Diffusion approximation revisited

Manabu Machida,[1,*] George Yu. Panasyuk,[1] John C. Schotland,[1] and Vadim A. Markel[2]

[1]*Department of Bioengineering, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA*
[2]*Departments of Radiology and Bioengineering, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA*
**Corresponding author: mmachida@seas.upenn.edu*

Received January 29, 2009; revised March 17, 2009; accepted March 19, 2009; posted March 23, 2009 (Doc. ID 106888); published April 30, 2009

We study the diffusion approximation (DA) to the radiative transport equation (RTE) in infinite homogeneous space. Different definitions of the reduced intensity I_r that satisy a simplified RTE (without accounting for scattering) and that are often used in the derivation of the DA are examined. By comparing the results of the DA with exact solutions to the RTE, we come to the conclusion that the best accuracy in the DA is achieved if we choose the definition of the reduced intensity (from a family of possible definitions) that results in $I_r=0$. Thus, the separation of the specific intensity into reduced and diffuse components is unnecessary. We also discuss the conditions under which the DA is applicable. © 2009 Optical Society of America

OCIS codes: 030.5620, 290.1990, 170.5280.

1. INTRODUCTION

The theoretical analysis of problems involving multiple scattering of waves is frequently based on the radiative transport equation (RTE) [1–4]. In many practical applications, the diffusion approximation (DA) to the RTE can be used. The advantage of the DA is its relative mathematical simplicity. It is of particular interest in biomedical imaging of tissues with multiply scattered light [5–7]. The diffusion equation (DE) depends on a number of parameters and functions that are inherited from the RTE. These include the diffusion and absorption coefficients, the extrapolation distance (which appears in the boundary conditions) and the source function. While the definition of the absorption coefficient is quite straightforward, a significant effort has been devoted to obtaining an accurate expression that relates the diffusion coefficient to the parameters of the RTE [8–13]. The boundary conditions for the DE and the extrapolation distance have been considered in detail [14–16]. The source function of the the DE has also been discussed in [17–20]. The optimal shape of the source function was deduced by comparison of the analytical solution to the DE with experimental measurements [17–19] or to the results of Monte Carlo simulations [20]. However, the question was not studied systematically from the theoretical point of view. The main purpose of this paper is to address this gap.

In the transport theory, the primary physical quantity of interest is the specific intensity $I(\mathbf{r},\hat{\mathbf{s}})$ where $\mathbf{r}$ is the position and the unit vector $\hat{\mathbf{s}}$ specifies the direction in which light propagates. In the DA one is interested only in the angularly averaged quantities $u(\mathbf{r})$ and $\mathbf{J}(\mathbf{r})$ [defined below in Eqs. (6) and (7)], which have the physical meaning of electromagnetic energy density and current. Correspondingly, the mathematical form of the source term in the RTE is markedly different from that in the DE. Namely, in the case of RTE, the source term is a function of both $\mathbf{r}$ and $\hat{\mathbf{s}}$ and can be written as $\epsilon(\mathbf{r},\hat{\mathbf{s}})$. For example, a point unidirectional source of the form

$$\epsilon(\mathbf{r},\hat{\mathbf{s}}) = \mathcal{A}\delta(\mathbf{r})\delta(\hat{\mathbf{s}}-\hat{\mathbf{s}}_0), \tag{1}$$

describes a narrow beam of light of total power $\mathcal{A}$ injected into a scattering medium in the direction of $\hat{\mathbf{s}}_0$. Without loss of generality, we assume that the injection point coincides with the origin of the laboratory frame. In the case of the DE, the source term is a function of position only and is denoted by $S(\mathbf{r})$. Obviously, a transition from the RTE with a point unidirectional source of the form (1) to the DE requires that $S(\mathbf{r})$ cannot be spherically symmetric. This deviation from spherical symmetry is required to capture the dependence of $\epsilon(\mathbf{r},\hat{\mathbf{s}})$ on the direction $\hat{\mathbf{s}}$.

The above fact has been widely recognized in the literature. The frequently accepted approach is to use the source function for the DE in the form of a "dipole." For example, in [17], the source was assumed to be constant inside a half-sphere. In this geometry, the source is characterized not only by location but also by direction. In [18,19], the problem was considered in which a narrow collimated beam is injected into a highly scattering medium from vacuum. The source function was modeled in these references as a point that is displaced from the boundary into the medium by a distance Δx, which must be determined experimentally. In [20], the DA (in the 1D geometry) was interpreted as the P_1 approximation in which the higher angular moments of the RTE source function (1) appear naturally. Yet another commonly used alternative is to assume that $S(\mathbf{r})$ is an exponentially decaying function along the ray that points from the origin in the direction $\hat{\mathbf{s}}_0$ and zero everywhere else [3]. This form of $S(\mathbf{r})$ can be obtained by decomposing the specific intensity into its reduced and diffuse components [3,21] $I_r(\mathbf{r},\hat{\mathbf{s}})$ and $I_d(\mathbf{r},\hat{\mathbf{s}})$, respectively. The DA is then made for the diffuse component. However, this approach contains a step that is quite arbitrary. We will see below that a family of different DAs can be obtained by using different definitions of I_r. The correct DA must be chosen by comparing the solutions to the DE to those of the RTE at large distances from the source. We will show that the most logical choice is $I_r=0$; correspondingly, the DA must be applied to the total specific intensity. Thus, we find that introduction of the reduced intensity in the context of the DA is unnecessary.

This paper is organized as follows. In Section 2 we briefly review the RTE and define the density and current of electromagnetic energy. In Section 3, we use the RTE to derive asymptotic expressions for these two quantities. In Section 4, we propose a systematic method for deriving the DA. The method accounts for the ambiguity in the definition of I_r by introducing a new adjustable parameter that determines the rate of exponential decay of the source function on the ray defined above. In Section 5 we compare the density of electromagnetic radiation computed from the RTE and from the DE. We come to the conclusion that the introduction of the reduced intensity is not justified and that the most logical choice is $I_r=0$. The DA is then made for the total specific intensity. In this section, we also discuss the conditions under which the DA is applicable. Section 6 contains a summary of obtained results. Some mathematical properties of the diffuse mode of the RTE are given in Appendix A.

2. RADIATIVE TRANSPORT EQUATION

We consider the RTE in an infinite, spatially uniform, isotropic medium, which is of the form

$$(\hat{\mathbf{s}}\cdot\nabla+\mu_t)I(\mathbf{r},\hat{\mathbf{s}})=\mu_s\int A(\hat{\mathbf{s}},\hat{\mathbf{s}}')I(\mathbf{r},\hat{\mathbf{s}}')\mathrm{d}^2s'+\epsilon(\mathbf{r},\hat{\mathbf{s}}), \quad (2)$$

where $\mu_t=\mu_a+\mu_s$ is the total attenuation coefficient; μ_a and μ_s are the absorption and scattering coefficients, respectively; $A(\hat{\mathbf{s}},\hat{\mathbf{s}}')$ is the phase function normalized by the condition

$$\int \mathrm{d}^2s'A(\hat{\mathbf{s}},\hat{\mathbf{s}}')=1; \quad (3)$$

and $\epsilon(\mathbf{r},\hat{\mathbf{s}})$ is the source term of the form (1). The isotropy of space implies that the phase function can be expanded as

$$A(\hat{\mathbf{s}},\hat{\mathbf{s}}')=\sum_{lm}A_lY_{lm}(\hat{\mathbf{s}})Y^*_{lm}(\hat{\mathbf{s}}'), \quad (4)$$

where $Y_{lm}(\hat{\mathbf{s}})$ are the spherical functions (viewed here as functions of the unit vector $\hat{\mathbf{s}}$) and $A_0=1$. The first moment of the phase function is the so-called scattering asymmetry parameter g,

$$A_1=g=\hat{\mathbf{s}}\cdot\int \mathrm{d}^2s'\hat{\mathbf{s}}'A(\hat{\mathbf{s}},\hat{\mathbf{s}}')\leqslant 1. \quad (5)$$

Several physical models for the higher-order coefficients A_l have been proposed. In the commonly used Henyey–Greenstein model [22], $A_l=g^l$. The density of the electromagnetic energy $u(\mathbf{r})$ and the current of energy $\mathbf{J}(\mathbf{r})$ are expressed as the zeroth and the first angular moments of the specific intensity

$$u(\mathbf{r})=\frac{1}{c}\int \mathrm{d}^2sI(\mathbf{r},\hat{\mathbf{s}}), \quad (6)$$

$$\mathbf{J}(\mathbf{r})=\int \mathrm{d}^2s\hat{\mathbf{s}}I(\mathbf{r},\hat{\mathbf{s}}), \quad (7)$$

where c is the average speed of light in the medium.

