

ПРЕЦИЗИОННЫЕ АППРОКСИМАЦИИ ФУНКЦИЙ ФЕРМИ-ДИРАКА ЦЕЛОГО ИНДЕКСА

© 2016 г. *Н.Н. Калиткин, С.А. Колганов*

Институт прикладной математики им. М.В. Келдыша РАН, Москва
Национальный исследовательский университет «МИЭТ», Зеленоград
kalitkin@imamod.ru, mkandds2012@gmail.com

Работа поддержана грантом РФФИ 14-01-00161.

Функции Ферми-Дирака целого индекса играют важную роль в задачах электронного переноса в плотной среде. Для их быстрого вычисления построены аппроксимации, использующие отношение многочленов. Построен простой алгоритм нахождения коэффициентов этих аппроксимаций, основанный на интерполяции со специальной линейно-тригонометрической сеткой узлов. Показано, что такая сетка обеспечивает результаты, близкие к оптимальным. Для функции индексов 1, 2, 3 найдены коэффициенты таких интерполяций, дающие относительную погрешность $2 \cdot 10^{-16}$ при 9 свободных параметрах.

Ключевые слова: функции Ферми-Дирака, прецизионные аппроксимации, рациональная аппроксимация, линейно-тригонометрическая сетка.

PRECISION APPROXIMATIONS FOR FERMI-DIRAC FUNCTIONS OF INTEGER INDEX

N.N. Kalitkin, S.A. Kolganov

Keldysh Institute of Applied Mathematics of Rus. Acad. Sci., Moscow
National Research University of Electronic Technology, Zelenograd
kalitkin@imamod.ru, mkandds2012@gmail.com

Fermi-Dirac functions of integer index are widely used in problems of electronic transport in dense substances. Polynomial approximations are constructed for its quick computation. A simple algorithm is built for finding the coefficients of these approximations based on interpolation with a special linear-trigonometric grid nodes. Represented that this grid provides results close to optimal. Such coefficients are founded for functions of index 1, 2, 3, which provide ratio error $2 \cdot 10^{-16}$ with 9 free parametrs.

Key words: Fermi-Dirac functions, precision approximations, rational approximation, linear-trigonometric grid.

1. Функции Ферми-Дирака

1.1. Прикладные задачи. Функции Ферми-Дирака возникают в задачах квантовой механики при описании свойств вещества, обусловленных поведением электронов (или

других фермионов). При достаточно высоких плотностях или низких температурах функция распределения электронов имеет вид $\{1 + \exp[(0.5p^2 - \mu)/T]\}^{-1}$, где p есть импульс электрона, μ – химический потенциал, а T – температура (предполагается атомная система единиц, где масса электрона равна 1). При решении квантово-механических задач используются различные моменты фермиевского распределения; они равны сверткам различных степеней импульса p с этим распределением. В таких свертках принято выбирать величину $t = 0.5p^2/T$ в качестве переменной интегрирования. Тогда эти моменты приобретают следующий вид:

$$I_k(x) = \int_0^\infty \frac{t^k dt}{1 + \exp(t - x)}, \quad x \in (-\infty; +\infty). \quad (1)$$

Здесь величина $x = \mu/T$. Индекс k принимает целые значения для нечетных степеней p и полуцелые значения – для четных. В физических задачах нужны только целые и полуцелые индексы, хотя в математической теории этих функций рассматриваются произвольные k .

Укажем физические величины, соответствующие различным индексам. Плотности электронов соответствует $k = 1/2$, кинетической энергии – $k = 3/2$, электронной проводимости – $k = 1$, электронной теплопроводности – $k = 2$, электронной вязкости – $k = 3$. В ряде приложений возникают меньшие индексы (например, $k = -1/2$ и даже $k = -3/2$); но более высокие индексы до сих пор не требовались.

