ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
21,2 (1991), pp. 289–301

S. PASZKOWSKI (Wrocław and Lille)

EVALUATION OF THE FERMI–DIRAC INTEGRAL OF HALF-INTEGER ORDER

1. Introduction. The Fermi–Dirac integral F_μ is defined by

$$F_\mu(z) = \int_0^\infty \frac{x^\mu dx}{1+e^{x-z}} \quad (\mu > -1).$$

Cody and Thacher [2] have obtained rational approximants (in z or another variable) to this integral for $\mu = -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}$. To find them the authors computed first a table of the values $F_\mu(z)$. All computations were carried out in 25-decimal arithmetic. Methods which were used (in particular, a Sommerfeld–Dingle expansion for $z \geq 4$) gave, however, rather poor accuracies, about 9–13 significant digits. An accuracy of approximants based on these values $F_\mu(z)$ was a priori still more limited, up to 8–9 significant digits.

In [10], Section 3, the author of this paper gave a new method of computing F_μ for half-integer μ and for rather small z, say for $z \leq 2$. (N.B. In the abstract of [10] and at the beginning of Section 1 the upper limit of integration in the definition of F_μ is evidently erroneous.) The goal of this paper is to show that some formulae from [10], Section 4, with complementary techniques (continued fractions, convergence acceleration) are better than those used by Cody and Thacher. In fact, it turns out that it is possible to evaluate $F_\mu(z)$ for half-integer μ and for sufficiently great z, say for $z \geq 2$, with accuracy only a bit worse than that which is guaranteed by the arithmetic in use. The same method permits us to evaluate some linear combinations of F_μ important in certain physical applications. The method is rather time-consuming and it can be recommended when a table of the values $F_\mu(z)$ is needed to construct an approximant of the Fermi–Dirac integral.

1991 *Mathematics Subject Classification*: Primary 65D20, 41A20, 30B70.

Key words and phrases: Fermi–Dirac integral, continued fraction, series expansion, convergence acceleration, best rational approximation.

2. Continued fractions. In [10], Section 4, the following expansion was proved:

$$(1)\qquad F_{m-1/2}(z)= -\frac{4}{2m+1}\sum_{j=0}^{[m/2]}(-1)^j\left(m+\frac{3}{2}-2j\right)_{2j}(2^{2j-1})\frac{\pi^{2j}B_{2j}}{(2j)!}z^{m+1/2-2j} + 2(-1)^{m-1}\left(\frac{1}{2}\right)_m\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k^{m+1/2}}G_{\sigma(m)}(\sqrt{kz})\quad (m=1,2,\ldots),$$

where $(t)_j$ denotes the Pochhammer symbol (i.e., $(t)_0 := 1$, $(t)_j := t(t+1)\ldots(t+j-1)$ for $j=1,2,\ldots$), B_{2j} is the $(2j)$th Bernoulli number, $\sigma(m) := (-1)^m$,

$$(2)\qquad G_{\pm1}(t) := e^{-t^2}\operatorname{Erfi}(t) + e^{t^2}\operatorname{Erfc}(t),$$

and Erfi, Erfc are known special functions:

$$\operatorname{Erfi}(t) := \int_0^t e^{y^2}\,dy,\qquad \operatorname{Erfc}(t) := \int_t^\infty e^{-y^2}\,dy.$$

There are several continued fraction expansions for these functions. For sufficiently great positive t the following continued fractions are recommended:

$$e^{-t^2}\operatorname{Erfi}(t) = \frac{\frac{1}{2}t}{\frac{1}{2}+t^2}-\frac{t^2}{\frac{3}{2}+t^2}-\frac{2t^2}{\frac{5}{2}+t^2}-\frac{3t^2}{\frac{7}{2}+t^2}-\ldots$$

([7], (7.3.108); it converges for every complex t) and

$$e^{t^2}\operatorname{Erfc}(t) = \frac{\frac{1}{2}}{t}+\frac{\frac{1}{2}}{t}+\frac{1}{t}+\frac{\frac{3}{2}}{t}+\frac{2}{t}+\frac{\frac{5}{2}}{t}+\ldots$$

([7], (6.2.24); it converges if $\operatorname{Re} t > 0$).

