Rigorous near- to far-field transformation for vectorial diffraction calculations and its numerical implementation

Peter Török and Peter R. T. Munro

Department of Physics, Imperial College London, Prince Consort Road, London SW7 2BW, UK

Emmanouil E. Kriezis

Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Thessaloniki 54 124, Greece

Received May 19, 2005; accepted July 10, 2005

A rigorous method for transforming an electromagnetic near-field distribution to the far field is presented. We start by deriving a set of self-consistent integral equations that can be used to represent the electromagnetic field rigorously everywhere in homogeneous space apart from the closed interior of a volume encompassing all charges and sinks. The representation is derived by imposing a condition analogous to Sommerfeld's radiation condition. We then examine the accuracy of our numerical implementation of the formula, also on a parallel computer cluster, by comparing the results with a case when the analytical solution is also available. Finally, an application example is shown for a nonanalytical case. © 2006 Optical Society of America

OCIS codes: 050.1960, 260.2110, 260.5430, 000.4430.

1. INTRODUCTION

Since the publication of the now famous paper by Stratton and Chu,[1] a number of papers have been published that chose the Stratton–Chu integral as starting point.[2–6] However, most of these works concentrated on diffraction and/or focusing problems. The Stratton–Chu formula has been derived for the specific case when all the sources of electromagnetic radiation are situated outside a closed volume. It then provides a rigorous solution for the boundary value problem when the field on the surface of the volume is known and one seeks to determine the value of the field within the volume.

In computational electromagnetics one frequently faces the problem of having to calculate the field in a rigorous manner either far from the object that causes scattering or on a surface that is much larger in lateral dimensions than the wavelength. In certain specific cases it is possible to obtain analytical solutions so that the above problem does not occur. However, in numerical methods, such as the finite difference time domain (FDTD)[7] method or the finite element method (FEM),[8] this is not possible even if a supercomputer is used, due to memory and computational time restrictions.

When numerical methods are used to solve a problem, both the electric and magnetic fields are usually known either anywhere within a volume or on the closed surface of it, and one faces the task of propagating this field to an arbitrary position in space. Since the Stratton–Chu formula is an exact representation, it would in principle be possible to use it to obtain the field from the numerical data. However, since one is required to integrate the field on a closed surface, the Stratton–Chu formula is not directly applicable. Hence one needs to find an alternative representation when the integration may be carried out over the closed surface encompassing *all* sources of electromagnetic radiation.

In this paper we first derive such a formula. We then use numerical integration to demonstrate the usefulness of the results. Finally, the conclusions are drawn. Note that throughout this paper the $\exp(-i\omega t)$ sign convention and the SI system are used.

It is important to point out that even though we refer to our method as "near- to far-field transformation," the solution equally applies to propagation to near and intermediate fields. Conversely, the "usual" transformation used in conjunction with the FDTD method[7] is strictly a near- to far-field transform, where the necessary approximations have been made to the surface integrals.

2. THEORY

Consider a volume V with its closed boundary S. $\mathfrak{M}(\mathbf{r})$ and $\mathfrak{N}(\mathbf{r})$ are two vector functions that are, together with their first derivatives, continuous on S and within V, and $\mathbf{r}=(x,y,z)$. From the divergence theorem of Gauss we have

$$\int\int\int_V \nabla\cdot\mathfrak{B}\,\mathrm{d}V = \oiint_S \mathfrak{B}\cdot\hat{\mathbf{m}}\,\mathrm{d}S, \tag{1}$$

where $\mathfrak{B}(\mathbf{r})$ is any well-behaved vector function and $\hat{\mathbf{m}}$ is

the *outward* surface normal. With the substitution $\mathfrak{B} = \mathfrak{M} \times (\nabla \times \mathfrak{N})$ one obtains

$$\int\int\int_V \nabla \cdot [\mathfrak{M} \times (\nabla \times \mathfrak{N})] \mathrm{d}V = \oiint_S [\mathfrak{M} \times (\nabla \times \mathfrak{N})] \cdot \hat{\mathbf{m}} \mathrm{d}S. \tag{2}$$

This yields the vector analog of Green's first identity:

$$\int\int\int_V [(\nabla \times \mathfrak{M}) \cdot (\nabla \times \mathfrak{N}) - \mathfrak{M} \cdot \nabla \times (\nabla \times \mathfrak{N})] \mathrm{d}V = \oiint_S [\mathfrak{M} \times (\nabla \times \mathfrak{N})] \cdot \hat{\mathbf{m}} \mathrm{d}S. \tag{3}$$

If we now interchange $\mathfrak{M}$ and $\mathfrak{N}$ in Eq. (3) and subtract the resulting equation from Eq. (3), we obtain

$$\int\int\int_V [\mathfrak{N} \cdot \nabla \times (\nabla \times \mathfrak{M}) - \mathfrak{M} \cdot \nabla \times (\nabla \times \mathfrak{N})] \mathrm{d}V = \oiint_S [\mathfrak{M} \times (\nabla \times \mathfrak{N}) - \mathfrak{N} \times (\nabla \times \mathfrak{M})] \cdot \hat{\mathbf{m}} \mathrm{d}S, \tag{4}$$

which is the vector analog of Green's second identity. These two integral theorems will be the basis of further derivation. Let us assume that the homogeneous and isotropic medium filling V is nonconducting and that it contains no charge or current within its interior. We now set[1] $\mathfrak{M} = \mathbf{E}$ and $\mathfrak{N} = G\mathfrak{a}$ in Eq. (4) where $\mathbf{E}$ represents the electric field, G is the scalar function required to satisfy the scalar wave equation on and within S, and $\mathfrak{a}$ is an arbitrary constant unit vector. If we substitute for $\mathfrak{M}$ and $\mathfrak{N}$ in the left hand side of Eq. (4), the kernel of the integral reads as

$$\begin{aligned} G\mathfrak{a} \cdot [\nabla \times (\nabla \times \mathbf{E})] - \mathbf{E} \cdot \{\nabla \times [\nabla \times (G\mathfrak{a})]\} &= k^2 G\mathfrak{a} \cdot \mathbf{E} - \mathbf{E} \cdot [k^2 G\mathfrak{a} + \nabla(\mathfrak{a} \cdot \nabla G)] \\ &= -\mathbf{E} \cdot [\nabla(\mathfrak{a} \cdot \nabla G)] = -\nabla \cdot [(\mathfrak{a} \cdot \nabla G)\mathbf{E}], \end{aligned}$$

with k being the wavenumber and where we used that $\nabla \cdot \mathbf{E} = 0$. Hence the left hand side of Eq. (4) now reads as

$$-\int\int\int_V \nabla \cdot [(\mathfrak{a} \cdot \nabla G)\mathbf{E}] \mathrm{d}V = -\mathfrak{a} \cdot \oiint_S (\mathbf{E} \cdot \hat{\mathbf{m}}) \nabla G \mathrm{d}S, \tag{5}$$

