5

LIGHT IN VACUUM

Theory of optical polarization

Introduction. In regions empty of matter—empty more particularly of *charged* matter—the electromagnetic field is described by equations that we have learned to write in various ways:

$$\left.\begin{aligned} \boldsymbol{\nabla}\cdot\boldsymbol{E} &= 0 \\ \boldsymbol{\nabla}\times\boldsymbol{B} - \tfrac{1}{c}\tfrac{\partial}{\partial t}\boldsymbol{E} &= \boldsymbol{0} \\ \boldsymbol{\nabla}\cdot\boldsymbol{B} &= 0 \\ \boldsymbol{\nabla}\times\boldsymbol{E} + \tfrac{1}{c}\tfrac{\partial}{\partial t}\boldsymbol{B} &= \boldsymbol{0} \end{aligned}\right\} \tag{65}$$

$$\left.\begin{aligned} \partial_\mu F^{\mu\nu} &= 0 \\ \partial_\alpha \epsilon^{\alpha\rho\sigma\nu} F_{\rho\sigma} &= 0 \end{aligned}\right\} \tag{168}$$

$$\left.\begin{aligned} F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu \\ \Box A^\nu - \partial^\nu(\partial_\mu A^\mu) &= 0 \quad : \quad \text{arbitrary gauge} \\ &\downarrow \\ \Box A^\nu &= 0 \quad : \quad \text{Lorentz gauge} \end{aligned}\right\} \tag{371}$$

And we have learned that, whichever language we adopt, multiple instances of the wave equation hover close by. It was Maxwell himself who first noticed that equations (65) can be "decoupled by differentiation" to yield six copies of the wave equation:

$$\Box\boldsymbol{E} = \Box\boldsymbol{B} = \boldsymbol{0}$$

The manifestly covariant version of Maxwell's argument is less familiar: to

$$\partial^a \epsilon_{ars\nu} \cdot \partial_\alpha \epsilon^{\alpha\rho\sigma\nu} F_{\rho\sigma} = 0$$

bring the identity[233]

$$\begin{aligned}
\epsilon_{ars\nu}\epsilon^{\alpha\rho\sigma\nu} = \tfrac{1}{g}\delta^{\alpha\rho\sigma}{}_{ars} \equiv \tfrac{1}{g}\begin{vmatrix} \delta^\alpha{}_a & \delta^\alpha{}_r & \delta^\alpha{}_s \\ \delta^\rho{}_a & \delta^\rho{}_r & \delta^\rho{}_s \\ \delta^\sigma{}_a & \delta^\sigma{}_r & \delta^\sigma{}_s \end{vmatrix} \\
= \tfrac{1}{g}\big\{ \quad \delta^\alpha{}_a(\delta^\rho{}_r\delta^\sigma{}_s - \delta^\sigma{}_r\delta^\rho{}_s) \\
+ \delta^\alpha{}_r(\delta^\rho{}_s\delta^\sigma{}_a - \delta^\sigma{}_s\delta^\rho{}_a) \\
+ \delta^\alpha{}_s(\delta^\rho{}_a\delta^\sigma{}_r - \delta^\sigma{}_a\delta^\rho{}_r)\big\}
\end{aligned}$$

and obtain

$$\Box(F_{rs} - F_{sr}) + \partial^a\big\{\partial_r(F_{sa} - F_{as}) + \partial_s(F_{ar} - F_{ra})\big\} = 0$$

whence (by the antisymmetry of $F_{\mu\nu}$)

$$\begin{aligned}
\Box F_{\mu\nu} &= \tfrac{1}{c}(\partial_\mu j_\nu - \partial_\nu j_\mu) \\
&\downarrow \\
&= 0 \quad \text{in charge-free space: } j_\mu = 0
\end{aligned}$$

Finally, at (371) we obtained four copies of the wave equation by covariant specialization of the gauge.

We will be concerned in these pages with certain particular solutions of the preceding free-field equations that bear on the classical physics of light. Two points should be born in mind:

- All of the equations ennumerated above are satisfied by the Coulomb field of an isolated charge *except at the location of the charge itself*. They are satisfied by the Lorentz transforms of such a field (field of a charge drifting by), by the field of a static population of such charges, by the magnetic field of a current-carrying wire *except at the location of the wire itself*,

[233] For discussion of the "generalized Kronecker deltas" see pages 7–8 in "Electrodynamical applications of the exterior calculus" (1996). The notational resources of the exterior calculus render the following argument—though it looks here a little contrived—entirely and transparently natural. Incidentally, g has recently signified magnetic charge, and before that was the name of a coupling constant: $g \equiv e/\hbar c$. In the following lines g is restored to its original meaning: $g \equiv \det\|g_{\mu\nu}\|$.

by the fields produced by drifting populations of such wires. In none of those situations are the fields detectable by the apparatus of optics (photometers, etc.); none of them present the diffraction/interference phenomena characteristic of wave physics; to each of them the language of optics would appear alien (except quantum mechanically, where one attributes electrostatic interaction to an "exchange of photons"). What we at present lack is a sharp criterion for distinguishing "light-like" from "other" solutions of the free-field equations.

- We will be studying the physics of light-in-the-absence-of-matter, of light *in vacuuo*. But such light is invisible, an inferential abstraction! For it is only by its interaction with matter (production by radiative processes, transmission through media, manipulation by lenses/mirrors/filters and other such devices,detection by eyes/photometers) that we "see" light, that we become aware of its existence as a fact of Nature—reportedly the *first* fact.[234] But before we can construct a theory of the light-matter interaction we must possess a theory of (the electromagnetic properties of) matter . . . and toward that objective—since matter and most production/ absorption processes are profoundly quantum mechanical—classical physics can carry us only a short part of the way (yet far enough to account phenomenologically for most of classical optics).

Nevertheless . . . the ideas to which we will be led are absolutely fundamental to the physics of light, whatever the depth of the physical detail and conceptual sophistication with which we elect to pursue that subject.

The physics of light is in several important (but too seldom remarked) respects "exceptional, surprising." In order to highlight the points at issue, which remain invisible until placed in broader context, I will (as I have several times already) draw occasionally on Proca's theory of "massive light."

1. Fourier decomposition of the wave field. On pages 291 & 292 we encountered several instances of the wave equation

$$\Box\varphi = 0 \qquad i.e., \qquad \big\{\tfrac{1}{C^2}\partial_t^2 - \nabla^2\big\}\varphi(t,\boldsymbol{x}) = 0$$

It is mathematically natural—alien to the spirit of relativity, but an option available to every *particular* inertial observer—to "split off the time variable,"

---

[234] "In the beginning God created the heavens and the earth. The earth was without form, and void, and darkness was on the face of the deep. Then God said, 'Let there be light'; and there was light. And God saw the light, that it was good; and God divided the light from the darkness. . . " (Genesis I: 1–4). For an absorbing account of the philosophical contemplation of relationships among God, Good and Light that, after more than two millennia, had led by the 16$^{\text{th}}$ Century to the conception of physical space—the non-obvious one we now take for granted—that "made physics possible" see Max Jammer's slim masterpiece *Concepts of Space: The History of Theories of Space in Physics* (1954), with forward by Albert Einstein.

writing $\varphi(t,\boldsymbol{x}) = f(t)\cdot\phi(\boldsymbol{x})$. Then

$$\tfrac{1}{c^2}\ddot{f} = -k^2 f \quad\text{and}\quad (\nabla^2 + k^2)\phi = 0$$

where k^2 is a positive separation constant, with the physical dimension of $(\text{length})^{-2}$. We are led thus to solutions of the *monochromatically oscillatory* form

$$\varphi_\omega(t,\boldsymbol{x}) = e^{i\omega t}\cdot\phi_\omega(\boldsymbol{x}) \quad\text{with}\quad \omega \equiv kc$$

where ω can assume any (positive or negative) real value.[235]

In Cartesian coordinates the

$$\text{HELMHOLTZ EQUATION :}\qquad (\nabla^2 + k^2)\phi = 0$$

reads

$$\Big\{(\tfrac{\partial}{\partial x})^2 + (\tfrac{\partial}{\partial y})^2 + (\tfrac{\partial}{\partial z})^2 + k^2\Big\}\phi(x,y,z) = 0$$

The separation of variables technique can be carried to completion, and yields solutions of the form

$$\phi(x,y,z) = (\text{constant})\cdot e^{i\,k_1 x}\cdot e^{i\,k_2 y}\cdot e^{i\,k_3 z}$$

with $k_1^2 + k_2^2 + k_3^2 = k^2$.[236] But it has been known since 1934 that separation can be carried to completion in a total of *eleven* coordinate systems; namely,

1. Cartesian (or rectangular) coordinates
2. Circular-cylinder (or polar) coordinates
3. Elliptic-cylinder coordinates
4. Parabolic-cylinder coordinates
5. Spherical coordinates
6. Prolate spheroidal coordinates
7. Oblate spheroidal coordinates
8. Parabolic coordinates
9. Conical coordinates
10. Ellipsoidal coordinates
11. Paraboloidal coordinates

---

[235] We make casual use here and henceforth of the familiar “complex variable trick,” with the understanding that one has direct *physical* interest only in the real/imaginary parts of φ_ω.

[236] Separation of three variables brings only two separation constants into play. Why, therefore, do we appear in the present instance to encounter three? By notational illusion. Look upon (say) k_2 and k_3 as separation constants, and regard $k_1 \equiv \sqrt{k^2 - k_2^2 - k_3^2}$ as an enforced definition.

so the question arises: Why are all but the first largely absent from literature pertaining to the physics of light? Why do theorists in this area so readily capitulate to "Cartesian tyranny." For *several* reasons:

- In non-Cartesian coordinates the description of ∇^2 becomes complicated, so separation of the Helmholtz equation leads to a system of three typically fairly complicated ordinary differential equations, the solutions of which are typically "higher functions" (Bessel functions, Legendre functions, Mathieu functions, etc.).[237] For example (looking only to the simplest case): in circular-cylinder coordinates

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \\ z &= z \end{aligned}$$

the Helmholtz equation becomes

$$\Big\{\big(\tfrac{\partial}{\partial r}\big)^2 + \tfrac{1}{r}\tfrac{\partial}{\partial r} + \tfrac{1}{r^2}\big(\tfrac{\partial}{\partial\theta}\big)^2 + \big(\tfrac{\partial}{\partial z}\big)^2 + k^2\Big\}\phi = 0$$

We write $\phi = R(r)\cdot\Theta(\theta)\cdot Z(z)$ and obtain

$$\left.\begin{aligned} \tfrac{d^2R}{dr^2} + \tfrac{1}{r}\tfrac{dR}{dr} - \big(\tfrac{\alpha}{r^2}+\beta\big)R &= 0 \\ \tfrac{d^2\Theta}{d\theta^2} + \alpha\Theta &= 0 \\ \tfrac{d^2Z}{dz^2} + (k^2+\beta)Z &= 0 \end{aligned}\right\} \quad \alpha \text{ and } \beta \text{ are separation constants}$$

The second equation gives

$$\Theta(\theta) = a_2\sin\sqrt{\alpha}\,\theta + b_2\cos\sqrt{\alpha}\,\theta$$

which by a single-valuedness requirement enforces

$$\sqrt{\alpha} = n \quad : \quad 0, \pm1, \pm2, \ldots$$

The third equation (no single-valuedness requirement is here in force, since z is not a periodic variable) gives

$$Z(z) = a_3\sin\sqrt{k^2+\beta}\,z + b_3\cos\sqrt{k^2+\beta}\,z$$

For the first equation *Mathematica* supplies

$$R(r) = a_1\texttt{BesselI[}n, r\sqrt{\beta}\texttt{]} + b_1\texttt{BesselI[}-n, r\sqrt{\beta}\texttt{]}$$

[237] Details are spelled out in various mathematical handbooks, of which my favorite in this connection is P. Moon & D. E. Spencer, *Field Theory Handbook* (1961).

- All the coordinate systems listed—with the sole exception of the Cartesian coordinate system(s)—possess *singularities* (recall the behavior of the circular-cylinder and spherical coordinate systems on the z-axis).
- Description of the translations/rotations/Lorentz transformations of physical interest is *awkward except in Cartesian coordinates*. Notice in particular that

$$\begin{aligned}\phi &= (\text{constant})\cdot e^{i\omega t}\cdot e^{i\,k_1 x}\cdot e^{i\,k_2 y}\cdot e^{i\,k_3 z}\\ &= (\text{constant})\cdot e^{i(k_0x^0+k_1x^1+k_2x^2+k_3x^3)}\quad\text{with}\quad k_0\equiv\omega/c\\ &= (\text{constant})\cdot e^{ikx}\end{aligned}$$

where $kx \equiv k_\alpha x^\alpha$ becomes Lorentz invariant if we stipulate that

$$k\equiv\begin{pmatrix}k^0\equiv\omega/c\\ k^1\\ k^2\\ k^3\end{pmatrix}\equiv\begin{pmatrix}k^0\\ \boldsymbol{k}\end{pmatrix}\quad\text{transforms as a covariant 4-vector}$$

Notice also that

$$\begin{aligned}\Box e^{ikx} &= i^2 g^{\alpha\beta}k_\alpha k_\beta\, e^{ikx}\\ &= 0\quad\text{if and only if } k \text{ is null: } k_\alpha k^\alpha = 0\end{aligned}$$

It is impossible to argue so neatly in non-Cartesian coordinates.
- In Cartesian coordinates—uniquely—we *gain direct access to the powerful techniques of Fourier transform theory*. . . for by superposition of the plane waves just described we obtain

$$\begin{aligned}\phi(x) &= \tfrac{1}{(2\pi)^2}\iiiint a(k)\delta(k_\alpha k^\alpha - 0)e^{ikx}\,dk^0dk^1dk^2dk^3\\ &= \text{Fourier transform of } a(k)\delta(k_\alpha k^\alpha - 0)\end{aligned}$$

- Last but most important: When we write (say) $\Box F^{\mu\nu} = 0$ we have interest not in independent ${}^{\mu\nu}$-indexed solutions of the wave equation, but in solutions so interrelated that they satisfy the ${}^{\nu}$-indexed side-conditions $\partial_\mu F^{\mu\nu} = 0$ and $\partial_\alpha \epsilon^{\alpha\rho\sigma\nu}F_{\rho\sigma} = 0$. Similarly, when we write $\Box A^\mu = 0$ we have interest not in independent ${}^{\mu}$-indexed solutions of the wave equation, but in solutions so interrelated that they satisfy the side-condition $\partial_\mu A^\mu = 0$. *Implications of the side conditions are far easier to work out in Cartesian coordinates than in any other coordinate system.*

So we yield uncomplainingly to "Cartesian tyranny," and expect soon to see concrete evidence of the advantages of doing so.

One further point merits preparatory comment. If solutions

$$\phi_n(x) = a_n e^{ik_n x}$$

of the wave equation are required to satisfy linear side conditions

$$\sum_n \phi_n(x) = 0$$

then pretty clearly it is essential that $k_1 = k_2 = \ldots$; *i.e.*, that they *buzz in synchrony*.

Look now to these plane wave solutions

$$\begin{aligned}
E_1(x) &= E_1 \cdot e^{ik_1 x}\\
E_2(x) &= E_2 \cdot e^{ik_2 x}\\
E_3(x) &= E_3 \cdot e^{ik_3 x}\\
B_1(x) &= B_1 \cdot e^{ik_4 x}\\
B_2(x) &= B_2 \cdot e^{ik_5 x}\\
B_3(x) &= B_3 \cdot e^{ik_6 x}
\end{aligned}$$

$\uparrow$ constants

upon Maxwell's equations (65) impose what amount to a set of eight linear side conditions, which there is no hope of satisfying unless the components of $\boldsymbol{E}$ and $\boldsymbol{B}$ "buzz in synchrony":

$$k_{1\alpha} = k_{2\alpha} = k_{3\alpha} = k_{4\alpha} = k_{5\alpha} = k_{6\alpha}$$

So we adopt this sharpened hypothesis:

$$\left.\begin{aligned}
\boldsymbol{E}(x) &= \boldsymbol{E}\cdot e^{ikx} = \boldsymbol{E}\cdot \exp\big\{i(\omega t - \boldsymbol{k}\cdot\boldsymbol{x})\big\}\\
\boldsymbol{B}(x) &= \boldsymbol{B}\cdot e^{ikx} = \boldsymbol{B}\cdot \exp\big\{i(\omega t - \boldsymbol{k}\cdot\boldsymbol{x})\big\}
\end{aligned}\right\} \tag{393}$$

Maxwell's equations (65) now become a set of conditions

$$\begin{aligned}
\boldsymbol{k}\cdot\boldsymbol{E} &= 0\\
\boldsymbol{k}\times\boldsymbol{B} + \tfrac{\omega}{c}\boldsymbol{E} &= \mathbf{0}\\[4pt]
\boldsymbol{k}\cdot\boldsymbol{B} &= 0\\
\boldsymbol{k}\times\boldsymbol{E} - \tfrac{\omega}{c}\boldsymbol{B} &= \mathbf{0}
\end{aligned}$$

that serve to constrain the relationships among $\boldsymbol{E}$, $\boldsymbol{B}$ and the propagation vector $\boldsymbol{k}$. The 1$^{\text{st}}$ and 3$^{\text{rd}}$ conditions tell us that

$\boldsymbol{E}$ and $\boldsymbol{B}$ lie necessarily in the plane normal to $\boldsymbol{k}$

Crossing $\boldsymbol{k}$ into the 2$^{\text{nd}}$ equation gives

$$\tfrac{\omega}{c}\boldsymbol{k}\times\boldsymbol{E} + \underbrace{\boldsymbol{k}\times(\boldsymbol{k}\times\boldsymbol{B})}_{} = \mathbf{0}$$
$$= (\boldsymbol{k}\cdot\boldsymbol{B})\boldsymbol{k} - (\boldsymbol{k}\cdot\boldsymbol{k})\boldsymbol{B} = \mathbf{0} - \big(\tfrac{\omega}{c}\big)^2\boldsymbol{B}$$

which is redundant with the 4$^{\text{th}}$ equation. Dotting $\boldsymbol{E}$ into the 4$^{\text{th}}$ equation we discover that

$\boldsymbol{E}$ and $\boldsymbol{B}$ are normal to each other

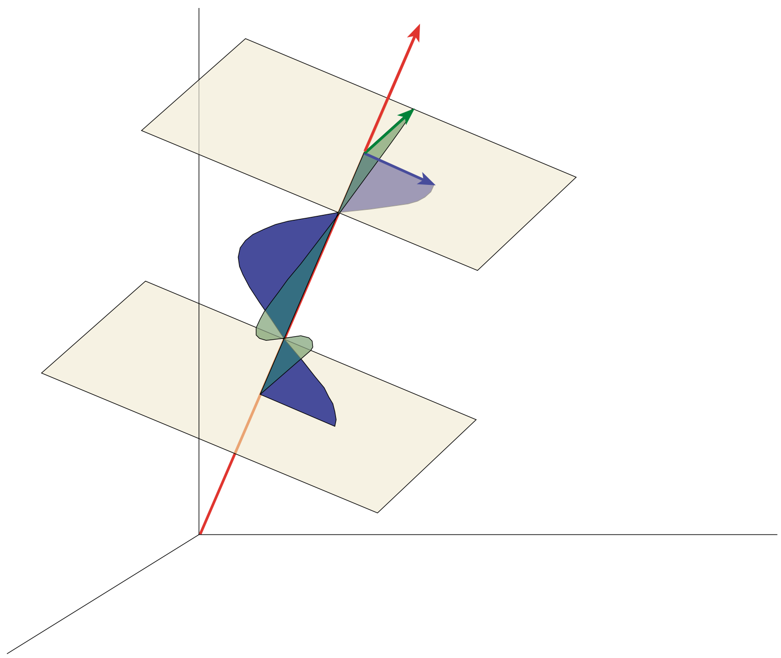


FIGURE 92: *Snapshot of a monochromatic electromagnetic plane wave. Normal to all planes-of-constant-phase (two are shown) is the "propagation or wave vector"* $\boldsymbol{k}$. *The blue sinusoid represents the* $\boldsymbol{E}$*-vector. Normal to it (and of the same amplitude and phase) is the green* $\boldsymbol{B}$*-vector. In animation the electric/magnetic waves would be seen to slide rigidly along* $\boldsymbol{k}$ *with phase speed* c.