Substituting above $u := \frac{1}{2}t^{-2}$ we obtain

$$(3)\qquad 2te^{-t^2}\operatorname{Erfi}(t) = \frac{1}{1+u}-\frac{2u}{1+3u}-\frac{4u}{1+5u}-\frac{6u}{1+7u}-\ldots,$$

$$(4)\qquad 2te^{t^2}\operatorname{Erfc}(t) = \frac{1}{1}+\frac{u}{1}+\frac{2u}{1}+\frac{3u}{1}+\frac{4u}{1}+\ldots,$$

It can be proved that the nth approximant of (3) is equal to $A_n(u)/B_n(u)$, where

$$B_n(u) = \sum_{k=0}^{n}\frac{(\frac{1}{2})_k(-n)_k}{k!}(-2u)^k\quad (n=0,1,\ldots).$$

It is a polynomial in u, with positive coefficients. Then a known formula implies that for $u > 0$, $n > 0$

$$\frac{A_n(u)}{B_n(u)} - \frac{A_{n-1}(u)}{B_{n-1}(u)} = (-1)^{n-1}\frac{(-2u)(-4u)\dots[-(2n-2)u]}{B_{n-1}(u)B_n(u)} > 0$$

and the sequence $\{A_n(u)/B_n(u)\}$ increases. The Nth approximant, calculated by the three-term recurrence formula for the numerators $A_n(u)$ and the denominators $B_n(u)$, can be accepted as the best one if the computed values satisfy the inequality $A_N(u)/B_N(u) \geq A_{N+1}(u)/B_{N+1}(u)$ (contrary to an inequality satisfied by the exact values). To check the accuracy of the results the values $A_N(u)/B_N(u)$ with the same N were computed by the backward algorithm. The results obtained by the two methods were identical. The speed of convergence of the continued fraction (3) turned out to be quite satisfactory for any $u > 0$. In fact, when Turbo Pascal 4 for IBM PC and extended type variables are used, the N required to obtain $2t\exp(-t)^2\,\mathrm{Erfi}(t)$ with 17–18 significant digits increases from 13 for $u = 0.002$ to 67 for $u = 0.014, 0.016, 0.018$ and then decreases slower and slower to 37 for $u = 0.1$, 28 for $u = 0.2$ and 25 for $u = 0.3$.

Let us now turn to the formula (4). For $u > 0$ the approximants of this continued fraction are alternately greater and smaller than its value. Therefore we can accept, as the best computed value of $2t\exp(t^2)\,\mathrm{Erfc}(t)$, the first Nth approximant (computed via numerators and denominators) which, together with the $(N+1)$st one, violates the property stated above. In this case N increases a little faster than u. Using Turbo Pascal 4 for IBM PC and extended type variables the values 20, 119, 142, 280, 601, 1095, 1490 and 2008 of N were obtained for $u = 0.01, 0.1, 0.2, 0.5, 1, 2, 3$ and 4, respectively. If, for example, $z \geq 2$ in (1), then $t \geq \sqrt{2}$, $u \leq \frac{1}{4}$ and the maximal N needed is about 160.

Because of great values of N it is natural to seek a method of convergence acceleration suitable for the continued fraction (4). Some methods were tested. (4) is a particular case of the continued fraction

$$\frac{a_1}{1} + \frac{a_2}{1} + \dots \tag{5}$$

such that $\{a_n\} \to \infty$; namely, $a_n = (n-1)u$ for $n > 1$. Jacobsen, Jones and Waadeland [6] suggest that the approximants

$$S_n := \frac{a_1}{1} + \frac{a_2}{1} + \dots + \frac{a_n}{1}$$

of (5) should be replaced by the modified ones:

$$S_n(w_n) := \frac{a_1}{1} + \frac{a_2}{1} + \dots + \frac{a_n}{1 + w_n}$$

with

$$w_n \equiv w'_n := \sqrt{a_{n+1} + \frac{1}{4}} - \frac{1}{2},$$

i.e., in the case (4), with

$$w'_n := \sqrt{nu + \frac{1}{4}} - \frac{1}{2}. \tag{6}$$

A more sophisticated choice of w_n results from Hautot's [5] considerations. In his notations $\alpha_k = a_{k+1}$. In general, α_k should be expanded in powers of k^{-1}:

$$\alpha_k = k^\gamma(a_0 + a_1/k + a_2/k^2 + \ldots)$$

(a_k used here has clearly nothing in common with the notations in (5)). In the case (4) this expansion is extremely simple: $\gamma = 1$, $a_0 = u$, $a_j = 0$ for $j > 0$. Using the corresponding entry of Table 3 in [5] the expression

$$w_n \equiv w''_n := -1 + \sqrt{n}\sum_{i=0}^{\infty} p_i/(\sqrt{n})^i \tag{7}$$

can be obtained, with

$$\begin{aligned} &p_0 = \sqrt{u}, \quad p_1 = \frac{1}{2}, \quad p_2 = \frac{-2u+1}{8\sqrt{u}}, \\ &p_3 = \frac{1}{8}, \quad p_4 = \frac{4u^2+4u-1}{128u\sqrt{u}}, \quad p_5 = \frac{2u-1}{32u}, \ldots \end{aligned} \tag{8}$$

