

Н.П. Романов

Исследование методов и погрешностей вычисления функций Риккати–Бесселя

НПО «ТАЙФУН», г. Обнинск Калужской обл.

Поступила в редакцию 20.04.2007 г.

Аналитическими и численными методами исследуются зависимости функций Риккати–Бесселя (РБ) 1, 2, 3 и 4-го родов действительного x и комплексного $z = x + iy$ аргумента от их порядка ℓ . Выявляются области увеличения модуля функций РБ с ростом ℓ , устойчивые к погрешностям прямой, и области уменьшения модуля, устойчивые к погрешностям обратной рекурсии. Для функции РБ1 область устойчивости прямой рекурсии при $x \gg 1$ определяется соотношением $0 \leq \ell \leq \ell_{\max} = x - 0,5 - 0,80861x^{1/3} - 0,1635x^{-1/3}$ при условии $|y| \leq 0,4 \lg x + 0,5$. В этой области относительная погрешность прямой рекурсии возрастает с ростом ℓ пропорционально $\ell^{1/2}$. В области $\ell > \ell_{\max}$ прямая рекурсия приводит к формированию вместо РБ1 функции, равной сумме функций РБ1 и РБ2. Относительная погрешность обратной рекурсии для функции РБ1 во всем диапазоне ℓ практически не зависит от ℓ , возрастает с ростом модуля z по закону $|z|^{1/2}$ и сравнивается при указанных выше ограничениях на y с погрешностью прямой рекурсии при $\ell = \ell_{\max}$. Для инициализации начальных значений РБ1 при обратной рекурсии предложены упрощенная процедура вычисления отношения этих функций соседних порядков с использованием непрерывной дроби, дополнительное вычисление функций РБ2 при $y = 0$ или РБ3 при $y > 0$ методом прямой рекурсии и использование вронскиана соответствующих функций. Приводится программа на языке ФОРТРАН для вычисления указанного выше отношения функций РБ1. В устойчивых областях основная погрешность вычисления функций при задании аргументов в десятичной системе может определяться погрешностью их перевода в двоичную, используемую в компьютерах.

Введение

Расчет характеристик взаимодействия электромагнитного излучения со сферическими частицами является основой атмосферной оптики и сводится к суммированию рядов Ми [1–3]. Радиальные коэффициенты этих рядов выражаются через функции Риккати–Бесселя первого $\psi_\ell(z)$ и второго $\chi_\ell(z)$ родов. На основе выполненных в настоящее время исследований общей рекомендацией для вычисления функций второго рода является применение прямой, а функций первого рода — обратной рекурсий [3–9]. Однако нам неизвестны работы с аналитическим обоснованием этого положения, что всегда оставляет сомнения в корректности результатов вычислений. Так, в [3] для функций первого рода от действительного аргумента используется прямая рекурсия с использованием двойной точности, которая, по мнению авторов этой монографии, не застраховывает от возникновения больших ошибок вычислений в неконтролируемой области $\ell \gg x$. Особое предостережение они высказывают по поводу вычисления функций $\psi_\ell(z)$ в неисследованных областях $\ell - z$. В настоящей статье проводится оценка погрешностей вычисления функций Риккати–Бесселя рекурсионным методом и делаются выводы об областях устойчивости рекурсионной процедуры. На основе анализа свойств функций РБ определяются эти области и предлагается методика совместного вычисле-

ния функций РБ первого и второго рода. Для инициализации обратной рекурсии разрабатывается упрощенная схема цепных дробей для вычисления отношения функций РБ первого рода соседних порядков и использование вронскиана для получения самих функций. Проводится проверка вновь предложенных алгоритмов численными методами. В приложении приводится текст подпрограммы на языке ФОРТРАН для вычисления отношения функций РБ1.

1. Основные соотношения

В теории Ми [1–3] используются функции Риккати–Бесселя первого $\psi_\ell(z)$ (РБ1) и второго $\chi_\ell(z)$ (РБ2) рода, которые являются двумя независимыми решениями дифференциального уравнения [4]:

$$\Phi'_\ell(z) - \left(\frac{\ell(\ell+1)}{z^2} - 1 \right) \Phi_\ell(z) = 0 \quad (1)$$

и определяются рядами вида

$$\psi_\ell(z) = \frac{z^{\ell+1}}{(2\ell+1)!!} \times \left[1 - \frac{z^2/2}{1!(2\ell+3)} + \frac{(z^2/2)^2}{2!(2\ell+3)(2\ell+5)} - \dots \right]; \quad (2)$$

2. Оценка погрешностей рекурсионных процедур

$$\chi_\ell(z) = \frac{(2\ell-1)!!}{z^\ell} \times \left[1 + \frac{z^2/2}{1!(2\ell-1)} + \frac{(z^2/2)^2}{2!(2\ell-1)(2\ell-3)} + \dots \right]. \quad (3)$$

Другой парой независимых решений являются функции $\eta_\ell(z)$ (РБЗ) и $\zeta_\ell(z)$ (РБ4), определяемые соотношениями

$$\eta_\ell(z) = \psi_\ell(z) - i\chi_\ell(z); \quad (4)$$

$$\zeta_\ell(z) = \psi_\ell(z) + i\chi_\ell(z). \quad (5)$$

Рекуррентные соотношения между функциями разного порядка имеют вид

$$\Phi_{\ell+1}(z) = \frac{2\ell+1}{z}\Phi_\ell(z) - \Phi_{\ell-1}(z), \quad (6)$$

$$\Phi'_\ell(z) = \Phi_{\ell-1}(z) - \frac{\ell}{z}\Phi_\ell(z). \quad (7)$$

Вронскиан двух независимых функций $\psi_\ell(z)$ и $\chi_\ell(z)$

$$W(\psi, \chi) = [\psi'_\ell(z)\chi_\ell(z) - \psi_\ell(z)\chi'_\ell(z)] = [\psi_{\ell-1}(z)\chi_\ell(z) - \psi_\ell(z)\chi_{\ell-1}(z)] = 1. \quad (8)$$

Для дальнейшего полезны выражения для вронскианов следующих пар функций:

$$W(\psi, \eta) = [\psi_{\ell-1}(z)\eta_\ell(z) - \psi_\ell(z)\eta_{\ell-1}(z)] = -i, \quad (8a)$$

$$W(\psi, \zeta) = [\psi_{\ell-1}(z)\zeta_\ell(z) - \psi_\ell(z)\zeta_{\ell-1}(z)] = i. \quad (8b)$$

