Angular Momentum in Circular Waveguides

Kirk T. McDonald
Joseph Henry Laboratories, Princeton University, Princeton, NJ 08544
(June 29, 2013); updated June 1, 2014)

1 Problem

In view of recent interest in orbital angular momentum modes in optical fibers, discuss angular momentum in the modes of a circular waveguide of radius \( a \), with perfectly conducting walls and filled with a medium of index of refraction \( n \).

2 Solution

Circular-waveguide modes were first discussed in 1893 by J.J. Thomson in sec. 300 of [2] and then by Rayleigh in 1897 [3]. Although these modes are “well known,” discussion of them seldom includes mention of angular momentum.

2.1 General Relations

We first recall the general formalism appropriate for discussion of waves in pipes along the \( z \)-axis. The dependence on \( z \) and \( t \) of all field components is taken to have the form \( e^{i(k_z z - \omega t)} \).

Inside the waveguide these field components \( \psi \) obey the Helmholtz wave equation

\[
0 = \left( \nabla^2 + \frac{n_{\text{index}}^2 \omega^2}{c^2} \right) \psi = \left( \nabla^2_{\perp} + \frac{k^2}{k_0^2} \right) \psi = \left( \nabla^2_{\perp} - k_z^2 + k_0^2 \right) \psi = \left( \nabla^2_{\perp} + k_z^2 \right) \psi, \tag{1}
\]

where \( c \) is the speed of light in vacuum, \( n_{\text{index}} = \sqrt{\epsilon_{\text{relative}} \mu_{\text{relative}}} \) is the index of refraction at angular frequency \( \omega \) of the (linear, isotropic) medium inside the guide, and

\[
k_0 = \frac{n_{\text{index}} \omega}{c} = \sqrt{\epsilon \mu \omega}, \quad k^2 = k_0^2 - k_z^2. \tag{2}
\]

Note that \( k_z \) and not \( k \) is propagation constant.

The use of the decomposition \( \nabla^2 = \nabla^2_{\perp} + \partial^2/\partial z^2 \) in the wave equation (1) suggests a use of a similar decomposition of the gradient operator, \( \nabla = \nabla_{\perp} + \nabla_z \), such that two of Maxwell’s equations can be written

\[
-\frac{\partial \mathbf{B}}{\partial t} = i \omega \mu \mathbf{H} = i \omega \mu (\mathbf{H}_{\perp} + \mathbf{H}_z) = \nabla \times \mathbf{E} = \nabla_{\perp} \times (\mathbf{E}_{\perp} + \mathbf{E}_z) + \nabla_z \times (\mathbf{E}_{\perp} + \mathbf{E}_z), \tag{3}
\]

\[
\frac{\partial \mathbf{D}}{\partial t} = -i \omega \epsilon \mathbf{E} = i \omega \epsilon (\mathbf{E}_{\perp} + \mathbf{E}_z) = \nabla \times \mathbf{H} = \nabla_{\perp} \times (\mathbf{H}_{\perp} + \mathbf{H}_z) + \nabla_z \times (\mathbf{H}_{\perp} + \mathbf{H}_z). \tag{4}
\]

1See, for example, [1].
2Technical development of waveguides was precipitated by two important papers from Bell Labs in 1936 [5, 6]. The latter paper also discussed what would now be called optical fibers.
3This topic has recently been considered in [4].
4See, for example, chap. 13 of [7] or sec. 8.2 of [8].
In particular, $\nabla_\perp \times E_\perp = i\omega \mu H_z$ and $\nabla_\perp \times H_\perp = -\omega \epsilon E_z$ and $\nabla_z \times E_z = 0 = \nabla_z \times H_z$, such that

\begin{align}
  i\omega \mu H_\perp &= \nabla_\perp \times E_z + \nabla_z \times E_\perp = \nabla_\perp \times E_z + ik_z \hat{z} \times E_\perp, \\
  -i\omega \epsilon E_\perp &= \nabla_\perp \times H_z + \nabla_z \times H_\perp = \nabla_\perp \times H_z + ik_z \hat{z} \times H_\perp.
\end{align}

Substituting eqs. (5) and (6) into one another, we obtain,

\begin{align}
  i\omega^2 \epsilon \mu H_\perp &= \omega \epsilon \nabla_\perp \times E_z - k_z \hat{z} \times (\nabla_\perp \times H_z + ik_z \hat{z} \times H_\perp), \\
  ik_0^2 H_\perp &= \omega \epsilon \nabla_\perp \times E_z - k_z(\nabla_\perp H_z - ik_z H_\perp), \\
  H_\perp &= \frac{ik_z \nabla_\perp H_z - \omega \epsilon \nabla_\perp \times E_z}{k_0^2 - k_z^2} = \frac{ik_z \nabla_\perp H_z - \omega \epsilon \nabla_\perp \times E_z}{k^2},
\end{align}

and similarly,

\begin{align}
  E_\perp &= \frac{ik_z \nabla_\perp E_z + \omega \mu \nabla_\perp \times H_z}{k^2}.
\end{align}

That is, the transverse fields can be deduced from the longitudinal fields.