Аппроксимациям функций Ферми-Дирака полуцелого индекса посвящена обширная литература. Однако для функций целого индекса удовлетворительных аппроксимаций до сих пор не было. Данная работа посвящена исправлению этого недостатка.

1.2. Важнейшие свойства. Теоретические свойства функций Ферми-Дирака детально исследованы в литературе [1-7]. Приведем те свойства, которые потребуются для данной работы. Для произвольных индексов k справедливо соотношение

$$I'_k(x) = k \cdot I_{k-1}(x). \quad (2)$$

Для **целых** индексов k функции положительных и отрицательных аргументов связаны простым соотношением

$$I_k(x) = (-1)^k I_k(-x) + P_{k+1}(x), \quad x \geq 0, \quad (3)$$

где $P_{k+1}(x)$ есть многочлен степени $k+1$. Для любого k при отрицательных значениях аргумента справедливо разложение

$$I_k(x) = \Gamma(k+1) \sum_{n=1}^{\infty} (-1)^{n-1} \frac{e^{nx}}{n^{k+1}}; \quad (4)$$

этот ряд абсолютно (хотя и неравномерно) сходится при $x < 0$. При $x \rightarrow +\infty$ существует асимптотически сходящийся ряд, главный член которого есть $I_k(x) \rightarrow x^{(k+1)}/(k+1)$, а следующий член имеет относительный порядок малости $O(x^{-2})$.

Есть единственный индекс $k = 0$, когда интеграл (1) берется в элементарных функциях:

$$I_0(x) = \ln(1 + e^x). \quad (5)$$

Из (5) непосредственно следует соотношение

$$I_0(x) = x + I_0(-x), \quad x \geq 0; \quad (6)$$

оно является частным случаем соотношения (3). Отсюда видно, что $P_1(x) = x$.

1.3. Целый индекс. Далее увидим, что вычислять и аппроксимировать функции (1) при отрицательном аргументе x намного проще, чем при положительном. Поэтому для функций целого индекса достаточно строить аппроксимации лишь при $x \leq 0$, а при $x \geq 0$ следует пользоваться соотношением (3). Подставляя (3) в (2), получим следующие соотношения между многочленами:

$$P'_{k+1}(x) = k \cdot P_k(x). \quad (7)$$

Поскольку известно $P_1(x) = x$, многочлены высших степеней можно получить последовательным интегрированием (7). При этом возникают неопределенные константы интегрирования. Их определяем из условия совпадения правой и левой частей (3) при $x = 0$. Это дает следующие выражения для требующихся нам функций:

$$I_1(x) = \frac{x^2}{2} + 2I_1(0) - I_1(-x), \quad x \geq 0; \quad (8)$$

$$I_2(x) = \frac{x^3}{3} + 4I_1(0)x + I_2(-x), \quad x \geq 0; \quad (9)$$

$$I_3(x) = x^4 / 4 + 6I_1(0)x^2 + 2I_3(0) - I_3(-x), \quad x \geq 0. \quad (10)$$

Видно, что при $x \rightarrow +\infty$ главный член формул (8)-(10) есть $x^{(k+1)} / (k+1)$, что соответствует упомянутой выше асимптотике.

2. Прямое вычисление функций Ферми-Дирака

2.1. Точность. В настоящее время наиболее массовыми являются вычисления с 64-битовыми числами. В расчетах с плавающей точкой при этом относительная погрешность единичного округления составляет $\sim 10^{-16}$. Поэтому наша конечная цель – построение таких аппроксимаций, при которых относительная ошибка вычисления функции Ферми-Дирака была бы возможно ближе к 10^{-16} .

Очевидно, для построения таких аппроксимаций и проверки получаемой точности надо использовать значения функций Ферми-Дирака в контрольных точках, вычисляемые с несколько лучшей точностью. У нас не было возможности проводить расчеты со 128-битовыми числами. Мы могли использовать только систему MATLAB с отдельными

вставками на языке C++. При достаточной аккуратности это позволяло получать до 16 достоверных десятичных знаков. Опишем процедуру вычисления базовых значений функций (1) при $x \leq 0$ с указанием важных деталей, обеспечивающих указанную точность.