При целых ℓ функции Риккати–Бесселя выражаются через тригонометрические функции. В частности:

$$\begin{aligned} \psi_0(z) &= -\chi_{-1}(z) = \sin(z), \\ \psi_{-1}(z) &= -\chi_0(z) = \cos(z), \\ \eta_{-1}(z) &= i\eta_0(z) = \exp(iz), \\ \zeta_{-1}(z) &= -i\zeta_0(z) = \exp(-iz). \end{aligned} \quad (9)$$

Соотношения (9) позволяют с использованием (6) образовать прямую рекурсию для вычисления функций Риккати–Бесселя для целых $\ell \geq 1$, используемых в рядах Ми. Соотношения (6) используются и для построения обратной рекурсии, однако выбор начальных значений при этом более сложен. Обзор и реализация различных схем построения обратных рекурсий имеются в [5–10]. Большую роль при этом играет отношение функций двух соседних порядков $R_\ell^\Phi(z) = \Phi_{\ell-1}(z)/\Phi_\ell(z)$, прямые и обратные рекуррентные соотношения для которого непосредственно следуют из (6) и имеют вид

$$R_{\ell+1}(z) = \left[\frac{2\ell+1}{z} - R_\ell(z) \right]^{-1}, \quad (10a)$$

$$R_\ell(z) = \frac{2\ell+1}{z} - \frac{1}{R_{\ell+1}(z)}. \quad (10b)$$

Для оценок погрешностей вычислений функций Риккати–Бесселя будем рассматривать отношение функций $R_\ell(z)$, поскольку оно зависит только от одной величины соседнего порядка, а сами функции зависят от двух функций двух соседних порядков. В этом случае можно получить аналитические соотношения для погрешностей рекуррентного процесса. Введем относительную погрешность g_ℓ вычисления отношения $R_\ell(z)$. При этом расчетная величина $\bar{R}_\ell(z)$ будет представляться в виде

$$\bar{R}_\ell = R_\ell(1 + g_\ell). \quad (11)$$

Подставляя (11) в (10б) и пренебрегая погрешностью вычисления величины $(2\ell+1)/z$, получаем

$$g_{\ell+1}(1 + g_{\ell+1})^{-1} = R_\ell R_{\ell+1} g_\ell,$$

а предполагая условие $|g_{\ell+1}| \ll 1$, имеем

$$g_{\ell+1} = g_\ell R_\ell R_{\ell+1} \equiv g_\ell \frac{\Phi_{\ell-1}}{\Phi_{\ell+1}}. \quad (12)$$

Подобное соотношение было получено в [8] и использовано для обоснования необходимости обратной рекурсии для величины $R_\ell^\Psi(z)$ при комплексном z . Применяя (12) k раз и проводя необходимые сокращения, получаем

$$g_{\ell+k} = g_\ell \frac{\Phi_\ell \Phi_{\ell-1}}{\Phi_{\ell+k} \Phi_{\ell+k-1}}, \quad g_{\ell+1} \dots g_\ell \ll 1. \quad (13)$$

Соотношения (12), (13) справедливы для оценки погрешности как прямой, так и обратной рекурсии. Отметим, что эти соотношения не являются точными, так как при их получении не учитывались погрешности выполнения самих операций сложения, деления и умножения, обусловленные ограниченностью памяти, предоставляемой для записи чисел. Однако они позволяют сделать качественный вывод о том, что погрешность прямого рекуррентного процесса должна демпфироваться в области увеличения модуля функций с ростом ℓ . Можно сказать, что эта ситуация является отражением более общего положения о том, что любое разностное уравнение, примененное в прямом направлении, эффективно только при вычислении «наибольшего» решения этого уравнения [7]. Естественно, что для обратного процесса область устойчивости находится в области уменьшения модуля функции при росте ℓ . Из вышеизложенного следует, что для выбора метода вычисления функций Риккати–Бесселя необходимо знать их качественную зависимость от ℓ в различных областях комплексных значений величины z . Ввиду более простых зависимостей функций РБ от действительного аргумента сначала рассмотрим этот случай.

3. Свойства функций Рикати–Бесселя действительного аргумента

Характер зависимости функций Рикати–Бесселя от действительного аргумента x при фиксированном значении ℓ , а также от ℓ при фиксированном значении x в аналитическом и графическом виде описан в [4] и кратко изложен ниже. При фиксированном значении ℓ величина $\psi_\ell(x)$ с ростом x от нуля до бесконечности сначала возрастает от бесконечно малых величин, определяемых первым членом ряда (2), до величин больше единицы в первом максимуме, определяемом первым нулем производной $\psi'_{\ell,1}$, далее проходит через первый нуль самой функции $\psi_{\ell,1}$ и затем осциллирует около оси x , приближаясь к гармонической зависимости от x с амплитудой, стремящейся к единице. Асимптотическое разложение нулей $\psi_{\ell,s}$ для больших значений ℓ имеются в [4], а для $\psi'_{\ell,s}$ получено автором [11]. При больших значениях ℓ функция $\psi_\ell(x)$ в первом максимуме приблизительно равна [4]:

$$\psi_\ell(\psi'_{\ell,1}) \approx 0,8458 \left(\ell + \frac{1}{2} \right)^{1/6}. \quad (14)$$

Функция $\chi_\ell(x)$ с ростом x от нуля до бесконечности сначала уменьшается от бесконечно больших значений, определяемых первым членом ряда (3), до первого нуля $\chi_{\ell,1}$, проходит через первый максимум абсолютной величины, определяемый первым нулем производной $\chi'_{\ell,1}$, и затем осциллирует около оси x с амплитудой, стремящейся к единице, и фазой, смещенной относительно фазы функции $\psi_\ell(x)$ на π . Значение функции $\chi_\ell(x)$ в первом максимуме абсолютной величины при больших значениях ℓ описывается аналогичной зависимостью

$$\chi_\ell(\chi'_{\ell,1}) \approx -0,7184 \left(\ell + \frac{1}{2} \right)^{1/6}. \quad (15)$$

Модуль функций $\eta_\ell(x)$ и $\zeta_\ell(x)$ с ростом x от нуля до бесконечности монотонно уменьшается от значений, определяемых первым членом ряда (3), до единицы.