Furthermore, since the field equations are linear, we can decompose the fields into transverse electric modes (TE, with $E_z = 0$) and transverse magnetic modes (TM, with $H_z = 0$). Then, TE waves can be deduced from $H_z^{TE}$ according to

\begin{align}
  H_\perp^{TE} &= \frac{ik_z \nabla_\perp H_z^{TE}}{k^2}, \\
  E_\perp^{TE} &= \frac{\omega \mu \nabla_\perp \times H_z^{TE}}{k^2} = -\frac{\omega \mu}{k_z} \hat{z} \times H_\perp^{TE},
\end{align}

where the last form follows from eq. (5). Similarly, TM waves can be deduced from $E_z^{TM}$ according to

\begin{align}
  E_\perp^{TM} &= \frac{ik_z \nabla_\perp E_z^{TM}}{k^2}, \\
  H_\perp^{TM} &= -\frac{\omega \epsilon \nabla_\perp \times E_z^{TM}}{k^2} = \frac{\omega \epsilon}{k_z} \hat{z} \times E_\perp^{TM}, \\
  E_\perp^{TM} &= -\frac{k_z}{\omega \epsilon} \hat{z} \times H_\perp^{TM}.
\end{align}

Recall that the quantity

\begin{align}
  Z = \sqrt{\frac{\mu}{\epsilon}} = \sqrt{\frac{\mu_0}{\epsilon_0} \sqrt{\frac{\mu_{\text{relative}}}{\epsilon_{\text{relative}}}}} = 377 \sqrt{\frac{\mu_{\text{relative}}}{\epsilon_{\text{relative}}}} \text{ Ohms}
\end{align}

is often called the impedance of the medium.\(^5\) However, in waveguide literature one also finds the definitions,

\begin{align}
  Z^{TE} &= \sqrt{\frac{\mu k_0}{\epsilon k_z}} = \frac{\omega \mu}{k_z}, \\
  Z^{TM} &= \sqrt{\frac{\mu k_z}{\epsilon k_0}} = \frac{k_z}{\omega \epsilon},
\end{align}

\(^5\)This usage typically refers to “free space” plane waves, far from any conductors, where for wave propagation in the $z$-direction, $k_z = k_0$ and $E = ZH \times \hat{z}$. Waves inside hollow conductors are not “free,” but rather are “inhomogeneous/evanescent” in the sense of these terms as used in a decomposition of the waves into electromagnetic plane waves [10]. As seen in sec. 2.2 below, guided waves with very small wavelengths have $k_z \approx k_0$ so in this limit, where in some sense much of the wave is far from the conductors, the definitions (13) and (14) coincide.
used in conjunction with the relations
\[ \mathbf{E}_{\perp}^{\text{TE}} = Z_{\text{TE}}^{*} \mathbf{H}_{\perp}^{\text{TE}} \times \hat{z}, \quad \mathbf{E}_{\perp}^{\text{TM}} = Z_{\text{TM}}^{*} \mathbf{H}_{\perp}^{\text{TM}} \times \hat{z}. \]  

(15)

The time-average flow of electromagnetic energy in the guide is described by the Poynting vector,

\[ \langle \mathbf{S} \rangle = \frac{1}{2} \text{Re}(\mathbf{E} \times \mathbf{H}^{*}) = \frac{1}{2} \text{Re}(\mathbf{E}_{\perp} \times \mathbf{H}_{\perp}^{*}) + \frac{1}{2} \text{Re}(\mathbf{E}_{z} \times \mathbf{H}_{z}^{*}) \]

\[ \langle \mathbf{S}^{\text{TE}} \rangle = \frac{Z_{\text{TE}}^{*}}{2} \text{Re} \left[ (\mathbf{H}_{\perp}^{\text{TE}} \times \hat{z}) \times \mathbf{H}_{\perp}^{\text{TE}} \right] + \frac{Z_{\text{TE}}^{*}}{2} \text{Re} \left[ (\mathbf{H}_{z}^{\text{TE}} \times \hat{z}) \times \mathbf{H}_{z}^{\text{TE}} \right] \]

(16)

\[ \langle \mathbf{S} \rangle = \frac{1}{2} \text{Re} \left( \mathbf{E}_{\perp}^{\text{TM}} \times (\hat{z} \times \mathbf{E}_{\perp}^{\text{TM}}) \right) + \frac{1}{2} \text{Re} \left( \mathbf{E}_{z}^{\text{TM}} \times (\hat{z} \times \mathbf{E}_{\perp}^{\text{TM}}) \right) \]

(17)

As expected, energy flows down the guide (parallel to the z-axis), but lines of (time-average) energy flow can have components in the transverse plane, such that these “streamlines” follow twisted paths corresponding to the presence of angular momentum in the wave.6

The time-average (Minkowksi) density of momentum in the wave is given by

\[ \langle \mathbf{p} \rangle = \frac{1}{2} \text{Re}(\mathbf{D} \times \mathbf{B}^{*}) = \langle \mathbf{S} \rangle \frac{\varepsilon \mu}{c^2} = n_{\text{index}}^{2} \frac{\langle \mathbf{S} \rangle}{c^2}, \]

(19)

and the time-average density of angular momentum is

\[ \langle \mathbf{l} \rangle = \mathbf{r} \times \langle \mathbf{p} \rangle = \mathbf{r} \times n_{\text{index}}^{2} \frac{\langle \mathbf{S} \rangle}{c^2}. \]

(20)