The expressions (6) and (7) are, of course, nearly identical:

$$\sqrt{nu + \frac{1}{4}} - \frac{1}{2} = nu - \frac{1}{2} + \frac{1}{8\sqrt{nu}} + \ldots,$$

$$-1 + \sqrt{n}\left(p_0 + \frac{p_1}{\sqrt{n}} + \frac{p_2}{n} + \ldots\right) = nu - \frac{1}{2} + \frac{-2u+1}{8\sqrt{nu}} + \ldots$$

If $z \geq 2$ in (1), then $t \geq \sqrt{2}$ in (2) and $u \leq 0.25$ in (4). Nevertheless, in order to investigate in detail the two methods of convergence acceleration for the continued fraction (4), a larger range of u, namely the interval $[0.05, 5]$, was examined. The results of numerical computations lead us to the following conclusions:

(i) The sequence $\{S'_n\}$ where $S'_n := S_n(w'_n)$ (cf. (6)) behaves quite regularly:

$$S'_1 < S'_3 < S'_5 < \ldots < S'_6 < S'_4 < S'_2.$$

Thus, we can stop the calculations when the evaluated values violate this property, i.e. for N' such that

$$(-1)^{N'}(S'_{N'} - S'_{N'+1}) \leq 0. \tag{9}$$

Let us recall here that for the sequence of approximants S_n an analogous stopping criterion has the form

$$(-1)^N(S_N - S_{N+1}) \geq 0\,.$$

The ratio N'/N changes rather regularly from 0.83 to 0.58 when u ranges between 0.05 and 5.

It is well known that $S_n(w_n)$ (whatever the w_n are) can be calculated by the formula

$$S_n(w_n) = \frac{A_n + w_n A_{n-1}}{B_n + w_n B_{n-1}}\,,$$

where A_n and B_n denote the numerator and the denominator, respectively, of S_n. Taking this into account, together with the necessity to evaluate the expression $(a_{n+1} + \frac{1}{4})^{1/2} - \frac{1}{2}$, one can assert that evaluation of the sequence $\{S'_n\}$ is at least twice as time-consuming as evaluation of the sequence $\{S_n\}$. Therefore, in the case of the continued fraction (4), the use of modifying factors w'_n cannot be recommended for $u \in [0.05, 5]$.

(ii) The modifying factors w''_n (cf. (7), (8)) are better than the w'_n in the sense that $S''_n := S_n(w''_n)$ approximates the value of (4) more accurately than S'_n does. On the other hand, the sequence $\{S''_n\}$ behaves, locally, rather irregularly. In particular, it is possible that S''_n, while being identical with S''_{n+1}, differs from the value of (4). For example, if $u = 0.180667$ then the numbers S''_n for $n = 2, 3, \ldots, 7$ equal

$$\begin{aligned} 0.8749&009,\\ &009,\\ &271,\\ &214,\\ &227,\\ &223, \end{aligned}$$

respectively. Thus, in the Hautot method, it is more difficult to decide whether S''_n sufficiently well approaches the value of (4) or not. The following test turned out to work well:

$$\left|1 - \frac{S''_{N''-1}}{S''_{N''}}\right| + \left|1 - \frac{S''_{N''}}{S''_{N''+1}}\right| < \varepsilon\,.$$

ε denotes here a small number exceeding several times ϱ, i.e. the relative precision of numbers in use. For $\varrho = 1.4 \times 10^{-17}$ the value $\varepsilon = 10^{-16}$ was chosen with success.