Finally, dot the 4$^{\text{th}}$ equation into itself to obtain

$$\begin{aligned}\big(\tfrac{\omega}{c}\big)^2\boldsymbol{B}\cdot\boldsymbol{B} &= \underbrace{(\boldsymbol{k}\times\boldsymbol{E})\cdot(\boldsymbol{k}\times\boldsymbol{E})}\\ &= (\boldsymbol{k}\cdot\boldsymbol{k})(\boldsymbol{E}\cdot\boldsymbol{E}) - (\boldsymbol{k}\cdot\boldsymbol{E})^2 = \big(\tfrac{\omega}{c}\big)^2\boldsymbol{E}\cdot\boldsymbol{E} - 0\end{aligned}$$

$\boldsymbol{E}$ and $\boldsymbol{B}$ are of equal magnitude

It now follows that if $\boldsymbol{k}$ and $\boldsymbol{E}$ are given/known, then $\boldsymbol{B}$ can be computed from

$$\boldsymbol{B} = \hat{\boldsymbol{k}}\times\boldsymbol{E} \tag{394}$$

We saw already on page 264 that

$$\text{phase} = \boldsymbol{k}\cdot\boldsymbol{x} - \omega t$$

is constant on planes $\perp\ \boldsymbol{k}$ that slide along with

$$\text{phase speed } \omega/k = c$$

so are led to the image of an *electromagnetic plane wave* shown in Figure 92.

The vector $\boldsymbol{E}$ can be inscribed in two linearly independent ways on the phase plane. With that fact in mind . . .

- go to some arbitrary "inspection point,"
- face into the onrushing plane wave,
- inscribe an arbitrarily unit vector $\boldsymbol{e}_1$ on the phase plane,
- construct $\boldsymbol{e}_2 \equiv \hat{\boldsymbol{k}} \times \boldsymbol{e}_1$, a unit vector $\perp \boldsymbol{e}_1$.

The "flying $\boldsymbol{E}$-vector" can by these conventions be described

$$\boldsymbol{E}(t) = E_1\,\boldsymbol{e}_1 + E_2\,\boldsymbol{e}_2 = \begin{pmatrix} E_1(t) \\ E_2(t) \end{pmatrix} \tag{395.1}$$

with

$$\left.\begin{aligned} E_1(t) &= \mathcal{E}_1 \cos(\omega t + \delta_1) \\ E_2(t) &= \mathcal{E}_2 \cos(\omega t + \delta_2) \end{aligned}\right\} \tag{395.2}$$

Equations (395) will provide the point of departure for the main work of this chapter.

Suppose we had elected to work in the language of potential theory; *i.e.*, from[238]

$$A^\mu(x) = A^\mu \cdot e^{ik\,x}$$
$$\uparrow\!\!\text{—constant 4-vector}$$

where

$$\Box A^\mu(x) = 0 \quad \text{requires } k^\mu \text{ to be null: } k_\mu k^\mu = 0$$
$$\text{The Lorentz gauge condition } \partial_\mu A^\mu = 0 \text{ requires } k_\mu A^\mu = 0$$

Borrowing notation from pages 296 and 259

$$\|k^\mu\| = \begin{pmatrix} k^0 \\ \boldsymbol{k} \end{pmatrix} \quad \text{with} \quad k^0 \equiv \sqrt{\boldsymbol{k}\cdot\boldsymbol{k}} = \omega/c$$
$$\|A^\mu\| = \begin{pmatrix} \varphi \\ \boldsymbol{A} \end{pmatrix}$$

we find that

$$k_\mu A^\mu = 0 \quad \Longleftrightarrow \quad \varphi = \hat{\boldsymbol{k}}\cdot\boldsymbol{A}$$

so our potential plane wave can be described

$$A_\mu(x) = A_\mu \cdot e^{ik\,x} \quad \text{with} \quad \|A_\mu\| = \begin{pmatrix} \hat{\boldsymbol{k}}\cdot\boldsymbol{A} \\ -\boldsymbol{A} \end{pmatrix}$$

[238] See again page 268. We employ the "complex variable trick" to simplify the writing: extract the real part to obtain the physics.

This we use to obtain

$$\begin{aligned}
E_1 &= F_{01} = \partial_0 A_1 - \partial_1 A_0 = i(k_0 A_1 - k_1 A_0)\cdot e^{ikx}\\
E_2 &= F_{02} = \partial_0 A_2 - \partial_2 A_0 = i(k_0 A_2 - k_2 A_0)\cdot e^{ikx}\\
E_3 &= F_{03} = \partial_0 A_3 - \partial_3 A_0 = i(k_0 A_3 - k_3 A_0)\cdot e^{ikx}\\
B_1 &= F_{32} = \partial_3 A_2 - \partial_2 A_3 = i(k_3 A_2 - k_2 A_3)\cdot e^{ikx}\\
B_2 &= F_{13} = \partial_1 A_3 - \partial_3 A_1 = i(k_1 A_3 - k_3 A_1)\cdot e^{ikx}\\
B_3 &= F_{21} = \partial_2 A_1 - \partial_1 A_2 = i(k_2 A_1 - k_1 A_2)\cdot e^{ikx}
\end{aligned}$$

whence

$$\begin{aligned}
\boldsymbol{E} &= -(\omega/c)\big[\boldsymbol{A} - (\hat{\boldsymbol{k}}\cdot\boldsymbol{A})\hat{\boldsymbol{k}}\big]\cdot ie^{ikx}\\
&= -(\omega/c)\boldsymbol{A}_\perp\cdot ie^{ikx} \qquad (396.1)\\
\boldsymbol{B} &= -(\omega/c)\big[\hat{\boldsymbol{k}}\times\boldsymbol{A}\big]\cdot ie^{ikx}\\
&= \hat{\boldsymbol{k}}\times\boldsymbol{E} \qquad (396.2)
\end{aligned}$$

Notice that

- there are two linearly independent ways to inscribe $\boldsymbol{A}_\perp$ on the plane normal to $\boldsymbol{k}$
- $\boldsymbol{A}_\parallel$ makes no contribution to $\boldsymbol{E}$ or $\boldsymbol{B}$, no contribution therefore to the physics . . . so can be discarded, the reason being that
- $\boldsymbol{A}_\parallel$ can be very simply *gauged away*: take $\chi = e^{ikx}$ and notice that

$$\partial^\mu\chi = ik^\mu\chi \text{ is parallel to } k^\mu$$

Moreover

$$\partial_\mu(\partial^\mu\chi) = -(k_\mu k^\mu)\chi = 0 \text{ because } k^\mu \text{ is null}$$

so such a gauge transformation respects the Lorentz gauge condition.

The argument just completed has led us back again—but rather more swiftly/luminously—to precisely the physical results obtained earlier by other means.

It is instructive to consider how electromagnetic plane wave physics would be altered "if the photon had mass." According to Proca,[239] we would have interest then the plane wave solutions

$$A_\mu(x) = A_\mu\cdot e^{ikx}$$

of

$$(\Box + \varkappa^2)A_\mu = 0 \quad\text{and}\quad \partial_\mu A^\mu = 0$$

[239] We borrow here from §5 in Chapter 4, but use A^μ rather than U^μ to denote the "massive vector Proca field."

The first condition supplies $k_\mu k^\mu = \varkappa^2$ or $k^0 = \sqrt{\boldsymbol{k}\cdot\boldsymbol{k} + \varkappa^2}$, while the second condition supplies $A^0 = (\boldsymbol{k}\cdot\boldsymbol{A})/k^0$. The argument that led to (396) now leads to

$$\begin{aligned}
\boldsymbol{E} &= -\big[k_0\boldsymbol{A} - A_0\boldsymbol{k}\big]\cdot ie^{ik\,x}\\
&= -k_0\bigg[\boldsymbol{A} - \frac{(\boldsymbol{k}\cdot\boldsymbol{A})\boldsymbol{k}}{k_0^2}\bigg]\cdot ie^{ik\,x}\\
&= -k_0\big[\boldsymbol{A} - \wp^2(\hat{\boldsymbol{k}}\cdot\boldsymbol{A})\hat{\boldsymbol{k}}\big]\cdot ie^{ik\,x} \qquad (397.1)\\
\boldsymbol{B} &= -k\big[\hat{\boldsymbol{k}}\times\boldsymbol{A}\big]\cdot ie^{ik\,x}\\
&= \wp\cdot(\hat{\boldsymbol{k}}\times\boldsymbol{E}) \qquad (397.2)
\end{aligned}$$

where what I call the "Proca factor"

$$\begin{aligned}
\wp \equiv k/k_0 &= \frac{\sqrt{\boldsymbol{k}\cdot\boldsymbol{k}}}{\sqrt{\boldsymbol{k}\cdot\boldsymbol{k}+\varkappa^2}} \quad \text{with}\begin{cases} k \;\equiv \sqrt{\boldsymbol{k}\cdot\boldsymbol{k}}\\ k_0 = \omega/c\end{cases}\\
&\downarrow\\
&= 1 \quad \text{in the Maxwellian limit } \varkappa^2\downarrow 0
\end{aligned}$$

We distinguish two cases:

CASE $\boldsymbol{A}\perp\boldsymbol{k}$ This can happen in two ways. Because $\hat{\boldsymbol{k}}\cdot\boldsymbol{A} = 0$ we have

$$\begin{aligned}
\boldsymbol{E} &= -(\omega/c)\boldsymbol{A}_\perp\cdot ie^{ik\,x}\\
\boldsymbol{B} &= \wp\cdot(\hat{\boldsymbol{k}}\times\boldsymbol{E})
\end{aligned}$$

which differs from (396) only in the presence of the $\wp$-factor, which diminishes the strength of the $\boldsymbol{B}$-field.

CASE $\boldsymbol{A}\parallel\boldsymbol{k}$ Writing $\boldsymbol{A} = A_\parallel\hat{\boldsymbol{k}}$ we have

$$\begin{aligned}
\boldsymbol{E} &= -(\omega/c)(1-\wp^2)A_\parallel\hat{\boldsymbol{k}}\cdot ie^{ik\,x}\\
\boldsymbol{B} &= \wp\cdot(\hat{\boldsymbol{k}}\times\boldsymbol{E})\\
&= \boldsymbol{0} \quad \text{because } \boldsymbol{E}\parallel\hat{\boldsymbol{k}}
\end{aligned}$$

The *electric field has acquired an oscillatory longitudinal component which possesses no magnetic counterpart*, and both longitudinal fields vanish in the Maxwellian limit.

2. Stokes parameters. The "flying $\boldsymbol{E}$-vector" of (395) traces/retraces the simplest of Lissajous figures—an ellipse—on the (E_1, E_2)-plane. The flight of $\boldsymbol{E}(t)$ is, at optical frequencies ($\omega \sim 10^{15}$ Hz), much too brisk to be observed, but the *figure* of the ellipse (size, shape, orientation) and the ↺ / ↻ sense in which it is pursued *are* observable—detectable by the "slow" devices of classical optics (eyes, photometers, filters of various types). They give rise to the phenomenology of **optical polarization**, the theory of which will concern us in this and the next few sections.

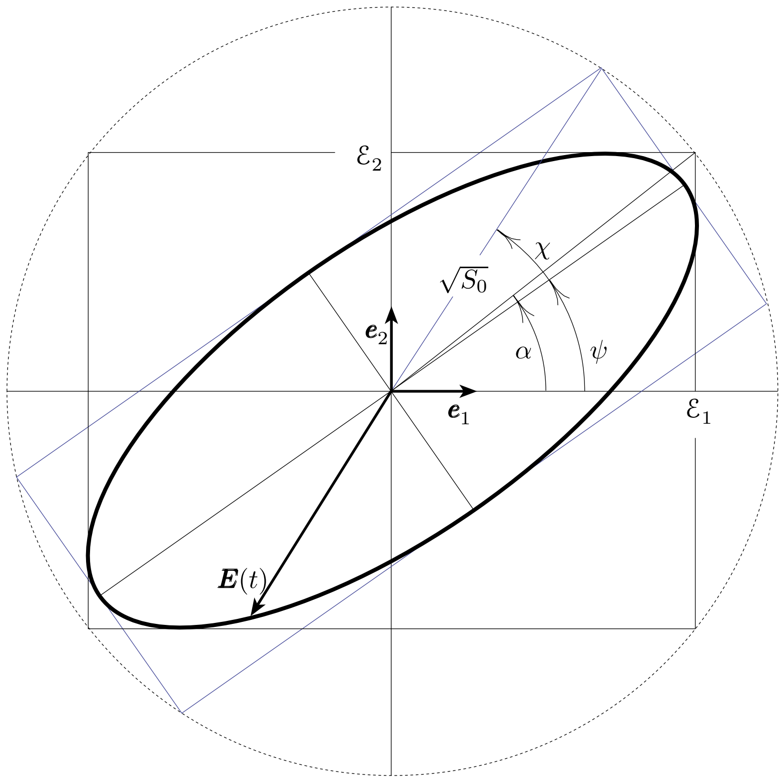


FIGURE 93: *Ellipse traced by the* $\boldsymbol{E}$*-vector of an electromagnetic plane wave, with* $\boldsymbol{k}$ *up out of the page. It is a remarkable property of ellipses that all circumscribing rectangles (two are shown) have the same diagonal measure, which can be taken to set the* size *of the ellipse. The angle* ψ *describes the* orientation *of the principal rectangle, which is of long dimension* $2a$*, short dimension* $2b$*. The* shape *of the ellipse is usually described in terms of the*

$$\text{ellipticity} \equiv \sqrt{1-(b/a)^2}$$

but—as Stokes appreciated—is equally well described by

$$\chi \equiv \arctan(b/a)$$

Helicity information is absent from (398)*, but from* (395.2) *we discover—look to* $\frac{d}{dt}\boldsymbol{E}(t)$ *at conveniently chosen points, or argue that if* $E_2(t)$ *leads* $E_1(t)$ *(i.e., if* $\delta_2 > \delta_1$*) the circulation is clockwise, and in the contrary case counterclockwise—that the circulation is* ↻ *or* ↺ *according as* $0 < \delta \equiv \delta_2 - \delta_1 < \pi$ *or* $-\pi < \delta < 0$.

Eliminating t between equations (395.2) we obtain[240]

$$\mathcal{E}_2^2\cdot E_1^2 - 2\mathcal{E}_1\mathcal{E}_2\cos\delta\cdot E_1E_2 + \mathcal{E}_1^2\cdot E_2^2 = \mathcal{E}_1^2\mathcal{E}_2^2\sin^2\delta \tag{398}$$
$$\delta\equiv\delta_2-\delta_1\equiv\text{phase difference}$$

Equations (395.2) provide a parametric description, and (398) an implicit description ... of the ellipse[241] shown in Figure 93. Some elementary analytical geometry—the details are fun but uninformative, and (since they have nothing specifically to do with electrodynamics) will be omitted—leads to the following conclusions:

$$S_0 = \mathcal{E}_1^2+\mathcal{E}_2^2$$
$$\sin 2\chi = \sin 2\alpha\cdot\sin\delta = \frac{2\mathcal{E}_1\mathcal{E}_2\sin\delta}{\mathcal{E}_1^2+\mathcal{E}_2^2}\equiv\frac{S_3}{S_0}$$
$$\tan 2\psi = \tan 2\alpha\cdot\cos\delta = \frac{2\mathcal{E}_1\mathcal{E}_2\cos\delta}{\mathcal{E}_1^2-\mathcal{E}_2^2}\equiv\frac{S_2}{S_1}$$

where

$$S_1\equiv\mathcal{E}_1^2-\mathcal{E}_2^2$$

Notice that helicity—which was observed above to be controlled by the sign of δ—could as well be said (since $\mathcal{E}_1$ and $\mathcal{E}_2$ are non-negative) to be controlled by the sign of χ, and that (as is clear from the figure) χ ranges on the restricted interval $\left[-\frac{\pi}{2},+\frac{\pi}{2}\right]$. Recasting and extending the results summarized above, we have

$$\left.\begin{aligned} S_0 &= \mathcal{E}_1^2+\mathcal{E}_2^2 \\ S_1 &= \mathcal{E}_1^2-\mathcal{E}_2^2 = S_0\cos 2\chi\cos 2\psi \\ S_2 &= 2\mathcal{E}_1\mathcal{E}_2\cos\delta = S_0\cos 2\chi\sin 2\psi \\ S_3 &= 2\mathcal{E}_1\mathcal{E}_2\sin\delta = S_0\sin 2\chi \end{aligned}\right\} \tag{399}$$

These equations define the so-called **Stokes parameters**, which were introduced by G. G. Stokes in 1852 to facilitate the discussion of some experimental results. There is reason to think that Stokes himself was unaware of the extraordinary power of his creation ... which took nearly a century, and the work of many hands, to be revealed. Today his lovely idea is recognized to be central to every classical/statistical/quantum account of the phenomenology of optical polarization.