which follows from the divergence theorem [Eq. (1)]. The first and the second terms on the right hand side integrand of Eq. (4) give

$$[\mathbf{E} \times (\nabla G \times \mathfrak{a})] \cdot \hat{\mathbf{m}} = \mathfrak{a} \cdot [\nabla G \times (\mathbf{E} \times \hat{\mathbf{m}})], \tag{6}$$

$$[G\mathfrak{a} \times (\nabla \times \mathbf{E})] \cdot \hat{\mathbf{m}} = \mathfrak{a} \cdot [(\nabla \times \mathbf{E}) \times \hat{\mathbf{m}}]G = i\omega\mu\mathfrak{a} \cdot (\mathbf{H} \times \hat{\mathbf{m}})G, \tag{7}$$

where we have used Maxwell's equation to relate $\mathbf{E}$ and $\mathbf{H}$. It is seen that $\mathfrak{a}$ is a factor common to all terms in Eqs. (5)–(7). Hence, if the results are gathered from these expressions and substituted into Eq. (4), then, as $\mathfrak{a}$ is arbitrary, it is a straightforward matter to write

$$\oiint_S [i\omega\mu(\hat{\mathbf{m}} \times \mathbf{H})G + (\hat{\mathbf{m}} \times \mathbf{E}) \times \nabla G + (\hat{\mathbf{m}} \cdot \mathbf{E}) \nabla G] \mathrm{d}S = \mathbf{0}. \tag{8}$$

Hitherto we introduced the function G and required that it be a solution of the scalar wave equation. For practical cases we set G to be the free-space Green's function:

$$G = \frac{\exp(ikr)}{r}, \tag{9a}$$

where

$$r = |r_s - r_p| = \sqrt{(x - x_p)^2 + (y - y_p)^2 + (z - z_p)^2}, \tag{9b}$$

with $\mathbf{r}_s = (x, y, z)$ being the coordinates of an infinitesimal surface element $\mathrm{d}S$ on S and $\mathbf{r}_p = (x_p, y_p, z_p)$ being the (fixed) observation point (see Fig. 1).

It has been assumed for the derivation of Eq. (4) that both $\mathfrak{M}$ and $\mathfrak{N}$ are continuous and that V is homogeneous. As clearly shown by Eq. (9a), we encounter a singularity in G for $r = 0$, i.e., when the observation point coincides with the boundary. Furthermore, there must be sources of electromagnetic radiation located somewhere in space to give rise to diffraction. Thus particular attention needs to be paid to the definition of S. Let us first shift the origin of the Cartesian coordinate system to P and circumscribe the point P by a small sphere of surface S^i and radius δ, as shown in Fig. 2. Note that due to the linearity of Maxwell's equations the choice of the origin of the coordinate system is arbitrary. We also circumscribe *all* sources and sinks in space by a volume S^{iii}. Let us now construct a large spherical surface S^v, centered on P, of radius ϵ and connect S^i and S^{iii} by a small diameter tube of surface S^{ii} and S^{iii} and S^v by another small diameter tube of surface S^{iv}. The original surface S is now made up of elementary surfaces $S^i, \ldots, S^v$ resulting in a piecewise smooth orientable and simply connected surface. The originally *outward* normal $\hat{\mathbf{m}}$ of S is now an *inward* nor-

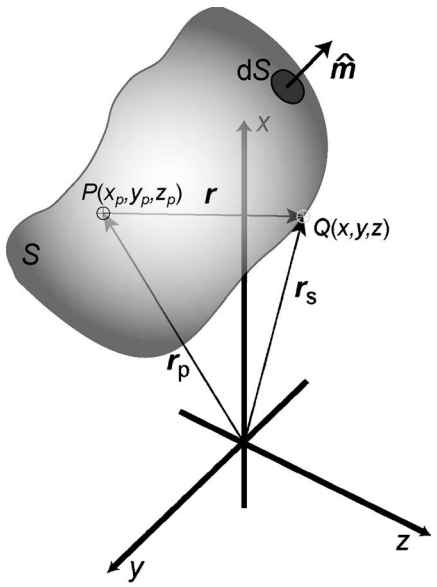

Fig. 1. Geometry of the problem showing the closed surface of the integration, the surface normal, the integral surface element, and the observation point.

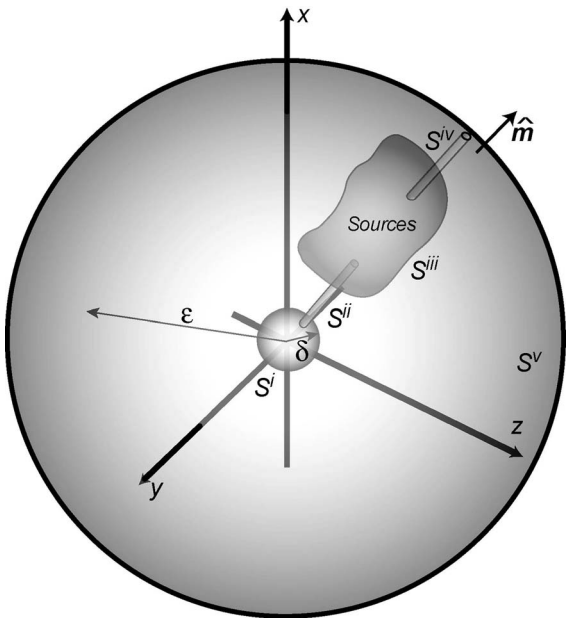

Fig. 2. Closed surface of integration for the first version of the Stratton–Chu formula.

mal for $S^i,\ldots,S^{iv}$ and an outward normal for S^v. Consequently, we have from Eq. (8) that

$$\int\int_{S^i}[\cdots]\mathrm{d}S^i+\int\int_{S^{ii}}[\cdots]\mathrm{d}S^{ii}+\int\int_{S^{iii}}[\cdots]\mathrm{d}S^{iii}+\int\int_{S^{iv}}[\cdots]\mathrm{d}S^{iv}+\int\int_{S^v}[\cdots]\mathrm{d}S^v=\mathbf{0}. \tag{10}$$

If the diameters of tubes S^{ii} and S^{iv} are made infinitesimally small, then the contribution from these surfaces becomes vanishingly small; hence we may write

$$-\oiint_{S^i}[\cdots]\mathrm{d}S^i-\oiint_{S^{iii}}[\cdots]\mathrm{d}S^{iii}+\oiint_{S^v}[\cdots]\mathrm{d}S^v=\mathbf{0}, \tag{11}$$

noting that now the integrals are again closed because S^i, S^{iii}, and S^v are separate closed surfaces. Also note that the signs of the first and second terms have been changed, so that for all surfaces $\hat{\mathbf{m}}$ is an *outward* normal. We first concentrate on the integration performed over S^i. If we refer to Figs. 1 and 2, it is clear that, since the sphere was circumscribed about P, $\hat{\mathbf{r}}=\hat{\mathbf{m}}$. Whence for S^i we have

$$-\oiint_{S^i}\left[i\omega\mu(\hat{\mathbf{m}}\times\mathbf{H})G+G\left(ik-\frac{1}{r}\right)(\hat{\mathbf{m}}\times\mathbf{E})\times\hat{\mathbf{m}}+G\left(ik-\frac{1}{r}\right)(\hat{\mathbf{m}}\cdot\mathbf{E})\hat{\mathbf{m}}\right]\mathrm{d}S^i. \tag{12}$$