3. DENSITY AND CURRENT FROM THE RTE

An explicit solution to Eq. (2) with the source (1) can be derived by the method of rotated reference frames [23,24]. The solution can be written as

$$I(\mathbf{r},\hat{\mathbf{s}})=\mathcal{A}\sum_{m=-\infty}^{\infty}\sum_{l,l'=|m|}^{\infty}\chi^m_{ll'}(r)Y_{lm}(\hat{\mathbf{s}};\hat{\mathbf{r}})Y_{l'm}(\hat{\mathbf{s}}_0;\hat{\mathbf{r}}). \quad (8)$$

In this expansion, $Y_{lm}(\hat{\mathbf{s}};\hat{\mathbf{r}})$ are the spherical functions defined in a reference frame whose z-axis coincides with the direction of the unit vector $\hat{\mathbf{r}}$. A detailed definition of these functions is given in [23,24], but is not needed here. Indeed, to compute the density and the current according to Eqs. (6) and (7), we require only the two integrals

$$\int \mathrm{d}^2sY_{lm}(\hat{\mathbf{s}};\hat{\mathbf{r}})=\sqrt{4\pi}\delta_{l0}\delta_{m0}, \quad (9)$$

$$\int \mathrm{d}^2s\hat{\mathbf{s}}Y_{lm}(\hat{\mathbf{s}};\hat{\mathbf{r}})=\hat{\mathbf{r}}\sqrt{\frac{4\pi}{3}}\delta_{l1}\delta_{m0}. \quad (10)$$

The expansion coefficients $\chi^m_{ll'}(r)$ are given by the expression

$$\chi^m_{ll'}(r)=\frac{(-1)^m}{2\pi\sqrt{\sigma_l\sigma_{l'}}}\sum_{M=-\bar{l}}^{\bar{l}}(-1)^M \times{\sum_n}'\frac{\langle l|y_n(M)\rangle\langle y_n(M)|l'\rangle}{\lambda_n^3(M)} \times\sum_{j=0}^{\bar{l}}C^{|l-l'|+2j,0}_{l,M,l',-M}C^{|l-l'|+2j,0}_{l,m,l',-m}k_{|l-l'|+2j}\left(\frac{r}{\lambda_n(M)}\right), \quad (11)$$

where $\bar{l}=\min(l,l')$ and various quantities appearing in Eq. (11) are explained below in Eqs. (12)–(15). First, the coefficients σ_l ($l=0,1,2,\ldots$) are given by

$$\sigma_l=\mu_a+\mu_s(1-A_l), \quad (12)$$

where A_l are the expansion coefficients in Eq. (4). Second, $\lambda_n(M)$ and $|y_n(M)\rangle$ are the eigenvalues and the eigenvectors of a set of infinite, tridiagonal, symmetric matrices $B(M)$ (parameterized by the integer M) whose elements are of the form

$$\langle l|B(M)|l'\rangle=\frac{b_l(M)\delta_{l',l-1}+b_{l+1}(M)\delta_{l',l+1}}{\sqrt{\sigma_l\sigma_{l'}}}, \quad (13)$$

where $l,l'\geqslant|M|$ and

$$b_l(M) = \sqrt{\frac{l^2 - M^2}{4l^2 - 1}}. \tag{14}$$

Third, $C_{j_1 m_1 j_2 m_2}^{j_3 m_3}$ are the Clebsch–Gordan coefficients. Fourth, $k_l(x) = -\mathrm{i}^l h_l^{(1)}(\mathrm{i}x)$ are the modified spherical Bessel functions of the second kind (defined here without the $\pi/2$ factor). Finally, the spectrum of eigenvalues of all matrices $B(M)$ is symmetric. That is, for each eigenvalue λ, there is also an eigenvalue $-\lambda$, and the summation in Eq. (11) is carried out only over such indices n that $\lambda_n > 0$. Thus, the notation Σ_n' should be understood as

$$\sum_n{}' f_n = \sum_{\lambda_n > 0} f_n. \tag{15}$$

We now use the integrals (9) and (10) to compute the density and the current as

$$u(\mathbf{r}) = \frac{\mathcal{A}}{c} \sum_{l=0}^{\infty} \sqrt{2l+1}\, \chi_{0l}^{0}(r) P_l(\hat{\mathbf{s}}_0 \cdot \hat{\mathbf{r}}), \tag{16}$$

$$\mathbf{J}(\mathbf{r}) = \frac{\mathcal{A}\hat{\mathbf{r}}}{\sqrt{3}} \sum_{l=0}^{\infty} \sqrt{2l+1}\, \chi_{1l}^{0}(r) P_l(\hat{\mathbf{s}}_0 \cdot \hat{\mathbf{r}}), \tag{17}$$

where $P_l(x)$ are the Legendre polynomials. The functions $\chi_{0l}^0(r)$ and $\chi_{1l}^0(r)$ are significantly simplified compared to the more general functions $\chi_{ll'}^m(r)$. Thus, for $\chi_{0l}^0(r)$, we have

$$\chi_{0l}^0(r) = \frac{1}{2\pi\sqrt{\sigma_0\sigma_l}} \sum_n{}' \frac{\langle 0|y_n(0)\rangle\langle y_n(0)|l\rangle}{\lambda_n^3(0)} k_l\left(\frac{r}{\lambda_n(0)}\right). \tag{18}$$

The expression for $\chi_{1l}^0(r)$ is more involved. It is possible to show that

$$\chi_{1l}^0(r) = \frac{1}{2\pi\sqrt{\sigma_1\sigma_l}} \left[-\sum_n{}' \frac{\langle 1|y_n(0)\rangle\langle y_n(0)|l\rangle}{\lambda_n^3(0)} k_l'\left(\frac{r}{\lambda_n(0)}\right) - \frac{\sqrt{2l(l+1)}}{r} \sum_n{}' \frac{\langle 1|y_n(1)\rangle\langle y_n(1)|l\rangle}{\lambda_n^2(1)} k_l\left(\frac{r}{\lambda_n(1)}\right) \right], \tag{19}$$

where $k_l'(z) = \mathrm{d}k_l(z)/\mathrm{d}z$.

We seek an expression for the density and current far from the source, so that r in the above expressions is sufficiently large. A more precise mathematical formulation of this condition will become clear momentarily. The functions $k_l(x)$ decay exponentially as $\exp(-x)$. Let the largest eigenvalue of the matrix $B(0)$ be λ_d and the corresponding eigenvector be $|y_d\rangle$. We will refer to this eigenstate as the diffusion mode. As was shown in [24], the spectrum of eigenvalues is discrete for $\lambda > 1/\mu_t$, and the largest eigenvalue λ_d is in the discrete spectrum. Therefore, there is a finite gap $\Delta\lambda$ between the largest and the second largest eigenvalues of $B(0)$. In many practical situations, this gap is quite significant, as is illustrated in Fig. 1(a). In addition, λ_d is typically much larger than the maximum eigenvalue of all the matrices $B(M)$ where $|M| > 0$. This is illustrated in Fig. 1(b). Thus, at sufficiently large distances from the source, we can neglect the second term in the square brackets in Eq. (19). Further, we can keep only one

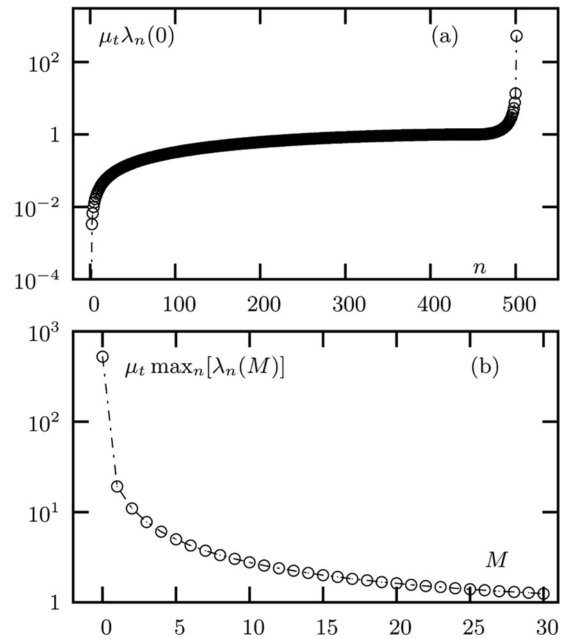

Fig. 1. Eigenvalues of the matrices $B(M)$ for the following parameters: $\mu_a = 0.03\ \mathrm{cm}^{-1}$, $\mu_s = 500\ \mathrm{cm}^{-1}$, $g = 0.98$. These parameters are typical for biological tissues in the near-IR spectral range. (a) All eigenvalues of $B(0)$ versus the eigenvalue number n. (b) The maximum eigenvalues of the matrices $B(M)$ versus M. In simulations, infinite matrices $B(M)$ were truncated so that the size of each matrix was $N = 10^3$.

