  и дискретное  $Z$  – преобразование связаны так

$$\dot{S}(\omega) = X(e^{i\omega T}), \quad X(z) = \dot{S}\left(\frac{1}{i \cdot T} \cdot \ln(z)\right).$$

Можно получить  $Z$  – преобразования для некоторых простых последовательностей (значок  $\Leftrightarrow$  означает “сопоставить”)

1. Дискретная экспоненциальная функция

$$\begin{cases} x(k) = a^k, k \geq 0 \\ x(k) = 0, k < 0 \end{cases} \Leftrightarrow X(z) = \frac{z}{z-a}.$$

2. Единичный скачок

$$\begin{cases} x(k) = 1, k \geq 0 \\ x(k) = 0, k < 0 \end{cases} \Leftrightarrow X(z) = \frac{z}{z-1}.$$

3. Единичная импульсная функция

$$\begin{cases} x_0(k) = 1, k = 0 \\ x_0(k) = 0, k \neq 0 \end{cases} \Leftrightarrow X(z) = 1. \quad (12.20)$$

4. Знакопеременный единичный скачок

$$\begin{cases} x(k) = (-1)^k, k \geq 0 \\ x(k) = 0, k < 0 \end{cases} \Leftrightarrow X(z) = \frac{z}{z+1}.$$

5. Гиперболические и тригонометрические дискретные функции

$$\begin{cases} x(k) = \text{ch}(\alpha \cdot k), k \geq 0 \\ x(k) = 0, k < 0 \end{cases} \Leftrightarrow X(z) = \frac{z \cdot (z - \text{ch}(\alpha))}{z^2 - 2 \cdot z \cdot \text{ch}(\alpha) + 1},$$

$$\begin{cases} x(k) = \sin(\alpha \cdot k), k \geq 0 \\ x(k) = 0, k < 0 \end{cases} \Leftrightarrow X(z) = \frac{z \cdot \sin(\alpha)}{z^2 - 2 \cdot z \cdot \cos(\alpha) + 1},$$

$$\begin{cases} x(k) = \cos(\alpha \cdot k), k \geq 0 \\ x(k) = 0, k < 0 \end{cases} \Leftrightarrow X(z) = \frac{z \cdot (z - \cos(\alpha))}{z^2 - 2 \cdot z \cdot \cos(\alpha) + 1}.$$

6. Дискретная затухающая синусоида

$$\begin{cases} x(k) = a^k \cdot \cos(\omega \cdot k + \varphi), k \geq 0 \\ x(k) = 0, k < 0 \end{cases} \Leftrightarrow X(z) = \frac{z \cdot (z \cdot \cos(\varphi) - a \cdot \cos(\omega - \varphi))}{z^2 - 2 \cdot a \cdot z \cdot \cos(\omega) + 1}.$$

Дискретное  $Z$  – преобразование имеет ряд полезных свойств. Имеют место следующие теоремы:

1. Теорема запаздывания. Если  $x(k) \Leftrightarrow X(z)$  то  $x(k-m) \Leftrightarrow \frac{X(z)}{z^m}$ .

2. Теорема опережения. Если  $x(k) \Leftrightarrow X(z)$  то

$$x(k+m) \Leftrightarrow z^m \cdot \left( X(z) - \sum_{k=0}^{m-1} x(k) \cdot z^{-k} \right).$$

3. Теорема подобия. Если  $x(k) \Leftrightarrow X(z)$  то  $a^{-k} \cdot x(k) \Leftrightarrow X(a \cdot z)$ .

4.  $Z$  – преобразование суммы. Если  $x(k) \Leftrightarrow X(z)$  то

$$\tilde{x}(n) = \sum_{k=0}^n x(k) \Leftrightarrow \tilde{X}(z) = \frac{z}{z-1} \cdot X(z). \quad (12.21)$$

