Taylor & Francis
Taylor & Francis Group

Evaluation of the validity of the scalar approximation in optical wave propagation using a systems approach and an accurate digital electromagnetic model

Onur Kulce, Levent Onural and Haldun M. Ozaktas

Department of Electrical and Electronics Engineering, Bilkent University, Ankara, Turkey

ABSTRACT

The cause and amount of error arising from the use of the scalar approximation in monochromatic optical wave propagation are discussed using a signals and systems formulation. Based on Gauss's Law, the longitudinal component of an electric field is computed from the transverse components by passing the latter through a two input single output linear shift-invariant system. The system is analytically characterized both in the space and frequency domains. For propagating waves, the large response for the frequencies near the limiting wave number indicates the small angle requirement for the validity of the scalar approximation. Also, a discrete simulator is developed to compute the longitudinal component from the transverse components for monochromatic propagating electric fields. The simulator output helps to evaluate the validity of the scalar approximation when the system output cannot be analytically calculated.

1. Introduction

When quantum behaviour can be ignored, light can be treated as a vector electromagnetic wave. Some optical phenomena, such as interference, diffraction and holography, are often treated using a scalar approximation, since it is analytically and computationally simpler (1, 2). The main concern of the present work is to discuss the limitations of commonly used scalar approaches with the help of a signals and systems model, for monochromatic waves. Such a model yields a quantitative evaluation of the deviation of the scalar treatment from the vector case.

Since an arbitrary electric field in free space satisfies Gauss's Law, the total outward flux passing through the surface enclosing an infinitesimal volume at each point in space should be zero, provided there is no source or sink within that volume; in other words, the divergence of the electric field, denoted by $\nabla \cdot \mathbf{E}$, should be zero (3). As a consequence, one of the three components of the electric field becomes dependent on the others. This dependency appears as the orthogonality requirement for a plane wave, i.e. the electric field vector and the propagation direction must be orthogonal to each other (3). The linear-shift invariant system model that we employ is based on this orthogonality requirement: the x and y component functions of the electric field are fed to the system as the inputs, and the z component function is obtained as the output. Since the angle of propagation of a plane wave corresponds to the spatial frequency over

the transverse plane, the frequency response characteristics of the system paves the way for the evaluation of the validity of the scalar approximation in a given application.

In Section 2, we present the mathematical preliminaries and in Section 3, we introduce the filters which are used in the system model and derive their impulse responses valid for both propagating and evanescent cases. In Section 4, the properties of the filters are provided. In Section 5, we interpret the scalar approximation to the vector theory based on the frequency characteristics of the system and mention related studies found in the literature. A digital simulator of the system is developed in Section 6. The aim of the digital simulator is, when analytical calculations are not possible, to be able to evaluate the validity of the scalar approximation based on the discussions in Section 5. Finally, in Section 7, we draw the conclusions.

2. Preliminaries

We assume the field is monochromatic and omit the time dependence in our analysis. Let the electric field in a three-dimensional (3D) space be $\mathbf{E}(\mathbf{r})$, where $\mathbf{r} = \begin{bmatrix} x & y & z \end{bmatrix}^T \in \mathbb{R}^3$ and $\mathbf{E}(\mathbf{r}) = \begin{bmatrix} E_x(\mathbf{r}) & E_y(\mathbf{r}) & E_z(\mathbf{r}) \end{bmatrix}^T \in \mathbb{C}^3$, is the electric field vector. $\mathbf{E}(\mathbf{r})$ can be written via the plane wave decomposition including both propagating and evanescent components as (also called the 2D inverse Fourier transform (IFT)):

$$\mathbf{E}\left(\mathbf{r}\right) = \frac{1}{4\pi^2} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \boldsymbol{\mathcal{E}}\left(k_x, k_y\right) e^{j\mathbf{k}^T\mathbf{r}} dk_x dk_y, \quad (1)$$

where $\mathbf{k} = \begin{bmatrix} k_x & k_y & k_z \end{bmatrix}^T$ is the spatial frequency vector for plane waves and also indicates the direction of propagation. As a result of the wave equation in a source-free space, the magnitude of $\mathbf{k}$, which is denoted as k and called the wave number, is equal to $2\pi/\lambda$, where λ is the wavelength (3). We choose the z component of the propagation direction, k_z, to be positive for each propagating plane wave, and evanescent waves decay along the positive z direction. So, k_x and k_y are real numbers and

$$k_z = \sqrt{k^2 - k_x^2 - k_y^2}. \quad (2)$$

$\boldsymbol{\mathcal{E}}\left(k_x, k_y\right) = \begin{bmatrix} \mathcal{E}_x\left(k_x, k_y\right) & \mathcal{E}_y\left(k_x, k_y\right) & \mathcal{E}_z\left(k_x, k_y\right) \end{bmatrix}^T \in \mathbb{C}^3$ is the 2D Fourier transform (FT), from the (x, y) domain to the (k_x, k_y) domain, of the electric field pattern at the $z = 0$ plane and can be found as,

$$\boldsymbol{\mathcal{E}}\left(k_x, k_y\right) = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \mathbf{E}\left(x, y, 0\right) e^{-j(k_x x + k_y y)} dx dy. \quad (3)$$

Thus, the vector amplitude of each plane wave component propagating along the $\mathbf{k}$ direction is $\left[\boldsymbol{\mathcal{E}}\left(k_x, k_y\right) / \left(4\pi^2\right)\right] dk_x dk_y$.

Equations 1 and 3 indicate that if the E_x, E_y, E_z components of the electric field are known over the $z = 0$ plane, then the electric field over the entire 3D space is also known. Thus, each component can be treated separately as scalar fields, as known in the literature (2). However, using an additional constraint due to $\nabla \cdot \mathbf{E} = 0$ condition, it is possible to uniquely deduce the E_z component from the E_x and E_y components as given below.

