Computer Physics Communications 55 (1989) 127–136
North-Holland, Amsterdam

NUMERICAL CALCULATION OF THE GENERALIZED FERMI–DIRAC INTEGRALS

Bernard PICHON

Département d'Astrophysique Relativiste et de Cosmologie, Observatoire de Paris – Meudon, 92195 Meudon Principal Cedex, France

and

Institut d'Astronomie, d'Astrophysique et de Géophysique, Université Libre de Bruxelles, CP 165, 1050 Brussels, Belgium

Received 8 July 1988

The aim of this paper is to provide the best compromise between accuracy and CPU-time for calculating the generalized Fermi–Dirac integrals. After a brief review of the existing methods, with emphasis on two of them, we present our method valid for all values of the degeneracy parameter and of the temperature, with an accuracy in the range 10^{-5}–10^{-7} and a short CPU-time. We have taken care of the computer implementation (including round-off errors, underflows or overflows) and we give the weights and abcissas of the Gauss–Laguerre quadrature rule needed for an accuracy of $\epsilon = 10^{-6}$.

1. Introduction

Quantum statistical distributions are frequently used in physics and, in particular, the Fermi–Dirac distribution must be used to describe many astrophysical situations such as the equilibrium structure of white dwarfs, the helium flash in moderately massive stars (see ref. [1]), the gravitational collapse of massive stars and formation of supernovae, in models describing gamma ray bursts [2] or the tidal pinching of white dwarfs by a massive black hole [3].

For the study of white dwarfs structure the gas of electrons is generally considered as a (completely) degenerate gas at *zero* temperature and in this case, this approximation (which is formally equivalent to an *infinite* degeneracy) is a good one as long as the entropy is not considered. But then, non-adiabatic evolution cannot be described. Such flaws are naturally removed when we use a non-zero temperature formalism.

For the other astrophysical applications quoted above, the non-zero temperature formalism is the only one which is relevant since in the medium in which these various phenomena occur, the electron gas is semi-degenerate semi-relativistic (SD-SR) and the temperature must be considered explicitly as a physical variable. Whereas the numerical computation of the ordinary Fermi integrals (involved in the zero temperature formalism) seems to be extensively treated in the literature, the generalized ones, which appear in the SD-SR case, are more seldom considered. As far as only pressure and energy are considered, a few approximation formulae are available, but when non-adiabatic evolutions or chemical evolutions are considered, the degeneracy parameter (which is related to the entropy and the chemical potential) remains unknown or is hard to be obtained when necessary from these approximation formulae. Again, the natural way to deal with this parameter is to consider the generalized Fermi integrals.

Section 2 presents these generalized Fermi integrals and the basic equations used to derive the main thermodynamical quantities such as pressure, energy or entropy. In section 3, we review briefly the different existing methods (in an exhaustive way as the author hopes it) with emphasis on two of them. In section 4, we suggest a new method with particular attention to computer implementation. In an appendix, we present the case of the Bose–Einstein distribution whose numerical calculation is very close to that of Fermi–Dirac distribution.

2. Generalized Fermi integrals and related quantities

The generalized Fermi integrals are defined by the following equation,

$$F_k(\eta, \theta) = \int_0^\infty \frac{x^k \sqrt{1 + \frac{1}{2}\theta x}}{e^{-\eta + x} + 1} \, dx, \tag{1}$$

where θ is a non-negative parameter and η a real parameter. For physical applications, the only values to be considered are $k = 1/2$, $3/2$, and $5/2$. The case of electrons is a typical application and in practice θ is related to the temperature by $\theta = kT/m_e c^2$, while η is the degeneracy parameter defined as $\eta = \mu_e^F/kT$, where m_e is the rest mass of the electron, k the Boltzmann constant and μ_e^F is the chemical potential (non-including the rest mass) of the electrons. For $\theta \equiv 0$, $F_k(\eta, \theta)$ reduce to

$$F_k(\eta) = \int_0^\infty \frac{x^k \, dx}{e^{-\eta + x} + 1}, \tag{2}$$

which are the Fermi integrals usually considered in the zero temperature formalism.

In fact, the number density, the pressure and the internal energy of the electron gas are given by [4]:

$$n_e = \frac{8\pi\sqrt{2}}{\lambda_e^3} \theta^{3/2} \left(F_{1/2}(\eta, \theta) + \theta F_{3/2}(\eta, \theta) \right), \tag{3a}$$

$$P_e = \frac{16\pi\sqrt{2}}{3} \frac{m_e c^2}{\lambda_e^3} \times \theta^{5/2} \left(F_{3/2}(\eta, \theta) + \tfrac{1}{2}\theta F_{5/2}(\eta, \theta) \right), \tag{3b}$$

$$U_e = 8\pi\sqrt{2} \frac{m_e c^2}{\lambda_e^3} \theta^{5/2} \left(F_{3/2}(\eta, \theta) + \theta F_{5/2}(\eta, \theta) \right), \tag{3c}$$

where $\lambda_e = h/m_e c$ is the Compton wavelength of the electron.

The entropy per baryon (in units of k) is equal to

$$s_e = \left(\tilde{p}_e + \tfrac{3}{2}\tilde{u}_e - \eta \right), \tag{3d}$$

where $\tilde{p}_e = P_e/n_e kT$ and $\tilde{u}_e = U_e/\frac{3}{2}n_e kT$. These factors are the increments over the related quantities for a classical (perfect) gas.