It is evident that

$$S_1^2+S_2^2+S_3^2=S_0^2 \tag{400}$$

and that *rotational sense (helicity) can be read from the sign of* S_3.

---

[240] PROBLEM 61.
[241] PROBLEM 62.

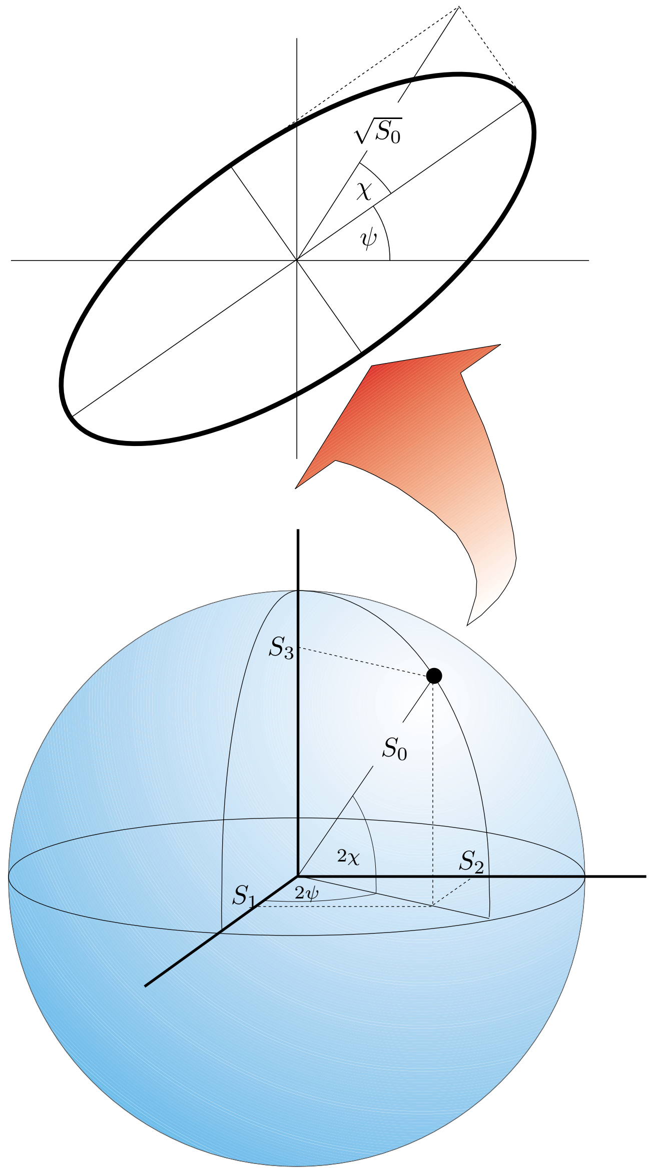
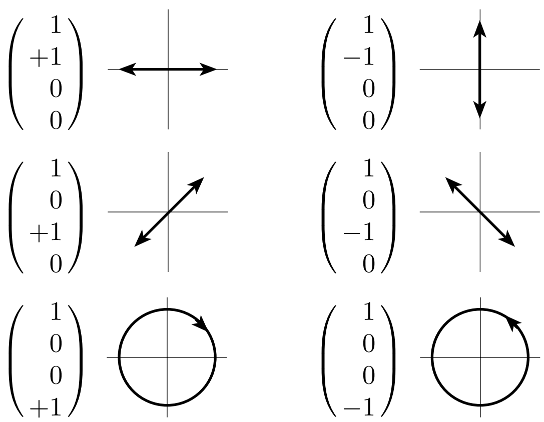


FIGURE 94: *Equations* (399) *serve to associate points on the Stokes sphere of radius S_0 with centered ellipses of fixed size and all possible figures & orientations. Points in the northern hemisphere* ($S_3 > 0$) *are assigned ↻ helicity, points in the southern hemisphere are assigned ↺ helicity. In the case $S_0 = 1$ the Stokes sphere becomes the Poincaré sphere.*

Henri Poincaré (1892) observed that, in view of the structure second stack of equalities in (399), it is natural to *place the polarizational states of electromagnetic plane waves in one-one association with the points* $\boldsymbol{S}$ *that comprise the surface of a sphere of radius* S_0 *in 3-dimensional "Stokes space,"* as indicated in Figure 94. It becomes obvious from the figure that specification of $\{S_0, S_1, S_2, S_3\}$ is equivalent to specification of the intuitively more immediate parameters $\{S_0, \psi, \chi\}$. We need $\boldsymbol{k}$ to describe the *direction of propagation* and *frequency/wavelength* of the monochromatic plane wave, but if we have only "slow detectors" to work with then $\{S_0, S_1, S_2, S_3\}$ summarize <u>all that we can experimentally verify</u> concerning the polarizational state of the wave.[242]

Reading from Figure 94, we find the polarizational states which correspond to (for example) the axial positions on the Poincaré sphere to be those illustrated below:



$$\begin{pmatrix}1\\+1\\0\\0\end{pmatrix}\quad\begin{pmatrix}1\\-1\\0\\0\end{pmatrix}\quad\begin{pmatrix}1\\0\\+1\\0\end{pmatrix}\quad\begin{pmatrix}1\\0\\-1\\0\end{pmatrix}\quad\begin{pmatrix}1\\0\\0\\+1\end{pmatrix}\quad\begin{pmatrix}1\\0\\0\\-1\end{pmatrix}$$

It becomes in this light natural to say (with Stokes) of a pair of plane waves that they are *"oppositely polarized" if and only if their Stokes*

$$\boldsymbol{S} \equiv \begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix} \quad \text{and} \quad \boldsymbol{S} \equiv \begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix}$$

vectors point in diametrically opposite directions:

$$\boldsymbol{S} = -\lambda^2 \boldsymbol{S}$$

of which

$$S_0 = +\lambda^2 S_0$$

[242] It is because they *relate so directly to the observational realities* that Stokes parameters become central to the quantum theory of photon spin. See §2–8 in J. M. Jauch & F. Rohrlich, *The Theory of Photons & Electrons* (1955) where, by the way, I was first introduced to this pretty subject.

is—by (400)—a corollary. From (399) we see that

$$\boldsymbol{S} \longrightarrow \boldsymbol{S} = -\lambda^2 \boldsymbol{S}$$

can, in more physical terms, be described

$$\begin{aligned} \mathcal{E}_1 &\longrightarrow \mathcal{E}_1 = +\lambda\mathcal{E}_2 \\ \mathcal{E}_2 &\longrightarrow \mathcal{E}_2 = -\lambda\mathcal{E}_1 \\ \delta &\longrightarrow \delta = \delta \end{aligned}$$

so the "oppositely polarized" associates of

$$\boldsymbol{E}(t) = \boldsymbol{e}_1\mathcal{E}_1 \cos\omega t + \boldsymbol{e}_2\mathcal{E}_2 \cos(\omega t + \delta)$$

have the form

$$\boldsymbol{E}(t) = \boldsymbol{e}_1\lambda\mathcal{E}_2 \cos(\omega t + \alpha) - \boldsymbol{e}_2\lambda\mathcal{E}_1 \cos(\omega t + \delta + \alpha)$$

where λ and α are arbitrary. As is intuitively evident, as Fresnel (∼1816) demonstrated experimentally,[243] and as we will soon be in position to prove, *oppositely polarized plane waves to not interfere*.

I propose now to make more secure the recent claim[242] that Stokes parameters pertain *directly* to the *observational* properties of plane waves. *Energy flux* is described (see again page 216) by the

$$\text{Poynting vector } \boldsymbol{S}(t) = c(\boldsymbol{E}\times\boldsymbol{B})$$

For a plane wave

$$\boldsymbol{B} = \hat{\boldsymbol{k}}\times\boldsymbol{E}$$

so

$$= cE^2(t)\hat{\boldsymbol{k}}$$

The *magnitude* of the Poynting vector is given therefore by

$$S(t) = cE^2(t) = c\big\{\mathcal{E}_1^2 \cos^2\omega t + \mathcal{E}_2^2 \cos^2(\omega t + \delta)\big\}$$

and the *intensity* of the wave ($S(t)$ averaged over a period τ) by

$$I \equiv \tfrac{1}{\tau}\int_0^\tau S(t)\,dt = \tfrac{1}{2}c\big\{\mathcal{E}_1^2 + \mathcal{E}_2^2\big\}$$

So

$$S_0 \equiv \mathcal{E}_1^2 + \mathcal{E}_2^2 = \tfrac{2}{c}I \tag{401}$$

<u>can be measured directly by a "J-meter,"</u> *i.e.*, by a photometer that has been re-scaled so that it displays

$$J \equiv \tfrac{2}{c}\cdot(\text{intensity})$$

[243] Augustin Jean Fresnel (1788–1827) was an engineer who took up optics while a political exile with time on his hands. It was his study of polarization that led him to propose that light was to be understood in terms of *transverse* waves, not the longitudinal waves postulated by Huygens, Young and others. Practical problems of lighthouse design led him to the invention of the Fresnel lens and to fundamental contributions to theoretical optics (diffraction).

If an arbitrarily polarized wave

$$\boldsymbol{E}_{\text{in}}(t) \;= \boldsymbol{e}_1\mathcal{E}_1\cos\omega t + \boldsymbol{e}_2\mathcal{E}_2\cos(\omega t + \delta)$$

is incident upon a $\longleftrightarrow$ linear polarizer then the exit beam can be described

$$\boldsymbol{E}_{\text{out}}(t) = \boldsymbol{e}_1\mathcal{E}_1\cos\omega t$$

so—arguing from (399)—we have

$$\begin{pmatrix} S_0\\ S_1\\ S_2\\ S_3\end{pmatrix}_{\text{out}} = \begin{pmatrix} \mathcal{E}_1^2\\ \mathcal{E}_1^2\\ 0\\ 0\end{pmatrix} = \begin{pmatrix} \frac{1}{2}(S_0+S_1)\\ \frac{1}{2}(S_0+S_1)\\ 0\\ 0\end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0\\ \frac{1}{2} & \frac{1}{2} & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}\begin{pmatrix} S_0\\ S_1\\ S_2\\ S_3\end{pmatrix}_{\text{in}}$$

We are led thus to these descriptions of *the action of some typical polarizers*:

$$\longleftrightarrow \text{ polarizer}: \begin{pmatrix} S_0\\ S_1\\ S_2\\ S_3\end{pmatrix}_{\text{out}} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0\\ \frac{1}{2} & \frac{1}{2} & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}\begin{pmatrix} S_0\\ S_1\\ S_2\\ S_3\end{pmatrix}_{\text{in}} \tag{402.1}$$

$$\updownarrow \text{ polarizer}: \begin{pmatrix} S_0\\ S_1\\ S_2\\ S_3\end{pmatrix}_{\text{out}} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & 0\\ -\frac{1}{2} & \frac{1}{2} & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}\begin{pmatrix} S_0\\ S_1\\ S_2\\ S_3\end{pmatrix}_{\text{in}} \tag{402.2}$$

$$\swarrow\!\!\!\nearrow \text{ polarizer}: \begin{pmatrix} S_0\\ S_1\\ S_2\\ S_3\end{pmatrix}_{\text{out}} = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0\\ 0 & 0 & 0 & 0\\ \frac{1}{2} & 0 & \frac{1}{2} & 0\\ 0 & 0 & 0 & 0\end{pmatrix}\begin{pmatrix} S_0\\ S_1\\ S_2\\ S_3\end{pmatrix}_{\text{in}} \tag{402.3}$$

$$\nwarrow\!\!\!\searrow \text{ polarizer}: \begin{pmatrix} S_0\\ S_1\\ S_2\\ S_3\end{pmatrix}_{\text{out}} = \begin{pmatrix} \frac{1}{2} & 0 & -\frac{1}{2} & 0\\ 0 & 0 & 0 & 0\\ -\frac{1}{2} & 0 & \frac{1}{2} & 0\\ 0 & 0 & 0 & 0\end{pmatrix}\begin{pmatrix} S_0\\ S_1\\ S_2\\ S_3\end{pmatrix}_{\text{in}} \tag{402.4}$$

$$\circlearrowright \text{ polarizer}: \begin{pmatrix} S_0\\ S_1\\ S_2\\ S_3\end{pmatrix}_{\text{out}} = \begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2}\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ \frac{1}{2} & 0 & 0 & \frac{1}{2}\end{pmatrix}\begin{pmatrix} S_0\\ S_1\\ S_2\\ S_3\end{pmatrix}_{\text{in}} \tag{402.5}$$

$$\circlearrowleft \text{ polarizer}: \begin{pmatrix} S_0\\ S_1\\ S_2\\ S_3\end{pmatrix}_{\text{out}} = \begin{pmatrix} \frac{1}{2} & 0 & 0 & -\frac{1}{2}\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ -\frac{1}{2} & 0 & 0 & \frac{1}{2}\end{pmatrix}\begin{pmatrix} S_0\\ S_1\\ S_2\\ S_3\end{pmatrix}_{\text{in}} \tag{402.6}$$

Arguing again from (399), we find that the action

$$\begin{aligned} \boldsymbol{E}_{\text{in}}(t) \;&= \boldsymbol{e}_1\mathcal{E}_1\cos\omega t + \boldsymbol{e}_2\mathcal{E}_2\cos(\omega t + \delta)\\ &\downarrow\\ \boldsymbol{E}_{\text{out}}(t) &= \boldsymbol{e}_1 e^{-\alpha}\mathcal{E}_1\cos\omega t + \boldsymbol{e}_2 e^{-\alpha}\mathcal{E}_2\cos(\omega t + \delta)\end{aligned}$$

of a neutral filter can be described

$$\text{neutral filter}:\quad \begin{pmatrix} S_0\\ S_1\\ S_2\\ S_3\end{pmatrix}_{\text{out}} = e^{-2\alpha}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}\begin{pmatrix} S_0\\ S_1\\ S_2\\ S_3\end{pmatrix}_{\text{in}} \tag{402.7}$$

Suppose, now, that we present a plane wave serially to

0) a neutral filter F_0 with $e^{-2\alpha} = \frac{1}{2}$,
1) a $\leftrightarrow$ polarizer F_1,
2) a $\swarrow\!\!\!\nearrow$ polarizer F_2,
3) a $\circlearrowright$ polarizer F_3

and in each case use a J-meter to measure the intensity of the output, obtaining

$$\big[S_0\big]_{\text{out}} = \begin{cases} J_0 = \frac{1}{2}S_0 & \text{when } \mathsf{F}_0 \text{ used}\\ J_1 = \frac{1}{2}(S_0+S_1) & \text{when } \mathsf{F}_1 \text{ used}\\ J_2 = \frac{1}{2}(S_0+S_2) & \text{when } \mathsf{F}_2 \text{ used}\\ J_3 = \frac{1}{2}(S_0+S_3) & \text{when } \mathsf{F}_3 \text{ used}\end{cases}$$

Algebraically deconvolving the output data, we obtain

$$\left.\begin{aligned} S_0 &= 2J_0\\ S_1 &= 2J_1 - 2J_0\\ S_2 &= 2J_2 - 2J_0\\ S_3 &= 2J_3 - 2J_0\end{aligned}\right\} \tag{403}$$

Alternative sets of filters would serve as well, but would require some algebraic adjustment at (403). The implication is that

> With four suitably selected filters and a photometer one can *measure* Stokes' parameters, and thus fully characterize the intensity/polarization/helicity of a (coherent monochromatic) plane wave.

3. Mueller calculus. A light beam—modeled, for the moment, as a plane wave—with attributes

$$\big\{\boldsymbol{k}, S_0, S_1, S_2, S_3\big\}_{\text{in}}$$

is presented to a passive device, from which a beam with attributes

$$\big\{\boldsymbol{k}, S_0, S_1, S_2, S_3\big\}_{\text{out}}$$

emerges. A description of how the output variables depend upon the input variables would comprise a *characterization of the device*. In view of the fact that

- mirrors/lenses typically *change the direction* of the beam, and scatterers typically spray a beam in multiple directions
- some crystals *change the frequency* of a monochromatic beam
- some materials/devices *alter the coherence properties* of an incident beam, others alter the *degree of polarization* (of which more later)

we recognize that some physical restriction is involved when agree to limit our concern to devices that conform to the following scheme:

$$\begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix}_{\text{in}} \longrightarrow \boxed{\quad\text{device}\quad} \longrightarrow \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix}_{\text{out}}$$

Since (400) pertains *generally* to monochromatic plane waves, we see that for *every* such device

$$\big[S_0^2 - S_1^2 - S_2^2 - S_3^2\big]_{\text{out}} = \big[S_0^2 - S_1^2 - S_2^2 - S_3^2\big]_{\text{in}} = 0 \tag{403.1}$$

while for every *passive* device (since passive devices are—unlike lasers—not connected to an external energy source, and therefore may absorb energy from, but cannot inject energy into... the transmitted light beam) energy conservation requires

$$0 \leqslant \big[S_0\big]_{\text{out}} \leqslant \big[S_0\big]_{\text{in}} \tag{403.2}$$

A <u>general theory of passive devices</u> would result from an effort to describe the functional relationships

$$S_{\mu\,\text{out}} = \mathcal{D}_\mu(S_{0\,\text{in}}, S_{1\,\text{in}}, S_{2\,\text{in}}, S_{3\,\text{in}}) \quad : \quad \mu = 0, 1, 2, 3$$

permitted by (403). Remarkably, such an effort, if based upon (403.1) alone, would lead back again to the *conformal group*, which was encountered earlier in quite another connection.[244] When (403.2) is brought into play certain group elements are excluded: one is left with what might be called the "device semigroup."[245]

A far simpler theory—which is, however, adequate to most practical needs —is obtained if one imposes the additional assumption that the parameters $S_{\mu\,\text{out}}$ are *linear* functions of $S_{\mu\,\text{in}}$:

$$\begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix}_{\text{in}} \longrightarrow \boxed{\text{linear passive device}} \longrightarrow \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix}_{\text{out}} = \mathbb{M} \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix}_{\text{in}}$$

One is led then to the linear fragment of the conformal group; *i.e.*, to the condition (compare (185.2) on pages 129 & 164)

$$\mathbb{M}^{\mathsf{T}} g\, \mathbb{M} = m^2 g \tag{404.1}$$

subject to the proviso that one must exclude cases that place one in violation of (403.2). Evidently $\det \mathbb{M} = m^4$, so in non-singular cases one can state that $\mathbb{M}/m$ is Lorentzian:

$$\mathbb{M} \text{ (if non-singular) possesses the structure } \mathbb{M} = m \cdot \mathbb{\Lambda} \tag{404.2}$$

---

[244] See again Chapter 2, §6. For a brief sketch of the resulting theory of optical devices see pages 353–354 in CLASSICAL ELECTRODYNAMICS (1980).