In source-free space, Gauss's Law, which is one of Maxwell's equations, states that (3):

$$\nabla \cdot \mathbf{E}\left(\mathbf{r}\right) = \frac{\partial E_x\left(\mathbf{r}\right)}{\partial x} + \frac{\partial E_y\left(\mathbf{r}\right)}{\partial y} + \frac{\partial E_z\left(\mathbf{r}\right)}{\partial z}$$
$$= \frac{j}{4\pi^2} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \mathbf{k}^T \boldsymbol{\mathcal{E}}\left(k_x, k_y\right) e^{j\mathbf{k}^T\mathbf{r}} dk_x dk_y$$
$$= 0. \quad (4)$$

In Equation 4, $\nabla \cdot \mathbf{E}\left(\mathbf{r}\right)$ is expressed as the linear combination of 2D complex sinusoidal basis functions $e^{j(k_x x + k_y y)}$ with the Fourier coefficients $j\mathbf{k}^T \boldsymbol{\mathcal{E}}\left(k_x, k_y\right) e^{jk_z z}$. Since the basic functions $e^{j(k_x x + k_y y)}$ constitute an orthogonal, and

hence, linearly independent basis set, Equation 4 implies that,

$$\frac{j}{4\pi^2} \mathbf{k}^T \boldsymbol{\mathcal{E}}\left(k_x, k_y\right) e^{jk_z z} = 0. \quad (5)$$

After simplifications, Equation 5 turns out to be

$$\mathbf{k}^T \boldsymbol{\mathcal{E}}\left(k_x, k_y\right) = k_x \mathcal{E}_x\left(k_x, k_y\right) + k_y \mathcal{E}_y\left(k_x, k_y\right)$$
$$+ k_z \mathcal{E}_z\left(k_x, k_y\right) = 0. \quad (6)$$

Equation 6 states that the electric field coefficient and propagation direction vectors are orthogonal to each other. From Equations 2 and 6, we find,

$$\mathcal{E}_z\left(k_x, k_y\right) = -\frac{k_x \mathcal{E}_x\left(k_x, k_y\right) + k_y \mathcal{E}_y\left(k_x, k_y\right)}{\sqrt{k^2 - k_x^2 - k_y^2}}. \quad (7)$$

Finally, we obtain $E_z\left(\mathbf{r}\right)$ from $E_x\left(\mathbf{r}\right)$ and $E_y\left(\mathbf{r}\right)$ as,

$$E_z\left(\mathbf{r}\right) = \frac{1}{4\pi^2} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \frac{-k_x \mathcal{E}_x\left(k_x, k_y\right) - k_y \mathcal{E}_y\left(k_x, k_y\right)}{\sqrt{k^2 - k_x^2 - k_y^2}}$$
$$e^{j\mathbf{k}^T\mathbf{r}} dk_x dk_y. \quad (8)$$

Equation 8 is also given in the literature and used for aperture antenna problems (4–6) and for vector beam solutions to Maxwell's equations (7–9). Please note that, for the special case $k_x^2 + k_y^2 = k^2$, we assume that $\mathcal{E}_x\left(k_x, k_y\right)$, $\mathcal{E}_y\left(k_x, k_y\right)$ and $\mathcal{E}_z\left(k_x, k_y\right)$ are 0.

The electric field given in Equation 1 can also be decomposed into propagating and evanescent components as

$$\mathbf{E}\left(\mathbf{r}\right) = \mathbf{E}_P\left(\mathbf{r}\right) + \mathbf{E}_{P'}\left(\mathbf{r}\right), \quad (9)$$

where the subscripts P and P' denote the propagating and evanescent components, respectively. These components can be written as

$$\mathbf{E}_P\left(\mathbf{r}\right) = \frac{1}{4\pi^2} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \boldsymbol{\mathcal{E}}\left(k_x, k_y\right)$$
$$\times \text{rect}\left(\frac{\sqrt{k_x^2 + k_y^2}}{k}\right) e^{j\mathbf{k}^T\mathbf{r}} dk_x dk_y, \quad (10)$$

$$\mathbf{E}_{P'}\left(\mathbf{r}\right) = \frac{1}{4\pi^2} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \mathcal{E}\left(k_x, k_y\right)$$

$$\times \left[1 - \mathrm{rect}\left(\frac{\sqrt{k_x^2 + k_y^2}}{k}\right)\right] e^{j\mathbf{k}^T\mathbf{r}} dk_x dk_y, \quad (11)$$

where

$$\mathrm{rect}\left(\chi\right) = \begin{cases} 1 & \text{if } \chi < 1 \\ 0 & \text{if } \chi > 1 \end{cases} \quad (12)$$

for $0 \le \chi < \infty$. Decomposing the field into the propagating and evanescent components as in Equation 9 simplifies the computation of the z component if the field is known to be propagating or evanescent, as will be shown in Section 3.

3. System characterization

In this section, we put the computation of the longitudinal component in a signals and systems framework. In the following sections, the frequency response of this system will serve as a tool to conduct a quantitative analysis of the deviation of the scalar approximation of optics from the vector case.

We define the filter transfer functions $H_x\left(k_x, k_y\right)$ and $H_y\left(k_x, k_y\right)$ which multiply $\mathcal{E}_x\left(k_x, k_y\right)$ and $\mathcal{E}_y\left(k_x, k_y\right)$ in Equation 8, respectively, as

$$H_x\left(k_x, k_y\right) = \begin{cases} -\frac{k_x}{\sqrt{k^2 - k_x^2 - k_y^2}} & \text{if } k_x^2 + k_y^2 \neq k^2 \\ 0 & \text{otherwise,} \end{cases} \quad (13)$$

$$H_y\left(k_x, k_y\right) = \begin{cases} -\frac{k_y}{\sqrt{k^2 - k_x^2 - k_y^2}} & \text{if } k_x^2 + k_y^2 \neq k^2 \\ 0 & \text{otherwise.} \end{cases} \quad (14)$$

These filters can be written in the polar coordinates ($k_x = \kappa \cos\phi$, $k_y = \kappa \sin\phi$) as

$$H_x^{polar}\left(\kappa, \phi\right) = \begin{cases} -\frac{\kappa \cos\phi}{\sqrt{k^2 - \kappa^2}} & \text{if } \kappa^2 \neq k^2 \\ 0 & \text{otherwise,} \end{cases} \quad (15)$$

$$H_y^{polar}\left(\kappa, \phi\right) = \begin{cases} -\frac{\kappa \sin\phi}{\sqrt{k^2 - \kappa^2}} & \text{if } \kappa^2 \neq k^2 \\ 0 & \text{otherwise.} \end{cases} \quad (16)$$

In order to find the 2D IFT of the filters, we introduce an auxiliary filter defined as

$$H_0\left(k_x, k_y\right) = \begin{cases} -\frac{1}{\sqrt{k^2 - k_x^2 - k_y^2}} & \text{if } k_x^2 + k_y^2 \neq k^2 \\ 0 & \text{otherwise.} \end{cases} \quad (17)$$

Due to circular symmetry, the 2D IFT of $H_0\left(k_x, k_y\right)$ becomes the inverse Hankel transform of order zero and can be written as (10),

$$\mathcal{F}_{2D}^{-1}\left\{H_0\left(k_x, k_y\right)\right\} = h_0\left(x, y\right) = \frac{j}{2\pi} \frac{e^{jk\sqrt{x^2+y^2}}}{\sqrt{x^2+y^2}}. \quad (18)$$

