



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

Препринты ИПМ • Препринт № 47 за 2023 г.



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

В.П. Варин

Спектральные методы решения дифференциальных и функциональных уравнений

Статья доступна по лицензии
Creative Commons Attribution 4.0 International



Рекомендуемая форма библиографической ссылки: Варин В.П. Спектральные методы решения дифференциальных и функциональных уравнений // Препринты ИПМ им. М.В.Келдыша. 2023. № 47. 30 с. <https://doi.org/10.20948/prepr-2023-47>
<https://library.keldysh.ru/preprint.asp?id=2023-47>

РОССИЙСКАЯ АКАДЕМИЯ НАУК
ОРДЕНА ЛЕНИНА ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
ИМЕНИ М.В. КЕЛДЫША

В.П. Варин

СПЕКТРАЛЬНЫЕ МЕТОДЫ РЕШЕНИЯ
ДИФФЕРЕНЦИАЛЬНЫХ И
ФУНКЦИОНАЛЬНЫХ УРАВНЕНИЙ

Москва, 2023

УДК 521.1+531.314

В.П. Варин. Спектральные методы решения дифференциальных и функциональных уравнений. Препринт Института прикладной математики им. М.В. Келдыша РАН, Москва, 2023.

Операторный подход, развитый ранее для спектрального метода, использующего полиномы Лежандра, здесь обобщается на любые системы базисных функций (необязательно ортогональных), удовлетворяющих всего двум условиям: результат операции умножения на x либо дифференцирования по x выражается в тех же функциях. Все системы классических ортогональных полиномов удовлетворяют этим условиям. В частности, построен спектральный метод, использующий полиномы Чебышева, который наиболее эффективен для численных расчетов. Этот метод применяется для численного решения линейных функциональных уравнений, которые возникают в задачах обобщенного суммирования рядов, а также в задачах аналитического продолжения дискретных отображений. Показано также, как этими методами решаются нестандартные и нелинейные краевые задачи, для которых обычные алгоритмы не применимы.

Ключевые слова. Спектральные методы, полиномы Чебышева, краевые задачи, функциональные уравнения, высокоточные вычисления.

V.P. Varin. Spectral methods for solution of differential and functional equations. Preprint of the Keldysh Institute of Applied Mathematics of RAS, Moscow, 2023.

An operational approach developed earlier for the spectral method that uses Legendre polynomials is generalized here for arbitrary systems of basis functions (not necessarily orthogonal) that satisfy only two conditions: the result of multiplication by x or of differentiation with respect to x is expressed in the same functions. All systems of classical orthogonal polynomials satisfy these conditions. In particular, we construct a spectral method that uses Chebyshev polynomials, which is most effective for numerical computations. This method is applied for numerical solution of the linear functional equations that appear in problems of generalized summation of series as well as in the problems of analytical continuation of discrete maps. We also demonstrate how these methods are used for solution of nonstandard and nonlinear boundary value problems for which ordinary algorithms are not applicable.

Key words. Spectral methods, Chebyshev polynomials, boundary value problems, functional equations, high precision computations.

© ИПМ им. М.В. Келдыша РАН.
Москва, 2023 г.

1. ВВЕДЕНИЕ

Спектральные методы решения краевых задач подразумевают разложение решений в ряды по некоторым наборам базисных функций (или пробных, в контексте методов Галеркина), в качестве которых часто используются полиномы, ортогональные на данном интервале с некоторым весом.

В [1] был предложен спектральный метод решения краевых задач для голономных ОДУ на интервале $[0, 1]$, в котором неизвестные функции раскладываются в ряды по смещенным полиномам Лежандра.

Было показано, что любая линейная краевая задача для голономного ОДУ аппроксимируется с помощью всего двух операторов, действующих в конечномерном линейном пространстве коэффициентов Фурье–Лежандра решений этих ОДУ: оператора X умножения на независимую переменную x и оператора D дифференцирования по x . Некоторые дополнительные операторы, введенные в [1], играли лишь вспомогательную роль.

Представление функций в виде разложений по (практически) произвольным наборам базисных функций оказывается вполне аналогично такому представлению в виде разложений по полиномам Лежандра или Чебышева, если для этих наборов функций существуют аналогичные операторы X и D , действующие в линейном пространстве коэффициентов разложений функций.

Для периодических функций, которые аппроксимируются с помощью тригонометрических полиномов, роль оператора X играет оператор умножения на основную гармонику.

В этой статье рассматриваются аналитические функции на отрезке $[0, 1]$, которые могут иметь особенности на концах интервала. Назовем это множество функций \mathcal{H} .

Предположим, что имеется алгоритм, который каждой функции $y(x) \in \mathcal{H}$ ставит в соответствие ее формальное разложение по некоторому набору базисных функций $\{p_n(x) \in \mathcal{H}, n = 0, 1, \dots\}$. При этом всегда $p_0(x) = 1$ (или $p_0(x) = \text{const}$), так как единица принадлежит \mathcal{H} .

Существование такого алгоритма означает, что определено линейное отображение множества \mathcal{H} в пространство коэффициентов разложений этих функций, которое мы назовем \mathcal{A} . Это отображение аналогично дискретному преобразованию Фурье периодических функций, а коэффициенты разложения $\{a_n, n = 0, 1, \dots\}$ – это аналог обычных коэффициентов Фурье. Например, в [1] рассматривалось преобразование Фурье–Лежандра.

Отличие от обычного преобразования Фурье состоит в самом алгоритме, который теперь не обязан опираться на ортогональность функций в каком-либо пространстве, а также в том, что никаких условий сходимости полученных разложений (пока) не требуется.

Поскольку реально вычисления всегда проводятся с конечными отрезками разложений, то наряду с отображением $\mathcal{H} \rightarrow \mathcal{A}$ рассматриваются проекции $\mathcal{H} \rightarrow \mathcal{A}_N$, где N – это размерность аппроксимации. То есть функция $y(x)$ аппроксимируется (в каком-то пока не определенном смысле) своим разложением

$$y(x) = \sum_{n=0}^{N-1} a_n p_n(x), \quad (1)$$

где равенство понимается в проективном смысле и чисто формально, т.е. никакой близости функции и ее разложения (в среднем, поточечно, равномерно, и т.п.) априори не требуется.

Иными словами, мы разделяем задачи вычисления разложения функции и интерпретации полученного разложения как аппроксимации этой функции. Например (см. разд. 2), ряд может быть асимптотическим и расходящимся, но породившая его аналитическая функция (решение голономного ОДУ) при этом вполне определена.

Множество функций \mathcal{H} замкнуто относительно операций дифференцирования и умножения на независимую переменную x , поэтому необходимо определить, какие линейные отображения эти операции индуцируют в пространствах коэффициентов конечномерных аппроксимаций \mathcal{A}_N . Иными словами, для аппроксимации (1) необходимо определить конечномерные отображения

$$X: \{a_n\} \rightarrow \{b_n\}, \quad \text{и} \quad D: \{a_n\} \rightarrow \{c_n\},$$

где

$$x y(x) = \sum_{n=0}^{N-1} b_n p_n(x), \quad \text{и} \quad y'(x) = \sum_{n=0}^{N-1} c_n p_n(x).$$

Существование таких операторов – это как раз те два требования к набору базисных функций $\{p_n(x)\}$, о которых говорилось в аннотации.

В случае, когда наборы базисных функций – это (классические) ортогональные полиномы, матрица X – это всегда транспонированная матрица Якоби, ассоциированная с данной системой полиномов (см. [1, 3]).

В разд. 2 мы покажем, как эта общая конструкция работает при решении голономных ОДУ для некоторых конкретных наборов базисных функций. Это делается так же, как и в [1], с использованием полиномов Лежандра.

В случае если задача не сводится к решению голономного ОДУ, предложенного формализма недостаточно для решения задачи. Это видно уже для линейного неголономного ОДУ, так как операция умножения на известную функцию $v(x)$ соответствует оператору $V = v(X)$, где $v(X)$ – это функция от матрицы (предполагая, что она существует).

Для рациональных функций $r(x)$ функция от матрицы $r(X)$ получается формальной подстановкой матрицы X вместо переменной x (см. [4]). Именно поэтому голономные ОДУ решаются относительно просто. Но для линейных неголономных ОДУ, для нелинейных ОДУ и для функциональных уравнений функции от матриц придется вычислять другим способом.

В случае если набор базисных функций – это ортогональные полиномы, функция от матрицы $v(X)$ однозначно определяется значениями функции $v(x)$ на спектре матрицы X (т.е. на спектре матрицы Якоби) с помощью интерполяционной формулы Лагранжа–Сильвестра (см. [4]), так как спектр матрицы X размерности N – это корни полинома $p_N(x)$.

Поэтому корни ортогональных полиномов – это, как правило, самые удобные узлы коллокации. Однако в общем случае это не так. Поэтому мы накладываем единственное очевидное ограничение на узлы коллокации $x_n \in [0, 1]$, $n = 1, 2, \dots, N$ – их несовпадение, т.е. $x_n \neq x_m$ при $n \neq m$.

Таким образом, имеем еще одно конечномерное представление функций $y(x) \in \mathcal{H}$,

$$y(x) = \{y_1, y_2, \dots, y_N\}, \quad y_n = y(x_n) = \sum_{k=0}^{N-1} a_k p_k(x_n). \quad (2)$$

Здесь, как и ранее в (1), мы используем способ обозначения, который в программировании называется «overloading», т.е. когда смысл символа (в данном случае $y(x)$) определяется в зависимости от контекста. Так же как оператор и его матрица обычно обозначаются одним символом.