2.2. Схема Горнера. Ряд (4) быстро сходится при $x \ll -1$. Его сходимость быстро ухудшается при возрастании x . Поскольку ряд абсолютно сходящийся и знакопеременный, то ошибка не превышает первого отброшенного члена. Тем самым, для получения относительной точности ε число членов суммы N должно удовлетворять условию

$$N|x| > -\ln \left[\varepsilon (N+1)^{k+1} \right], \quad \varepsilon = 10^{-17}, \quad x < 0. \quad (11)$$

Видно, что при $x \rightarrow -0$ требуемое число членов ряда (4) быстро возрастает. Ряд (4) мы использовали при $x \leq -0.1$. При значении $x = -0.1$ требовалось $N = 260$ для $k = 1$, $N = 214$ для $k = 2$, $N = 165$ для $k = 3$.

Визуально контролировалась сходимость ряда при увеличении числа членов. Поскольку ряд знакопеременный, сходимость была двусторонней. Вблизи расчетного N усатанавливалось 17 десятичных знаков, и при дальнейшем увеличении N эти значащие цифры не менялись.

Обратим внимание на важную деталь. Требуется точность на пределе ошибок округления. При прямом суммировании ряда (4) $n = 1, 2, \dots$ последние малые члены становятся меньше ошибок единичного округления и фактически не прибавляются; такие ряды следует суммировать в обратном порядке от малых членов к большим.

На практике вычисления от конечных членов к начальным легко реализуются по схеме Горнера:

$$I_k(x) = \Gamma(k+1)e^x(1/1^{(k+1)} - e^x(1/2^{(k+1)} - e^x(1/3^{(k+1)} - e^x(\dots - e^x 1/N^{(k+1)}))))). \quad (12)$$

Схема (12) требует предварительного определения N из (11).

2.3. Квадратуры. При $x > -0.1$ количество членов схемы (12) становится неприемлемо большим. Поэтому на отрезке $-0.1 < x < 0$ функции (1) выгодно вычислять квадратурами. Для этого интеграл (1) был несколько преобразован, а бесконечный промежуток интегрирования заменен конечным:

$$I_k(x) \approx e^x \int_0^T \frac{t^k dt}{\exp(x) + \exp(t)}. \quad (13)$$

Для требуемой точности на отрезке $-0.1 < x < 0$ следует брать $T = 50$ для $k = 0$, $T = 60$ для $k = 1$, $T = 75$ для $k = 2$, $T = 100$ для $k = 3$.

Интеграл в (13) вычислялся по следующей квадратурной формуле. Отрезок $0 \leq t \leq T$ разбивался на N равных интервалов. На каждом интервале использовалась квадратурная формула Гаусса, основанная на нулях полиномов Лежандра 5-й степени. Это означает, что погрешность такой сеточно-гауссовой квадратуры есть $O(N^{-10})$. Начальное дробление бралось небольшим: $N = 16$. Затем N последовательно удваивалось, и

к полученным значениям применялся метод Ричардсона (см. [8, 9]). Требуемая точность $\varepsilon = 10^{-17}$ обычно достигалась уже при $N = 128$.

Здесь также есть существенная деталь. Проводить суммирование интервалов надо не с левого конца (малых t), а с правого ($t \approx T$), так как именно правые члены являются наименьшими. Если соблюдать эту предосторожность, то погрешность округлений не превышает $2 \cdot 10^{-16}$ (до уровня 10^{-17} ее довести не удастся). Если же суммировать интервалы с левого конца, то погрешность округлений возрастает до 10^{-15} и более.