[245] A *semigroup* is a "group without inversion."

REMARK: One must carefully resist any temptation to conclude from the design of (404) that the Stokes parameters S_μ transform as the components of a 4-vector. Their Lorentz transformation properties are inherited—*via* the definitions (399)—from those of the electromagnetic field, and are in fact quite intricate. The subject is treated on pages 436 *et seq* in my ELECTRODYNAMICS (1972).

The idea of using 4×4 matrices to describe the action of linear passive optical devices was first developed in a report by Hans Mueller . . . which, however, he never published. Such matrices are called "Mueller matrices," and their use (discussed below) is the subject matter of the "Mueller calculus."

The 4×4 matrices encountered in (402.1–6) are readily shown to satisfy

$$\mathbb{M}^{\mathsf{T}} g\, \mathbb{M} = \mathbb{O}, \text{ which is (404.1) with } m = 0 \tag{405}$$

and to be always in compliance with (403.2).[246] So each is a Mueller matrix. Each is found, moreover, to possess[247] the "projection property"[248]

$$\mathbb{M}^2 = \mathbb{M} \tag{406}$$

Calculation shows, moreover, that in each case

$$\det(\mathbb{M} - \lambda\mathbb{I}) = \lambda^3(\lambda - 1) \tag{407}$$

so

$\mathbb{M}\,S_{\text{in}} = 0$	has three linearly independent solutions; the device extinguishes such beams
$\mathbb{M}\,S_{\text{in}} = S_{\text{in}}$	has but one; the device is transparent to such beams (scalar multiples of one another)

EXAMPLE: Noting that $3^2 + 4^2 + 12^2 = 13^2$ let us, by contrivance, take

$$S_{\text{in}} = \begin{pmatrix} 13 \\ 3 \\ 4 \\ 12 \end{pmatrix}$$

and let us take $\mathbb{M}$ to be the Mueller matrix of (402.1) that describes the action of a $\longleftrightarrow$ polarizer. Then (by quick calculation)

---

[246] PROBLEM 63.

[247] PROBLEM 64.

[248] From (406) it follows, by the way, that $(\det\mathbb{M})^2 = \det\mathbb{M}$ whence

$$\det\mathbb{M} = \begin{cases} 1 & \text{if } \mathbb{M} \text{ is the trivial projector } \mathbb{I} \\ 0 & \text{otherwise} \end{cases}$$

The zero on the right side of (405) can be therefore be looked upon as a forced consequence of projectivity.

$$S_{\text{out}} = \mathbb{M} S_{\text{in}} = \begin{pmatrix} 8 \\ 8 \\ 0 \\ 0 \end{pmatrix}, \text{ projected component of } \begin{pmatrix} 13 \\ 3 \\ 4 \\ 12 \end{pmatrix}$$

The exit beam is 100% $\longleftrightarrow$ polarized, but dimmer:

$$S_{0\text{out}} = 8 \ < \ S_{0\text{in}} = 13$$

A second pass through the device (second such projection) has no effect (that being the upshot of $\mathbb{M}^2 = \mathbb{M}$):

$$\mathbb{M} \begin{pmatrix} 8 \\ 8 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 8 \\ 8 \\ 0 \\ 0 \end{pmatrix}$$

To describe the action of an *arbitrary polarizer*: let $\boldsymbol{\sigma}$ be an arbitrary unit 3-vector and construct

$$\mathbb{M}(\boldsymbol{\sigma}) \equiv \tfrac{1}{2} \cdot \begin{pmatrix} 1 & \sigma_1 & \sigma_2 & \sigma_3 \\ \sigma_1 & \sigma_1\sigma_1 & \sigma_1\sigma_2 & \sigma_1\sigma_3 \\ \sigma_2 & \sigma_2\sigma_1 & \sigma_2\sigma_2 & \sigma_2\sigma_3 \\ \sigma_3 & \sigma_3\sigma_1 & \sigma_3\sigma_2 & \sigma_3\sigma_3 \end{pmatrix} \tag{408.1}$$

One can show[249] that $\mathbb{M}(\boldsymbol{\sigma})$ satisfies (405/6/7) and that

$$\mathbb{M}(\boldsymbol{\sigma}) \begin{pmatrix} 1 \\ \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix} = \begin{pmatrix} 1 \\ \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix} \tag{408.2}$$

Moreover

$$\mathbb{M}(-\boldsymbol{\sigma})\mathbb{M}(+\boldsymbol{\sigma}) = \mathbb{O} \quad : \quad \text{all } \boldsymbol{\sigma} \tag{409}$$

which supplies neat support for Stokes' claim (page 305) that diametrically opposite points on the Stokes sphere refer to "opposite polarizations," and conforms precisely to the pattern evident when one compares (402.2) with (402.1), (402.4) with (402.3), (402.6) with (402.5). In the case

$$\boldsymbol{\sigma} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Equation (409) might be notated

$$\mathbb{M}(\updownarrow)\mathbb{M}(\leftrightarrow) = \mathbb{O}$$

---

[249] PROBLEM 65.

and interpreted to express the familiar fact that *no light passes through crossed polarizers*.

Suppose, however, we were to interpose (between $\mathbb{M}(\updownarrow)$ and $\mathbb{M}(\leftrightarrow)$) a third device: let it be (say) the linear polarizer represented (see again Figure 94) by

$$\mathbb{M}(\psi) \equiv \mathbb{M}(\boldsymbol{\sigma}) \quad \text{with} \quad \boldsymbol{\sigma} = \begin{pmatrix} \cos 2\psi \\ \sin 2\psi \\ 0 \end{pmatrix}$$

With the assistance of *Mathematica* we compute

$$\mathbb{M}(\updownarrow)\mathbb{M}(\psi)\mathbb{M}(\leftrightarrow) = \begin{pmatrix} \frac{1}{8}\sin^2 2\psi & \frac{1}{8}\sin^2 2\psi & 0 & 0 \\ -\frac{1}{8}\sin^2 2\psi & -\frac{1}{8}\sin^2 2\psi & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \neq \mathbb{O}$$

which illustrates the basis of an experimental technique standard to microscopy and engineering: one places a microscope slide or the stressed Lucite model of a machine part between crossed polarizers, and examines the transmitted image.

The preceding calculation also illustrates the central idea of the "Mueller calculus": To determine the net effect of cascaded optical devices one simply *multiplies the corresponding Mueller matrices*.

"Optical devices" exist in considerable variety. At (402.7) we encountered the Mueller matrices

$$\mathbb{M} = e^{-2\alpha} \cdot \mathbb{I} \tag{410}$$

that describe the action of "neutral filters." Such a device is transparent at $\alpha = 0$, and becomes progressively more absorptive (optically dense) as α increases.

Mueller matrices of major practical importance arise if at (404.2) we set $m = 1$ and assume $\mathbb{M} = \mathbb{\Lambda}$ to have (see again (208) on page 155) the rotational design

$$\mathbb{M} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & \mathbb{R} & \\ 0 & & & \end{pmatrix} \tag{411}$$

$$\mathbb{R} \equiv \exp\left\{ 2\theta \begin{pmatrix} 0 & -\sigma_3 & \sigma_2 \\ \sigma_3 & 0 & -\sigma_1 \\ -\sigma_2 & \sigma_1 & 0 \end{pmatrix} \right\} \;:\; \text{a rotation matrix}$$

Such an $\mathbb{M}$ leaves S_0 invariant (no absorption) but causes

$$\boldsymbol{S} \equiv \begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix}$$

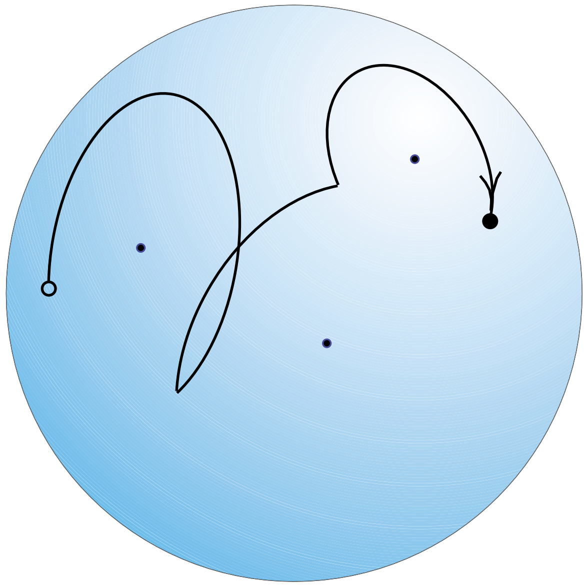


FIGURE 95: *The input beam in Stokes state* ∘ *is passes through three successive devices of type* (411) *to produce an output beam in Stokes state* •. *Dots mark the centers of rotation (ends of the* $\boldsymbol{\sigma}$ *vectors). Because rotations possess the group property, the net effect of the three rotational beam transformations could have been achieved by a* single *such transformation.*

to experience righthanded (↻) rotation through the angle 2θ about the axis defined by the unit vector $\boldsymbol{\sigma}$. In the special case

$$\boldsymbol{\sigma} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

(411) gives

$$\mathbb{M} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\theta & -\sin 2\theta & 0 \\ 0 & \sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

the action of which (see again Figure 94) is to rotate the plane of polarization:

$$\psi \rightarrow \psi + \theta$$

Such devices exploit the *optical activity* phenomenon, and are called "rotators." The case

$$\boldsymbol{\sigma} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

gives

$$\mathbb{M} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos 2\theta & -\sin 2\theta \\ 0 & 0 & \sin 2\theta & \cos 2\theta \end{pmatrix}$$

which achieves

$$\delta \to \delta + 2\theta$$

Such devices are called “compensators” or “phase shifters.” It is clear that Mueller matrices of type (411) are non-singular: $\mathbb{M}^{-1}$ is again a Mueller matrix, which means that the action of such a device could be undone by a suitably chosen second such device. Projection, on the other hand, is a *non*-invertible operation: the action of a polarizer, when undone by subsequent polarizers, always entails attenuation of the beam. To illustrate the point, we return to the example of page 312 and by computation find that

$$\mathbb{M}(\leftrightarrow)\mathbb{M}(\psi)\mathbb{M}(\leftrightarrow) = \tfrac{1}{2}\cos^4\psi \cdot \mathbb{M}(\leftrightarrow)$$

Looking back again to (404.2), it becomes natural in view of the foregoing to assign $\mathbb{\Lambda}$ the “boost” design of (209), writing

$$\mathbb{M} = m\cdot\begin{pmatrix} \gamma & \beta_1\gamma & \beta_2\gamma & \beta_3\gamma \\ \beta_1\gamma & 1{+}(\gamma-1)\beta_1\beta_1/\beta^2 & (\gamma-1)\beta_1\beta_2/\beta^2 & (\gamma-1)\beta_1\beta_3/\beta^2 \\ \beta_2\gamma & (\gamma-1)\beta_2\beta_1/\beta^2 & 1{+}(\gamma-1)\beta_2\beta_2/\beta^2 & (\gamma-1)\beta_2\beta_3/\beta^2 \\ \beta_3\gamma & (\gamma-1)\beta_3\beta_1/\beta^2 & (\gamma-1)\beta_3\beta_2/\beta^2 & 1{+}(\gamma-1)\beta_3\beta_3/\beta^2 \end{pmatrix}$$

where the β’s are “device parameters” that have now *nothing to do with velocity*. Immediately

$$\begin{aligned} S_{0\text{out}} &= m\gamma(S_{0\text{in}} + \boldsymbol{\beta}\cdot\boldsymbol{S}_\text{in}) \\ &\qquad\qquad \boldsymbol{S}_\text{in} = S_{0\text{in}}\hat{\boldsymbol{S}}_\text{in} \text{ by (400)} \\ &= m\gamma(1 + \boldsymbol{\beta}\cdot\hat{\boldsymbol{S}}_\text{in})S_{0\text{in}} \\ &= m\gamma(1 + \beta\cos\omega)S_{0\text{in}} \quad : \quad \omega \text{ is the angle between } \boldsymbol{\beta} \text{ and } \boldsymbol{S}_\text{in} \end{aligned}$$

so to achieve universal compliance with the passivity condition (403.2) we must have

$$0 < m \leqslant \sqrt{\tfrac{1-\beta}{1+\beta}} \leqslant 1$$

where it is understood that $0 \leqslant \beta < 1$. It is not at all difficult to show of such Mueller matrices that though $\mathbb{M}^{-1}$ exists—and is, in fact, easy to describe

$$[m\mathbb{\Lambda}(\boldsymbol{\beta})]^{-1} = m^{-1}\mathbb{\Lambda}(-\boldsymbol{\beta})$$

—it stands in violation of the passivity condition, so cannot be realized by a passive device. On pages 361/2 of some notes already cited[244] I explore some of the finer details of this subject, and argue that *it should be possible* *to mimic 4-dimensional relativity (composition of non-colinear boosts, Thomas precession, etc.) by experiments performed on a linear optical bench*!

In some respects more elegantly efficient—but in other physical respects more limited—than the Mueller calculus is the "Jones calculus," devised by R. Clark Jones one summer in the early 1940's while he was employed in the laboratory of Edwin Land as a Harvard undergraduate. In Jones' formalism Stokes' parameters are folded into the design of a complex 2-vector, and devices are represented by complex 2×2 matrices. The formalism is developed in elaborate detail in my "Ellipsometry" (1999) and in the literature cited there, but it would take us too far afield to attempt to treat the subject here.

4. Partially polarized plane waves. The "plane waves" considered thus far are highly idealized abstractions: they

- are of infinite temporal duration
- are of infinite spatial extent . . . and therefore
- carry infinite energy and momentum, and moreover
- are spatially/temporally perfectly coherent.

But so also—and in much the same way—is the Euclidean plane an idealized abstraction. Euclidean geometry becomes relevant to physical geometry only in contexts (very numerous indeed!) in which it is sensible to conflate the local geometry of the curved surface with the local geometry of the tangent plane. So it is in classical electrodynamics: ideas borrowed from the idealized physics of plane waves become relevant to the physics of realistic radiation fields only as local approximants,[250] and can be expected to lose their utility "in the large," as also in the vicinity of charges, caustics, "kinks" in the field.

But radiation fields the gross properties of which display any degree of spatial/temporal variability cannot be precisely monochromatic. We expect natural fields to acquire also some degree of spatial/temporal incoherence from the radiation production mechanism, whatever it might be. We are led thus to the concept of a **quasi-monochromatic plane wave**—led, that is, to the replacement

$$\boldsymbol{E}(t) = \big\{\boldsymbol{e}_1\mathcal{E}_1e^{i\delta_1} + \boldsymbol{e}_2\mathcal{E}_2e^{i\delta_2}\big\}e^{i\omega t} \tag{395}$$

$$\downarrow$$

$$\boldsymbol{E}(t) = \big\{\boldsymbol{e}_1\mathcal{E}_1(t)e^{i\delta_1(t)} + \boldsymbol{e}_2\mathcal{E}_2(t)e^{i\delta_2(t)}\big\}e^{i\omega t} \tag{412}$$

where ω sets the nominal frequency and $\mathcal{E}_1(t)$, $\mathcal{E}_2(t)$, $\delta_1(t)$ and $\delta_2(t)$ are assumed to change

- slowly with respect to $e^{i\omega t}$ but (in typical cases)
- rapidly with respect to the response time of our photometers.

[250] Beware! Plane waves are, in one critical respect, *not* representative of the typical local facts. I refer to the circumstance that, while $\boldsymbol{E}\perp\boldsymbol{B}$ is characteristic of plane waves, it is <u>not a property of fields in general</u> (superimposed plane waves). See below, page 332.

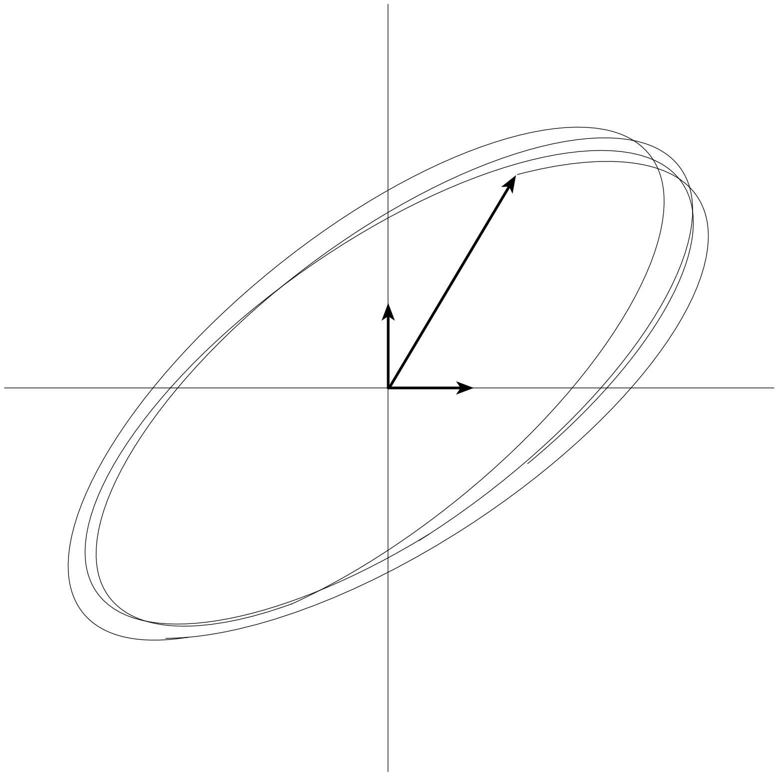


FIGURE 96: *Imperfectly elliptical flight (compare Figure 93) of the* ***E****-vector when the plane wave is only* quasi-*monochromatic*.

