Nondiffuse Radiative Transfer 2: Planar Area Sources and Receivers

D.L. DiLaura

THIS PAPER IS AN IESNA PAPER OF DISTINCTION. IT WAS ORIGINALLY PRESENTED AT THE 1995 IESNA ANNUAL CONFERENCE.

Introduction

The most common form of nondiffuse radiative transfer calculation in lighting is determining the initial flux from luminaires onto discrete elements of architectural surface. The fundamental expression for the flux transferred from emitter to receiver is a double area integral over the areas of emitter and receiver. In general, such integrals do not have an analytic solution and various approximations have been developed.[1]

The first paper in this series derived a contour integral taken around the perimeter of a nondiffuse emitter, which gives the illuminance at a point from the emitter.[2] This work extends this to an illuminated area; the result being a double contour integral, taken around the perimeters of emitter and receiver, which gives the total flux transferred directly from emitter to receiver. This is similar to the double contour integral first obtained for diffuse emitters by Fock.[3] The double contour integral expression for area receivers and diffuse area emitters is well known, though only recently has an analytic expression been obtained for polygonal emitters and receivers.

The present results are useful in the design and analysis of luminaires where the reflectance properties of materials are not diffuse and the discrete elements that constitute reflecting surfaces must be treated as non-diffuse emitters.[4] Additionally, the present results can be used to calculate nondiffuse form factors for use in non-diffuse radiative transfer analysis.

Strategy

The strategy used in this development is as follows. The light vector produced at a point by a nondiffuse plane area emitter is determined. Using a form of Stokes's theorem, the resulting area integral is transformed to a contour integral around the emitter's perimeter. This total light vector is then used to determine the flux transferred to a differential area of a receiver. Area integration over the receiving surface gives the total flux. This second area integral is transformed to a contour integral, leaving a double contour integral around the perimeters of the emitter and receiver. The spherical coordinate system is used throughout having variables (r, θ, φ) and unit vectors $(\hat{\mathbf{r}}, \hat{\theta}, \hat{\varphi})$. It is convenient for light field manipulations and for expressing nondiffuse luminance distributions.

Author's affiliation: Civil and Architectural Engineering Dept., University of Colorado at Boulder, Boulder, CO.

Tensors, dyads, and the light field

The vector light field is a powerful and simplifying formalism for computations and developments like those attempted here.[5,6] Consequently, vectors, tensors, and a little field theory are used in the present work. The tensors used in this development are of second order, and thus are simple extensions of the idea of the product of two vectors. Therefore, although modern tensor subscript and superscript notation is in general use, what is used here is the older, and somewhat more accessible, notation of dyads and dyadics.[7–9]

Vector multiplication can be defined by the dot product, $\mathbf{a} \circ \mathbf{b}$, and the cross product, $\mathbf{a} \times \mathbf{b}$. Along with these two operations, an operationless product exists; vectors are simply bound together forming a single entity. We will indicate this as two vectors placed together: $\mathbf{ab}$. Insistence on a geometric interpretation gives **Figure 1**. The single entity $\mathbf{ab}$ can be thought of as having potential or capability in two directions, that of $\mathbf{a}$ and that of $\mathbf{b}$; the magnitude of these potentials or capabilities being $|\mathbf{a}|$ and $|\mathbf{b}|$, respectively. It is in this sense that $\mathbf{ab}$ is an extension of the idea of a vector, which has potential or capability in one direction.

The entity $\mathbf{ab}$ is a tensor of order 2 and is also called a dyad. It can be thought of as a single operator, capable of transforming a vector $\mathbf{t}$ into a vector $\mathbf{u}$. The left and right components (sometimes called the antecedent and consequent) of a dyad are not necessarily the same, and in this case the order of the antecedent and consequent vectors is important. That is, it is usually the case that $\mathbf{ab} \neq \mathbf{ba}$. Given that $\mathbf{a}$ and $\mathbf{b}$ are defined in Cartesian coordinates as

$$\mathbf{a} = \hat{\mathbf{x}}\,a_x + \hat{\mathbf{y}}\,a_y + \hat{\mathbf{z}}\,a_z$$
$$\mathbf{b} = \hat{\mathbf{x}}\,b_x + \hat{\mathbf{y}}\,b_y + \hat{\mathbf{z}}\,b_z$$

$\mathbf{ab}$ can be interpreted as $(\hat{\mathbf{x}}a_x + \hat{\mathbf{y}}a_y + \hat{\mathbf{z}}a_z)(\hat{\mathbf{x}}b_x + \hat{\mathbf{y}}b_y + \hat{\mathbf{z}}b_z)$ and is seen to have nine components:

$$\mathbf{ab} \equiv \begin{matrix} \hat{\mathbf{x}}\hat{\mathbf{x}}\,a_x b_x & \hat{\mathbf{x}}\hat{\mathbf{y}}\,a_x b_y & \hat{\mathbf{x}}\hat{\mathbf{z}}\,a_x b_z \\ \hat{\mathbf{y}}\hat{\mathbf{x}}\,a_y b_x & \hat{\mathbf{y}}\hat{\mathbf{y}}\,a_y b_y & \hat{\mathbf{y}}\hat{\mathbf{z}}\,a_y b_z \\ \hat{\mathbf{z}}\hat{\mathbf{x}}\,a_z b_x & \hat{\mathbf{z}}\hat{\mathbf{y}}\,a_z b_y & \hat{\mathbf{z}}\hat{\mathbf{z}}\,a_z b_z \end{matrix}$$

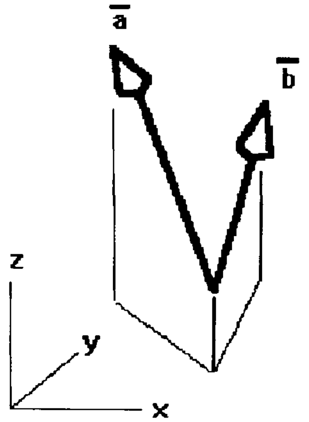

Figure 1—Geometric interpretation of the dyad ab

The scalar quantities (e.g., $a_x b_z$) that weight the components are said to be the coefficients of the dyad. They can, of course, have the value 0. It can be seen that the simplest dyads in a given coordinate system are formed from pairs of its unit coordinate vectors.