Был проведен дополнительный контроль точности. Для $k = 0$ известно точное выражение (5); квадратурная формула воспроизводила эти значения с 17 десятичными знаками. Для $k = 1, 2, 3$ расчет по квадратурам сравнивался со схемой (12) при $x = -0.1$; отличие не превышало $2 \cdot 10^{-16}$.

Выводы: Из сказанного выше следует, что указанными двумя способами мы вычисляем значения $I_k(x)$ при $x < 0$ с относительной точностью не хуже $2 \cdot 10^{-16}$. Поскольку $I_k(x)$ монотонно возрастают при $k \geq 0$, их вычисления по формулам (8)-(10) для $x > 0$ обеспечивают более высокую относительную точность (тем лучшую, чем больше x).

3. Аппроксимации функции целого индекса

3.1. Вид аппроксимации. Используем следующие соображения. При $x \rightarrow -\infty$ главный член асимптотики есть $I_k(x) \approx \Gamma(k+1) \cdot e^x$. При $x \rightarrow +\infty$ асимптотика имеет не экспоненциальное, а степенное поведение: $I_k(x) \approx x^{k+1} / (k+1)$. Поэтому переменная x не слишком удобна в качестве аргумента аппроксимации.

Попробуем взять в качестве аргумента $y \equiv I_0(x) = \ln(1 + e^x)$. Асимптотики этой переменной суть $y \rightarrow e^x$ при $x \rightarrow -\infty$ и $y \rightarrow x$ при $x \rightarrow +\infty$. Такой аргумент лучше соответствует качественному поведению $I_k(x)$ во всем диапазоне $-\infty < x < +\infty$.

Аппроксимация многочленом редко бывает удачной, хотя она широко используется в литературе. Обычно лучшие результаты дает рациональная аппроксимация, то есть приближение отношением многочленов. Такая аппроксимация может передать разные асимптотики функций. В данном случае мы выбрали следующую аппроксимацию:

$$I_k(x) \approx \Gamma(k+1) y \left(\frac{\sum_{n=0}^{N+1} a_n y^n}{\sum_{m=0}^N b_m y^m} \right)^k, \quad a_0 = 1, \quad b_0 = 1, \quad x \leq 0. \quad (14)$$

При $x \rightarrow -\infty$ аппроксимация (14) точно передает первый член ряда (4). Если провести разложение (14) по степеням e^x при $x \rightarrow -\infty$, то оно качественно будет подобно ряду (4). При $x \rightarrow +\infty$ главный член разложения (14) будет $I_k(x) \sim x^{k+1}$, хотя коэффициент при нем не будет совпадать с точным. Однако даже такая передача правой асимптоты улучшает точность аппроксимации.

3.2. Нахождение коэффициентов. Алгоритмы вычисления коэффициентов рациональных аппроксимаций обычно достаточно сложны, если добиваться минимизации некоторой нормы погрешности. В данном случае мы ограничились несложным эвристическим алгоритмом, дающим хорошие результаты. Преобразуем (14) к следующей форме:

$$w = \left[\frac{I_k(x)}{\Gamma(k+1)y} \right]^{1/k}, \quad w \approx \left(\frac{\sum_{n=0}^{N+1} a_n y^n}{\sum_{m=0}^N b_m y^m} \right). \quad (15)$$

Аппроксимация (15) содержит $2N+1$ свободный коэффициент: a_n , $1 \leq n \leq N+1$, и b_m , $1 \leq m \leq N$. Мы ищем аппроксимацию на интервале $-\infty < x \leq 0$, то есть $0 < y \leq y_{\max} = \ln 2$. Выберем $2N+1$ базовую точку $y_j: 0 < y_1 < y_2 < \dots < y_{2N+1} = y_{\max}$. По величинам y_j вычислим соответствующие значения x_j , $I_k(x_j)$ и w_j . Потребуем, чтобы в точках y_j приближенное равенство (15) становилось точным, то есть поставим задачу интерполяции по выбранным точкам. Тогда для нахождения свободных коэффициентов получается система линейных уравнений; мы её не приводим в виду очевидности.