5. Теорема о дифференцировании  $Z$  – преобразования. Если  $x(k) \Leftrightarrow X(z)$

$$\text{то } k \cdot x(k) \Leftrightarrow -z \cdot \frac{d}{dz} X(z).$$

6. Теорема о  $Z$  – преобразовании конечной разности. Если  $x(k) \Leftrightarrow X(z)$  то  $\Delta x(k) \Leftrightarrow (z-1) \cdot X(z) - z \cdot x(0)$ . С помощью этого правила получаем  $Z$  – преобразование факториального полинома степени  $n$  от целочисленного аргумента  $k$ :  $\varphi(n, k) = k \cdot (k-1) \cdot \dots \cdot (k-n+1)$

$$x(k) = \varphi(n, k) \Leftrightarrow X(z) = \frac{n! \cdot z}{(1-z)^{n+1}}.$$

7. Теорема о свертке двух последовательностей. Сверткой  $(f * g)$  двух

последовательностей  $\left\{ \begin{array}{l} f(k), k=0,1,\dots \\ f(k)=0, k<0 \end{array} \right\}$  и  $\left\{ \begin{array}{l} g(k), k=0,1,\dots \\ g(k)=0, k<0 \end{array} \right\}$  называется

$$(f * g) = s(k) = \sum_{m=0}^k f(m) \cdot g(k-m).$$

Если  $f(k) \Leftrightarrow F(z)$ , а  $g(k) \Leftrightarrow G(z)$ , то

$$s(k) \Leftrightarrow S(z) = F(z) \cdot G(z). \quad (12.22)$$

Соответствие между дискретной последовательностью и ее  $Z$  – преобразованием является взаимно обратным. Обратное  $Z$  – преобразование осуществляет переход от прямого  $Z$  – преобразования к последовательности чисел. Формула обратного  $Z$  – преобразования

$$x(k) = \frac{1}{2 \cdot \pi \cdot i} \cdot \oint X(z) \cdot z^{k-1} dz = \sum_{\{p\}} \text{Res}(X(z) \cdot z^{k-1}). \quad (12.23)$$

Контур интегрирования охватывает все полюсы функции  $X(z) \cdot z^{k-1}$ , множество которых обозначено  $\{p\}$ . Интеграл вычисляется по вычетам. Напомним, что вычет порядка  $r$  в особой точке  $z=p$  функции  $Y(z)$  вычисляется по формуле

$$\text{Res}(Y(z))\Big|_{z=p} = \frac{1}{(r-1)!} \cdot \lim_{z \rightarrow p} \frac{d^{(r-1)}}{dz^{(r-1)}} \left\{ (z-p)^r \cdot Y(z) \right\}.$$

Например, функция  $X(z) \cdot z^{k-1} = \frac{z^{k-1}}{(z-a)^2}$ ,  $k \geq 0$  имеет полюс второго порядка в точке  $z=a$  и полюс первого порядка в точке  $z=0$  при  $k=0$ . Тогда последовательность  $x(k)$  получится

$$x(k) = \sum_{\{p\}} \text{Res}(X(z) \cdot z^{k-1}) = (k-1) \cdot a^{k-2} + a^{-2} \cdot x_0(k).$$

Здесь  $x_0(k)$  определено в (12.20).

Вернемся к разностному уравнению фильтра (12.16). В случае дискретных линейных систем с постоянными параметрами для анализа прохождения любого сигнала достаточно знать выходную реакцию фильтра на единичный импульс  $x_0(k)$  (12.20) (при нулевых начальных условиях). Другими словами, надо знать *импульсную характеристику* дискретной системы  $h(k)$ . Чтобы убедиться в этом, представим произвольный дискретный сигнал  $x(k)$  в виде суммы единичных сигналов

$$x(k) = \sum_{m=-\infty}^{\infty} x(m) \cdot x_0(k-m).$$

Тогда, в силу линейности, выходной сигнал  $y(k)$  фильтра запишется в виде линейной комбинации импульсных характеристик

$$y(k) = \sum_{m=-\infty}^k x(m) \cdot h(k-m). \quad (12.24)$$

Используя формулы (12.17), (12.22), применим  $Z$  - преобразование к (12.24)

$$Y(z) = X(z) \cdot H(z).$$

Здесь дано определение *функции передачи*

$$H(z) = \frac{Y(z)}{X(z)} = \sum_{k=0}^{\infty} h(k) \cdot z^{-k}. \quad (12.25)$$