Notice that we make no attempt to tinker with the *spatial* properties of the wave (our photometer looks, after all, to only a *local sample* of the physical wave), and that the procedure we have adopted is frankly "phenomenological" in the sense that we do not ask how $\mathcal{E}_1(t)$, $\mathcal{E}_2(t)$, $\delta_1(t)$ and $\delta_2(t)$ might be constrained by Maxwell's equations.

From (412) we conclude that, as illustrated above, $\boldsymbol{E}(t)$ traces an ellipse only in the shortrun—an ellipse with "instantaneous" Stokes parameters given (see again (399)) by

$$\left.\begin{aligned} S_0(t) &= \mathcal{E}_1^2(t) + \mathcal{E}_2^2(t)\\ S_1(t) &= \mathcal{E}_1^2(t) - \mathcal{E}_2^2(t)\\ S_2(t) &= 2\mathcal{E}_1(t)\mathcal{E}_2(t)\cos\delta(t)\\ S_3(t) &= 2\mathcal{E}_1(t)\mathcal{E}_2(t)\sin\delta(t) \end{aligned}\right\} \tag{413}$$

$$\delta(t) \equiv \delta_2(t) - \delta_1(t)$$

The ellipse jiggles about, constantly changing is figure/orientation, in a manner determined by the (let us say steady) *statistical properties of the wave*. The functions $\mathcal{E}_1(t), \mathcal{E}_2(t)$ and $\delta(t)$—whence also $S_0(t)$, $S_1(t)$, $S_2(t)$ and $S_3(t)$—have, in other words, assumed the character of random variables. Our filters and (slow) J-meters, used as described on page 308, supply information not about the functions $S_\mu(t)$ but about their *mean values*:

$$\mathcal{S}_\mu \equiv \langle S_\mu(t)\rangle \equiv \tfrac{1}{T}\int_0^T S_\mu(t)\,dt \quad : \quad \begin{cases} T \text{ might refer to the response} \\ \text{time of the instrument} \end{cases}$$

Proceeding in this light from (403) and (413) we have

$$\left.\begin{aligned} \mathcal{S}_0 &= 2J_0 &&= \langle \mathcal{E}_1^2\rangle + \langle \mathcal{E}_2^2\rangle \\ \mathcal{S}_1 &= 2J_1 - 2J_0 &&= \langle \mathcal{E}_1^2\rangle - \langle \mathcal{E}_2^2\rangle \\ \mathcal{S}_2 &= 2J_2 - 2J_0 &&= 2\langle \mathcal{E}_1\mathcal{E}_2\cos\delta\rangle \\ \mathcal{S}_3 &= 2J_3 - 2J_0 &&= 2\langle \mathcal{E}_1\mathcal{E}_2\sin\delta\rangle \end{aligned}\right\} \tag{414}$$

Evidence that Stokes' parameters are, if not by initial intent, nevertheless wonderfully well-adapted to discussion of the dominant statistical properties of physical lightbeams emerges from the following little argument: working from (414) we have

$$\mathcal{S}_0^2 = \langle \mathcal{E}_1^2\rangle^2 + 2\langle \mathcal{E}_1^2\rangle\langle \mathcal{E}_2^2\rangle + \langle \mathcal{E}_2^2\rangle^2 \tag{415.1}$$

$$\begin{aligned} \mathcal{S}_1^2 + \mathcal{S}_2^2 + \mathcal{S}_3^2 &= \langle \mathcal{E}_1^2\rangle^2 - 2\langle \mathcal{E}_1^2\rangle\langle \mathcal{E}_2^2\rangle + \langle \mathcal{E}_2^2\rangle^2 + \langle 2\mathcal{E}_1\mathcal{E}_2\cos\delta\rangle^2 + \langle 2\mathcal{E}_1\mathcal{E}_2\sin\delta\rangle^2 \\ &= \mathcal{S}_0^2 + 4\big\{\langle \mathcal{E}_1\mathcal{E}_2\cos\delta\rangle^2 + \langle \mathcal{E}_1\mathcal{E}_2\sin\delta\rangle^2 - \langle \mathcal{E}_1^2\rangle\langle \mathcal{E}_2^2\rangle\big\} \end{aligned} \tag{415.2}$$

But if x and y are any random variables (however distributed) then from $\langle(\lambda x + y)^2\rangle = \lambda^2\langle x\rangle^2 + 2\lambda\langle xy\rangle + \langle y\rangle^2 \geqslant 0$ (all λ) it follows that in all cases $\langle xy\rangle^2 \leqslant \langle x^2\rangle\langle y^2\rangle$, so we have

$$\begin{aligned} \langle \mathcal{E}_1\mathcal{E}_2\cos\delta\rangle^2 &\leqslant \langle \mathcal{E}_1^2\rangle\langle \mathcal{E}_2^2\cos^2\delta\rangle \\ \langle \mathcal{E}_1\mathcal{E}_2\sin\delta\rangle^2 &\leqslant \langle \mathcal{E}_1^2\rangle\langle \mathcal{E}_2^2\sin^2\delta\rangle \end{aligned}$$

giving

$$\mathcal{S}_1^2 + \mathcal{S}_2^2 + \mathcal{S}_3^2 \leq \mathcal{S}_0^2 + 4\underbrace{\big\{\langle \mathcal{E}_1^2\rangle\langle \mathcal{E}_2(\cos^2\delta + \sin^2\delta)^2\rangle - \langle \mathcal{E}_1^2\rangle\langle \mathcal{E}_2^2\rangle\big\}}_{0}$$

We are led thus to the important inequality

$$\mathcal{S}_0^2 - \mathcal{S}_1^2 - \mathcal{S}_2^2 - \mathcal{S}_3^2 \geqslant 0 \tag{416}$$

with—according to (400)—equality if (but not only if!) the beam is literally monochromatic. Looking back again to Figure 94, we see that (416) serves to place the vector

$$\boldsymbol{\mathcal{S}} \equiv \begin{pmatrix} \mathcal{S}_1 \\ \mathcal{S}_2 \\ \mathcal{S}_3 \end{pmatrix}$$

inside the Stokes sphere of radius $\mathcal{S}_0$, and that $\boldsymbol{\mathcal{S}}$ reaches all the way to the surface of the Stokes sphere if and only if the beam is, in a fairly evident sense, *statistically equivalent to a monochromatic beam*.

If $\mathcal{E}_1$, $\mathcal{E}_2$ and δ are *statistically independent* random variables then we can in place of (414) write

$$\begin{aligned} \mathcal{S}_0 &= \ \langle \mathcal{E}_1^2 \rangle + \langle \mathcal{E}_2^2 \rangle \\ \mathcal{S}_1 &= \ \langle \mathcal{E}_1^2 \rangle - \langle \mathcal{E}_2^2 \rangle \\ \mathcal{S}_2 &= 2\langle \mathcal{E}_1 \rangle \langle \mathcal{E}_2 \rangle \langle \cos\delta \rangle \\ \mathcal{S}_3 &= 2\langle \mathcal{E}_1 \rangle \langle \mathcal{E}_2 \rangle \langle \sin\delta \rangle \end{aligned}$$

If, moreover, all δ-values are equally likely, then $\langle \cos\delta \rangle = \langle \sin\delta \rangle = 0$, and we have $\mathcal{S}_2 = \mathcal{S}_3 = 0$. If, moreover, $\langle \mathcal{E}_1 \rangle = \langle \mathcal{E}_2 \rangle$ then $\mathcal{S}_1 = 0$. The resulting beam

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{is said to be } \textit{unpolarized} : \boldsymbol{\mathcal{S}} = \mathbf{0}$$

It becomes on this basis natural to introduce the

$$\text{“degree of polarization”} \ P \equiv \frac{\sqrt{\mathcal{S}_1^2 + \mathcal{S}_2^2 + \mathcal{S}_3^2}}{\mathcal{S}_0} \quad : \quad 0 \leqslant P \leqslant 1 \tag{417}$$

and to write

$$\begin{pmatrix} \mathcal{S}_0 \\ \mathcal{S}_1 \\ \mathcal{S}_2 \\ \mathcal{S}_3 \end{pmatrix} = \begin{pmatrix} P\mathcal{S}_0 \\ \mathcal{S}_1 \\ \mathcal{S}_2 \\ \mathcal{S}_3 \end{pmatrix} + \begin{pmatrix} (1-P)\mathcal{S}_0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$= 100\% \text{ polarized component} \quad + \quad \text{unpolarized component}$$

When an unpolarized beam is presented to (for example) the linear polarizer of (402.1) one obtains

$$\begin{pmatrix} \mathcal{S}_0 \\ \mathcal{S}_1 \\ \mathcal{S}_2 \\ \mathcal{S}_3 \end{pmatrix}_{\!\text{in}} \xrightarrow[\text{linear polarizer at } 0^\circ]{} \begin{pmatrix} \mathcal{S}_0 \\ \mathcal{S}_1 \\ \mathcal{S}_2 \\ \mathcal{S}_3 \end{pmatrix}_{\!\text{out}} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathcal{S}_0 \\ 0 \\ 0 \\ 0 \end{pmatrix}_{\!\text{in}} = \begin{pmatrix} \frac{1}{2}\mathcal{S}_0 \\ \frac{1}{2}\mathcal{S}_0 \\ 0 \\ 0 \end{pmatrix}$$

Here

$$\begin{aligned} P_{\text{in}} = 0 \quad &: \quad \text{the entry beam is unpolarized, but} \\ P_{\text{out}} = 1 \quad &: \quad \text{the exit beam is 100\% polarized} \end{aligned}$$

And when the exit beam is presented to a second linear polarizer, described by the $\mathbb{M}(\psi)$ of page 312, one obtains[251] the "Law of Malus":

$$\frac{\text{output intensity}}{\text{input intensity}} = \tfrac{1}{4}(1+\cos 2\psi) = \tfrac{1}{2}\cos^2\psi$$

A quasi-monochromatic beam is said to be

$$\left.\begin{array}{r} \text{unpolarized} \\ \text{partially polarized} \\ \text{completely polarized} \end{array}\right\} \text{ according as } \left\{\begin{array}{l} 0 = P \\ 0 < P < 1 \\ \quad\;\; P = 1 \end{array}\right.$$

An unpolarized beam necessarily *is* polarized in the shortrun, but in the longer term the $\boldsymbol{E}$-vector traces an orientation-free scribble. Partial polarization results when the scribble is *somewhat* oriented (fuzzy): this requires that $\mathcal{E}_1(t)$, $\mathcal{E}_2(t)$, $\delta_1(t)$ and $\delta_2(t)$ more somewhat in concert; *i.e.*, that they be *statistically correlated*. It is important to note that the numbers $\mathcal{S}_\mu$ provide a very incomplete description of the beam statistics, and that even complete knowledge of the statistical properties of the beam would leave the actual t-dependence of $\boldsymbol{E}$ indeterminate. *Many* beams are—even in the case of complete polarization—consistent with any prescribed/measured set of $\mathcal{S}_\mu$-values.

We are by those remarks into position to appreciate the import of Stokes'

> **Principle of Optical Equivalence**: Lightbeams with identical Stokes parameters are "equivalent" in the sense that they interact identically with devices which detect or alter the intensity and/or polarizational state of the incident beam.

and the depth of his insight into the physics of light. But one does not say of objects that they are, in designated respects, "equivalent" unless there exist other respects—whether overt or covert—in which they are at the same time inequivalent; implicit in the formulation of Stokes' principle is an assertion that physical light beams possess properties beyond those to which the Stokes parameters allude, properties to which photometer-like devices are insensitive. There are many ways to render a page gray with featureless squiggles, many ways to assemble an unpolarized light beam. What such beams, such statistical assemblages share is, according to (414), not "identity" but only the property that a certain quartet of numbers arising from their low-order moments and correlation coefficients are equi-valued.

---

[251] PROBLEM 66. Étiènne Louis Malus (1775–1812) was a French engineer/ physicist.

We have, in effect, been alerted by Stokes to the existence of a "statistical optics"—to the possibility that instruments (more subtle in their action than photometers) might be devised which are sensitive to higher moments of an incident optical beam. And we have been alerted to the possible existence and potential usefulness of an *ascending hierarchy* of "higher order analogs" of the parameters that bear Stokes' name, formal devices that serve to capture successively more refined statistical properties of optical beams. Examination of the literature[252] shows all those expectations to be borne out by fairly recent developments. It becomes interesting in the light of these remarks to recall the title of the paper in which the Stokes parameters were first described: "On the composition and resolution of streams of polarized light from different sources" (Trans. Camb. Phil. Soc. **9**, 399 (1852)). Stokes brought the theory of physical light beams to a state somewhat analogous to that encountered in thermodynamics, where a few operationally defined variables mask a rich time-dependent microphysics, yet serve to support a formalism which is—surprisingly—closed/self-consistent/complete . . . and which accounts accurately for the phenomenological facts.

Already on page 318 we began to accumulate evidence that the Mueller calculus is as "robust" as the Stokes formalism upon which it is based. To our former population of Mueller matrices $\mathbb{M}$ it might now seem appropriate to add (for example)

$$\mathbb{M} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{-u} & 0 & 0 \\ 0 & 0 & e^{-u} & 0 \\ 0 & 0 & 0 & e^{-u} \end{pmatrix} \quad : \quad u \geqslant 0 \tag{418}$$

which evidently describes the action of an *isotropic depolarizer*, where the adjective refers to isotropy not in physical space but in Stokes space. The interesting point—which stands as an open invitation to formal/physical invention—is that the $\mathbb{M}$ described above *does not satisfy the fundamental Mueller condition* (404.1). Relatedly: I am informed by Morgan Mitchell, my optical colleague, that while active "polarization scramblers" do exist, a "*passive* depolarization device" would be a "tall order." [253,254]

5. Optical beams. Listed at the beginning are several respects in which "plane waves are highly idealized abstractions." With the introduction of the notion of "quasi-monochromaticity" we were able to introduce an element of realism into the discussion, but

- infinite temporal duration
- infinite spatial extent
- infinite energy/momentum

[252] See, for example, E. L. O'Neill, *Introduction to Statistical Optics* (1963); J. W. Simmons & M. J. Guttmann, *States, Waves and Photons: A Modern Introduction to Light* (1970); C. Brosseau,*Fundamentals of Polarized Light: A Statistical Optics Approach* (1998).

[253] PROBLEM 67.

[254] PROBLEM 68.

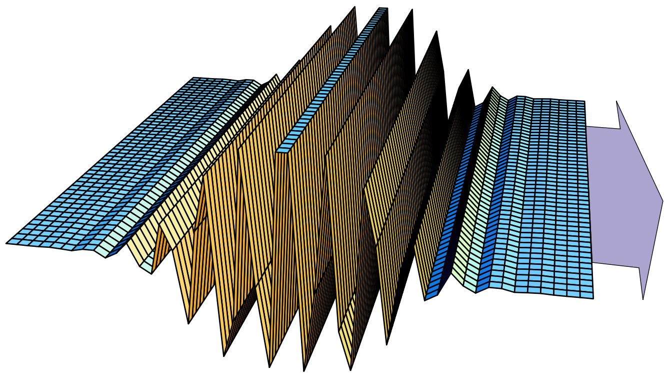


FIGURE 97: *Representation of the function $\varphi(t,x,0,z)$ described at (419) below. The Gaussian wavepacket glides rigidly, as indicated by the arrow. It is temporally confined, but spatially unconfined.*

are unphysical abstractions that survived untouched in the ensuing discussion of beam statistics and imperfect polarization. Temporal confinement is fairly easy to achieve, as the following remark makes clear:

Write $e^{i(kct+0x+0y-kz)}$ to describe a plane wave running up the z-axis. Write

$$\begin{aligned}\varphi(t,x,y,z) &= \int_{-\infty}^{+\infty} f(k)e^{ik(ct-z)}\,dk\\ &= \int_{-\infty}^{+\infty} g(\omega)e^{i\omega(t-z/c)}\,d\omega\end{aligned}$$

to describe a weighted superposition of such waves. Take $g(\omega)$ to have, in particular, the form of a normalized Gaussian centered at Ω:

$$\begin{aligned}g(\omega) &\equiv \tfrac{1}{\sqrt{2\pi}}Te^{-\frac{1}{2}T^2(\omega-\Omega)^2} \quad:\ T>0 \text{ has the physical dimension of TIME}\\ &\downarrow\\ &= \delta(\omega-\Omega) \quad \text{as } T\uparrow\infty\end{aligned}$$

Then

$$\begin{aligned}\varphi(t,x,y,z) &= \int_{-\infty}^{+\infty} \tfrac{1}{\sqrt{2\pi}}Te^{-\frac{1}{2}T^2(\omega-\Omega)^2}e^{i\omega(t-z/c)}\,d\omega\\ &= e^{-\frac{1}{2}T^{-2}(t-z/c)^2}\cdot e^{i\Omega(t-z/c)} \qquad (419)\\ &\downarrow\\ &= e^{i\Omega(t-z/c)} \quad \text{as } T\uparrow\infty\end{aligned}$$

The physical (*i.e.*, the real) part of the expression on the right side of (419) is plotted in Figure 97.