Since, on the surface S^i, $\mathrm{d}S^i=\delta^2\mathrm{d}\Omega$, with Ω being a solid angle, Eq. (12) gives the following for a vanishing radius:

$$\lim_{\delta\to 0}-\oiint_{\Omega}\left[i\omega\mu(\hat{\mathbf{m}}\times\mathbf{H})\frac{\exp(ikr)}{r}+\mathbf{E}\frac{\exp(ikr)}{r}\left(ik-\frac{1}{r}\right)\right]_{r=\delta}\delta^2\mathrm{d}\Omega=4\pi\mathbf{E}. \tag{13}$$

We now turn to discuss the integral over S^v. In a manner similar to that before, we have

$$\oiint_{S^v}\left[i\omega\mu(\hat{\mathbf{m}}\times\mathbf{H})\frac{\exp(ikr)}{r}+\mathbf{E}\frac{\exp(ikr)}{r}\left(ik-\frac{1}{r}\right)\right]\mathrm{d}S^v=\oiint_{S^v}\left\{[\hat{\mathbf{m}}\times(\boldsymbol{\nabla}\times\mathbf{E})+ik\mathbf{E}]\frac{1}{r}-\mathbf{E}\frac{1}{r^2}\right\}\exp(ikr)\mathrm{d}S^v, \tag{14}$$

which can be written for integration over the solid angle Ω as

$$\int\int_{\Omega}\{[\hat{\mathbf{m}}\times(\boldsymbol{\nabla}\times\mathbf{E})+ik\mathbf{E}]r-\mathbf{E}\}\exp(ikr)\mathrm{d}\Omega=\int\int_{\Omega}\{[\mathbf{r}\times(\boldsymbol{\nabla}\times\mathbf{E})+ikr\mathbf{E}]-\mathbf{E}\}\exp(ikr)\mathrm{d}\Omega, \tag{15}$$

with the function values taken on the surface of the sphere.

Because the surface S^v limits the volume within which our calculation of the electromagnetic field is valid, we would like to extend this volume such that it encompasses the entire space, that is, $\mathbf{r}\to\infty$. Since S^{iii} is a surface encompassing all sources and sinks of electromagnetic radiation and the surface S^v is merely a mathematical construction, it is not physically tenable to expect its presence to perturb the field at P. Consequently, as $\epsilon\to\infty$, the contribution of Eq. (15) to Eq. (11) should be vanishingly small. For this to happen we must have

$$|r\mathbf{E}|<K, \tag{16a}$$

$$\mathbf{r}\times(\boldsymbol{\nabla}\times\mathbf{E})+ikr\mathbf{E}\to 0 \tag{16b}$$

as $r\to\infty$, with K being an arbitrary finite constant. The conditions given by relations (16) shall be termed the *vectorial* radiation condition in analogy to the *scalar* radiation condition of Sommerfeld.[9] Note that the second condition is identical to that of Nédélec.[10]

After substituting from Eq. (13) into Eq. (11), Eq. (8) the electric field is given on the *open exterior* of S^{iii} by

$$\mathbf{E}(x_p,y_p,z_p)=\frac{1}{4\pi}\oiint_{S^{iii}}[i\omega\mu(\hat{\mathbf{m}}\times\mathbf{H})G+(\hat{\mathbf{m}}\times\mathbf{E})\times\nabla G+(\hat{\mathbf{m}}\cdot\mathbf{E})\nabla G]\mathrm{d}S^{iii}. \tag{17a}$$

Due to the symmetry of Maxwell's equations the substitution $\mathbf{E}\to\mu\mathbf{H}$ and $\mathbf{H}\to-\epsilon\mathbf{E}$ yields the following for the magnetic field:

$$\mathbf{H}(x_p,y_p,z_p)=-\frac{1}{4\pi}\oiint_{S^{iii}}[i\omega\epsilon(\hat{\mathbf{m}}\times\mathbf{E})G-(\hat{\mathbf{m}}\times\mathbf{H})\times\nabla G-(\hat{\mathbf{m}}\cdot\mathbf{H})\nabla G]\mathrm{d}S^{iii}. \tag{17b}$$

This formula is different from that published in the literature,[1,11] where workers defined the problem slightly differently, as we shall now discuss.

Consider the volume shown in Fig. 3. The surface S is now being put together as the surface of the sphere surrounding the point of observation P, denoted by S^i; the

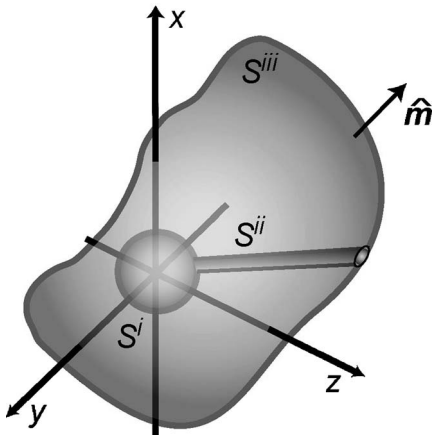

Fig. 3. Closed surface of integration for the second version of the Stratton–Chu formula.

small diameter tube of surface S^{ii}; and, finally, the surface of the outer shell S^{iii}. Since the volume where the integral is valid is the exterior of S^i and S^{ii} and the interior of S^{iii}, all sources and sinks must be located *outside* this surface. The contribution from surface S^{ii} vanishes as the diameter of the tube is made zero, and the contribution from S^i is just $4\pi\mathbf{E}$ as before. Consequently, the electric field is given, within the *closed interior* of S^{iii}, by

$$\mathbf{E}(x_p,y_p,z_p) = -\frac{1}{4\pi}\oiint_{S^{iii}} [i\omega\mu(\hat{\mathbf{m}}\times\mathbf{H})G + (\hat{\mathbf{m}}\times\mathbf{E})\times\nabla G + (\hat{\mathbf{m}}\cdot\mathbf{E})\nabla G]\mathrm{d}S^{iii}, \tag{18a}$$

and the expression for the magnetic field is given by

$$\mathbf{H}(x_p,y_p,z_p) = \frac{1}{4\pi}\oiint_{S^{iii}} [i\omega\epsilon(\hat{\mathbf{m}}\times\mathbf{E})G - (\hat{\mathbf{m}}\times\mathbf{H})\times\nabla G - (\hat{\mathbf{m}}\cdot\mathbf{H})\nabla G]\mathrm{d}S^{iii}. \tag{18b}$$

Equations (18) are usually referred to as the Stratton–Chu integral theorem.

