PAPER • OPEN ACCESS

Field theory of monochromatic optical beams: I. Classical fields

Field theory of monochromatic optical beams: I. Classical fields

Andrea Aiello

Max Planck Institute for the Science of Light, Staudtstr. 2, D-91058 Erlangen, Germany

E-mail: andrea.aiello@mpl.mpg.de

Abstract

We study monochromatic, scalar solutions of the Helmholtz and paraxial wave equations (PWEs) from a field-theoretic point of view. We introduce appropriate time-independent Lagrangian densities for which the Euler–Lagrange equations reproduces either Helmholtz and PWEs with the z-coordinate, associated with the main direction of propagation of the fields, playing the same role of time in standard Lagrangian theory. For both Helmholtz and paraxial scalar fields, we calculate the canonical energy-momentum tensor and determine the continuity equations relating 'energy' and 'momentum' of the fields. Eventually, the reduction of the Helmholtz wave equation to a useful first-order Dirac form, is presented. This work sheds some light on the intriguing and not so acknowledged connections between angular spectrum representation of optical wavefields, cosmological models and physics of black holes.

Keywords: Helmholtz wave equation, paraxial wave equation, field theory

1. Introduction

Light is an electromagnetic phenomenon which can be described by a field theory governed by Maxwell's equations. These are a set of first-order partial differential equations that relates electric and magnetic *vector* fields each other and to possible sources. However, in many practical instances, a vector field representation of light appears redundant and a simpler *scalar* field description results appropriate. This is not the case for structured-light beams with nontrivial polarization patterns as the ones considered, for example, in [1]. In this paper, we will not consider such a kind of light. Thus, when a scalar description is appropriate, according to the characteristics of the phenomenon under investigation, *monochromatic light* propagating in free space can be described either by a field $\psi(\mathbf{x}, z) = \psi(x, y, z)$ obeying the Helmholtz wave equation (HWE)

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} + k_0^2\right)\psi(\mathbf{x}, z) = 0, \qquad k_0 > 0, \quad (1)$$

or by a field $\phi(\mathbf{x}, z) = \phi(x, y, z)$ satisfying the paraxial wave equation (PWE)

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + 2ik_0\frac{\partial}{\partial z}\right)\phi(\mathbf{x}, z) = 0, \qquad k_0 > 0, \quad (2)$$

where here and hereafter we use the notation $\mathbf{x} = (x, y) \in \mathbb{R}^2$.

In the appendix XI of their book '*Principles of Optics*', Born and Wolf derive the energy conservation law for a *real*, time-dependent scalar wavefield $V^{(r)}(\mathbf{r}, t)$ in free space, with $\mathbf{r} = (x, y, z) \in \mathbb{R}^3$ [2]. Because of the explicit time dependence of $V^{(r)}(\mathbf{r}, t)$, a continuity equation expressing the local energy conservation law could be deduced from the Lagrangian form of field equations. For the case of a monochromatic field of frequency ω, Born and Wolf first rewrite the field as the real part of a time-harmonic complex amplitude, that is $V^{(r)}(\mathbf{r}, t) = \mathrm{Re}\{U(\mathbf{r}, \omega)e^{-i\omega t}\}$. Then, they take the time averages of the energy density and the energy flux vector to obtain conservation laws involving only the time-independent *complex* field $U(\mathbf{r}, \omega)$.

In this work, we pursue the same goal of Born and Wolf, yet following an entirely different and new approach. Instead of considering time-dependent monochromatic fields and erasing such dependence via time averages, we develop an *ab initio* time-independent field theory taking the monochromatic Helmholtz and PWEs (1) and (2) as the central points around which we build a time-independent Lagrangian field theory [3]. The idea is to deal with *action functionals* of

the form

$$S = \int_{z_1}^{z_2} d\mathbf{r}\, \mathcal{L}, \tag{3}$$

where z_1 and z_2 are the limits of integration for the variable z which is associated with the main propagation direction of the field, $d\mathbf{r} = dx\, dy\, dz$ is the volume measure and $\mathcal{L}$ denotes the Lagrange density (Lagrangian, for short). We require S to be stationary for arbitrary variations of the field quantities that vanish at the end points, namely

$$\delta S = 0, \tag{4}$$

in order to infer the Euler–Lagrange equations reproducing (1) and (2). Thus, in our *nonstandard* approach, propagation along the z-axis of a time-independent field obeying either HWE or PWE, is formally described in the same manner the time evolution of a time-dependent field is depicted in the *standard* Lagrangian formalism.

2. Nonstandard Lagrangian formalism for Helmholtz fields

In this section, we discuss the classical mechanics of a *complex* scalar field $\psi = \psi(\mathbf{x}, z)$, which is a solution of the HWE (1)

$$(\partial^2 + k_0^2)\psi = 0, \tag{5}$$

where the Laplacian ∂^2 in the three-dimensional Euclidean space $\mathbb{R}^3$ is written as (throughout this paper, we use Einstein's summation convention)

$$\partial^2 = \partial_\mu \partial^\mu = g^{\mu\nu}\partial_\mu\partial_\nu, \qquad (\mu, \nu = 1, 2, 3), \tag{6}$$

with $g^{\mu\nu} = \delta^{\mu\nu}$. The three-dimensional gradient is expressed as $\partial\psi/\partial x^\mu \equiv \partial_\mu\psi = (\boldsymbol{\nabla}, \partial_z)\psi$, where a point in $\mathbb{R}^3$ is labeled by the three coordinates x^μ, with $x^3 = z$ the *longitudinal* coordinate and x^k, $k = 1, 2$ the *transverse* coordinates. The *two-dimensional* gradient of a scalar function $f(x, y, z)$ is denoted $\boldsymbol{\nabla} f$ and is defined as

$$\boldsymbol{\nabla} f = \frac{\partial f}{\partial x}\boldsymbol{\epsilon}_1 + \frac{\partial f}{\partial y}\boldsymbol{\epsilon}_2, \tag{7}$$

where $\boldsymbol{\epsilon}_1$ and $\boldsymbol{\epsilon}_2$ are the orthogonal unit vectors pointing in the x and y Cartesian coordinate directions, respectively. From now on, Greek indexes $\mu, \nu, \alpha, \beta, \ldots$, run from 1 to 3, while Latin indexes $i, j, k, l, m, n, \ldots$, take the values 1 and 2.