Then,

$$\begin{aligned} h_x\left(x, y\right) &= \mathcal{F}_{2D}^{-1}\left\{H_x\left(k_x, k_y\right)\right\} \\ &= \mathcal{F}_{2D}^{-1}\left\{k_x H_0\left(k_x, k_y\right)\right\} = \frac{1}{j} \frac{\partial h_o\left(x, y\right)}{\partial x} \\ &= \frac{1}{2\pi} \frac{x}{x^2+y^2} e^{jk\sqrt{x^2+y^2}} \\ &\quad \times \left(jk - \frac{1}{\sqrt{x^2+y^2}}\right). \end{aligned} \quad (19)$$

Similarly,

$$\begin{aligned} h_y\left(x, y\right) &= \mathcal{F}_{2D}^{-1}\left\{H_y\left(k_x, k_y\right)\right\} \\ &= \mathcal{F}_{2D}^{-1}\left\{k_y H_0\left(k_x, k_y\right)\right\} = \frac{1}{j} \frac{\partial h_o\left(x, y\right)}{\partial y} \\ &= \frac{1}{2\pi} \frac{y}{x^2+y^2} e^{jk\sqrt{x^2+y^2}} \\ &\quad \times \left(jk - \frac{1}{\sqrt{x^2+y^2}}\right). \end{aligned} \quad (20)$$

The impulse response of the filters can be written in the polar coordinates ($x = \rho \cos\theta$, $y = \rho \sin\theta$) as

$$h_x^{polar}\left(\rho, \theta\right) = \frac{\cos\theta}{2\pi} \frac{e^{jk\rho}}{\rho}\left(jk - \frac{1}{\rho}\right), \quad (21)$$

$$h_y^{polar}\left(\rho, \theta\right) = \frac{\sin\theta}{2\pi} \frac{e^{jk\rho}}{\rho}\left(jk - \frac{1}{\rho}\right). \quad (22)$$

Please note that, the impulse responses in Equations 21 and 22 that we derive using the properties of the Fourier transform were also obtained using the Green's function of the free space propagation in (11). In this respect, this equivalence can be viewed as similar to the relation between the transfer function and the Green's function of the free space propagation that is developed by Sherman in (12).

If the field is composed only of propagating waves, i.e. $\mathbf{E}_{P'}\left(\mathbf{r}\right) = 0$, then the transfer functions of the filters can be written as

$$H_{x,p}\left(k_x, k_y\right) = H_x\left(k_x, k_y\right) \mathrm{rect}\left(\frac{\sqrt{k_x^2 + k_y^2}}{k}\right), \quad (23)$$

$$H_{y,p}\left(k_x, k_y\right) = H_y\left(k_x, k_y\right) \operatorname{rect}\left(\frac{\sqrt{k_x^2 + k_y^2}}{k}\right), \quad (24)$$

and in polar coordinates as

$$H_{x,p}^{polar}\left(\kappa, \phi\right) = H_x^{polar}\left(\kappa, \phi\right) \operatorname{rect}\left(\frac{\kappa}{k}\right), \quad (25)$$

$$H_{y,p}^{polar}\left(\kappa, \phi\right) = H_y^{polar}\left(\kappa, \phi\right) \operatorname{rect}\left(\frac{\kappa}{k}\right). \quad (26)$$

The auxiliary filter for this case, $H_{0,p}\left(k_x, k_y\right)$, then becomes $H_0\left(k_x, k_y\right) \operatorname{rect}\left(\frac{\sqrt{k_x^2 + k_y^2}}{k}\right)$ and its 2D IFT is (10)

$$\mathcal{F}_{2D}^{-1}\left\{H_{0,p}\left(k_x, k_y\right)\right\} = h_{0,p}\left(x, y\right)$$
$$= -\frac{\sin\left(k\sqrt{x^2 + y^2}\right)}{2\pi\sqrt{x^2 + y^2}}. \quad (27)$$

Again using the derivative property of the FT, the impulse responses of the filters are written as

$$h_{x,p}\left(x, y\right) = \frac{j}{2\pi} \frac{x}{x^2 + y^2}\left[k\cos\left(k\sqrt{x^2 + y^2}\right)\right.$$
$$\left. - \frac{1}{\sqrt{x^2 + y^2}}\sin\left(k\sqrt{x^2 + y^2}\right)\right], \quad (28)$$

$$h_{y,p}\left(x, y\right) = \frac{j}{2\pi} \frac{y}{x^2 + y^2}\left[k\cos\left(k\sqrt{x^2 + y^2}\right)\right.$$
$$\left. - \frac{1}{\sqrt{x^2 + y^2}}\sin\left(k\sqrt{x^2 + y^2}\right)\right]. \quad (29)$$

In polar coordinates, Equations 28 and 29 become

$$h_{x,p}^{polar}\left(\rho, \theta\right)$$
$$= \frac{j}{2\pi} \frac{\cos\theta}{\rho}\left[k\cos\left(k\rho\right) - \frac{1}{\rho}\sin\left(k\rho\right)\right], \quad (30)$$
$$h_{y,p}^{polar}\left(\rho, \theta\right)$$
$$= \frac{j}{2\pi} \frac{\sin\theta}{\rho}\left[k\cos\left(k\rho\right) - \frac{1}{\rho}\sin\left(k\rho\right)\right]. \quad (31)$$

Please note that, since k_z is always real for propagating waves, the impulse response given for the propagating fields in Equation 27, $h_{0,p}\left(x, y\right)$, is the real part of the impulse response given for a general field in Equation 18, $h_0\left(x, y\right)$. In order to show this, we first define the filter

corresponding to evanescent components as

$$H_{0,p'}\left(k_x, k_y\right)$$
$$= H_0\left(k_x, k_y\right)\left[1 - \operatorname{rect}\left(\frac{\sqrt{k_x^2 + k_y^2}}{k}\right)\right]. \quad (32)$$

Then,

$$\mathcal{F}_{2D}^{-1}\left\{H_0\left(k_x, k_y\right)\right\}$$
$$= \mathcal{F}_{2D}^{-1}\left\{H_{0,p}\left(k_x, k_y\right) + H_{0,p'}\left(k_x, k_y\right)\right\}$$
$$= \mathcal{F}_{2D}^{-1}\left\{H_{0,p}\left(k_x, k_y\right)\right\} + \mathcal{F}_{2D}^{-1}\left\{H_{0,p'}\left(k_x, k_y\right)\right\}$$
$$= h_{0,p}\left(x, y\right) + h_{0,p'}\left(x, y\right)$$
$$= h_0\left(x, y\right). \quad (33)$$