2.2 Circular Waveguides

Turning at last to the special case of waves inside a circular cylinder with a perfect conductor at radius \( r = a \) in a cylindrical coordinate system \( (r, \theta, z) \), we can deduce the waveforms from either the scalar function \( \psi = H_{z} \) (TE waves) or \( \psi = E_{z} \) (TM waves), where \( \psi \) obeys the wave equation (1),

\[ \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial \psi}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 \psi}{\partial \theta^2} + \left( k_{0}^{2} - k_{z}^{2} \right) \psi = 0 \]

(21)

As we seek solutions that could exhibit angular momentum, we consider the form

\[ \psi = R_{mn}(r) e^{\pm im\theta} e^{i(k_{z}mnz - \omega t)}, \]

(22)

6The nonzero, time-average, transverse flow of energy should not be called “reactive” as in sec. 8.5 of [8], as this term is reserved for instantaneous flow of energy whose time average is zero (for example, in the fields outside the conductors of an LC circuit, or those near an antenna where one speaks of the “reactive near field” as that part of the field that does not contribute to the flow of energy to “infinity”). See [9] for discussion of some subtleties to the use of the term “reactance” for antennas.
where here \( n \) is not the index of refraction, and \( m \) is a non-negative integer. Then, the wave equation (21) reduces to

\[
\frac{1}{r} \frac{d}{dr} \left( r \frac{dR_{mn}}{dr} \right) + \left( k_0^2 - k_z^2 - \frac{m^2}{r^2} \right) R_{mn} = 0. \tag{23}
\]

This is a form of Bessel’s equation, and we write the relevant solutions as

\[
R_{mn}(r) = J_m(k_{mn}r), \quad \text{where} \quad k_{mn} = \sqrt{k_0^2 - k_{z,\text{mn}}^2}, \tag{24}
\]

and \( J_m \) is an “ordinary” Bessel function of the first kind.

The waves must obey the perfect-conductor boundary condition at \( r = a \) that the tangential component of the electric field vanish there. Hence, for TM waves, where \( \psi = E_z^{\text{TM}} \), we have that

\[
J_m(k_{\text{TM,\text{mn}}a}) = 0. \tag{25}
\]

We adopt the notation

\[
J_m(u_{\text{mn}}) = 0, \tag{26}
\]

for the zeroes of the Bessel functions, such that

\[
\begin{align*}
    u_{01} &\approx 2.405, \quad u_{02} \approx 5.520, \quad u_{03} \approx 8.564, \quad u_{04} \approx 11.791, \quad \cdots \\
    u_{11} &\approx 3.832, \quad u_{12} \approx 7.016, \quad u_{13} \approx 10.174, \quad u_{14} \approx 13.324, \quad \cdots \\
    u_{21} &\approx 5.136, \quad u_{22} \approx 8.417, \quad u_{23} \approx 11.620, \quad u_{24} \approx 14.796, \quad \cdots
\end{align*} \tag{27}
\]

Then,

\[
k_{\text{TM,\text{mn}}}^2 = \frac{u_{\text{mn}}^2}{a^2} = k_0^2 - k_{z,\text{mn}}^2, \tag{28}
\]

and the (so-called) guide wavelength in \( z \) is given by

\[
\lambda_{z,\text{mn}}^{\text{TM}} = \frac{2\pi}{k_{\text{z,\text{mn}}}^{\text{TM}}} = \frac{2\pi}{\sqrt{k_0^2 - k_{z,\text{mn}}^2}} = \frac{2\pi a}{\sqrt{(2\pi a/\lambda_0)^2 - u_{\text{mn}}^2}} = \frac{\lambda_0}{\sqrt{1 - (\lambda_0 u_{\text{mn}}/2\pi a)^2}}, \tag{29}
\]

where the “free space” wavelength at frequency \( \nu = 2\pi \omega \) is

\[
\lambda_0 = \frac{2\pi}{k_0} = \sqrt{\epsilon_{\text{relative}} \mu_{\text{relative}}} \frac{\nu}{c}. \tag{30}
\]

Waves propagate only for sufficiently short wavelengths (high frequencies); the longest possible TM wavelength (in free space) is for the 01 mode, with \( \lambda_{\text{max}}^{\text{TM}} = 2\pi/k_0^{\text{TM}} = 2.61a \).

The TM wave fields can now be written as, using eq. (12) (with \( E_0 \) complex), noting eq. (28),\(^7\)

\[
E_z^{\text{TM}} = E_0 J_m \left( \frac{r}{a} u_{\text{mn}} \right) e^{\pm im\theta} e^{i(k_{z,\text{mn}}z - \omega t)}, \tag{31}
\]

\(^7\)We follow the notation of sec. 9.18 of [11], although that section used \( \cos m\theta \) and \( \sin m\theta \) rather than \( e^{\pm im\theta} \) in the wavefunctions, which implies that the angular momentum is zero.
\[ E_r^{TM} = \frac{i E_0 u_{mn} k_{z, mn} a^2}{a} J'_m \left( \frac{r}{a} u_{mn} \right) e^{\pm i m \theta} e^{i(k_{z, mn} z - \omega t)}, \]
\[ E_\theta^{TM} = \frac{-m E_0 k_{z, mn} a^2}{r} J_m \left( \frac{r}{a} u_{mn} \right) e^{\pm i m \theta} e^{i(k_{z, mn} z - \omega t)}, \]
\[ H_r^{TM} = -\frac{E_\theta^{TM}}{Z_{TM}}, \]
\[ H_\theta^{TM} = \frac{E_r^{TM}}{Z_{TM}}. \]