Operations on tensors may raise or lower their order. If a dyad **ab** is dotted on the right with a vector **t**, the result is a vector in the direction of **a**, having a magnitude equal to that of **a**, multiplied by the dot product of the vector **t** and **b**. That is, $\mathbf{ab}\circ\mathbf{t} = \mathbf{a}(\mathbf{b}\circ\mathbf{t})$. Dotting the dyad on the left with a vector t gives, in general, a different result in both magnitude and direction: $\mathbf{t}\circ\mathbf{ab} = (\mathbf{t}\circ\mathbf{a})\mathbf{b}$. In each case, the result is a vector. That is, the dot product operation on the dyad, a tensor of order 2, has reduced it to a vector, a tensor of order 1. Not all operations reduce a dyad to a vector. For the cross product of a vector and a dyad we have either $\mathbf{abt} = \mathbf{a}(\mathbf{b}\times\mathbf{t})$ or $\mathbf{t}\times\mathbf{ab} = (\mathbf{t}\times\mathbf{a})\mathbf{b}$. In each case the result is a dyad.

Dyads can be operated on with vector operators, including partial differentiation, ∇. The curl of a dyad can be obtained by extension of the curl applied to vectors. The result is somewhat more elaborate, the ∇ operator being applied to both the antecedent and the consequent: $\nabla\times\mathbf{ab} = (\nabla\times\mathbf{a})\mathbf{b} - \mathbf{a}\times(\nabla\mathbf{b})$.

The sum of dyads is a dyadic. A dyadic, $\Im$, of three dyads can be written as $\Im = \mathbf{au} + \mathbf{bv} + \mathbf{cw}$. It is always possible to transform such a dyadic into one having three dyads, each with a unit orthogonal coordinate vector as its antecedent. Similarly for consequents. That is, vectors $\mathbf{u}'$, $\mathbf{v}'$, $\mathbf{w}'$, $\mathbf{a}'$, $\mathbf{b}'$, and $\mathbf{c}'$ can always be found such that $\Im = \hat{\mathbf{r}}\mathbf{u}' + \hat{\theta}\mathbf{v}' + \hat{\varphi}\mathbf{w}' = \mathbf{a}'\hat{\mathbf{r}} + \mathbf{b}'\hat{\theta} + \mathbf{c}'\hat{\varphi}$ Gibbs and Wilson give a particularly clear introduction to dyads and dyadics.[8]

Dyad formulation of the light vector

We use the vector light field as defined and described by Moon.[10] Yamauti[11] and Gershun[12] provide a more abstract development. The excellent monograph by Gershun is particularly recommended.

A differential source creates a vector light field. At any point p in that field, the light vector, $\mathbf{dE_p}$, has a direction equal to that of the radial line from the emitter to point p. The magnitude of this vector is equal to the luminous intensity, $I(\theta_p,\varphi_p)$, exhibited by the differential source in the direction of p, divided by the square of the distance from source to point p. That is, $\mathbf{dE_p} = I(\theta_p,\varphi_p)\hat{\mathbf{r}}/r^2$. The magnitude is the spatial flux density at point p, and the direction is radial from the emitter. Multiple emitters create a vector light field that is the vector sum of the individual fields. That is, a light vector, **E**, can result from the vector sum of many light vectors; including the integration of differential light vectors.

We begin with the fundamental equation for the differential light vector, **dE**, generated at some point p by the luminance of a differential element, $\mathbf{A_1}$ of a nondiffuse emitter, $\mathbf{A_1}$, assuming radiative transfer in non absorbing and non-scattering media:

$$\mathbf{dE} = -\frac{L(\theta,\varphi)\cos(\theta)\, dA_1}{r^2}\hat{\mathbf{r}} \tag{1}$$

where **dE** = differential light vector produced by dA_1

$L(\theta,\varphi)$ = scalar field equal to the luminance of the emitter in direction p, i.e., (θ,φ)

dA_1 = differential area element of emitter, A_1

r = distance between element dA_1 and point p

$\hat{\mathbf{r}}$ = unit radial vector of the spherical coordinate system, origin at dA_1

θ = angle between surface normal at dA_1 and r

The minus sign reverses the direction of $\hat{\mathbf{r}}$ so that **dE** is a vector from point p to dA_1.

If $\mathbf{dA_1}$ is the differential axial vector having magnitude dA_1 and direction of the outward surface normal, then $\hat{\mathbf{r}}\circ\mathbf{dA_1} = -\cos(\theta)dA_1$ and

$$\mathbf{dE} = \frac{L(\theta,\varphi)\hat{\mathbf{r}}\circ\mathbf{dA_1}}{r^2}\hat{\mathbf{r}} \tag{2}$$

$$\mathbf{dE} = \frac{L(\theta,\varphi)\hat{\mathbf{r}}\hat{\mathbf{r}}}{r^2}\circ\mathbf{dA_1} \tag{3}$$

The quantity $\hat{\mathbf{r}}\hat{\mathbf{r}}$ is a dyad, with antecedent and consequent equal to the unit radial vector $\hat{\mathbf{r}}$. The product of $\hat{\mathbf{r}}\hat{\mathbf{r}}$ and the scalar field $L(\theta,\varphi)/r^2$ produces a dyad field, which we define as **L**, having directional components $\hat{\mathbf{r}}\hat{\mathbf{r}}$ and magnitude given by the scalar field $L(\theta,\varphi)/r^2$. That is,

$$\mathbf{L}(r,\theta,\varphi) = \hat{\mathbf{r}}\hat{\mathbf{r}}L(\theta,\varphi)/r^2 \tag{4}$$

The dyadic field has a magnitude varying with (r,θ,φ) and, when necessary to avoid ambiguity, will be indicated as $\mathbf{L}(r,\theta,\varphi)$; otherwise **L** suffices. It has units of candela/area/area. **L** is the dyadic field produced by the luminance of an emitter. The dot product of it and the axial vector of the emitter surface gives the resulting vector light field. That is, substituting **Equation 4** into **Equation 3** gives

$$\mathbf{dE} = \mathbf{L}\circ\mathbf{dA_1} \tag{5}$$

Since the dyad **L** is symmetric, the dot product with $\mathbf{dA_1}$ can be on either side of **L**. The total light vector produced at point p is obtained by integrating this dot product over the emitter:

$$\mathbf{E} = \int_{A_1} \mathbf{L} \circ \mathbf{dA_1} \tag{6}$$

Transformation to a contour integral

The area integral of **Equation 6** can be converted to a contour integral by the application of Stokes's theorem. We use a dyadic generalization of its more common vector form:

$$\int_A (\nabla \times \mathfrak{I}) \circ \mathbf{dA} = \oint_\Gamma \mathfrak{I} \circ \mathbf{d\gamma} \tag{7}$$

where $\mathbf{dA} = dA\,\hat{\mathbf{n}}$ = differential axial vector with magnitude equal to the differential element of area and direction equal to the outward directed normal of the surface