Заметим, что можно формально взять $j = 0$ и положить $y_0 = 0$. В этой точке приближенное равенство (15) автоматически становится точным, поскольку аппроксимация (14) точно передает главный член левой асимптотики.

3.3. Точки интерполяции. Положение выбранных точек сильно влияет на качество полученной интерполяции. При неудачном положении этих точек могут возникать отрицательные коэффициенты a_n, b_m ; это опасно, особенно если знаменатель или числитель обращаются в нуль внутри требуемого диапазона y . В последних случаях полученная аппроксимация совершенно неприемлема.

Мы опробовали на практике некоторые способы выбора точек интерполяции. Проиллюстрируем их на примере $k = 2$ и $N = 3$ (это означает 7 свободных коэффициентов); результаты для других k и N были аналогичными.

Простейшим способом было линейное расположение точек

$$y_j = \frac{j y_{\max}}{2N+1}, \quad 0 \leq j \leq 2N+1; \quad (16)$$

напомним, что формально мы можем вводить y_0 , хотя в расчетах оно не используется. Профиль относительной погрешности d для этого случая приведен на рис.1. Видно, что погрешность обращается в нуль во всех базовых точках интерполяции, а между ними имеет вид полуволн. Амплитуды этих полуволн невелики в средних интервалах, а в крайних они в несколько раз больше. Это указывает, что в середине отрезка следовало бы увеличить расстояние между базовыми узлами, а на краях отрезка уменьшить.

Хорошо разработана теория аппроксимаций многочленами, наилучшая в норме C . В ней положение узлов интерполяции точно не вычисляется. Однако оно близко к распределению, описываемому тригонометрической функцией:

$$y_j = \frac{1}{2} y_{\max} \left(1 - \cos \left(\frac{\pi j}{2N+1} \right) \right), \quad 0 \leq j \leq 2N+1. \quad (17)$$

Профиль погрешности d для выбора узлов (17) так же показан на рис.1. Между узлами интерполяции он имеет вид полуволн, амплитуды которых велики в середине отрезка и малы по краям. Поэтому для распределения (17) надо сблизить узлы интерполяции в середине отрезка и раздвинуть вблизи границ отрезка.

Заметим, что это не противоречит теоретическим результатам для чебышевских приближений: они относятся к аппроксимации многочленами, а мы используем аппроксимацию рациональными функциями.

Представляется естественным построить распределение узлов интерполяции, промежуточное между линейным и тригонометрическим. Такая задача уже возникала в сверхбыстром итерационном методе решения систем эллиптических уравнений на прямоугольной сетке. Воспользуемся предложенным в [10] линейно-тригонометрическим распределением:

$$y_j = \frac{y_{\max}}{2} \left[\frac{2\alpha j}{2N+1} + (1-\alpha) \left(1 - \cos \left(\frac{\pi j}{2N+1} \right) \right) \right], \quad \alpha = \frac{\pi}{2+\pi}, \quad 0 \leq j \leq 2N+1. \quad (18)$$

Узлы (18) были построены для функции, которая фактически является отношением многочленов одинаковой степени. Поскольку у нас степени многочленов в числителе и знаменателе отличаются всего на 1, можно ожидать их хорошей применимости в нашем случае. Результаты расчета также показаны на рис.1. Видно, что теперь экстремумы в центральных и краевых интервалах почти одинаковы: отличие составляет ~ 15%. Поэтому эвристическое распределение (18) можно считать практически неулучшаемым и можно пользоваться им для любых N и k .

Видно также, что погрешность для линейно-тригонометрического распределения меньше в 7.5 раз, чем для линейного, и в 4 раза по сравнению с тригонометрическим. Это достаточно существенный выигрыш в точности.