При фиксированном x функция $\psi_\ell(x)$ с ростом ℓ , которая в данном случае рассматривается как непрерывная переменная, осциллирует с возрастающей амплитудой, достигает максимальной величины в точке $\ell = \ell_{\max}^\psi = \ell_1^{\psi'}$, определяемой соотношением $\psi'_{\ell_s^\psi}(x) = 0$, и затем стремится к нулю как первый член ряда (2). Численное исследование зависимости $\ell_{\max}^\psi(x)$ на основании имеющейся в [11] зависимости $\psi'_{\ell,s}$ привело к соотношению

$$\ell_{\max}^\psi(x) = x - \frac{1}{2} - 0,808616x^{1/3} - 0,1635x^{-1/3}, \quad x \geq \psi'_{\ell,1} \approx 2,74. \quad (16)$$

Подобная формула для первого нуля самой функции имеет вид

$$\ell_1^\psi(x) = x - \frac{1}{2} - 1,855757x^{1/3} + 0,1146x^{-1/3}, \quad x \geq \psi_{\ell,1} \approx 4,49. \quad (17)$$

Для дальнейшего полезно знать также зависимость $\ell_1^\chi(x)$, определенную аналогично. Из численного анализа получаем следующую зависимость:

$$\ell_1^\chi(x) = x - \frac{1}{2} - 0,931577x^{1/3} + 0,0256x^{-1/3}, \quad x \geq \chi_{\ell,1} \approx 2,80. \quad (18)$$

Для $\ell_{\max}^\chi = \ell_1^{\chi'}$ аналогичная процедура приводит к формуле

$$\ell_{\max}^\chi(x) = x - \frac{1}{2} - 1,821090x^{1/3} + 0,03045x^{-1/3}, \quad x \geq \chi'_{\ell,1} \approx 4,48. \quad (19)$$

В соответствии с вышеизложенным вырисовывается следующая оптимальная с точки зрения достижения минимальных погрешностей схема вычисления функций $\psi_\ell(x)$ действительного аргумента. А именно: для $\ell \leq \ell_{\max}^\psi$ вычисление проводится методом прямой, а при $\ell > \ell_{\max}^\psi$ — методом обратной рекурсии. Для комплексного аргумента z вопрос о разделении областей требует отдельного исследования. А сейчас рассмотрим методику вычисления функций $\psi_\ell(x)$ для произвольных соотношений x и ℓ .

4. Методика вычисления функций РБ действительного аргумента

Начальные значения функций для обратной рекурсии получались следующим образом. С использованием непрерывных дробей упрощенным по сравнению с алгоритмом Ленца [12] способом, описание которого дается ниже, вычисляются $R_\ell^\psi(x)$, а методом прямой рекурсии — $\chi_{\ell-1}(x)$ и $\chi_\ell(x)$ для граничных значений $\ell = \ell_{\text{гр}} > \ell_{\max}^\psi$. Далее с использованием вронскиана (6) вычисляется функция $\psi_\ell(x)$ по соотношению

$$\psi_\ell(x) = \frac{1}{R_\ell^\psi(x)\chi_\ell(x) - \chi_{\ell-1}(x)}, \quad (20)$$

а $\psi_{\ell-1}(x) = \psi_\ell(x)R_\ell^\psi(x)$ — по определению. Отметим, что соотношение (20) не дает погрешностей в образовании малой величины $\psi_\ell(x)$ при $\ell = \ell_{\text{гр}}$, поскольку она образуется делением единицы на большую величину.

Численное исследование погрешностей этой схемы проводилось на компьютере IBM PC, в котором для выражения мантиссы числа с плавающей точкой отводятся 24 бита. Основным методом было сравнение вычислений функций с использованием одинарной и двойной точности. Программа вычислений была написана на языке ФОРТРАН.

На первом этапе сравнительных вычислений выяснилось, что характер погрешности вычислений кардинально зависит от того, может ли дробная часть аргумента x , заданного в десятичной системе, быть

выраженной в конечном виде в используемой в компьютерах двоичной системе. Отметим, что все целые числа переводятся в двоичную систему. В то же время дробная часть числа, равная, например, $0,1N$ ($N \neq 5, 10\dots$), представляется только в виде периодической двоичной дроби [13]. Погрешность перевода числа из одной системы счисления в другую называется погрешностью системы счисления. Максимальное значение относительной величины этой погрешности для двоичной системы

$$\max \left| \frac{\Delta x}{x} \right|_{10-2} = 2^{-q}, \quad (21)$$

где q — количество используемых для мантиссы бит. Для нашего случая в десятичном представлении $\max |\Delta x/x|_{10-2} = 6 \cdot 10^{-8}$. Как показали исследования, в области устойчивости рекуррентной процедуры погрешности вычисления всех функций РБ с использованием одинарной точности в случаях непредставимости аргументов в двоичной системе в конечном виде обусловлены в основном погрешностью системы счисления и определяются соотношением

$$\Delta \Phi = \Phi' \Delta x_{10-2}. \quad (22)$$

В случае представимости дробной части x в двоичной системе в конечном виде проявляются только погрешности действий [13], исследование которых и является целью данной работы. Здесь этот случай реализовывался заданием дробной части x в виде $1/2^n$ при целом значении n . При этом оказалось, что абсолютная погрешность вычисления основных функций $\sin x$ и $\cos x$ не имеет трендов с ростом x и при одинарной точности не превышает $3 \cdot 10^{-8}$ в исследованном диапазоне $x \leq 10^4$.

На втором этапе была проведена проверка зависимости погрешности от порядка выполнения действия в (4). При этом оказалось, что результат вычислений отличается только в случае, когда перед началом рекуррентного процесса образуется величина $1/x$ и потом на нее проводится умножение. Как показывают численные эксперименты, этот вариант занимает в 1,5 раза меньше времени из-за того, что операция деления намного длительнее остальных алгебраических операций, а в этом варианте она проводится только один раз перед началом рекуррентного цикла. Во всех остальных случаях время и результат вычисления в (4) остаются одинаковыми. При этом погрешности вычислений обычно меньше, чем в варианте с образованием величины $1/x$. Поэтому основные исследования погрешности вычислений функций Риккати–Бесселя проводились с порядком вычислений, соответствующим записи в соотношении (4).

Полученные в результате сравнительных расчетов общие закономерности погрешности вычисления функций $\psi_\ell(x)$ для $\ell \leq \ell_{\max}^v$ и $\chi_\ell(x)$ для $\ell \leq \ell_{\max}^z$ при прямой рекурсии состоят в том, что абсолютная погрешность осциллирует с ростом ℓ со слабо возрастающей амплитудой от величин, указанных выше для $\sin x$ и $\cos x$, до величин, определяемых для обеих функций одним соотношением

$$|\Delta \psi_\ell(x), \Delta \chi_\ell(x)| \approx 2 \cdot 10^{-8} x^{2/3}, \quad \ell \approx \ell_{\max}^v, \quad \ell \approx \ell_{\max}^z. \quad (23)$$

При дальнейшем увеличении ℓ абсолютное значение $\Delta \chi_\ell$ начинает резко возрастать, однако относительная погрешность остается практически постоянной, а ее среднее значение описывается соотношением

$$|\Delta \chi_\ell / \chi_\ell| \approx 2,5 \cdot 10^{-8} x^{1/2}, \quad \ell \geq \ell_1^z, \quad (24)$$

которое согласуется с соотношениями (14), (15) и (23).