I have occasionally allowed myself to speak informally of "beams" when the objects to which I referred were actually plane waves. We confront now the mathematical force of the distinction. While the waves sampled by astronomers are good approximations to plane waves, when we go into the laboratory to perform optical experiments we deal most commonly with laterally confined light *beams*.[255] The mathematical description of lateral confinement poses a number of delicate problems entirely absent from the theory of temporal confinement. The subject acquired new urgency from the invention of the laser, and it is from a classic contribution to that literature[256] that I have adapted the following remarks:

Setting aside, for the moment, the fact that electromagnetic radiation is properly described by a *transverse vector* field, we look for laterally confined monochromatic solutions $\varphi(t,x,y,z) = e^{i\omega t}\cdot\phi(x,y,z)$ of the *scalar* wave equation $\Box\varphi = 0$. Which is to say (see again page 294): we look for laterally confined solutions of the Helmholtz equation

$$\Big\{(\tfrac{\partial}{\partial x})^2 + (\tfrac{\partial}{\partial y})^2 + (\tfrac{\partial}{\partial z})^2 + k^2\Big\}\phi(x,y,z) = 0$$

We have interest in laterally confined waves propagating in the z-direction, so look for solutions of the form

$$\phi(x,y,z) = e^{-ikz}\cdot\psi(x,y,z) \quad : \quad k = \omega/c$$

Which is to say: we look for laterally confined solutions of

$$\Big\{(\tfrac{\partial}{\partial x})^2 + (\tfrac{\partial}{\partial y})^2 + (\tfrac{\partial}{\partial z})^2\Big\}\psi(x,y,z) = 2ik\tfrac{\partial}{\partial z}\psi(x,y,z)$$

We agree to work in the **approximation** that $\psi(x,y,z)$ changes so gradually in the z-direction that the red $(\frac{\partial}{\partial z})^2$-term can be dropped. We arrive then at an equation

$$\tfrac{1}{2k}\Big\{(\tfrac{\partial}{\partial x})^2 + (\tfrac{\partial}{\partial y})^2\Big\}\psi(x,y,z) = \quad i\tfrac{\partial}{\partial z}\psi(x,y,z) \tag{420}$$

which is structurally reminiscent of the Schrödinger equation for a particle free to move in two dimensions:

$$\tfrac{\hbar}{2m}\Big\{(\tfrac{\partial}{\partial x})^2 + (\tfrac{\partial}{\partial y})^2\Big\}\psi(x,y,t) = -i\tfrac{\partial}{\partial t}\psi(x,y,t)$$

Both equations have unlimitedly many solutions, depending

- in quantum mechanics upon the form assigned to $\psi(x,y,t)$ at an initial time t_0, commonly taken to be $t_0 = 0$

[255] We do not speak of "star beams," and it is only for local meteorological reasons that we speak sometimes of "sun beams."

[256] H. Kogelnik & T. Li, "Laser beams and resonators," Applied Optics **5**, 1550 (1966). See also §4.5 in O. Svelto, *Principles of Lasers* (3rd edition 1989).

- in beam theory upon the form assigned to $\psi(x, y, z)$ at some prescribed axial point z_0; we will find it convenient to take $z_0 = 0$.

To illustrate the point, the authors of quantum texts[257] often take $\psi(x, y, t_0)$ to be Gaussian

$$\psi(x, y, 0) = Ae^{-a(x^2+y^2)}$$

and by one or another of the available computational techniques obtain

$$\psi(x, y, 0) \xrightarrow[t]{} \psi(x, y, t) = A\tfrac{1}{1+i(t/T)}\exp\Big\{-\tfrac{a(x^2+y^2)}{1+i(t/T)}\Big\} \quad : \quad T \equiv m/2a\hbar$$

which they use to demonstrate the characteristic *temporal diffusion of initially localized quantum states*. Exactly the same mathematics lies at the base of the "theory of Gaussiam beams." Suppose it to be the case that

$$\psi(x, y, z) = Be^{-a(x^2+y^2)} \quad \text{at } z = 0$$

The exact solution of (420) is given then at other axial points z by[258]

$$\begin{aligned}\psi(x, y, z) &= B\tfrac{1}{1-iz/Z}e^{-a(x^2+y^2)/(1-iz/Z)} \quad : \quad Z \equiv k/2a\\ &= B\tfrac{1}{1+(z/Z)^2}[1 + i(z/Z)]\exp\Big\{-ar^2\tfrac{1}{1+(z/Z)^2}[1 + i(z/Z)]\Big\}\end{aligned}$$

$$[1 + i(z/Z)] = \sqrt{1 + (z/Z)^2}\, e^{i\Phi} \text{ with } \Phi \equiv \arctan(z/Z)$$

$$= \frac{B}{\sqrt{1+(z/Z)^2}}\exp\Big\{-ar^2\tfrac{1}{1+(z/Z)^2}\Big\}\exp\Big\{i\big[\Phi(z) - ar^2\tfrac{z/Z}{1+(z/Z)^2}\big]\Big\}$$

$$r^2 \equiv x^2 + y^2$$

We are brought thus to a beam of the design

$$\varphi(t, x, y, z) \sim \tfrac{1}{\rho(z)}\exp\Big\{-\big[\tfrac{r}{\rho(z)}\big]^2\Big\}\cdot e^{i\big[\omega t - kz + \Phi(z) - (r/\rho)^2(z/Z)\big]} \tag{421}$$

where

$$\rho(z) \equiv \sqrt{\frac{1 + (z/Z)^2}{a}} \quad \text{describes the "spot radius" at } z$$

Evidently

$$\rho_{\min} \equiv \rho_0 = \rho(0) = \sqrt{1/a} \quad : \quad \text{called the "beam waist"}$$

and at this point the a-notation—a relic of Griffiths' discussion of another subject—has outworn its usefulness: we agree henceforth to write $1/\rho_0^2$ in place of a. In this new notation we have

$$\rho(z) = \rho_0\sqrt{1 + (z/Z)^2} \qquad i.e., \qquad (\rho/\rho_0)^2 - (z/Z)^2 = 1 \tag{422}$$

---

[257] See, for example, David Griffiths, *Introduction to Quantum Mechanics* (1995), page 50: Problem 2.22.

[258] PROBLEM 69.

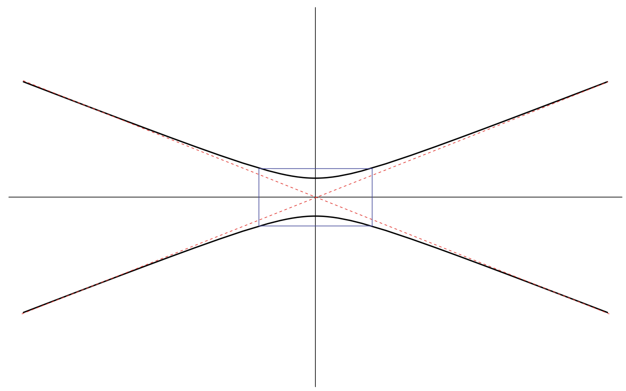


FIGURE 98: *Graph of the function $\rho(z) = \rho_0\sqrt{1 + (z/Z)^2}$ that lends the Gaussian beam its hyperbolic profile. The asymptotes are shown in red. The blue box is of length L. Its ends are positioned at $z = \pm Z$, where the spot radius has grown from ρ_0 to $\sqrt{2}\rho_0$. The figure was drawn with $Z/\rho_0 = 10$, and is in that respect misleading: in realistic cases $Z/\rho_0 \sim 10^4$ and the angle between the asymptotes (beam divergence) is* much(!) *reduced.*

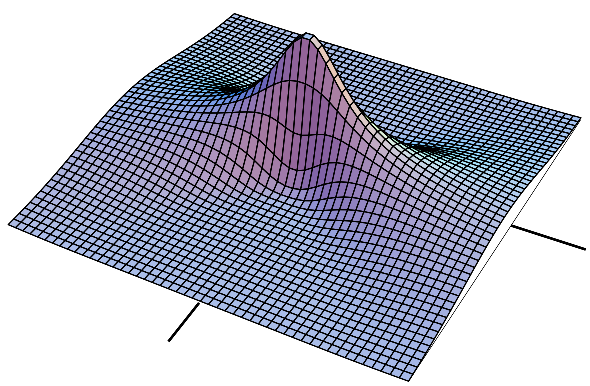


FIGURE 99: *Graph of the factor that, according to* (421), *controls the amplitude of a Gaussian beam. The values assigned to ρ_0 and Z are the same as those in the preceding figure, and are unrealistic in the sense already explained. The running-wave modulation would be much too finely detailed to be displayed at the same scale.*

which shows that the growth of the spot radius is hyperbolic (see Figure 98 & Figure 99), with asymptotes

$$\rho_{\text{asymptotic}} = \pm(\rho_0/Z)z$$

the slopes of which are typically very shallow: from the definition

$$Z \equiv \tfrac{1}{2}L \equiv k\rho_0^2/2 = \pi\rho_0^2/\lambda$$

we have

$$\left.\begin{aligned}\text{beam waist} &= 0.3989\sqrt{L\lambda}\\ \text{beam divergence} &= 0.7978\sqrt{\lambda/L}\end{aligned}\right\} \tag{423}$$

where the numerics arise from $\sqrt{1/2\pi}$ and $\sqrt{2/\pi}$ respectively. In a typical case $L \sim 1$ meter and $\lambda \sim 7.0\times10^{-7}$ meter, giving

$$\begin{aligned}\text{beam waist} &= 0.33 \text{ mm}\\ \text{beam divergence} &= 6.67\times10^{-4} \text{ (dimensionless)}\end{aligned}$$

Such a beam must travel about 15 meters for the spot radius to grow to 1 cm.

Looking back again to (421), we set $r = 0$ and find that the

$$\text{axial phase at } z = \omega t - kz + \arctan(z/Z)$$

Arguing from $\frac{d}{dt}(\text{axial phase at } z) = 0$ we compute

$$\begin{aligned}\text{phase velocity at } z &= \left[k - \tfrac{Z}{Z^2+z^2}\right]^{-1}\cdot\omega\\ &= \left[1 - \tfrac{Z\lambda}{2\pi(Z^2+z^2)}\right]^{-1}\cdot c \quad \text{by } Z/k = Z\lambda/2\pi = \tfrac{1}{2}\rho_0^2\\ &= \left\{1 + \tfrac{\lambda}{2\pi Z} + \left(\tfrac{\lambda}{2\pi Z}\right)^2 + \cdots\right\}\cdot c \gtrsim c \quad \text{at } z = 0\\ &\downarrow\\ &= c \quad \text{as } z \to \infty\end{aligned}$$

—the interesting point being that as z becomes large the axial phase velocity approaches c *from above*. Looking next to the *geometry of the near-axial equiphase surfaces*, we study

$$kz - \arctan(z/Z) + \frac{r^2}{\rho_0^2[1 + (z/Z)^2]}(z/Z) = kz_0 - \arctan(z_0/Z) \tag{424}$$

where z_0 marks the point at which the surface in question intersects the z-axis. Taking both r^2 and $z_0 - z$ to be small and asking *Mathematica* to develop the arctan as a power series in $(z - z_0)$, we obtain

$$\frac{r^2}{\rho_0^2[1 + (z_0/Z)^2]}(z_0/Z) = \left\{k - \frac{1}{Z[1 + (z_0/Z)^2]}\right\}(z_0 - z) + \cdots \tag{425}$$

Define R in such a way that

$$\frac{k}{2R} \equiv \frac{1}{\rho_0^2[1 + (z_0/Z)^2]}(z_0/Z)$$

which is to say: let $R \equiv z[1 + (Z/z)^2]$, so that the expression on the left side of

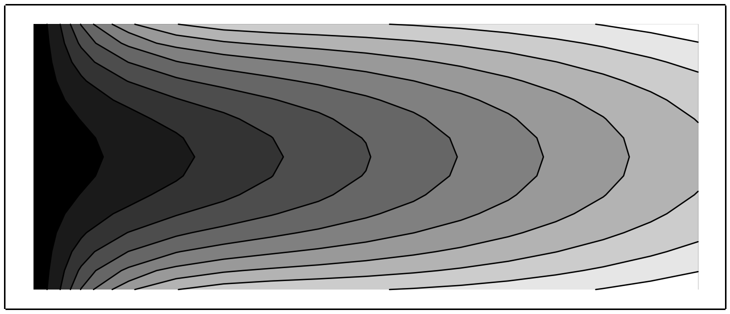


FIGURE 100: *Equiphase contours, taken from the expression on the left side of* (424).

(425) can be written $kr^2/2R$. Next, notice that

$$\frac{1}{Z[1+(z_0/Z)^2]} < \frac{1}{Z} = \frac{2}{L} \ll \pi\frac{2}{\lambda} = k$$

so the second term in braces can be abandoned, giving (see Figure 100)

$$z_0 - z = (1/2R)(x^2+y^2) \;:\; \begin{cases}\text{parabola-of-revolution, opening}\\ \text{to the left, with apex at } z_0\end{cases} \tag{425}$$

That

$$R = \text{radius of curvature at the apex}$$

follows from the observations (i) that

$$[z-(z_0-R)]^2 + x^2 + y^2 = R^2 \tag{426}$$

describes a sphere of radius R that is centered on the z-axis and intersects that axis at $z = z_0$ and $z = z_0 - 2R$, and (ii) that expansion of (426) gives back (425) if a small $(z_0 - z)^2$-term is abandoned. This information might be used to design the concave mirrors placed at the ends of a "Gaussian laser."

The Gaussian beam discussed above can be used as the "seed" from which to grow an infinite population of "Gaussian beams of higher order." These (at least those of lower order) are of physical importance when taken individually, and collectively enable one (by weighted superposition) to fabricate beams of unlimited variety. The generative idea is quite elementary

> If φ is a solution of $\Box\varphi = 0$ and if $\mathcal{D}$ is a differential operator that *commutes* with $\Box$
>
> $$\mathcal{D}\Box = \Box\mathcal{D}$$
>
> then so also is $\mathcal{D}\varphi$ a solution.

but must be adapted to the approximation scheme that was seen on page 322 to lie at the base of Gaussian beam theory: we write

$$\varphi(t,x,y,z) = e^{i(\omega t - kx)}\cdot\psi(x,y,z)$$

and require that ψ be an exact solution of the "Schrödinger equation"[259]

$$\Big\{(\tfrac{\partial}{\partial x})^2+(\tfrac{\partial}{\partial y})^2\Big\}\psi(x,y,z)=i(4Z/\rho_0^2)\tfrac{\partial}{\partial z}\psi(x,y,z) \tag{427}$$

Taking from page 323 the demonstrably exact solution

$$\psi_{00}=\rho_0\,\frac{1}{\rho_0[1-iz/Z]}\,\exp\Big\{-\frac{x^2+y^2}{\rho_0^2[1-iz/Z]}\Big\}$$

—which by $\rho_0[1-iz/Z]=\rho_0\sqrt{1+(z/Z)^2}\,e^{-i\arctan(z/Z)}\equiv\rho(z)e^{-i\Phi(z)}$ can also be written

$$\begin{aligned}&=\rho_0\,\frac{1}{\rho(z)}e^{i\Phi(z)}\cdot\exp\Big\{-\Big[\frac{x}{\sigma(z)}\Big]^2-\Big[\frac{y}{\sigma(z)}\Big]^2\Big\}\\ &\qquad\qquad\qquad\qquad\sigma(z)\equiv\sqrt{\rho_0\rho(z)}\,e^{-i\frac{1}{2}\Phi}\\ &=\rho_0\,\tfrac{1}{\rho}e^{i\Phi}\cdot e^{-\xi^2-\eta^2}\quad:\quad \xi\equiv x/\sigma \text{ and } \eta\equiv y/\sigma\end{aligned}$$

—as our "seed," we harvest this fairly natural fruit:

$$\begin{aligned}\psi_{mn}&\equiv\big(-\rho_0\tfrac{\partial}{\partial x}\big)^m\big(-\rho_0\tfrac{\partial}{\partial y}\big)^n\,\psi_{00}\\ &=\rho_0\tfrac{1}{\rho}e^{i\Phi}\big(-\rho_0\tfrac{1}{\sigma}\tfrac{\partial}{\partial\xi}\big)^m\big(-\rho_0\tfrac{1}{\sigma}\tfrac{\partial}{\partial\eta}\big)^n e^{-\xi^2-\eta^2}\\ &=\rho_0\tfrac{1}{\rho}e^{i\Phi}\big(\rho_0\tfrac{1}{\sigma}\big)^{m+n}\big(-\tfrac{\partial}{\partial\xi}\big)^m\big(-\tfrac{\partial}{\partial\eta}\big)^n e^{-\xi^2-\eta^2}\\ &=\big(\rho_0\tfrac{1}{\rho}\big)^{1+\frac{1}{2}(m+n)}\,e^{i[1+\frac{1}{2}(m+n)]\Phi}H_m(\xi)H_n(\eta)\cdot e^{-\xi^2-\eta^2}\end{aligned} \tag{428}$$

In the final line we have recalled[260] Rodrigues' construction

$$H_m(\xi)=e^{\xi^2}\big(-\tfrac{\partial}{\partial\xi}\big)^m e^{-\xi^2}$$

of the Hermite polynomials:

$$\begin{aligned}H_0(\xi)&=1\\ H_1(\xi)&=2\xi\\ H_2(\xi)&=4\xi^2-2\\ H_3(\xi)&=8\xi^3-12\xi\\ H_4(\xi)&=16\xi^4-48\xi^2+12\\ &\ \vdots\end{aligned}$$

That the functions ψ_{mn} constructed in this way do in fact exactly satisfy (427) can be demonstrated (for small m, n) by *Mathematica*-assisted calculation, but that they *must* do so follows transparently from the observation that

$$\tfrac{\partial}{\partial x}\text{ and }\tfrac{\partial}{\partial y}\text{ commute with }\Big\{(\tfrac{\partial}{\partial x})^2+(\tfrac{\partial}{\partial y})^2\Big\}-i(4Z/\rho_0^2)\tfrac{\partial}{\partial z}$$

---

259 This is just (420) with $k\mapsto 2aZ=2Z/\rho_0^2$.

260 See, for example, Chapter 24 in J. Spanier & K. B. Oldham, *An Atlas of Functions* (1987).

The Gaussian factor

$$e^{-\xi^2-\eta^2} = \exp\left\{-\frac{x^2+y^2}{\rho^2(z)}[1+i(z/Z)]\right\}$$

is a shared feature of *all* the ψ_{mn}-functions, which give rise therefore to identical populations of equiphase surfaces (Figure 100). Using *Mathematica*'s

`HermiteH[n,x]`

command to evaluate the complex prefactors

$$g_{mn}(x,y) \equiv \big(\rho_0\tfrac{1}{\rho}\big)^{1+\frac{1}{2}(m+n)}\, e^{i[1+\frac{1}{2}(m+n)]\Phi} H_m\big(\tfrac{x}{\sqrt{\rho_0\rho}}e^{i\frac{1}{2}\Phi}\big) H_n\big(\tfrac{y}{\sqrt{\rho_0\rho}}e^{i\frac{1}{2}\Phi}\big)$$

in some low-order cases, we find

$$\begin{aligned}
g_{00} &= (\rho_0/\rho)e^{i\Phi}\\
g_{10} &= (\rho_0/\rho)2(x/\rho)e^{2i\Phi}\\
&\ \vdots\\
g_{20} &= (\rho_0/\rho)4(x/\rho)^2e^{3i\Phi} - 2(\rho_0/\rho)^2e^{2i\Phi}\\
g_{11} &= (\rho_0/\rho)\big\{2(x/\rho)\big\}\big\{2(y/\rho)\big\}e^{3i\Phi}\\
&\ \vdots\\
g_{30} &= (\rho_0/\rho)8(x/\rho)^3e^{4i\Phi} - 12(\rho_0/\rho)^2(x/\rho)e^{3i\Phi}\\
g_{21} &= (\rho_0/\rho)\big\{4(x/\rho)^2 2(y/\rho)e^{4i\Phi} - 2(\rho_0/\rho)^2e^{3i\Phi}\big\}\big\{2(y/\rho)\big\}\\
&\ \vdots
\end{aligned}$$

The red terms depart from the result asserted by Kogelnik & Li and quoted by Svelto[256]:

$$g_{mn} = (\rho_0/\rho)H_m(x/\rho)H_n(y/\rho)\, e^{i[1+m+n]\Phi}$$

Their results[261] and mine are, however, in precise agreement at $z = 0$, where $\rho = \rho_0$ and $\Phi = 0$ give

$$\psi_{mn}(x,y,0) = H_m(x/\rho_0)H_n(y/\rho_0)\exp\left\{-\frac{x^2+y^2}{\rho_0^2}\right\}$$

This striking result acquires special interest from the orthogonality relation

$$\int_{-\infty}^{+\infty} H_\mu(u)H_\nu(u)e^{-u^2}\,du = \sqrt{\pi}\mu!2^\mu\delta_{\mu\nu}$$

---

[261] . . . which are not incorrect (as I for awhile supposed) but refer to a *distinct population* of beam modes: the point is developed in §§3 & 4 of a companion essay "Toward an exact theory of lightbeams" (2002).