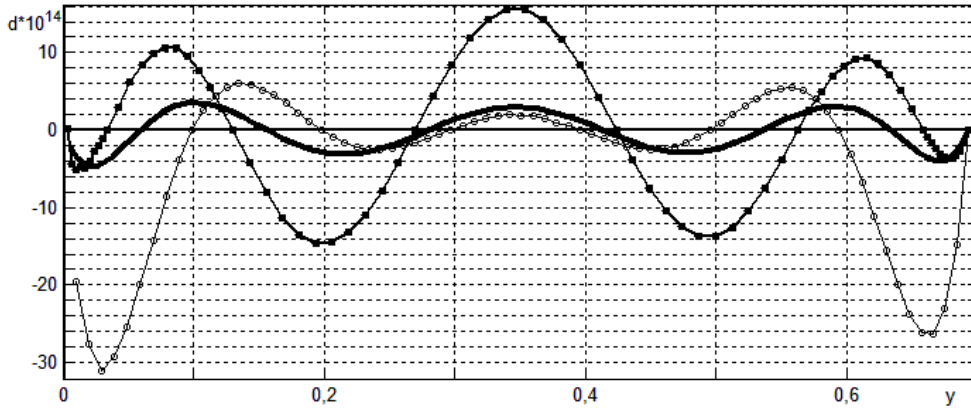


Рис.1. Относительная погрешность d для $k=2$ и $N=3$. Узлы интерполяции: жирная – линейно-тригонометрические, \circ – линейные, \bullet – тригонометрические.

3.4. Влияние числа параметров. Подробней исследуем погрешность аппроксимации для линейно-тригонометрических узлов. При $N \leq 3$ профили погрешности для всех k и N имеют тот же качественный вид, как и жирная линия на рис.1. Погрешность обра-

щается в нуль в узлах интерполяции, между ними имеет вид полуволн (число которых равно $2N+1$), а экстремумы всех полуволн примерно одинаковы. Это показывает, что ошибки округления не влияют на полученные результаты.

Картина меняется при $N = 4$. Профиль относительной погрешности становится хаотическим с амплитудой $\sim 2 \cdot 10^{-16}$ (см. рис.2). Это показывает, что расчет вышел на ошибки округления и дальнейшее увеличение числа параметров бессмысленно.

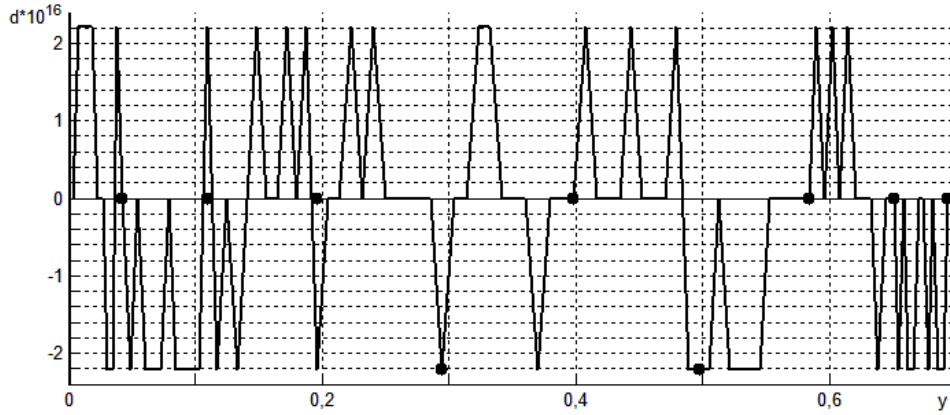


Рис.2. Профиль погрешности для $k = 2$ и $N = 4$.

Погрешность аппроксимации можно характеризовать нормой C : $d_C = \max|d(y)|$. В табл.1 приведены значения этой величины (точнее, $\lg(d_C)$ для различных k и N). Видно, что максимальная погрешность слабо зависит от k , но быстро убывает с увеличением N . Для $N = 1$ аппроксимация дает 6 верных десятичных знаков, для $N=2$ это 10 знаков, для $N=3$ это 14 знаков; для $N = 4$ следовало бы ожидать 18 знаков, но ошибки округления позволяют выйти всего лишь на ~ 16 знаков.