$H_{0,p}\left(k_x, k_y\right)$ is a real and even function of k_x and k_y. From the symmetry property of the FT, the 2D IFT of $H_{0,p}\left(k_x, k_y\right)$ becomes an even and real function of x and y. Similarly, $H_{0,p'}\left(k_x, k_y\right)$ is a purely imaginary and even function and hence, its 2D IFT also becomes even and purely imaginary. Therefore, $h_{0,p}\left(x, y\right) = Re\left\{h_0\left(x, y\right)\right\}$, where the $Re\{\cdot\}$ operator gives the real part of its input. Then since the derivative of a real valued function remains real valued,

$$\frac{1}{j}\frac{\partial h_{0,p}\left(x, y\right)}{\partial x}$$
$$= \frac{1}{j}\frac{\partial Re\left\{h_0\left(x, y\right)\right\}}{\partial x} = jIm\left\{\frac{1}{j}\frac{\partial h_0\left(x, y\right)}{\partial x}\right\}, \quad (34)$$

where the $Im\{\cdot\}$ operator gives the imaginary part of its input. So, $h_{x,p}\left(x, y\right) = jIm\left\{h_x\left(x, y\right)\right\}$, and similarly $h_{y,p}\left(x, y\right) = jIm\left\{h_y\left(x, y\right)\right\}$.

If the field is composed only of evanescent waves, i.e. if k_z is always purely imaginary, then by following a similar approach, it can be shown that

$$h_{x,p'}\left(x, y\right) = Re\left\{h_x\left(x, y\right)\right\} \text{ and}$$
$$h_{y,p'}\left(x, y\right) = Re\left\{h_y\left(x, y\right)\right\}, \quad (35)$$

where $h_{x,p'}\left(x, y\right)$ and $h_{y,p'}\left(x, y\right)$ are the filter impulse responses for fields which involve only evanescent waves.

Although the impulse responses corresponding to the arbitrary monochromatic fields are complex valued, as given in Equations 19 and 20, if the field is known to be propagating or evanescent, then the impulse responses of the filters become purely imaginary or real, respectively. Therefore, such a knowledge reduces the computational
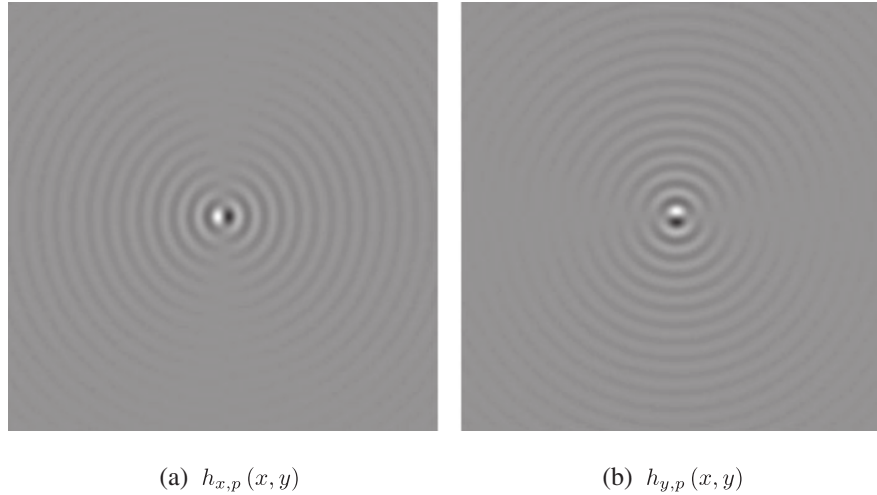
(a) $h_{x,p}(x,y)$ (b) $h_{y,p}(x,y)$

Figure 1. For the propagating field case, a finite portion of the imaginary parts of the impulse responses of the filters are shown. (x; y) = (0; 0) corresponds to the centre of the images; the horizontal axis is x and the vertical axis is y.

complexity in the computation of the z component from the x and y components.

A finite portion of the imaginary part of the impulse response of the filters $h_{x,p}(x,y)$ and $h_{y,p}(x,y)$ are shown in Figure 1. The block diagram of the computation of $E_z(x,y,0)$ is shown in Figure 2. In the figure, propagation in free space is also included as a block.

4. Properties of the filters

In this section, we describe the frequency selectivity characteristics of the filters defined in Section 3 for the propagating case. By inspection, it can be seen that the 2D filters $H_{x,p}(k_x,k_y)$ and $H_{y,p}(k_x,k_y)$ are band-pass filters with a circular discontinuity at those values of k_x and k_y which satisfy $k_x^2 + k_y^2 = k^2$. However, within their pass-bands, they show high-pass characteristics.

Firstly, for the sake of illustration, we plot the magnitude of the radial part of $H_{x,p}^{polar}(\kappa,\phi)$ or $H_{y,p}^{polar}(\kappa,\phi)$, which is given by $-\frac{\kappa}{\sqrt{k^2-\kappa^2}}\text{rect}\left(\frac{\kappa}{k}\right)$ from Equations 25 and 26, with the normalized frequency variable $\nu = \kappa/k$, in Figure 3.

As can be seen from Figure 3, the magnitude response of the filters increases rather slowly from the center in both positive and negative directions until $|\nu| \approx 0.75$. However, beyond these points, the magnitude response of the filters increases sharply, and as ν goes to -1 from the right and $+1$ from the left, the filter magnitude tends to infinity.

In Figure 4, the 2D magnitude response of both filters with the normalized frequency variables ν_x and ν_y are shown as 3D surface plots.

The characteristics of the filters observed in Figures 3 and 4 also help to show that the scalar theory can be used

as an approximation to the vector theory under some conditions. For example, it is clear that if the field is restricted only to those small angles around zero (the paraxial approximation), then the magnitude of the filters will be small enough to justify neglecting the E_z component. Under this circumstance, the scalar field we are working with may be interpreted as either E_x or E_y, or any linear combination of those. Other valid approximations are also possible. For example, any narrow-band of angles around a given direction will also yield an accurate enough scalar approximation together with a constant E_z component. In Section 5, we discuss this topic in more detail.

As an additional property, in Appendix 1, we prove that the integrals $\frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|H_{x,p}(k_x,k_y)|^2\,dk_x dk_y$ and $\frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|H_{y,p}(k_x,k_y)|^2\,dk_x dk_y$ diverge. Please note that, being the magnitude square integrals of the transfer functions divergent does not mean that the monochromatic field possesses infinite energy. Since the actual spatial bandwidth of the monochromatic field is restricted by the spatial bandwidth of the input field components, the energy of the z component is also restricted by the energy of the multiplication of the 2D FTs of the x and y components by the transfer functions of the filters given by Equations 23 and 24.