∇= partial differential operator, with $\nabla \times$ giving the curl

Γ = directed perimeter of surface A. Direction is that of the right-hand rule applied to the surface normal

$\mathbf{d\gamma}$ = differential vector with magnitude equal to differential element of the perimeter, Γ, and direction equal to the local tangent, and

$\mathfrak{I}$ = arbitrary dyadic field

There is a difficulty in applying this to **Equation 6**; the dyadic field of **Equation 7** must be fixed and single valued throughout the region of interest. This is not to say constant; but, given an arbitrary point on surface A, the value of $\mathfrak{I}$ at that point cannot depend on the orientation of A at that point. Concomitantly, the value of $\nabla\times\mathfrak{I}$ cannot depend on the orientation of A. In the present application, the surface generates the field, and thus, for a nondiffuse emitter, the field depends on the orientation of A at that point. Concomitantly, the value of cannot depend on the orientation of A. In the present application, the surface generates the field, and thus, for a nondiffuse emitter, the field depends on the orientation of the surface not just on its perimeter. To make the field fixed, single valued and dependent only on the perimeter of the emitting surface, it is necessary to fix the shape of the emitting surface. Though any shape is permitted, it is practical to assume the emitting surface is planar, making it simple to specify the field. If the emitting surface is planar than the coordinate system can have its origin at point p or at dA. In either case, the zenith direction (θ=0) is normal to the planar emitter.

The development is also simplified if we assume that the emitter is homogeneous, that is, $L(\theta,\varphi)$ is the luminance distribution everywhere on the planar surface of the emitter. We note that in the diffuse case, the field produced by the emitter is not a function of the emitter's shape, and fixing it is not necessary.

To transform **Equation 6** to a contour integral using **Equation 7**, it is necessary to find a dyadic, $\mathfrak{I}$, such that its curl yields the dyadic, **L**, defined by **Equation 4.** Thus, the required dyadic is defined by

$$\nabla \times \mathfrak{I} = \mathbf{L} = \hat{\mathbf{r}}\hat{\mathbf{r}}L(\theta,\varphi)/r^2 \tag{8}$$

If $\mathfrak{I}$ can be found, then **Equation 6** becomes

$$\mathbf{E} = \int_{A_1} \mathbf{L} \circ \mathbf{dA_1} = \int_{A_1} (\nabla\times\mathfrak{I}) \circ \mathbf{dA_1} = \oint_{\Gamma_1} \mathfrak{I} \circ \mathbf{d\gamma_1} \tag{9}$$

where $\nabla\times\mathfrak{I} = \mathbf{L} = \hat{\mathbf{r}}\hat{\mathbf{r}}L(\theta,\varphi)/r^2$.

To solve for $\mathfrak{I}$, we assume it has a form given by $\mathfrak{I} = \mathbf{a}\hat{\mathbf{r}} + \mathbf{b}\hat{\theta} + \hat{\varphi}$. That is, the dyadic is the sum of three dyads, each consisting of one of the unit vectors of the spherical coordinate system as the consequent, and a vector to be determined as the antecedent. The curl of $\mathfrak{I}$, in spherical coordinates, is

$$\nabla \times \mathfrak{I} = \begin{vmatrix} \hat{\mathbf{r}} & r\hat{\theta} & r\sin(\theta)\hat{\varphi} \\ \dfrac{\partial}{\partial r} & \dfrac{\partial}{\partial \theta} & \dfrac{\partial}{\partial \varphi} \\ \mathbf{a} & r\mathbf{b} & r\sin(\theta)\mathbf{c} \end{vmatrix} \frac{1}{r^2\sin(\theta)} \tag{10}$$

Note that the partial derivatives are of the vectors **a**, **b**, and **c**, and that the result of taking the curl of the dyadic is another dyadic. Expanding **Equation 10** and substituting into **Equation 8** gives

$$\frac{\hat{\mathbf{r}}\hat{\mathbf{r}}L(\theta,\varphi)}{r^2} = \hat{\mathbf{r}}\frac{1}{r\sin(\theta)}\left\{\frac{\partial}{\partial\theta}\sin(\theta)\mathbf{c} - \frac{\partial}{\partial\varphi}\mathbf{b}\right\} \tag{11}$$

$$+\hat{\theta}\frac{1}{r}\left\{\frac{1}{\sin(\theta)}\frac{\partial}{\partial\varphi}\mathbf{a} - \frac{\partial}{\partial r}r\mathbf{c}\right\} + \hat{\varphi}\frac{1}{r}\left\{\frac{\partial}{\partial r}r\mathbf{b} - \frac{\partial}{\partial\theta}\mathbf{a}\right\}$$

This must be solved for vectors **a**, **b**, and **c** which make any $\hat{\theta}$ or $\hat{\varphi}$ component of the dyadic on the right-hand side vanish. There must remain a single component equal $\hat{\mathbf{r}}\hat{\mathbf{r}}L(\theta,\varphi)/r^2$. Solutions are not obvious.

A general solution to **Equation 11** for the nondiffuse case can be obtained as follows. We recognize that only the first term on the right side can remain since the second and third can never yield dyads with as $\hat{\mathbf{r}}$ antecedent, as the left side demands. The second and third terms on the right hand side vanish if vectors **a** and **b** are null, and vector **c** has as one of its factors, $1/r$. Assuming this, **Equation 11** becomes

$$\frac{\hat{r}\hat{r}L(\theta,\varphi)}{r^2} = \hat{r}\frac{1}{r\sin(\theta)}\frac{\partial}{\partial\theta}\sin(\theta)\mathbf{c} \tag{12}$$

Assuming vectors **a** and **b** are null means the dyadic we seek will be incomplete. It remains to be seen whether the general solution will allow this; though we note that the dyad on the left is also incomplete. Taking the dot product with $\hat{r}$ of both sides of **Equation 12** we obtain

$$\hat{r}\circ\frac{\hat{r}\hat{r}L(\theta,\varphi)}{r^2} = \hat{r}\circ\hat{r}\frac{1}{r\sin(\theta)}\frac{\partial}{\partial\theta}\sin(\theta)\mathbf{c}$$

$$\hat{r}\frac{L(\theta,\varphi)\sin(\theta)}{r} = \frac{\partial}{\partial\theta}\sin(\theta)\mathbf{c} \tag{13}$$

We first determine a form of **c** which yields the correct direction: $\hat{r}$. Unlike Cartesian coordinates, the unit vectors of spherical coordinates can have nonzero derivatives. We note that the partial derivative in **Equation 13** is with respect to θ, and observe that

$$\frac{\partial}{\partial\theta}\hat{r} = \hat{\theta}\ ;\ \frac{\partial}{\partial\theta}\hat{\theta} = -\hat{r} \tag{14}$$