3. Brief review of different methods

Various authors have already computed the generalized Fermi integrals. Other ones provide approximation formulae for the related thermodynamical quantities. In the first case we can quote the papers of Divine [5], Diaz-Alonso [6], Bonazzola [7], Tooper [8], Bludman and Van Riper [9] and in the second case Service [10] (who treats only the limit $-\eta \gg 1$), Beaudet and Tassoul [11], Eggleton, Faulkner and Flannery [12]. The main characteristics of some of these methods (domain of validity, accuracy, possibility to deal with the degeneracy parameter or the entropy, speed) are summarized in table 1.

Most of them present either approximations with poor accuracy or formulae involving a prohibitive computation time for problems where these integrals must be evaluated hundreds thou-

Table 1
Comparison of the domain of validity, accuracy, possibility to deal with the degeneracy parameter or the entropy and speed for different methods. In the item "Domain of validity" "all" means all values of η and θ. The speed is indicated from + + + for the fastest method to – – for the slowest

Method	Domain of validity	Accuracy	Entropy	Speed
Divine	$\eta \neq 0$	≈ 0.005	yes	+ +
Tooper	all	10^{-15}–10^{-5}	yes	– to +
Bludman and Van Riper	all	0.001–0.1	yes	+
Beaudet and Tassoul	all	0.001–0.1	yes	+ +
Eggleton et al.	all	$\approx 7 \times 10^{-4}$	no	+ +
Eggleton et al.	all	$\approx 7 \times 10^{-4}$	yes	– –
Paczynski	$\eta \approx 0$, $\theta \approx 1$	≈ 0.05	no	+ + +
Service	$-\eta \gg 1$	$\approx 5 \times 10^{-5}$	no	+ +

sands times (as in ref. [3], which describes the tidal compression of a white dwarf in the relativistic gravitational field of a black hole). However, two of these methods may be very useful for tests or short runs and therefore will be discussed in details below.

3.1. The method of Eggleton, Faulkner and Flannery

When the degeneracy factor need not be known and only a relatively poor accuracy is required, the method introduced by Eggleton, Faulkner and Flannery (hereafter EFF, ref. [12]) can be used. The notations used by these authors (viz. $\hat{\rho}$, $\hat{p}$, $\hat{u}$) are related to ours by the following relations:

$$n_e = \frac{8\pi}{\lambda_c^3}\hat{\rho}, \tag{4a}$$

$$P_e = m_e c^2\left(\frac{8\pi}{\lambda_c^3}\right)\hat{p}, \tag{4b}$$

$$U_e = m_e c^2\left(\frac{8\pi}{\lambda_c^3}\right)\hat{u}, \tag{4c}$$

and hence, with their notation, the entropy reads now:

$$s_e = \frac{1}{\mu_e}\left(\frac{1}{\theta}\frac{\hat{u}+\hat{p}}{\hat{\rho}} - \eta\right). \tag{4d}$$

The functions $F_{5/2}$, $F_{3/2}$, and $F_{1/2}$ are related to $\hat{\rho}$, $\hat{p}$, $\hat{u}$ by:

$$F_{5/2}(\eta, \theta) = (\sqrt{2}\,\theta^{7/2})^{-1}(2\hat{u} - 3\hat{p}), \tag{5a}$$

$$F_{3/2}(\eta, \theta) = (\sqrt{2}\,\theta^{5/2})^{-1}(3\hat{p} - \hat{u}), \tag{5b}$$

$$F_{1/2}(\eta, \theta) = (\sqrt{2}\,\theta^{3/2})^{-1}(\hat{\rho} + \hat{u} - 3\hat{p}). \tag{5c}$$

The principle of the EFF method consists, for the computation of a function $F(x, y)$ where x and y are positive real values, in rewriting $F(x, y)$ in the following form,

$$F(x, y) = \tilde{F}(x, y) L_{M,N}[\sigma_F](x, y), \tag{6}$$

where $\tilde{F}(x, y)$ is a simple algebraic function with the same behaviour as $F(x, y)$, when x (or y) tends to 0 or $+\infty$ and

$$L_{M,N}[\sigma_F](x, y) = \frac{\sum_{m=0}^{M}\sum_{n=0}^{N} \sigma_{m,n} x^m y^n}{(1+x)^M (1+y)^N}. \tag{7}$$

As the authors suggest it, $M = N = 4$ can be chosen, whereas

$$\tilde{F}(\eta, \theta) \equiv K(f, g) = \frac{f}{1+f}(g(1+g))^{3/2}, \tag{8}$$

where

$$\eta = 2u + \ln\left(\frac{u-1}{u+1}\right) = 2(u - \operatorname{arctanh}(1/u)), \tag{9}$$

$$u = \sqrt{1+f} \text{ and } g = \theta u. \tag{10}$$

The values of the coefficients $\sigma_{m,n}$ can be found in the quoted paper. These values have been fitted, by mean squares adjustments, in the physically interesting region of the (ρ, T)-plane, where ρ is the density.

The non-universality of the EFF method lies in the difficulty of changing the variables $(f, g) \leftrightarrow (\eta, \theta)$, more precisely in the correspondence $f \leftrightarrow \eta$, since trivially

$$\theta = g/\sqrt{1+f}. \tag{11}$$

If the degeneracy factor does not have to be known, we can deal directly with f instead of η, as suggested by the quoted authors. In this case this method, in spite of its relatively poor accuracy, seems to be the fastest.

Calculation of the degeneracy factor

If we want to determine f from η, the CPU-time increases and, moreover, we must take care of the accuracy of the result, in view of the round off errors for instance.