On comparing Eqs. (17) and (18), we initially note a sign difference for both the electric and magnetic fields. On closer inspection we note that there is also a difference in the surface over which the integration is performed. In Eqs. (18) the surface encompasses a source- and sink-free volume of space, and $\hat{\mathbf{m}}$ is the outward surface normal. The point of observation P is inside the volume. In the case of Eq. (17) the surface S^{iii} encompasses all sources and sinks in space, and $\hat{\mathbf{m}}$ is an outward normal. The point of observation is outside the enclosed volume. Consequently, we recognize that Eqs. (17) and (18) are *solutions of two different problems*. Appendix A contains a proof showing that Eqs. (17) and (18) are exact solutions of Maxwell's equations. The solution is unique, which follows from the fact that Maxwell's equations lead to a unique solution for a given set of boundary conditions.

In conclusion, Eqs. (18) are a unique, self-consistent, and rigorous solution of Maxwell's equations for the electric and magnetic field at a point within any closed surface S bounding a charge- and current-free, homogeneous, and isotropic volume. Also, Eqs. (17) are a unique, self-consistent, and rigorous solution of Maxwell's equations for the electric and magnetic field at any exterior point of a closed surface bounding all sources and sinks. It can also be shown,[12] though it is well outside the scope of this paper, that this integral can be the basis for deriving other, well-known diffraction theories.

3. PRACTICAL USE OF THE STRATTON–CHU THEORY

When seeking the solution of an electromagnetic scattering or diffraction problem by means of an arbitrary shaped object, it is, in general, necessary to employ a rigorous numerical method. This is because it is very difficult to analytically solve Maxwell's equations in arbitrary inhomogeneous regions with even the simplest of boundary conditions. Numerical methods such as the FDTD technique and the FEM are commonly used for this purpose. These methods, however, use computer memory proportional to the volume of the computational space being analyzed. It is thus often impossible to include far-field regions in the computational space. If, however, a numerical method is used to calculate the field on a boundary surrounding the scatterer, the field anywhere outside the boundary may be found by applying the Stratton–Chu integral [Eqs. (17)].

There are other methods of obtaining the far field from near-field data. However, as stated at the end of Section 2, the other major vectorial diffraction theories such as the vectorial integral theorem of Kirchhoff and the vectorial Rayleigh–Sommerfeld theory may be obtained from the Stratton–Chu integral.[12] Thus the Stratton–Chu integral is if anything more general than the other theories and is well suited to our application.

A. Numerical Implementation of the Stratton–Chu Integral

Since the near-field data are known numerically, the Stratton–Chu integral must be evaluated numerically. The surface of integration and associated complex amplitudes are defined by using a mesh of triangles. Such a mesh is represented by a set of vertices $V=\{\mathbf{r}_{s,i}=(r_{s,i}^1,r_{s,i}^2,r_{s,i}^3)\in\mathbb{R}^3\}$, a set of facets $F=\{(v_i^1,v_i^2,v_i^3)\in\mathbb{N}^3, 1\leqslant v_i^j\leqslant N_v\}$, where N_v is the number of vertices, and two sets of complex amplitudes $E=\{\mathbf{E}_i=(e_i^1,e_i^2,e_i^3)\in\mathbb{C}^3\}$ and $\mathbf{H}=\{\mathbf{H}_i=(h_i^1,h_i^2,h_i^3)\in\mathbb{C}^3\}$. Each element of each triple in F is an index into V, E, and H. In this way each triangle is constructed from three vertices, and the field at each vertex is also known. The orientation of the surface is stored according to the order in which the facet indices are stored. Such a representation minimizes computer storage and provides an efficient way to traverse the surface. This representation can be used to represent any polyhedral surface and so is very general.

Integration is performed over each facet, and the results summed to give the final result. Gaussian quadrature is commonly used for integration over a triangle. In general, a high order Gaussian quadrature would be employed to improve integration accuracy. However, since the field is known only at the triangle vertices, only first order Gaussian quadrature may be employed. The first order scheme provides an exact result if the integral kernel varies no worse than a polynomial of first order.[13] This

is why it is important to use a fine mesh to represent the closed surface of integration.

By employing first order Gaussian quadrature integration, we may evaluate the Stratton–Chu integrals according to

$$\mathbf{U}(\mathbf{r}_p) = \sum_{i=1}^{N_{facets}} \left[\frac{1}{3} \sum_{j=1}^{3} \mathbf{I}(\mathbf{r}_p, \hat{\mathbf{m}}_i, \mathbf{r}_{s,v_i^j}, \mathbf{E}_{v_i^j}, \mathbf{H}_{v_i^j}) \Delta_i \right], \quad (19)$$

where $\mathbf{U}$ is the field of interest, N_{facets} is the number of facets, $\mathbf{I}$ is the kernel of the integral being evaluated, $\mathbf{r}_p = (x_p, y_p, z_p)$ is the observation point, $\hat{\mathbf{m}}_i$ is the surface normal of facet i, $\mathbf{r}_{s,v_i^j}$ is the jth vertex of facet i, $\mathbf{E}_{v_i^j}$ is the complex electric field at the jth vertex of facet i, $\mathbf{H}_{v_i^j}$ is the complex magnetic field at the jth vertex of facet i, and Δ_i is the area of facet i.

B. Parallel Implementation of the Stratton–Chu Integral

The numerical evaluation of the Stratton–Chu integral at a number of observation points is quite a computationally intensive task. This is because the argument of the Green's function and its gradient is dependent on both the observation point and surface coordinates. Thus the Green's function and its gradient must be evaluated for every pairing of observation and surface points. To speed up computation time, the program is implemented in parallel.

The message passing interface is used as the basis for interprocess communication. The program is parallelized by dividing up the set of observation points where the field is to be calculated. Each processor then calculates the field at a subset of observation points. The program runs on our 36 processor Beowulf cluster.

4. EXAMPLES AND ANALYSIS

A. Accuracy of Implementation

The implementation of the Stratton–Chu integral was tested by using the field due to a harmonically oscillating dipole as a test field. This was done by first calculating the field on a test surface surrounding the dipole. This field was then used as the input to the Stratton–Chu program to evaluate the field at a test plane some distance from the test surface. The field at the test plane was then compared with the field obtained analytically.

The electromagnetic field at point $r\hat{\mathbf{n}}$ due to a harmonically oscillating dipole with moment $\mathbf{p}$, situated at the origin, may be calculated according to[14]

$$\mathbf{H} = \frac{ck^2}{4\pi}(\hat{\mathbf{n}} \times \mathbf{p}) \frac{\exp(ikr)}{r} \left(1 - \frac{1}{ikr}\right),$$

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0} \left\{ k^2(\hat{\mathbf{n}} \times \mathbf{p}) \times \hat{\mathbf{n}} \frac{\exp(ikr)}{r} + [3\hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \mathbf{p}) - \mathbf{p}] \left(\frac{1}{r^3} - \frac{ik}{r^2}\right) \exp(ikr) \right\}, \quad (20)$$

where $\hat{\mathbf{n}}$ is a unit vector directed from the dipole to the point of observation and r is the distance from the dipole to the point of observation.