Таблица 1. Зависимость погрешности $\lg(d_C)$ для линейно-тригонометрических узлов от числа параметров $2N+1$ при различных k .

$k \backslash 2N+1$	3	5	7	9
1	-6.17	-9.88	-13.58	-15.65
2	-5.95	-9.64	-13.36	-15.65
3	-5.84	-9.49	-13.05	-15.65

3.5. Глобальные аппроксимации. Можно построить такие аппроксимирующие формулы, которые будут лучше передавать асимптотическое поведение при $x \rightarrow \pm\infty$. Тем самым, единую аппроксимацию сразу можно будет использовать во всем диапазоне $-\infty < x < +\infty$. Разумеется, для высокой точности такие формулы потребуют существенно большего числа свободных параметров, а само нахождение параметров окажется более сложным. Поэтому приведем только две простейшие формулы.

Первая формула правильно передает главные члены асимптотик при $x \rightarrow \pm\infty$:

$$I_k(x) \approx \Gamma(k+1)y(1+a_1y)^k, \quad a_1 = [\Gamma(k+2)]^{-1/k}. \quad (19)$$

Вторая формула передает ещё по одному члену асимптотических разложений при $x \rightarrow \pm\infty$:

$$I_k(x) \approx \Gamma(k+1)y \left[\frac{1+a_1y+a_2y^2}{1+b_1y} \right]^k, \quad 0 < y < +\infty \quad (-\infty < x < +\infty);$$

$$a_1 = [\Gamma(k+2)]^{-1/k}, \quad b_1 = a_1 - \frac{1}{2k}(1-2^{-k}), \quad a_2 = a_1b_1. \quad (20)$$

Формулы (19), (20) можно использовать для любых $k>0$, необязательно целых или полу-целых (за исключением $k=0$). При этом все коэффициенты a_1, a_2, b_1 оказываются положительными.

4. Рекомендуемые аппроксимации

Окончательно мы рекомендуем для функции Ферми-Дирака целого индекса k аппроксимацию вида (14). Трудоемкость расчета по этой формуле почти не зависит от N . Поэтому приведем два набора коэффициентов. Первый для $N=3$ дан в табл.2, где коэффициенты a_n, b_m представлены с 15 знаками после точки; эти коэффициенты обеспечивают расчет с относительной погрешностью $\sim 10^{-14}$. Второй набор приведен в табл.3 и обеспечивает погрешность $\sim 10^{-16}$; для обеспечения такой точности коэффициенты приведены с 16 значащими цифрами. Поскольку Matlab разрешает выдачу только 15 знаков после точки, значения коэффициентов умножены на 10.

Таблица 2. Коэффициенты a_n, b_m для $N=3$.

$k \backslash a_n, b_m$	1	2	3
a_1	0.266352676322699	0.206100260111457	0.139953685646763
a_2	0.049001409501337	0.043157971511600	0.035516724691922
a_3	0.004902384670602	0.004214705779461	0.003133713899982
a_4	0.000239283842563	0.000210745245200	0.000163991842901
b_1	0.016352676325596	0.018600260114022	-0.005879647683364
b_2	0.017135462543372	0.017896695800770	0.018804085754461
b_3	0.000164279201975	0.000196375385473	0.000046330229928

Напомним, что $a_0=b_0=1$. С учетом этого коэффициенты a_n, b_m быстро убывают с увеличением индекса. Такое поведение коэффициентов разумно. Это показывает, что вид аппроксимации (14) выбран удачно.