This suggests that if **c** has both $\hat{\theta}$ and $\hat{r}$ components then, when differentiated and subtracted, only an $\hat{r}$ component will remain. Assume that the $\hat{\theta}$ and $\hat{r}$ components have magnitudes derived from the same scalar field. Application of the operation $\frac{\partial}{\partial\theta}\sin(\theta)$ results in a vector with only an $\hat{r}$ component if

$$\mathbf{c} = \frac{\hat{r}\frac{\partial}{\partial\theta}\mathcal{L}(\theta,\varphi) - \hat{\theta}\mathcal{L}(\theta,\varphi)}{r\sin(\theta)} \tag{15}$$

where $\mathcal{L}(\theta,\varphi)$ is a scalar field to be determined. Substituting this definition for **c** gives

$$\frac{\partial}{\partial\theta}\sin(\theta)\mathbf{c} = \frac{\partial}{\partial\theta}\sin(\theta)\frac{\hat{r}\frac{\partial}{\partial\theta}\mathcal{L}(\theta,\varphi) - \hat{\theta}\mathcal{L}(\theta,\varphi)}{r\sin(\theta)}$$

$$= \left[\left(\frac{\partial}{\partial\theta}\hat{r}\right)\left(\frac{\partial}{\partial\theta}\mathcal{L}(\theta,\varphi)\right) + \hat{r}\left(\frac{\partial}{\partial\theta}\frac{\partial}{\partial\theta}\mathcal{L}(\theta,\varphi)\right) - \left(\frac{\partial}{\partial\theta}\hat{\theta}\right)\mathcal{L}(\theta,\varphi) - \hat{\theta}\left(\frac{\partial}{\partial\theta}\mathcal{L}(\theta,\varphi)\right)\right]\Big/ r$$

Substituting this result for $\frac{\partial}{\partial\theta}\sin(\theta)\mathbf{c}$ in **Equation 13** gives

$$\hat{r}\frac{\left(\mathcal{L}(\theta,\varphi) + \frac{\partial^2}{\partial\theta^2}\mathcal{L}(\theta,\varphi)\right)}{r} = \hat{r}\frac{L(\theta,\varphi)\sin(\theta)}{r}$$

Evidently the scalar field, $\mathcal{L}(\theta,\varphi)$, must be such that

$$\left(\mathcal{L}(\theta,\varphi) + \frac{\partial^2}{\partial\theta^2}\mathcal{L}(\theta,\varphi)\right) = L(\theta,\varphi)\sin(\theta).$$

Solving gives

$$\mathcal{L}(\theta,\varphi) = \int_0^\theta L(\theta',\varphi)\sin(\theta-\theta')\sin(\theta')d\theta' \tag{16}$$

Equations 15 and **16** define the vector **c**, which is the $\hat{\varphi}$–component of the dyad, $\mathfrak{I}$, which, in turn, satisfies **Equation 8**. Substituting into **Equation 9** allows the total light vector produced at point p by a planar nondiffuse emitter with luminance distribution $L(\theta,\varphi)$, to be expressed as a contour integral:

$$\mathbf{E} = \oint_{\Gamma_1}\frac{\hat{r}\hat{\varphi}\frac{\partial}{\partial\theta}\mathcal{L}(\theta,\varphi) - \hat{\theta}\hat{\varphi}\mathcal{L}(\theta,\varphi)}{r\sin(\theta)}\circ\mathbf{d}\gamma_1$$

$$= \oint_{\Gamma_1}\frac{\hat{r}\frac{\partial}{\partial\theta}\mathcal{L}(\theta,\varphi) - \hat{\theta}\mathcal{L}(\theta,\varphi)}{r\sin(\theta)}\hat{\varphi}\circ\mathbf{d}\gamma_1 \tag{17}$$

where $\mathcal{L}(\theta,\varphi) = \int_0^\theta L(\theta',\varphi)\sin(\theta-\theta')\sin(\theta')d\theta'$

$$\frac{\partial}{\partial\theta}\mathcal{L}(\theta,\varphi) = \int_0^\theta L(\theta',\varphi)\cos(\theta-\theta')\sin(\theta')d\theta'$$

Γ_1 = directed perimeter of emitting surface,
$\mathbf{d}\gamma_1$ = differential vector element of Γ_1
r = distance from $\mathbf{d}\gamma_1$, to point p
(θ,φ) = spherical coordinates with origin at p and zenith direction (θ=0) normal to the planar surface of the emitter

Examination of **Equation 17** shows that the total light vector, **E**, has components in the $\hat{r}$ and $\hat{\theta}$ directions. A diffuse emitter is the special case where $L(\theta,\varphi) = L$. The integrals for $\mathcal{L}(\theta,\varphi)$ and $\frac{\partial}{\partial\theta}\mathcal{L}(\theta,\varphi)$ have simple analytic expressions. Substituting them into **Equation 17** we obtain

$$\mathbf{E} = \frac{L}{2}\oint_{\Gamma_1}\frac{\hat{r}\theta - \hat{\theta}(1-\theta\cot(\theta))}{r}\hat{\varphi}\circ\mathbf{d}\gamma_1 \tag{18}$$

This formulation appears to be new and can be compared with the result obtained by Yamauti and others:[11,12]

$$\mathbf{E} = \frac{L}{2} \oint_{\Gamma_1} \frac{\hat{r} \times \mathbf{d}\gamma_1}{r}$$

Dyad formulation of total flux transfer

We now assume that the point p is on the surface of an area receiver. The light vector at this point is given by **Equation 17**. The flux incident on a differential element of area, dA_2, is

$$d\Phi = \hat{n} \circ \mathbf{E}\, dA_2,$$

where $\hat{n}$ is the unit vector, normal to dA_2.