When the complex scalar field ψ has two *independent* real components ψ_1 and ψ_2, we may put

$$\psi = (\psi_1 + i\psi_2)/\sqrt{2}, \tag{8}$$

$$\psi^* = (\psi_1 - i\psi_2)/\sqrt{2}, \tag{9}$$

and regard to ψ and ψ^* (instead of ψ_1 and ψ_2) as independent fields. In this case, we expect that the two Euler–Lagrange equations

$$\frac{\partial\mathcal{L}}{\partial\psi^*} - \frac{\partial}{\partial x^\mu}\frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi^*)} = 0, \tag{10}$$

$$\frac{\partial\mathcal{L}}{\partial\psi} - \frac{\partial}{\partial x^\mu}\frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)} = 0, \tag{11}$$

yield to the two Helmholtz equations

$$(\partial^2 + k_0^2)\psi = 0, \tag{12}$$

$$(\partial^2 + k_0^2)\psi^* = 0, \tag{13}$$

respectively, when an appropriate Lagrangian $\mathcal{L} = \mathcal{L}_{\text{HWE}}$ is chosen. Thus, our goal now is to find a proper Lagrangian $\mathcal{L}_{\text{HWE}}$. To this end, suppose that

$$\mathcal{L}_{\text{HWE}} = A\,\delta^{\mu\nu}\partial_\mu\psi^*\partial_\nu\psi + B\,\psi^*\psi, \tag{14}$$

where A, B are real constants to be determined. Substitution of equation (14) into (10) gives

$$B\,\psi - A\,\partial^2\psi = 0, \tag{15}$$

which coincides with (12) if one chooses $B = -A\,k_0^2$ (A remains undetermined, therefore we are free to choose $A = 1$). Thus, the sought Lagrangian is

$$\mathcal{L}_{\text{HWE}} = \delta^{\mu\nu}\partial_\mu\psi^*\partial_\nu\psi - k_0^2\,\psi^*\psi. \tag{16}$$

From the action principle, it follows that is always possible to add a three-divergence to the Lagrangian (16) without altering the equation of motion (12). Therefore, since

$$\partial_\mu(\delta^{\mu\nu}\psi^*\partial_\nu\psi) = \delta^{\mu\nu}\partial_\mu\psi^*\partial_\nu\psi + \psi^*\partial^2\psi, \tag{17}$$

then we can rewrite equation (16) in the following *equivalent* form:

$$\mathcal{L}_{\text{HWE}} = \psi^*(\partial^2 + k_0^2)\psi. \tag{18}$$

The Euler–Lagrange equation (10) now simply becomes

$$\frac{\partial\mathcal{L}}{\partial\psi^*} = (\partial^2 + k_0^2)\psi = 0. \tag{19}$$

2.1. Structural aspects

Equation (12) admits *separable* solutions of the form

$$\psi(\mathbf{x}, z) = \varphi(\mathbf{x})e^{i\zeta z}, \tag{20}$$

where by definition $\varphi(\mathbf{x})$ satisfies

$$\zeta^2\varphi(\mathbf{x}) = (\nabla^2 + k_0^2)\varphi(\mathbf{x}). \tag{21}$$

Elementary examples thereof are given by plane wave fields

$$\psi_{\text{PW}}(\mathbf{x}, z) = \exp(i\mathbf{p}\cdot\mathbf{x})\exp[iz(k_0^2 - \mathbf{p}\cdot\mathbf{p})^{1/2}], \tag{22}$$

with $\mathbf{p} = (p_1, p_2) \in \mathbb{R}^2$, and by Bessel fields

$$\psi_B(\mathbf{x}, z) = J_0[k_0\sin\vartheta_0(x^2 + y^2)^{1/2}]\exp(izk_0\cos\vartheta_0), \tag{23}$$

where $J_0(z)$ denotes the zeroth-order Bessel function of first kind. The 'frequency' ζ can be either real and positive, or *purely imaginary*, namely $\zeta^* = -\zeta$ and $\zeta^2 < 0$. This can be seen multiplying equation (21) by $\varphi^*(\mathbf{x})$ and integrating over the xy-plane, thus obtaining [4]

$$\zeta^2 = \frac{k_0^2\int d\mathbf{x}\,|\varphi(\mathbf{x})|^2 - \int d\mathbf{x}\,|\boldsymbol{\nabla}\varphi(\mathbf{x})|^2}{\int d\mathbf{x}\,|\varphi(\mathbf{x})|^2}, \tag{24}$$

where $d\mathbf{x} = dx\,dy$ is the surface element and we assumed that the field $\varphi(\mathbf{x})$ vanishes for $x, y \to \infty$ in order to neglect surface terms. From equation (24), it follows that either $\zeta^2 \geqslant 0$, or $\zeta^2 < 0$ whenever

$$\int d\mathbf{x}\ |\boldsymbol{\nabla}\varphi(\mathbf{x})|^2 > k_0^2 \int d\mathbf{x}\ |\varphi(\mathbf{x})|^2. \tag{25}$$

This relation imposes a constraint upon the Fourier spectrum $\widetilde{\varphi}(\mathbf{p})$ of the field $\varphi(\mathbf{x})$, where $\mathbf{p} = (p_1, p_2) \in \mathbb{R}^2$. Substituting the Fourier representation

$$\varphi(\mathbf{x}) = \frac{1}{2\pi}\int d\mathbf{p}\ \widetilde{\varphi}(\mathbf{p})e^{-i\mathbf{p}\cdot\mathbf{x}} \tag{26}$$

into equation (25) yields to

$$\int d\mathbf{p}(p^2 - k_0^2)\ |\widetilde{\varphi}(\mathbf{p})|^2 > 0, \tag{27}$$

where $p^2 = p_i p_i = \mathbf{p}\cdot\mathbf{p}$. Therefore, whenever the support of $\widetilde{\varphi}(\mathbf{p})$ is not entirely contained within the circle of equation $p^2 = p_1^2 + p_2^2 = k_0^2$, the field develops purely imaginary frequencies. This fact will have profound consequences upon the quantization of $\psi(\mathbf{x}, z)$. Since the right side of equation (24) is always *real*, it follows that

$$\zeta^2 = (\zeta^2)^* = (\zeta^*)^2. \tag{28}$$

Using this result in equation (21), we see that $\varphi(\mathbf{x})$ and $\varphi^*(\mathbf{x})$ satisfy the *same* equation. Therefore, if $\varphi(\mathbf{x})$ is a given solution of equation (21), then $\varphi^*(\mathbf{x})$ is also a solution.

We conclude this part by noticing that, irrespective of the either positive or purely imaginary value taken by ζ, there are four linearly independent separable solutions of equation (12), namely

$$\begin{aligned}\psi_+ &= \varphi(\mathbf{x})e^{i\zeta z}, \qquad \psi_+^* = \varphi^*(\mathbf{x})e^{-i\zeta^* z},\\ \psi_-^* &= \varphi^*(\mathbf{x})e^{i\zeta^* z}, \qquad \psi_- = \varphi(\mathbf{x})e^{-i\zeta z}.\end{aligned} \tag{29}$$

When $\zeta = \zeta^*$ all solutions are oscillatory and therefore physically acceptable. However, when $\zeta = i|\zeta|$ it has

$$\begin{aligned}\psi_+ &= \varphi(\mathbf{x})e^{-|\zeta|z}, \qquad \psi_+^* = \varphi^*(\mathbf{x})e^{-|\zeta|z},\\ \psi_-^* &= \varphi^*(\mathbf{x})e^{|\zeta|z}, \qquad \psi_- = \varphi(\mathbf{x})e^{|\zeta|z}\end{aligned} \tag{30}$$

and the ψ_-'s solutions are *exponentially* growing as z increases. Therefore they represent physically acceptable solutions only for $z < 0$. Vice versa, the ψ_+'s solutions are exponentially decaying and physically acceptable only for $z > 0$. For fields associated to optical beams, ψ_+ (ψ_-) and ψ_+^* (ψ_-^*) are called *evanescent waves* when $\zeta = i|\zeta|$ and $z > 0$ ($z < 0$).