A z-polarized dipole situated at the origin and radiating at wavelength $\lambda = 632.8$ nm was used in all tests. A regular triangular mesh fitted to the surface of a cube, centered on the origin and of side 2λ, was used as the closed surface of integration.

An aggregate error metric was defined to measure the accuracy of the Stratton–Chu integral. This results in a measure of relative error that does not produce the spurious results that some pointwise relative error measures produce. It also has the added advantage of producing only a single error value, thus making it useful for comparison purposes. The aggregate error metric is thus defined as

$$\epsilon_{\mathbf{U}} = \frac{\sum_{i=1}^{N} |\mathbf{U}^{SC}(\mathbf{r}_i) - \mathbf{U}^{An}(\mathbf{r}_i)|^2}{\sum_{i=1}^{N} |\mathbf{U}^{An}(\mathbf{r}_i)|^2}, \quad (21)$$

where $\mathbf{U}$ can represent either the electric field $\mathbf{E}$ or the magnetic field $\mathbf{H}$, N is the number of observation points, $\mathbf{U}^{An}$ is the field calculated analytically with Eqs. (20), $\mathbf{U}^{SC}$ is the field calculated with the Stratton–Chu integral [Eqs. (17)], and $\mathbf{r}_i$ is the ith point on the test plane where the field was calculated.

The first test examined the dependence of error on the density of the mesh used to represent the closed surface of integration. This was performed by using the Stratton–Chu integral [Eqs. (17)] to propagate the dipole field from a series of test surfaces with increasingly dense meshes to a fixed test plane. The test plane was parallel to the xy plane and situated at 200λ along the z axis. It had dimension $4000\lambda \times 4000\lambda$ so as to contain the majority of the dipole field pattern. The vertex spacing on the surface of the integration mesh was varied from $\lambda/1$ to $\lambda/40$. The mesh was constructed by first partitioning each cube face of the integration surface into a minimal number of squares of the desired vertex spacing. Each square was then divided into two right angled isosceles triangles. Vertex spacing is defined as the length of the side of the squares.

Figure 4 shows a plot of aggregate error versus mesh

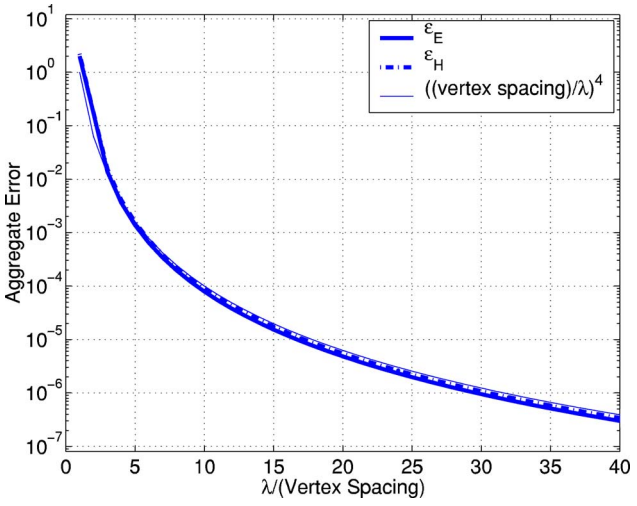

Fig. 4. (Color online) Graph showing aggregate error versus mesh density for a single dipole scatterer.

density for **E** and **H**. Unsurprisingly, the error reduces as the vertex spacing is reduced. For large vertex spacing the error is unacceptably high; however, it improves rapidly as the vertex spacing reduces. In fact, as shown on the plot, the error follows approximately a (vertex spacing)^4 relationship. The appropriate vertex spacing depends very much on the application. For example, when the field data for the closed surface of integration are calculated with a rigorous numerical method, they will already have an error associated with them attributable to the numerical method. The vertex spacing should thus be chosen to result in error less than that of the numerical method.

Numerical methods require grid spacing to be kept below a certain upper limit in order to obtain accurate results. The Stratton–Chu surface is obtained directly from sample points on the boundary of the grid of the numerical method. Thus the grid spacing of the numerical method determines the vertex spacing of the Stratton–Chu surface. This is, in general, $\lambda/20$ or finer if the FDTD method is used to compute the near field. At this vertex spacing the error associated with the Stratton–Chu integral is likely to be less than that due to the FDTD method. Thus the Stratton–Chu integral is well suited for propagating fields obtained from numerical calculations.

The second test examined the dependence of error on distance between the dipole and test plane. The test surface from the mesh density error calculation with mesh vertex spacing of $\lambda/20$ was reused for this calculation. The test plane was again parallel to the xy plane; however, its position was varied along the z axis. The dimensions of the test plane were altered proportionately with the plane's position along the z axis such that each vertex retained its relative position within the field distribution. This ensures that the aggregate error measures between two plane positions are comparable. The distance along the z axis from the dipole to the observation plane was varied from 1λ to 1000λ in order to test the efficacy of the Stratton–Chu integral as a near to far-field transform. The results of this test are shown in Fig. 5; note, however,

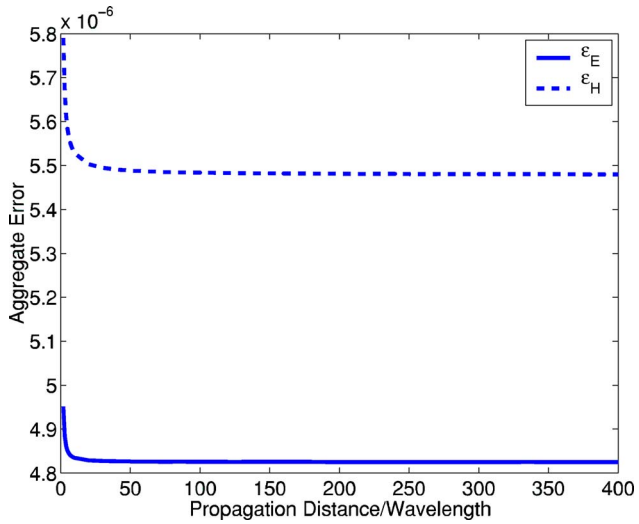

Fig. 5. (Color online) Graph showing aggregate error versus propagation distance for a single dipole scatterer when a vertex spacing of $\lambda/20$ is used.