For if we introduce the "normalized Gaussian beam functions"

$$\Psi_{mn}(x,y,z) \equiv \tfrac{1}{\rho_0\sqrt{m!2^m n!2^n\pi}}\psi_{mn}(x,y,z)$$

then we have

$$\int_{-\infty}^{+\infty}\!\int_{-\infty}^{+\infty}\Psi_{\mu\nu}(x,y,0)\Psi_{mn}(x,y,0)\exp\left\{\frac{x^2+y^2}{\rho_0^2}\right\}dxdy=\delta_{\mu m}\delta_{\nu n}$$

which we can use to evaluate the coefficients c_{mn} that enter into the description

$$\psi(x,y,0)=\sum_{m,n}c_{mn}\Psi_{mn}(x,y,0)$$

of *beam structure at the waist*. We then write

$$\begin{aligned}\varphi(t,x,y,z) &= e^{i(\omega t-kz)}\cdot\sum_{m,n}c_{mn}\Psi_{mn}(x,y,z) \qquad (429)\\ &= \sum_{\text{modes}}\text{Gaussian beams of various "modes" (identified by } m,n)\end{aligned}$$

to describe the generalized Gaussian beam possessing that prescribed structure at the waist.

Physically more realistic beam models would be obtained if we

- used the mechanism described on page 321 to *turn the beam on/off* (this would entail loss of strict monochromaticity)
- constructed *statistical linear combinations* of such beams.

But the beams thus constructed could not possibly describe laser beams: they are *scalar* beams ("acoustic" beams), whereas physical laser beams must be endowed with the transverse vectorial properties known to be characteristic of all electromagnetic radiation. This is a circumstance we were content to set aside on page 322, but would like now to find some way to accommodate. I invite you to turn on *Mathematica* and follow along...

Exponential solutions of the "Schrödinger equation" (427) can be described

$$\exp\left\{i\Big[-px-qy+\frac{p^2+q^2}{4Z/\rho_0^2}z\Big]\right\} \quad : \quad \text{all real } p,q$$

and minimal tinkering leads to the discovery that

$$\begin{aligned}\iint_{-\infty}^{+\infty}\tfrac{\rho_0^2}{4\pi}e^{-\frac{1}{4}\rho_0^2(p^2+q^2)}\cdot\exp\left\{i\Big[-px-qy+\frac{p^2+q^2}{4Z/\rho_0^2}z\Big]\right\}dpdq \qquad (430.1)\\ =\frac{1}{[1-iz/Z]}\exp\left\{-\frac{x^2+y^2}{\rho_0^2[1-iz/Z]}\right\}\\ =\psi_{00}(x,y,z)\text{ of page 327}\end{aligned}$$

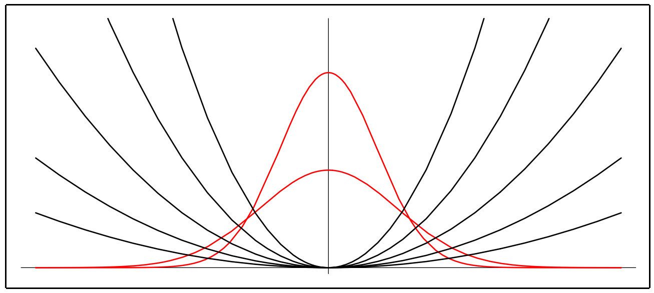


FIGURE 101: *As* ρ_0 *increases the Gaussian* $g = \frac{\rho_0^2}{4\pi}e^{-\frac{1}{4}\rho_0^2(p^2+q^2)}$ *becomes narrower, while at higher frequencies the parabolic term* $f = \frac{1}{4\pi}(p^2+q^2)\lambda$ *becomes shallower. At sufficiently high frequencies the Gaussian discriminates against the* (p,q)*-values where* f *departs significantly from zero, and it is this circumstance that justifies the approximation upon which Gaussian beam theory is based.*

But (see again the bottom of page 323) $4Z/\rho_0^2 = 2k$ so we have

$$\begin{aligned}\varphi_{00}(t,x,y,z) &= e^{i(\omega t-kz)}\cdot\psi_{00}(x,y,z) \quad\text{with}\quad \omega = kc \qquad (430.2)\\ &= \iint_{-\infty}^{+\infty}\tfrac{\rho_0^2}{4\pi}e^{-\frac{1}{4}\rho_0^2(p^2+q^2)}\cdot\exp\Big\{i\Big[\omega t - px - qy - \Big(k-\frac{p^2+q^2}{2k}\Big)z\Big]\Big\}\,dpdq\end{aligned}$$

From

$$(\omega/c = k)^2 - p^2 - q^2 - \Big(k - \frac{p^2+q^2}{2k}\Big)^2 = -\Big(\frac{p^2+q^2}{2k}\Big)^2$$

we see that the wave vector

$$\begin{pmatrix}k^0\\k^1\\k^2\\k^3\end{pmatrix} = \begin{pmatrix}\omega/c\\p\\q\\k-[(p^2+q^2)/2k]\end{pmatrix}$$

is not null (as the wave equation $\Box\varphi_{00} = 0$ requires) but spacelike: we encounter here the **force of the approximation** made on page 322. Notice in this connection that (because $k = 2\pi/\lambda$)

$$\frac{p^2+q^2}{2k} = \frac{p^2+q^2}{4\pi}\cdot\lambda \quad : \quad \text{vanishes at high frequencies}$$

so for given ρ_0 the approximation becomes better and better as $\lambda\uparrow\infty$, while for given λ the approximation becomes progressively better as the Gaussian $e^{-\rho_0^2(p^2+q^2)}$ becomes narrower; *i.e.*, as ρ_0 becomes larger (see the figure).

We want now to extract from (430) the description of a Gaussian *light* beam. To that end we must replace the scalar plane waves encountered at (430) with *electromagnetic* plane waves, and that effort presents certain problems. I will carry this discussion only far enough to expose the problems and some

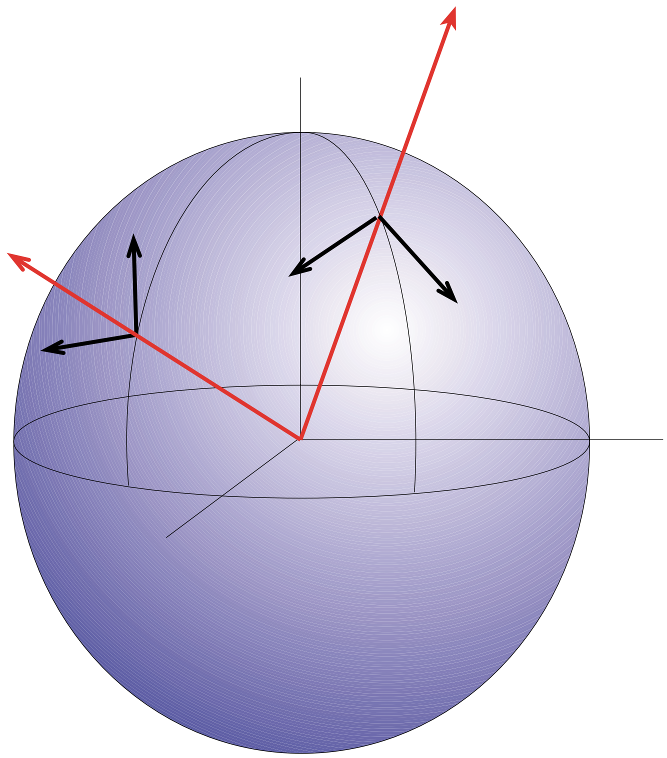


FIGURE 102: *To each vector* $\boldsymbol{k}(p,q)$ *we associate a pair of unit vectors* $\boldsymbol{e}(p,q)$ *and* $\boldsymbol{f}(p,q)$ *in such a way that* $\{\hat{\boldsymbol{k}}, \boldsymbol{e}, \boldsymbol{f} \equiv \hat{\boldsymbol{k}}\times\boldsymbol{e}\}$ *comprise an orthonormal triad. Let* $\boldsymbol{k}_1$ *and* $\boldsymbol{k}_2$ *be two such wave vectors. If* $\boldsymbol{k}_1$ *and* $\boldsymbol{k}_2$ *are not parallel then specification of* $\boldsymbol{e}_1$ *exerts no geometrically compelling constraint on the selection of* $\boldsymbol{e}_2$.

points of principle: to carry it farther would to risk becoming lost in bewildering detail.

Let (430.2) be notated

$$\varphi_{00}(t,\boldsymbol{x}) = \iint_{-\infty}^{+\infty} \tfrac{\rho_0^2}{4\pi} e^{-\frac{1}{4}\rho_0^2(p^2+q^2)} \cdot \exp\Big\{ i\Big[\omega t - \boldsymbol{k}(p,q)\cdot\boldsymbol{x}\Big]\Big\}\, dpdq$$

with

$$\boldsymbol{k}(p,q) \equiv \begin{pmatrix} p \\ q \\ k - [(p^2+q^2)/2k] \end{pmatrix} = k\sqrt{1 + \left(\tfrac{p^2+q^2}{2k^2}\right)^2} \cdot \hat{\boldsymbol{k}}(p,q)$$

To every such $\boldsymbol{k}(p,q)$ assign unit vectors $\boldsymbol{e}(p,q)$ and $\boldsymbol{f}(p,q)$ in such a way that $\{\hat{\boldsymbol{k}}(p,q), \boldsymbol{e}(p,q), \boldsymbol{f}(p,q)\}$ is an orthonormal triad. The first point of interest is that *this can be accomplished in infinitely many ways*: the triads erected at the points of (p,q)-space are independent creations (see Figure 102). Given such an assignment, form

$$\boldsymbol{E}(t,\boldsymbol{x}) = \iint_{-\infty}^{+\infty} \tfrac{\rho_0^2}{4\pi} e^{-\frac{1}{4}\rho_0^2(p^2+q^2)} \cdot \boldsymbol{\mathcal{E}}(p,q)\exp\Big\{i\Big[\omega t - \boldsymbol{k}(p,q)\boldsymbol{\cdot}\boldsymbol{x}\Big]\Big\}\,dpdq \tag{431}$$

with

$$\boldsymbol{\mathcal{E}}(p,q) \equiv \mathcal{E}_1(p,q)\boldsymbol{e}(p,q) + \mathcal{E}_2(p,q)e^{i\delta(p,q)}\boldsymbol{f}(p,q)$$

where further arbitrariness enters into the design of the functions $\mathcal{E}_1(p,q)$, $\mathcal{E}_2(p,q)$ and $\delta(p,q)$. The constructions

$$\begin{aligned}
\boldsymbol{E}_{p,q}(t,\boldsymbol{x}) &= \big\{\mathcal{E}_1(p,q)\boldsymbol{e}(p,q) + \mathcal{E}_2(p,q)e^{i\delta(p,q)}\boldsymbol{f}(p,q)\big\}\exp\Big\{i\Big[\omega t - \boldsymbol{k}(p,q)\boldsymbol{\cdot}\boldsymbol{x}\Big]\Big\}\\
\boldsymbol{B}_{p,q}(t,\boldsymbol{x}) &= \boldsymbol{k}(p,q)\times\boldsymbol{E}_{p,q}(t,\boldsymbol{x})\\
&= \big\{\mathcal{E}_1(p,q)\boldsymbol{f}(p,q) - \mathcal{E}_2(p,q)e^{i\delta(p,q)}\boldsymbol{e}(p,q)\big\}\exp\Big\{i\Big[\omega t - \boldsymbol{k}(p,q)\boldsymbol{\cdot}\boldsymbol{x}\Big]\Big\}
\end{aligned}$$

serve—in the approximation $(p^2+q^2)/2k \approx 0$—to associate a monochromatic polarized electromagnetic plane wave (propagating in the direction $\hat{\boldsymbol{k}}(p,q)$) with each point of (p,q)-space, and (431) describes a Gauss-weighted superposition of such plane waves. The "bewildering detail" to which I have referred arises (even in the simplest of the cases I have studied) when one undertakes to do the integration.

"Electromagnetic Gaussian beams" exist, by this account, in infinite variety. Evidently one must look to the finer particulars of laser design to discover how the physical device "selects among options," how to construct acceptable models of the laser beams encountered in laboratories.

The fields

$$\begin{aligned}
\boldsymbol{E}(t,\boldsymbol{x}) &= \iint_{-\infty}^{+\infty} \tfrac{\rho_0^2}{4\pi} e^{-\frac{1}{4}\rho_0^2(p^2+q^2)} \cdot \boldsymbol{E}_{p,q}(t,\boldsymbol{x})\,dpdq\\
\boldsymbol{B}(t,\boldsymbol{x}) &= \iint_{-\infty}^{+\infty} \tfrac{\rho_0^2}{4\pi} e^{-\frac{1}{4}\rho_0^2(p^2+q^2)} \cdot \boldsymbol{B}_{p,q}(t,\boldsymbol{x})\,dpdq
\end{aligned}$$

possess a property worthy of notice which I will expose by considering the superposition of only *two* electromagnetic plane waves. Let

$$\begin{aligned}
\boldsymbol{E}_1(t,\boldsymbol{x}) &= \boldsymbol{\mathcal{E}}_1\exp\Big\{i(\omega t - \boldsymbol{k}_1\boldsymbol{\cdot}\boldsymbol{x})\Big\} &&:\quad \boldsymbol{\mathcal{E}}_1\perp\boldsymbol{k}_1\\
\boldsymbol{B}_1(t,\boldsymbol{x}) &= \hat{\boldsymbol{k}}_1\times\boldsymbol{\mathcal{E}}_1\exp\Big\{i(\omega t - \boldsymbol{k}_1\boldsymbol{\cdot}\boldsymbol{x})\Big\}
\end{aligned}$$

describe one monochromatic plane wave, and

$$\begin{aligned}
\boldsymbol{E}_2(t,\boldsymbol{x}) &= \boldsymbol{\mathcal{E}}_2\exp\Big\{i(\omega t - \boldsymbol{k}_2\boldsymbol{\cdot}\boldsymbol{x}+\delta)\Big\} &&:\quad \boldsymbol{\mathcal{E}}_2\perp\boldsymbol{k}_2\\
\boldsymbol{B}_2(t,\boldsymbol{x}) &= \hat{\boldsymbol{k}}_2\times\boldsymbol{\mathcal{E}}_2\exp\Big\{i(\omega t - \boldsymbol{k}_2\boldsymbol{\cdot}\boldsymbol{x}+\delta)\Big\}
\end{aligned}$$

describe another. Let $\boldsymbol{E} = \boldsymbol{E}_1 + \boldsymbol{E}_2$ and $\boldsymbol{B} = \boldsymbol{B}_1 + \boldsymbol{B}_2$. Then

$$\boldsymbol{E}\cdot\boldsymbol{B} = \underbrace{\big\{\boldsymbol{\mathcal{E}}_1\cdot(\hat{\boldsymbol{k}}_2\times\boldsymbol{\mathcal{E}}_2) + \boldsymbol{\mathcal{E}}_2\cdot(\hat{\boldsymbol{k}}_1\times\boldsymbol{\mathcal{E}}_1)\big\}}_{\substack{= (\hat{\boldsymbol{k}}_1 - \hat{\boldsymbol{k}}_2)\cdot(\boldsymbol{\mathcal{E}}_1\times\boldsymbol{\mathcal{E}}_2)\\ \neq \boldsymbol{0}\text{ except under obvious special conditions}}} \exp\Big\{i(2\omega t - [\boldsymbol{k}_1 + \boldsymbol{k}_2]\cdot\boldsymbol{x} + \delta)\Big\}$$

shows that, in general, *superimposed plane waves do not share the* $\boldsymbol{E}\perp\boldsymbol{B}$ *condition characteristic of individual plane waves*. In particular: $\boldsymbol{E}\perp\boldsymbol{B}$ will not be found in the superpositions that produce "beams."