Интерполяционная формула (1), примененная в (2), – это не что иное, как аналог обратного (дискретного) преобразования Фурье (которое мы обозначим F^{-1}), так как это линейный оператор, действующий в пространстве «коэффициентов Фурье» \mathcal{A}_N и восстанавливающий функцию $y(x)$, как таблицу ее (приближенных) значений в выбранных узлах. Таким образом, матрица этого преобразования всегда известна в явном виде для любого набора узлов x_n ,

$$F^{-1} = [p_{k-1}(x_j)]_{1 \leq j, k \leq N}, \quad F^{-1}: \mathcal{A}_N \rightarrow \mathcal{H}_N, \quad (3)$$

где \mathcal{H}_N обозначает пространство конечномерных аппроксимаций функций $y(x) \in \mathcal{H}$ в виде таблиц их (приближенных) значений в узлах x_n .

Набор базисных функций $\{p_n(x) \in \mathcal{H}, n \in \mathbb{N}_0\}$ теперь должен удовлетворять еще одному очевидному требованию: матрица F^{-1} должна быть обратимой для любой размерности аппроксимации N . Тогда мы всегда имеем аналог обычного (дискретного) преобразования Фурье, $F = (F^{-1})^{-1}$, которое преобразует функцию $y(x)$, представленную таблицей ее значений в узлах x_n , в таблицу ее «коэффициентов Фурье» $\{a_0, \dots, a_{N-1}\}$.

Представление функций $y(x) \in \mathcal{H}$ двумя способами, т.е. в виде таблиц их значений в узлах коллокации и в виде таблиц их *коэффициентов Фурье* (кавычки далее опускаем), с возможностью менять эти представления по мере надобности, позволяет использовать преимущества обоих представлений.

Голономные дифференциальные операторы, действующие в пространстве коэффициентов Фурье, допускают весьма простой учет произвольных краевых условий, осуществляемый точно так же, как это делалось в [1] для аппроксимаций полиномами Лежандра. В то же время коллокационный подход позволяет обобщить эти результаты на произвольные линейные дифференциальные операторы, а также на линейные функциональные уравнения.

Например, функция от матрицы $v(X)$, нужная для оператора умножения на функцию $v(x)$, вычисляется (аппроксимируется) как

$$v(X) = F \cdot \text{Diag}[v(x_1), \dots, v(x_N)] \cdot F^{-1}, \quad (4)$$

где точка означает умножение матриц, а $\text{Diag}[]$ обозначает диагональную матрицу.

Преимущество ортогональных полиномов над любыми другими наборами базисных функций состоит в том, что представление функции от матрицы (4) является точным, если в качестве узлов используется спектр матрицы X , т.е. корни N -го ортогонального полинома. А преимущество полиномов Чебышева над любыми другими системами ортогональных полиномов состоит в том, что спектр матрицы X известен в явном виде, а также в том, что обе матрицы, F и F^{-1} , даются явными формулами (см. разд. 3).

Это, как мы полагаем, закрывает дискуссию о том, какие полиномы, Чебышева или Лежандра, лучше в численных расчетах (см. [5]).

В разд. 2 мы покажем, как предложенный формализм работает при вычислении формальных степенных разложений решений голономных ОДУ. Эти разложения всегда асимптотические в силу самого способа их вычисления (так как метод неопределенных коэффициентов по сути асимптотический).

В то же время функция (решение ОДУ), породившая степенное разложение, является аналитической и может быть восстановлена, например, с помощью преобразования степенного ряда в факториальный (см. [6]). Оказывается (см. разд. 2), факториальное разложение решения можно вычислять непосредственно из ОДУ при подходящем наборе базисных функций.

В разд. 3 даются явные формулы для операторов X , D , F и F^{-1} , а также для некоторых других для аппроксимаций полиномами Чебышева. Эти результаты применяются для численного решения некоторых линейных функциональных уравнений в разд. 4,5.

В разд. 6 показано, как результаты разд. 3 применяются при решении нестандартных и/или нелинейных краевых задач, для которых обычные численные методы неприменимы.

2. СТЕПЕННЫЕ И ФАКТОРИАЛЬНЫЕ РАЗЛОЖЕНИЯ

Рассмотрим простейший набор базисных функций,

$$p_n(x) = x^n, \quad n \in \mathbb{N}_0, \quad (5)$$

и покажем, что наш минималистский подход к вычислению разложений является содержательным.

В данном случае пространство коэффициентов Фурье \mathcal{A} – это просто коэффициенты формальных степенных разложений в нуле тех функций из \mathcal{H} , которые имеют эти разложения.

Операторы X и D здесь очевидно существуют, а их матрицы устроены особенно просто. Приведем эти матрицы для размерности $N = 5$.

$$X = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Структура этих матриц для любой размерности N очевидна.

Фиксируем размерность аппроксимации N и введем вектор $e \in \mathcal{A}_N$, соответствующий функции $e(x) = 1$, т.е. $e = \langle 1, 0, \dots, 0 \rangle^t$.

Предложение 1. *Для любой регулярной в нуле функции $f(x)$ коэффициенты ее степенного разложения в нуле до номера $N - 1$ включительно даются вектором $f(X).e$.*

Доказательство, на самом деле, очевидно. Вектор $X.e$ соответствует функции x , т.е. это коэффициенты Фурье функции x . Это же справедливо для любой степени X^n , т.е. вектор $X^n.e$ соответствует функции x^n . Но регулярная функция $f()$ от матрицы X дается отрезком ее тейлоровского разложения в нуле до номера $N - 1$ включительно, так как матрица X нильпотентна. Поэтому $f(X).e$ – это коэффициенты Фурье функции $f(x)$, т.е. коэффициенты ее степенного разложения. Ч.Т.Д.

Аналогичное утверждение справедливо для решения любого ОДУ, которое сводится к голономному ОДУ, регулярному в нуле. Например, рассмотрим линейное ОДУ

$$(1 + x) y''(x) + x^2 y'(x) - y(x) = 1 + x, \quad (6)$$

которое не интегрируемо (по крайней мере в CAS Maple), и найдем степенное разложение в нуле его решения $y(x)$, такого, что $y(0) = a$ и $y'(0) = b$.

Для этого составим матрицу $A = (E + X).D^2 + X^2.D - E$, где E – это единичная матрица. Затем вычислим вектор правой части, $r = (E + X).e$. Затем заменим две последние строки матрицы A на строки

$$A[N, n] = p(n - 1, 0), \quad A[N - 1, n] = p'(n - 1, 0), \quad n = 1, 2, \dots, N,$$

для учета начальных значений функции $y(x)$. Затем положим $r[N] = a$ и $r[N - 1] = b$ (в том же порядке, как это делалось для матрицы A). Тогда вектор $y = A^{-1}.r$ дает коэффициенты степенного разложения (до $(N - 1)$ -го включительно) решения $y(x)$ ОДУ (6) с этими начальными данными, т.е.

$$y(x) = a + bx + \frac{1}{2}(1 + a)x^2 - \frac{1}{6}(a - b)x^3 + \frac{1}{24}(1 + 3a - 4b)x^4 + \dots$$

Гибкость этого подхода к вычислению (регулярного) степенного разложения решения ОДУ состоит в том, что мы можем наложить на решение произвольные (линейные) начальные условия. Например, потребовать, чтобы третья производная функции в нуле была равна нулю. Тогда окажется, что необходимо $a = b$.

Если же потребовать выполнения несовместных начальных условий, то матрица A просто окажется вырожденной.

Наложение краевых условий в точке $x = 1$, как мы это делали для полиномов Лежандра в [1], возможно, но довольно бессмысленно, так как аппроксимация решения краевой задачи отрезками степенных разложений крайне неэффективна даже в случае их сходимости. В разд. 3 мы используем для этого полиномы Чебышева.

По той же причине коллокационный подход здесь не имеет смысла, хотя матрица (3) здесь – это матрица Вандермонда и всегда обратима. То есть базисные функции (5) нужны только для вычисления формальных степенных разложений в нуле решений голономных ОДУ.

Также можно вычислять степенные разложения решений ОДУ, сингулярных в нуле. Такие разложения могут как сходиться, так и расходиться. Общим для них является то, что наложение n начальных условий на решение ОДУ n -го порядка (вообще говоря) невозможно, так как не все решения в нуле раскладываются в степенные ряды. Приведем два таких примера.

Рассмотрим сингулярную задачу Коши $y(0) = 0$ для уравнения

$$x^2 y'(x) + y(x) = x. \quad (7)$$

Его общим решением является сингулярная в нуле функция, имеющая при

$x > 0$ всюду расходящееся степенное разложение,

$$y(x) = \exp\left(\frac{1}{x}\right) \left(C + \text{Ei}\left(1, \frac{1}{x}\right)\right), \quad y(x) = \sum_{n=1}^{\infty} (-1)^{n-1} (n-1)! x^n, \quad (8)$$

где $\text{Ei}()$ – это интегральная экспонента.

Единственное ограниченное в нуле ($x > 0$) решение ОДУ (7) имеет константу интегрирования $C = 0$ и степенное разложение (8), поэтому конечномерная аппроксимация строится просто: берутся матрица $A = X^2.D + E$ и правая часть $r = X.e$. Никакой корректировки матрицы A и вектора r делать не нужно, так как это лишь приведет к тому, что полученная матрица будет сингулярна.

В результате для размерности аппроксимации N получим разложение (8) до степени $N - 1$ включительно.