Now $\hat{n} \bullet \mathbf{E}$ is, of course, the illuminance at point p, and its product with dA_2 gives the flux onto dA_2. We use the axial vector, $\mathbf{dA_2}$, and integration over the receiving area, A_2, to obtain the total flux onto the receiver:

$$\Phi = \int_{A_2} \mathbf{dA_2} \circ \mathbf{E}$$

Substituting for **E** from **Equation 17** gives,

$$\Phi = \int_{A_2} \mathbf{dA_2} \circ \oint_{\Gamma_1} \frac{\hat{r}\hat{\varphi}\frac{\partial}{\partial\theta}\mathcal{L}(\theta,\varphi) - \hat{\theta}\hat{\varphi}\mathcal{L}(\theta,\varphi)}{r\sin(\theta)} \circ \mathbf{d}\gamma_1 \tag{19}$$

$$\Phi = \oint_{\Gamma_1} \left\{ \int_{A_2} \mathbf{dA_2} \circ \frac{\hat{r}\hat{\varphi}\frac{\partial}{\partial\theta}\mathcal{L}(\theta,\varphi) - \hat{\theta}\hat{\varphi}\mathcal{L}(\theta,\varphi)}{r\sin(\theta)} \right\} \circ \mathbf{d}\gamma_1$$

Once again, we transform the area integral to a contour integral by finding the appropriate dyadic and invoking Stokes's theorem. We note that the field produced by the emitter in the region containing the receiver is fixed and no assumptions about the area of the receiver are necessary.

We seek a dyadic, $\aleph$, which satisfies the equation

$$\nabla \times \aleph = \frac{\hat{r}\hat{\varphi}\frac{\partial}{\partial\theta}\mathcal{L}(\theta,\varphi) - \hat{\theta}\hat{\varphi}\mathcal{L}(\theta,\varphi)}{r\sin(\theta)}$$

Let $\aleph = \mathbf{u}\hat{r} + \mathbf{v}\hat{\theta} + \mathbf{w}\hat{\varphi}$. Expanding $\nabla \times \aleph$ and substituting gives

$$\frac{\hat{r}\hat{\varphi}\frac{\partial}{\partial\theta}\mathcal{L}(\theta,\varphi) - \hat{\theta}\hat{\varphi}\mathcal{L}(\theta,\varphi)}{r\sin(\theta)} =$$

$$\hat{r}\frac{1}{r\sin(\theta)}\left\{\frac{\partial}{\partial\theta}\sin(\theta)\mathbf{w} - \frac{\partial}{\partial\varphi}\mathbf{v}\right\}$$

$$+\hat{\theta}\frac{1}{r}\left\{\frac{1}{\sin(\theta)}\frac{\partial}{\partial\varphi}\mathbf{u} - \frac{\partial}{\partial r}r\,\mathbf{w}\right\} + \hat{\varphi}\frac{1}{r}\left\{\frac{\partial}{\partial r}r\,\mathbf{v} - \frac{\partial}{\partial\theta}\mathbf{u}\right\}.$$

We recognize that the third term on the right side must vanish since it can never result in a dyad with antecedents of $\hat{r}$ or $\hat{\theta}$, which the left side demands. The third term vanishes if vectors **v** and **u** are null, leaving

$$\frac{\hat{r}\hat{\varphi}\frac{\partial}{\partial\theta}\mathcal{L}(\theta,\varphi) - \hat{\theta}\hat{\varphi}\mathcal{L}(\theta,\varphi)}{r\sin(\theta)} =$$

$$= \hat{r}\frac{1}{r\sin(\theta)}\left\{\frac{\partial}{\partial\theta}\sin(\theta)\mathbf{w}\right\} - \hat{\theta}\frac{1}{r}\left\{\frac{\partial}{\partial r}r\,\mathbf{w}\right\}$$

Considering the terms on the right hand side and noting that $\frac{\partial}{\partial r}\hat{\varphi}=0$ and $\frac{\partial}{\partial\theta}\hat{\varphi}=0$ reveals that $\mathbf{w} = \frac{\hat{\varphi}\mathcal{L}(\theta,\varphi)}{\sin(\theta)}$ is a solution, giving the remarkably simple expression for $\aleph$: $\aleph = \frac{\hat{\varphi}\hat{\varphi}\mathcal{L}(\theta,\varphi)}{\sin(\theta)}$.

Substituting in **Equation 19** and invoking the dyadic form of Stoke's theorem gives for the total flux transfer between the nondiffuse planar emitter, A_1, and the receiving surface A_2:

$$\Phi = \oint_{\Gamma_1}\oint_{\Gamma_2} \mathbf{d}\gamma_2 \circ \hat{\varphi}\hat{\varphi} \circ \mathbf{d}\gamma_1 \mathcal{L}(\theta,\varphi) \tag{20}$$

where

$$\mathcal{L}(\theta,\varphi) = \frac{1}{\sin(\theta)} \int_0^{\theta} L(\theta',\varphi)\sin(\theta-\theta')\sin(\theta')\,d\theta' \tag{21}$$

Note that the factor of $1/\sin(\theta)$ has been subsumed into the definition of $\mathcal{L}(\theta,\varphi)$. **Equations 20** and **21** give a complete solution to the general radiative transfer between surfaces. We note that the spherical coordinate system used in **Equation 20** has its zenith direction ($\theta=0$) normal to the planar emitter surface and its origin at $\mathbf{d}\gamma_1$.

Once again, diffuse emission is a special case and the scalar function $\mathcal{L}(\theta,\varphi)$ has an analytic form; we obtain

$$\Phi = \frac{L}{2} \oint_{\Gamma} \oint_{\Gamma} d\gamma_2 \circ \hat{\varphi}\hat{\varphi} \circ d\gamma_1 \left(1 - \theta \cot(\theta)\right). \tag{22}$$

This can be compared to the result obtained by Fock:[13]

$$\Phi = \frac{L}{2} \oint_{\Gamma_1} \oint_{\Gamma_2} d\gamma_2 \circ d\gamma_1 \ln(r) \tag{23}$$

We observe that Fock's result can be written as

$$\Phi = \frac{L}{2} \oint_{\Gamma_1} \oint_{\Gamma_2} d\gamma_2 \circ \left(\hat{r}\hat{r} + \hat{\theta}\hat{\theta} + \hat{\varphi}\hat{\varphi}\right) \circ d\gamma_1 \ln(r),$$

since $(\hat{r}\hat{r}+\hat{\theta}\hat{\theta}+\hat{\varphi}\hat{\varphi})\circ d\gamma_1 = (\hat{r}\hat{r}+\hat{\theta}\hat{\theta}+\hat{\varphi}\hat{\varphi})\circ(\hat{r}d\gamma_r+\hat{\theta}d\gamma_\theta+\hat{\varphi}d\gamma_\varphi) = d\gamma_1$. The special dyadic, $\hat{r}\hat{r}+\hat{\theta}\hat{\theta}+\hat{\varphi}\hat{\varphi}$, is referred to as the idemfactor, and functions as 1 does in ordinary algebra. The dyadic fields of **Equations 22** and **23**, $\hat{\varphi}\hat{\varphi}$ $[1-\cot(\theta)]$ and $(\hat{r}\hat{r}+\hat{\theta}\hat{\theta}+\hat{\varphi}\hat{\varphi})\ln(r)$, produce the same double contour integral values.