First, let us remark that when f tends towards 0, η tends to minus infinity according to

$$\eta \approx \ln f + 2(1 - \ln 2) + \tfrac{1}{2}f - \tfrac{3}{16}f^2 + \mathcal{O}(f^3), \tag{12}$$

and in the same way, when f (and η) tends to infinity,

$$\tfrac{1}{2} f^{-1/2}\eta \approx 1 - \tfrac{1}{2} f^{-1} + \tfrac{1}{24} f^{-2} + \tfrac{3}{16} f^{-3} - \tfrac{45}{128} f^{-4} + \mathcal{O}(f^{-5}). \tag{13}$$

Hence, when η tends towards $-\infty$, in practice when $\eta \leq -20$, we can write

$$f \approx 4\,\exp(\eta - 2), \tag{14}$$

when η tends towards $+\infty$, in practice when $\eta \geq 2000$, we can write

$$f \approx \left(\tfrac{1}{2}\eta\right)^2, \tag{15}$$

and in the other cases, let us put x as an auxiliary variable

$$x = \ln(u - 1) = \ln\left(\sqrt{1 + f} - 1\right). \tag{16}$$

Then the equation in f becomes

$$F(x) = 2\,\mathrm{e}^{x} + 2 - \ln(1 + 2\,\mathrm{e}^{-x}) = \eta. \tag{17}$$

The dummy variable x, previously introduced, allows us to solve easily this new equation by the Newton method. In fact, the $(x \leftrightarrow f)$-change of variables removes the stiff behaviour of f in the interval under consideration. Using the Newton method, let us remark that

$$F'(x) = 2\left(\mathrm{e}^{x} + \frac{1}{\mathrm{e}^{x} + 2}\right) = 4\frac{1 + \cosh x}{1 + 2e^{-x}}, \tag{18}$$

and we can choose as guess value of x, the following value x_0

$$x_0 = \begin{cases} 0.5(\eta - 3) & \text{for } \eta \leq 0, \\ 0.5(2\ln(1 + \eta) - 3) & \text{for } \eta \geq 0. \end{cases} \tag{19}$$

When x is known, f and g can be computed easily without round off errors, because x performs a good variable scaling in the region of interest, by

$$\begin{cases} f = u^2 - 1 = \mathrm{e}^{x}(\mathrm{e}^{x} + 2), \\ g = \theta u = \theta(1 + \mathrm{e}^{x}). \end{cases} \tag{20}$$

3.2. The approximation of Paczynski

B. Paczynski [13] provided an elegant approximation formula applicable in the region of semi-degeneracy.

First, the pressure is computed in three cases. In each one, the electron gas is considered as forming:

(1) a pure perfect non-degenerate gas, the corresponding pressure is noted $P_{\mathrm{e,nd}}$;

(2) a pure entirely degenerate non-relativistic gas, the corresponding pressure is noted $P_{\mathrm{e,dnr}}$;

(3) a pure entirely degenerate ultra-relativistic gas, the corresponding pressure is noted $P_{\mathrm{e,dr}}$.

The resulting pressure is then

$$P_{\mathrm{e}} \equiv \left(P_{\mathrm{e,d}}^2 + P_{\mathrm{e,nd}}^2\right)^{1/2} = \frac{m_{\mathrm{e}}c^2}{\mu_{\mathrm{e}}}\sqrt{\theta^2 + \frac{y^4}{25 + 16y^2}}, \tag{21}$$

where

$$P_{\mathrm{e,d}} \equiv \left(P_{\mathrm{e,dnr}}^{-2} + P_{\mathrm{e,dr}}^{-2}\right)^{-1/2} = \frac{m_{\mathrm{e}}c^2}{\mu_{\mathrm{e}}}\frac{y^2}{\sqrt{25 + 16y^2}} \tag{22}$$

can be interpreted as the pressure of a pure completely degenerate but semi-relativistic gas of electrons. The variable y is related to η and θ by $1 + y^2 = (1 + \theta\eta)^2$ and, for physical interpretation, y is (sometimes) called the relativistic parameter, since it satisfies the equality

$$1 + y^2 = \left(1 + \frac{\mu_{\mathrm{e}}^{\mathrm{F}}}{m_{\mathrm{e}}c^2}\right)^2. \tag{23}$$

Let us recall the well known expression for $P_{\mathrm{e,nd}}$, $P_{\mathrm{e,dnr}}$ and $P_{\mathrm{e,dr}}$ (see ref. [4]).

$$P_{\mathrm{e,nd}} = \frac{1}{\mu_{\mathrm{e}}}kT, \tag{24a}$$

$$P_{\mathrm{e,dnr}} = \frac{m_{\mathrm{e}}c^2}{\mu_{\mathrm{e}}}\frac{1}{5}y^2, \tag{24b}$$

$$P_{\mathrm{e,dr}} = \frac{m_{\mathrm{e}}c^2}{\mu_{\mathrm{e}}}\frac{1}{4}y, \tag{24c}$$

where μ_{e} is the mean molecular weight by electrons. Here, the pressures are expressed per nucleon (i.e. the pressures are divided by the density of nucleons, which is equal to $\mathcal{N}\rho$ when the density ρ is expressed in g cm^{-3} and $\mathcal{N}$ is the Avogadro number).

This approximation provides a good description for the pressure (and the energy) as a function of the temperature (θ). But the overall accuracy is poor (see table 1) and hence this formula seems to be useful only for numerical tests.