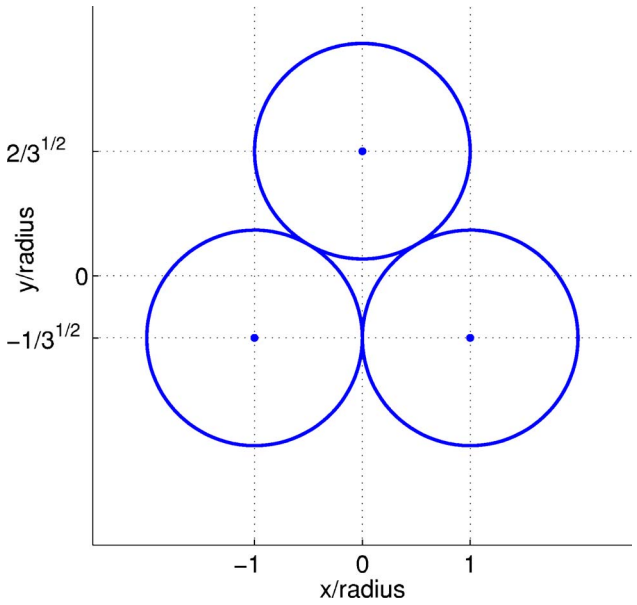

Fig. 6. (Color online) Diagram showing the position of the three spheres in the plane $z=0$.

that the axes have been truncated to a maximum propagation distance of 400λ, as the error remains constant for greater propagation distances.

Figure 5 shows that the aggregate errors for the electric and magnetic fields differ slightly, as is also evident in Fig. 4. Note, however, that this difference is very small and is most probably due to the differing complexity of the dipole field components. The most important feature of Fig. 5 is that the error reduces rapidly as the propagation distance increases into the far-field region. This is to be expected, as the field becomes smoother and better behaved in the far field. Furthermore, it may be shown that the change in r, the argument of the Green's function, due to a small change in the x surface coordinate is given by

$$\delta r \approx \frac{x - x_p}{r}\delta x, \tag{22}$$

where the result is identical for the y and z coordinates. Thus, as the propagation distance increases, the phase and amplitude of the Green's function vary more slowly as a function of the surface coordinates. It is thus to be expected that the aggregate error diminishes with propagation distance.

B. Diffraction by Three Spheres

As a practical example we consider calculating the far-field diffraction pattern of a complex object for which the scattered field cannot be calculated analytically in a simple manner: a collection of three perfectly conducting spheres, each 0.405 μm in diameter located in the plane $z=0$ (Fig. 6). A linearly polarized plane wave, traveling in the positive z direction, is incident on the spheres. Suppose that we wish to know the scattered field at a plane intersecting the negative z axis a long way from the spheres for both x- and y-polarized incident plane waves.

The field scattered by the spheres can be calculated routinely with the FDTD method. However, even on powerful workstations this technique could not be used to cal-

culate the field beyond about ten wavelengths from the spheres. On supercomputers it may be possible to go as far as over 100 wavelengths, but this would require a very expensive machine. Conversely, the first version of the Stratton–Chu integral theorem [Eqs. (17)] can be run on a medium-sized workstation and is perfectly suited to such a problem, as it may be used to calculate the field at arbitrarily large distances from the spheres by using the FDTD data.

To illustrate how this can be done, we first show the intensity as given by the FDTD method on the surface that surrounded the computational volume in Fig. 7 for $\lambda = 405$ nm. The FDTD method produced these results on a cube of side 2.5λ meshed by using a triangular mesh of 120,000 facets and 60,002 vertices. The vertex spacing was $\lambda/40$. Then the Stratton–Chu integral is employed, and the resulting intensity distribution in the plane $z=-2$ m is shown in Fig. 8 for x- and y-polarized incident plane waves.

These results show that even in the far field there are differences in the diffraction pattern caused by the polarization of the incident wave. The true value of this example is that it demonstrates how, with the scattered field stored on a surface enclosing the scatterer, the field anywhere outside that surface can be calculated in a memory- and time-efficient manner. The practical benefit of this is that it allows numerical modeling of scatterers without needing to worry about where the field should be propagated to. It would be very inefficient and in many cases impossible to, for example, rerun an FDTD simulation simply to propagate the scattered field further than was originally done.

C. Parallel Implementation of the Stratton–Chu Integral

To assess the time saving due to parallelization, we ran a calculation, encountered frequently when modeling the readout of optical disks, many times by using an increasing number of processors. In this calculation the Stratton–Chu integral was used to evaluate the scattered field on a square of width 4 μm positioned 253 nm (one wavelength in plastic with refractive index 1.6) from the data layer of the disk. A uniform grid of 4624 observation points was constructed on this square. The closed surface of integration, on which the field was determined by the FDTD method, was composed of a mesh of 195,072 facets and 97,538 vertices. The vertex spacing was $\lambda/40$.

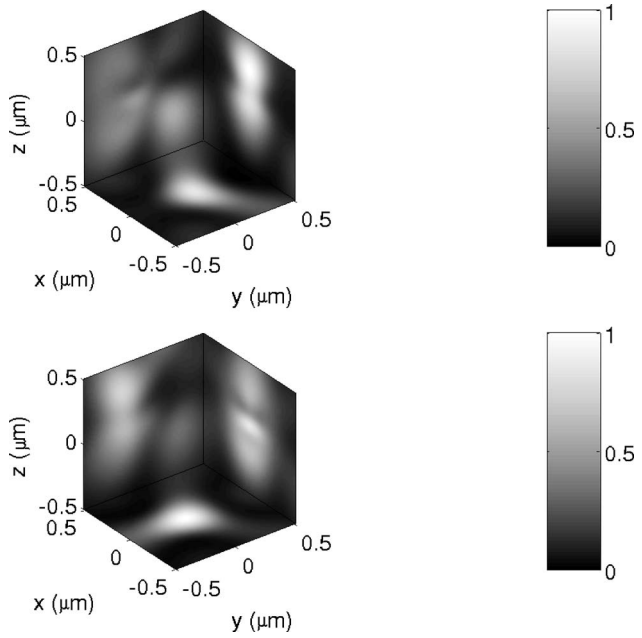

Fig. 7. Intensity of the electric field scattered by three spheres on the FDTD surface for x- (top) and y- (lower) polarized incident waves. Images have been individually normalized.

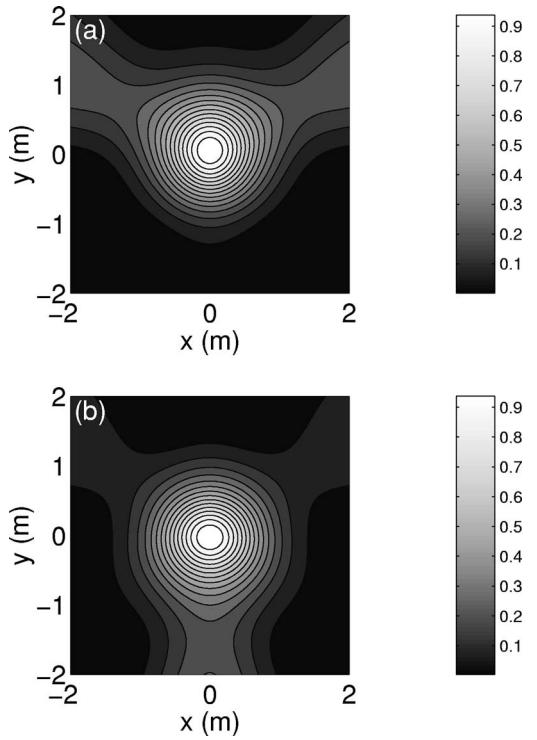

Fig. 8. Intensity of scattered field in the plane $z=-2$ m for (a) x-polarized incident waves and (b) y-polarized incident waves. Images have been individually normalized.