Применим  $Z$  - преобразование к уравнению (12.16), с помощью теоремы (12.21) получаем

$$H(z) = \kappa \cdot \frac{1 + b_1 \cdot z^{-1} \dots + b_m \cdot z^{-m}}{1 + a_1 \cdot z^{-1} + \dots + a_n \cdot z^{-n}} = \kappa \cdot \frac{\prod_{q=0}^{m-1} (1 - N_q \cdot z^{-1})}{\prod_{q=0}^{n-1} (1 - P_q \cdot z^{-1})}. \quad (12.26)$$

*Частотная характеристика* (комплексный коэффициент передачи) дискретной системы определяется по формуле

$$\dot{K}(\omega) = H(e^{i\omega T}) = \sum_{k=0}^{\infty} h(k) \cdot e^{-i\omega k T}.$$

Модуль частотной характеристики  $|\dot{K}(\omega)|$  называют *амплитудно-частотной характеристикой* системы. Очевидно, нули числителя и нули знаменателя рациональной функции (12.26) определяют характерные свойства функции передачи. Сконструируем простейший фильтр и проверим его свойства в численном эксперименте. Для этого зададим параметры функции передачи (12.26), выбрав для простоты степень полинома числителя  $m=2$ , и степень полинома знаменателя  $n=2$ . Коэффициенты полиномов выберем вещественными. Полиномы числителя и знаменателя (12.26) разбиты на элементарные сомножители с, возможно, комплексными коэффициентами  $N_k$  и  $P_k$ , которые однозначно связаны с коэффициентами полиномов  $b_k$  и  $a_k$ . Для конструирования фильтра с заданными свойствами удобно задать  $N_k$  и  $P_k$  и по ним найти  $b_k$  и  $a_k$  для уравнения фильтра (12.16). Зададим параметры

$$(12.27) \quad \begin{aligned} \delta &:= \blacklozenge & \varphi &:= \blacklozenge \\ \gamma &:= \blacklozenge & \psi &:= \blacklozenge \end{aligned}$$

и через них векторы коэффициентов

$$\begin{aligned} N_0 &:= \exp\{-\delta + i \cdot \varphi\} \\ N_1 &:= \exp\{-\delta - i \cdot \varphi\} \\ P_0 &:= \exp\{-\gamma + i \cdot \psi\} \\ P_1 &:= \exp\{-\gamma - i \cdot \psi\} \end{aligned}$$

Экспоненциальная форма удобна для определения свойств фильтра. Зададим *коэффициент усиления* и шаг дискретизации временной сетки

$$\kappa := \blacklozenge \quad T := \blacklozenge.$$

Зададим параметры уравнения фильтра

$$\begin{aligned} b_0 &:= 1 & b_1 &:= -(N_0 + N_1) & b_2 &:= N_0 \cdot N_1 \\ a_0 &:= 1 & a_1 &:= -(P_0 + P_1) & a_2 &:= P_0 \cdot P_1 \end{aligned}$$

Получим *функцию передачи* фильтра

$$(12.28) \quad H(z) := \kappa \cdot \frac{\prod_{q=0}^1 (1 - N_q \cdot z^{-1})}{\prod_{q=0}^1 (1 - P_q \cdot z^{-1})}.$$

*Частотная характеристика* фильтра имеет вид

$$CHX(\omega) := H(\exp\{i \cdot \omega \cdot T\}).$$

*Амплитудно-частотная характеристика* фильтра имеет вид

$$ACHX(\omega) := |CHX(\omega)|.$$

Получим *импульсную характеристику*  $h(k)$  фильтра, совершив обратное  $Z$  – преобразование (12.23) над функцией передачи (12.28)

$$h(k) := \diamond .$$

Для проверки результата удобно подставить  $h(k)$  в правую часть (12.25) и сравнить с (12.28) графически. Зададим дискретный входной сигнал  $x(k)$

$$\chi(k) := \diamond ,$$

$$x(k) := \begin{cases} s \leftarrow 0 \\ s \leftarrow \chi(k) \text{ if } k \geq 0. \\ s \end{cases}$$