Obviously, the above test is more complicated (and less elegant, too) than (9). The same is true for the modifying factors. The numbers p_0, $p_1 - 1, p_2, \ldots, p_5$ may be calculated only once but the evaluation of w_n, for example by the formula

$$w_n = (p_5/n + p_3)/n + p_1 - 1 + \sqrt{n}[(p_4/n + p_2)/n + p_0],$$

remains to be done for each n. It requires one square extraction, four divisions, one multiplication and five additions. Thus, the cost of evaluating S''_n can be estimated as 3.5 times greater than for S_n. In other words, in the case of the continued fraction (4) the Hautot method can be recommended only if $N'' < \frac{2}{7}N$. This inequality holds only for $u \geq 0.9$, which rather excludes the use of the Hautot method in our main problem, namely that of the Fermi–Dirac integral evaluation.

Taking into account a regular behaviour of the sequence $\{S'_n\}$, an attempt was made to accelerate its convergence by Levin's t transformation. More precisely, the following formulae were applied for $n = 1, 2, \ldots$:

$$\xi_n := S'_{n+1} - S'_n, \quad P_n^{(0)} := \frac{S'_n}{\xi_n}, \quad Q_n^{(0)} := \frac{1}{\xi_n},$$

$$P_{n-1}^{(1)} := P_{n-1}^{(0)} - P_n^{(0)}, \quad Q_{n-1}^{(1)} := Q_{n-1}^{(0)} - Q_n^{(0)} \quad (\text{if } n > 1),$$

$$\left.\begin{aligned} P_{n-j}^{(j)} &:= (n-j+1)P_{n-j}^{(j-1)} - (n+1)P_{n-j+1}^{(j-1)} \\ Q_{n-j}^{(j)} &:= (n-j+1)Q_{n-j}^{(j-1)} - (n+1)Q_{n-j+1}^{(j-1)} \end{aligned}\right\} (\text{if } n > 2;\ j = 2, \ldots, n-1),$$

$$S_{n-j}^{(j)} := \frac{P_{n-j}^{(j)}}{Q_{n-j}^{(j)}} \quad (j = 0, \ldots, n-1)$$

(cf. [4]). For an easily accelerable sequence converging to S each antidiagonal

$$S_n^{(0)},\ S_{n-1}^{(1)}, \ldots, S_1^{(n-1)} \tag{10}$$

ends with the best approximation of S. This is not the case for the sequence $\{S'_n\}$, as shown in the following table. Its each entry, j, indicates the number $S_{n-j}^{(j)}$ which is, on the antidiagonal (10), the best approximation of the value of (4). Obviously $j > 1$, except for small values of u or n:

u \ n	2	3	4	5	6	7	8	9	10	11	12
0.05	1	1	1	1	1	1					
0.25	1	1	1	1	1	1	1	1	3	3	3
0.5	1	1	1	1	2	2	4	4	6		
1	1	1	2	3	4	4	5	6	7	8	
5	1	1	2	3	4	5	6	7	8	9	

Briefly, it is not a good idea to apply Levin's t transformation to the sequence $\{S'_n\}$.

Some other methods of convergence acceleration were also tested in the case (4). The final conclusions are rather discouraging: probably the cheapest method of evaluating the values of (4) consists in using this continued fraction without any changes or transformations.

3. Summation of the series (1). To find a value of the Fermi–Dirac integral $F_{m-1/2}$ we can use the expansion (1). We will show that owing to the particular form of the series in (1), a considerable acceleration of its convergence is possible. This is important because, as we already know, the calculation of each term of this series is very expensive.

Let us remark first that every term $G_{\pm 1}(\sqrt{kz})$ in (1) requires the evaluation of the continued fractions (3) and (4) for $u = 1/(2kz)$. As $k \to \infty$, the fractions both tend to 1. Therefore, to improve slightly the accuracy of the results, it is worthwhile to introduce the auxiliary functions:

$$\mathrm{Fi}(t) := 2te^{-t^2}\,\mathrm{Erfi}(t) - 1\,, \quad \mathrm{Fc}(t) := 2te^{t^2}\,\mathrm{Erfc}(t) - 1\,.$$

Then

$$G_1(t) = \frac{1}{2t}[\mathrm{Fi}(t) + \mathrm{Fc}(t) + 2]\,, \quad G_{-1}(t) = \frac{1}{2t}[\mathrm{Fi}(t) - \mathrm{Fc}(t)]\,. \tag{11}$$

Fi and Fc can be calculated similarly to the fractions (3) and (4). In fact, instead of the numerator A_n of each approximant A_n/B_n, it suffices to evaluate the difference $A_n - B_n$ obeying the same three-term recurrence formula as A_n and B_n.