4. The suggested method

4.1. The case of a negative degeneracy factor

Introducing $\tilde{K}_\nu(x) = e^x K_\nu(x)$, K_ν being the modified Bessel functions, the formulae presented by Cox and Giuli [4] are well applicable. We have then:

$$F_{1/2}(\eta, \theta) = \frac{1}{\sqrt{2\theta}} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} e^{n\eta} \tilde{K}_1\left(\frac{n}{\theta}\right), \tag{25a}$$

$$F_{3/2}(\eta, \theta) = \frac{1}{\sqrt{2\theta^3}} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} e^{n\eta} \times\left(\tilde{K}_2\left(\frac{n}{\theta}\right) - \tilde{K}_1\left(\frac{n}{\theta}\right)\right), \tag{25b}$$

$$F_{5/2}(\eta, \theta) = \frac{1}{\sqrt{2\theta^5}} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} e^{n\eta} \times\left\{2\left(\tilde{K}_1\left(\frac{n}{\theta}\right) - \tilde{K}_2\left(\frac{n}{\theta}\right)\right) + \frac{3\theta}{n}\tilde{K}_2\left(\frac{n}{\theta}\right)\right\}. \tag{25c}$$

Only $\tilde{K}_0$(x) and $\tilde{K}_1$(x) must be evaluated, which can be done quite easily at any precision by polynomial expansion, since $\tilde{K}_2(x) = \tilde{K}_0(x) + (2/x)\tilde{K}_1(x)$.

The series expansions (25a)–(25c) are convergent for all positive (non-zero) values of θ and for all negative (even zero) values of η, but the summation must be truncated at some value of $N(\epsilon)$, depending on the desired accuracy ϵ. We have found, by several numerical trials, that

$$N(\epsilon) = \operatorname{int}\left(\frac{-2\log(\epsilon) + 3}{|\eta|}\right). \tag{26}$$

We can remark that our expression generalizes the one by Tooper [8] who gives only the value for $\epsilon = 10^{-15}$. However, for reasons of calculation speed, it is convenient to use this method for values of η sufficiently negative, say $\eta \leq -1$. Hence, for an accuracy of $\epsilon = 10^{-6}$ there are less than 15 terms to sum.

The functions $\tilde{K}_0(x)$ and $\tilde{K}_1(x)$ are computed by means of expansions in terms of Chebyshev polynomials. We have used the NAG routines S18CCF and S18CDF (see ref. [14]), respectively, in which the various expansions have been truncated. With the goal of reducing, as much as possible, the computation-time, only those terms which are needed to guarantee an accuracy of $10^{-2}\epsilon$ have been kept

4.2. The case of a great degeneracy

For $\eta > 0$ there exists, to may knowledge, no exact formula for these integrals, but only asymptotic expansions for which the rest is in the order to $e^{-\eta}$ (see ref. [4], eqs. (24.179), (24.180) and (24.181), p. 826):

$$F_{1/2}(\eta, \theta) = \frac{1}{\sqrt{2\theta^3}}\left(f_{1/2}(y) + (1+\eta\theta) \times\left(C_1 + C_2 + C_3(4y^2+7)\right)\right) + \mathcal{O}(e^{-\eta}), \tag{27a}$$

$$F_{3/2}(\eta, \theta) = \frac{1}{\sqrt{2\theta^5}}\left(f_{3/2}(y) + C_1(3+2\eta\theta) - C_2 - C_3((4\eta\theta+6)\eta\theta+3)\right) + \mathcal{O}(e^{-\eta}), \tag{27b}$$

$$F_{5/2}(\eta, \theta) = \frac{1}{\sqrt{2\theta^7}}\left(f_{5/2}(y) + C_1(5+\eta\theta) + C_2(((2\eta\theta+10)\eta\theta+15)\eta\theta+5) + C_3(3+5\eta\theta)\right) + \mathcal{O}(e^{-\eta}), \tag{27c}$$

where

$$C_1 = \frac{\pi^2}{6}\frac{\theta^2}{y}, \quad C_2 = \frac{7}{20}C_1\left(\frac{\pi\theta}{y^2}\right)^2, \quad C_3 = \frac{31}{168}C_1\left(\frac{\pi\theta}{y^2}\right)^4, \tag{28}$$

and

$$f_{1/2}(y)=\tfrac{1}{2}\left(y\sqrt{1+y^2}-\operatorname{arcsinh}(y)\right), \tag{29a}$$

$$f_{3/2}(y)=\tfrac{1}{3}y^3-f_{1/2}(y), \tag{29b}$$

$$f_{5/2}(y)=\tfrac{5}{8}y\left(1+\tfrac{2}{5}y^2\right)\sqrt{1+y^2}-\tfrac{2}{3}y^3 - \tfrac{5}{8}\operatorname{arcsinh}(y). \tag{29c}$$

The relativistic parameter y has been defined previously (see eq. (23)) but, we can remark, that it can be also defined by $y=p_0/mc$ where p_0 is the Fermi momentum (see ref. [4], p. 825).

By several numerical tests, we have found that, given an accuracy ϵ, these asymptotic formulae hold (and hence produce the desired result with the wanted accuracy) for

$$\eta \geq \eta_0(\epsilon) = -\ln(0.2\epsilon). \tag{30}$$

In other words, the remainder in $\mathcal{O}(e^{-\eta})$ of these asymptotic expansions has been estimated to be approximately $5e^{-\eta}$.

For values of y greater than 0.1 the above expressions can be used but it is more convenient for numerical computation to rewrite the formula of $f_{5/2}(y)$ in the form of

$$f_{5/2}(y)=\tfrac{1}{24}y(((6\eta\theta+2)\eta\theta-5)\eta\theta+15) - \tfrac{5}{8}\ln(1+y+\eta\theta), \tag{31}$$

and we have always used, when possible, the Horner scheme for the evaluation of polynomials.