term in the summations over n in both Eqs. (18) and (19), namely, the term corresponding to the diffusion mode. We then obtain

$$\chi_{0l}^0(r) \approx \frac{1}{2\pi\sqrt{\sigma_0\sigma_l}} \frac{\langle 0|y_d\rangle\langle y_d|l\rangle}{\lambda_d^3} k_l\left(\frac{r}{\lambda_d}\right), \tag{20}$$

$$\chi_{1l}^0(r) \approx \frac{-1}{2\pi\sqrt{\sigma_1\sigma_l}} \frac{\langle 1|y_d\rangle\langle y_d|l\rangle}{\lambda_d^3} k_l'\left(\frac{r}{\lambda_d}\right). \tag{21}$$

The above equalities are exponentially accurate when $\exp(-\kappa) \ll 1$, where

$$\kappa = \frac{r}{\lambda_d} \frac{\Delta\lambda}{\lambda_d - \Delta\lambda}. \tag{22}$$

It is now possible to state the condition under which Fick's law,

$$\mathbf{J} = -D\nabla u, \tag{23}$$

is applicable. This condition is $r/\lambda_d \gg 1$. Indeed, if this inequality holds, we can write approximately $\nabla u = \hat{\mathbf{r}}\partial u/\partial r$ and neglect the tangential derivative of the factors $P_l(\hat{\mathbf{s}}_0 \cdot \hat{\mathbf{r}})$ in Eq. (16). The accuracy of this approximation is algebraic, and the error is of the order of $O(r/\lambda_d)$, while the error in Eqs. (20) and (21) is of the order of $O[\exp(-\kappa)]$. Note, however, that for isotropic sources of the form $\epsilon = \mathcal{A}\delta(\mathbf{r})$, Fick's law is exponentially accurate. Indeed, if the expression (16) is integrated over $\hat{\mathbf{s}}_0$, only the $l = 0$ term remains nonzero. Therefore, the equality $\nabla u = \hat{\mathbf{r}}\partial u/\partial r$ is, in this case, exact.

The diffusion coefficient D in Eq. (23) can easily be found from Eqs. (16), (17), (20), and (21). As was mentioned above, in evaluating the gradient of Eq. (16), the tangential derivative can be neglected. We then obtain

$$D = c\frac{\langle 1|y_d\rangle}{\langle 0|y_d\rangle}\sqrt{\frac{\sigma_0}{3\sigma_1}}\lambda_d. \tag{24}$$

If we also take into account that $\sigma_0=\mu_a$ and that the characteristic equation for the diffusion mode implies $[b_1/\sqrt{\sigma_0\sigma_1}]\langle 1|y_d\rangle=\lambda_d\langle 0|y_d\rangle$ with $b_1=1/\sqrt{3}$, we arrive at the result

$$D = c\mu_a\lambda_d^2. \tag{25}$$

Thus, the diffusion coefficient is defined by the average speed of light in the medium, the absorption coefficient, and the largest eigenvalue of $B(0)$. Some mathematical properties of the diffusion eigenvalue λ_d and of the corresponding eigenvector $|y_d\rangle$ are given in Appendix A.

Introducing the diffuse wave number $k_d=1/\lambda_d$ and using Eqs. (16) and (20), we can write the density $u(\mathbf{r})$ in the form

$$u(\mathbf{r}) = \mathcal{A}\frac{k_d}{D}\sum_{lm} S_l^{(\mathrm{RTE})} k_l(k_d r) Y_{lm}(\hat{\mathbf{r}}) Y_{lm}^*(\hat{\mathbf{s}}_0), \tag{26}$$

where

$$S_l^{(\mathrm{RTE})} = \sqrt{\frac{\sigma_0}{(2l+1)\sigma_l}}2\langle 0|y_d\rangle\langle y_d|l\rangle. \tag{27}$$

In particular,

$$S_0^{(\mathrm{RTE})} = 2\langle 0|y_d\rangle\langle y_d|0\rangle, \quad S_1^{(\mathrm{RTE})} = 2\mu_a\lambda_d\langle 0|y_d\rangle\langle y_d|0\rangle. \tag{28}$$

We will see below that a similar expansion (but with different coefficients $S_{lm}^{(\mathrm{DA})}$) can be obtained in the DA.

4. DIFFUSION APPROXIMATION

The diffusion approximation (DA) to the RTE (2) is usually obtained as follows. We first expand the specific intensity as

$$I(\mathbf{r},\hat{\mathbf{s}}) = I_r(\mathbf{r},\hat{\mathbf{s}}) + I_d(\mathbf{r},\hat{\mathbf{s}}), \tag{29}$$

where I_r is the reduced intensity, which is defined to satisfy a reduced RTE, and I_d is the "diffuse" component of the specific intensity. It is usually assumed that I_d obeys (approximately) a DE, while I_r is highly singular and, therefore, must be considered separately [3].

In this paper, we point out that the reduced intensity can be defined in more than one way. Thus, for example, we can require that the reduced intensity satisfy the RTE (2) in which we set $\mu_s=0$. This leads to the equation

$$(\hat{\mathbf{s}}\cdot\nabla + \mu_a)I_r(\mathbf{r},\hat{\mathbf{s}}) = \epsilon(\mathbf{r},\hat{\mathbf{s}}). \tag{30}$$

We can also formally set $A(\hat{\mathbf{s}},\hat{\mathbf{s}}')=0$ in Eq. (2) and obtain

$$(\hat{\mathbf{s}}\cdot\nabla + \mu_t)I_r(\mathbf{r},\hat{\mathbf{s}}) = \epsilon(\mathbf{r},\hat{\mathbf{s}}). \tag{31}$$

In highly scattering media such that $\mu_t \gg \mu_a$, the above two definitions result in reduced intensities that decay with very different exponential rates.

In the standard approach to the DA, however, neither Eq. (30) nor Eq. (31) is used. Instead, the reduced intensity is defined by

$$(\hat{\mathbf{s}}\cdot\nabla + \mu^*)I_r(\mathbf{r},\hat{\mathbf{s}}) = \epsilon(\mathbf{r},\hat{\mathbf{s}}), \tag{32}$$

where $\mu^*=1/\ell^*=\mu_a+(1-g)\mu_s$ is the reciprocal of the transport mean free path ℓ^*. We note the obvious inequality $\mu_a \leqslant \mu^* \leqslant \mu_t$. The definition (32) utilizes a more physically relevant quantity compared with Eqs. (30) and (31), namely, the reduced scattering coefficient $\mu_s'=(1-g)\mu_s$. Thus Eq. (32) predicts that the reduced intensity in a medium with strictly forward-peaked scattering ($g=1$) is no different from that in a medium with no scattering at all ($\mu_s=0$), as one could expect on physical grounds. The deficiency of the definitions (30) and (31) is that the first is completely independent of μ_s while the second is independent of g.

In spite of the above, all three definitions (30)–(32) are *ad hoc* and need justification. Such justification can be obtained only by direct comparison with the RTE. This will be done below. However, there is no compelling reason to restrict consideration to the three discrete values of the coefficient that enters the definition of reduced intensity (either μ_a, μ_t, or μ^*). Instead, we will pose a more general problem. Suppose that I_r is defined by

$$(\hat{\mathbf{s}}\cdot\nabla + \bar{\mu})I_r(\mathbf{r},\hat{\mathbf{s}}) = \epsilon(\mathbf{r},\hat{\mathbf{s}}), \tag{33}$$

where $\bar{\mu}$ is an arbitrary positive constant (additional constraints on the values of $\bar{\mu}$ will be imposed later). We will view $\bar{\mu}$ as a free parameter and compare the results obtained by making the DA for a given value of $\bar{\mu}$ with the predictions of radiative transport theory.