Let us remark next that if m in (1) is an odd integer, $m = 2l - 1$, then there the function G_1 occurs. According to (11) we decompose the series from (1) into two series. The first contains the functions Fi and Fc and the second is

$$\frac{1}{\sqrt{z}}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k^{2l}} = \frac{(2^{2l}-1)\pi^{2l}}{(2l)!}B_{2l}\frac{1}{\sqrt{z}}\,.$$

In particular,

$$\begin{aligned} F_{1/2}(z) &= \frac{2}{3}z^{3/2} + \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k^{3/2}}G_1(\sqrt{kz}) \\ &= \frac{2}{3}z^{3/2} + \frac{1}{12}\pi^2 z^{-1/2} + \frac{1}{2}z^{-1/2}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k^2}[\mathrm{Fi}(\sqrt{kz}) + \mathrm{Fc}(\sqrt{kz})]\,, \end{aligned}$$

$$\begin{aligned} F_{3/2}(z) &= \frac{2}{5}z^{5/2} + \frac{1}{4}\pi^2 z^{1/2} - \frac{3}{2}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k^{5/2}}G_{-1}(\sqrt{kz}) \\ &= \frac{2}{5}z^{5/2} + \frac{1}{4}\pi^2 z^{1/2} - \frac{3}{4}z^{-1/2}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k^3}[\mathrm{Fi}(\sqrt{kz}) - \mathrm{Fc}(\sqrt{kz})]\,, \end{aligned}$$

$$F_{5/2}(z) = \frac{2}{7}z^{7/2} + \frac{5}{12}\pi^2 z^{3/2} + \frac{15}{4}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k^{7/2}}G_1(\sqrt{kz})$$

$$= \frac{2}{7}z^{7/2} + \frac{5}{12}\pi^2 z^{3/2} + \frac{7}{192}\pi^4 z^{-1/2} + \frac{15}{8}z^{-1/2}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k^4}[\mathrm{Fi}(\sqrt{kz}) + \mathrm{Fc}(\sqrt{kz})].$$

In some physical applications the following linear combinations of these integrals are also needed:

$$(12)\quad F_{1/2;1}(z) := \frac{5}{2}F_{3/2}(z) - \frac{3}{2}zF_{1/2}(z) = \frac{1}{2}\pi^2 z^{1/2} - \frac{3}{4}z^{1/2}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k^2}[\mathrm{Fi}(\sqrt{kz}) + \mathrm{Fc}(\sqrt{kz})] - \frac{15}{8}z^{-1/2}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k^3}[\mathrm{Fi}(\sqrt{kz}) - \mathrm{Fc}(\sqrt{kz})],$$

$$(13)\quad F_{1/2;2}(z) := \frac{7}{2}F_{5/2}(z) - 5zF_{3/2}(z) + \frac{3}{2}z^2F_{1/2}(z) = \frac{1}{3}\pi^2 z^{3/2} + \frac{49}{384}\pi^4 z^{-1/2} + \frac{3}{4}z^{3/2}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k^2} \times [\mathrm{Fi}(\sqrt{kz}) + \mathrm{Fc}(\sqrt{kz})] + \frac{15}{4}z^{1/2}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k^3}[\mathrm{Fi}(\sqrt{kz}) - \mathrm{Fc}(\sqrt{kz})] + \frac{105}{16}z^{-1/2}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k^4}[\mathrm{Fi}(\sqrt{kz}) + \mathrm{Fc}(\sqrt{kz})].$$

For great z the values $F_{3/2}(z)$ and $F_{5/2}(z)$ are of order $z^{5/2}$ and $z^{7/2}$, respectively, whereas the values $F_{1/2;1}(z)$ and $F_{1/2;2}(z)$ are of order $z^{1/2}$ and $z^{3/2}$, respectively. Therefore it is evident that evaluating these linear combinations by means of the tabulated values of the Fermi–Dirac integrals (cf. [1], [8]) may cause a considerable loss of accuracy. The formulae (12), (13) do not have this drawback.