В области $\ell > \ell_{\max}^v$ с ростом ℓ начинают быстро возрастать относительная и даже абсолютная погрешности вычислений функции $\psi_\ell(x)$. Однако при этом значение вронскиана (8) сохраняется равным единице. Это означает, что при $\ell > \ell_{\max}^v$ вычисляемая величина $\tilde{\psi}_\ell(x)$ является суммой

$$\tilde{\psi}_\ell(x) = \psi_\ell(x) + b \chi_\ell(x), \quad \ell > \ell_{\max}^v. \quad (25)$$

При изменении x может меняться как знак, так и величина b . Для средних величин поведение коэффициента b может быть описано следующим соотношением:

$$|b| \approx 2 \cdot 10^{-8} x^{1/2}. \quad (26)$$

Уже при небольшом превышении ℓ над ℓ_{\max}^v второй член в (25) является определяющим, т.е. вычисляемая в этом случае по прямой рекурсии величина $\tilde{\psi}_\ell(x)$ фактически пропорциональна $\chi_\ell(x)$, а полученная при этом величина R_ℓ является $R_\ell^z(x)$. Полученный, казалось бы, парадоксальный результат является подтверждением уже упомянутого выше положения, что любое разностное уравнение, примененное в прямом направлении, эффективно только при вычислении «наибольшего» решения этого уравнения [7]. Рост абсолютной (23) и относительной погрешностей [что следует из сопоставления (23) с (14), (15)] в области устойчивости можно объяснить тем, что демпфирующее соотношение (13) только частично подавляет погрешности действий и в результате погрешность вычислений пропорциональна квадратному корню из количества операций.

Абсолютная погрешность обратной рекурсии исследовалась для $\ell_{\text{гр}}$, определяемого соотношением $\ell_{\text{гр}} = x + 4x^{1/3} + 8$, которое согласно [3, 6] обеспечивает корректное суммирование рядов Ми. При этом оказалось, что относительная погрешность вычисления функции $\psi_\ell(x)$ по обратной рекурсии для фиксированного x остается практически постоянной во всем диапазоне ℓ . Ее зависимость от x приблизительно описывается следующим соотношением:

$$|\Delta \psi_\ell / \psi_\ell| \approx (3 \div 5) \cdot 10^{-8} x^{1/2}, \quad (27)$$

которое практически совпадает с таковым для функций $\psi_\ell(x)$ при $\ell = \ell_{\max}^v$ и $\chi_\ell(x)$ при $\ell \geq \ell_{\max}^z$ [см. (24)] для прямой рекурсии. Иллюстрация указанных выше положений приводится на рис. 1, на котором для $x = 1000$ (целое число, представимое в двоичной системе) и $x = 1000,1$ (не представимое

в двоичной системе число) приведены абсолютные погрешности прямой и обратной рекурсий для $\psi_\ell(x)$, которые для $\ell < \ell_{\max}^{\psi}$ представлены областями их изменения.

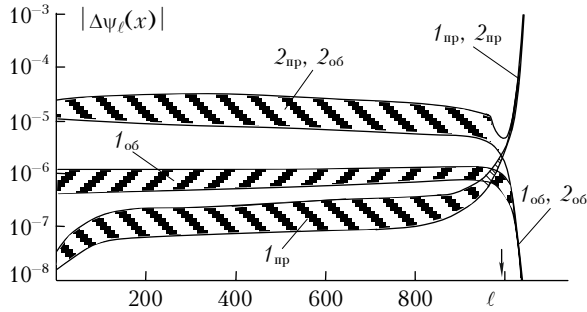


Рис. 1. Абсолютные погрешности вычисления функций РБ1 для $x = 1000$ (кривые 1) и $x = 1000,1$ (кривые 2) методом прямой (кривые 1пр, 2пр) и обратной (кривые 1об, 2об) рекурсий

Положение $\ell = \ell_{\max}^{\psi}$ указано на рис. 1 стрелкой. Из рис. 1 хорошо видно, что для представимых в двоичной системе чисел прямая рекурсия дает лучшую точность при $\ell < \ell_{\max}^{\psi}$, а обратная — при $\ell > \ell_{\max}^{\psi}$ и что погрешности обеих рекурсий сравниваются в районе $\ell = \ell_{\max}^{\psi}$. В то же время погрешность преобразования чисел практически сравнивается погрешности обеих рекурсий в области $\ell < \ell_{\max}^{\psi}$, которые для приведенного на рис. 1 случая больше чем на порядок превышают погрешности вычислений. Резкое нарастание погрешности при $\ell > \ell_{\max}^{\psi}$ обуславливается, как указывалось выше, переходом рекуррентного процесса к формированию функции, пропорциональной $\chi_\ell(x)$.

5. Свойства и погрешности методики вычисления функций РБ комплексного аргумента z

При описании свойств функций РБ комплексного аргумента ограничимся областью $0 \leq \theta = \arg z \leq \pi/2$, поскольку в других секторах функции РБ выражаются через функции этого сектора посредством операций комплексного сопряжения или изменения знака. Так, $\psi_\ell(z^*) = \psi_\ell^*(z)$, а $|\eta_\ell(z^*)| = |\zeta_\ell(z)|$. Будем использовать две эквивалентные формы записи комплексного числа $z = x + iy = r(\cos\theta + i\sin\theta)$. В последней записи r является модулем, а θ — аргументом комплексного числа.

Из выражений для функций РБ в виде рядов (2), (3) следует, что зависимость модулей этих функций от модуля аргумента при $\ell \gg r$ должна быть подобна таковому для действительного аргумента. В то же время для малых ℓ , как следует из выражений (9) для функций РБ при $\ell = 0$, модуль этих функций имеет экспоненциальную зависимость от y , а именно:

$$|\psi_0(z)| = |\chi_0(z)| = \exp(y)/2, \quad (28a)$$

$$|\eta_0(z)| = \exp(-y), \quad (28б)$$

$$|\zeta_0(z)| = \exp(y). \quad (28в)$$

Функциональная зависимость при промежуточных значениях ℓ простому анализу не поддается и может быть получена из численных расчетов. Примеры таких расчетов для $x, y = 10$ даются в работе [14], в которой исследуется область применимости прямой рекурсии при использовании одинарной, двойной, тройной и четверной точности представления чисел.

В настоящей статье для иллюстрации поведения функции РБ1 при больших значениях аргумента на рис. 2 приводятся результаты расчетов зависимости функции РБ1 от ℓ для $r = 10$ и нескольких значений аргумента.