We now proceed with the derivation of the DA. First, we substitute the decomposition (29) into Eq. (2) and obtain the equation for the diffuse component,

$$(\hat{\mathbf{s}}\cdot\nabla + \mu_t)I_d = \mu_s A I_d + \epsilon_r, \tag{34}$$

where the reduced source function ϵ_r is given by

$$\epsilon_r(\mathbf{r},\hat{\mathbf{s}}) = (\mu_s A - \mu_t + \bar{\mu})I_r \tag{35}$$

and I_r satisfies Eq. (33). In the above two equations, A denotes the linear operator defined by the integral in the right-hand side of Eq. (2).

We seek an approximate solution to Eq. (34) in the form

$$I_d(\mathbf{r},\hat{\mathbf{s}}) = (c/4\pi)u_d(\mathbf{r}) + (3/4\pi)\hat{\mathbf{s}}\cdot\mathbf{J}_d(\mathbf{r}). \tag{36}$$

It can be verified by direct integration of Eq. (36), according to Eqs. (6) and (7), that u_d and $\mathbf{J}_d$ are the energy density and current associated with the diffuse intensity I_d.

The DA is obtained by substitution of the ansatz (36) into Eq. (34) and by considering the first two moments (with respect to the angular variable) of the resultant equation. The substitution yields

$$(\hat{\mathbf{s}} \cdot \nabla + \mu_a)u_d + 3/c(\hat{\mathbf{s}} \cdot \nabla + \mu^*)\hat{\mathbf{s}} \cdot \mathbf{J}_d = (4\pi/c)\epsilon_r. \quad (37)$$

We now evaluate the zeroth and the first moments of the above equation with respect to $\hat{\mathbf{s}}$. In the first case, we integrate Eq. (37) over d^2s and, in the second case, over $\hat{\mathbf{s}}d^2s$. This leads to the two equations

$$\nabla \cdot \mathbf{J}_d + c\mu_a u_d = E, \quad (38)$$

$$\nabla u_d + (3\mu^*/c)\mathbf{J}_d = (3/c)\mathbf{Q}, \quad (39)$$

where the scalar and the vector source terms E and $\mathbf{Q}$ are given by the expressions

$$E(\mathbf{r}) = \int d^2s\epsilon_r(\mathbf{r},\hat{\mathbf{s}}), \quad (40)$$

$$\mathbf{Q}(\mathbf{r}) = \int d^2s\hat{\mathbf{s}}\epsilon_r(\mathbf{r},\hat{\mathbf{s}}). \quad (41)$$

We can use Eq. (35) to express the above two functions in terms of the reduced intensity I_r:

$$E(\mathbf{r}) = (\bar{\mu} - \mu_a) \int d^2sI_r(\mathbf{r},\hat{\mathbf{s}}), \quad (42)$$

$$\mathbf{Q}(\mathbf{r}) = (\bar{\mu} - \mu^*) \int d^2s\hat{\mathbf{s}}I_r(\mathbf{r},\hat{\mathbf{s}}). \quad (43)$$

It can now be seen why the choice $\bar{\mu}=\mu^*$ is special: it causes $\mathbf{Q}(\mathbf{r})$ to vanish. We will show, however, that accounting for this term in the DE does not lead to additional difficulties.

At this point, we require a specific expression for the reduced intensity. It can be easily seen that the solution to Eq. (33) with the source (1) is

$$I_r(\mathbf{r},\hat{\mathbf{s}}) = \mathcal{A}\frac{\exp(-\bar{\mu}r)}{r^2}\delta(\hat{\mathbf{s}} - \hat{\mathbf{s}}_0)\delta(\hat{\mathbf{r}} - \hat{\mathbf{s}}_0). \quad (44)$$

We then substitute Eq. (44) into Eqs. (42) and (43) and obtain

$$E(\mathbf{r}) = \mathcal{A}(\bar{\mu} - \mu_a)\frac{\exp(-\bar{\mu}r)}{r^2}\delta(\hat{\mathbf{r}} - \hat{\mathbf{s}}_0), \quad (45)$$

$$\mathbf{Q}(\mathbf{r}) = \mathcal{A}\hat{\mathbf{s}}_0(\bar{\mu} - \mu^*)\frac{\exp(-\bar{\mu}r)}{r^2}\delta(\hat{\mathbf{r}} - \hat{\mathbf{s}}_0). \quad (46)$$

At the next step, we exclude the current from Eqs. (38) and (39) and obtain a DE with respect to the density $u_d(\mathbf{r})$, namely,

$$(-D\nabla^2 + c\mu_a)u_d(\mathbf{r}) = S(\mathbf{r}), \quad (47)$$

where

$$D = c/3\mu^* \quad (48)$$

is the diffusion coefficient obtained in the approximation specified by Eqs. (29) and (36), and the source term is given by

$$S(\mathbf{r}) = E(\mathbf{r}) - \ell^* \nabla \cdot \mathbf{Q}(\mathbf{r}). \quad (49)$$

The diffusion coefficient (48) and the similar quantity (25) obtained from the RTE under more general conditions differ, but approach each other in the limit $\mu_a/\mu_s \to 0$. In this limit, λ_d is sharply bounded by the inequalities (A8) of Appendix A. If only the lowest-order (in μ_a/μ_s) non-vanishing term is retained, $\lambda_d \approx 1/\sqrt{3\mu_a\mu^*}$. Substitution of this expression into Eq. (25) results in $D \approx c/3\mu^*$ in agreement with Eq. (48). However, the solution to the DE (47) depends on D exponentially. Therefore, even a small error in D can result in an exponentially large discrepancy between the solution to the DE and the solution to the RTE. It has been suggested that the DA can be "corrected" by using a more accurate expression (25) for the diffusion coefficient [11–13]. This correction does not follow in a mathematically consistent way from the DA. Nevertheless, it has proved to be useful. We will adopt this approach below. Specifically, we will use the expression (25) instead of Eq. (48) for the coefficient D that appears in the DE (47).

The second term in (49) contains a derivative of a delta function, namely, $\nabla \cdot \mathbf{Q}$. Although this has not been stated explicitly in the literature, it seems plausible that the choice $\bar{\mu}=\mu^*$ in the definition of I_r has been made because it results in $\mathbf{Q}=0$. However, the presence of the term $\nabla \cdot \mathbf{Q}$ does not pose a mathematical problem. To evaluate the latter, we act with the operator ∇ on Eq. (43) and use Eq. (33). A straightforward calculation yields

$$\nabla \cdot \mathbf{Q} = \mathcal{A}(\bar{\mu} - \mu^*)\left[\delta(\mathbf{r}) - \bar{\mu}\frac{\exp(-\bar{\mu}r)}{r^2}\delta(\hat{\mathbf{r}} - \hat{\mathbf{s}}_0)\right]. \quad (50)$$

Combining this expression with Eq. (45), we obtain

$$S(\mathbf{r}) = \mathcal{A}\left\{-\frac{\bar{\mu} - \mu^*}{\mu^*}\delta(\mathbf{r}) + \frac{\bar{\mu}^2 - \mu_a\mu^*}{\mu^*}\frac{\exp(-\bar{\mu}r)}{r^2}\delta(\hat{\mathbf{r}} - \hat{\mathbf{s}}_0)\right\}. \quad (51)$$

It can be seen that the source term contains two contributions. The first term describes a point source located at the origin; this term vanishes if $\bar{\mu}=\mu^*$. The second term is a function that exponentially decays along the ray $\hat{\mathbf{r}}=\hat{\mathbf{s}}_0$; this term vanishes if $\bar{\mu}=\sqrt{\mu_a\mu^*}$. The integral source is

$$\int d^3rS(\mathbf{r}) = \mathcal{A}(1 - \mu_a/\bar{\mu}). \quad (52)$$

We expect on physical grounds the above expression to be positive and conclude that $\bar{\mu}$ must satisfy $\bar{\mu} > \mu_a$. It also can be expected that the integral (52) should be equal to $\mathcal{A}$, which is the overall power of the source. This already suggests that the proper choice for $\bar{\mu}$ is $\bar{\mu}=\infty$. We will come to the same conclusion using more rigorous arguments below. At this point, we note that under the conditions when the DA is typically used, $\mu_a \ll \mu^*$ and the choice $\bar{\mu}=\mu^*$ may result in a relatively small error. Nevertheless, this error is both non-negligible and easily avoidable. We will also show that the choice $\bar{\mu}=\infty$ results in a more natural and transparent theory.