It follows from (3) and (4) that for great t (i.e., for small u)

$$2te^{-t^2}\,\mathrm{Erfi}(t) \approx 1 + u + 3u^2, \quad 2te^{t^2}\,\mathrm{Erfc}(t) \approx 1 - u + 3u^2.$$

Then, as $k \to \infty$,

$$\mathrm{Fi}(\sqrt{kz}) + \mathrm{Fc}(\sqrt{kz}) \sim \frac{3}{2k^2z^2}, \quad \mathrm{Fi}(\sqrt{kz}) - \mathrm{Fc}(\sqrt{kz}) \sim \frac{1}{kz}$$

and hence all the series occurring in the formulae for the Fermi–Dirac integrals and their linear combinations (12), (13) behave like the series

$$\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k^m} \quad (m = 4, 6).$$

They can be accelerated by the following algorithm applicable, more generally, to sequences $\{s_k\}$ such that

$$s_k = s + (-1)^k \sum_{j=1}^{\infty} \frac{c_j}{k^{j+3}} \quad (k = 1, 2, \ldots).$$

The algorithm belongs to a wide class of series transformations including, among others, Levin's transformation, but, contrary to the latter, is linear with respect to the elements s_k: for $n = 1, 2, \ldots$

$$P_n^{(0)} := n^3 s_n, \quad Q_n^{(0)} := n^3,$$

$$\left.\begin{aligned} P_{n-j}^{(j)} &:= (n-j)P_{n-j}^{(j-1)} + nP_{n-j+1}^{(j-1)} \\ Q_{n-j}^{(j)} &:= (n-j)Q_{n-j}^{(j-1)} + nQ_{n-j+1}^{(j-1)} \end{aligned}\right\} \quad (j = 1, \ldots, n-1),$$

$$s_{n-j}^{(j)} := \frac{P_{n-j}^{(j)}}{Q_{n-j}^{(j)}} \quad (j = 0, \ldots, n-1).$$

It has been experimentally checked that for each $z \geq 2$ at most 15 terms of the series suffice to guarantee about 16–17 significant digits in the values of $F_{1/2}$, $F_{3/2}$, $F_{5/2}$ and of their linear combinations $F_{1/2;1}$, $F_{1/2;2}$. This is rather surprising bearing in mind the difficulties caused by the function Erfc (cf. Section 2).

4. Rational approximations. In [2] approximate expressions for the functions F_μ ($\mu = -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}$) were given. Similar approximation forms are applicable to the linear combinations $F_{1/2;1}(z)$ and $F_{1/2;2}(z)$ (cf. (12), (13)). Their approximation is necessary at least for sufficiently great z because these combinations are then considerably less than their sums.

Consequently, one should solve some problems of uniform (Chebyshev) weighted rational approximation. For example, bearing in mind (13) and the conclusions of [2], we seek a rational function R such that for a sufficiently great $a > 0$ (for $a = 7$, say)

$$F_{1/2;2}(z) \approx \frac{1}{3}\pi^2 z^{3/2} + z^{-1/2} R(z^{-2}) \quad (z \geq a).$$

It is natural to minimize here the relative error of approximation, defined as

$$\left\| \frac{R(z^{-2}) - z^{1/2} F_{1/2;2}(z) - \frac{1}{3}\pi^2 z^2}{z^{1/2} F_{1/2;2}(z)} \right\|_{[a,\infty)},$$

where

$$\|D\|_I := \sup_{z\in I} |D(z)| .$$

In other words, $R(y)$ should approximate the function

$$F(y) := y^{-1/4} F_{1/2;2}(y^{-1/2}) - \frac{1}{3}\pi^2 y^{-1}$$

in $I := [0, a^{-1/2}]$, with the weight function $G(y) := y^{-1/4} F_{1/2;2}(y^{-1/2})$. The minimum of $\|(R-F)/G\|_I$ is taken over all the rational functions $R = P/Q$ with numerators P of degree $\le l$ and denominators Q of degree $\le m$, for some fixed l, m. Such a problem is well known (cf., for example, [9] or [11]). To solve it, the so-called second Remes algorithm is often recommended. Applying this method one should solve many systems of equations like this one:

$$\frac{R(x_k) - F(x_k)}{G(x_k)} = (-1)^k \varepsilon \quad (k = 0, \ldots, l+m+1), \tag{14}$$

where $x_k \in I$, $x_0 < x_1 < \ldots < x_{l+m+1}$. Besides the coefficients of P and Q, there is an extra unknown here, namely ε. Several methods of solving such nonlinear systems were developed. Therefore we will describe only a detail of our Remes algorithm implementation, possibly distinguishing it from earlier ones.