2.2. Symmetries and conservation laws

In this part, we discuss the symmetries of the Lagrangian (16). A different discussion can be found in [5]. To begin with, let us note that such Lagrangian is manifestly invariant under the transformation

$$\psi \to e^{-i\Lambda}\psi, \qquad \psi^* \to e^{i\Lambda}\psi^*, \tag{31}$$

where Λ is a real constant. From Noether's theorem, it follows that there exists a current

$$\begin{aligned}\mathscr{I}_\mu &= \frac{1}{i}\left[\psi\frac{\partial\mathcal{L}}{\partial(\partial^\mu\psi)} - \psi^*\frac{\partial\mathcal{L}}{\partial(\partial^\mu\psi^*)}\right]\\ &= \frac{1}{i}(\psi\partial_\mu\psi^* - \psi^*\partial_\mu\psi),\end{aligned} \tag{32}$$

which has a vanishing three-divergence [6–8]

$$\partial^\mu\mathscr{I}_\mu = \partial_z\mathscr{I}_z + \boldsymbol{\nabla}\cdot\mathscr{I} = 0, \tag{33}$$

namely

$$\frac{\partial}{\partial z}\left(\psi\frac{\partial\psi^*}{\partial z} - \psi^*\frac{\partial\psi}{\partial z}\right) = -\boldsymbol{\nabla}\cdot(\psi\boldsymbol{\nabla}\psi^* - \psi^*\boldsymbol{\nabla}\psi). \tag{34}$$

Integrating both sides of this equation over all the xy-plane, we obtain

$$\begin{aligned}\partial_z\int d\mathbf{x}\mathscr{I}_z &= -\int d\mathbf{x}\ \boldsymbol{\nabla}\cdot\mathscr{I}\\ &= 0,\end{aligned} \tag{35}$$

where the right side amounts to the two-dimensional integral of a two-divergence and then vanishes for fields localized within a *finite* region of the xy-plane. This equation states that during propagation of a monochromatic optical field along the z-axis, the 'charge' $\mathcal{C}$ defined as

$$\mathcal{C} = \frac{1}{i}\int d\mathbf{x}\left(\psi\frac{\partial\psi^*}{\partial z} - \psi^*\frac{\partial\psi}{\partial z}\right), \tag{36}$$

is conserved, that is

$$\frac{\partial\mathcal{C}}{\partial z} = 0. \tag{37}$$

It is instructive to evaluate $\mathcal{C}$ for the four fundamental solutions (29). A straightforward calculation gives

$$\psi_+ \to \mathcal{C}_+ = -2\,C\,e^{-2z\,\mathrm{Im}\,\zeta}\,\mathrm{Re}\,\zeta, \tag{38}$$

$$\psi_+^* \to -\mathcal{C}_+, \tag{39}$$

$$\psi_-^* \to -\mathcal{C}_-, \tag{40}$$

$$\psi_- \to \mathcal{C}_- = 2\,C\,e^{2z\,\mathrm{Im}\,\zeta}\,\mathrm{Re}\,\zeta, \tag{41}$$

where

$$C = \int d\mathbf{x}\ |\varphi(\mathbf{x})|^2. \tag{42}$$

At first sight, the charges $\mathcal{C}_\pm$ seems to depend on z, thus contradicting the conservation law (36). However, one should remember that there are only two possibilities for ζ: either $\zeta = \zeta^* \Rightarrow \mathrm{Im}\,\zeta = 0$, or $\zeta = -\zeta^* \Rightarrow \mathrm{Re}\,\zeta = 0$. In the first case, we have $\mathcal{C}_\pm = \mp 2C\zeta$ and the conservation law (36) is satisfied. In the latter case $\mathcal{C}_+ = 0 = \mathcal{C}_-$ and there is no conserved charge. Thus, there is no charge associated to evanescent waves.

Additional conserved quantities can be calculated in a straightforward manner from the *canonical energy-momentum tensor*, which is calculated from the Lagrangian as

$$\mathscr{T}_{\mu\nu} = \frac{\partial\mathcal{L}}{\partial(\partial^\mu\psi)}\partial_\nu\psi + \frac{\partial\mathcal{L}}{\partial(\partial^\mu\psi^*)}\partial_\nu\psi^* - \delta_{\mu\nu}\mathcal{L}$$

$$= \partial_\mu\psi^*\partial_\nu\psi + \partial_\mu\psi\partial_\nu\psi^* - \delta_{\mu\nu}(\delta^{\alpha\beta}\partial_\alpha\psi^*\partial_\beta\psi - k_0^2\psi^*\psi). \tag{43}$$

Note that this tensor is automatically symmetric, namely $\mathscr{T}_{\mu\nu} = \mathscr{T}_{\nu\mu}$ and, by definition

$$\frac{\partial}{\partial x_\mu}\mathscr{T}_{\mu\nu} = 0. \tag{44}$$

Since $\nu = 1, 2, 3$ this means that there are three conserved quantities, the 'energy' H and the linear momentum vector $\mathbf{P}$ of the field, defined respectively as

$$H = \int d\mathbf{x}\,\mathscr{T}^{33} = \int d\mathbf{x}(|\partial_z\psi|^2 - |\boldsymbol{\nabla}\psi|^2 + k_0^2\,|\psi|^2), \tag{45}$$

and

$$\mathbf{P} = \int d\mathbf{x}\,\mathscr{T}^{3l}\epsilon_l = \int d\mathbf{x}\left(\frac{\partial\psi^*}{\partial z}\boldsymbol{\nabla}\psi + \frac{\partial\psi}{\partial z}\boldsymbol{\nabla}\psi^*\right), \tag{46}$$

where $P^\mu = (H, \mathbf{P})$ denotes the full three-momentum of the beam. Therefore, as the beam propagates along the z-axis, the 'energy' H and the two components of the linear momentum $\mathbf{P}$ remain constant as a consequence of equation (44), namely

$$\frac{\partial H}{\partial z} = 0 = \frac{\partial\mathbf{P}}{\partial z}. \tag{47}$$

As will be shown later, the 'energy' H coincides with the Hamiltonian of the system. However, the expression in equation (45) is not manifestly *positive semidefinite*, as a physically realizable Hamiltonian should be, because of the negative 'kinetic energy' term $-|\boldsymbol{\nabla}\psi|^2$. We will discuss this point at length later, when proceeding with the quantization of the field. For the moment, we verify that H is actually positive semidefinite for the four fundamental solutions (29). After a straightforward calculation, one finds

$$\psi_\pm \to H = e^{\mp 2z\,\mathrm{Im}\,\zeta}\int d\mathbf{x}(|\zeta|^2|\varphi(\mathbf{x})|^2$$
$$- |\boldsymbol{\nabla}\varphi(\mathbf{x})|^2 + k_0^2\,|\varphi(\mathbf{x})|^2)$$
$$= (\zeta^2 + |\zeta|^2)e^{\mp 2z\,\mathrm{Im}\,\zeta}\int d\mathbf{x}\,\,|\varphi(\mathbf{x})|^2, \tag{48}$$

where we used the equation of motion equation (21) and integration by part (discarding surface terms) to pass from the first to the second expression. The conjugate fields $\psi_\pm^*$ yield the same H. Again, equation (48) seems to depend upon the propagation distance z, but this is not the case because $\zeta^2 + |\zeta|^2 = \zeta(\zeta + \zeta^*) = 2\,\zeta\,\mathrm{Re}\,\zeta$ which implies that

$$\begin{cases} \zeta = \zeta^* \quad\Rightarrow\quad \zeta^2 = |\zeta|^2 \quad\text{when}\quad \mathrm{Im}\,\zeta = 0, \\ \zeta = -\zeta^* \Rightarrow \zeta^2 = -|\zeta|^2 \quad\text{when}\quad \mathrm{Re}\,\zeta = 0. \end{cases} \tag{49}$$

Therefore, using equation (49) into (48), we obtain

$$H = \begin{cases} 2\,C\,\zeta^2 \geqslant 0, & \zeta = \zeta^*, \\ 0, & \zeta = -\zeta^*, \end{cases} \tag{50}$$

where C is given again by equation (42). This nice result shows that the evanescent waves *do not* contribute to the total energy of the field.