Рассмотрим уравнение

$$x^2 y''(x) + (1 + 2x^2) y'(x) + 2y(x) = x, \quad (9)$$

которое интегрируемо в Maple. Однако общая формула первого интеграла весьма громоздка и не позволяет определить (встроенными средствами Maple) асимптотики решений в нуле. Также имеющиеся в Maple процедуры не позволяют найти степенное разложение решения, которое легко находится нашим способом.

Для этого составим матрицу $A = X^2.D^2 + (E + 2X^2).D + 2E$ и вектор правой части, $r = X.e$. Затем заменим последнюю строку матрицы A на строку $\langle 1, 0, \dots, 0 \rangle$ и положим $r[N] = a$ для учета начального значения $y(0) = a$. Тогда вектор $y = A^{-1}.r$ дает коэффициенты степенного разложения (до $(N - 1)$ -го включительно) решения ОДУ (9), т.е.

$$y(x) = a - 2ax + \frac{1}{2}(1+4a)x^2 - \frac{2}{3}(1+2a)x^3 + \frac{1}{6}(5+4a)x^4 - \frac{1}{15}(23+4a)x^5 + \dots$$

Сравнение этого решения с полученной общей квадратурой (которую мы опускаем), а также с решением однородного уравнения (9) дает асимптотику в нуле той части квадратуры, для которой имеющиеся средства Maple не применимы. Нужно просто подставить $a = 0$ в это решение.

Рассмотрим теперь набор базисных функций,

$$\left\{ q_n(x) = \frac{1}{(1 + 1/x)_n} \right\} = \left\{ 1, \frac{x}{x+1}, \frac{x^2}{(x+1)(2x+1)}, \dots \right\}, \quad n \in \mathbb{N}_0, \quad (10)$$

где $()_n$ – это символ Почхаммера. Ряд Фурье по этим функциям – это классическое (но малоизвестное) факториальное разложение.

Эти ряды асимптотически эквивалентны степенным в нуле, но, как правило, сходятся на всем интервале $x \in (0, 1]$ (см. [6] и ссылки там).

Существует простое преобразование ряда Фурье по функциям (5) в ряд Фурье по функциям (10) и наоборот. Это нижнетреугольные матрицы

$$P = \left[(-1)^{n+m} S_{n-1, m-1}^{(1)} \right]_{1 \leq m \leq n \leq N}, \quad Q = \left[(-1)^{n+m} S_{n-1, m-1}^{(2)} \right]_{1 \leq m \leq n \leq N},$$

где P преобразует степенной ряд в факториальный, а $Q = P^{-1}$ – наоборот, и $S^{(1)}$ и $S^{(2)}$ – это числа Стирлинга первого и второго рода.

Можно легко проверить, что строить график функции (8) по ее степенному разложению совершенно бесполезно, в то время как применение оператора P к вектору коэффициентов этого разложения дает вектор коэффициентов разложения по функциям (10). Это разложение приближает функцию (8) на всем интервале $[0, 1]$ достаточно хорошо уже при небольших N .

Правда, диагональные Паде-аппроксимации в данном случае лучше. Однако разложения по функциям (10) дают рациональные приближения с заранее известными полюсами.

Обозначим \tilde{X} и \tilde{D} матрицы соответствующих преобразований для функций (10), в то время как X и D – это матрицы для функций (5), как и ранее в этом разделе. Тогда можно проверить, что

$$\tilde{X} = P.X.Q, \quad \tilde{D} \approx P.D.Q,$$

т.е. матрица \tilde{X} вычисляется точно, а матрица \tilde{D} – с точностью до последней строки. Существуют и явные формулы для этих матриц, которые мы опускаем.

Таким образом, факториальные разложения решений голономных ОДУ можно получать точно так же, как мы получали степенные в этом разделе.

В заключение этого раздела отметим, что наш подход работает и для других типов разложений, например, по полиномам Бернулли. Хотя здесь пока даже неясно, в каком смысле понимать близость функции и ее разложения.

3. АППРОКСИМАЦИЯ ПОЛИНОМАМИ ЧЕБЫШЕВА

Здесь мы даем сводку формул, необходимых (и достаточных) для численного решения задач из обозначенных классов. Часть этих формул в той или иной форме уже встречалась в литературе, однако в различных других контекстах.

Под численным решением задачи мы понимаем предъявление алгоритма, позволяющего вычислить решение, в принципе, с произвольной заданной точностью. А точность полученного решения оценивается с помощью анализа коэффициентов Фурье решения.

Само же решение задачи понимается как таблица значений функции $y(x)$ в выбранных узлах либо таблица коэффициентов Фурье функции $y(x)$, по которым сама функция может быть восстановлена (интерполирована) с заданной точностью в любой точке интервала $[0, 1]$ (или $(0, 1)$).

Классические полиномы Чебышева, обозначаемые $T(n, x)$, относятся к классу ортогональных полиномов Якоби и имеют весовую функцию $w(x) = (1 - x)^\alpha (1 + x)^\beta$, $\alpha = \beta = -1/2$. Эти полиномы наиболее популярны, и мы будем использовать именно их. Полиномы Якоби с показателями $\alpha = \beta = 1/2$ обозначаются $U(n, x)$ и также используются в численном анализе. Значительно реже, но также используются полиномы Якоби с показателями $\alpha = 1/2, \beta = -1/2$, и $\alpha = -1/2, \beta = 1/2$.

Все эти четыре типа полиномов Якоби называют полиномами Чебышева, и все они имеют тригонометрическое представление и все преимущества, перечисленные в разд. 1. Так что явные формулы, приведенные в этом разделе для полиномов $T(n, x)$, существуют и для всех остальных типов полиномов Чебышева.

Однако численные эксперименты не выявили существенных преимуществ полиномов Чебышева над полиномами Лежандра (т.е. $\alpha = \beta = 0$) при решении ряда модельных задач. Поэтому мы ограничимся здесь одним набором формул для полиномов $T(n, x)$. Аналогичные формулы имеются для полиномов Лежандра (см. [1]), но получаются со значительно большими затратами.

Далее в этом разделе (и статье) мы используем систему базисных функций, состоящую из смещенных полиномов Чебышева, т.е.

$$p_0(x) = \frac{1}{2}, \quad p_n(x) = T(n, 2x - 1), \quad n \in \mathbb{N}. \quad (11)$$

Коэффициент $1/2$ здесь наследуется от обычных рядов Фурье и их тесной связи с рядами Чебышева. Избавиться от него никак невозможно, так как если взять $p_0(x) = 1$, то коэффициент $1/2$ появится далее в других местах.

Это свойство базисных функций (11) приводит к тому, что единичный вектор в пространстве коэффициентов Фурье \mathcal{A} , соответствующий функции $e(x) = 1$, имеет вид

$$e = \langle 2, 0, \dots, 0 \rangle^t.$$

Матрицы операторов X , D , а также $X^2 \approx X.X$ для системы базисных функций (11) устроены очень просто. Матрица X^2 – это несколько улучшенный вариант матрицы X^2 . Эти матрицы отличаются только в одном элементе, $16(X^2 - X.X)[N, N] = 1$. Однако, как и для полиномов Лежандра в [1], оператор X^2 все же лучше в численных расчетах, чем $X.X$, но не настолько, чтобы вводить отдельные операторы X^3, X^4 и т.д.

Структура этих матриц совершенно очевидна для любых N и не требует формализации. Для $N = 6$ они имеют вид

$$X = \frac{1}{4} \begin{bmatrix} 2 & 2 & 0 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 4 & 0 & 12 & 0 & 20 \\ 0 & 0 & 8 & 0 & 16 & 0 \\ 0 & 0 & 0 & 12 & 0 & 20 \\ 0 & 0 & 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & 0 & 0 & 20 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad X2 = \frac{1}{16} \begin{bmatrix} 6 & 8 & 2 & 0 & 0 & 0 \\ 4 & 7 & 4 & 1 & 0 & 0 \\ 1 & 4 & 6 & 4 & 1 & 0 \\ 0 & 1 & 4 & 6 & 4 & 1 \\ 0 & 0 & 1 & 4 & 6 & 4 \\ 0 & 0 & 0 & 1 & 4 & 6 \end{bmatrix}.$$

Матрица X – это транспонированная матрица Якоби, ассоциированная с системой полиномов (11), а структура матрицы D следует из свойств полиномов Чебышева (см. [7]).

Осталось привести векторы-строки для учета краевых условий, как мы это делали в [1]. Эти векторы размерности N имеют вид

$$i_m = \langle p_{n-1}^{(m)}(0) \rangle, \quad b_m = \langle p_{n-1}^{(m)}(1) \rangle, \quad n = 1, \dots, N, \quad m \in \mathbb{N}_0,$$

где m обозначает номер производной полинома $p_n(x)$ по x . Векторы i_m отвечают за граничные значения функции при $x = 0$, а b_m – за граничные значения при $x = 1$. Эти векторы устроены весьма просто:

$$\begin{aligned} i_0 &= \langle \frac{1}{2}, -1, 1, -1, \dots \rangle, & b_0 &= \langle \frac{1}{2}, 1, 1, 1, \dots \rangle, \\ i_1 &= \langle 2(-1)^n (n-1)^2 \rangle, & b_1 &= \langle 2(n-1)^2 \rangle, \\ i_2 &= \langle \frac{4}{3}(-1)^{n-1} n(n-1)^2(n-2) \rangle, & b_2 &= \langle \frac{4}{3}n(n-1)^2(n-2) \rangle. \\ & \dots & & \end{aligned}$$

Как и в [1], последние строки матриц формальных дифференциальных или разностных операторов заполняются нужными строками i_m или b_m , а также их линейными комбинациями для учета краевых условий. В том же порядке последние элементы вектора правой части краевой задачи заменяются на нужные краевые условия.