Let, for the moment, P, Q, R be arbitrary functions defined on $\{x_0, x_1, \ldots\}$ and such that $P = QR$. Their divided differences satisfy the identity

$$P(x_0, \ldots, x_n) = \sum_{j=0}^{n} Q(x_0, \ldots, x_j) R(x_j, \ldots, x_n) \quad (n = 0, 1, \ldots)$$

(cf., for example, [3], p. 158). If P and Q are polynomials of degrees defined as above then

$$P(x_i, \ldots, x_{i+s}) = 0 \ \ (s \ge l+1), \quad Q(x_i, \ldots x_{i+s}) = 0 \ \ (s \ge m+1)$$

and hence

$$\sum_{j=0}^{\min\{n,m\}} Q(x_0, \ldots, x_j) R(x_j, \ldots, x_n) = 0 \quad (n = l+1, \ldots, l+m+1). \tag{15}$$

Let

$$Q = \sum_{i=0}^{m} c_i Q_i, \quad \text{where } Q_i(x) := (x - x_0) \ldots (x - x_{i-1}) \ \ (i = 0, \ldots, m).$$

Then

$$Q_i(x_0, \ldots, x_j) = \begin{cases} 0 & (j \ne i), \\ 1 & (j = i) \end{cases}$$

and (15) implies the following linear homogeneous system satisfied by the coefficients c_i:

$$\sum_{j=0}^{\min\{n,m\}} R(x_j,\ldots,x_n)c_j = 0 \quad (n = l+1,\ldots,l+m+1).$$

Hence, provided that R does not vanish identically,

$$\Delta(\varepsilon) := \begin{vmatrix} R(x_0,\ldots,x_{l+1}) & \ldots & R(x_m,\ldots,x_{l+1}) \\ \ldots & \ldots & \ldots \\ R(x_0,\ldots,x_{l+m+1}) & \ldots & R(x_m,\ldots,x_{l+m+1}) \end{vmatrix} = 0 \tag{16}$$

(cf. [3], p. 158, (3.16b)), where $R(x_i,\ldots,x_j) := 0$ for $i > j$. It is worthwhile to remark that (16) is a difference analogue of the differential condition

$$\begin{vmatrix} [R(x)]^{(l+1)} & \ldots & [x^m R(x)]^{(l+1)} \\ \ldots & \ldots & \ldots \\ [R(x)]^{(l+m+1)} & \ldots & [x^m R(x)]^{(l+m+1)} \end{vmatrix} = 0$$

characterizing rational functions R of the type defined above ([9], §9.2).

According to (14),

$$R(x_k) = F(x_k) + \varepsilon H(x_k) \quad (k = 0,\ldots,l+m+1), \tag{17}$$

where

$$H(x_k) := (-1)^k G(x_k) \quad (k = 0,\ldots,l+m+1).$$

Hence in the determinant $\Delta(\varepsilon)$ each entry $R(x_i,\ldots,x_j)$ equals $F(x_i,\ldots,x_j) + \varepsilon H(x_i,\ldots,x_j)$. The quantity ε can be evaluated by solving the equation $\Delta(\varepsilon) = 0$ (i.e., a generalized eigenproblem). In computations the regula falsi method turned out to be very efficient.

After finding ε the rational function R can be evaluated from the interpolation conditions (17) for $k = 0,\ldots,l+m$. The resulting value of $(R-F)/G$ at x_{l+m+1} permits one to verify the precision of the results.

There are well known difficulties associated with rational interpolation and approximation, such as nonexistence of an interpolant, rather complicated characterization of the best approximant and so on. Fortunately in none of the considered cases did such phenomena occur.

Finally, four rational functions R_1, R_1^∞, R_2, R_2^∞ were obtained. They are such that

$$\begin{aligned} F_{1/2;1}(z) &\approx \begin{cases} R_1(z) & (2 \le z \le 7), \\ \frac{1}{2}\pi^2 z^{1/2} + z^{-3/2} R_1^\infty(z^{-2}) & (7 \le z < \infty), \end{cases} \\ F_{1/2;2}(z) &\approx \begin{cases} R_2(z) & (2 \le z \le 7), \\ \frac{1}{3}\pi^2 z^{3/2} + z^{-1/2} R_2^\infty(z^{-2}) & (7 \le z < \infty). \end{cases} \end{aligned} \tag{18}$$

The numerator (N) and the denominator (D) of each rational function $R(x)$ is a polynomial of degree four. The coefficients of $x^0,\ldots,x^4$ in these polynomials are given below; the notation $m\mathrm{E}e$ is used there for $m \times 10^e$. The last row of the table contains the maximal relative errors (ME) of the approximation forms (18). It should be noted that the interval $[2,\infty)$ was split into two subintervals [2,7] and $[7,\infty)$ in such a manner that these four errors be approximately equal.