In a similar manner, we can now calculate $\mathbf{P}$ and the outcome is

$$\psi_\pm \to \mathbf{P} = \pm 2\,\mathrm{Re}\,\zeta\,e^{\mp 2z\,\mathrm{Im}\,\zeta}\int d\mathbf{x}\,\varphi^*(\mathbf{x})\left(\frac{1}{i}\boldsymbol{\nabla}\right)\varphi(\mathbf{x})$$

$$= \pm 2\,\mathrm{Re}\,\zeta\,e^{\mp 2z\,\mathrm{Im}\,\zeta}\int d\mathbf{p}\,\mathbf{p}\,|\widetilde{\varphi}(\mathbf{p})|^2, \tag{51}$$

where the Fourier representation equation (26) has been used. The conjugate fields $\psi_\pm^*$ yield the same $\mathbf{P}$. Also in this case the z-dependence of $\mathbf{P}$ is only deceptive because we can always rewrite equation (51) as

$$\mathbf{P} = \begin{cases} \pm 2\,\zeta\int d\mathbf{p}\,\mathbf{p}\,|\widetilde{\varphi}(\mathbf{p})|^2, & \zeta = \zeta^*, \\ 0, & \zeta = -\zeta^*. \end{cases} \tag{52}$$

Once again, the *physical* transverse linear momentum $\mathbf{P}$ does not receive contributions from the evanescent waves.

Finally, it is of some interest to note that equation (44) also follows an interesting *continuity equation* connecting the linear momentum density $\mathscr{P}^i = \mathscr{T}^{3i}$ with the linear momentum *transverse flux density* $\mathscr{T}^{ij}$ [9]:

$$\frac{\partial}{\partial z}\mathscr{P}^i + \frac{\partial}{\partial x^j}\mathscr{T}^{ij} = 0, \tag{53}$$

where

$$\mathscr{T}^{ij} = \partial^i\psi^*\,\partial^j\psi + \partial^j\psi\,\partial^i\psi^* - \delta^{ij}(|\partial_\alpha\psi|^2 - k_0^2\,|\psi|^2). \tag{54}$$

This equation (53) simply tells us that variations of the transverse linear momentum density during possibly occurring during propagation, are compensated by the transverse variations of $\mathscr{T}^{ml}$.

Since our canonical energy-momentum tensor is symmetric, we can construct an additional conserved tensor density:

$$\mathscr{M}^{\lambda\mu\nu} \equiv x^\mu\mathscr{T}^{\lambda\nu} - x^\nu\mathscr{T}^{\lambda\mu}, \tag{55}$$

such that

$$\partial_\lambda\mathscr{M}^{\lambda\mu\nu} = \mathscr{T}^{\mu\nu} - \mathscr{T}^{\nu\mu} = 0. \tag{56}$$

To proceed further, first we rewrite explicitly equation (56) as

$$\partial_\lambda\mathscr{M}^{\lambda\mu\nu} = \frac{\partial}{\partial x}\mathscr{M}^{1\mu\nu} + \frac{\partial}{\partial y}\mathscr{M}^{2\mu\nu} + \frac{\partial}{\partial z}\mathscr{M}^{3\mu\nu} = 0. \tag{57}$$

Then, we integrate term by term over the xy-plane obtaining

$$\frac{\partial}{\partial z}\int d\mathbf{x}\,\mathscr{M}^{3\mu\nu} = -\int d\mathbf{x}\left(\frac{\partial}{\partial x}\mathscr{M}^{1\mu\nu}\right) - \int d\mathbf{x}\left(\frac{\partial}{\partial y}\mathscr{M}^{2\mu\nu}\right). \tag{58}$$

The two terms at the right side of this equation can be discarded under the assumption that the fields and their derivative fall off sufficiently fast at infinity. Thus, we recover the

well-known conservation law [6]:

$$\frac{\partial}{\partial z} \int d\mathbf{x} \, \mathscr{M}^{3\mu\nu} = 0. \tag{59}$$

However, it should be reminded that particular care must be taken when handling equation (58) because of the risk of improper manipulation of the surface terms [10]. To proceed further, it is useful to define the conserved *angular momentum tensor* as

$$J^{\mu\nu} \equiv \int d\mathbf{x} \, \mathscr{M}^{3\mu\nu} \equiv \int d\mathbf{x} \, \mathscr{J}^{\mu\nu}, \tag{60}$$

where

$$\mathscr{J}^{\mu\nu} = x^\mu \mathscr{P}^\nu - x^\nu \mathscr{P}^\mu, \tag{61}$$

with $\mathscr{P}^\mu \equiv \mathscr{T}^{3\mu}$. From the definitions above, it follows that $\mathscr{J}^{\mu\mu}$ is antisymmetric, namely $\mathscr{J}^{\mu\mu} = 0$ (no sum over repeated indices) and $\mathscr{J}^{\mu\nu} = -\mathscr{J}^{\nu\mu}$. Moreover, using the only three independent Cartesian components of $\mathscr{J}^{\mu\nu}$, we define the '*angular momentum density vector*' as

$$\mathscr{J}_\lambda = \frac{1}{2} \epsilon_{\lambda\mu\nu} \mathscr{J}^{\mu\nu} = (\mathscr{J}^{23}, \mathscr{J}^{31}, \mathscr{J}^{12}), \tag{62}$$

where $\epsilon_{\lambda\mu\nu}$ denotes the totally antisymmetric three-dimensional Levi-Civita symbol. Using equations (55), (61) into (62) we obtain

$$\mathscr{J}_x = y\mathscr{P}_z - z\mathscr{P}_y, \tag{63}$$

$$\mathscr{J}_y = z\mathscr{P}_x - x\mathscr{P}_z, \tag{64}$$

$$\mathscr{J}_z = x\mathscr{P}_y - y\mathscr{P}_x. \tag{65}$$

The corresponding *total* angular momentum components are straightforwardly calculated by integrating the relations above to obtain $J_\lambda = \epsilon_{\lambda\mu\nu} J^{\mu\nu}/2$, where equation (60) has been used. Explicitly, integration both sides of equations (63) and (64) over the *xy*-plane gives

$$J_x = \bar{y}(z) - zP_y, \tag{66}$$

$$J_y = zP_x - \bar{x}(z), \tag{67}$$

where we have defined

$$\bar{y}(z) \equiv \int d\mathbf{x}(y\mathscr{P}_z), \qquad \text{and} \qquad \bar{x}(z) \equiv \int d\mathbf{x}(x\mathscr{P}_z), \quad (68)$$

and equation (46) has been used. For fields such that $\mathscr{P}_z \geqslant 0$, the vector $\bar{\mathbf{x}}(z) = \epsilon_1 \bar{x}(z) + \epsilon_2 \bar{y}(z)$ can be interpreted as the *centroid* of the energy distribution on the *xy*-plane. Then, deriving both sides of equations (66), (67) with respect to z and using the conservation laws (47) and (59) we obtain the equations of motions of the centroid of the field

$$\frac{d\,\bar{y}(z)}{dz} = P_y, \qquad \text{and} \qquad \frac{d\,\bar{x}(z)}{dz} = P_x, \tag{69}$$

which reproduce the laws of *rays* propagation in geometrical optics [11].