When y is small, it is better to compute the $f_k(y)$ with their power series expansions (up to the 19^{th} order) owing to cancellation and round off errors. Then, these new expressions were rewritten according the Horner scheme, which is the more suitable for computational purposes.

$$f_{1/2}(y)=\left(\left(\left(\left(\left(\left(\left(-\tfrac{58773}{1114112}y^2+\tfrac{77}{5120}\right)y^2 - \tfrac{63}{3328}\right)y^2+\tfrac{35}{1408}\right)y^2-\tfrac{5}{144}\right)y^2 + \tfrac{3}{56}\right)y^2-\tfrac{1}{10}\right)y^2+\tfrac{1}{3}\right)y^3+\mathcal{O}(y^{19}), \tag{32a}$$

$$f_{3/2}(y)=\left(\left(\left(\left(\left(\left(\tfrac{58773}{1114112}y^2-\tfrac{77}{5120}\right)y^2+\tfrac{63}{3328}\right)y^2 - \tfrac{35}{1408}\right)y^2+\tfrac{5}{144}\right)y^2-\tfrac{3}{56}\right)y^2+\tfrac{1}{10}\right)y^5 +\mathcal{O}(y^{19}), \tag{32b}$$

$$f_{5/2}(y)=\left(\left(\left(\left(\left(-\tfrac{275913}{4456448}y^2+\tfrac{7}{512}\right)y^2-\tfrac{7}{416}\right)y^2 + \tfrac{15}{704}\right)y^2-\tfrac{1}{36}\right)y^2+\tfrac{1}{28}\right)y^7+\mathcal{O}(y^{19}). \tag{32c}$$

With respect to the evaluation of the relative magnitude of each term in the previous power series expansions, we present in table 2, for the two limiting cases already mentioned, the behaviour of the different involved terms.

At this point, we can point out some mistakes in the appendix of the paper of Edwards and Merilan [15], who present an algorithm to derive additional terms for the power series expansions. The last terms in their eq. (A.38a)–(A.38c) correspond neither to their previous formulation, or to the exact formulae, which we can found elsewhere (e.g. eqs. (27) or ref. [4], p. 826).

Table 2
Order of magnitude of the different terms in the asymptotic expansions of $F_k(\eta, \theta)$ for $\eta \gg 1$

	Dominant term i.e. $f_k(y)$	1st corr. term	2nd corr. term	3rd corr. term
Relativistic limit ($y \gg 1$), $y \approx \eta\theta$				
$\theta^{3/2}F_{1/2}(\eta,\theta)$	y^2	$\eta^{-2}y^2$	η^{-4}	$\eta^{-6}y^{-2}$
$\theta^{5/2}F_{3/2}(\eta,\theta)$	y^3	$\eta^{-2}y^3$	η^{-4}	$\eta^{-6}y^{-1}$
$\theta^{7/2}F_{5/2}(\eta,\theta)$	y^4	$\eta^{-2}y^4$	$\eta^{-6}y^4$	η^{-6}
Non-relativistic limit ($y \ll 1$), $y^2 \approx 2\eta\theta$				
$F_k(\eta,\theta)$	η^{k+1}	η^{k-1}	η^{k-3}	η^{k-5}

4.3. Intermediate degeneracy

For intermediate values of η, i.e. $-1 \leq \eta \leq \eta_0(\epsilon)$, we must perform a direct quadrature of the integrand $(x^k\sqrt{1+\frac{1}{2}\theta x})/(e^{-\eta+x}+1)$ over $[0, +\infty[$. In view of the requested accuracy, there are some possibilities:

– We can use a quadrature subroutine which warrants the result to the desired accuracy (e.g. the D01AMF subroutine of the NAG library [14]). Although very CPU-time consuming, this seems the only possibility for very accurate computation (say $\epsilon \leq 10^{-10}$).

Table 3
Abscissas and weights for the Gauss–Jacobi–Legendre 24 points quadrature formulae

	Abscissas	Weights
1) for the computation of $F_{1/2}$ $(k=1/2)$		
1	0.02849181901211425612	8.708587123952392133E-03
2	0.1140823953415001514	2.588239336080700032E-02
3	0.2571196607654419903	3.544909376759483930E-02
4	0.4581922214392973504	3.136491522445054680E-02
5	0.7181429425482090201	1.988436558407921550E-02
6	1.038089029843696041	9.434998721307272737E-03
7	1.419449804484678023	3.419850049429588749E-03
8	1.863983925481609230	9.555329428130830533E-04
9	2.373838585576534944	2.062955559902515674E-04
10	2.951614313169913921	3.433654074066855722E-05
11	3.600450651741771930	4.379600444913362890E-06
12	4.324140488959726354	4.240385407038400264E-07
13	5.127284742831941503	3.076018284902174265E-08
14	6.015505513813138140	1.643383533457525834E-09
15	6.995746620354544013	6.325185875970369282E-11
16	8.076709482470633067	1.704842486377056832E-12
17	9.269507611292598112	3.101817790525326374E-14
18	10.58869241385666403	3.629077676477144465E-16
19	12.05395014429822367	2.556163240566112559E-18
20	13.69311182128868830	9.873226804757046188E-21
21	15.54801556261556146	1.817864820219574227E-23
22	17.68754663804027985	1.268090348112044032E-26
23	20.24330437352278005	2.169210449373242921E-30
24	23.55302923725045460	3.039858528529206592E-35
2) for the computation of $F_{3/2}$ $(k=3/2)$		
1	0.05715903814507704487	9.391039109364949266E-04
2	0.1691691337466636962	5.395060892756637293E-03
3	0.3376961274854422335	1.185604647175828626E-02
4	0.5634281922970708672	1.498727911832798787E-02
5	0.8472862423305308693	1.259200700140959199E-02
6	1.190457155906641785	7.517817834723649452E-03
7	1.594424264068745150	3.301889289393359145E-03
8	2.061005946067815622	1.086597705172058257E-03
9	2.592405877884360298	2.702708661455162480E-04
10	3.191278791656152039	5.092566771087720233E-05
11	3.860817110306774041	7.249687936498855808E-06
12	4.604866300433345506	7.743310993978272513E-07
13	5.428080731300453654	6.136643895030836018E-08
14	6.336138259586000403	3.552722706759996572E-09
15	7.336042625742008587	1.471575290929083610E-10
16	8.436561900678225587	4.243638727023408471E-12
17	9.648886714786056247	8.219691994616413877E-14
18	10.98766185128874102	1.019514801600291550E-15
19	12.47269276188876817	7.586099982823628041E-18
20	14.13197255778337397	3.086650576946254470E-20
21	16.00757798316519261	5.973965114759314238E-23
22	18.16878543402941250	4.375012580807258086E-26
23	20.74794941948531924	7.857961729842841089E-30
24	24.08479843708068601	1.160069244427643163E-34