Note that the current is given in the DA by the formula

$$\mathbf{J}_d(\mathbf{r}) = -D\nabla u_d(\mathbf{r}) + \ell^*\mathbf{Q}(\mathbf{r}). \quad (53)$$

We now use the above results to compute the density $u_d(\mathbf{r})$ due to the source function $S(\mathbf{r})$ (51). We have

$$u_d(\mathbf{r}) = \int \mathrm{d}^3 r' G(\mathbf{r},\mathbf{r}') S(\mathbf{r}'), \tag{54}$$

where

$$G(\mathbf{r},\mathbf{r}') = \frac{\exp(-k_d|\mathbf{r}-\mathbf{r}'|)}{4\pi D|\mathbf{r}-\mathbf{r}'|} \tag{55}$$

is the free-space Green's function of the DE (47) and

$$k_d = \sqrt{c\mu_a/D} = 1/\lambda_d \tag{56}$$

is the diffuse wave number. Note that, as discussed above, we use the expression (25) for the diffusion coefficient.

It is convenient to expand the Green's function as

$$G(\mathbf{r},\mathbf{r}') = \frac{k_d}{D}\sum_{lm} i_l(k_d r_<) k_l(k_d r_>) Y_{lm}(\hat{\mathbf{r}}) Y^*_{lm}(\hat{\mathbf{r}}'), \tag{57}$$

where $i_\ell(x)$ and $k_\ell(x)$ are the modified spherical Bessel functions of the first and second kind, respectively, and $r_<$ and $r_>$ are the lesser and greater of r and r'. Note that $k_\ell(x)$ is defined here without the $\pi/2$ factor, so that $k_0(x) = \exp(-x)/x$, etc. From this, we obtain

$$u_d(\mathbf{r}) = \frac{k_d}{D}\sum_{lm} Y_{lm}(\hat{\mathbf{r}}) \int_0^\infty \mathrm{d}r'(r')^2 i_l(k_d r_<) k_l(k_d r_>) \times \int \mathrm{d}^2\hat{\mathbf{r}}' S(\mathbf{r}') Y^*_{lm}(\hat{\mathbf{r}}'), \tag{58}$$

We now note the following. The Green's function (55) decays exponentially as $\exp(-k_d|\mathbf{r}-\mathbf{r}'|)$, while the the source function (51) decays as $\exp(-\bar{\mu}r')$. On physical grounds, we expect that the density $u_d(\mathbf{r})$ should decay slower than the source function $S(\mathbf{r})$ when $r\to\infty$. This leads to the requirement $\bar{\mu} > k_d$ which is in addition to the previously imposed requirement $\bar{\mu} > \mu_a$. If we assume that the point of observation $\mathbf{r}$ is sufficiently far from the origin, the radial integral in Eq. (58) converges while $r_< = r'$ and $r_> = r$. The density can then be written with exponentially high precision in the form (26), namely, as

$$u_d(\mathbf{r}) = \mathcal{A}\frac{k_d}{D}\sum_{lm} S_l^{(\mathrm{DE})} k_l(k_d r) Y_{lm}(\hat{\mathbf{r}}) Y^*_{lm}(\hat{\mathbf{s}}_0), \tag{59}$$

where

$$S_l^{(\mathrm{DE})} = -\frac{\bar{\mu}-\mu^*}{\mu^*}\delta_{l0} + \frac{\bar{\mu}^2-\mu_a\mu^*}{\mu^* k_d} Q_l(\bar{\mu}/k_d). \tag{60}$$

Here

$$Q_l(p) = \int_0^\infty i_l(x)\exp(-px)\mathrm{d}x, \tag{61}$$

is the Legendre function of the second kind, and we have used the specific form of the source function (51). Note that the Legendre function must be evaluated for arguments that are larger than unity. In particular,

$$Q_0(p) = \frac{1}{2}\ln\frac{p+1}{p-1}, \quad Q_1(p) = -1 + \frac{p}{2}\ln\frac{p+1}{p-1}. \tag{62}$$

5. RESULTS AND DISCUSSION

We now compare the results derived in Sections 3 and 4. In Fig. 2, we plot the absolute difference between the expansion coefficients $S_l^{(\mathrm{RTE})}$ that are defined in Eq. (27) and $S_l^{(\mathrm{DE})}$ that are defined in Eq. (60). Comparison is made only for $l=0$ and $l=1$. The higher-order coefficients $S_l^{(\mathrm{RTE})}$

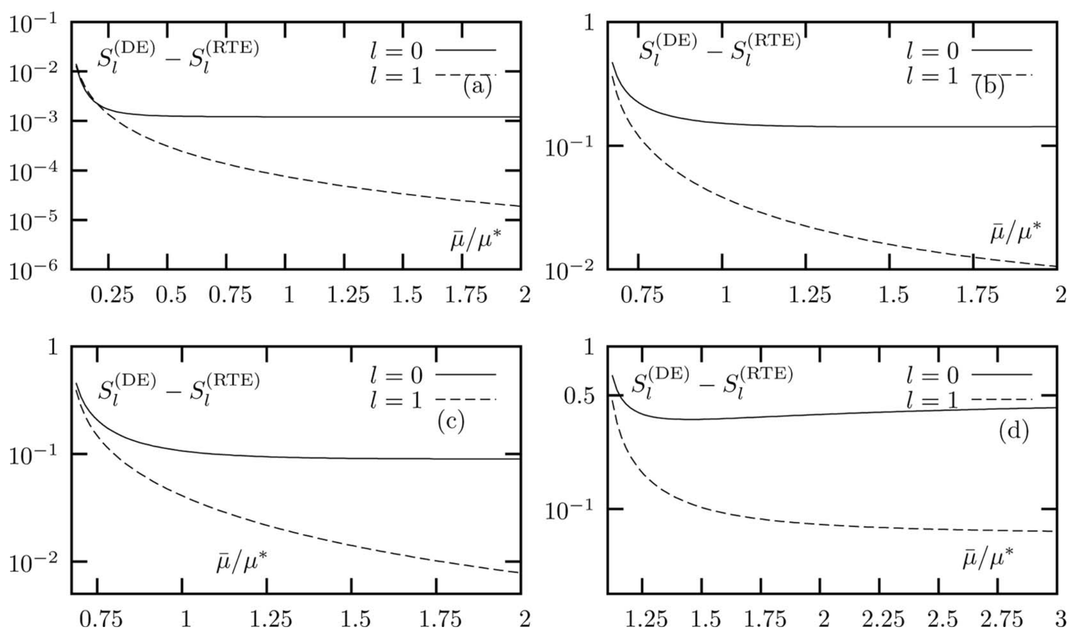

Fig. 2. Absolute difference $S_l^{(\mathrm{DE})} - S_l^{(\mathrm{RTE})}$ as a function of $\bar{\mu}/\mu^*$ for $l=0,1$ and for the following sets of parameters: (a) $\mu_a = 0.03\ \mathrm{cm}^{-1}$, $\mu_s = 500\ \mathrm{cm}^{-1}$, $g=0.98$ (these are the "physiological" parameters typical of biological tissues in the near-IR spectral range; same parameters have been used in Fig. 1). (b) $\mu_a = 1\ \mathrm{cm}^{-1}$, $\mu_s = 5\ \mathrm{cm}^{-1}$, $g=0$ (isotropic scattering). (c) $\mu_a = 1\ \mathrm{cm}^{-1}$, $\mu_s = 10\ \mathrm{cm}^{-1}$, $g=0.5$ [anisotropic scattering but same μ^* as in (b)]. (d) $\mu_a = \mu_s = 1\ \mathrm{cm}^{-1}$, $g=0.5$.

depend on the higher-order coefficients A_l (in addition to $g=A_1$). The DA is independent of these higher-order coefficients and, therefore, direct comparison for $l>1$ is inappropriate.