Гибкость этого подхода к учету краевых условий задачи состоит в том, что он полностью формализован и универсален, т.е. любые линейные краевые условия учитываются одинаково. А также можно учитывать произвольные линейные условия, наложенные на решение внутри интервала (этого мы не нашли в литературе). Например, для того чтобы решение принимало заданное значение $y(x) = y_0$ в какой-либо точке $x_0 \in [0, 1]$, нужно заменить одну из последних строк матрицы формального дифференциального оператора, составленного из матриц X , D и единичной матрицы E , на строку

$$\langle p_{n-1}(x_0) \rangle, \quad n = 1, \dots, N,$$

и заменить соответствующий элемент вектора правой части задачи на y_0 .

Решение голономных ОДУ этим методом полностью идентично тому, что мы делали в [1] с использованием полиномов Лежандра. Результаты по точности аппроксимации также вполне аналогичны. В частности, метод дает точные коэффициенты разложения, где это возможно. Рассмотрим краевую задачу

$$x y'(x) - y(x) = x, \quad y(1) = 0,$$

решением которой является функция $y(x) = x \ln x$.

Действуя описанным способом, для каждого N получим все коэффициенты разложения решения по смещенным полиномам Чебышева, кроме последнего, $y[N]$, что связано с методом вычисления, и кроме двух первых, поскольку они иррациональны. Таким образом, получаем

$$x \ln x = -\frac{1}{4} + \left(\frac{3}{2} - 2 \ln 2\right) x + \sum_{n=2}^{\infty} \frac{(-1)^n p_n(x)}{n(n^2 - 1)}, \quad x \in [+0, 1].$$

Перейдем к изложению коллокационного подхода, который необходим для решения более общих задач, чем решение голономных ОДУ.

В качестве узлов коллокации, очевидно, следует использовать корни смещенных полиномов Чебышева. То есть для размерности аппроксимации N ,

$$x_n = \frac{1}{2} + \frac{1}{2} \cos \left(\frac{(2n-1)\pi}{2N} \right), \quad n = 1, 2, \dots, N. \quad (12)$$

Тот факт, что ноль и единица не входят в число узлов, является, на самом деле, преимуществом, так как это позволяет решать задачи с особенностями по краям интервала.

Матрицы преобразований Фурье, т.е. оператора (3) и его обращения, здесь известны в явном виде:

$$F^{-1} = \left[\varkappa(n) \cos \left(\frac{(n-1)(2m-1)\pi}{2N} \right) \right], \quad F = \frac{2}{N} \left[\cos \left(\frac{(m-1)(2n-1)\pi}{2N} \right) \right],$$

где m и n нумеруют строки и столбцы соответственно, и

$$\varkappa(n) = \begin{cases} \frac{1}{2}, & n = 1, \\ 1, & n > 1. \end{cases}$$

Заметим, что F и F^{-1} – это практически одна и та же матрица с точки зрения вычислительных затрат. И они даны явными формулами. В то время как для полиномов Лежандра необходимо сперва вычислить корни N -го полинома численно, затем вычислить матрицу (3), а затем ее обратить.

Уже приведенных формул вполне достаточно для того, чтобы решать (теперь только в плавающей арифметике) произвольные линейные ОДУ (имеющие решения в \mathcal{H}) точно так же, как мы решали голономные, так как функции от матрицы X теперь вычисляются точно по формуле (4).

Альтернативный способ вычисления функций от матрицы X опирается на общую формулу Лагранжа–Сильвестра (см. [4]). Для функции $v(x)$, определенной в узлах (12), функция от матрицы X дается формулой

$$v(X) = \sum_{n=1}^N v(x_n) l_{N,n}(X), \quad l_{N,n}(x) = \prod_{k \neq n} \frac{x - x_k}{x_n - x_k}, \quad (13)$$

где $\{x_n, n = 1, \dots, N\}$ даны в (12), а $l_{N,n}(x)$ – это фундаментальные полиномы лагранжевой интерполяции.

Предложение 2. Для любой функции $v(x) \in \mathcal{H}$ матрицы $v(X)$, рассчитанные по формулам (4) и (13), совпадают.

Доказательство. Любая функция $v(x) \in \mathcal{H}$ определена на спектре матрицы X (12). Матрица $v(X): \mathcal{A}_N \rightarrow \mathcal{A}_N$ в (4) преобразуется в матрицу $F^{-1} \cdot v(X) \cdot F = \text{Diag}[v(x_1), \dots, v(x_N)]: \mathcal{H}_N \rightarrow \mathcal{H}_N$, т.е. матрица $v(X)$ всегда диагонализируема. Оператор умножения на функцию $v(x)$ в пространстве \mathcal{H}_N – это поточечное умножение функций. Но в пространстве \mathcal{A}_N ему соответствует оператор $v(X)$, т.е. каноническая функция от матрицы (13), полученная по формуле Лагранжа–Сильвестра (см. [4]). Ч.Т.Д.

Ну и, разумеется, еще один способ вычислить функцию от матрицы – это просуммировать матричный ряд Тейлора этой функции.

Перейдем к изложению метода решения линейных функциональных уравнений, которые возникают в задачах обобщенного суммирования рядов и в задачах аналитического продолжения дискретных отображений. Эти уравнения имеют вид

$$y(f(x)) \pm y(x) = g(x), \quad (14)$$

где $x \in [0, 1]$, $f: [0, 1] \rightarrow [0, 1]$, и $g(x) \in \mathcal{H}$. Возможны также некоторые модификации уравнения (14), которые существенно не влияют на способ его решения.

Здесь необходимо применять как спектральное, т.е. $y(x) \in \mathcal{A}_N$, так и коллокационное, т.е. $y(x) \in \mathcal{H}_N$, представления функций.

Дело в том, что при знаке минус в (14) функция $y(x)$ определена лишь с точностью до константы. Поэтому конечномерная аппроксимация уравнения (14) даст вырожденную матрицу, так же как и при решении задачи Коши для линейного ОДУ. Поэтому необходимо учитывать краевые условия решения, а это возможно сделать эффективно только при спектральном представлении функций.

С другой стороны, необходимо вычислить значения функции $f(x)$ в узлах (12) для определения линейного оператора $L: \mathcal{A}_N \rightarrow \mathcal{A}_N$, соответствующего

функции $y(f(x))$. Этот оператор имеет вид

$$L = F \cdot [p_{k-1}(f(x_j))]_{1 \leq j, k \leq N}, \quad (15)$$

т.е. представляется в виде суперпозиции оператора, аналогичного (3), и преобразования Фурье.

Далее очевидным образом составляется матрица конечномерной аппроксимации левой части (14) и вектор правой части

$$r = g(X) \cdot e = F \cdot \langle g(x_1), \dots, g(x_n) \rangle^t.$$

Затем учитываются краевые условия (или условия внутри интервала), как это делалось ранее.

Заметим, что решения $y(x)$ уравнений (14) обычно принадлежат \mathcal{H} и даже могут иметь степенное разложение в нуле. Однако эти разложения, как правило, всюду расходятся и непригодны для аппроксимации функций. Так что наш метод (или его аналоги) численного решения уравнения (14), по видимому, является безальтернативным. В следующих разделах мы приводим ряд примеров решения этих уравнений.

В заключение этого раздела заметим, что приведенный формализм имеет более широкие приложения, чем те, которые мы обозначили. В частности, мы не обсуждали функции от матрицы D . Например, экспонента от этой матрицы, которая равна

$$\exp(D) = \sum_{n=0}^{N-1} \frac{D^n}{n!}$$

в силу нильпотентности оператора D , дает оператор сдвига аргумента на единицу, т.е. дает коэффициенты Фурье функции $y(x+1)$. Таким образом, уравнение

$$(\exp(D) - E) \cdot y = n X^{n-1} \cdot e$$

имеет в качестве решения полином Бернулли $B(n, x)$ при $n < N$, при условии, что учтено начальное значение $y(0) = B_n$, как это делалось при решении голономных ОДУ.

4. АНАЛИТИЧЕСКОЕ ПРОДОЛЖЕНИЕ ДИСКРЕТНЫХ ОТОБРАЖЕНИЙ

Здесь мы рассмотрим один пример численного решения функциональных уравнений, возникающих в задачах аналитического продолжения дискретных отображений (см. [8]). Как мы полагаем, этот пример вполне типичен.

В [8] было показано, что существует аналитический первый интеграл

$$H(x, y) = -x + \log(y - y^2) + \frac{1}{y - y^2} - V(y), \quad (16)$$

дающий решение задачи об аналитическом продолжении логистического отображения

$$y_{n+1} = y_n - y_n^2, \quad y_0 \in \mathbb{C}, \quad n \in \mathbb{N}_0, \quad (17)$$

где непрерывная переменная x соответствует дискретному «времени» n , а функция $V(y)$ является голоморфной в области Фату \mathcal{F}_0 , лежащей внутри множества Жюлиа отображения (17), и удовлетворяет там уравнению

$$V(y - y^2) - V(y) = \log(y^2 - y + 1) + \frac{y(1 - y)}{y^2 - y + 1}, \quad V(0) = 0. \quad (18)$$

Уравнение (17) приводилось в [8] в качестве примера интегрируемой динамической системы, в которой реализуется динамический хаос.