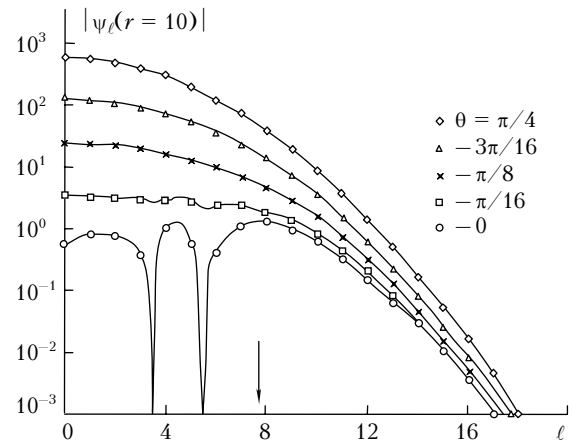


Рис. 2. Зависимость модуля функции РБ1 при значении модуля комплексной переменной $r = 10$ от ℓ при различных значениях аргумента этой переменной θ

Точками отмечены расчетные значения при целых значениях ℓ , а сплошные кривые, соединяющие эти точки, представляют величины РБ1 при промежуточных значениях ℓ . Стрелкой отмечено значение ℓ , при котором достигается максимум функции РБ1 при действительном аргументе, т.е. $\ell_{\max}^{\psi}(10 + i \cdot 0) = 7,68$. Видно резкое уменьшение модуля функции РБ1 с увеличением порядка, что согласно вышеизложенным соображениям [соотношение (13)] отвергает применимость прямой рекурсии для больших значений θ . Поведение функций РБ1 и РБ2 при больших значениях реальной части и малых величинах θ приводится на рис. 3, а функций РБ3 и РБ4 — на рис. 4.

Ввиду осциллирующего характера функций РБ1 и РБ2 на рис. 3 штриховкой указаны области изменения их абсолютных значений. Из рис. 3, 4 видно, что с ростом аргумента амплитуда осцилляций функций РБ1 и РБ2 уменьшается, а их среднее значение увеличивается. Явно проявляется отмеченный выше характер поведения функций при малых [выполнение соотношений (28)] и больших (одинаковая зависимость при фиксированном r) значениях ℓ .

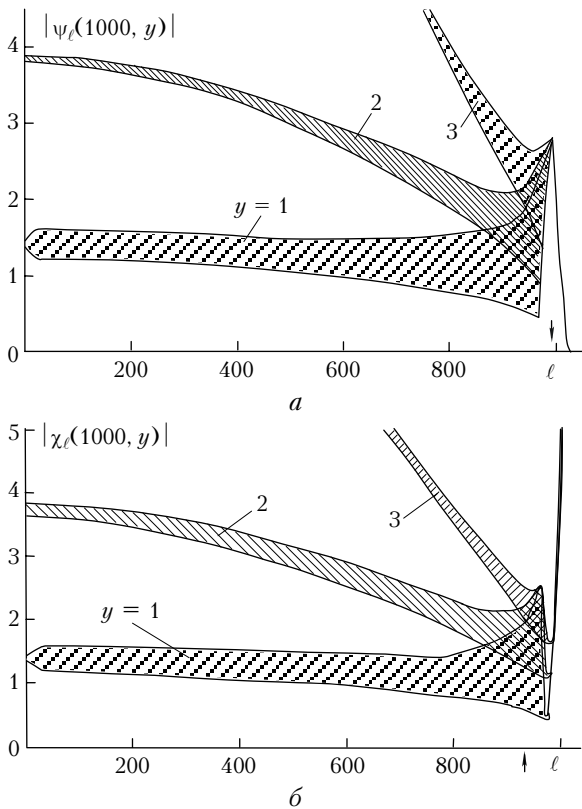


Рис. 3. Зависимость модуля функций РБ1 и РБ2 от ℓ при фиксированном значении $x = 10^3$ и различных значениях y . Стрелками отмечены значения ℓ_{\max}^{ψ} и ℓ_{\max}^{χ} для $x = r$ по соотношениям (16) и (19) соответственно

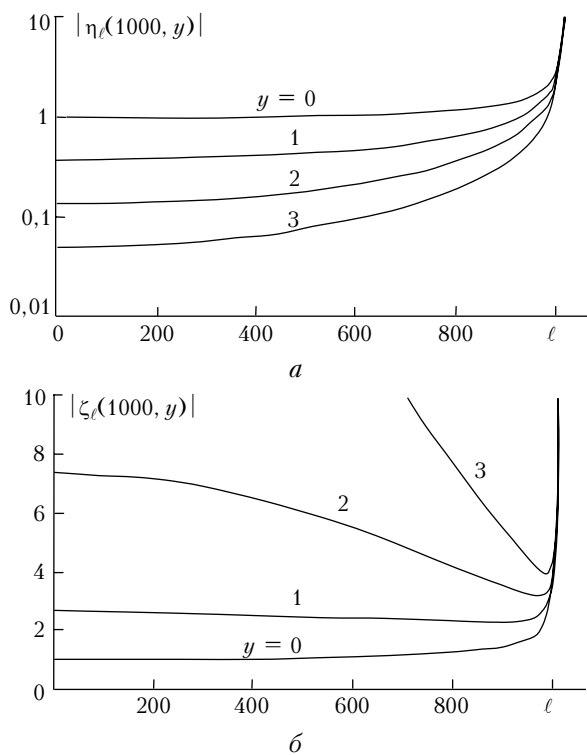


Рис. 4. Зависимость модуля функций РБ3 и РБ4 от ℓ при фиксированном значении $x = 10^3$ и различных значениях y (номера на кривых)

В частности, мало меняется значение максимума модуля в точках ℓ_{\max}^{ψ} и ℓ_{\max}^{χ} , определяемых соотношениями (16) и (19) соответственно при замене x на r . Приравнявая (28а) и (14), находим условие равенства модуля функции РБ1 в точках $\ell = 0$ и $\ell = \ell_{\max}^{\psi}$, которое определяет крайнюю точку устойчивости прямой рекурсии при $\ell \leq \ell_{\max}^{\psi}$:

$$y \leq 0,4 \lg x + 0,5. \quad (29)$$

Для функции РБ3 имеет место уменьшение, а для функции РБ4 – увеличение модуля с ростом мнимой части z . Из рис. 4, б видно, что модуль функции РБ4 при $y > 0$ имеет минимум вблизи $\ell = x$, глубина которого увеличивается с ростом y . Отсюда с учетом соотношений (13) можно сделать вывод, что прямая рекурсия для этой функции неустойчива. Модуль функции РБ3, как следует из рис. 4, а, монотонно увеличивается с ростом ℓ при всех $y \geq 0$, поэтому прямая рекурсия для функции $\eta_{\ell}(z)$ при $0 \leq \arg z \leq \pi/2$ должна быть устойчивой.