Application

One obvious application of these results is the calculation of the flux input from luminaires to discrete surfaces. In most applications, a far field candela distribution for a luminaire is available. From this, the average luminance distribution can be determined:

$$L(\theta,\varphi) = I(\theta,\varphi)/ A \cos(\theta)$$

where A = luminous area of the luminaire

$I(\theta,\varphi)$ = far-field candela distribution

This luminance distribution function is then substituted into **Equation 21**. In all practical cases, a functional form for $L(\theta,\varphi)$ is not available; only an array of values determined from the array of far-field candela values resulting from photometric measurement. This means that a set of discrete values of the scalar field $\mathcal{L}(\theta,\varphi)$ will be used. In addition, the double contour integrals of **Equation 20** will be approximated by a double summation over discrete pieces of the emitter's and receiver's contour.

$$\Phi = \sum_{i=1}^{M} \sum_{j=1}^{N} \Delta\Gamma_{1i} \circ \hat{\varphi}\hat{\varphi} \circ \Delta\Gamma_{2j}\, \mathcal{L}\left(\theta_{ij},\varphi_{ij}\right) \tag{24}$$

where M = total number of discrete pieces of emitter contour

N = total number of discrete pieces of receiver contour

$\Delta\Gamma_{1i}$ = discrete vector formed by ith piece of contour of emitter with direction given by the right hand rule,

$\Delta\Gamma_{2j}$ = discrete vector formed by jth piece of contour of receiver with direction given by the right hand rule,

$(\theta_{ij},\varphi_{ij})$ = spherical coordinates of jth piece of receiver contour, with respect to the ith piece of emitter contour

$\mathcal{L}(\theta,\varphi)$ = scalar function given by

$$\mathcal{L}(\theta,\varphi) = \frac{1}{\sin(\theta)} \int_0^{\theta} \frac{I(\theta',\varphi)\sin(\theta-\theta')\sin(\theta')}{A\cos(\theta')} d\theta' \tag{25}$$

or approximately

$$\mathcal{L}(\theta_j,\varphi_k) = \frac{1}{\sin(\theta_j)} \sum_{i}^{N_j} \frac{I(\theta_i',\varphi_k)\sin(\theta_j-\theta_i')\sin(\theta_i')}{A\cos(\theta_i')} \Delta\theta \tag{26}$$

where $N_j = \theta_j/\Delta\theta$ = number of elements in the sum for θ_j. If a convenient interpolating function is built from the values of $I(\theta,\varphi)\sin(\theta)/\cos(\theta)$, then the integration of **Equation 25** can be performed analytically and a better approximation than that of **Equation 26** is obtained. A convenient example of this is to use a finite Fourier series to functionalize the discrete data available for $I(\theta,\varphi)\sin(\theta)/\cos(\theta)$.[14] That is,

$$I(\theta,\varphi)\sin(\theta)/\cos(\theta) = \sum_{m=1}^{M}\sum_{n=0}^{N} a_{mn}\cos(m\theta)\cos(n\varphi)$$

where M and N depend on the number of discrete values of $I(\theta,\varphi)$, and the coefficients amn are calculated from values of $I(\theta,\varphi)\sin(\theta)/\cos(\theta)$. We have a simple continuous function allowing analytic integration to obtain the function $\mathcal{L}(\theta,\varphi)$.

Equation 25 becomes

$$\mathcal{L}(\theta,\varphi) = \frac{1}{A\sin(\theta)} \int_0^{\theta} \left\{\sum_{m=1}^{M}\sum_{n=0}^{N} a_{mn}\cos(m\theta')\cos(n\varphi)\right\} \bullet \sin(\theta-\theta')\, d\theta'$$

Term-by-term integration is allowed, and we obtain

$$\mathcal{L}(\theta,\varphi) = \frac{-2}{A\sin(\theta)} \sum_{m=1}^{M}\sum_{n=0}^{N} a_{mn} \bullet \frac{\left[\sin\left(\left(\frac{m+1}{2}\right)\theta\right)\sin\left(\left(\frac{m-1}{2}\right)\theta\right)\right]}{\left(m^2-1\right)}\cos(n\varphi) \tag{27}$$

Convergence is rapid and significantly fewer than M and N terms are sufficient. **Equation 27** is an accurate way to build a table of values of $\mathcal{L}(\theta,\varphi)$ for use in **Equation 24**. Of course, this computation need be done only once. This procedure can be applied to any luminance distribution.

Computational economy

The computational utility of **Equation 24** is demonstrated by application to two axially symmetric luminance distributions:

$$L_1(\theta,\varphi) = L_n \cos^3(\theta)$$

$$L_2(\theta,\varphi) = L_n \cos^{-2}(\theta)$$

These correspond to narrow and wide candela distributions, respectively. Use of $L_2(\theta,\varphi)$ is limited to θ less than $\pi/2$. L_n is set equal to 1.0 and **Equation 21** yields

$$\mathcal{L}_1(\theta,\varphi) = \frac{1}{15}\left(3 - 2\cos(\theta) - \cos^3(\theta)\right)$$

$$\mathcal{L}_2(\theta,\varphi) = \left(-1 + \cot(\theta) - \ln(\cos(\theta))\right)$$

The geometry used for these calculations consists of two parallel squares, separated by three distances; the first equal to three times the size of the squares, the second equal to the size of the squares, and the third equal to one-third the size of the squares. The flux transferred is calculated two ways: using **Equation 24** and the values of $\mathcal{L}_1(\theta,\varphi)$ and $\mathcal{L}_2(\theta,\varphi)$, and the other using a finite summation approximation to the double area integrals. For both computational procedures, values for the flux transferred is calculated as a function of the number of discrete element evaluations used. The evaluations are approximately the same amount of computational work for the two procedures consisting of table look-ups for values of either $\mathcal{L}_1(\theta,\varphi)$ or $\mathcal{L}_2(\theta,\varphi)$, or $I(\theta,\varphi)$. **Figures 2, 3,** and **4** show the results for the $L_n\cos^3(\theta)$ distribution at the three separating distances. **Figures 5, 6,** and **7** show the results for the $L_n\cos^{-2}(\theta)$ distribution at the three separating distances. The superior convergence properties of the contour integration method are apparent.

References

1. Rea, M.S. 1993. Editor, *Illuminating Engineering Society Lighting Handbook,* eighth edition. New York: Illuminating Engineering Society of North America.