Функция $V(y)$ является четной относительно точки $y = 1/2$, что видно после замены $y \rightarrow 1 - y$ в уравнении (18), но мы не будем этим пользоваться.

В [8] был предложен алгоритм, позволяющий вычислять функцию $V(y)$ в любой точке $y \in \mathcal{F}_0$ с заданной точностью. Однако этот алгоритм опирался на вычисления с использованием всюду расходящегося ряда функции $V(y)$ в нуле, что вполне аналогично применению формулы Эйлера–Маклорена. Так что альтернативный метод решения задачи здесь служит двойкой цели: демонстрирует эффективность предложенных методов и взаимно их проверяет.

Решение функционального уравнения (18) на интервале $y \in [0, 1]$ получается описанным в предыдущем разделе методом. Мы выбрали для этого QD -арифметику (т.е. примерно 64 десятичных разряда), чтобы погрешности вычислений не влияли на оценку точности аппроксимации.

В [8] мы привели значение функции $V(y)$ в точке ее глобального минимума при $y = 1/2$,

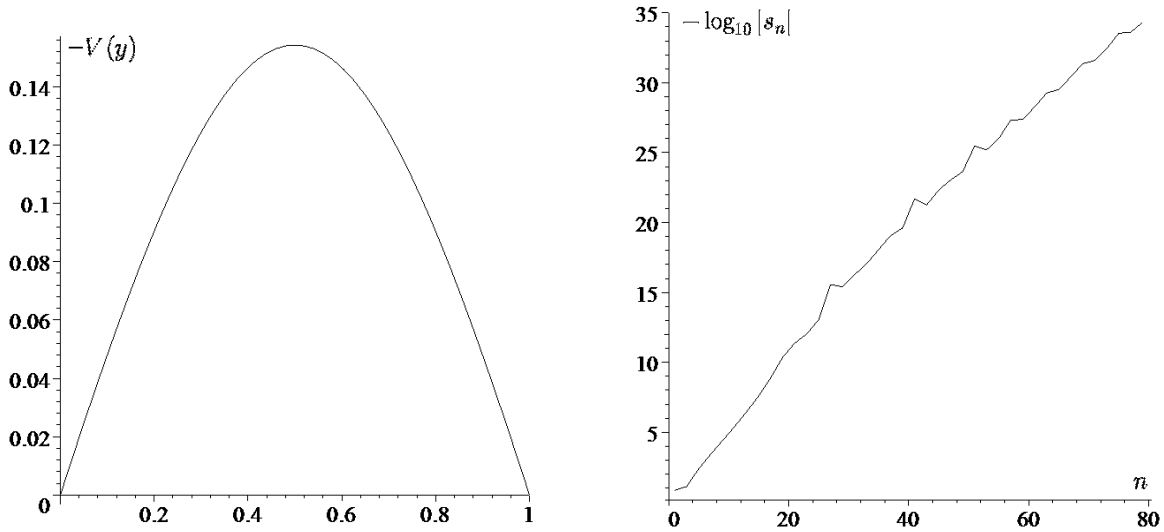
$$V(0.5) = -0.1542881472560446692780986496974,$$

с большой долей уверенности, что все десятичные знаки здесь верны. Теперь мы можем подтвердить это предположение.

На фиг. 1 слева показан график функции $-V(y)$, а на фиг. 1 справа приведены абсолютные величины четных коэффициентов Фурье полученного разложения

$$V(y) = \sum_{n=0}^{N-1} s_n p_n(y)$$

в логарифмической шкале для $N = 80$. Все нечетные коэффициенты s_{2n-1} близки к машинному нулю в силу четности функции $V(y)$ относительно точки $y = 1/2$.



Фиг. 1. Функция $V(y)$ и ее коэффициенты Фурье–Чебышева.

Фиг. 1 справа содержит информацию о достигнутой равномерной погрешности полученного решения ($\approx 10^{-35}$), а также указывает на то, что функция $V(y)$ не может быть голоморфной в полной окрестности отрезка $[0, 1]$. Иначе этот график приближался бы к некоторой прямой (экспоненциальное убывание коэффициентов Фурье). В то время как на рисунке видно, что это не так. Таким образом, (некоторые) аналитические свойства решений, полученных нашим методом, выводятся из анализа убывания их коэффициентов Фурье.

5. ФУНКЦИОНАЛЬНОЕ СУММИРОВАНИЕ РЯДОВ

Перейдем к задачам, которые возникают при функциональном суммировании рядов (см. [9]). Напомним некоторые формулы, которые нам понадобятся.

При суммировании числового ряда его частичные суммы

$$s(n) = \sum_{k=1}^n a(k) \quad (19)$$

удовлетворяют разностному уравнению

$$s(n) - s(n - 1) = a(n), \quad (20)$$

где $a(x)$ (по предположению) – это некоторая аналитическая функция при $x \in \mathbb{R}$, $x > \text{const}$. При этом ряд может быть и знакопеременным.

Сделаем замену переменных в уравнении (20), $n = 1/x$, $S(x) = s(1/x)$, $A(x) = a(1/x)$. Тогда получим функциональное уравнение

$$S(x) - S\left(\frac{x}{1-x}\right) = A(x). \quad (21)$$

Наконец, сделаем замену $x \rightarrow x/(1+x)$ в уравнении (21) и получим функциональное уравнение на интервале $[0,1]$,

$$S\left(\frac{x}{1+x}\right) - S(x) = A\left(\frac{x}{1+x}\right). \quad (22)$$

При суммировании знакопеременных рядов вместо уравнения (22) получим (см. [9]) уравнение

$$U\left(\frac{x}{1+x}\right) + U(x) = A\left(\frac{x}{1+x}\right). \quad (23)$$

Пусть исходный ряд (19) сходится. Тогда его сумма

$$C = \lim_{n \rightarrow \infty} s(n)$$

однозначно восстанавливается условием согласования двух решений разностного уравнения (20), т.е. дискретного, $s(n)$, и непрерывного, $s(1/x) = S(x)$:

$$C = s(n) - S(1/n), \quad n \in \mathbb{N}, \quad (24)$$

где решение $S(x)$ уравнения (22) получается описанным в разд. 3 способом с учетом начального значения $S(0) = 0$.

Рассмотрим Базельскую проблему в качестве примера, т.е. $A(x) = x^2$ и $C = \pi^2/6$. Действуя описанным способом (и в QD -арифметике), для размерностей аппроксимации $N = 10, 20, 40$ получим, соответственно,

$$\left|C_{10} - \frac{\pi^2}{6}\right| \approx 4.5 \times 10^{-8}, \quad \left|C_{20} - \frac{\pi^2}{6}\right| \approx 3.4 \times 10^{-14}, \quad \left|C_{40} - \frac{\pi^2}{6}\right| \approx 9.4 \times 10^{-23},$$

где погрешности получаются по формуле (24) при $n = 1$. При $n = 2, 3, \dots$ погрешности аналогичны, что также дает оценку равномерной аппроксимации функции $S(x)$, $x \in [0, 1]$.

Суммирование сходящихся рядов со сложно устроенной правой частью $A(x)$ ничем не отличается от суммирования ряда для Базельской проблемы, но только если функция $A(x/(1+x))$ хорошо аппроксимируется полиномами Чебышева на интервале $[0, 1]$. В случае, когда это не так, а это может быть только при достаточно сложной особенности функции $A(x)$ в нуле, необходимо сделать замены переменных в уравнении (22), которые учитывают эту сингулярность (см. ниже).

В случае если ряд (19) расходится в обычном смысле, формула (24) (вообще говоря) однозначно определяет его обобщенную сумму, так как асимптотика частичных сумм $s(n)$ и асимптотика решения $S(x)$ могут сократить

друг друга как разрешение неопределенности $\infty - \infty$. В этом случае, очевидно, функция $S(x)$ неограниченна на интервале $[0, 1]$ и уравнение (22) решается по-другому.

Возьмем в качестве примера обычный ряд для ζ -функции Римана

$$\zeta(\sigma + 1) = \sum_{n=1}^{\infty} \frac{1}{n^{\sigma+1}}, \quad \operatorname{Re}(\sigma) > 0, \quad (25)$$

т.е. $A(x) = x^{\sigma+1}$ в уравнении (21).

Следуя [9], сделаем в уравнении (21) замену $S(x) = x^{\sigma} V(x)$, а затем подстановку $x \rightarrow x/(1+x)$. Тогда вместо уравнения (22) получим уравнение

$$V\left(\frac{x}{1+x}\right) - (1+x)^{\sigma} V(x) = \frac{x}{1+x}, \quad (26)$$

которое не имеет логарифмической особенности в нуле.

Уравнение (26) при $\sigma \neq 0$ имеет явное решение

$$V(x) = -\frac{1}{x^{\sigma}} \zeta\left(\sigma + 1, 1 + \frac{1}{x}\right),$$

где $\zeta()$ – это ζ -функция Гурвица.

Уравнение (26) теперь аппроксимируется без учета начального значения. Это объясняется тем, что однородное уравнение (26) имеет решение

$$V_0(x) = \operatorname{const} x^{-\sigma}, \quad (27)$$

которое является полиномом только при $\sigma = 0, -1, -2, \dots$, т.е. в особой точке и там, где функция Римана определена явно и принимает рациональные значения. Там матрица аппроксимации левой части (26) будет вырождена.