Численная проверка показала, что для этой функции поведение абсолютного значения относительной погрешности при прямой рекурсии слабо зависит от θ и практически повторяет закономерности поведения погрешностей для функции РБ2 от действительного аргумента. Поэтому для получения функций $\psi_{\ell}(z)$ для больших значений ℓ при известной величине $R_{\ell}^{\psi}(z)$ следует выбрать соотношение (8а), с использованием которого выражение для $\psi_{\ell}(z)$ примет вид

$$\psi_{\ell}(z) = \frac{-i}{R_{\ell}^{\psi}(z)\eta_{\ell}(z) - \eta_{\ell-1}(z)}. \quad (30)$$

Примеры погрешностей прямой и обратной [при использовании соотношения (30)] рекурсий приведены на рис. 5.

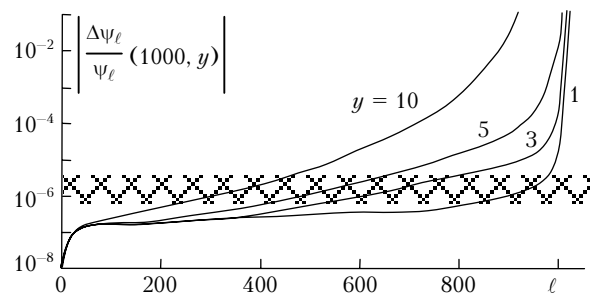


Рис. 5. Зависимость модуля относительной погрешности вычисления функции РБ1 от ℓ при $x = 1000$ и различных значениях y для прямой (сплошные кривые, номера которых соответствуют значениям y) и обратной (заштрихованная область) рекурсий

Для прямой рекурсии номера кривых соответствуют значению y . Погрешности обратной рекурсии для тех же значений y находятся в заштрихованной горизонтальной области. Из рис. 5, а также из расчетов при других величинах z и ℓ выявляются следующие закономерности поведения погрешностей прямой и обратной рекурсий при вычислении

функций РБ1 для комплексных значений аргумента. А именно: погрешность обратной рекурсии при фиксированном значении z практически постоянна для всех ℓ , слабо зависит от θ и находится в пределах

$$\left| \frac{\Delta \psi_\ell(z)}{\psi_\ell(z)} \right| = (3 \div 6) \cdot 10^{-8} r^{1/2}; \quad 0 \leq \theta \leq \frac{\pi}{2}, \quad \ell \leq \ell_{\text{гр}}. \quad (31)$$

Это соотношение фактически включает в себя и соотношение (27), полученное для действительных аргументов. При выполнении условия (29) прямая рекурсия дает меньшую погрешность при $\ell < \ell_{\text{max}}^{\text{пр}}$, которая сравнивается с таковой для обратной рекурсии при $\ell = \ell_{\text{max}}^{\text{пр}}$. При увеличении y эта граница сдвигается влево. Однако получение аналитических соотношений для этого случая в данной работе не проводится, поскольку оно уже сделано в [14] в рамках поставленной там задачи. В заключение этого раздела отметим, что все выводы об устойчивости рекурсий для функции РБ1 автоматически относятся и к их отношению, процедура получения которого обосновывается ниже.

6. Выбор и обоснование алгоритма вычисления отношения функций РБ1

Использование ряда (2) для нахождения начальных значений $\psi_{\ell-1}(z)$ и $\psi_\ell(z)$ или их отношения корректно только при условии

$$2\ell > |z|^2, \quad (32)$$

при котором его члены убывают по абсолютной величине. При меньших ℓ члены ряда вначале возрастают по абсолютной величине до величин много больше единицы, поэтому значение суммы в квадратных скобках (2) определяется разностью величин, намного превосходящих саму разность. Члены ряда Тейлора для отношения могут быть получены делением рядов для $\psi_{\ell-1}(z)$ и $\psi_\ell(z)$ и имеют вид

$$R_\ell^{\text{пр}}(z) = \frac{2\ell + 1}{z} \times \left[1 - \frac{z^2}{(2\ell + 1)(2\ell + 3)} - \frac{z^4}{(2\ell + 1)(2\ell + 3)^2(2\ell + 5)} - \frac{2z^6}{(2\ell + 1)(2\ell + 3)^3(2\ell + 5)(2\ell + 7)} - \frac{(10\ell + 27)z^8}{(2\ell + 1)(2\ell + 3)^4(2\ell + 5)^2(2\ell + 7)(2\ell + 9)} - \dots \right]. \quad (33)$$

Из рассмотрения ряда (33) видно, что сходимость его лучше, чем ряда (2). Условием корректного вычисления этого ряда является неравенство

$$2\ell > |z|. \quad (34)$$

Однако трудность его использования состоит в отсутствии аналитического выражения для последующих членов. В то же время известно [10], и это непосредственно следует из соотношения (10б), что

обратная рекурсия для R_ℓ эквивалентна бесконечной цепной дробью, которая обычно записывается в виде [15]:

$$R_\ell(z) = a_0 + \frac{b_1}{a_1 + \frac{b_2}{a_2 + \frac{b_3}{a_3 + \dots}}}, \quad (35)$$

где

$$a_k = \frac{2\ell + 2k + 1}{z}; \quad b_k = -1.$$

Для цепных дробей рекуррентные уравнения

$$P_n = a_n P_{n-1} + b_n P_{n-2}; \quad (36)$$

$$Q_n = a_n Q_{n-1} + b_n Q_{n-2}$$

с начальными условиями

$$b_0 = 1, \quad P_{-2} = 0, \quad P_{-1} = 1, \quad Q_{-2} = 1, \quad Q_{-1} = 0$$

определяют числитель и знаменатель несократимой n -й подходящей дроби

$$\delta_n = P_n / Q_n, \quad (37)$$

которая эквивалентна конечной цепной дроби, где отброшены все члены с $k > n$. При этом разность двух соседних подходящих дробей определяется соотношением

$$\delta_n - \delta_{n-1} = \frac{(-1)^{n-1} b_1 \dots b_n}{Q_{n-1} Q_n}. \quad (38)$$