Тогда *сигнал на выходе* фильтра получим в виде

$$y(k) := \sum_{m=0}^k x(m) \cdot h(k-m).$$

Аналогичный выходной сигнал  $y1(k)$  можно сгенерировать с помощью разностного уравнения фильтра

$$y1(k) := \begin{cases} s_0 \leftarrow y(0) \\ s_1 \leftarrow y(1) \\ \text{for } m \in 2..100 \\ s_m \leftarrow -\sum_{p=1}^2 a_p \cdot s_{m-p} + \kappa \cdot \sum_{p=0}^2 b_p \cdot x_{m-p} \\ s \end{cases} .$$

Подбором параметров (12.27) можно менять форму амплитудно-частотной характеристики фильтра. Если параметр  $\delta > 0$  близок (или равен) нулю, то, меняя фазу  $\varphi$ , можно в области частоты  $\omega = \varphi/T$  получить величину характеристики, близкую к нулю. Если параметр  $\gamma > 0$  близок к нулю, то меняя фазу  $\psi$ , можно в области частоты  $\omega = \psi/T$  резонансно поднять значение пропускания фильтра. Если параметры  $\delta$  и  $\gamma$  отрицательны, то из-за положительной обратной связи фильтр становится неустойчивым. Эти свойства можно проверить численно, например, выбрав входной сигнал в виде суммы двух гармоник и убедиться, что одна гармоника может быть сильно подавлена, а вторая усилена.

### 12.3. Дискретный WKВ метод

Теория линейных разностных уравнений с переменными коэффициентами значительно сложнее теории линейных разностных уравнений с постоянными коэффициентами. Для анализа эрмитовских трехчленных линейных рекуррентных соотношений с переменными коэффициентами в работе [18] развит так называемый дискретный WKВ метод. Аббревиатура содержит начальные буквы фамилий авторов (Wentzel, Kramers, Brillouin), развивших похожий метод для обычных дифференциальных уравнений второго порядка. Эрмитовское однородное трехчленное рекуррентное соотношение записывается в виде

$$p(j) \cdot C_{j-1} + (w(j) - E) \cdot C_j + p(j+1) \cdot C_{j+1} = 0. \quad (12.29)$$

Здесь  $\{C_j, j = 0, \pm 1, \pm 2, \dots\}$  - искомая последовательность,  $\{p(j), w(j) \mid j = 0, \pm 1, \pm 2, \dots\}$  - вещественные коэффициенты, “плавно” зависящие от номера,  $E$  - вещественный спектральный параметр. Для примера рассмотрим трехчленное рекуррентное соотношение для *функций Бесселя*

$$J(\nu + 1, z) = \frac{2 \cdot \nu}{z} \cdot J(\nu, z) - J(\nu - 1, z). \quad (12.30)$$

Здесь  $\nu$  - произвольное число – порядок функции,  $z$  - ее аргумент. Ради простоты положим  $\nu$  - не целое,  $z$  - вещественное, положительное. Соотношение (12.30) генерирует последовательность функций Бесселя при заданном  $z$  и  $\nu$ , порядки которых отличаются друг от друга на произвольное целое положительное или отрицательное число. Дифференциальное уравнение для функций Бесселя имеет вид

$$\left( z^2 \cdot \frac{d^2}{dz^2} + z \cdot \frac{d}{dz} + z^2 - \nu^2 \right) J(\nu, z) = 0.$$

Функции Бесселя находят широкое применение в механике, например, для описания движения планет, в оптике, например, для описания поля в цилиндрическом волноводе и в целом ряде других случаев. Соотношение (12.30) имеет много решений, среди которых выделяют *функции Бесселя первого рода*, которые можно выразить степенным рядом

$$(12.31) \quad J(\mu, z) := \sum_{n=0}^{\infty} \frac{(-1)^n \cdot \left(\frac{z}{2}\right)^{2-n+\mu}}{n! \cdot \Gamma(n + \mu + 1)}.$$

Здесь  $\Gamma(\nu)$  - известная гамма-функция. Перепишем (12.30) в форме соотношения (12.29). Для этого обозначим

$$(12.32) \quad w(j) := -\frac{2 \cdot j}{z},$$

$$(12.33) \quad p(j) := 1.$$

Зададим спектральный параметр  $E > 0$  и аргумент  $z > 0$

$$E := \diamond,$$

$$z := \diamond.$$

Коэффициент  $w(j)$  будет “плавно” зависеть от номера, если  $z \ll 1$ .