The time-average flow of energy inside the guide follows from eqs. (18) as
\[
\langle S_{TM} \rangle = \frac{1}{2Z_{TM}} |E_{\perp}^{TM}|^2 \hat{z} + \frac{1}{2Z_{TM}} \text{Re} \left( E_{\perp}^{TM} E_{\perp}^{TM*} \right) = \left\{ \frac{|E_0|^2 k_{z, mn} a^2}{2Z_{TM}} \left[ \frac{m^2}{r^2} J_m^2 \left( \frac{r}{a} u_{mn} \right) + \frac{u_{mn}^2}{a^2} J'_m \left( \frac{r}{a} u_{mn} \right) \right] \right\} \hat{z} + \frac{m}{r} J_m^2 \left( \frac{r}{a} u_{mn} \right) \hat{\theta}. \quad (36)
\]

There is no radial component to the time-average energy flow, whose streamlines follow helices of constant radius; however, the pitch of these helices is a function of radius, so the flow pattern is complex.

The time-average momentum density per unit length in the guide is, for TM waves,
\[
\langle P_{TM} \rangle = \frac{n^2 \pi}{c^2} \frac{|E_0|^2 k_{z, mn} a^2}{u_{mn}^2} \int_0^a r \, dr \left[ \frac{m^2}{r^2} J_m^2 \left( \frac{r}{a} u_{mn} \right) + J'_m \left( \frac{r}{a} u_{mn} \right) \right] \hat{z} = \frac{n^2 \pi}{c^2} \frac{|E_0|^2 k_{z, mn} a^2}{u_{mn}} J_m^2 \left( u_{mn} \right) \hat{z}, \quad (37)
\]
recalling eq. (19), noting that \( J_{m-1} + J_{m+1} = 2m J_m(x)/x, J_{m-1} - J_{m+1} = 2dJ_m(x)/dx, \)
\( J_m(u_{mn}) = 0, J_{m-1}(u_{mn}) = -J_m(u_{mn}) = J'_m(u_{mn}), \) and using eq. 5.11(11), p. 135 of [12] (due to Lommel),
\[
\int x \, dx J_m^2(kx) = \frac{x^2}{2} \left[ J_m^2(kx) - J_{m-1}(kx) J_{m+1}(kx) \right] = \frac{x^2}{2} \left[ 1 - \frac{m^2}{k^2 x^2} \right] J_m^2(kx) + J'_m^2(kx). \quad (38)
\]

The time-average angular-momentum density per unit length is, recalling eq. (20) and using eq. (38) for \( J_m(u_{mn}) = 0, \)
\[
\langle L_{TM} \rangle = \pm \frac{n^2 \pi}{c^2} \frac{m |E_0|^2 k_{z, mn} a^2}{Z_{TM}} \int_0^a r \, dr \, J_m^2 \left( \frac{r}{a} u_{mn} \right) \hat{z}
\]
\[
\begin{align*}
&= \pm \frac{n_{\text{index}}^2 m \pi |E_0|^2}{c^2 Z_{\text{TM}}} \frac{k_{z,\text{mn}} a^4}{u_{mn}^2} \int_0^1 x \, dx \, J_m^2(u_{mn}x) \, \hat{z} \\
&= \pm \frac{n_{\text{index}}^2 m \pi |E_0|^2}{c^2 Z_{\text{TM}}} \frac{k_{z,\text{mn}} a^4}{u_{mn}^2} J_m^2(u_{mn}) \, \hat{z} = \pm m \frac{\langle P_{\text{TM}} \rangle}{k_{z,\text{mn}}}. 
\end{align*}
\] (39)

For TE waves where \( \psi = H_{z,\text{TE}} \) and \( E_{\theta,\text{TE}} \propto \partial H_{z,\text{TE}} / \partial r \) according to eq. (11), the condition that \( E_{\theta,\text{TE}}(r = a) = 0 \) implies that

\[
J'_m(k_{m,\text{mn}} a) = 0. \tag{40}
\]

We adopt the notation

\[
J'_m(u'_{mn}) = 0, \tag{41}
\]

for the zeroes of the derivatives of the Bessel functions, such that

\[
\begin{align*}
u'_{01} & \approx 3.832, & u'_{02} & \approx 7.016, & u'_{03} & \approx 10.174, & u'_{04} & \approx 13.324, & \cdots \\
u'_{11} & \approx 1.841, & u'_{12} & \approx 5.331, & u'_{13} & \approx 8.536, & u'_{14} & \approx 8.536, & \cdots \\
u'_{21} & \approx 3.054, & u'_{22} & \approx 6.706, & u'_{23} & \approx 9.970, & u'_{24} & \approx 9.970, & \cdots
\end{align*}
\] (42)

Then,

\[
k_{\text{TE}}^2 = \frac{u_{mn}^2}{a^2} = k_0^2 - k_{z,\text{mn}}^2, \tag{43}
\]

and the guide wavelength (in \( z \)) is given by

\[
\lambda_{\text{TE}}^{z,\text{mn}} = \frac{2\pi}{k_{z,\text{mn}}^2} = \frac{2\pi}{\sqrt{k_0^2 - k_{\text{TE}}^2}} = \frac{2\pi a}{\sqrt{(2\pi a/\lambda_0)^2 - u_{mn}^2}}, \tag{44}
\]

The maximum (free-space) TE wavelength is for the 11 mode, with \( \lambda_{\text{max}}^{\text{TE}} = 3.41a \), which is longer than \( \lambda_{\text{max}}^{\text{TM}} \).