It can be seen that in all cases, the point $\bar{\mu}=\mu^*$ has no special significance. The difference $S_0^{(\mathrm{DE})}-S_0^{(\mathrm{RTE})}$ is rather flat, especially in the three cases (a)–(c) for which the DA is typically considered to be applicable. If only the zeroth-order expansion coefficients are examined, one can conclude that the choice of $\bar{\mu}$ does not matter at all, and one can choose without loss of accuracy any value of $\bar{\mu}$, provided that $\bar{\mu}\gtrsim\mu^*$. However, the first-order difference $S_1^{(\mathrm{DE})}-S_1^{(\mathrm{RTE})}$ monotonically decreases with $\bar{\mu}$. This suggests that the optimal choice is $\bar{\mu}=\infty$. It may seem that this choice is paradoxical and mathematically ill defined. However, this is not so. The quantity $u_d(\mathbf{r})$ is well defined in the limit $\bar{\mu}\to\infty$. Indeed, we have

$$\lim_{\bar{\mu}\to\infty} S_0^{(\mathrm{DE})}=1, \quad \lim_{\bar{\mu}\to\infty} S_1^{(\mathrm{DE})}=k_d/3\mu^*, \quad S_l^{(\mathrm{DE})}=0 \text{ for } l>1. \tag{63}$$

From this, we find

$$\begin{aligned} u_d(\mathbf{r}) &= (1+\ell^* k_d \hat{\mathbf{r}}\cdot\hat{\mathbf{s}}_0)\exp(-k_d r)/4\pi D r \\ &= (1+\ell^* k_d \hat{\mathbf{r}}\cdot\hat{\mathbf{s}}_0)G(\mathbf{r},0). \end{aligned} \tag{64}$$

Alternatively, we can take the $\bar{\mu}\to\infty$ limit of Eq. (54) with $S(\mathbf{r})$ given by Eq. (51). We obtain

$$\lim_{\bar{\mu}\to\infty} S(\mathbf{r}) = \mathcal{A}(1-\ell^*\hat{\mathbf{s}}_0\cdot\nabla)\delta(\mathbf{r}), \tag{65}$$

$$\lim_{\bar{\mu}\to\infty} u_d(\mathbf{r}) = \mathcal{A}(1+\ell^*\hat{\mathbf{s}}_0\cdot\nabla_{\mathbf{r}'})G(\mathbf{r},\mathbf{r}')|_{\mathbf{r}'=0}. \tag{66}$$

The same result can be obtained even more directly by noting that the limit $\bar{\mu}\to\infty$ corresponds to $I_r=0$. We thus can apply the DA to Eq. (34) in which the reduced source ϵ_r must be substituted by the original source of the RTE, ϵ, given in Eq. (1). This yields $E(\mathbf{r})=\mathcal{A}\delta(\mathbf{r})$, $\mathbf{Q}(\mathbf{r})=\mathcal{A}\hat{\mathbf{s}}_0\delta(\mathbf{r})$ and, according to Eq. (49), $S(\mathbf{r})=\mathcal{A}(1-\ell^*\hat{\mathbf{s}}_0\cdot\nabla)\delta(\mathbf{r})$. This is in agreement with Eq. (65). Thus, it can be seen that the use of the reduced intensity I_r is unnecessary. Note that its introduction actually decreases the accuracy of the DA. Also, the reduced intensity cannot be interpreted as the correct specific intensity in the vicinity of the source. This is clear already from the ambiguity in the definition of I_r, as discussed in Section 4.

We now compare the density $u(\mathbf{r})$ computed using the RTE [according to Eq. (16)] without any approximations and the density $u_d(\mathbf{r})$ obtained in the DA. The results are shown in Fig. 3. The relative error Δ, defined as

$$\Delta = \frac{|u-u_d|}{u}, \tag{67}$$

is plotted as a function of $\bar{\mu}$ for various distances to the source ($r=10\ell^*$ and $r=20\ell^*$) and various orientations of the point of observation relative to the direction of incidence $\hat{\mathbf{s}}_0$. It can be seen that the point $\bar{\mu}=\mu^*$ is in no way special or optimal. Depending on the angle $\theta=\arccos(\hat{\mathbf{r}}\cdot\hat{\mathbf{s}}_0)$, the discrepancy has minima at certain seemingly random values of $\bar{\mu}$. None of these values can be universally used because each of them minimizes the error only for a specific distance r, a specific angle θ, and a specific set of parameters μ_a, μ_s, and g. At the same time, it is obvious that the choice $\bar{\mu}=\infty$ is almost always preferable to $\bar{\mu}=\mu^*$.

To conclude this section, we discuss the choice $\bar{\mu}=\mu^*$ in more detail. We will denote the source function that corresponds to this case by $S^*(\mathbf{r})$ and the corresponding moments $S_l^{(\mathrm{DE})}$ by S_l^*. As can be seen from Eqs. (51) and (60),

$$S^*(\mathbf{r}) = \mathcal{A}\mu_s' \frac{\exp(-\mu^* r)}{r^2}\delta(\hat{\mathbf{r}}-\hat{\mathbf{s}}_0), \tag{68}$$

$$S_l^* = (\mu_s'/k_d)Q_l(\mu^*/k_d), \tag{69}$$

where $\mu_s'=(1-g)\mu_s$ is the reduced scattering coefficient. In biomedical optics, it is typical to work in the regime in which $\mu^*\gg k_d$, although this inequality is never very strong. Thus, for the physiological parameters used in Figs. 1, 2(a), and 3(a), $\mu^*/k_d\sim 10$. Nevertheless, for this ratio of μ^*/k_d, the asymptotic formulas for $Q_0(\mu^*/k_d)$ and $Q_1(\mu^*/k_d)$ are already very accurate and we obtain

$$S_0^* \approx \mu_s'/\mu^*, \quad S_1^* \approx (\mu_s'/\mu^*)(k_d/3\mu^*). \tag{70}$$

These expressions differ from the first two equations in Eqs. (63) by the constant factor μ_s'/μ^*.

We can make an additional approximation and assume that all higher-order coefficients S_l^* ($l>1$) are zero. Then it turns out that the density u_d obtained in the DA with $\bar{\mu}=\mu^*$ differs from the density obtained for $\bar{\mu}=\infty$ (which we deem here to be more accurate) by the overall factor of μ_s'/μ^*. In the case of physiological parameters, this factor is close to unity. Essentially, this approach was adopted by us earlier in [25] where the factor $\mu_s'/\mu^*=\mu_s'\ell^*$ appears, for example, in Eq. (16). Note that in [25], we have implicitly assumed that $S_l^*=0$ for $l>1$ by truncating the Taylor expansion of the Green's function $G(\mathbf{r},\mathbf{r}')$ at the first order, namely, by writing $G(\mathbf{r},\mathbf{r}')\approx G(\mathbf{r},\mathbf{r}_0)+(\mathbf{r}'-\mathbf{r}_0)\cdot\nabla_{\mathbf{r}'}G(\mathbf{r},\mathbf{r}')|_{\mathbf{r}'=0}$, where $\mathbf{r}_0$ is the location of the source. This approximation was then used in the integral of the form (54).

Thus, in [25], we have utilized the choice $\bar{\mu}=\mu^*$ and obtained the result for u_d the differs from Eq. (65) or (66) by the overall factor μ_s'/μ^*, which may seem to be an insignificant correction. However, an additional approximation was used in [25] (neglecting the higher-order moments S_l^* as discussed above) that was not well justified. In a sense, there were two inaccuracies in [25] that canceled each other. The first inaccuracy was neglecting the higher-order moments, and the second inaccuracy was the incorrect choice of $\bar{\mu}$. Further, in less ideal cases, the ratio μ_s'/μ^* may be significantly different from unity, yet the DA could still be applicable sufficiently far from the source. In such cases, an incorrect choice of $\bar{\mu}$ can lead to significant errors. Finally, the derivation of the DA presented in this paper is mathematically consistent and based on direct comparison with the RTE rather than on an *ad hoc* choice of $\bar{\mu}$.