It is enlightening to calculate explicitly J_z for the fields (29). After a lengthy but straightforward calculation, one finds

$$\psi_\pm \to J_z$$

$$= \pm 2 \, \mathrm{Re} \, \zeta \, e^{\mp 2z \, \mathrm{Im} \, \zeta} \int d\mathbf{x} \, \varphi^*(\mathbf{x}) \left(x\frac{1}{i}\frac{\partial}{\partial y} - y\frac{1}{i}\frac{\partial}{\partial x} \right) \varphi(\mathbf{x}), \tag{70}$$

where $2 \, \mathrm{Re} \, \zeta \exp(-2z \, \mathrm{Im} \, \zeta) = 2\zeta$ for $\zeta = \zeta^*$ and it is equal to zero for $\zeta = -\zeta^*$. Once again, the 'unphysical' evanescent waves generated by the angular spectrum representation do not carry angular momentum. From equation (70), it follows that J_z is conserved along with propagation because $\partial J_z/\partial z \propto \mathrm{Re} \, \zeta \, \mathrm{Im} \, \zeta = 0$, the latter equality being a consequence of the fact that ζ is either real or purely imaginary. The conjugate fields $\psi_\pm^*$ produce the same J_z.

One more continuity equation may be derived by rewriting equation (56) with the help of equation (60), as

$$\frac{\partial}{\partial z} \mathscr{J}^{\mu\nu} + \frac{\partial}{\partial x^i} \mathscr{M}^{i\mu\nu} = 0. \tag{71}$$

Multiplying both sides of this equation by $\epsilon_{\lambda\mu\nu}/2$ and then summing over repeated indices, we obtain

$$\frac{\partial}{\partial z} \mathscr{J}_\lambda + \frac{\partial}{\partial x^i} \mathscr{L}_\lambda^i = 0, \tag{72}$$

where we have defined the *angular momentum flux density* as

$$\mathscr{L}_\lambda^i \equiv \frac{1}{2} \epsilon_{\lambda\mu\nu} \mathscr{M}^{i\mu\nu}. \tag{73}$$

This quantity differs from the homonym one introduced by Barnett [9] in that $\mathscr{L}_\lambda^i$ is time-independent.

In conclusion, in this section we have shown that all the relevant physical quantities as energy, linear and angular momenta vanishes for evanescent fields. Thus, the latter appears more as *virtual* fields that do not correspond to any real physical field. However, as we shall see later, they are necessary to preserve unitarity [12].

3. Hamiltonian formalism

Propaedeutical to the quantization procedure, which will be carried out in part II, is the introduction of the Hamiltonian formalism for equation (1). The procedure for passing from the Lagrangian to the Hamiltonian representation of the field is standard [6]. First, we write down our Lagrangian (16) as

$$\mathcal{L} = \partial_\mu \psi^* \partial_\mu \psi - k_0^2 \, \psi^* \psi, \tag{74}$$

where, for the sake of clarity, we omitted the subscript 'HWE'. Then, we determine the fields $\Pi(\mathbf{x}, z)$ and $\Pi^*(\mathbf{x}, z)$ canonically conjugate to $\psi(\mathbf{x}, z)$ and $\psi^*(\mathbf{x}, z)$, respectively

$$\Pi = \frac{\partial \mathcal{L}}{\partial(\partial_3 \psi)} = \partial_3 \psi^*, \qquad \Pi^* = \frac{\partial \mathcal{L}}{\partial(\partial_3 \psi^*)} = \partial_3 \psi. \quad (75)$$

The *Hamiltonian density* $\mathscr{H}$ is defined in terms of the four fields ψ, Π, ψ^*, Π^* as usual:

$$\mathscr{H} = \Pi \, \partial_3 \psi + \Pi^* \partial_3 \psi^* - \mathcal{L}$$
$$= \Pi^* \Pi - \boldsymbol{\nabla} \psi^* \cdot \boldsymbol{\nabla} \psi + k_0^2 \psi^* \psi. \tag{76}$$

Inverting the relations (75) to express the field derivatives in terms of the conjugate momenta, and using this result in equation (76), we obtain

$$\mathscr{H} = \partial_3 \psi^* \partial_3 \psi - \delta^{ij} \partial_i \psi^* \partial_j \psi + k_0^2 \psi^* \psi, \qquad (77)$$

which shows that actually $\mathscr{H} = \mathscr{T}^{33}$, as previously stated. The reader familiar with quantum field theory of tachyons, will appreciate the similarity between equation (77) and the Hamiltonian of a Klein–Gordon field with purely imaginary mass [13, 14].

Now we are going to show that the canonical Hamiltonian

$$H = \int d\mathbf{x}\, \mathscr{H} \qquad (78)$$

is naturally partitioned in a 'propagating' and an 'evanescent' part. Interestingly, the same phenomenon manifests in the quantization of scalar fields near rapidly rotating stars [4] and in cosmological models of universes with unstable modes [15]. In the latter case, the propagating and evanescent parts are quite suggestively dubbed 'light' and 'dark' components of the Hamiltonian, respectively [16].

At any position z the fields $\Pi(\mathbf{x}, z)$ and $\psi(\mathbf{x}, z)$ can be expanded in terms of the Fourier transform representations:

$$\psi(\mathbf{x}, z) = \frac{1}{2\pi} \int d\mathbf{p}\, Q(\mathbf{p}, z) e^{i\mathbf{p}\cdot\mathbf{x}} \qquad (79)$$

and

$$\Pi(\mathbf{x}, z) = \frac{1}{2\pi} \int d\mathbf{p}\, P(\mathbf{p}, z) e^{-i\mathbf{p}\cdot\mathbf{x}}, \qquad (80)$$

where the complex amplitudes $Q(\mathbf{p}, z)$ and $P(\mathbf{p}, z)$ are z-dependent. The minus sign in the exponential in equation (80) is *not* a typo. Substituting equations (79), (80) into (78) we obtain, after some manipulation

$$H = \int d\mathbf{p}[P^*(\mathbf{p}, z)P(\mathbf{p}, z) + (k_0^2 - p^2)Q^*(\mathbf{p}, z)Q(\mathbf{p}, z)], \qquad (81)$$

where $p^2 = \mathbf{p} \cdot \mathbf{p}$. It is clear that either $\zeta_p^2 \equiv k_0^2 - p^2 \geqslant 0$ for $p^2 \leqslant k_0^2$, or $\zeta_p^2 < 0$ for $p^2 > k_0^2$. In any case, the z-derivatives of P and Q are given by the Hamilton equations

$$\frac{\partial}{\partial z}P(\mathbf{p}, z) = -\frac{\delta H}{\delta Q(\mathbf{p}, z)} = -\zeta_p^2\, Q^*(\mathbf{p}, z) \qquad (82)$$

and

$$\frac{\partial}{\partial z}Q^*(\mathbf{p}, z) = \frac{\delta H}{\delta P^*(\mathbf{p}, z)} = P(\mathbf{p}, z), \qquad (83)$$

and their conjugates. Here the symbol $\delta F[f]/\delta f[t]$ denotes the functional derivative of the functional $F[f]$ [17]. Deriving equation (82) with respect to z and using equation (83) yields the equation of motion of $Q(\mathbf{p}, z)$ and $Q^*(\mathbf{p}, z)$:

$$\left(\frac{\partial^2}{\partial z^2} + \zeta_p^2\right)Q(\mathbf{p}, z) = 0. \qquad (84)$$