Порядок (начальный)  $\nu$  функции Бесселя определится

$$\nu := \frac{E \cdot z}{2}.$$

Обозначим множество значений функций Бесселя (для аргумента  $z$  и порядка  $\nu$ ), которое генерируется с помощью (12.30), через

$$\{C_j = J(\nu + j, z), j = 0, \pm 1, \pm 2, \dots\}.$$

В дискретном WKВ методе вводят две так называемые “потенциальные функции”  $U^+(j)$  и  $U^-(j)$

$$(12.34) \quad \begin{aligned} U^+(j) &:= w(j) + 2 \cdot p(j), \\ U^-(j) &:= w(j) - 2 \cdot p(j). \end{aligned}$$

“Потенциальные функции” по своим свойствам напоминают потенциальную энергию в задачах одномерного движения частицы в силовом поле. Отметим следующие свойства “потенциальных функций”:

а) В задачах механики классическая скорость обращается в ноль в точках на оси  $x$ , в которых полная энергия равна потенциальной энергии (точки поворота). В дискретном случае под “скоростью” следует понимать

$$v(j) := \sqrt{(U^+(j) - E) \cdot (E - U^-(j))}.$$

Две “точки поворота” - левая  $jtL$  и правая  $jtR$ , ограничивающие область “классически разрешенного движения”, могут принадлежать либо той же самой “потенциальной функции”, либо двум разным “потенциальным функциям”. Для (12.34) “точки поворота” (левая и правая) на оси  $j$  находятся из двух уравнений

$$U^+(j) = E,$$

$$U^-(j) = E.$$

Из (12.32), (12.33), (12.34) получаем

$$jtL := \text{floor} \left( \frac{-z \cdot (E + 2)}{2} \right),$$

$$jtR := \text{floor} \left( \frac{-z \cdot (E - 2)}{2} \right).$$

б) Определим функцию  $B(j)$



$$B(j) := \frac{E - w(j)}{2 \cdot p(j)}.$$

В области изменения  $j$ , где выполнено неравенство (левее  $jtL$ )

$$E < U^-(j)$$

имеем

$$B(j) < -1.$$

Назовем эту область *первой*. Определим в этой “классически запрещенной области” “действие”  $S1(j)$  по формуле

$$S1(j) := \int_j^{jtL} \operatorname{arccch}(-B(j)) dj.$$

Согласно теории, в первой области WKB решение  $C1_j$  имеет вид

$$N1 := \diamond \quad A1 := \diamond,$$

$$NL := jtL - N1,$$

$$j1 := 0..N1,$$

$$C1_{j1} := (-1)^{NL+j1} \frac{A1}{\sqrt{|v(NL + j1)|}} \cdot \exp\{S1(NL + j1)\}.$$

WKB решение  $C1_{j1}$  построено для номеров  $j$ , изменяющихся в интервале  $NL \leq j \leq jtL$ .

Здесь выбрано *растущее по модулю решение* при удалении от левой “точки поворота” влево в согласии с поведением ряда (12.31) в первой области. Второе, линейно независимое решение, уменьшается при удалении от  $jtL$  влево. Здесь число  $N1$  определяет ширину интервала, на котором построено решение. Параметр  $A1$  подбирается оптимальным образом по графику.

с) *Вторая* область изменения номеров  $j$  определяется неравенством  $jtL + 1 \leq j \leq jtR$  (область “разрешенного движения” между “точками поворота”). Для “потенциальных функций” (12.34) во второй области выполнено

$$U^-(j) \leq E \leq U^+(j).$$

При фиксированном  $E$  зависимость решений  $C2_j$  от  $j$  носит осцилляционный характер. Определим “действие”  $S2(j)$  для “классически разрешенной области” соотношением

$$S2(j) := \int_{jtL+1}^j \arccos(B(j)) dj.$$

В этой области WKВ решение  $C2_j$  по модулю порядка единицы и оценивается по формуле

$$\begin{aligned} A2 &:= \diamond \quad \chi := \diamond, \\ N2 &:= jtR - jtL, \\ j2 &:= 0..N2 - 1, \end{aligned}$$

$$C2_{j2} := \frac{A2}{\sqrt{v(jtL + 1 + j2)}} \cdot \cos(S2(jtL + 1 + j2) + \chi).$$