Одним из признаков сходимости δ_n к конечному пределу при $n \rightarrow \infty$ для действительных чисел b_n и a_n и положительности a_n после некоторых значений n_0 является выполнение неравенств [15]:

$$a_n - |b_n| \geq 1, \quad n > n_0. \quad (39)$$

Из последнего соотношения следует, что для действительных аргументов цепная дробь (35) сходится, так как при увеличении k значение $(\ell + k + 1)/z$ всегда становится больше единицы, т.е. выполняется условие (39). При этом встает вопрос, к отношению каких функций сходится (35), так как соотношения (35)–(39) формально имеют место для любых функций Риккати–Бесселя. Ответом на этот в общем-то риторический вопрос может служить сравнение первых членов ряда (33) для $R_\ell^{\text{пр}}(z)$ и ряда, полученного делением множителя $P_n(z)$ на $Q_n(z)$. Для этого выписываем выражения для первых двух величин $P_n(z)$ и $Q_n(z)$, полученные из (35) и (36):

$$P_0 = \frac{2\ell + 1}{z}, \quad Q_0 = 1;$$

$$P_1 = \frac{2\ell + 1}{z} \frac{2\ell + 3}{z} - 1, \quad Q_1 = \frac{2\ell + 3}{z}. \quad (40)$$

Из (37), (40) видно, что с увеличением порядка подходящих дробей они начинают описывать все большее количество членов ряда (33). Таким образом, можно сделать вывод, что бесконечная дробь (35)

сходится к функции $R_\ell^\psi(z)$. Для выяснения характера сходимости подставим в (38) значение $b_n = -1$. В результате получим

$$R_{\ell,n} - R_{\ell,n-1} = -\frac{1}{Q_{n-1}Q_n}. \quad (41)$$

Из уравнения (36) видно, что при $x \leq 2(\ell + k + 1)$ все $Q_n > 0$, т.е. разность между двумя соседними подходящими дробями отрицательна. Это означает, что последовательность $R_{\ell,n}$ является монотонно убывающей с ростом n . Разность между величиной R_ℓ и $R_{\ell,n}$ с учетом (41) выражается в виде ряда

$$R - R_{\ell,n} = -\frac{1}{Q_{n+1}Q_n} \left(1 + \frac{Q_n}{Q_{n+2}} + \frac{Q_n Q_{n+1}}{Q_{n+2} Q_{n+3}} + \dots \right). \quad (42)$$

При условии $n + \ell > z$ отношение $Q_{n+k}/Q_{n+k+2} \approx 1/4$ и сумма в скобках (42), согласно выражению для суммы геометрической прогрессии, приблизительно равна $4/3$. Отсюда следует, что произведение $Q_n Q_{n+1}$, которое может быть получено при рекуррентном вычислении числителя и знаменателя подходящих дробей, может служить для выработки критерия останова рекуррентного процесса. Поэтому в разрабатываемом алгоритме вычисления R_ℓ^ψ по соотношениям (36), (37), который для действительных аргументов реализован в приведенной в приложении к данной статье подпрограмме RPSIRE для одинарной точности был использован следующий критерий останова рекуррентного процесса:

$$|Q_k Q_{k-1}| \geq C. \quad (43)$$

Для оценки C отметим, что R_ℓ^ψ меняется от $\ell/x \approx 1$ при $\ell = \ell_1^\psi$ до $(2\ell + 1)/x$ при $\ell \rightarrow \infty$. Поэтому для получения 7 верных знаков можно выбрать $C = 10^6$. Численное исследование C путем сравнения вычислений R_ℓ^ψ с одинарной и двойной точностью для $x \leq 10^5$ и $1 \leq \ell \leq x + 50$ показало, что это значение C является достаточным, поскольку при дальнейшем его увеличении для вычисления с одинарной точностью не происходит увеличения точности вычислений R_ℓ^ψ . Из результатов численных исследований следует также, что при выбранной выше величине C количество рекуррентных циклов хорошо описывается следующим соотношением:

$$k = 4 + 3,6x^{1/3} \begin{cases} -0,5(\ell - \ell_{\max}^\psi), & \ell \geq \ell_{\max}^\psi, \\ +(\ell_{\max}^\psi - \ell), & \ell < \ell_{\max}^\psi. \end{cases} \quad (44)$$

Приведенное выше соотношение дает ответ и на вопрос о выборе начальных величин n в алгоритме Миллера [4, с. 270], в котором $\psi_{\ell+n+1}(z)$ полагается равным нулю, а $\psi_{\ell+n}(z)$ — единице. Действительно, как следует из сопоставления (6) и (35), этот алгоритм эквивалентен прямому вычислению n -подходящей непрерывной дроби. Поэтому для компьютеров с 24-битовым представлением мантииссы значение n в алгоритме Миллера должно просто равняться k по (44). Однако в связи с предложенным в настоящей работе алгоритмом получения функций

$\psi_\ell(z)$ необходимость в алгоритме Миллера отпадает. Преимуществом нашего алгоритма является автоматическое определение k , а необходимость вычисления функции РБЗ в нашем алгоритме по объему вычислений сопоставима с вычислениями рядов при нормировке в алгоритме Миллера.

В алгоритме Ленца [12] процедура вычислений подходящих дробей включает в себя рекуррентное вычисление отношений

$$RP_k = P_k/P_{k-1} \text{ и } RQ_k = Q_k/Q_{k-1},$$

которые легко получаются из (36) и поэтому здесь не приводятся, а последующая подходящая дробь получается из предыдущей:

$$\delta_k = \delta_{k-1} \frac{RP_k}{RQ_k}. \quad (45)$$

Условием окончания рекуррентной процедуры в этом случае является равенство с машинной точностью величин RP_k и RQ_k . Отметим, что в этой процедуре на каждом шаге имеет место одна дополнительная операция умножения и деления, что увеличивает время вычислений. В этом же алгоритме имеется вероятность деления на 0, для исключения которой в [12] предусматривается специальная процедура. Для сравнения работы подпрограммы RPSIRE с алгоритмом Ленца была составлена соответствующая программа без указанной выше процедуры исключения деления на нуль. При этом оказалось, что условие равенства RP_k и RQ_k при работе программы [12] не могло быть выполнено, поэтому для останова рекуррентного процесса было использовано соотношение

$$\frac{|RP_k - RQ_k|}{RQ_k} \leq \varepsilon. \quad (46)$$

Значение ε , равное 10^{-6} , обеспечивало при одинарной точности предельно малую погрешность. При этом количество рекурсий и погрешность вычисления практически совпадали с таковыми для RPSIRE при $C = 10^6$. Однако время вычислений, усредненное по диапазону $x = 10 \div 1000$, для RPSIRE оказалось в 2 раза меньше.

Заключение

По результатам проведенных исследований можно сделать следующие основные выводы об устойчивости рекуррентных процедур и методах их реализации при вычислении функций Риккати–Бесселя.