Before solving this equation, we turn back to equation (81) to rewrite it as $H = H_L + H_D$, where the subscripts 'L' and 'D' stand for light and dark, respectively, with

$$H_L = \int d\mathbf{p}[P^*(\mathbf{p}, z)P(\mathbf{p}, z) + \zeta_p^2\, Q^*(\mathbf{p}, z)Q(\mathbf{p}, z)] \\ \times \Theta(k_0^2 - p^2), \qquad (85)$$

and

$$H_D = \int d\mathbf{p}[P^*(\mathbf{p}, z)P(\mathbf{p}, z) \\ - |\zeta_p|^2 Q^*(\mathbf{p}, z)Q(\mathbf{p}, z)]\Theta(p^2 - k_0^2), \qquad (86)$$

where $\Theta(z)$ denotes the Heaviside step function. It will be clear later that the 'dark' part of the Hamiltonian does not contribute to physical quantities. Equation (85) clearly represents the Hamiltonian of a continuum set of harmonic oscillators, because $\Theta(k_0^2 - p^2)\zeta_p^2 \geqslant 0$. However, equation (86) expresses the Hamiltonian of a continuum set of *repulsive* (or, inverted) harmonic oscillators which are known, in quantum mechanics, to do not possess neither square-integrable eigenstates, nor a lower energy vacuum state (H_D is not bounded from below, as we shall see soon) [15]. Therefore, equation (84) naturally splits in two independent equations of the form

$$\left(\frac{\partial^2}{\partial z^2} \pm |\zeta_p|^2\right)Q_\pm(\mathbf{p}, z) = 0, \qquad (87)$$

whose solutions are

$$Q_+(\mathbf{p}, z) = c_{+,1}(\mathbf{p})e^{i|\zeta_p|z} + c_{+,2}(\mathbf{p})e^{-i|\zeta_p|z}, \qquad (88)$$

and

$$Q_-(\mathbf{p}, z) = c_{-,1}(\mathbf{p})e^{-|\zeta_p|z} + c_{-,2}(\mathbf{p})e^{|\zeta_p|z}, \qquad (89)$$

where $c_{\pm,1}(\mathbf{p})$ and $c_{\pm,2}(\mathbf{p})$ are arbitrary functions of $\mathbf{p}$ solely. The corresponding $P_\pm(\mathbf{p}, z)$ are straightforwardly calculated from equation (83):

$$P_+(\mathbf{p}, z) = \frac{\partial}{\partial z}Q_+^*(\mathbf{p}, z) = -i\zeta_p(c_{+,1}^*(\mathbf{p})\, e^{-i\zeta_p z} - c_{+,2}^*(\mathbf{p})e^{i\zeta_p z}), \qquad (90)$$

$$P_-(\mathbf{p}, z) = \frac{\partial}{\partial z}Q_-^*(\mathbf{p}, z) \\ = -|\zeta_p|(c_{-,1}^*(\mathbf{p})\, e^{-|\zeta_p|z} - c_{-,2}^*(\mathbf{p})e^{|\zeta_p|z}), \qquad (91)$$

where from now on we drop the redundant symbol $|\zeta_p|$ whenever $\zeta_p = |\zeta_p|$. Substituting equations (16)–(19) into equations (13) and (14) one obtains

$$H_L = 2\int d\mathbf{p}\, \zeta_p^2(|c_{+,1}(\mathbf{p})|^2 + |c_{+,2}(\mathbf{p})|^2)\Theta(k_0^2 - p^2), \qquad (92)$$

and

$$H_D = -2\int d\mathbf{p}\, |\zeta_p|^2(c_{-,1}^*(\mathbf{p})c_{-,2}(\mathbf{p}) \\ + c_{-,2}^*(\mathbf{p})c_{-,1}(\mathbf{p}))\Theta(p^2 - k_0^2). \qquad (93)$$

Since $c_{-,1}^*c_{-,2} + c_{-,1}c_{-,2}^* = 2\,\text{Re}\, c_{-,1}^*c_{-,2}$ has not a definite sign and can take any value, it is clear that H_D is not bounded from below.

3.1. Recovering the angular spectrum representation

To make a connection with the *angular spectrum* theory in classical optics [18], let us begin by remarking that in free space, the diverging and converging exponential functions in equation (89) cannot *both* represent physical solutions of equation (87) for z either positive or negative, since unbounded exponentially growing functions cannot belong to the spectrum of a realistic physical theory. Therefore, we must require that

$$\begin{cases} c_{-,2}(\mathbf{p}) = 0, & z > 0, \\ c_{-,1}(\mathbf{p}) = 0, & z \leqslant 0. \end{cases} \tag{94}$$

However, this means that $c_{-,1}^*(\mathbf{p})c_{-,2}(\mathbf{p}) = 0$ *everywhere* and, consequently, $H_D = 0$. Therefore, the 'dark' component of the Hamiltonian does not contribute to the *physical* energy of the field.

As shown through a wealth of details by Mandel and Wolf in [18], the angular spectrum of a wavefield is *uniquely* determined in the half-space $z \geqslant 0$ ($z \leqslant 0$) for well-behaving forward (backward) propagating fields. Let us fix, from now on, $z \geqslant 0$ for the sake of definiteness. In order to recover standard the angular spectrum from our equations (88) and (89), we have to choose $c_{+,1}(\mathbf{p}) = c_{-,1}(\mathbf{p}) \equiv a(\mathbf{p})$ and $c_{\pm,2}(\mathbf{p}) = 0$, where the amplitude $a(\mathbf{p})$ is the quantity customarily dubbed 'angular spectrum'. In this case, we can rewrite the original field $\psi(\mathbf{x}, z)$ in the following form:

$$\begin{aligned} \psi(\mathbf{x}, z \geqslant 0) &= \frac{1}{2\pi} \int d\mathbf{p} \, Q(\mathbf{p}, z) e^{i\mathbf{p} \cdot \mathbf{x}} \\ &= \frac{1}{2\pi} \int d\mathbf{p} \, e^{i\mathbf{p} \cdot \mathbf{x}} [Q_+(\mathbf{p}, z)\Theta(k_0^2 - p^2) \\ &\quad + Q_-(\mathbf{p}, z)\Theta(p^2 - k_0^2)] \\ &= \frac{1}{2\pi} \int d\mathbf{p} \, a(\mathbf{p}) \, e^{i\zeta_p z} \, e^{i\mathbf{p} \cdot \mathbf{x}}, \end{aligned} \tag{95}$$