5. Vector vs. scalar modelling of optical diffraction

In order to reduce the computational cost, vector fields may be mapped to scalar fields under some restrictions. In the following paragraphs, we describe such mappings found in the literature that we are aware of.
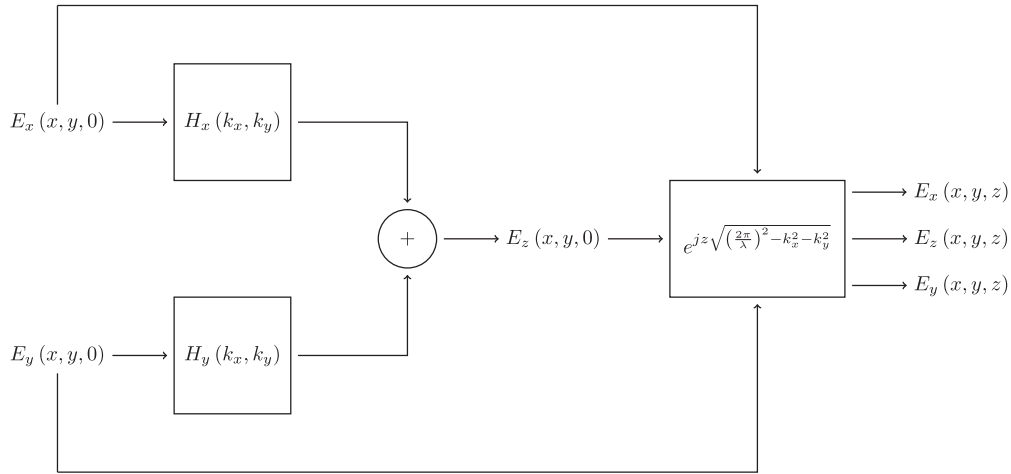
Figure 2. The computation steps of the z component of the electric field, including free space propagation, are shown as a block diagram.

The most common approach is to solve for a scalar field and then map this scalar field to one of the transverse components of the vector electromagnetic field (2, 13) and assume the longitudinal component to be zero. An immediate extension is to map the scalar field to the complex amplitude of the electric field along a fixed direction within the (x, y) plane; such a field satisfies the same Helmholtz's equation (2). When dealing with free space propagation, if the main propagation direction of the light beam is chosen to lie within a small cone around the z axis and the primary goal is to compute the intensity of the wave, then such a map from a scalar field to a vector electromagnetic field gives the intensity value with negligible error. If the paraxial condition is satisfied, the main contribution to the intensity of the vector wave comes from the dominant component which is orthogonal to the propagation direction. In (14), for example, the validity of such a scalar approximation was showed for a paraxial imaging setup. Since k_x and k_y for each plane wave component are small in the paraxial case, by the virtue of the high-pass characteristics within the pass bands of the filters $H_{x,p}(k_x, k_y)$ and $H_{y,p}(k_x, k_y)$, the z component becomes small and its contribution to the intensity becomes negligible. Otherwise, if k_x or k_y is large, then the z component cannot be neglected and the scalar approximation becomes erroneous during the intensity computation.

Another approach is to map the complex amplitude of the scalar plane wave to a vector plane wave whose intensity is proportional to the intensity of the scalar wave (1, 15, 16). With this approach, the intensity of a single vector plane wave and the corresponding scalar plane wave can be exactly matched at each point in space as it is constant everywhere. Moreover, if a vector plane wave is linearly polarized, that wave can be described exactly
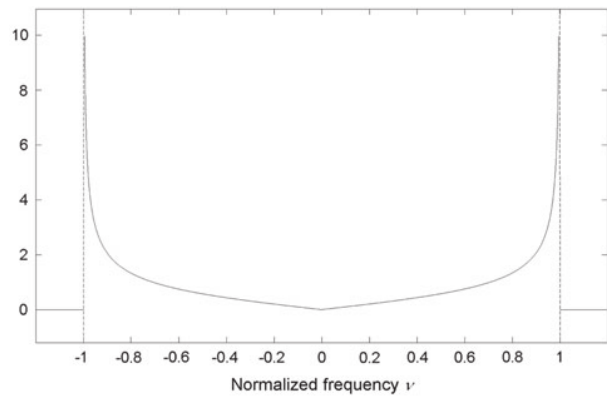

Figure 3. The magnitude of the radial variation of $H_{x,p}^{polar}(\kappa, \phi)$ or $H_{y,p}^{polar}(\kappa, \phi)$ with the normalized frequency variable $v = \kappa/k$ is shown.

by the corresponding scalar plane wave in terms of the magnitude and phase even if the wave is not paraxial. For example, for a given (k_x, k_z) pair, the vector wave $(\hat{a}_x E_x + \hat{a}_z E_z) e^{j(k_x x + k_z z)}$ and the scalar wave $\sqrt{|E_x|^2 + |E_z|^2} e^{j(k_x x + k_z z + \phi)}$, give the same intensity value up to a constant multiplier at each point in space. Here, $\hat{a}_x$ and $\hat{a}_z$ are the unit vectors pointing along the x and z directions, $E_z = E_x H_{x,p}(k_x, k_y)$ is the z component of the field, $|E_x|$ and $|E_z|$ are the respective magnitudes, and ϕ denotes the angular component of the complex number E_x (and hence describes the phase of the linearly polarized wave). However, if we cannot make the narrow-band assumption, plane waves propagating along different directions will interfere, and consequently, the total intensity of the field computed by the scalar theory will not be negligible. In other words, if the vector plane wave amplitude is mapped to the corresponding scalar plane wave amplitude and if the z component is large, then the computed intensity in the scalar domain highly
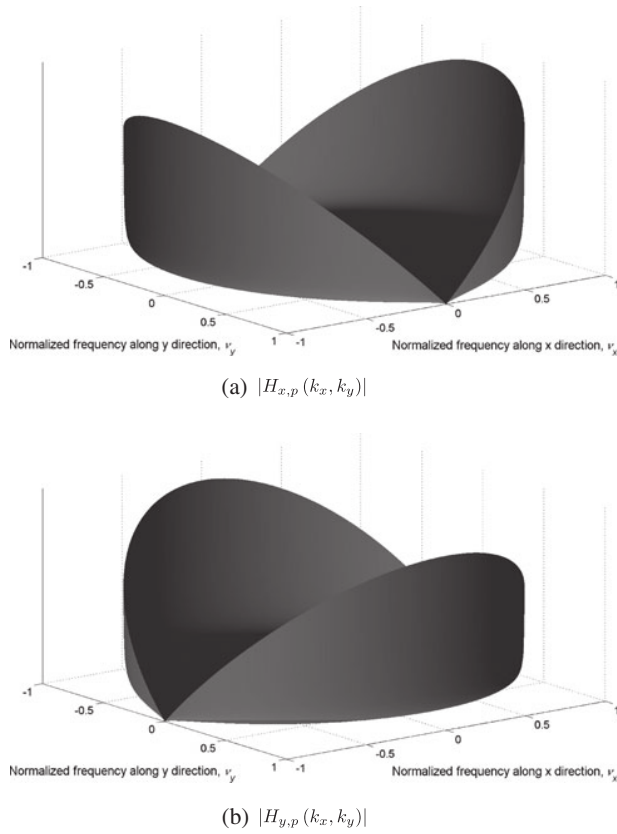
(a) $|H_{x,p}(k_x, k_y)|$

(b) $|H_{y,p}(k_x, k_y)|$

Figure 4. The magnitude response of the filters are shown as 3D surface plots, where v_x and v_y are the normalized frequency variables: $v_x = k_x/k$ and $v_y = k_y/k$.

deviates from the actual intensity computed according to the vector theory.