В точке $x = 0$ при $\sigma \neq 0$ функция $V(x)$ имеет асимптотическое разложение в полуплоскости $\operatorname{Re}(x) > 0$,

$$V(x) = -\frac{1}{\sigma} + \frac{x}{2} - \frac{x^2(\sigma+1)}{12} + \dots = -\frac{1}{\sigma} \sum_{n=0}^{\infty} B(n) C(n+\sigma-1, n) x^n, \quad (28)$$

где $B(n)$ – это числа Бернулли, и $C()$ – биномиальный коэффициент. Этот ряд обрывается при $\sigma = -1, -2, \dots$ и дает решение уравнения (26) в полиномах Бернулли, т.е.

$$V(x) = -\frac{(-1)^{-\sigma}}{\sigma} x^{-\sigma} B(-\sigma, -1/x), \quad \sigma = -1, -2, \dots$$

Таким образом, в результате несложных преобразований мы фактически осуществили аналитическое продолжение ζ -функции Римана, определенной рядом (25), на всю комплексную плоскость, кроме особой точки $\sigma = 0$.

Уравнение (26) позволяет дать характеристику всего множества нулей ζ -функции Римана (как тривиальных, так и нетривиальных) в терминах свойств аналитического решения функционального уравнения, зависящего от параметра.

Предложение 3. *Функция $\zeta(\sigma + 1) = 0$ для тех и только тех $\sigma \in \mathbb{C}$, для которых решение уравнения (26) удовлетворяет набору тождеств*

$$x^\sigma V(x) = \sum_{k=1}^n \frac{1}{k^{\sigma+1}}, \quad x = 1/n, \quad n \in \mathbb{N},$$

т.е. когда функция $x^\sigma V(x)$ интерполирует данную сумму на всем интервале $[0, 1]$, при условии, что $V(x)$ имеет асимптотику (28).

Доказательство очевидно, так как это всего лишь переформулировка формулы суммирования (24). Условие на асимптотику нельзя опустить, так как при $\operatorname{Re}(\sigma) < 0$ всегда существует решение уравнения (26), такое, что $V(1) = 1$ и $V(+0) = -1/\sigma$, в силу (27). Ч.Т.Д.

Например, для $\sigma = -1, -2, \dots$ имеем согласно (28), $V(0) = -1/\sigma$, и

$$V(1) = 1 - \zeta(\sigma + 1) \in \left\{ \frac{3}{2}, \frac{13}{12}, 1, \frac{119}{120}, 1, \frac{253}{252}, 1, \frac{239}{240}, 1, \frac{133}{132}, 1, \frac{32069}{32760}, 1, \frac{13}{12}, \dots \right\},$$

где единицы соответствуют тривиальным нулям ζ -функции Римана.

При этих значениях σ функция $V(x)$, будучи полиномом, находится точно нашим методом (при $N \geq -\sigma$), если выбрать соответствующее краевое значение $V(1)$. Термин «точно» здесь понимается в обозначенном ранее смысле «сколь угодно точно», так как вычисления проводятся в плавающей арифметике, а коэффициенты решения являются рациональными числами.

Здесь мы обращаем внимание на тот факт, что иррациональности, которые использовались в конструировании матриц аппроксимации нашим методом, в результате сокращают друг друга. Это же справедливо для вычислений с полиномами Лежандра.

Если для $\sigma = -1, -2, \dots$ назначить краевое значение $V(1) = 1$, то функция $V(x)$ по-прежнему находится точно в виде полинома, и с тем же значением $V(0) = -1/\sigma$, так как отличие от предыдущего полинома только в мономе $\operatorname{const} x^{-\sigma}$. Однако асимптотика (28) будет нарушена.

Для σ , соответствующих нетривиальным нулям ζ -функции Римана, краевые условия функции $V(x)$ не могут быть назначены, так как моном $\operatorname{const} x^{-\sigma}$ уже не будет полиномом, т.е. формально существующее решение $V(x) \in \mathcal{H}$ уравнения (26) с этими свойствами не может быть найдено нашим методом (что скорее преимущество, чем недостаток).

Напомним, что аппроксимация уравнения (26) в пространстве \mathcal{A}_N согласно формулам разд. 3 имеет вид

$$A.s = F. \left([p_{k-1}(f_j)]_{1 \leq j, k \leq N} - \text{Diag} [(1+x_1)^\sigma, \dots, (1+x_N)^\sigma] \cdot F^{-1} \right) \cdot s = r,$$

где вектор $f \in \mathcal{H}_N$ имеет вид

$$f = \left\langle \frac{x_1}{1+x_1}, \dots, \frac{x_N}{1+x_N} \right\rangle^t,$$

а вектор $r = F.f = X.(E+X)^{-1}.e \in \mathcal{A}_N$.

Тогда вектор коэффициентов Фурье–Чебышева решения получается как $s = A^{-1}.r$, причем никакой модификации матрицы A и вектора r для учета краевых условий не требуется (так как мы уже разобрались с тривиальными нулями).

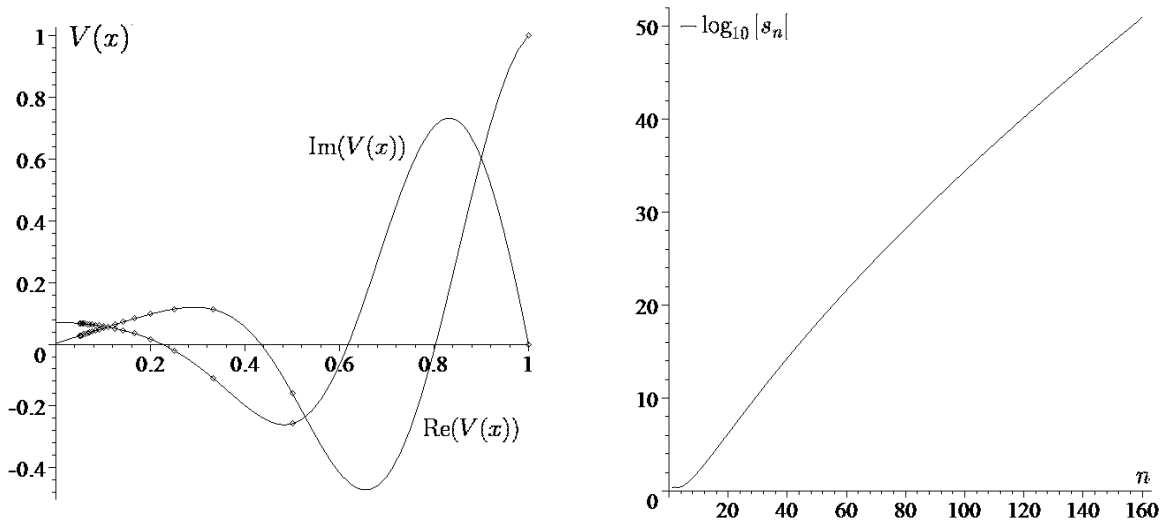
Альтернативный (и более эффективный) способ вычислить то же самое – это использовать коллокационное представление функций и операторов. Но это работает только потому, что краевые условия здесь можно не учитывать.

В пространстве \mathcal{H}_N эта задача аппроксимируется как

$$B.v = \left([p_{k-1}(f_j)]_{1 \leq j, k \leq N} \cdot F - \text{Diag} [(1+x_1)^\sigma, \dots, (1+x_N)^\sigma] \right) \cdot v = f,$$

где вектор f определен выше, а неизвестный вектор $v = \langle V(x_1), \dots, V(x_N) \rangle^t$. Затем находим $v = B^{-1}.f$, а затем вычисляем вектор коэффициентов Фурье решения, $s = F.v$.

Можно проверить, что для любой размерности аппроксимации N результаты, полученные в \mathcal{A}_N или в \mathcal{H}_N , всегда различаются лишь в пределах машинной точности. Это является следствием хорошей обусловленности матриц F, F^{-1} .



Фиг. 2. Функция $V(x)$ и ее коэффициенты Фурье–Чебышева.

Выберем в качестве примера первый нетривиальный нуль ζ -функции Римана, $\zeta(\sigma + 1) = 0$, который в QD -арифметике (в CAS Maple) имеет вид

$$\sigma = -0.5 + i 14.13472514173469379045725198356247027078425711569924317568556746,$$

и найдем численное решение задачи для $N = 160$. Результат представлен на фиг. 2.

На фиг. 2 слева показаны графики вещественной и мнимой части функции $V(x)$, посчитанные по ее интерполяционному полиному Чебышева

$$V(x) = \sum_{n=0}^{N-1} s_n p_n(x), \quad (29)$$

а также дискретные точки для $n = 1, 2, \dots, 20$, в которых функция $V(x)$ известна явно, согласно предложению 3.

На фиг. 2 справа приведены абсолютные величины коэффициентов s_n в логарифмической шкале. Фиг. 2 справа содержит информацию о достигнутой равномерной погрешности полученного решения ($\approx 10^{-50}$), что также подтверждается краевыми значениями, вычисленными по формуле (29), т.е. $V(1) = 1 - \zeta(\sigma + 1) = 1$ и $V(0) = -1/\sigma$ с данной точностью.

Расчеты при $N = 80$ дают точность $\approx 7 \times 10^{-29}$, что также считывается с фиг. 2.

Заметим, что точно так же и с той же точностью вычисляются значения ζ -функции Римана $\zeta(\sigma + 1) \approx 1 - V(1)$ для всех близких значений σ . Для других σ размерность аппроксимации и величину разрядной сетки, возможно, придется изменить с учетом требуемой точности.

Таким образом, мы имеем альтернативный способ вычисления ζ -функции Римана с контролируемой точностью (хотя и не самый эффективный).