This calculation was performed 36 times, initially by using a single processor and finally by using 36 processors. Results are shown in Fig. 9. Results have been presented in two different ways. The upper plot in Fig. 9 shows the total time taken to evaluate the field at all observation points on the grid. This may be considered the real time required to complete the simulation. This plot shows that, as expected, significant time can be saved by employing more processors. The lower plot shows the total processor time per observation point evaluated, taking into account the number of processors employed. This plot is of more interest, as it shows the penalty associated with using additional processors. This penalty is caused by the additional initialization and communication that results from using more processors. In the absence of such a penalty one would expect the lower plot in Fig. 9 to be horizontal.

A linear trend line has been fitted to the data and reveals that the penalty is approximately 10^{-4} s per observation point per processor. Note that the plot exhibits apparently random variations often associated with benchmarking measurements. Thus, of the 151.4 s required to perform the calculation on 36 processors, just

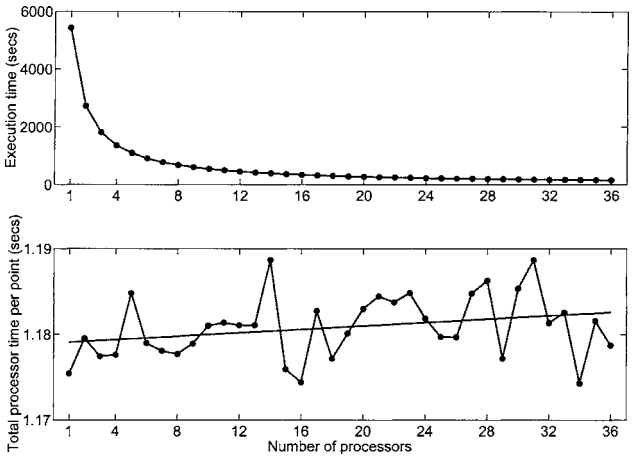

Fig. 9. Timing data for parallel implementation of the Stratton–Chu code. The upper plot shows total execution time as a function of the number of processors, and the lower plot shows the average total amount of processor time per observation point as a function of the number of processors employed.

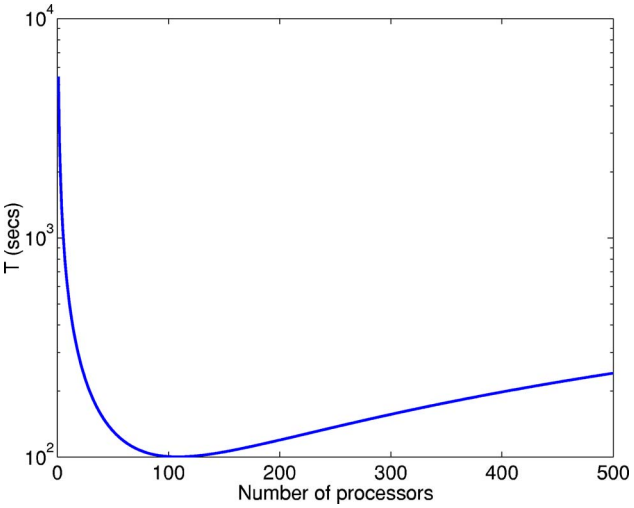

Fig. 10. (Color online) Theoretical prediction for the total execution time as a function of the number of processors employed.

under 17 s (or 11%) of this was consumed by additional initialization and communication. Although we have only 36 processors, it is of interest to know at what number of processors the total computation time will begin to rise. If it is assumed that the penalty per processor remains constant, the total time to complete a simulation can be approximated as

$$T(N) = T_1/N + Nt_p, \tag{23}$$

where T_1+t_p is the time taken to complete the calculation on a single processor, N is the number of processors, and t_p is the penalty per processor. Assuming a value for t_p of $10^{-4}\times 4624 = 0.4624$, this T is plotted in Fig. 10. The turning point of the plot occurs at approximately $N=108$ processors. It should be noted that this result depends on the specific simulation being run and is quite approximate due to the nature of measuring code execution time. This result does reveal, however, that for this typical simulation there is little point in exceeding approximately 100 processors.

Without parallelization, evaluation of such problems would become prohibitively time consuming. For example, we need to perform 42 of these typical calculations in order to analyze a portion of a single track on an optical disk. On a single processor this would take over 63 h. Using 36 processors requires only 1.8 h. We need to simulate many of these tracks, so, given the often iterative nature of such simulation, parallelization of this problem is essential.

5. CONCLUSIONS

This paper presented a rigorous method for transforming the near field of an electromagnetic distribution to the far field. We derived a set of self-consistent integral equations that were used to represent the electromagnetic field rigorously everywhere in a homogeneous space apart from the closed interior of a volume encompassing all charges and sinks. The representation was derived by imposing a condition analogous to Sommerfeld's radiation condition that we termed the vectorial radiation condition.

We examined the accuracy of our numerical implementation of the formula, also on a parallel computer cluster, by comparing the results with a problem when the analytical solution is known. We have found that the aggregate computational error decreases with increasing distance between the volume and the observation point. We also have found that in the case of our parallel computer cluster there is an optimum number of processors beyond which, due to interprocessor communication latency, the time it takes to evaluate the integral is longer than it is with fewer processors.

Our basic equation is of fundamental relevance in electromagnetic propagation, and the numerical implementation will find applications in computational electromagnetism.

APPENDIX A

Now we discuss how Eqs. (18) are a rigorous solution of Maxwell's equation. It trivially follows that Eqs. (17) and also a solution of Maxwell's equations. First, we introduce the notation ∇_p to distinguish the vector differential operator acting on the observation point P from the one, ∇, acting on the surface and use S to denote the closed surface of integration for simplicity. We have to show that (i) $\nabla_p\cdot\mathbf{E}=0$ and (ii) $\nabla_p\times\mathbf{H}=-i\omega\epsilon\mathbf{E}$. Let us start with (i):

$$4\pi\nabla_p\cdot\mathbf{E} = -i\omega\mu\oiint_S \nabla_p\cdot[(\hat{\mathbf{m}}\times\mathbf{H})G]\mathrm{d}S - \oiint_S \nabla_p\cdot[(\hat{\mathbf{m}}\times\mathbf{E}) \times\nabla G]\mathrm{d}S - \oiint_S \nabla_p\cdot[(\hat{\mathbf{m}}\cdot\mathbf{E})\,\nabla G]\mathrm{d}S. \tag{A1}$$