Напомним, что табл.2,3 следует использовать при $x \leq 0, y \leq \ln 2$. Для аргументов $x > 0$ следует пользоваться формулами (8)-(10), в которые входят величины $I_k(0)$. Последние величины следует брать с точностью до ошибок округления; вычислять их по 15-значным коэффициентам нежелательно. Эти значения приведены в табл.3.

Таблица 3. Коэффициенты a_n, b_m для $N = 4$.

$k \backslash a_n, b_m$	1	2	3
$10a_1$	3.126028287472988	2.588025680820918	1.751249480400745
$10a_2$	0.673008212829461	0.601284498924688	0.484611862591945
$10a_3$	0.087798043423074	0.077052021557577	0.054886614994638
$10a_4$	0.007222414330882	0.006416284842287	0.004875355489602
$10a_5$	0.000295873218273	0.000259595076916	0.000201815238332
$10b_1$	0.626028287472659	0.713025680820707	0.292916147067307
$10b_2$	0.238723363198067	0.249854915262277	0.266194049997825
$10b_3$	0.010727527758408	0.012101958452386	0.006435803052724
$10b_4$	0.000687107172921	0.000728669232953	0.000833646424907
$I_k(0)$	0.82246703342411465	1.8030853547393952	5.6821969769834864

Заключение. 1. Для функций Ферми-Дирака целых индексов $k=1,2,3$ построены несложные аппроксимации, погрешность которых равна ошибке единичного округления.

2. Предложен эвристический метод построения рациональных аппроксимаций с помощью интерполяции по узлам линейно-тригонометрической сетки. Он дает почти не улучшаемую погрешность и поэтому имеет самостоятельную математическую ценность.

СПИСОК ЛИТЕРАТУРЫ

1. E.C. Stoner, J. McDougall. The computation of Fermi-Dirac functions // Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences, 1938, 237(773), p.67-104.
2. Jr.H.C. Thacher, W.J. Cody. Rational Chebyshev approximations for Fermi-Dirac integrals of orders $-1/2$, $1/2$ and $3/2$ // Mathematics of Computation, 1967, p.30-40.
3. M. Lundstrom, R. Kim. Notes on Fermi-Dirac integrals. Arxiv preprint arXiv:0811.0116, 2008.
4. Н.Н. Калиткин. О вычислении функций Ферми-Дирака // Журнал вычислительной математики и математической физики, 1968, 8(1), с.173-175.
N.N. Kalitkin. About computation of functions the Fermi-Dirac // Magazine of computational mathematics and mathematical physics, 1968, 8(1), p.173-175.
5. L.D. Cloutman. Numerical evaluation of the Fermi-Dirac integrals // The Astrophysical Journal Supplement Series, 1989, 71, 677p.
6. M. Goano. Algorithm 745: computation of the complete and incomplete Fermi-Dirac integral // ACM Transactions on Mathematical Software (TOMS), 1995, 21(3), p.221-232.
7. A.J. MacLeod. Algorithm 779: Fermi-Dirac functions of order $-1/2$, $1/2$, $3/2$, $5/2$ // ACM Transactions on Mathematical Software (TOMS), 1998, 24(1), p.1-12.
8. Н.Н. Калиткин. Численные методы. – М.: Физматлит, 1978.
N.N. Kalitkin. Chislennyye metody. – M.: 1978.
9. Н.Н. Калиткин, Е.А. Альшина. Численные методы. Кн.1. Численный анализ. – М.: Академия, 2013.
N.N. Kalitkin, E.A. Alshina. Chislennyye metody. Кн.1. Chislennyi analis. M.: Akademiia, 2013.
10. Н.Н. Калиткин, А.А. Белов. Аналог метода Ричардсона для логарифмически сходящегося счета на установление // ДАН, 2013, т.452, № 3, с.261-265.
N.N. Kalitkin, A.A. Belov. Analogue of the Richardson method for logarithmically converging time marching // Doklady Mathematics, 2013, v.88, №2, p.596-600.

Поступила в редакцию 21.04.2015