Table 3 (continued)

	Abscissas	Weights
3) for the computation of $F_{5/2}$ $(k=5/2)$		
1	0.09226964901078394532	2.071063088505366675E-04
2	0.2301330622033462354	1.877822582532298638E-03
3	0.4234062206866813589	5.865300612172511508E-03
4	0.6730526278462398295	9.802948061231375216E-03
5	0.9801047549043842122	1.034976046218783665E-02
6	1.345826430906904518	7.481224655789864808E-03
7	1.771765191059100042	3.867716232737840024E-03
8	2.259797098310934291	1.465702083140081271E-03
9	2.812179586536520288	4.125482564509925770E-04
10	3.431618732467618560	8.672782534952122686E-05
11	4.121357053976206369	1.361516047122477181E-05
12	4.885289958806278508	1.588151474110357971E-06
13	5.728122789858624402	1.363348855352401851E-07
14	6.655586828308871788	8.490582412567194274E-09
15	7.674743520887705036	3.760904155849678447E-10
16	8.794425447353131225	1.153952714999818400E-11
17	10.02589823157352428	2.367942527175377107E-13
18	11.38389784186368062	3.100161110168569149E-15
19	12.88834654509108268	2.427473024103134183E-17
20	14.56739675567256287	1.036799160103740114E-19
21	16.46336023193037544	2.102542345866995361E-22
22	18.64589912658745909	1.611719304087483749E-25
23	21.24811339403773192	3.030805921094216037E-29
24	24.61169463440596677	4.699997859827737074E-34

– An other possibility lies in the use of various Gaussian quadratures, for example the Gauss–Laguerre quadrature formula viz. [16],

$$\int_0^{\infty} x^k \, e^{-ax} g(x) \, dx \equiv \sum_{i=1}^{n} w_i(k, a) g(x_i), \qquad (33)$$

where w_i and x_i are, respectively, the weights and the abscissas of the (Gaussian) quadrature formula. Here, we have

$$g(x) = \frac{\sqrt{1 + \frac{1}{2}\theta x}}{e^{-\eta-(a-1)x} + e^{-ax}}.$$

The parameter a (not necessary equal to 1.) is adjusted to obtain the best result. In fact, the exponentials which lie in the denominator of $g(x)$ must be computed without underflow or overflow.

Let us denote by $x_{\lim}$ the greatest real value for which the computation of $\exp(-x_{\lim})$ is permitted by the computer architecture and differs from zero (for a VAX computer $x_{\lim}$ is about equal to 88.7). Let us denote also by x_0 the greatest abscissa used in the Gaussian quadrature formula for $a=1$. This later value depends only upon the value of n, the number of points (abscissas) involved in the summation. When we change the value of a, the abscissas x_i become x_i/a, the weights w_i become w_i/a^{k+1}, but the products ax_i remain constant.

For all the values of a, the following two conditions are to be satisfied:

$$\eta_0(\epsilon)+(a-1)\frac{x_0}{a}\leq x_{\lim}, \tag{34a}$$

$$x_0\leq x_{\lim}. \tag{34b}$$

The second condition is not fulfilled if the number of points in the quadrature formula is greater than $n=24$. In view of the dependance of the final accuracy upon the number of points used, the accuracy ϵ is greater than $\epsilon=10^{-6}$ when the number of points n is less than $n=24$. Then, only the first condition must be satisfied; it is equivalent to

$$a\leq\frac{x_0}{\eta_0+(x_0-x_{\lim})}. \tag{35}$$

Numerical trials have shown that for $\epsilon=10^{-6}$ ($\eta_0=15$) the optimal value for a is $a=3.5$ for $n=24$ or $a=3$ for $n=30$.

In table 3, we present the abscissas and the weight of a 24 points Gauss–Laguerre quadrature formula with a value of a equal to 3.5 for the three cases under consideration here, namely $k=1/2$, $3/2$ and $5/2$. These weights and abscissas have been computed with the routine D01BCF of the NAG library [14]. But, with the intention of obtaining highest accuracy, the different involved subroutines have been adapted to quadruple precision (H floating of the Vax Fortran with about 33 significant decimal digits). This adaptation has been realized straightforwardly and quickly since, in our case, only the values at 1/2, 3/2 and 5/2 of the Gamma function (computed by the NAG-routine S14AAF) are involved and hence their values are exactly known, which allows us to bypass the extension of the polynomial expansions involved in this subroutine. Therefore, table 3 presents these results with only 19 decimal digits which then can be seen as all significant.