The TE wave fields can now be written using eq. (11) as

\[
\begin{align*}
H_{z,\text{TE}} &= H_0 J_m \left( \frac{r}{a} u'_{mn} \right) e^{\pm im\theta} e^{i(k_{z,\text{mn}} z - \omega t)}, \tag{45} \\
H_{r,\text{TE}} &= i H_0 u'_{mn} \frac{k_{z,\text{mn}} a^2}{a} J'_m \left( \frac{r}{a} u'_{mn} \right) e^{\pm im\theta} e^{i(k_{z,\text{mn}} z - \omega t)}, \tag{46} \\
H_{\theta,\text{TE}} &= m H_0 \frac{k_{z,\text{mn}} a^2}{a} J_m \left( \frac{r}{a} u'_{mn} \right) e^{\pm im\theta} e^{i(k_{z,\text{mn}} z - \omega t)}, \tag{47} \\
E_{r,\text{TE}} &= Z_{\text{TE}} H_{\theta,\text{TE}}, \tag{48} \\
E_{\theta,\text{TE}} &= -Z_{\text{TE}} H_{r,\text{TE}}. \tag{49}
\end{align*}
\]
The time-average flow of energy inside the guide follows from eqs. (17) as

\[
\left\langle S^{\text{TE}} \right\rangle = \frac{Z_{\text{TE}}}{2} |H_{\perp}^{\text{TE}}|^2 \hat{z} + \frac{Z_{\text{TE}}}{2} \text{Re} \left( H_{z}^{\text{TE}} H_{\perp}^{\text{TE}} \right)
\]

\[
= \frac{Z_{\text{TE}}}{2} |H_0|^2 \frac{k_{z,mm} a^2}{u_{mm}^2} \left\{ \frac{m^2}{r^2} J_m^2 \left( \frac{r}{a} u_{mm}' \right) + \frac{u_{mm}^2}{a^2} J_m^2 \left( \frac{r}{a} u_{mm}' \right) \right\} \hat{z}
\]

\[
\mp \frac{m}{r} J_m^2 \left( \frac{r}{a} u_{mm}' \right) \hat{\theta}.
\]

(50)

Again, the energy-flow lines are helices of constant radii.

The time-average momentum density per unit length in the guide is,

\[
\left\langle P^{\text{TE}} \right\rangle = \frac{n_{\text{index}}^2 \pi Z_{\text{TE}} |H_0|^2}{c^2} \left( \frac{k_{z,mm} a^2}{u_{mm}^2} \right)^2 \frac{u_{mm}^2}{a^2} \int_0^a r dr \left[ \frac{m^2}{r^2} u_{mm}^2 J_m^2 \left( \frac{r}{a} u_{mm}' \right) + J_m^2 \left( \frac{r}{a} u_{mm}' \right) \right] \hat{z}
\]

\[
= \frac{n_{\text{index}}^2 \pi Z_{\text{TE}} |H_0|^2}{c^2} \frac{k_{z,mm} a^4}{u_{mm}^4} \int_0^1 x dx \left[ \frac{m^2}{u_{mm}^2} J_m^2(u_{mm}',x) + J_m^2(u_{mm}',x) \right] \hat{z}
\]

\[
= \frac{n_{\text{index}}^2 \pi Z_{\text{TE}} |H_0|^2}{c^2} \frac{k_{z,mm} a^4}{u_{mm}^4} \left\{ J_{m-1}(u_{mm}') + J_{m+1}(u_{mm}') \right\} \hat{z}
\]

\[
= \frac{n_{\text{index}}^2 \pi Z_{\text{TE}} |H_0|^2}{c^2} \frac{k_{z,mm} a^4}{u_{mm}^4} \left( \frac{2 m^2}{u_{mm}^2} J_m^2(u_{mm}') - 2 \left( \frac{2 m^2}{u_{mm}^2} - 1 \right) J_m^2(u_{mm}') \right) \hat{z}
\]

\[
= \frac{n_{\text{index}}^2 \pi Z_{\text{TE}} |H_0|^2}{c^2} \frac{k_{z,mm} a^4}{u_{mm}^4} \left( 1 - \frac{m^2}{u_{mm}^2} \right) J_m^2(u_{mm}') \hat{z},
\]

(51)

using eq. (38) and that \( J_{m-1} + J_{m+1} = 2mJ_m(x)/x, \) \( J_{m-1} - J_{m+1} = 2J_m', \) so \( J_m' (u_{mm}') = 0 \) implies that \( J_{m-1}(u_{mm}') = J_{m+1}(u_{mm}') = mJ_m(u_{mm}')/u_{mm}' \). The time-average angular-momentum density per unit length is, using eq. (38) for \( J_m' (u_{mm}') = 0 \),

\[
\left\langle L^{\text{TE}} \right\rangle = \mp \frac{n_{\text{index}}^2}{c^2} \pi m Z_{\text{TE}} |H_0|^2 \frac{k_{z,mm} a^2}{u_{mm}^2} \int_0^a r dr \left( \frac{r}{a} u_{mm}' \right) \hat{z}
\]

\[
= \mp \frac{n_{\text{index}}^2}{c^2} \pi m Z_{\text{TE}} |H_0|^2 \frac{k_{z,mm} a^4}{u_{mm}^4} \int_0^1 x dx \left( \frac{r}{a} u_{mm}' \right) \hat{z}
\]

\[
= \mp \frac{n_{\text{index}}^2}{c^2} \frac{m \pi Z_{\text{TE}} |H_0|^2}{c^2} \frac{k_{z,mm} a^4}{u_{mm}^4} \left( 1 - \frac{m^2}{u_{mm}^2} \right) J_m^2(u_{mm}') \hat{z} = \mp m \frac{\left\langle P^{\text{TE}} \right\rangle}{k_{z,mm}^2}.
\]