Аналитическое исследование и численные результаты подтверждают общий принцип устойчивости разностных схем для получения решений, возрастающих в направлении их применения. Согласно этому принципу абсолютно устойчивыми для прямой рекурсии являются функции $\chi_\ell(x)$ для реального и $\eta_\ell(z)$ для комплексного аргументов. Функции $\psi_\ell(x)$ реального аргумента абсолютно устойчивы для прямой рекурсии только до $\ell = \ell_{\max}^\psi$, определяемых соотношением (16). При дальнейшем увеличении ℓ рекурсия становится абсолютно неустойчивой

и приводит к формированию суммы функций $\psi_\ell(x)$ и $\chi_\ell(x)$. Обратная рекурсия для функции $\psi_\ell(x)$ абсолютно устойчива при $\ell > \ell_{\max}^{\psi}$ и относительно устойчива при $\ell < \ell_{\max}^{\psi}$. Для комплексного аргумента устойчивость обратной рекурсии повышается во всем диапазоне ℓ .

При использовании рекуррентных процедур в устойчивых областях основная погрешность вычисления функций при задании аргументов в десятичной системе возникает за счет погрешностей их перевода в двоичную.

Для получения начальных значений для обратной рекурсии предлагается использовать отношение $RPSI = \psi_{\ell-1}(z)/\psi_\ell(z)$ и вычисленное по прямой рекурсии значение $\chi_\ell(x)$ или $\eta_\ell(z)$. С использованием вронскианов это приводит к соотношениям (20) и (30). Для вычисления RPSI эффективна процедура цепных дробей, упрощенная схема вычисления которых обоснована в данной статье и приводится в виде подпрограммы.

ПРИЛОЖЕНИЕ

Подпрограмма RPSIRE (ratio psi real E) служит для вычисления отношения функций Риккати–Бесселя первого рода $RPSI = \psi_{\ell-1}(z)/\psi_\ell(z)$ для реальных значений аргумента. Переход к комплексным аргументам проводится простым переописанием типа переменных, при этом соотношение (43) остается в силе. Входными параметрами подпрограммы являются порядок L и значение x , выходным параметром – значение функции RPSI.

```
SUBROUTINE RPSIRE(L,X,RPSI)
DX=2./X
W=(L+0.5)*DX
P2=1.
P1=W
Q2=0.
Q1=1.
DO 1 K=2,9999
DX=-DX
W=DX-W
P=W*P1+P2
```

N.P. Romanov. Study of methods and errors in Riccati–Bessel functions computations.

The dependence of the Riccati–Bessel function (RB) of the 1st, 2nd, 3rd and 4th kinds of the real x and complex $z = x + iy$ argument of their order ℓ are studied analytically and numerically. The regions of increasing RB functions module with growing ℓ stable to the errors of the direct recursion and the regions where the module decreased stable to the errors of the reverse recursion are being sought for. For RB1 function the stability region of the direct recursion at $x \gg 1$ is determined by the correlation $0 \leq \ell \leq \ell_{\max} = x - 0.5 - 0.80861x^{1/3} - 0.1635x^{-1/3}$ under the condition that $|y| \leq 0.4 \lg x + 0.5$. In this region a relative error of the direct recursion increases with increasing l proportionally to $\ell^{1/2}$. In the range of $\ell > \ell_{\max}$ the direct recursion results in the formation of the sum of the functions RB1 and RB2 instead of the function RB1. A relative error of the reverse recursion for the RB1 function in the whole range of ℓ does not practically depend on ℓ , it increases with the module z increase according to the law $|z|^{1/2}$ and is comparable under the above limitations on y with the direct recursion error at $\ell = \ell_{\max}$. To obtain the initial values of the RB1 at the reverse recursion a simplified procedure to calculate the ratios of these functions of the neighboring orders with the use of a continued fraction, an additional calculation of the RB2 function at $y = 0$ or RB3 at $y > 0$ by the method of the direct recursion and the use of the Wronskian of the corresponding functions are proposed. A program written in FORTRAN for computing the above-mentioned ratio of the RB1 function is presented. In the stable regions the main error in computing the functions at presetting the arguments in the decimal system can be determined by the error of their conversion into the binary computer system.

```
Q=W*Q1+Q2
IF(ABS(Q*Q1).GE.1.E+6) GO TO 2
P2=P1
Q2=Q1
P1=P
1 Q1=Q
2 RPSI=P/Q
END
```

1. Хюлст Г. Рассеяние света малыми частицами. М.: Изд-во иностр. лит-ры, 1961. 536 с.
2. Дейрменджан Д. Рассеяние электромагнитного излучения сферическими полидисперсными частицами. М.: Мир, 1971. 167 с.
3. Борен К., Хафмен Д. Поглощение и рассеяние света малыми частицами. М.: Мир, 1986. 660 с.
4. Справочник по специальным функциям / Под ред. М. Абрамовица и И. Стиган. М.: Наука, 1979. 832 с.
5. Акулинин А.А. Особенности применения алгоритма расчета характеристик рассеяния по теории Ми // Оптика атмосф. 1988. Т. 1. № 6. С. 127–129.
6. Акулинин А.А. Обзор некоторых схем расчета оптических характеристик рассеяния электромагнитного излучения одиночной однородной сферической частицей. Деп. в ВИНТИ. 16.04.1987, № 3388–В87.
7. Люк Ю. Специальные математические функции и их аппроксимации. М.: Мир, 1980. 608 с.
8. Kattawar G.W., Plass G.N. Electromagnetic scattering from absorbing spheres // J. Opt. Soc. Amer. 1967. V. 6. N 8. P. 1377–1382.
9. Wiscombe W.J. Improved Mie scattering algorithms // Appl. Opt. 1980. V. 19. N 9. P. 1505–1509.
10. Гаурис И.Б. Алгоритмы расчета сферических функций Бесселя, присоединенных функций Лежандра и реализация их на языке ФОРТРАН // Вестн. БГУ. 1974. Сер. 1. № 3. С. 8–16.
11. Романов Н.П. Классификация и свойства собственных частот электромагнитных колебаний шара // Тр. ИЭМ. 1987. Вып. 45 (135). С. 3–40.
12. Lentz W.J. Generating Bessel functions in Mie scattering calculations using continued fractions // Appl. Opt. 1976. V. 15. N 3. P. 668–671.
13. Любимский Э.З., Мартынюк В.В., Трифонов Н.П. Программирование. М.: Физматгиз, 1980. 608 с.
14. Белов Н.Н. Расчет функций Риккати–Бесселя восходящей рекурсией // Оптика атмосф. и океана. 1992. Т. 5. № 2. С. 165–168.
15. Демидович Б.П., Марон И.А. Основы вычислительной математики. М.: Физматгиз, 1960. 660 с.