2. DiLaura, D.L. and Quinlan, J. 1995. Nondiffuse radiative transfer I: planar area sources and point receivers. *J of the IES* 24(no. 2.):102–112.

3. Fock, V.A. 1924. Zur Berecnung der Belechtungsstarke. *Z. Physik* 27:102.

4. Florence, N. and DiLaura, D.L. 1993. On the accuracy of computer predicted luminaire photometric performance. *J of the IES* 22(no. 2): 8–15.

5. Spencer, D.E. 1964. Tensors in illuminating engineering. *Illum Eng* 59:345-356.

6. Moon, P. and Spencer, D.E. 1981. The photic field. Cambridge: MIT Press.

7. Morse, P.M. and Feshbach, H. 1953. *Methods of theoretical physics.* New York: McGraw-Hill.

8. Gibbs, J.W. and Wilson, E.B. 1909, 1960. *Vector analysis.* New York: Dover.

9. Danielson, D.A. 1992. *Vectors and tensors in engineering and physics.* New York: Addison Wesley.

10. Moon, P. 1961. *The scientific basis of illuminating engineering.* New York: Dover.

11. Yamauti, Z. 1932. *Theory of field of illumination, researches of the Electrotechnical Laboratory,* No. 339. Tokyo: Ministry of Communications.

12. Gershun, A. 1939. The light field. *J Mathematical Physics* 18:51.

14. DiLaura, D.L. 1975. On the computation of equivalent sphere illumination, *J of the IES* 4(no. 2)129–149.

Discussion

Whenever I see a paper with "Part 1" in the title, I always wonder whether there will ever be a "Part 2." I am pleased to see that in this case, the author has removed any doubt. "Nondiffuse Radiative Transfer"[a] was a welcome and impressive extension to V.A. Fock's contour integration procedure for calculating illuminance at a point from an ideal diffuse area source.[b] Nondiffuse Radiative Transfer 2" completes the work with yet another mathematical *tour de force.*

The problem of calculating the radiative flux transferred between two arbitrary polygons was first addressed by Johann Lambert in 1760. It took over 230 years to solve the problem for ideal diffuse sources. Schroder reports that it took Mathematica 15 mins to solve 90 percent of the problem, followed by nine months of intensive mathematics research to fill in the details of the remaining 10 percent.[c] It also took 15 pages of Mathematica source code to express the resultant algorithm.

In comparison, the author has developed an elegant algorithm for nondiffuse polygons (and which evidently subsumes Schroder's work) that can be expressed in a single equation.

I am not ashamed to admit that I was unable to follow the mathematical development in it entirety. I have also not had the time to program the algorithm, although I certainly hope to do so in the near future. From the results present in this paper, however, it certainly appears to be a practical and time-efficient approach.

It is doubtful whether the author's procedure will find practical use in day-to-day lighting calculations. On the other hand, it should prove very useful to the computer graphics community, where radiative transfer techniques are used to create photorealistic and photometrically accurate images. One problem in particular comes to mind: cove lighting. Accurately modeling this type of lighting in an architectural rendering typically requires that the wall and ceiling surfaces be finely discretized where they meet. The author's procedure resolves this issue by allowing coarser discretization of the surfaces, and consequently faster execution times for the rendering program. More to the point, these programs will eventually replace our pocket-calculator style of lighting calculations.

The situation in complex environments is more problematic. As the author noted in his previous paper, the procedure is only applicable where the illuminated surface does not "cut" (i.e., partially occlude) the emitter. In a complex environment, of course, this must be generalized to include any other intervening surface or object.

To quote my discussion of the author's previous paper:

> There are practical solutions to this problem. An efficient "area subdivision" algorithm such as Warnock's algorithm can be used to determine the precise extent to which objects in an environment partially occlude the emitter.[d] Relevant examples of the successful use of Warnock's algorithm in complex environments include Nishita and Nakamae and Sillion and Puech.[e,f] In other words, the author's procedure forms the basis of an eminently practical technique for calculating illuminance or irradiance in computer graphics, lighting design, and thermal engineering applications.

Area subdivision algorithms assume a specific viewpoint in the environment, which in this case translates into a point receiver. It is unclear whether these algorithms can be generalized to model the occlusion of a finite surface area—an interesting challenge for a graduate student in computer graphics or illumination engineering perhaps? There may even be relevant ideas to be derived from this paper.

In presenting his paper last year, the author lamented that the leadership role in theoretical lighting research had been assumed by the computer graphics community. I am pleased to say that this paper has unquestionably reclaimed that title for illumination engineering.

I. Ashdown,
Ledalite Architectural Products, Inc.

References

a. DiLaura, D.L. and Qinlan, J. 1995. Nondiffuse radiative transfer 1: Planar area sources and point receivers. *J of the IES* 24(no. 2):102–112.

b. Fock, V.A. 1924. Illumination from surfaces of arbitrary shape. *Trudy GOI* No. 28 (in Russian).

c. Schroder, P. and Hanrahan, P. 1993. On the form factor between two polygons, *Computer Graphics Proceedings.*

d. Sutherland, I.E.; Sproull, R.F.; and Schumacker, R. 1974. A characterization of ten hidden-surface algorithms. *Computing Surveys* 6(no. 1):1–55.

e. Mishita, T. and Nakamea, E. 1984. Continuous tome representation of three-dimension objects taking account shadows and interreflections. *Computer Graphics* 20(no. 4):125-132

f. Sillion, F.X. and Puech, C. 1989. A general two-pass method integrating specular and diffuse reflection. *Computer Graphics* 23(no. 3): 335–355.

The author has once again shown how advanced mathematics can be applied to a complex lighting situation to achieve a difficult and elusive result. This paper offers a very thorough, although highly complex, derivation of this procedure for the reader to follow. While the graphs provided in the paper shown the computational economy in terms of number of discrete elements, it is unclear where these discrete elements are being applied. Are they on the receiving or sending surface?

One of the primary concerns in applying this technique appears to be the fact that the procedure assumes that the light field is uniform across the receiving surface. In the case of a calculation for orthogonal surfaces, which are very close to each other, such as in a corner, the convergence situation is likely to change significantly from what occurs when the two surfaces are parallel. Has the author attempted to apply this approach to other than parallel surfaces?

It is interesting to note that the number of discrete elements which are required to get a close approximation is large when the elements are farther away and smaller when the elements are closer. One would expect this condition to peak at a particular distance, then fall off again at larger distances. Is this what occurs?