(52)

In a quantum view, photons of the waveguide modes carry momentum \( \mathbf{P} = \hbar k_z \hat{z} \), and hence these photons have angular momentum \( \mathbf{L} = \mp m \hbar \hat{z} \). The calculation here, \( \mathbf{L} = r \times \mathbf{P} \), corresponds to orbital angular momentum.

\[\text{Footnotes:}
\]

8 Some authors imply that the result (51) should be more evident, but I had to slog through the above.

9 As \( k \neq \omega/c \), the waveguide photons are “virtual” and the guided waves are “evanescent.” Regarding the latter, see sec. 2.8 of [10].
2.2.1 Phase, Group and Energy-Flow Velocities

The Poynting vector $\mathbf{S}$ has dimensions of energy density time velocity, which suggests that we define an energy-flow velocity as

$$v_E = \frac{\mathbf{S}}{u} = \frac{\mathbf{E} \times \mathbf{H}}{u} ,$$

(53)

where the electromagnetic energy density is

$$u = \frac{\mathbf{E} \cdot \mathbf{D} + \mathbf{B} \cdot \mathbf{H}}{2} = \frac{\epsilon E^2 + \mu H^2}{2} ,$$

(54)

with $\mathbf{E}$ and $\mathbf{H}$ are purely real in eqs. (53)-(54). However, both $\mathbf{S}$ and $u$ vary rapidly in space and time, so their use in eq. (53) would lead to a velocity field with dramatic fluctuations. These fluctuations could be smoothed by considering only the time-average densities,

$$\langle \mathbf{S} \rangle = \frac{Re(\mathbf{E} \times \mathbf{H}^*)}{2} , \quad \langle u \rangle = \frac{Re(\mathbf{E} \cdot \mathbf{D}^* + \mathbf{B} \cdot \mathbf{H}^*)}{4} = \frac{\epsilon |E|^2 + \mu |H|^2}{4} ,$$

(55)

where the time-average Poynting vectors $\langle \mathbf{S} \rangle$ for TM and TE modes are given in eqs. (36) and (50). The time-average field energy density for TM modes is

$$\langle u_{TM} \rangle = \frac{\epsilon |E_{TM}^*|^2 + \mu |H_{TM}^*|^2}{4} = \frac{\epsilon \left( |E_{TM}^z|^2 + |E_{TM}^\perp|^2 \right) + \mu |E_{TM}^\perp|^2 / Z_{TM}^2}{4}$$

$$= \frac{\epsilon}{4} \left| E_{TM}^z \right|^2 + \left( \frac{k_0^2}{k_{TM}^2} \right) \left( 1 + \frac{k_0^2}{k_{TM}^2} \right)$$

$$= \frac{\epsilon |E_0|^2}{4} \left\{ J_m^2 \left( \frac{r}{a} u_{mn} \right) + \frac{r}{a} u_{mn} \left[ \frac{m^2 a^2}{u_{mn}^2 r^2} J_m^2 \left( \frac{r}{a} u_{mn} \right) + J_m^2 \left( \frac{r}{a} u_{mn} \right) \right] \right\}$$

(56)

The ratio $\langle \mathbf{S}_{TM}^z \rangle / \langle u_{TM} \rangle$ is a complicated function of radius $r$, and seems still to be too fine-grained a definition for the energy-flow velocity of the wave.

A coarser-grained definition is to consider the ratio $\int \langle \mathbf{S}_{TM}^z \rangle \, d\text{Area} / \langle U_{TM} \rangle$ of the total power flowing down the guide to the energy density per unit length, where $\int \langle \mathbf{S}_{TM}^z \rangle \, d\text{Area}$ has been given to within a factor in eq. (37). Of course, this definition averages over the helical flow of the Poynting vector $\langle \mathbf{S} \rangle$.

The energy density per unit length is

$$\langle U_{TM} \rangle = \frac{\pi \epsilon |E_0|^2}{2} \int_0^a r \, dr \left\{ J_m^2 \left( \frac{r}{a} u_{mn} \right) + \left( k_{TM}^2 + k_0^2 \right) \frac{a^2}{u_{mn}^2} \left[ \frac{m^2 a^2}{u_{mn}^2 r^2} J_m^2 \left( \frac{r}{a} u_{mn} \right) + J_m^2 \left( \frac{r}{a} u_{mn} \right) \right] \right\}$$

$$= \frac{\pi \epsilon |E_0|^2}{2} \int_0^1 x \, dx \left\{ J_m^2 (u_{mn} x) + \left( k_{TM}^2 + k_0^2 \right) \frac{a^2}{u_{mn}^2} \left[ \frac{m^2 a^2}{u_{mn}^2 x^2} J_m^2 (u_{mn} x) + J_m^2 (u_{mn} x) \right] \right\}$$

$$= \frac{\pi \epsilon |E_0|^2}{4} J_m^2 (u_{mn}) \left[ 1 + \left( k_{TM}^2 + k_0^2 \right) \frac{a^2}{u_{mn}^2} \right] = \frac{\pi \epsilon |E_0|^2}{2} \frac{k_0^2 a^4}{u_{mn}^2} J_m^2 (u_{mn}).$$