where the last equality follows from the fact that

$$\begin{cases} e^{i\zeta_p z} = e^{i|\zeta_p|z}, & p^2 \leqslant k_0^2, \\ e^{i\zeta_p z} = e^{-|\zeta_p|z}, & p^2 > k_0^2, \end{cases} \tag{96}$$

where $\zeta_p = +(k_0^2 - p^2)^{1/2}$. Substituting equation (95) into (77) and integrating over the xy-plane, one obtains

$$\begin{aligned} H &= \int d\mathbf{x} \, \mathcal{H} \\ &= 2 \int d\mathbf{p} \, \zeta_p \, \text{Re} \, \zeta_p \, e^{-2z \, \text{Im} \, \zeta_p} \, |a(\mathbf{p})|^2 \\ &= 2 \int d\mathbf{p} \, |a(\mathbf{p})|^2 \zeta_p^2 \, \Theta(k_0^2 - p^2), \end{aligned} \tag{97}$$

where $\zeta_p^2 \, \Theta(k_0^2 - p^2) \geqslant 0$ by definition. This Hamiltonian is manifestly positive semidefinite and, therefore, physically acceptable. Moreover, equation (97) trivially implies that

$$\frac{\partial H}{\partial z} = 0, \tag{98}$$

which shows that energy is conserved during free propagation of wavefields represented in the angular spectrum form.

3.1.1. Effective-Lagrangian theory for the angular spectrum.

Comparing the first and the last lines of equation (95), one sees that

$$Q(\mathbf{p}, z) = a(\mathbf{p})e^{i\zeta_p z} \qquad \Rightarrow \qquad a(\mathbf{p}) = Q(\mathbf{p}, 0). \tag{99}$$

However, by definition $Q(\mathbf{p}, 0)$ is the inverse Fourier transform of $\psi(\mathbf{x}, 0)$. Therefore, the angular spectrum $a(\mathbf{p})$ is *uniquely* determined by the knowledge of $\psi(\mathbf{x}, 0)$ solely. This seems to be in contradiction with equation (84) which is a *second-order* differential equation whose uniqueness of the solution require the knowledge of both

$$Q(\mathbf{p}, z)|_{z=0} \qquad \text{and} \qquad \frac{\partial Q(\mathbf{p}, z)}{\partial z}\bigg|_{z=0} = P^*(\mathbf{p}, 0). \tag{100}$$

The clarification of the foregoing apparent paradox lies in the observation that if we derive with respect to z both sides of the leftmost part of equation (99), we obtain

$$\frac{\partial Q(\mathbf{p}, z)}{\partial z} = i\zeta_p Q(\mathbf{p}, z), \tag{101}$$

which is a *first-order* differential equation. For reasons that will be soon clear, let us define

$$Q(\mathbf{p}, z) = A_+^*(\mathbf{p}, z)\Theta(k_0^2 - p^2) + A_-(\mathbf{p}, z)\Theta(p^2 - k_0^2), \tag{102}$$

where the presence of A_- instead of A_-^* in the equation above is not accidental. Then, equation (101) yields to two independent first-order equations of motion

$$\begin{aligned} &\frac{\partial}{\partial z}A_+(\mathbf{p}, z) = -i\zeta_p A_+(\mathbf{p}, z), \quad \text{and} \\ &\frac{\partial}{\partial z}A_-(\mathbf{p}, z) = -|\zeta_p| \, A_-(\mathbf{p}, z), \end{aligned} \tag{103}$$

which can be derived from the *effective first-order Lagrangians* $\mathcal{L}_+$ and $\mathcal{L}_-$ defined as

$$\begin{aligned} &\mathcal{L}_+ = -\frac{i}{2}(A_+\partial_z A_+^* - A_+^*\partial_z A_+) - \zeta_p A_+^* A_+, \quad \text{for} \\ &\zeta_p = \zeta_p^*, \end{aligned} \tag{104}$$

and

$$\begin{aligned} &\mathcal{L}_- = \frac{1}{2}(A_-\partial_z \bar{A}_- - \bar{A}_-\partial_z A_-) - |\zeta_p| \, \bar{A}_- A_-, \quad \text{for} \\ &\zeta_p = -\zeta_p^*. \end{aligned} \tag{105}$$

From the latter equation follows that $\bar{A}_-(\mathbf{p}, z)$ obeys the equation of motion

$$\frac{\partial}{\partial z}\bar{A}_-(\mathbf{p}, z) = |\zeta_p| \, \bar{A}_-(\mathbf{p}, z). \tag{106}$$

Therefore, $\bar{A}_-(\mathbf{p}, z) \neq A_-^*(\mathbf{p}, z)$ and it must be regarded as an *independent variable*. If we choose both variables $A_-(\mathbf{p}, z)$ and $\bar{A}_-(\mathbf{p}, z)$ *reals*, then also $\mathcal{L}_-$ is real and it becomes a physically admissible Lagrangian. However, while the Hamiltonian density

$$\mathcal{H}_+ = \zeta_p \, |A_+|^2 \tag{107}$$

is positive semidefinite, the same is not true for

$$\mathscr{H}_- = |\zeta_p|\, \bar{A}_- A_-, \tag{108}$$

because $\bar{A}_-$ and A_- can take, *independently*, any real value.

Last but not least, it should be noticed that the solutions of equation (106) are, evidently, exponential functions diverging for $z \to \infty$:

$$\bar{A}_-(\mathbf{p}, z) = \bar{A}_-(\mathbf{p}, 0)\, e^{|\zeta_p|z}. \tag{109}$$

This is really curious: although we had removed ad hoc such 'unstable' solutions from the expression (95) of the field, they entered back in the game to ensure the existence of a proper Lagrangian $\mathcal{L}_-$. Indeed, if in equation (105) one replaced $\bar{A}_-$ with A_-^*, the corresponding Lagrangian would become complex.

4. Helmholtz equation in dirac form

In the foregoing section, we have seen that the angular spectrum of a field obeys a first-order differential equation. However, this is not the only way to reduce the Helmholtz equation to a first-order form. To show this, let us rewrite the HWE in the compact form

$$\frac{\partial^2 \psi}{\partial z^2} = -(\nabla^2 + k_0^2)\psi, \tag{110}$$

where, as usual in this work, $\nabla^2 = \partial_x^2 + \partial_y^2$. A *formal* solution of this equation can be written in an operator form as

$$\psi(\mathbf{x}, z) = e^{iz\sqrt{\nabla^2 + k_0^2}}\, \psi(\mathbf{x}, 0). \tag{111}$$

Here, the 'Hamiltonian' differential operator

$$H = \sqrt{\nabla^2 + k_0^2}, \tag{112}$$

is reminiscent of the Hamiltonian of a relativistic free particle $H = \sqrt{p^2 c^2 + m^2 c^4}$. The problems with our H in equation (112) are the same ones encountered in quantum mechanics when extending the Schödinger equation to the relativistic regime. The square root operator on the right in equation (112) contains all powers of the ∇ operator. Even worst, in our case $\nabla^2 \sim -p^2$. Therefore, $\sqrt{-p^2 + k_0^2}$ even becomes purely imaginary for $p^2 > k_0^2$ hence, apparently, breaking down the unitarity of the theory. In quantum mechanics, this problem was brilliantly solved by Dirac who managed to reduce the second-order differential equation in equation (110) to a first-order form, without altering the physics of the problem.