4.4. The problems of discontinuities

For the most part of the presented methods, there are always discontinuities in the manner to compute the integrals. Therefore, at a discontinuity point, noted η_c, there is no reason that the two methods involved (the first for $\eta<\eta_c$ and the second for $\eta>\eta_c$) give exactly the same results (but, for the desired accuracy, the two results must be the same). This small discontinuity may induce some troubles (for example, when used in a program which solves a differential equation). This technical point has been already pointed out by Tooper [8] in the framework of hydrodynamical calculations.

In order to eliminate this trouble, it is preferable to interpolate in the neighbourhood of the discontinuity η_c. Let us the notation

$$\begin{cases} f_-(\eta) & \text{for } \eta<\eta_c, \\ f_+(\eta) & \text{for } \eta>\eta_c, \end{cases} \tag{36}$$

for η_c in the range $|\eta-\eta_c|\leq\Delta\eta$. For example, we can choose $\Delta\eta=1$. The simplest approach is linear interpolation,

$$f(\eta)=f_+(\eta)h_1(x)+f_-(\eta)(1-h_1(x)), \tag{37}$$

where

$$x=\frac{\eta-\eta_c}{\Delta\eta}$$

and (38)

$$h_1(x)=\begin{cases} 0 & \text{for } x\leq -1, \\ \frac{1}{2}(1+x) & \text{for } |x|\leq 1, \\ 1 & \text{for } x\geq 1. \end{cases}$$

In spite of the possibility of building interpolation functions of all C^k classes, it seems that linear interpolation is sufficient to avoid the trouble of discontinuities in the applications described in this article. Moreover, it is the simplest and consequently the fastest.

5. Conclusions

We think that the method presented here represents the best compromise between accuracy and computational speed for the evaluation of the equation of state of a semi-degenerate semi-relativistic gas at non-zero temperature for an accuracy in the range $\epsilon \approx 10^{-5}$ to $\epsilon \approx 10^{-7}$ and in a CPU-time less than 2 ms (in the worst case) on a VAX 8600 (for which a standard Fortran function such as DSQRT requires 106 μs and DEXP 71.5 μs, the CPU-speed being 0.79 Mflps in double precision).

Moreover, the correction due to the non-zero temperature formalism in the equation of state is far from being a negligible correction, in comparison to other possible corrections, such as the Coulomb or the Thomas–Fermi electrostatic corrections. For instance, even in the case of the sun's interior, the pressure due to the electronic contribution is slightly modified when calculated with a semi-degenerate non-zero temperature electronic gas. In fact, in the case ($\theta \ll 1$ and $-\eta \gg 1$), we have the relation

$$\left(\frac{P_e}{\rho}\right)_{T\neq 0} \approx \left(\frac{P_e}{\rho}\right)_{T=0}\left(1+\frac{3}{4}\theta\right). \tag{39}$$

In the center of the sun, θ is nearly equal to 3×10^{-3} and therefore, for the same pressure, the temperature will be 2×10^{-3} lower.

Note: During the numerical tests of the different methods presented here, we have discovered some misprints in the tabulations of the generalized Fermi-integrals presented in the annex of the book of Cox and Giuli [4]. These are:
for $\eta = 15$ and for $\log\beta = -6.0$ we must read $F_{1/2} = 3.8943(+1)$,
for $\eta = 15$ and for $\log\beta = -4.0$ we must read $F_{3/2} = 3.5822(+1)$,
for $\eta = 40$ and for $\log\beta = -0.5$ we must read $F_{1/2} = 3.6348(+2)$,
for $\eta = 50$ and for $\log\beta = -1.0$ we must read $F_{1/2} = 3.6953(+2)$.

Acknowledgements

I am grateful to Jean-Pierre Luminet and Remi Hakim for precious advice. It is also a great pleasure to thank the anonymous referee for all his valuable comments and for bringing to my attention, among other things, the reference [16].

Appendix. Application to a gas of bosons

By analogy with the preceding case of Fermi–Dirac distributions, it is easy to treat the case of Bose–Einstein distributions at any temperature. In this paragraph, all the quantities presented are related to non-condensed particles. This is recalled, in the notations, by the use of "primes" (for instance, the density n' is equal to $(1-\zeta)n$ where ζ is the fraction of condensed particles and n is the total number density [17].

We have

$$n' = C\theta^{3/2}\left(G_{1/2}(\eta,\theta)+\theta G_{3/2}(\eta,\theta)\right), \tag{40a}$$

$$p' = \tfrac{2}{3}Cmc^2\theta^{5/2}\left(G_{3/2}(\eta,\theta)+\tfrac{1}{2}\theta G_{5/2}(\eta,\theta)\right), \tag{40b}$$

$$u' = Cmc^2\theta^{5/2}\left(G_{3/2}(\eta,\theta)+\theta G_{5/2}(\eta,\theta)\right), \tag{40c}$$

where

$$C=\frac{4\pi g\sqrt{2}}{\lambda_c^3},\quad \lambda_c=\frac{h}{mc},\quad \theta=\frac{kT}{mc^2}, \tag{41}$$

and

$$G_k(\eta,\theta)=\int_0^\infty \frac{x^k\sqrt{1+\frac{1}{2}\theta x}}{e^{-\eta+x}-1}\,dx. \tag{42}$$

Here, g is the degeneracy factor ($g = 1$ for scalar bosons) and m is the mass of the bosons.