Также на фиг. 2 видно, что экспоненциального убывания коэффициентов Фурье здесь ожидать не следует, что вполне аналогично фиг. 1. Отличие состоит в том, что здесь коэффициенты Фурье функции $V(x)$ ведут себя весьма регулярно.

Ну и, наконец, заметим, что если бы расчеты велись, например, в DD -арифметике (≈ 32 десятичных разряда), то фиг. 2 справа оказался бы обрезанным по высоте ординаты ≈ 32 , что также является полезной информацией, указывающей на бесполезность (в данном примере и этой арифметике) увеличения размерности аппроксимации больше $N \approx 100$.

В завершение этого раздела дадим пример суммирования расходящегося знакопеременного ряда нашим методом. Рассмотрим классическую задачу вычисления константы Эйлера–Гомперца $\delta = \exp(1) \text{Ei}(1, 1)$ как суммы всюду расходящегося ряда (8) при $x = 1$.

Мы показали в [9], что этот ряд суммируется точно методом функционального суммирования, но здесь нас интересует численное решение задачи.

Действуя описанными в [9] способом, получим функциональное уравнение

$$W\left(\frac{x}{1+x}\right) + xW(x) = 1, \quad x \in [0, 1], \quad (30)$$

а также набор тождеств

$$\delta = \sum_{n=1}^m (-1)^{n-1} (n-1)! + (-1)^m (m-1)! W\left(\frac{1}{m}\right), \quad m \in \mathbb{N}. \quad (31)$$

Уравнение (30) имеет точное решение

$$W(x) = 1 - \exp(1) \operatorname{Ei}\left(\frac{1}{x}, 1\right), \quad (32)$$

что легко проверяется. Так что численное решение уравнения (30) – это по сути задача аппроксимации этой функции полиномами Чебышева.

Здесь следует отметить одно обстоятельство, относящееся к отличию математически точного решения от численного, полученного даже, возможно, с очень большой точностью.

Дело в том, что уравнение (31) выполняется точно для функции $W(x)$ в (32) для $m \in \mathbb{N}$. Однако подстановка в (31) численной аппроксимации функции $W(x)$ приводит к умножению (даже очень малой) погрешности на $(m-1)!$, так что процесс быстро расходится. Но для небольших m в (31) это несущественно.

В результате для размерностей аппроксимации $N = 10, 20, 40$ получим, соответственно,

$$|\delta_{10} - \delta| \approx 1.8 \times 10^{-5}, \quad |\delta_{20} - \delta| \approx 6.9 \times 10^{-9}, \quad |\delta_{40} - \delta| \approx 1.1 \times 10^{-13},$$

где δ_N вычисляется по формуле (31) при $m = 1$.

Заметим, что диагональная $[N, N]$ Паде-аппроксимация ряда (8) дает приближение для δ на 2–3 порядка хуже.

6. РЕШЕНИЕ НЕЛИНЕЙНЫХ И НЕСТАНДАРТНЫХ ЗАДАЧ

Здесь мы дадим два примера решения нелинейных задач нашим методом, которые примечательны только тем, что стандартные давно известные численные методы (реализованные в CAS) здесь либо неприменимы, либо неэффективны и весьма трудоемки.

До сих пор мы рассматривали только линейные задачи, так что, казалось бы, решение нелинейных задач является существенным обобщением. Однако это не так.

Решение нелинейной задачи в смысле, определенном в начале разд. 3, – это всегда итерационный процесс, в котором участвуют только линейные операторы. Например, метод Ньютона можно охарактеризовать как метод последовательных линеаризаций.

Рассмотрим задачу Коши $y(0) = 0$ для уравнения

$$y'(x) = y^3(x) + x. \quad (33)$$

Это уравнение примечательно тем, что Maple зависает при попытке его проинтегрировать в явном виде (что крайне необычно).

Поставим задачу определить значение $y(1)$ (если оно существует) с большой (скажем, $\approx 10^{-50}$) и контролируемой точностью. Заметим, что стандартная процедура численного интегрирования в Maple (без дополнительных настроек) даст точность $\approx 1.3 \times 10^{-6}$, причем без указания, что точность именно такова.

В Maple существует способ проинтегрировать это уравнение методом тейлоровских разложений с данной точностью. Однако эта процедура для пользователя является «черным ящиком», и нет никаких подтверждений достигнутой точности.

Решение нелинейных задач итерационными методами возможно, в принципе, в пространстве коэффициентов Фурье \mathcal{A}_N . Но коллокационный подход, т.е. когда функции представлены таблицами их значений в узлах (12), здесь значительно удобнее. Однако и спектральное представление функций по-прежнему необходимо для учета краевых и иных условий.

Самый простой способ решить уравнение (33) – это применить итерации Пикара, которые здесь, очевидно, сходятся.

Выберем QD -арифметику и $N = 100$ (чтобы технические моменты не влияли на точность аппроксимации).

Возьмем матрицу D и заменим ее последнюю строку на вектор i_0 для учета начального значения. Тогда получим обратимый оператор дифференцирования $D_0: \mathcal{A}_N \rightarrow \mathcal{A}_N$. Поэтому матрицы

$$D_p = F^{-1} \cdot D_0 \cdot F: \mathcal{H}_N \rightarrow \mathcal{H}_N, \quad \text{и} \quad I_p = D_p^{-1} = F^{-1} \cdot D_0^{-1} \cdot F: \mathcal{H}_N \rightarrow \mathcal{H}_N$$

представляют, соответственно, операторы дифференцирования и интегрирования в пространстве \mathcal{H}_N с учетом начального значения.

Иными словами, применение оператора I_p к вектору $y(x) \in \mathcal{H}_N$ дает вектор $z(x) = \langle z(x_1), \dots, z(x_N) \rangle^t$, где

$$z(x) = \int_0^x y(t) dt.$$

Поэтому итерации Пикара в пространстве \mathcal{H}_N имеют вид

$$y_{n+1} = I_p \cdot \langle y_n^3(x_k) + x_k \rangle_{k=1, \dots, N}^t, \quad n \in \mathbb{N}_0,$$

где $y_0 = 0$ – это начальное приближение решения.

В данном случае (т.е. $y(0) = 0$) 35 итераций Пикара дают решение с машинной точностью, т.е. норма вектора $y_n - y_{n-1}$ близка к машинному нулю. Однако это лишь косвенная оценка достигнутой точности. Преобразование Фурье $F.y$ дает коэффициенты разложения функции $y(x)$ по полиномам Чебышева, по которым видно (как на фиг. 1, 2), что достигнута точность $y(1) \approx 0.5190566558$ порядка 10^{-63} .

Итерации Пикара имеют линейную скорость сходимости, но весьма мало-затратны в вычислительном плане и обладают двумя важными преимуществами.

Во-первых, они всегда дают правильный результат, если он есть. Это утверждение – всего лишь переформулировка стандартной теоремы о существовании и единственности решения ОДУ.

Второе преимущество состоит в том, что если нужное решение не существует, то итерации Пикара просто не будут сходиться и укажут тем самым на этот факт.

Например, решение задачи Коши $y(0) = a$ для уравнения (33) имеет вид

$$y_{n+1} = \langle a \rangle_{k=1, \dots, N}^t + I_p \cdot \langle y_n^3(x_k) + x_k \rangle_{k=1, \dots, N}^t, \quad n \in \mathbb{N}_0,$$

т.е. операторы дифференцирования и интегрирования в \mathcal{H}_N всегда строятся для однородных краевых (и иных) условий. Неоднородности учитываются в самом уравнении или в его аппроксимации.

В данном случае легко проверить, что решение задачи Коши $y(0) = 1/2$ для уравнения (33) существует на интервале $[0, 1]$, но требует большего количества итераций, а решение задачи Коши $y(0) = 1$ не продолжается до $x = 1$.

Метод Ньютона обладает, как известно, квадратичной сходимостью. Однако он весьма чувствителен к выбору начального приближения. Это всегда является отдельной задачей, требующей индивидуального подхода.

В данном случае начальное приближение можно выбрать как отрезок ряда Тейлора решения задачи Коши (33), $y_0(x) = x^2/2$.

Уравнение в вариациях уравнения (33) имеет вид

$$v'(x) = 3y^2(x)v(x), \quad (34)$$

поэтому итерации метода Ньютона здесь имеют вид $y_{n+1} = y_n - \Delta_n$, $n \in \mathbb{N}_0$,

$$\Delta_n = (D_p - 3 \text{Diag} [y_n^2(x_1), \dots, y_n^2(x_N)])^{-1} \cdot (D_p \cdot y_n - \langle y_n^3(x_k) + x_k \rangle_{k=1, \dots, N}^t).$$

Здесь мы обращаем внимание на тот факт, что линейный оператор в методе Ньютона вычисляется для конечномерной аппроксимации задачи, а не для исходного уравнения, т.е. уравнение в вариациях (34) играет лишь вспомогательную роль. Иными словами, оператор D_p здесь не может быть заменен на другой оператор дифференцирования.

В данном случае 6 итераций метода Ньютона дают то же, что ранее мы получили с помощью 35 итераций Пикара.

В завершение этого раздела (и статьи) приведем несколько искусственный пример, который иллюстрирует ограничения, существующие в стандартных численных методах. Рассмотрим краевую задачу для уравнения

$$(1 + y(x))y''(x) + x = 0$$

на интервале $[0, 1]$ с нестандартными краевыми условиями, $y'(0) = 2y(1)$, $2y'(1) = -y(1/2)$.