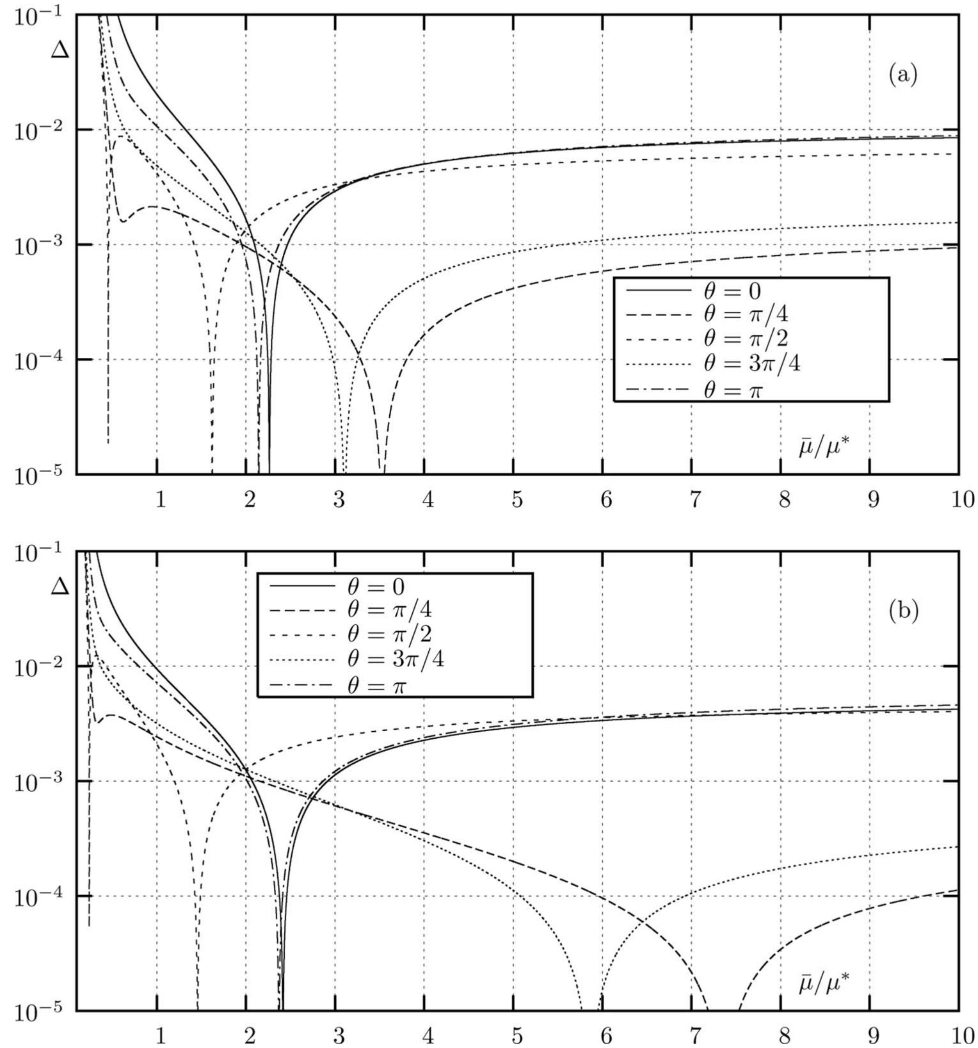

Fig. 3. Relative error Δ (67) as a function of $\bar{\mu}/\mu^*$ for different values of the distance from the source r and the angle θ, where $\cos\theta = \hat{\mathbf{s}}_0 \cdot \hat{\mathbf{r}}$. The distance to the source is r=(a) $10\ell^*$, (b) $20\ell^*$. All parameters are the same as in Fig. 1 and Fig. 2(a).

So far, we have not discussed the conditions under which the DA is applicable. This question is both simple and difficult. It is often remarked that the DA is valid when $\mu_a \ll \mu_s$. However, it is not always clear what exactly is meant by applicability of the DA.

If it is deemed that the DA is applicable if the exact density of electromagnetic energy u defined by Eq. (6) satisfies (approximately) the diffusion equation (47), then the inequality $\mu_a \ll \mu_s$ is not really required (for the reminder of this section, the subscript d in the referenced formulas should be omitted). Indeed, the diffusion equation for u follows from the continuity equation (38) and Fick's law (23). But the continuity equation that couples the density and the current is exact irrespective of parameters. As for Fick's law, the condition for its applicability was discussed at the end of Section 3 and reads $r \gg \lambda_d$. Thus, at sufficiently large distances from the source, Fick's law is accurate irrespective of the ratio μ_a/μ_s. It should be kept in mind though that when the ratio μ_a/μ_s is not small, expressions (25) and (48) can yield very different values of the diffusion coefficient, and the diffusion constant appearing in Fick's law is given by the former expression.

If, however, we deem the DA to be accurate when the specific intensity $I(\mathbf{r},\hat{\mathbf{s}})$ is well approximated by the expansion (36), then the applicability of the DA is more difficult to show. Indeed, for the expansion (36) to be accurate, it is required that all the coefficients $\chi_{ll'}^{m}(r)$ in the expansion (8) be negligibly small for $l,l' > 1$. Now consider the expression (11) for $\chi_{ll'}^{m}(r)$. If we make an approximation in which only one term in the summation over n is retained, namely, the term that corresponds to the diffusion mode (see Appendix A), the above condition can not be proved. Indeed, the Bessel functions $k_l(x)$ increase factorially with l (for large values of l and a fixed x), while the components of the diffuse eigenvector $\langle l|y_d\rangle$ decrease with l only exponentially, as is shown in Appendix A. Thus, the factorial growth of the first factor always overpowers the exponential decay of the second. In fact, if an accurate pointwise approximation to $I(\mathbf{r},\hat{\mathbf{s}})$ is being sought, it is not correct to retain only the diffuse mode in Eq. (11). Instead, summation over all modes must be performed, including the modes of the continuous spectrum. The terms in this summation have different sign. As a result, if the summation is performed with sufficient numerical accuracy, the resultant coefficients $\chi^m_{ll'}(r)$ decrease with l and l' so that the expansion (8) is, in fact, convergent. (When the summation is performed numerically, one should be conscious of the round-off errors, which can become large and even dominant for very large values of l and l'.) Therefore, the proof of pointwise convergence of the specific intensity to the "diffuse" value given by formula (36) is difficult to obtain, and we are not certain whether it is known. Such a proof could be, however, purely of academic interest. The reason is that all physical detectors always measure the specific intensity integrated over a finite area and a finite solid angle. This integration regularizes the formula (11) and can greatly improve numerical convergence. It therefore can be stated that the condition of validity of the DA depends on the type of detector used.

6. SUMMARY

We have systematically developed the diffusion approximation (DA) to the radiative transport equation (RTE) in infinite homogeneous space. We have examined different admissible definitions of the reduced intensity I_r that is commonly used in the derivation of the DA. The existing ambiguity in the definition of I_r affects the form of the source function for the DE. By comparing the results of the DA with more rigorous solutions to the RTE, we have found that the best accuracy is achieved if we set $I_r=0$. In this case the source function for the DE is given by Eq. (65) and the density of electromagnetic energy is given by Eq. (66) or, equivalently, by Eq. (64). We conclude that the separation of the total specific intensity into the reduced and the diffuse components is not justified and, in a typical derivation of the DA, leads to additional errors. Thus, the DA should be made for the total specific intensity.

The theory developed in this paper is not intended to significantly improve the accuracy of the DA, although some improvement can be anticipated. It is rather aimed at removing the existing ambiguities and formulating the DA in a mathematically consistent way. We also study the general limits of applicability of Fick's law and derive certain mathematical properties of the so-called diffuse mode of the RTE.

APPENDIX A: DIFFUSION MODE

In this appendix, we derive some mathematical properties of the diffusion mode, more specifically, of the eigenvalue λ_d and of the corresponding eigenvector $|y_d\rangle$. By definition, λ_d is the largest eigenvalue of the infinite, tridiagonal, real and symmetric matrix $B(0)$ that has been defined in Section 3, Eq. (13). We denote the elements of the first superdiagonal of this matrix by β_ℓ, where

$$\beta_l = \langle l-1|B(0)|l\rangle = \frac{b_l}{\sqrt{\sigma_{l-1}\sigma_l}}, \quad l=1,2,\ldots. \tag{A1}$$

The coefficients b_l are given in Eq. (14) in which we must set $M=0$, so that $b_l=l/\sqrt{4l^2-1}$. Note that

$$\beta_1 = 1/\sqrt{3\mu_a\mu^*}, \tag{A2}$$

$$\beta_2 = 2/\sqrt{15\mu^*[\mu_a+\mu_s(1-A_2)]}, \tag{A3}$$

$$\lim_{l\to\infty}\beta_l = 1/2\mu_t. \tag{A4}$$

In obtaining the limit (A4), we have assumed that $\lim_{l\to\infty}A_l=0$ and $\lim_{l\to\infty}\sigma_l=\mu_t$.

We start by deriving the lower and upper bounds of λ_d. The upper bound is obtained from the Gershgorin theorem, which states that

$$|\lambda_d| \leq \max_l \sum_{l'} (1-\delta_{ll'})|\langle l|B(0)|l'\rangle| = \beta_1+\beta_2. \tag{A5}$$

Here we have assumed that $\sigma_0<\sigma_1<\sigma_2\cdots$, as is typically the case, and therefore $\beta_1>\beta_2>\beta_3\ldots$.