Здесь  $A2$ ,  $\chi$  - произвольные постоянные амплитуда и фаза, которые необходимо подобрать по графику. “Время”  $T(E)$  движения от одной “точки поворота” до другой оценивается по формуле

$$T(E) = 2 \cdot \int_{jtL}^{jtR} \frac{dj}{v(j)}.$$

d) Для области изменения  $j$ , удовлетворяющей неравенству

$$E > U^+(j)$$

выполнено неравенство

$$B(j) > 1.$$

Назовем эту область *третьей*. Определим в этой “классически запрещенной области” “действие”  $S3(j)$  по формуле

$$S3(j) := \int_{jtR+1}^j \operatorname{arcc}h(B(j)) dj.$$

Согласно теории, в третьей области уменьшающееся по модулю WKВ решение  $C3_j$  имеет вид

$$\begin{aligned} N3 &:= \diamond \quad A3 := \diamond, \\ j3 &:= 0..N3, \end{aligned}$$

$$C3_{j3} := \frac{A3}{\sqrt{|v(jtR + 1 + j3)|}} \cdot \exp\{-S3(jtR + 1 + j3)\}.$$

Здесь выбрано уменьшающееся по модулю решение (второе решение растет при удалении от правой “точки поворота” вправо) в согласии с поведением ряда (12.31) в третьей области. Здесь число  $N3$  определяет ширину области, на которой построено решение  $C3_{j3}$ . Параметр  $A3$  подбирается оптимальным образом по графику. Таким образом, существуют три области изменения номера  $j$ , в каждой области дискретный WKВ метод позволяет построить приближенное решение рекуррентного соотношения (12.30). Решение строится в каждой области с

помощью соответствующей формулы, в окрестности “точек поворота” метод WKВ не работает, параметры решений следует подбирать с помощью графика.

Для проверки точности дискретного WKВ метода и для подбора параметров WKВ решения  $(A1, A2, A3, \chi)$  рассчитаем точные значения серии функций Бесселя для заданного выше порядка  $\nu$  и аргумента  $z$ , используя соответствующие функции из математического пакета, и с помощью ряда (12.31). Определим общую ширину  $N$  интервала изменения  $j$

$$N := N2 + N1 + N3 + 1, \\ j := 0..N,$$

и заполним вектор  $J1_j$  значениями функций Бесселя  $Jn(j + \nu, z)$  из используемого математического пакета

$$(12.35) \quad J1_j := Jn(j + \nu + NL, z)$$

и вектор  $J2_j$  значениями ряда (12.31)

$$(12.36) \quad J2_j := J(j + \nu + NL, z).$$

Полученные точные решения сравним с приближенными WKВ решениями графически.

Рекуррентные соотношения (12.2) можно использовать для вычисления элементов числовой последовательности. Вычислим с помощью рекуррентного соотношения (12.30) последовательные значения функций Бесселя для заданных выше значений  $\nu$  и аргумента  $z$ . Для этого зададим точные (начальные) значения функций  $Jn(\nu + NL, z)$ ,  $Jn(\nu + NL + 1, z)$  и, отталкиваясь от них, рассчитаем шаг за шагом остальные значения функций Бесселя до значения  $Jn(\nu + NL + N, z)$ . Это можно сделать с помощью программы

$$(12.37) \quad J3 := \left. \begin{array}{l} a_0 \leftarrow Jn(\nu + NL, z) \\ a_1 \leftarrow Jn(\nu + NL + 1, z) \\ \text{for } k \in 2..N \\ a_k \leftarrow \frac{2 \cdot (\nu + NL + k - 1)}{z} a_{k-1} - a_{k-2} \\ a \end{array} \right\} .$$

Из (12.30) видно, что при расчете  $J(\nu + m, z)$  по рекуррентной формуле ошибка округления (которая всегда присутствует в расчетах с плавающей запятой) будет нарастать приблизительно как  $m!$ . Это означает, что расчет по рекуррентным формулам следует выполнять с осторожностью. Данное утверждение следует проверить численно, сравнивая результаты расчетов

для  $J(\nu + m, z)$ , полученные по формулам (12.35) и (12.36), с расчетом по рекуррентному соотношению (12.37).