Если бы краевые условия задавались только по краям интервала, то, например, Maple (вообще говоря) генерирует численную процедуру для решения такой задачи. Иными словами, подобные нелинейные краевые задачи являются стандартными. Однако при выбранных нами краевых условиях Maple выдает ошибку типа «слишком много краевых условий».

Тем не менее метод Ньютона в нашей аппроксимации этой задачи дает нужное решение за несколько итераций.

ДОКАЗАТЕЛЬСТВО ПРЕДЛОЖЕНИЯ 3 В ПРЕПРИНТЕ N^o 77, 2022

Эта часть препринта не имеет прямого отношения к предыдущей и является дополнением препринта [2].

Теорема, которая здесь доказывается, вошла в статью [1], написанную по этому препринту, но в самом препринте [2] доказательство отсутствует.

В этом разделе мы используем полиномы Лежандра и все обозначения, введенные в [2], за исключением нумерации формул.

Возьмем произвольную рациональную функцию $r(x)$ с рациональными коэффициентами (можно и комплексно-рациональными или даже символьными). Рассмотрим задачу Коши

$$z'(x) = r(x) = \frac{p(x)}{q(x)}, \quad x \in [0, 1], \quad z(0) = 0. \quad (35)$$

Ее решением является функция

$$z(x) = \int_0^x r(t) dt,$$

если она существует. Но нас интересуют формальные символные вычисления в конечномерной аппроксимации.

Для задачи (35) в случае общего положения (см. ниже) вычислим следующие массивы для размерностей аппроксимации $N = 1, 2, \dots$

$S_1[N] = r_1$, где вектор $r = \langle r_1, \dots, r_N \rangle^t$ вычисляется по формуле

$$r = p(X).q(X)^{-1}.e = q(X)^{-1}.p(X).e,$$

где коммутирующие матрицы $p(X)$ и $q(X)$ получаются подстановкой матрицы X вместо независимой переменной x в полиномы $p(x)$ и $q(x)$. Напомним, что X^2 мы теперь не используем.

$S_2[N] = \sum_{k=1}^N z_k$, где вектор z вычисляется как и ранее. То есть берем матрицу $A = D$, заменяем последнюю (нулевую) строку на вектор i_0 . Далее заменяем последний (ненулевой) элемент вектора r на ноль, т.е. берем $r[N] = 0$, согласно начальному значению решения. Затем вычисляем вектор $z = A^{-1}.r$, который дает приближенное разложение функции $z(x)$, если она существует. Иначе это просто формальный вектор. Отметим, что $S_2[1] = 0$ в силу способа вычисления.

Этот массив можно также вычислять альтернативным способом с помощью оператора формального интегрирования, как это описано в разд. 2 в [2].

$S_3[N] = G_N(r)$, т.е. $S_3[N]$ – это значение N -точечной квадратуры Гаусса функции $r(x)$ на интервале $[0, 1]$.

Напомним, что корни x_n полиномов Лежандра $P(N, x)$ могут быть вычислены (в принципе) в виде радикалов только до $N = 9$ включительно. Это же относится к весам w_n квадратурной формулы Гаусса,

$$w_n = \frac{4x_n(1-x_n)}{N^2 P^2(N-1, x_n)}, \quad n = 1, \dots, N. \quad (36)$$

Разумеется, практически делать это бессмысленно, поэтому массив $S_3[N]$ (в отличие от предыдущих) дается в плавающей арифметике, но с любым (в принципе) числом десятичных разрядов.

Наложим очевидное ограничение на полюса функции $r(x)$, т.е. если $q(x)$ делится на полином Лежандра $P(M, x)$, то в массивах S_1, S_2, S_3 следует пропустить размерность $N = M$ (см. предложение 1 в [2]). А в случае если

$q(x)$ имеет корень $x = 1/2$, то эти массивы вычисляются только для четных N .

Заметим, что функция $r(x)$ не может иметь полюс ($\neq 1/2$) в одном из корней полинома $P(M, x)$ и не иметь полюсов во всех остальных корнях.

С учетом этих условий справедлива

Теорема. *Массивы S_k , $k = 1, 2, 3$, совпадают.*

Доказательство. Напомним, что $x = X.e$ по определению, т.е. вектор x , соответствующий функции $x \in \mathcal{P}_N$, получается применением матрицы X к вектору e . Аналогично получается вектор $p = p(X).e$, соответствующий любому полиному $p(x)$. Но эти же рассуждения применимы для общей функции от матрицы, $f(X)$, т.е. вектор $f = f(X).e$ – это коэффициенты приближенного разложения функции $f(x)$ по смещенным полиномам Лежандра при условии, что матрица $f(X)$ определена.

Функция от матрицы $f(X)$ однозначно определяется значениями функции $f(x)$ на спектре матрицы X (см. [4]), т.е. значениями этой функции в корнях полинома $P(N, x)$ (так как матрица X диагонализируема, т.е. производные $f(x)$ не нужны).

Для рациональной функции $r(x)$ не требуется вычислять ее значения на спектре матрицы X , а можно формально подставить матрицу X вместо переменной x (см. [4]). Поэтому совпадение массивов S_1 и S_2 следует из совпадения двух интегралов.

Величина $S_1[N]$ – это первая компонента вектора правой части (35), т.е. нулевой коэффициент ее аппроксимации полиномами $P(n, x)$. Иными словами, это аппроксимация интеграла $r(x)$ на отрезке $[0, 1]$. Величина $S_2[N]$ получается из аппроксимации интеграла $\int_0^x r(t) dt$ как функции на отрезке $[0, 1]$ подстановкой $x = 1$ в эту функцию, т.е. $S_1 = S_2$.

Вторая часть утверждения теоремы, т.е. $S_1 = S_3$, следует из представления матричной функции $r(X)$ интерполяционным полиномом Лагранжа–Сильвестра (см. [4]). Имеем

$$r(X) = \sum_{n=1}^N r(x_n) l_{N,n}(X), \quad l_{N,n}(x) = \prod_{k \neq n} \frac{x - x_k}{x_n - x_k},$$

где $\{x_n, n = 1, \dots, N\}$ – это спектр матрицы X (корни полинома $P(N, x)$), а $l_{N,n}(x)$ – это фундаментальные полиномы лагранжевой интерполяции.

Поскольку $l_{N,n}(x) \in \mathcal{P}_N$, то

$$l_{N,n}(X).e[1] = \int_0^1 l_{N,n}(x) dx = w_n,$$

где w_n – это веса (36) квадратурной формулы Лежандра–Гаусса (см. [10]). Таким образом, вектор правой части (35) представим в виде

$$r(X).e = \sum_{n=1}^N r(x_n) l_{N,n}(X).e,$$

поэтому $S_1 = r(X).e[1]$ – это квадратура Лежандра–Гаусса функции $r(x)$. Ч.Т.Д.

Напомним, что квадратура Лежандра–Гаусса порядка N (здесь на интервале $[0, 1]$) точна на полиномах степени $\leq 2N - 1$. В частности, равенство

$$G_N(x^n) = \frac{1}{n+1}, \quad n < 2N$$

выполняется точно.

Однако реально вычисления выполняются на компьютере в плавающей арифметике, поэтому термин «точно» здесь следует понимать как «сколь угодно точно» или с машинной точностью.

Если же рассматривать идеализированный компьютер с бесконечной разрядной сеткой, то термин «точно» опять приобретает строгий математический смысл.

Очевидным следствием доказанной теоремы является тот факт (ранее, насколько нам известно, не отмеченный в численном анализе), что любая квадратура Лежандра–Гаусса на интервале $[0, 1]$ от любой рациональной функции с рациональными коэффициентами является рациональным числом либо не определена (равна бесконечности). То есть в этом случае иррациональности всегда сокращают друг друга.

В практическом плане это означает, что для численного интегрирования любой рациональной функции методом Лежандра–Гаусса любого порядка нет необходимости вычислять корни полиномов Лежандра и веса квадратурной формулы. Достаточно выполнить ряд линейных операций с заранее известной матрицей X .

Список литературы

- [1] Варин В.П. Аппроксимация дифференциальных операторов с учетом граничных условий // Ж. вычисл. матем. и матем. физ. Т. 63. № 8. с. 1251-1271. (2023).
- [2] Варин В.П. Аппроксимация дифференциальных операторов с учетом граничных условий // Препринты ИПМ им. М.В.Келдыша. 2022. № 77.

- [3] *Wilf H.S.* Mathematics for the physical sciences. (Wiley. NewYork. 1962).
- [4] *Gantmacher F.R.* Application of the Theory of Matrices. (Chelsea Press. New-York. 1960).
- [5] *Boyd J.P., Patschek R.* The Relationships Between Chebyshev, Legendre and Jacobi Polynomials: The Generic Superiority of Chebyshev Polynomials and Three Important Exceptions // Journal of Scientific Computing. V.59. P. 1-27. (2014).
- [6] *Варин В.П.* Факториальное преобразование некоторых классических комбинаторных последовательностей // Ж. вычисл. матем. и матем. физ. Т.59. № 6. С. 1747-1770. (2018).
- [7] *Pashkovskii S.* Computational Application of Chebyshev Polynomials and Series (Nauka. Moscow. 1983) [in Russian].
- [8] *Варин В.П.* Инвариантные кривые некоторых дискретных динамических систем // Ж. вычисл. матем. и матем. физ. Т. 62. № 2. с. 199-216. (2022).
- [9] *Варин В.П.* Функциональное суммирование рядов // Ж. вычисл. матем. и матем. физ. Т.63. № 1. С. 3-17. (2023).
- [10] *Krylov V.I.* Approximate calculation of integrals. (Macmillan. New-York, London. 1962).