The lower bound can be derived by considering the sequence of truncated matrices $B_j(0)$ $(j=1,2,\ldots)$, which are obtained by keeping only the first j rows and columns of the infinite matrix $B(0)$. As a direct consequence of the Cauchy interlace theorem, we can state that the eigenvalues of $B_j(0)$ and $B_{j+1}(0)$ interlace. Consequently, the maximum eigenvalue of each matrix $B_j(0)$, λ_{dj}, increases monotonically with j. We thus have $\lambda_{d1}<\lambda_{d2}<\cdots<\lambda_d$ where $\lambda_d=\lim_{j\to\infty}\lambda_{dj}$. For $j=3$,

$$\lambda_{d3} = \sqrt{\beta_1^2+\beta_2^2}. \tag{A6}$$

Thus, we have proved the inequality

$$\sqrt{\beta_1^2+\beta_2^2} \leq \lambda_d \leq \beta_1+\beta_2. \tag{A7}$$

Substituting β_1 and β_2 from Eqs. (A2) and (A3), we can rewrite this as

$$\frac{\sqrt{1+\eta}}{\sqrt{3\mu_a\mu^*}} \leq \lambda_d \leq \frac{1+\sqrt{\eta}}{\sqrt{3\mu_a\mu^*}}, \tag{A8}$$

where

$$\eta = \frac{4}{5}\frac{\mu_a}{\mu_a+\mu_s(1-A_2)}. \tag{A9}$$

In the limit when $\mu_a\ll\mu_s$, η can be viewed as a small parameter.

A sharper estimate can be obtained if we use λ_{d4} as the lower bound for λ_d. The former is given by the formula

$$\lambda_{d4} = \sqrt{\frac{\beta_1^2+\beta_2^2+\beta_3^2}{2}\left[1+\sqrt{1-\frac{4\beta_1^2\beta_3^2}{(\beta_1^2+\beta_2^2+\beta_3^2)^2}}\right]}. \tag{A10}$$

Further, the gap between the largest and the second largest eigenvalues of $B(0)$ can be estimated as follows. Let the second largest eigenvalue be λ_s. From the Gershgorin theorem we have $\lambda_s\leq\max(\beta_1,\beta_2+\beta_3)$. Combining this with Eq. (A7), we find that

$$\Delta\lambda \equiv \lambda_d-\lambda_s \geq \max[\sqrt{\beta_1^2+\beta_2^2}-\max(\beta_1,\beta_2+\beta_3),0]. \tag{A11}$$

Note that in the limit of strong scattering we have $\beta_1\gg\beta_2,\beta_3$, and the first number in the square brackets is positive, so that the "max" can be omitted and we have

$$\Delta\lambda \geqslant \sqrt{\beta_1^2+\beta_2^2}-\beta_1. \tag{A12}$$

Finally, we consider the components of the eigenvector $|y_d\rangle$. The first two components ($l=0$ and $l=1$) can be expanded in powers of η as

$$\langle 0|y_d\rangle = 1/\sqrt{2}[1-1/2\,\eta+O(\eta^2)], \tag{A13}$$

$$\langle 1|y_d\rangle = 1/\sqrt{2}[1+O(\eta^2)]. \tag{A14}$$

At large values of the index l, the components $\langle l|y_d\rangle$ decay exponentially [24] as

$$\langle l|y_d\rangle \propto \exp(-\tau l), \quad (l\to\infty). \tag{A15}$$

By considering the three-term recurrence relation for $\langle l|y_d\rangle$ in the limit $l\to\infty$ when $\beta_l\to 1/2\mu_t$, we find

$$\tau = \ln[\mu_t\lambda_d+\sqrt{(\mu_t\lambda_d)^2-1}]. \tag{A16}$$

ACKNOWLEGMENT

This research was supported by the National Science Foundation (NSF) under grants EEC-0615857 and DMS-0554100.

REFERENCES

1. S. Chandrasekhar, *Radiative Transfer* (Dover, 1960).
2. V. V. Sobolev, *Light Scattering in Planetary Atmospheres*, Vol. 76 of International Series of Monographs in Natural Philosophy (Elsevier, 1975).
3. A. Ishimaru, *Wave Propagation and Scattering in Random Media* (IEEE Press, 1997).
4. M. C. W. van Rossum and Th. M. Nieuwenhuizen, “Multiple scattering of classical waves: microscopy, mesoscopy, and diffusion,” Rev. Mod. Phys. **71**, 313–371 (1999).
5. S. R. Arridge, “Optical tomography in medical imaging,” Inverse Probl. **15**, R41–R93 (1999).
6. D. A. Boas, D. H. Brooks, E. L. Miller, C. A. DiMarzio, M. Kilmer, R. J. Gaudette, and Q. Zhang, “Imaging the body with diffuse optical tomography,” IEEE Signal Process. Mag. **18**, 57–75 (2001).
7. A. P. Gibson, J. C. Hebden, and S. R. Arridge, “Recent advances in diffuse optical imaging,” Phys. Med. Biol. **50**, R1–R43 (2005).
8. G. Holte, “On a method of calculating the density of neutrons emitted from a point source in an infinite medium,” Ark. Mat., Astron. Fys. **35A**, 1–9 (1948).
9. K. M. Case and P. F. Zweifel, *Linear Transport Theory* (Addison-Wesley, 1967).
10. I. Kuscer and N. J. McCormick, “Some analytical results for the radiative transfer in thick atmospheres,” Transp. Theory Stat. Phys. **20**, 351–381 (1991).
11. R. Aronson and N. Corngold, “Photon diffusion coefficient in an absorbing medium,” J. Opt. Soc. Am. A **16**, 1066–1071 (1999).
12. R. Elaloufi, R. Carminanti, and J.-J. Greffet, “Definition of the diffusion coefficient in scattering and absorbing media,” J. Opt. Soc. Am. A **20**, 678–685 (2003).
13. R. Pierrat, J.-J. Greffet, and R. Carminati, “Photon diffusion coefficient in scattering and absorbing media,” J. Opt. Soc. Am. A **23**, 1106–1110 (2006).
14. I. Freund, “Surface reflections and boundary conditions for diffusive photon transport,” Phys. Rev. A **45**, 8854–8858 (1992).
15. R. C. Haskell, L. O. Svaasand, T. T. Tsay, T. C. Feng, and M. S. McAdams, “Boundary-conditions for the diffusion equation in radiative transfer,” J. Opt. Soc. Am. A **11**, 2727–2741 (1994).
16. R. Aronson, “Boundary conditions for diffuse light,” J. Opt. Soc. Am. A **12**, 2532–2539 (1995).
17. S. Fantini, M. A. Franceschini, and E. Gratton, “Effective source term in the diffusion equation for photon transport in turbid media,” Appl. Opt. **36**, 156–163 (1997).
18. A. F. Kostko and V. A. Pavlov, “Location of the diffusing-photon source in a strongly scattering medium,” Appl. Opt. **36**, 7577–7582 (1997).
19. X. Intes, B. Le Jeune, F. Pellen, Y. Guern, J. Cariou, and J. Lotrian, “Localization of the virtual point source used in the diffusion approximation to model a collimated beam source,” Waves Random Complex Media **9**, 489–499 (1999).
20. T. Spott and L. O. Svaasand, “Collimated light sources in the diffusion approximation,” Appl. Opt. **39**, 6453–6465 (2000).
21. A. D. Kim and A. Ishimaru, “Optical diffusion of continuou-wave, pulsed, and density waves in scattering media and comparisons with radiative transfer,” Appl. Opt. **37**, 5313–5319 (1998).
22. L. G. Henyey and J. L. Greenstein, “Diffuse radiation in the galaxy,” Astrophys. J. **93**, 70–83 (1941).
23. V. A. Markel, “Modified spherical harmonics method for solving the radiative transport equation,” Waves Random Complex Media **14**, L13–L19 (2004).
24. G. Panasyuk, J. C. Schotland, and V. A. Markel, “Radiative transport equation in rotated reference frames,” J. Phys. A **39**, 115–137 (2006).
25. V. A. Markel and J. C. Schotland, “Symmetries, inversion formulas, and image reconstruction for optical tomography,” Phys. Rev. E **70**, 056616 (2004).