R.G. Mistrick
Department of Architectural Engineering
The Pennsylvania State University

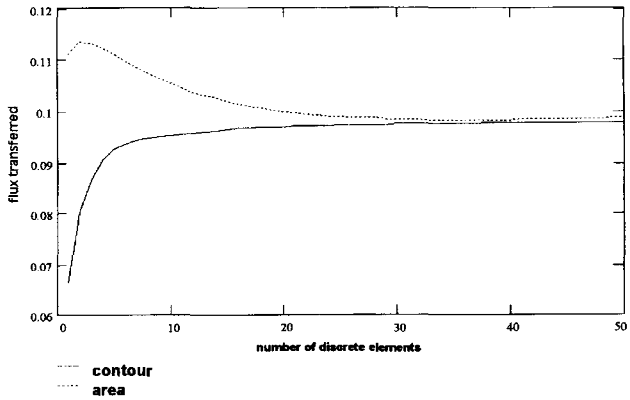

Figure 2—Flux transfer for $L_n \cos^3(\theta)$; separating distance =3

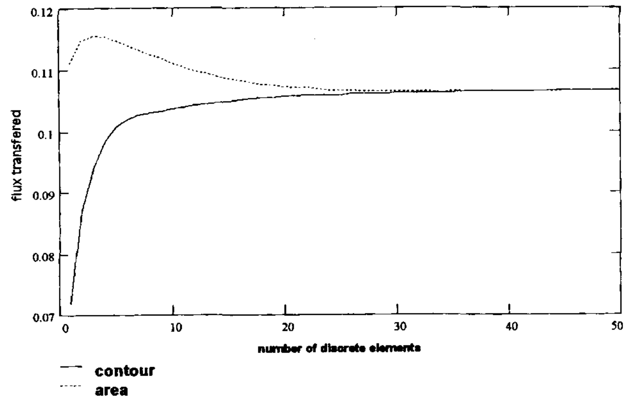

Figure 5—Flux transfer for $L_n \cos^2(\theta)$; separating distance =3

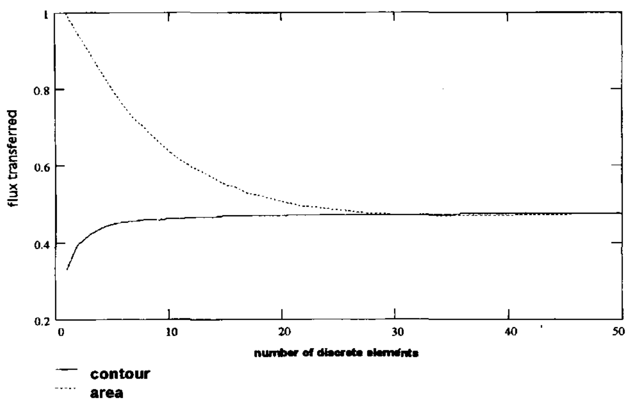

Figure 3—Flux transfer for $L_n \cos^3(\theta)$; separating distance =1

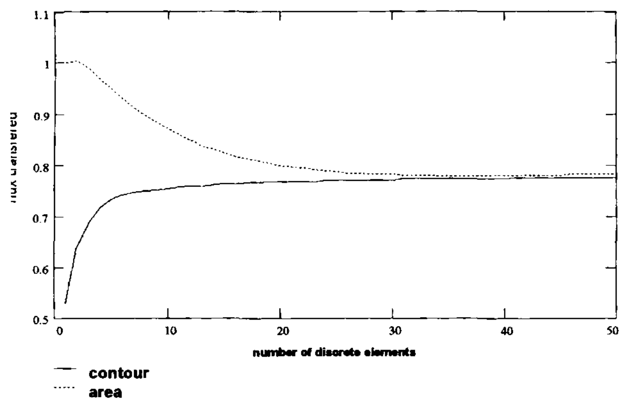

Figure 6—Flux transfer for $L_n \cos^2(\theta)$; separating distance =1

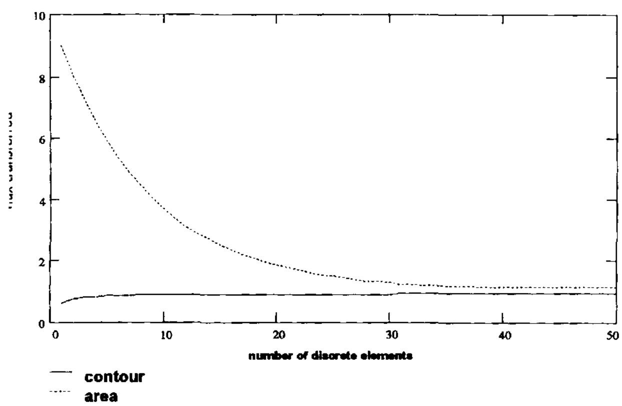

Figure 4—Flux transfer for $L_n \cos^3(\theta)$; separating distance =1/3

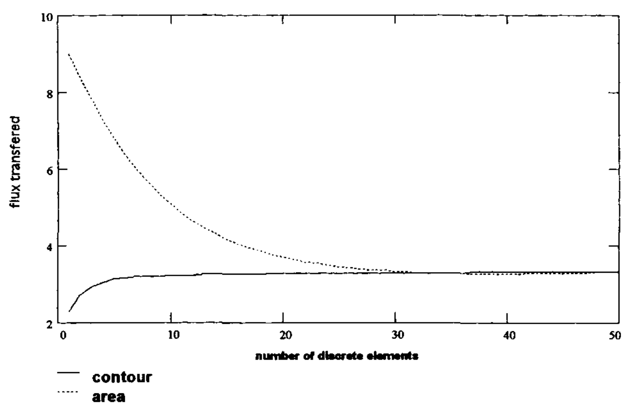

Figure 7—Flux transfer for $L_n \cos^2(\theta)$; separating distance =1/3

Author's response

I. Ashdown correctly identifies the most difficult problem in the implementation of this new development; the effect of occluding objects between surfaces exchanging flux. The references he sights are adequate for solving this problem from a single viewing point; but the elegant solution for the full three-dimension problem remains elusive. The author feels that the light-field approach to this problem will be very fruitful. This will shift the emphasis from what is happening at the surfaces of the radiative transfer system, to what is happening in the interstices between them.

R. Mistrick asks about the use of the discrete element. The element counts shown on the abscissae of the **Figures 2–7** are total numbers of discrete element used; that is, the sum of the number of elements on the receiving and sending surfaces. Regarding convergence, it is the author's experience that convergence appears to be governed by not only the proximity of the receiving surface, but its size and the method of approximating the function referred to as $\mathcal{L}$ in the paper. Convergence is not monotonic and the position of the peak of element count does, as suggested, depend on the separating distance of the emitting and receiving surfaces.