(57)
Hence, the coarse-grained energy-flow velocity for TM waves is

$$\bar{v}_E^{TM} \equiv \frac{\int \langle S_z^{TM} \rangle d\text{Area}}{\langle U^{TM} \rangle} = \frac{c^2}{n_{\text{index}}^2} \frac{\langle P^{TM} \rangle}{\langle U^{TM} \rangle} = \frac{1}{Z_{TM}} \frac{k_{z,mn}^{TM}}{k_0} \bar{z} = \frac{c}{n_{\text{index}}} \frac{k_{z,mn}^{TM}}{k_0} \leq \frac{c}{n_{\text{index}}} \hat{z}. \quad (58)$$

A similar calculation for TE waves gives the same form.

A different approach to characterization of wave velocities is the so-called eikonal method, introduced by Sommerfeld and Runge [13]. We follow two papers of Whitham [14, 15], which are variants of arguments by Landau [16, 17]. See also sec. 45 of [18], and the author’s note [19].

An argument applicable to waves of all types that are far from localized sources is based on the approximation that any scalar component of the wave function can be written as

$$\psi = A(\mathbf{r}, t)e^{i\varphi(\mathbf{r}, t)}, \quad (59)$$

where the $A$ is a slowly-varying (complex) amplitude and $\varphi$ is a rapidly varying (real) phase (sometimes also called the eikonal). In any small region (far from the source) the form (59) is nearly a plane wave $e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)}$ with wave vector $\mathbf{k}$ and angular frequency $\omega$ obtained from the first-order terms in a Taylor expansion of the phase $\varphi$,

$$\varphi(\mathbf{r}, t) = \varphi_0 + \nabla \varphi \cdot \mathbf{r} + \frac{\partial \varphi}{\partial t} t + ..., \quad (60)$$

such that we identify

$$\mathbf{k} = \nabla \varphi, \quad \text{and} \quad \omega = -\frac{\partial \varphi}{\partial t}. \quad (61)$$

The locally plane wave has phase velocity

$$\mathbf{v}_p = \frac{\omega}{k} \hat{\mathbf{k}} = \frac{\omega}{k^2} \mathbf{k}, \quad (62)$$

where $k = |\mathbf{k}|$. It also follows from eq. (61) that

$$\frac{\partial \mathbf{k}}{\partial t} = -\nabla \omega. \quad (63)$$

The plane wave $e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)}$ is at least an approximate solution to some wave equation. Using the plane wave as a trial solution to this wave equation leads to a functional relation between $\omega$ and $\mathbf{k}$ (and possibly $\mathbf{r}$ as well) called the dispersion relation, which we write as

$$\omega = \omega(\mathbf{k}, \mathbf{r}). \quad (64)$$

The dispersion relation can be used to generate the equations of the geometric or ray approximation as follows.

We first note that since $\mathbf{k} = \nabla \varphi$ we have that $\nabla \times \mathbf{k} = 0$, i.e., $\partial k_i/\partial x_j = \partial k_j/\partial x_i$. Then, if we use the dispersion relation in eq. (63), the $i$th component of that equation can be rewritten as

$$\frac{\partial k_i}{\partial t} = -\frac{\partial \omega}{\partial x_i} \sum_j \frac{\partial \omega}{\partial k_j} \frac{\partial k_j}{\partial x_i} = -\frac{\partial \omega}{\partial x_i} \sum_j \frac{\partial \omega}{\partial k_j} \frac{\partial k_j}{\partial x_i} = -\frac{\partial \omega}{\partial x_i} - \sum_j v_{g,j} \frac{\partial k_i}{\partial x_j}, \quad (65)$$
where
\[ v_g = \frac{\partial \omega}{\partial k} = \nabla_k \omega \] (66)
is the group velocity, so that
\[ \frac{dk}{dt} = \frac{\partial k}{\partial t} + (v_g \cdot \nabla)k = -\partial_x \omega. \] (67)

We can interpret eq. (67) as implying that for an observer who moves with velocity
\[ v_g = \frac{dr}{dt} = \nabla_k \omega \] (68)
in a homogeneous medium (i.e., one for which \( \partial_x \omega = 0 \)), the wave vector \( k \) remains constant.\(^{10}\)

This result leads us to introduce the concept of a ray (in ordinary space) whose direction is that of the group velocity \( v_g \). In a homogeneous medium the wave vector \( k \) is constant along a ray (although \( k \) is not necessarily parallel to \( v_g \)).\(^{11}\)

Furthermore, in a homogeneous medium the gradient \( \nabla_k \omega \) is constant along a ray, since \( \omega \) is only a function of \( k \) in such a medium, and \( k \) is constant along a ray. Hence, the group velocity vector \( v_g = \nabla_k \omega \) is constant along a ray, and the rays are straight lines in a homogeneous medium. This result holds even if the medium is anisotropic, and it holds whether or not the medium is linear.