Following in Dirac's footsteps, first we write a first-order equation of the form

$$\frac{\partial \psi}{\partial z} = \alpha_1 \frac{\partial \psi}{\partial x} + \alpha_2 \frac{\partial \psi}{\partial y} + \beta k_0^2 \psi, \tag{113}$$

and then we try to determine the unknown coefficients α_i, β by imposing that ψ must also satisfy the second-order HWE.

Therefore, iterating equation (113), we find

$$\begin{aligned}
\frac{\partial^2 \psi}{\partial z^2} &= \alpha_1^2 \frac{\partial^2 \psi}{\partial x^2} + \alpha_2^2 \frac{\partial^2 \psi}{\partial y^2} + (\alpha_1 \alpha_2 + \alpha_2 \alpha_1)\frac{\partial^2 \psi}{\partial x \partial y} \\
&\quad + k_0^2(\alpha_1 \beta + \beta \alpha_1)\frac{\partial \psi}{\partial x} \\
&\quad + k_0^2(\alpha_2 \beta + \beta \alpha_2)\frac{\partial \psi}{\partial y} + \beta^2 k_0^4 \psi \\
&\equiv -\frac{\partial^2 \psi}{\partial x^2} - \frac{\partial^2 \psi}{\partial y^2} - k_0^2 \psi.
\end{aligned} \tag{114}$$

The first two and the last lines of equation (114) coincide providing that

$$\begin{aligned}
\alpha_1^2 &= \alpha_2^2 = -1, \\
\alpha_1 \alpha_2 + \alpha_2 \alpha_1 &= 0, \\
\alpha_i \beta + \beta \alpha_i &= 0, \\
\beta^2 k_0^2 &= -1.
\end{aligned}$$

It is not difficult to verify that the equations above are satisfied by choosing

$$\alpha_1 = i\sigma_1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \ \alpha_2 = i\sigma_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$
$$\beta k_0 = i\sigma_3 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \tag{115}$$

where σ_μ are the Pauli matrices. Differently from the Dirac's theory, here all the αs and β matrices are *anti-Hermitean*. Then, we can rewrite equation (113) as

$$\frac{\partial}{\partial z}\psi(\mathbf{x}, z) = i\left(\sigma_1 \frac{\partial}{\partial x} + \sigma_2 \frac{\partial}{\partial y} + \sigma_3\, k_0\right)\psi(\mathbf{x}, z), \tag{116}$$

where the original scalar field $\psi(\mathbf{x}, z)$ must be now regarded as a doublet

$$\psi(\mathbf{x}, z) = \begin{pmatrix} \Psi_1(\mathbf{x}, z) \\ \Psi_2(\mathbf{x}, z) \end{pmatrix}. \tag{117}$$

The appearance of the two functions Ψ_i instead of the unique original one ψ, may be hardly surprising as clearly explained by Messiah [19]. Indeed, the solution of a second-order differential equation with respect to z, as the HWE is, requires the knowledge of both $\psi(\mathbf{x}, z)$ and $\partial_z \psi(\mathbf{x}, z)$ evaluated at the initial position $z = 0$. Therefore, converting the second-order HWE to a first-order Dirac form without loosing information, necessarily introduces a two-component wavefield.

In order to find a Lagrangian for equation (116) it is convenient first to rewrite it in the more suggestive form

$$(i\gamma^\mu \partial_\mu + k_0)\psi = 0, \tag{118}$$

where the three 2×2 matrices γ^μ are defined as:

$$\gamma^1 = -i\sigma_2, \qquad \gamma^2 = i\sigma_1, \qquad \gamma^3 = \sigma_3. \tag{119}$$

Then, if we define $\bar{\psi} \equiv \psi^\dagger \gamma^3$, with $(\psi^\dagger)_i = \Psi_i^*$, we can straightforwardly write down the Lagrangian as

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - k_0)\psi. \tag{120}$$

5. Conclusions

In this first part of our two-part work, we have mainly developed a Lagrangian theory for scalar, monochromatic optical fields. Our approach is unconventional in that we consider the coordinate z of a Cartesian reference frame attached to a beam of light and parallel to the propagation direction of the latter, in the same way the time t is considered in ordinary field theory. In this way, we are able to determine exact conservation laws for several physical quantities and we can highlight the different roles that propagating and evanescent waves have in the description of optical beams. Moreover, this theory also serves to establish the basic formalism before proceeding with the 'phenomenological' quantization of both Helmholtz and PWEs, which will be presented in part II.

ORCID iDs

Andrea Aiello ⦿ https://orcid.org/0000-0003-1647-0448

References

[1] Shvedov V and Krolikowski W 2018 Instantaneous field singularities in electromagnetic waves *New J. Phys.* **20** 103034

[2] Born M and Wolf E 2001 *Principles of Optics* 7th edn (Cambridge: Cambridge University Press)

[3] Dennis M R and Fenwick N A similar approach was proposed by these authors private communication

[4] Kang G 1996 Quantization of scalar field in the presence of imaginary *f* requency modes arXiv:hep-th/9603166

[5] Miller W 2012 Symmetry and separation of variables *Encyclopedia of Mathematics and Its Applications* vol 4 (Cambridge: Cambridge University Press)

[6] Ryder L H 1996 *Quantum Field Theory* 2nd edn (Cambridge: Cambridge University Press)

[7] Brown L S 1996 *Quantum Field Theory* (Cambridge: Cambridge University Press)

[8] Weinberg S 2010 *The Quantum Theory of Fields* vol I (Cambridge: Cambridge University Press)

[9] Barnett S M 2002 Optical angular-momentum flux *J. Opt. B: Quantum Semiclass. Opt.* **4** S7–16

[10] Ornigotti M and Aiello A 2014 Surface angular momentum of light beams *Opt. Express* **22** 6586–96

[11] Aiello A, Lindlein N, Marquardt Ch and Leuchs G 2009 Transverse angular momentum and geometric spin Hall effect of light *Phys. Rev. Lett.* **103** 100401

[12] Twareque Ali S 1972 Evanescent waves in quantum electrodynamics with unquantized sources *Phys. Rev.* D **7** 1668–75

[13] Schroer B 1971 Quantization of $m^2 < 0$ field equations *Phys. Rev.* D **3** 1764–70

[14] Kamoi K and Kamefuchi S 1971 Comments on quantum field theory of tachyons *Prog. Theor. Phys.* **45** 1646–61

[15] Barton G 1986 Quantum mechanics of the inverted oscillator potential *Ann. Phys.* **166** 322–63

[16] Broadbridge P and Zulkowski P 2006 Dark energy states from quantization of boson fields in a universe with unstable modes *Rep. Math. Phys.* **57** 27–40

[17] Feynman R P and Hibbs A R 2010 *Quantum Mechanics and Path Integrals* ed Emended (New York: Dover)

[18] Mandel L and Wolf E 1995 *Optical Coherence and Quantum Optics* (Cambridge: Cambridge University Press)

[19] Messiah A M 1999 *Quantum Mechanics* (New York: Dover)