The evaluation of these integrals is simpler that in the fermion case because, for bosons, we have always $\eta \le 0$ [17]. With this restriction, we can write in the same way as in eqs. (25a)–(25c):

$$G_{1/2}(\eta,\theta)=\frac{1}{\sqrt{2\theta}}\sum_{n=1}^{\infty}\frac{1}{n}e^{n\eta}\tilde{K}_1\left(\frac{n}{\theta}\right), \tag{43a}$$

$$G_{3/2}(\eta,\theta)=\frac{1}{\sqrt{2\theta^3}}\sum_{n=1}^{\infty}\frac{1}{n}e^{n\eta}\left(\tilde{K}_2\left(\frac{n}{\theta}\right)-\tilde{K}_1\left(\frac{n}{\theta}\right)\right), \tag{43b}$$

$$G_{5/2}(\eta, \theta) = \frac{1}{\sqrt{2\theta^5}} \sum_{n=1}^{\infty} \frac{2}{n} e^{n\eta} \left(\tilde{K}_1\left(\frac{n}{\theta}\right) - \tilde{K}_2\left(\frac{n}{\theta}\right) + \frac{3}{2}\frac{\theta}{n} \tilde{K}_2\left(\frac{n}{\theta}\right)\right). \quad (43c)$$

These expansions are always convergent for any values of η ($\eta \leq 0$) and the different integrals can be evaluated, at any accuracy, with the same algorithm as the Fermi integrals.

Moreover, in the bosonic case, there is a very important and interesting limiting case if the boson mass tends towards zero. The relevant physical applications are, for example, gluons and, especially, photons (with a zero chemical potential). Therefore, we can write

$$\text{For } \theta \to +\infty, \quad G_k(\eta, \theta) \to \sqrt{\frac{\theta}{2}}\, G_{k+1/2}(\eta), \quad (44)$$

where

$$G_n(\eta) = \int_0^{\infty} \frac{x^n}{e^{-\eta+x} - 1} dx, \quad (45)$$

and for $\eta = 0$ (e.g. for photons),

$$G_n(0) = \Gamma(n+1)\zeta(n+1), \quad (46)$$

$$G_1(0) = \frac{\pi^2}{6} = 1.64493406684822643647\ldots,$$

$$G_2(0) = 2.40411380631918857080\ldots, \quad (47)$$

$$G_3(0) = \frac{\pi^4}{15} = 6.49393940226682914909\ldots .$$

In this case, we have

$$n' = \tilde{C}T^3 G_2(\eta), \quad (48a)$$

$$p' = \tilde{C}kT^4 G_3(\eta), \quad (48b)$$

$$u' = 3p', \quad (48c)$$

where

$$\tilde{C} = 4\pi g \left(\frac{k}{hc}\right)^3. \quad (49)$$

Thus, we recover the black body radiation formulae.

Let us point out that in the quoted textbooks [17] we find the functions f_k and g_k, which are related to F_k ($F_k(x) \equiv F_k(x, \theta \equiv 0)$) and G_k by the following relations:

$$f_k(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1} x^n}{n^k} = \frac{1}{\Gamma(k+1)} F_k(x), \quad (50)$$

$$g_k(x) = \sum_{n=1}^{\infty} \frac{x^n}{n^k} = \frac{1}{\Gamma(k+1)} G_k(x). \quad (51)$$

References

[1] P.W. Cole and R. G. Deupree, Astrophys. J. 239 (1980) 284; Astrophys. J. 247 (1981) 607.
A.C. Edwards, Mon. Not. R. Astron. Soc. 146 (1969) 445, Appendix.

[2] S. Bonazzola, J.-M. Hameury, J. Heyvaerts and J.P. Lasota, Astron. Astrophys. 136 (1984) 89.

[3] J.-P. Luminet and B. Pichon, Astron. Astrophys. 209 (1989) 103. See also: J.-P. Luminet and B. Pichon, Astron. Astrophys. 209 (1989) 85.

[4] J.P. Cox and R.T. Giuli, Principles of Stellar Structure, Vol. 2: Applications to Stars (Gordon and Breach, London, 1968).

[5] N. Divine, Astrophys. J. 142 (1965) 1652.

[6] J. Diaz-Alonso, private communication.

[7] S. Bonazzola, private communication.

[8] R.F. Tooper, Astrophys. J. 156 (1969) 1075.

[9] S.A. Bludman and K.A. Van Riper, Astrophys. J. 212 (1977) 859.

[10] A.T. Service, Astrophys. J. 307 (1986) 60.

[11] G. Beaudet and M. Tassoul, Astron. Astrophys. 13 (1971) 209.

[12] P.P. Eggleton, J. Faulkner and B.P. Flannery, Astron. Astrophys. 23 (1973) 325.

[13] B. Paczynski, Astrophys. J. 267 (1983) 315.

[14] NAG library (Version Mk13, July 1983) available at the Numerical Algorithm Group Ltd. (Wilkinson house, Jordan Hill Road, Oxford OX2 8DR United Kingdom). For the computation of abscissas and weights of Gauss–Laguerre quadrature formula, see also a recent paper of T. Takemasa in Comput. Phys. Commun. 52 (1988) 133.

[15] T.W. Edwards and M.P. Merilan, Astrophys. J. 244 (1981) 600.

[16] P.J. Davis and P. Rabinowitz, Methods of Numerical Integration, 2nd ed. (Academic, New York, 1984).

[17] K. Huang, Statistical Mechanics (Wiley, New York, 1963).
L.E. Reichl, A Modern Course in Statistical Physics (Arnold, 1983).