To sum up, the above two mappings from scalar to vector waves are valid for narrow-band fields. That is, if the main propagation axis of the light beam is chosen as the z axis and if k_x and k_y are small, then the z component, which is the cause of the error in the scalar approximations, also becomes small. This narrow-band requirement for the scalar field, for instance, imposes that the angle of the reference beam cannot deviate too much from the object beam in off-axis holography. In this respect, the ratio of the energy of the z component to the energy of the total electric field at the $z = 0$ plane can be seen as an error measure of the scalar approximation. This measure can be formulated as

Moreover, since the energy of the field is preserved in the free-space propagation, the error measure given in Equation 36 does not change at different $z = d$ planes.

A detailed treatment of the scalar representations of different paraxial and nonparaxial beam solutions of Maxwell's equations together with further references is given in (*13*). The validity of the scalar and paraxial approximations for Gaussian beam is considered in (*7*, *17*, *18*).

Another, somewhat different approach is to consider the scalar field to be a physical quantity distinct from the electric or magnetic field components. For example, in (*19–21*), a scalar representation is developed by assuming the two independent components of the real magnetic vector potential as the real and imaginary parts of the complex scalar field. Then, the analytical form of the corresponding energy and momentum densities are computed based on the developed scalar representation.

6. Digital simulator

In this section, we present a digital simulator to compute the longitudinal component of the electric field from its transversal components all at $z = 0$ plane. So, there is no propagation involved in the simulator.

First of all, since the field components are assumed to be propagating, the spatial frequency content of the transversal field components are confined in a circular band whose radius is smaller than the wave number. The energy of the impulse response of the filters is found in Appendix 1 to be

$$-\frac{\kappa'^2}{8\pi} - \frac{k^2}{8\pi} \ln\left[1 - \left(\frac{\kappa'}{k}\right)^2\right], \qquad (37)$$

where $\kappa' < k$ is the radius of the imposed circular pass-band of the filters that can be taken as the larger one of the radius' of the pass-bands of the x and y components of the field. Therefore, although the resulting bandwidth of the related signals is smaller than k, in order to satisfy the Shannon-Nyquist criteria under regular rectangular sampling scheme, we use π/k as the sampling period, since the pass-band of the field components may vary, but cannot be larger than or equal to k.

$$\frac{\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\left|H_x\left(k_x, k_y\right)\mathcal{E}_x\left(k_x, k_y\right) + H_y\left(k_x, k_y\right)\mathcal{E}_y\left(k_x, k_y\right)\right|^2 dk_x dk_y}{\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\left(\left|\mathcal{E}_x\left(k_x, k_y\right)\right|^2 + \left|\mathcal{E}_y\left(k_x, k_y\right)\right|^2 + \left|H_x\left(k_x, k_y\right)\mathcal{E}_x\left(k_x, k_y\right) + H_y\left(k_x, k_y\right)\mathcal{E}_y\left(k_x, k_y\right)\right|^2\right) dk_x dk_y}. \qquad (36)$$

We form the x component of the discrete input field $\hat{E}_x[n, m, 0]$, from the continuous field $E_x(x, y, 0)$ with the integer variables n and m as

$$\hat{E}_x[n, m, 0] = \begin{cases} E_x(nX_s, mY_s, 0) & \text{if } 0 \leq n \leq N-1, \\ & 0 \leq m \leq M-1 \\ 0 & \text{otherwise,} \end{cases}$$

(38)

where the discrete input field size is assumed to be $N \times M$. Then, we convert this finite size input to the periodically replicated version as

$$\tilde{E}_x[n, m, 0] = \hat{E}_x\left[n_{\text{modN}}, \ m_{\text{modM}}\right].$$

(39)

Therefore, the finite size discrete input field is one period of the infinite size discrete periodic input field. $\tilde{E}_x[n, m, 0]$ has only finite number of frequency components and can be written as (22)

$$\tilde{E}_x[n, m, 0]$$
$$= \frac{1}{NM} \sum_{p=0}^{N-1} \sum_{q=0}^{M-1} D_x[p, q] e^{j2\pi\left(\frac{pn}{N} + \frac{qm}{M}\right)}.$$

(40)

So,

$$D_x[p, q] = IDFT_{N \times M}\left\{\hat{E}_x[n, m, 0]\right\},$$

(41)

where $IDFT_{N \times M}\{\cdot\}$ stands for the 2D, size $N \times M$ inverse discrete Fourier transform (DFT) (22).

The discrete filter corresponding to the x component in the DFT domain, $\hat{H}_{x,p}[p, q]$, is generated from the transfer function of the filter defined in Equation 23 as

$$\hat{H}_{x,p}[p, q]\Big|_{\substack{p=\hat{p}_{\text{modN}} \\ q=\hat{q}_{\text{modM}}}} = H_{x,p}\left(\frac{2k}{N}\hat{p}, \frac{2k}{M}\hat{q}\right),$$

(42)

where $\hat{p} \in \left[-\left\lceil\frac{N-1}{2}\right\rceil, \left\lceil\frac{N}{2}-1\right\rceil\right]$, $\hat{q} \in \left[-\left\lceil\frac{M-1}{2}\right\rceil, \left\lceil\frac{M}{2}-1\right\rceil\right]$, $\lceil\cdot\rceil$ is the ceiling operator which rounds a decimal number to the nearest larger integer and $k = 2\pi/\lambda$, as usual. The generation procedure of the y component and the corresponding filter in the DFT domain, $D_y[p, q]$ and $\hat{H}_{y,p}[p, q]$, is the same as the procedure described above.