### **Задания к работе**

1. Исследовать стохастическую авторегрессионную модель второго порядка  $AR(2)$  (12.7). Доказать формулу (12.15) для дисперсии дискретного случайного процесса и проверить ее численно с помощью методов математической статистики.
2. Получить  $Z$  – преобразования для простых последовательностей 1. – 6. и доказать теоремы  $Z$  – преобразований 1. – 7.
3. Выбрать параметры (12.27) фильтра, построить его амплитудно–частотную характеристику и проверить работу фильтра с помощью численного эксперимента.
4. Применить дискретный WKВ метод для построения приближенных решений для функций Бесселя первого рода для трех областей изменения их номеров. Точность метода WKВ оценить путем сравнения графиков для приближенных и точных функций Бесселя. Изучить закономерность нарастания ошибки в значениях функций Бесселя, получаемых с помощью рекуррентного соотношения (12.37).

## Литература

- [1]. Н.С. Бахвалов, Н.П. Жидков, Г.М. Кобельков. Численные методы. Москва, “Наука”, 1987.
- [2]. В.Я. Арсенин. Методы математической физики и специальные функции. Москва, “Наука”, 1984.
- [3]. Б.М. Будаков, С.В. Фомин. Кратные интегралы и ряды. Москва, “Наука”, 1965.
- [4]. В.И. Смирнов. Курс высшей математики. Том второй. Москва, “Наука”, 1965.
- [5]. П.И. Лизоркин. Курс дифференциальных и интегральных уравнений. Москва, “Наука”, 1981.
- [6]. Л.И. Турчак, П.В. Плотников. Основы численных методов. Москва, Физматлит, 2003.
- [7]. Дж. Ортега, У. Пул. Введение в численные методы решения дифференциальных уравнений. Москва, “Наука”, 1986.
- [8]. А.А. Самарский, П.Н. Вабищевич, Е.А. Самарская. Задачи и упражнения по численным методам. Эдиториал УРСС, Москва, 2000.
- [9]. Л.Э. Эльсгольц. Дифференциальные уравнения и вариационное исчисление. Москва, “Наука”, 1987.
- [10]. И.Г. Петровский. Лекции по теории обыкновенных дифференциальных уравнений. Изд. Московского университета, 1984.
- [11]. К.Ю. Богачев. Практикум на ЭВМ. Методы приближения функций. Учебное пособие.-М.: МГУ, Мех-мат. факультет, 1998.
- [12]. С.Б. Стечкин, Ю.Н. Субботин. Сплаины в вычислительной математике. Москва, “Наука”, 1976.
- [13]. Н.В. Бутенин, Ю.И. Неймарк, Н.А.Фуфаев. Введение в теорию нелинейных колебаний. Москва, “Наука”, 1976.
- [14]. Л.Я. Цлаф. Вариационное исчисление и интегральные уравнения. Москва, “Наука”, 1970.
- [15]. В.Л. Олейник. Рекуррентные соотношения и разностные уравнения. Соросовский образовательный журнал. 2001, №3, с.114-120.
- [16]. А.Н. Ширяев. Основы стохастической финансовой математики. ФАЗИС, Москва, 1998.
- [17]. А.Б. Сергиенко. Цифровая обработка сигналов. ПИТЕР, Москва, 2006.
- [18]. P.A. Braun. Discrete semiclassical methods in the theory of Rydberg atoms in external fields. Rev.Mod.Phys., v.65, №1, 1993, p.115-161.

В 2007 году СПбГУ ИТМО стал победителем конкурса инновационных образовательных программ вузов России на 2007–2008 годы. Реализация инновационной образовательной программы «Инновационная система подготовки специалистов нового поколения в области информационных и оптических технологий» позволит выйти на качественно новый уровень подготовки выпускников и удовлетворить возрастающий спрос на специалистов в информационной, оптической и других высокотехнологичных отраслях экономики.