If we suppose a ray to be associated with the Hamiltonian
\[ H = \omega(r, k) \] (70)
then Hamilton’s equations of motion,
\[ \frac{dr}{dt} = \nabla_k H = \nabla_k \omega, \quad \text{and} \quad \frac{dk}{dt} = -\nabla_H = -\nabla \omega, \] (71)
lead to the same forms as eqs. (67)-(68). Thus, the ray concept, which derives from the view of Fermat that light is a particle phenomenon, together with the principles of Hamiltonian mechanics, leads us to suppose that the energy of a particle of light is proportional to its frequency. Although this argument is perhaps the most compact derivation of so-called Hamiltonian optics, it was not made by Hamilton. Rather, it was Einstein [21] who first

\(^{10}\) For waveguide modes with nonzero \( m \) (which carry angular momentum), eqs. (2) and (28) can be combined to give
\[ \omega(k, r) = \frac{k_0 c}{n_{\text{index}}} = \frac{c}{n_{\text{index}}} \sqrt{k_{z,mn}^2 + \frac{u_{mn}^2}{a^2}} = \frac{c}{n_{\text{index}}} \sqrt{k_{\theta,m}^2 + k_{z,mn}^2 - \frac{m^2}{r^2} + \frac{u_{mn}^2}{a^2}}, \] (69)
which depends on the radius \( r \) if one uses the form that also depends on \( k_{\theta,m} \). Hence, the guide medium, even if vacuum, is formally not “homogeneous” in the present sense for waves that carry angular momentum. In this case, the wave vector \( k \) is not expected to be (and is not) constant along a group-velocity ray.

\(^{11}\) See, for example, the figure on p. 5 of [20].
noted the relation between frequency and energy for quanta of light, while the use of eqs. (70)-(71) as the basis for geometric optics appears to be due to Landau [16].

For the present case of guided waves with wavefunctions of the form (22), the eikonal is

$$\varphi = k_{z,mn} z \pm m \theta - \omega t, \quad k_{z,mn} = \sqrt{k_0^2 - \frac{u_{mn}^2}{a^2}} = \sqrt{\frac{n_{\text{index}}^2 \omega^2}{c^2} - \frac{u_{mn}^2}{a^2}},$$ (72)

so the wave vector $\mathbf{k}$ in the sense of the eikonal method is

$$\mathbf{k} = \nabla \varphi = k_{z,mn} \hat{z} \pm \frac{m}{r} \hat{\theta}, \quad k = \sqrt{k_{\theta,m}^2 + k_{z,mn}^2} = \sqrt{k_0^2 + \frac{m^2}{r^2} - \frac{u_{mn}^2}{a^2}}.$$ (73)

The eikonal phase velocity is

$$v_p = \frac{\omega}{k^2} \mathbf{k}, \quad v_p = \frac{\omega}{k} = \frac{c}{n_{\text{index}}} \frac{k_0}{k}.$$ (74)

such that lines of the phase-velocity field are helices with pitch that depends on radius $r$. The eikonal group velocity follows from the last form of eq. (69) as

$$\mathbf{v}_g = \frac{\partial \omega}{\partial \mathbf{k}} = \frac{\partial \omega}{\partial k_{z,mn}} \hat{z} + \frac{\partial \omega}{\partial k_{\theta,m}} \hat{\theta} = \frac{c}{n_{\text{index}}} \frac{k}{k_0} = \frac{k^2}{k_0^2} \mathbf{v}_p = \left(1 + \frac{m^2}{k_0^2 r^2} - \frac{u_{mn}^2}{k_0^2 a^2}\right) \mathbf{v}_p,$$ (75)

whose magnitude grows arbitrarily large for small $r$.

This result is similar to the awkward behavior of the ratio $\langle S_{\text{TM}}(r) \rangle / \langle u_{\text{TM}}(r) \rangle$ that was a candidate for the energy-flow velocity as discussed earlier in this section. We could avoid this behavior by using the second-to-last form of eq. (69), in which case we have

$$\mathbf{v}_g \equiv \frac{\partial \omega}{\partial k_{z,mn}} \hat{z} = \frac{c}{n_{\text{index}}} \frac{k_{z,mn}}{k_0} \hat{z} = \mathbf{v}_E,$$ (76)

which is the same as the energy-flow velocity found in eq. (58) (after a similar disregard of awkwardness associated with the azimuthal term in the Poynting vector). Lines of $\mathbf{v}_g$ are straight.

Equation (76) is an example of the general identity between group velocity and energy-flow velocity in homogeneous (linear) media, as reviewed in sec. 2.1 of [19].

\[12\] Although Schrödinger used Hamiltonian optics to motivate his equation for the quantum behavior of particles [22], he appears not to have considered the inverse notion of the quantum relation $E = \hbar \omega$ for the energy of particles of light as a starting point for Hamiltonian optics.

\[13\] In [16] Landau explicitly identifies the angular frequency $\omega$ as the Hamiltonian for geometrical optics, but the only medium he considers is vacuum. In [17] he considers anisotropic media that support mechanical waves and notes that if the medium is homogeneous then the rays are straight lines; but he does not explicitly identify the dispersion relation $\omega(k)$ with the Hamiltonian. In [23] he considers the optics/electrodynamics of anisotropic media but omits mention that the rays are straight lines in homogeneous anisotropic media, and of the connection between ray optics and Hamiltonian mechanics.
References