After computing the 2D DFT coefficients of the field components and the transfer functions of the filters, we compute the discrete system output as

$$\hat{E}_z[n, m, 0] = IDFT_{N \times M}\left\{D_x[p, q]\hat{H}_{x,p}[p, q]\right.$$
$$\left. + D_y[p, q]\hat{H}_{y,p}[p, q]\right\}$$

(43)

$$= IDFT_{N \times M}\left\{D_z[p, q]\right\}.$$

(44)

Similar to the input field components, the finite size discrete output field, $\hat{E}_z[n, m, 0]$, can be assumed as the one period of the periodic infinite size discrete field $\tilde{E}_z[n, m, 0]$.

As a result, if the continuous input fields include only the spatial frequency components $2k\hat{p}/N$ and $2k\hat{q}/M$ for $\hat{p} \in \left[-\left\lceil\frac{N-1}{2}\right\rceil, \left\lceil\frac{N}{2}-1\right\rceil\right]$ and $\hat{q} \in \left[-\left\lceil\frac{M-1}{2}\right\rceil, \left\lceil\frac{M}{2}-1\right\rceil\right]$ along k_x and k_y directions, respectively, then the simulator output gives the exact samples of the corresponding periodic continuous system output. Moreover, since the sampling period is chosen small enough, the corresponding output of the continuous system can be exactly found from the output of the discrete simulator. In other words, in an application which deals only with propagating waves, if N and M are chosen such that entire spatial frequency content of the periodic continuous field is covered in the discrete system, then the output of the digital simulator is useful to accurately simulate the continuous field.

We note that although here we use a regular rectangular sampling scheme, hexagonal sampling can also be used in an application which requires maximally efficient memory use and less computational complexity, since it is more efficient for signals with a circular band area (23).

We present a simulation where a test pattern that is commonly used to reveal the 2D frequency response of a 2D system is used as the input: The x and y components of the input are chosen to be the same as $E_x[n, m, 0] = E_y[n, m, 0] = \cos\left[\frac{\pi}{N}\left(n^2 + m^2\right)\right]$ as shown in Figure 5. Please note that for this 2D chirp signal, the parameter N specifies both the size and the rate of change of instantaneous frequency of the discrete field. When $N = M$ and N is an even number, the 2D DFT of this function becomes $N \sin\left[\frac{\pi}{N}\left(p^2 + q^2\right)\right]$ (24). As the instantaneous frequency of this chirp function linearly increases in the interval $n \in [0, \ N/2]$ and $m \in [0, \ M/2]$, we expect to observe the approximate frequency and orientation selectivity characteristics of the system shown in Figure 2, when N is sufficiently large. In our simulation, we choose $N = 16384 = 2^{14}$. As seen from the output shown in Figure 6, the system frequency response is observed as expected for this test input.

Figure 6 reveals the orientation characteristics of the system. When k_x and k_y have the same sign, the filters add coherently and the response of the filters becomes large, and when they have different signs, the filters cancel each other out and the response of the system becomes weak. As a consequence of the discussion in Section 5,
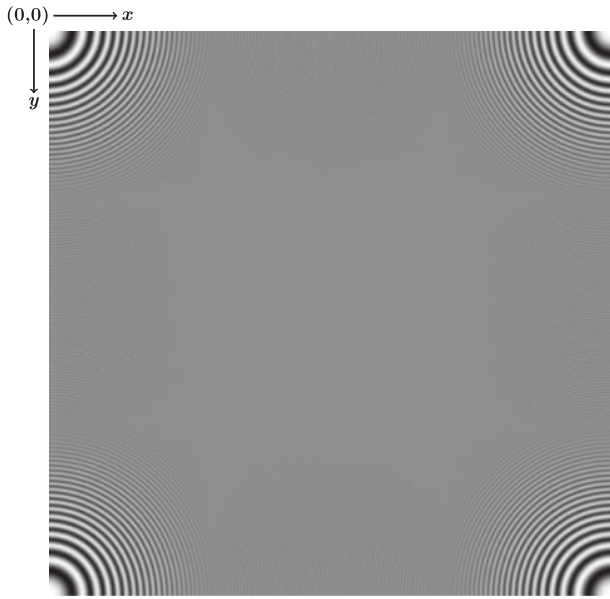
(0,0) ⟶ x

↓ y

Figure 5. One period of the x component of the input field for the simulation is shown. Since the x and y components are the same, this figure also represents the y component of the input field.
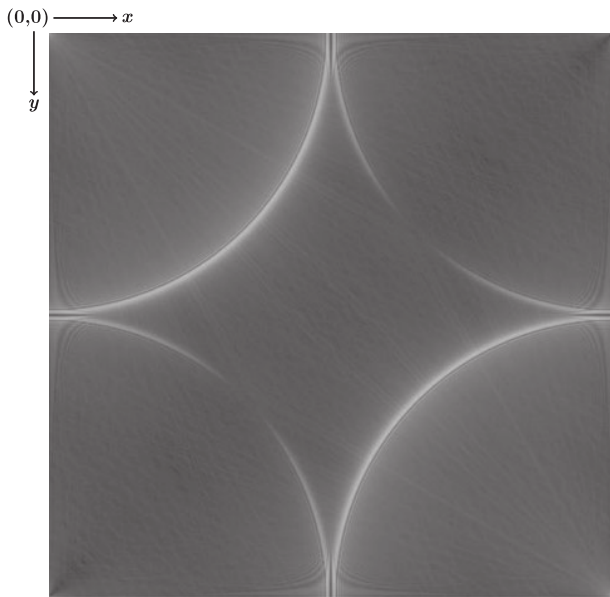
(0,0) ⟶ x

↓ y

Figure 6. For $N = 2^{14}$, the magnitude of the z component corresponding to the input fields $E_x[n, m, 0] = E_y[n, m, 0] = \cos\left[\frac{\pi}{N}\left(n^2 + m^2\right)\right]$ is shown at a logarithmic scale for better illustration.

it can be said that, if the field components were, for example, this chirp signal, the scalar theory would fail in the spatial positions in which the magnitude of the z component is high. Please note that, in this case, the response of the system is small not only on the points near the low frequencies, but also the frequencies in which the

magnitudes of k_x and k_y are close to each other, but their signs are opposite. Therefore, in these spatial points, the error in the scalar theory would be less, even if the field components include high frequency components.

Furthermore, this discrete simulator can be used when the continuous system inputs are such that the output cannot be calculated analytically. An evaluation of the amount of error, when the scalar approximation is used, is possible using the provided simulator.

7. Conclusions

The main purpose of this paper is to present a tool which is useful in the quantitative analysis of the deviation from the vector theory of optics when the scalar theory is used instead. This tool is based on signals and systems formulation of monochromatic electromagnetic wave propagation in free space. Our approach involves the computation of the z component of the electric field from the x and y components by a simple linear-shift invariant filtering operation. We present analytical expressions for the impulse responses of the filters, derived from their transfer functions. The expressions for the transfer functions and their impulse responses cover both propagating and evanescent cases.