---

## **Кафедра высшей математики**

Кафедра высшей математики осуществляет обучение студентов всех специальностей и специализаций института по дисциплине “Высшая математика”, самой большой по объему часов. Дисциплина “Высшая математика” является базовой для математического цикла, играет исключительно важную роль в обеспечении фундаментальной подготовки. Кафедра осуществляет подготовку бакалавров и магистров по направлению 010500 “Прикладная математика и информатика”. Для студентов естественнонаучного факультета и факультета компьютерных технологий и управления читается ряд специальных дисциплин математического цикла. Кафедра является самой многочисленной в институте по числу преподавателей. Среди её сотрудников 8 докторов и 19 кандидатов наук. Преподаватели кафедры активно участвуют как в фундаментальных исследованиях по математике, физике и теоретической механике, так и в прикладных исследованиях. За последние пять лет сотрудниками опубликовано более 300 работ в отечественных и зарубежных изданиях. Сотрудники принимают активное участие в работе российских и международных научных конференций, выступают с докладами и преподают за рубежом. Профессор В.Ю. Тертычный является членом Нью-йоркской академии и Соросовским профессором. Областью научных интересов профессора А.Г. Петрашени является теория взаимодействия излучения с веществом, оптика и спектроскопия. Профессор В.П. Смирнов – специалист по теории твёрдого тела и применению теории групп в квантовой механике. Профессор В.В. Жук – один из ведущих в мире ученых в области дифференциальных уравнений. Профессор В.Ю. Тертычный занимается теорией оптимального управления механическими системами. Профессор В.М. Уздин является

известным специалистом в физике магнитных наносистем. Профессор Г.П. Мирошниченко активно занимается изучением взаимодействия излучения с веществом. Область научных интересов профессора А.П. Качалова – современные методы теории дифракции.

Кафедра высшей математики (ВМ) была организована в 1931 году. Первым заведующим кафедрой был профессор Г.Д. Гродский. С конца 1936 года кафедрой ВМ заведовал профессор И.П. Натансон, известный специалист по теории функций действительной переменной. В 1944 году заведующим кафедрой ВМ становится профессор В.А. Тартаковский (1901–1973), замечательный математик и педагог. Владимир Абрамович Тартаковский является одним из крупнейших советских алгебраистов. Им получены пользующиеся мировой известностью результаты по проблеме тождества в теории бесконечных групп. Известность получили также его работы по использованию теоретико-числовых методов в теории изгибающей поверхностей, теории диофантовых уравнений. Обладая исключительной энергией, В.А.Тартаковский уделял много внимания научной и общественной работе. Ещё в тридцатые годы он в составе комиссии Наркомпроса участвовал в разработке программы по математике для средней школы. В течение долгого времени был членом президиума Учебно-методического совета при Министерстве высшего и среднего специального образования СССР, входил в комиссию по реформе математического образования в стране. Был одним из инициаторов проведения среди школьников Ленинграда первой математической олимпиады. В.А. Тартаковский участвовал в организации Ленинградского отделения математического института им. В.А. Стеклова и был первым его директором. В разное время на кафедре ВМ преподавали академик В.И. Смирнов, член-корреспондент АН СССР Д.К. Фаддеев, проф. И.С. Соминский, проф. Ф.И. Харшиладзе, проф. А.Ф. Андреев, проф. Ю.В. Аленицын, проф. И.А. Молотков. В 1979 году кафедру возглавил доктор технических наук, профессор В.Г. Дегтярёв, специалист по теории устойчивости и теории движения космических аппаратов. С 1997 года кафедрой руководит доктор физико-математических наук, профессор И.Ю. Попов, в область научных интересов которого входят теория рассеяния, теория операторов, моделирование сложных физических систем.

**Георгий Петрович Мирошниченко,  
Александр Георгиевич Петрашень**

**Численные методы  
Учебное пособие**

В авторской редакции

Дизайн обложки

Г.П. Мирошниченко

Редакционно-издательский отдел Санкт-Петербургского государственного  
университета информационных технологий, механики и оптики

Зав. РИО

Н.Ф. Гусарова

Лицензия ИД №00408 от 05.11.99

Подписано к печати 04.05.07

Отпечатано на ризографе

Тираж 100 экз. Заказ №1031