	R_1	R_1^∞	R_2	R_2^∞
N	7.8490060468E4	−2.2163308226E−7	−2.4803761104E5	1.2941864362E−6
	4.9370456818E4	1.2333674553E−5	−1.2196627916E5	−2.2844809340E−5
	1.4969689827E4	−2.8706825487E−3	−3.1694056504E4	1.7929705949E−2
	2.1640900717E3	2.2158396050E−2	−4.8961272394E3	7.5620665001E−2
	4.9084345912E1	−3.2759754440	−1.5941589997E3	1.8821262831E1
D	2.7193572913E4	7.8012049691E−8	−2.2974666173E4	7.5920649530E−8
	2.8278463515E3	−4.8022523179E−6	−2.1633269279E3	−1.4882593217E−6
	1.7193827267E3	1.0310564589E−3	−1.1508828042E3	1.0519838079E−3
	6.7216742252E1	−1.3856896622E−2	−8.2815885995E1	2.5543094061E−3
	1	1	1	1
ME	5.0E−9	3.0E−9	6.2E−9	2.5E−9

Acknowledgements. The author wishes to express his thanks to Dr. K. Durczewski for suggesting the problem. He is greatly indebted to Prof. C. Brezinski for the excellent working conditions at the University of Lille (where many numerical tests were performed and the article was finished), as well as for helpful remarks concerning some reference papers. The author wishes also to express his gratitude to Dr. E. J. Weniger for his valuable comments.

References

[1] A. C. Beer, M. N. Chase and P. F. Choquard, *Extension of McDougall–Stoner tables*, Helv. Phys. Acta 28 (1955), 529–541.

[2] W. J. Cody and H. C. Thacher, Jr., *Rational Chebyshev approximations for Fermi–Dirac integrals of orders* $-\frac{1}{2}$, $\frac{1}{2}$ *and* $\frac{3}{2}$, Math. Comput. 21 (1967), 30–40.

[3] A. Cuyt and L. Wuytack, *Nonlinear Methods in Numerical Analysis*, Stud. in Comput. Math. 1, North-Holland, Amsterdam 1987.

[4] T. Fessler, W. F. Ford and D. A. Smith, *HURRY: An acceleration algorithm for scalar sequences and series*, ACM Trans. Math. Software 9 (1983), 346–354.

[5] A. Hautot, *Convergence acceleration of continued fractions of Poincaré's type*, Appl. Numer. Math 4 (1988), 309–322.

[6] L. Jacobsen, W. B. Jones and H. Waadeland, *Convergence acceleration for continued fractions* $K(a_n/1)$, *where* $a_n \to \infty$, in: *Rational Approximation and Its Applications in Mathematics and Physics*, J. Gilewicz, M. Pindor and W. Siemaszko (eds.), Lecture Notes in Math. 1237, Springer, New York 1987, 177–187.

[7] W. B. Jones and W. J. Thron, *Continued Fractions: Analytic Theory and Applications*, Encyclopedia Math. Appl. 11, Addison-Wesley, Reading, Mass., 1980.
[8] J. McDougall and E. C. Stoner, *The computation of Fermi–Dirac functions*, Philos. Trans. Roy. Soc. London Ser. A 237 (1939), 67–104.
[9] G. Meinardus, *Approximation von Funktionen und ihre numerische Behandlung*, Springer, Berlin 1964.
[10] S. Paszkowski, *Evaluation of Fermi–Dirac integral*, in: *Nonlinear Numerical Methods and Rational Approximation*, A. Cuyt (ed.), Reidel, Dordrecht 1988, 435–444.
[11] J. R. Rice, *The Approximation of Functions*, Vol. 2: *Nonlinear and Multivariate Theory*, Addison-Wesley, Reading, Mass., 1969.

STEFAN PASZKOWSKI
INSTITUTE OF LOW TEMPERATURE AND STRUCTURE RESEARCH
POLISH ACADEMY OF SCIENCES
P.O. BOX 937
50-950 WROCLAW, POLAND;
UNIVERSITÉ DES SCIENCES ET TECHNIQUES DE LILLE
FLANDRES-ARTOIS
59655 VILLENEUVE D'ASCQ, FRANCE

Received on 14.2.1990