Interpretation of the system in the frequency domain indicates the source and amount of error when a scalar approximation is used, and provides a means to eliminate such errors by providing correct vector components. According to the vector to scalar mappings found in the literature, the source and amount of error is found to be related to the strength of the longitudinal component of the electric field, which is also directly related to the magnitude response of the system.

Finally, we developed an appropriate digital simulator that successfully computes the z component from the x and y components of the electric field for the propagating case. The simulator output gives the exact samples of the continuous longitudinal field if the filter size is chosen such that each frequency content of the continuous input fields is represented in the 2D DFT computations. Furthermore, if this condition is satisfied, the exact continuous longitudinal field can also be found from the discrete simulation output. By investigating the output field, it can be decided whether it is appropriate to use the scalar theory for the given input fields. Moreover, the simulator output in the space domain can be partitioned into local regions according to their magnitudes. Then, it can be said that the scalar theory highly deviates from the actual behaviour of the field in those local regions with high magnitude.

Acknowledgements

Disclosure statement

No potential conflict of interest was reported by the authors.

References

(1) Saleh, B.E.A.; Teich, M.C. *Fundamentals of Photonics*, 2nd ed.; Wiley: Hoboken, NJ, 2007.
(2) Goodman, J.W. *Introduction to Fourier Optics*, 2nd ed.; McGraw-Hill: New York, 1996.
(3) Balanis, C.A. *Advanced Engineering Electromagnetics*, 2nd ed.; Wiley: Hoboken, NJ, 2012.
(4) Balanis, C.A. *Antenna Theory: Analysis and Design*, 3rd ed.; Wiley: Hoboken, NJ, 2005.
(5) Collin, R.E. *Antennas and Radiowave Propagation*; McGraw-Hill: New York, 1985.
(6) Borgiotti, G. *Alta Frequenza* **1963,** *32,* 196–204.
(7) Chen, C.G.; Konkola, P.T.; Ferrera, J.; Heilmann, R.K.; Schattenburg, M.L. *J. Opt. Soc. Am. A* **2002,** *19,* 404–412.
(8) Chaumet, P.C.; . *J. Opt. Soc. Am. A* **2006,** *23,* 3197–3202.
(9) Carter, W.H. *J. Opt. Soc. Am. A* **1972,** *62,* 1195–1201.
(10) Erdélyi, A.; Magnus, W.; Oberhettinger, F. *Tables of Integral Transforms*; McGraw-Hill: New York, 1954.
(11) Luneburg, R.K.; Herzberger, M. *Mathematical Theory of Optics*; University of California Press: Los Angeles, CA, 1964.
(12) Sherman, G.C. *J. Opt. Soc. Am.* **1967,** *57,* 546–547.
(13) Martínez-Herrero, R.; Mejías, P.M.; Piquero, G. *Characterization of Partially Polarized Light Fields*; Springer: Berlin, Heidelberg, 2009.
(14) Theimer, O.; Wassermann, G.D.; Wolf, E. *Proc. Roc. Soc. A* **1952,** *212,* 426–437.
(15) Gaskill, J.D. *Linear Systems, Fourier Transforms, and Optics*; Wiley: Hoboken, NJ, 1978.
(16) Marathay, A.S.; Parrent, G.B. *J. Opt. Soc. Am.* **1970,** *60,* 243–245.
(17) Varga, P.; Török, P. *Opt. Commun.* **1998,** *152,* 108–118.
(18) Carnicer, A.; Juvells, I.; Maluenda, D.; Martinez-Herrero, R.; Mejías, P.M. *Eur. J. Phys.* **2012,** *33,* 1235–1247.
(19) Green, H.S.; Wolf, E. *Proc. Phys. Soc. A* **1953,** *66,* 1129–1137.
(20) Wolf, E. *Proc. Phys. Soc. London* **1959,** *74,* 269–280.
(21) Roman, P. *Proc. Phys. Soc.* **1959,** *74,* 281–289.
(22) Oppenheim, A.V.; Schafer, R.W.; Buck, J.R. *Discrete-Time Signal Processing*, 2nd ed.; Prentice Hall: Upper Saddle River, NJ, 1999.
(23) Mersereau, R. *Proc. IEEE* **1979,** *67,* 930–949.
(24) Onural, L. *Opt. Eng.* **2004,** *43,* 2557–2563.
(25) Gradshteyn, I.S.; Ryzhik, I.M. *Table of Integrals, Series, and Products* , 7th ed.; Academic Press: San Diego, CA, 2007.

Appendix 1.

Here, we prove that the energy of the filters in the propagating case is infinite. We give the analysis only for the filter $H_{x,p}(k_x, k_y)$ since the analysis for $H_{y,p}(k_x, k_y)$ is the same.

The energy, $\epsilon_{\kappa'}$, of $H_{x,p}^{polar}(\kappa, \phi)$, which is defined in Equation 25, for the interval $0 \leq \kappa \leq \kappa'$ can be written as

$$
\begin{aligned}
\varepsilon_{\kappa'} &= \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{\kappa'} \left| H_{x,p}^{polar}(\kappa, \phi) \right|^2 \kappa \, d\kappa \, d\phi \\
&= \frac{1}{4\pi^2} \int_0^{2\pi} \cos^2 \phi \int_0^{\kappa'} \frac{\kappa^3}{k^2 - \kappa^2} \, d\kappa \, d\phi
\end{aligned} \tag{A1}
$$

Then,

$$
\begin{aligned}
\int_0^{\kappa'} \frac{\kappa^3}{k^2 - \kappa^2} \, d\kappa &= -\frac{\kappa^2}{2} \Big|_{\kappa=0}^{\kappa=\kappa'} + k^2 \int_0^{\kappa'} \frac{\kappa}{k^2 - \kappa^2} \, d\kappa \\
&= -\frac{\kappa'^2}{2} - \frac{k^2}{2} \ln\left(1 - \kappa^2\right) \Big|_{\kappa=0}^{\kappa=\frac{\kappa'}{k}} \\
&= -\frac{\kappa'^2}{2} - \frac{k^2}{2} \ln\left[1 - \left(\frac{\kappa'}{k}\right)^2\right]
\end{aligned} \tag{A2}
$$

The integrals in Equation A2 are obtained from (25). When κ' tends to k, as a result of the logarithm in Equation A2, the energy of $H_{x,p}(k_x, k_y)$ tends to infinity. If we choose $\kappa' < k$, then the energy of the filter for $0 \leq \kappa \leq \kappa'$ becomes

$$
-\frac{\kappa'^2}{8\pi} - \frac{k^2}{8\pi} \ln\left[1 - \left(\frac{\kappa'}{k}\right)^2\right]. \tag{A3}
$$