The first integral can be rewritten to give

$$-i\omega\mu\oiint_S \nabla_p\cdot[(\hat{\mathbf{m}}\times\mathbf{H})G]\mathrm{d}S$$
$$=-i\omega\mu\oiint_S[\nabla_p\cdot(\hat{\mathbf{m}}\times\mathbf{H})]G\mathrm{d}S$$
$$-i\omega\mu\oiint_S(\hat{\mathbf{m}}\times\mathbf{H})\cdot\nabla_p G\mathrm{d}S$$
$$=i\omega\mu\oiint_S(\hat{\mathbf{m}}\times\mathbf{H})\cdot\nabla G\mathrm{d}S \quad \text{(A2)}$$

because $\nabla_p G=-\nabla G$. The second integral in Eq. (A1) yields identically zero after using the rule of triple scalar products on the kernel and by noting that $\nabla\times\nabla G\equiv\mathbf{0}$. The third term may be written as

$$-\oiint_S \nabla_p\cdot[(\hat{\mathbf{m}}\cdot\mathbf{E})\nabla G]\mathrm{d}S=\oiint_S(\hat{\mathbf{m}}\cdot\mathbf{E})\nabla^2 G\mathrm{d}S$$
$$=-\omega^2\mu\epsilon\oiint_S(\hat{\mathbf{m}}\cdot\mathbf{E})G\mathrm{d}S \quad \text{(A3)}$$

because it was assumed that G satisfies the scalar wave equation. We can write

$$-\omega^2\mu\epsilon\oiint_S(\hat{\mathbf{m}}\cdot\mathbf{E})G\mathrm{d}S=-i\omega\mu\oiint_S[G(\nabla\times\mathbf{H})]\cdot\hat{\mathbf{m}}\mathrm{d}S, \quad \text{(A4)}$$

so that now Eq. (A4) reads as

$$-i\omega\mu\oiint_S[G(\nabla\times\mathbf{H})]\cdot\hat{\mathbf{m}}\mathrm{d}S$$
$$=-i\omega\mu\oiint_S[\nabla\times(G\mathbf{H})]\cdot\hat{\mathbf{m}}\mathrm{d}S$$
$$+i\omega\mu\oiint_S(\nabla G\times\mathbf{H})\cdot\hat{\mathbf{m}}\mathrm{d}S$$
$$=-i\omega\mu\oiint_S[\nabla\times(G\mathbf{H})]\cdot\hat{\mathbf{m}}\mathrm{d}S$$
$$-i\omega\mu\oiint_S(\hat{\mathbf{m}}\times\mathbf{H})\cdot\nabla G\mathrm{d}S. \quad \text{(A5)}$$

A form of Stokes's theorem reads as, with our usual notation,

$$\int\int_S(\nabla\times\mathfrak{B})\cdot\hat{\mathbf{m}}\mathrm{d}S=\int\int_S(\nabla\times\mathfrak{B})\cdot\mathrm{d}\mathbf{S}=\oint_C\mathfrak{B}\cdot\mathrm{d}\mathbf{s}, \quad \text{(A6)}$$

where C is a closed contour around the open surface S. It is easy to see that if the surface is closed, the value of the line integral vanishes. Hence the first term on the right hand side of Eq. (A5) also vanishes, so we are left with

$$-\oiint_S\nabla_p\cdot[(\hat{\mathbf{m}}\cdot\mathbf{E})\nabla G]\mathrm{d}S=-i\omega\mu\oiint_S(\hat{\mathbf{m}}\times\mathbf{H})\cdot\nabla G\mathrm{d}S. \quad \text{(A7)}$$

When substituting back from Eqs. (A2) and (A7) into Eq. (A1), we obtain

$$\nabla_p\cdot\mathbf{E}=0, \quad \text{(A8)}$$

which completes the first part of the proof. We now prove that $\nabla_p\times\mathbf{H}=-i\omega\epsilon\mathbf{E}$. Let us consider

$$4\pi\nabla_p\times\mathbf{H}=i\omega\epsilon\oiint_S\nabla_p\times[(\hat{\mathbf{m}}\times\mathbf{E})G]\mathrm{d}S$$
$$-\oiint_S\nabla_p\times[(\hat{\mathbf{m}}\times\mathbf{H})\times\nabla G]\mathrm{d}S \quad \text{(A9)}$$

because, again, $\nabla\times\nabla G\equiv\mathbf{0}$. A procedure similar to that described above for the previous proof yields

$$\nabla_p\times\mathbf{H}=\frac{i\omega\epsilon}{4\pi}\oiint_S[i\omega\mu(\hat{\mathbf{m}}\times\mathbf{E})G+(\hat{\mathbf{m}}\times\mathbf{E})\times\nabla G$$
$$+(\hat{\mathbf{m}}\cdot\mathbf{E})\nabla G]\mathrm{d}S=-i\omega\epsilon\mathbf{E}. \quad \text{(A10)}$$

With this equation we have completed the proof.

ACKNOWLEDGMENT

The authors acknowledge help from Ben Eastley, who wrote parts of the Stratton–Chu evaluation code. Address correspondence to P. Török at peter.torok@imperial.ac.uk.

REFERENCES

1. J. A. Stratton and L. J. Chu, "Diffraction theory of electromagnetic waves," Phys. Rev. **56**, 99–107 (1939).
2. W. Hsu and R. Barakat, "Stratton–Chu vectorial diffraction of electromagnetic fields by apertures with application to small-Fresnel-number systems," J. Opt. Soc. Am. A **11**, 623–629 (1994).
3. P. Török, "Focusing of electromagnetic waves through a dielectric interface by lenses of finite Fresnel number," J. Opt. Soc. Am. A **15**, 3009–3015 (1998).
4. P. Varga and P. Török, "Electromagnetic focusing by a paraboloid mirror. I. Theory," J. Opt. Soc. Am. A **17**, 2081–2089 (2000).
5. P. Varga and P. Török, "Electromagnetic focusing by a paraboloid mirror. II. Numerical results," J. Opt. Soc. Am. A **17**, 2090–2095 (2000).
6. P. Varga, "Focusing of electromagnetic radiation by hyperboloidal and ellipsoidal lenses," J. Opt. Soc. Am. A **19**, 1658–1667 (2002).
7. A. Taflove, *Computational Electrodynamics: The Finite Difference Time Domain Method*, 1st ed. (Artech House, 1995).
8. J. Jin, *The Finite Element Method in Electromagnetics* (Wiley, 2002).

9. B. Baker and E. Copson, *The Mathematical Theory of Huygens' Principle*, 2nd ed. (Oxford U. Press, 1953).
10. J.-C. Nédélec, *Acoustic and Electromagnetic Equations: Integral Representations for Harmonic Problems* (Springer-Verlag, 2001).
11. J. Stratton, *Electromagnetic Theory* (McGraw-Hill, 1941).
12. P. Török and C. Sheppard, *High Numerical Aperture Focusing and Imaging* (Hilger, to be published).
13. G. Arfken, *Mathematical Methods for Physicists*, 3rd ed. (Academic, 1985).
14. J. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, 1999).