Let us look to a concrete example. Working from

$$\hat{\boldsymbol{k}}(p,q) \equiv k^{-1}\begin{pmatrix} p \\ q \\ k - [(p^2+q^2)/2k] \end{pmatrix} \quad \text{in the approximation } \left(\tfrac{p^2+q^2}{2k^2}\right)^2 \approx 0$$

we complete the dimensionless orthonormal triad by writing

$$\begin{aligned}
\boldsymbol{e}(p,q) &\equiv \tfrac{1}{\sqrt{p^2+q^2}}\begin{pmatrix} +q \\ -p \\ 0 \end{pmatrix}\\
\boldsymbol{f}(p,q) &\equiv \tfrac{1}{2k^2\sqrt{p^2+q^2}}\begin{pmatrix} p[2k^2-(p^2+q^2)] \\ q[2k^2-(p^2+q^2)] \\ -2k(p^2+q^2) \end{pmatrix} = \hat{\boldsymbol{k}}(p,q)\times\boldsymbol{e}(p,q)
\end{aligned}$$

and to achieve tractable integrals set

$$\begin{aligned}
\mathcal{E}_1(p,q) &= \mathcal{E}_1\ell\cdot\sqrt{p^2+q^2}\\
\mathcal{E}_2(p,q) &= \mathcal{E}_2\ell\cdot\sqrt{p^2+q^2}\\
\delta(p,q) &= \text{constant}
\end{aligned}$$

Here ℓ (introduced to cancel the physical dimension of $\sqrt{p^2+q^2}$) is a constant of arbitrary value and the dimensionality of length, so $\mathcal{E}_1\ell$ and $\mathcal{E}_2\ell$ have the dimensionality of electric potential. Working from (431) with $k = 2Z/\rho_0^2$ and

$$\begin{aligned}
\boldsymbol{\mathcal{E}}(p,q) &= \mathcal{E}_1\ell\begin{pmatrix} +q \\ -p \\ 0 \end{pmatrix} + \mathcal{E}_2\ell e^{i\delta}\begin{pmatrix} p[1-(p^2+q^2)/2k^2] \\ q[1-(p^2+q^2)/2k^2] \\ -(p^2+q^2)/k \end{pmatrix}\\
\boldsymbol{\mathcal{B}}(p,q) &= \mathcal{E}_1\ell\begin{pmatrix} p[1-(p^2+q^2)/2k^2] \\ q[1-(p^2+q^2)/2k^2] \\ -(p^2+q^2)/k \end{pmatrix} - \mathcal{E}_2\ell e^{i\delta}\begin{pmatrix} +q \\ -p \\ 0 \end{pmatrix}
\end{aligned}$$

we entrust the $\iint$'s to *Mathematica*, who supplies

$$\left.\begin{aligned}
\boldsymbol{E}(t,\boldsymbol{x}) &= \mathcal{E}_1\ell\ \boldsymbol{e}(t,\boldsymbol{x}) + \mathcal{E}_2\ell e^{i\delta}\,\boldsymbol{f}(t,\boldsymbol{x})\\
\boldsymbol{B}(t,\boldsymbol{x}) &= \mathcal{E}_1\ell\,\boldsymbol{f}(t,\boldsymbol{x}) - \mathcal{E}_2\ell e^{i\delta}\ \boldsymbol{e}(t,\boldsymbol{x})
\end{aligned}\right\} \tag{432}$$

with

$$\left.\begin{aligned} \boldsymbol{e}(t,\boldsymbol{x}) &= e^G\cdot\begin{pmatrix} -\mathcal{A}y\\ +\mathcal{A}x\\ 0\end{pmatrix}\\ \boldsymbol{f}(t,\boldsymbol{x}) &= e^G\cdot\begin{pmatrix} -\mathcal{B}x\\ -\mathcal{B}y\\ \mathcal{C}\end{pmatrix}\end{aligned}\right\} \tag{433}$$

where

$$e^G \equiv \exp\Big\{-\tfrac{x^2+y^2}{\rho^2}[1+i(z/Z)] + i[\omega t - kz]\Big\}$$

is familiar already (see again page 323) from the *scalar* theory of Gaussian beams, and where

$$\begin{aligned} \mathcal{A} &= \tfrac{2iZ^2}{\rho_0^2(Z-iz)^2}\\ &\equiv Ae^{i\alpha}\quad\text{with}\quad A = \sqrt{\tfrac{\{0\}^2+\{2Z^2\}^2}{\rho_0^4(Z^2+z^2)^2}}\\ \mathcal{B} &= \tfrac{-2Z\rho_0^2(z+iZ)+iZ^2[r^2-2(z+iZ)^2]}{\rho_0^2(z+iZ)^4}\\ &\equiv Be^{i\beta}\quad\text{with}\quad B = \sqrt{\tfrac{\{\text{stuff}\}^2+\{\text{more stuff}\}^2}{\rho_0^4(Z^2+z^2)^4}}\\ \mathcal{C} &= \tfrac{2Z^2r^2-2\rho_0^2Z(Z-iz)}{\rho_0^2(Z-iz)^3}\\ &\equiv Ce^{i\gamma}\quad\text{with}\quad C = \sqrt{\tfrac{\{\text{stuff}\}^2+\{\text{more stuff}\}^2}{\rho_0^4(Z^2+z^2)^3}}\end{aligned}$$

I have indicated how the invariable reality of A, B and C comes about, but have omitted details too complicated to be informative, and have also omitted (as irrelevant to the purposes at hand) explicit description of the phase factors α, β and γ (which could be expressed as the arctangents of the obvious ratios). Notice that the functions described above depend upon x and y only through $r^2 \equiv x^2+y^2$; they are, in short, *axially symmetric*. Notice also that $[A]=[B]=(\text{length})^{-2}$ while $[C]=(\text{length})^{-1}$.

Returning now with (433) to (432), we have $\boldsymbol{E}(t,\boldsymbol{x}) = \boldsymbol{E}_1(t,\boldsymbol{x}) + \boldsymbol{E}_2(t,\boldsymbol{x})$ with

$$\begin{aligned} \boldsymbol{E}_1 &= \mathcal{E}_1\ell e^{-(r/\rho)^2}\cdot e^{i\{(\omega t-kz)-(r/\rho)^2(z/Z)\}}\left\{ Ae^{i\alpha}\begin{pmatrix} -y\\ +x\\ 0\end{pmatrix}\right\}\\ \boldsymbol{E}_2 &= \mathcal{E}_2 e^{i\delta}\ell e^{-(r/\rho)^2}\cdot e^{i\{(\omega t-kz)-(r/\rho)^2(z/Z)\}}\left\{ Be^{i\beta}\begin{pmatrix} -x\\ -y\\ 0\end{pmatrix} + Ce^{i\gamma}\begin{pmatrix} 0\\ 0\\ 1\end{pmatrix}\right\}\end{aligned}$$

The associated magnetic fields are

$$\boldsymbol{B}_1 = \quad \mathcal{E}_1\ell e^{-(r/\rho)^2}\cdot e^{i\{(\omega t-kz)-(r/\rho)^2(z/Z)\}}\left\{\quad Be^{i\beta}\begin{pmatrix}-x\\-y\\0\end{pmatrix}+Ce^{i\gamma}\begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$$

$$\boldsymbol{B}_2 = \mathcal{E}_2 e^{i\delta}\ell e^{-(r/\rho)^2}\cdot e^{i\{(\omega t-kz)-(r/\rho)^2(z/Z)\}}\left\{-Ae^{i\alpha}\begin{pmatrix}-y\\+x\\0\end{pmatrix}\right\}$$

It is understood that to extract the *physical* fields, and before we assemble such quadratic constructions as (field)$\cdot$(field) and (field)$\times$(field), we must make the replacements

$$e^{i(\text{stuff})}\longmapsto\cos(\text{stuff})$$

That done, we obtain finally

$$\boldsymbol{E}_1 = \mathcal{E}_1\ell e^{-(r/\rho)^2}\left\{\quad A\cos(\vartheta+\alpha)\begin{pmatrix}-y\\+x\\0\end{pmatrix}\right\}$$

$$\boldsymbol{E}_2 = \mathcal{E}_2\ell e^{-(r/\rho)^2}\left\{\quad B\cos(\vartheta+\beta+\delta)\begin{pmatrix}-x\\-y\\0\end{pmatrix}+C\cos(\vartheta+\gamma+\delta)\begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$$

$$\boldsymbol{B}_1 = \mathcal{E}_1\ell e^{-(r/\rho)^2}\left\{\quad B\cos(\vartheta+\beta)\begin{pmatrix}-x\\-y\\0\end{pmatrix}\quad +C\cos(\vartheta+\gamma)\begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$$

$$\boldsymbol{B}_2 = \mathcal{E}_2\ell e^{-(r/\rho)^2}\left\{-A\cos(\vartheta+\alpha+\delta)\begin{pmatrix}-y\\+x\\0\end{pmatrix}\right\}$$

with $\vartheta \equiv \omega t - kz - (r/\rho)^2(z/Z)$. It is immediately evident that at every spacetime point

$$\begin{aligned}\boldsymbol{E}_1 \perp \boldsymbol{E}_2\,,\quad \boldsymbol{B}_1 \perp \boldsymbol{B}_2\\ \boldsymbol{E}_1 \perp \boldsymbol{B}_1\,,\quad \boldsymbol{E}_2 \perp \boldsymbol{B}_2\end{aligned}$$

but from

$$\begin{aligned}\boldsymbol{E}\cdot\boldsymbol{B} &= (\boldsymbol{E}_1+\boldsymbol{E}_2)\cdot(\boldsymbol{B}_1+\boldsymbol{B}_2)\\ &= \mathcal{E}_1\mathcal{E}_2\ell^2 e^{-2(r/\rho)^2}\Big\{r^2\big[B^2\cos(\vartheta+\beta)\cos(\vartheta+\beta+\delta)\\ &\qquad\qquad - A^2\cos(\vartheta+\alpha)\cos(\vartheta+\alpha+\delta)\big]\\ &\qquad\qquad + C^2\cos(\vartheta+\gamma)\cos(\vartheta+\gamma+\delta)\Big\}\\ &\neq 0 \text{ except under non-obvious special conditions: note, however, that}\\ &\downarrow\\ &= 0 \quad \text{as } r\to\infty \text{ because the fields die at points far from the beam axis}\end{aligned}$$

we see that—consistently with the remark developed on page 332—the net fields $\boldsymbol{E}$ and $\boldsymbol{B}$ are typically *not* perpendicular: at axial points ($r = 0$) they are, in fact, *parallel*! The *energy flux* and *momentum density* at the spacetime point are proportional to

$$\boldsymbol{E}\times\boldsymbol{B} = (\boldsymbol{E}_1+\boldsymbol{E}_2)\times(\boldsymbol{B}_1+\boldsymbol{B}_2) \equiv e^{-2(r/\rho)^2}\cdot\boldsymbol{F}$$

where according to *Mathematica*

$$\begin{aligned}
F_1 &= xAC\big[\mathcal{E}_1^2\cos(\vartheta+\alpha)\cos(\vartheta+\gamma)+\mathcal{E}_2^2\cos(\vartheta+\alpha+\delta)\cos(\vartheta+\gamma+\delta)\big]\\
&\quad -yBC\mathcal{E}_1\mathcal{E}_2\big[\cos(\vartheta+\beta+\delta)\cos(\vartheta+\gamma)-\cos(\vartheta+\beta)\cos(\vartheta+\gamma+\delta)\big]\\
F_2 &= yAC\big[\mathcal{E}_1^2\cos(\vartheta+\alpha)\cos(\vartheta+\gamma)+\mathcal{E}_2^2\cos(\vartheta+\alpha+\delta)\cos(\vartheta+\gamma+\delta)\big]\\
&\quad +xBC\mathcal{E}_1\mathcal{E}_2\big[\cos(\vartheta+\beta+\delta)\cos(\vartheta+\gamma)-\cos(\vartheta+\beta)\cos(\vartheta+\gamma+\delta)\big]\\
F_3 &= r^2AB\big[\mathcal{E}_1^2\cos(\vartheta+\alpha)\cos(\vartheta+\beta)+\mathcal{E}_2^2\cos(\vartheta+\alpha+\delta)\cos(\vartheta+\beta+\delta)\big]
\end{aligned}$$

This is of the design

$$\begin{aligned}
\boldsymbol{F} &= a\begin{pmatrix}x\\y\\0\end{pmatrix}+b\begin{pmatrix}-y\\+x\\0\end{pmatrix}+c\begin{pmatrix}0\\0\\1\end{pmatrix}\\
&= \boldsymbol{F}_{\text{radial}}+\boldsymbol{F}_{\text{tangential}}+\boldsymbol{F}_{\text{axial}}
\end{aligned}$$

where the vectors $\boldsymbol{F}_{\text{radial}}$ stand normal to the z-axis (beam-axis) and are of constant magnitude on circles concentric about that axis, the vectors $\boldsymbol{F}_{\text{tangential}}$ are (also constant on but) tangent to such circles and have ↺ or ↻ handedness according as $b \gtrless 0$, and the vectors $\boldsymbol{F}_{\text{axial}}$ (also constant on such circles) run parallel to the z-axis. The "constants" a, b and c are in fact horribly complicated functions of the variables $\{t, z, r\}$ and of the parameters $\{\omega, \rho_0, Z, \mathcal{E}_1, \mathcal{E}_2, \delta\}$.

We are in position now to state that the momentary *momentum density of the beam field* at any designated point $\boldsymbol{x}$ can be described (see again page 216)

$$\boldsymbol{\mathcal{P}} = \tfrac{1}{c}e^{-2(r/\rho)^2}\boldsymbol{F}$$

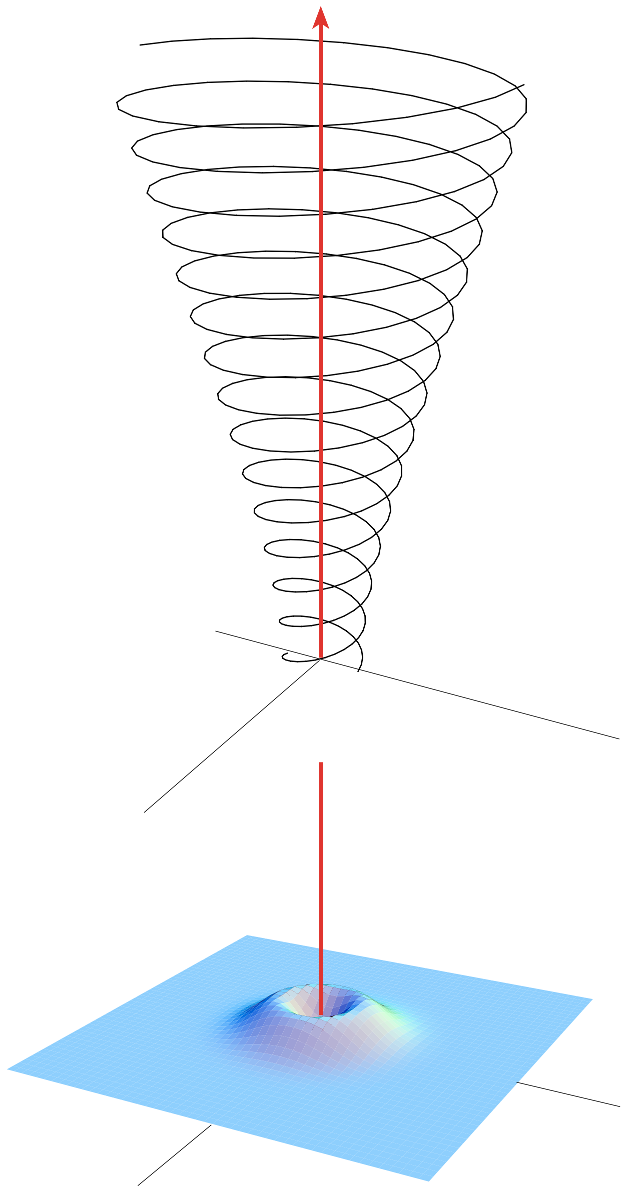
We observe that

- $\boldsymbol{\mathcal{P}}$ vanishes far from the beam axis because of Gaussian attenuation
- $\boldsymbol{\mathcal{P}}$ vanishes *on* the beam axis by the design of a, b and c
- <u>field momentum traces a divergent spiral</u> in the near neighborhood of the beam axis unless $b = 0$.

The *angular momentum density* of the beam field is given by[262]

$$\begin{aligned}
\boldsymbol{\mathcal{L}} = \boldsymbol{x}\times\boldsymbol{\mathcal{P}} &= \tfrac{1}{c}e^{-2(r/\rho)^2}\left\{-bz\begin{pmatrix}x\\y\\0\end{pmatrix}+(az-c)\begin{pmatrix}-y\\+x\\0\end{pmatrix}+br^2\begin{pmatrix}0\\0\\1\end{pmatrix}\right\}\\
&= \boldsymbol{\mathcal{L}}_{\text{radial}}+\boldsymbol{\mathcal{L}}_{\text{tangential}}+\boldsymbol{\mathcal{L}}_{\text{axial}}
\end{aligned}$$

[262] Don't be confused by the fact that c is used here to mean two entirely different things.



FIGURES 103 & 104: *The upper figure portrays the spiroform deployment of the momentum in the electromagnetic field of the Gaussian beam described in the text. Displayed below is the resulting angular momentum density (presented as a function of x and y at the beam waist: $z = 0$). The figures show that/why it makes sense to say that "the angular momentum lives at the fringes of the beam."*

The first two components (by an elementary symmetry argument) can make no net contribution to the total angular momentum of the beam, which is given therefore by

$$\boldsymbol{L} = L\begin{pmatrix}0\\0\\1\end{pmatrix} \quad \text{with} \quad L = \iiint \tfrac{1}{c}e^{-2(r/\rho)^2}br^2\,dxdydz$$

We notice that L vanishes if $b = 0$, and that this happens when $\delta = 0$, for in the latter circumstance the equations at near the top of page 336 assume the much-simplified form

$$\begin{aligned}
F_1 &= \ \ xAC\big[\mathcal{E}_1^2 + \mathcal{E}_2^2\big]\cos(\vartheta+\alpha)\cos(\vartheta+\gamma) + \text{no } y\text{-term}\\
F_2 &= \ \ yAC\big[\mathcal{E}_1^2 + \mathcal{E}_2^2\big]\cos(\vartheta+\alpha)\cos(\vartheta+\gamma) + \text{no } x\text{-term}\\
F_3 &= r^2AB\big[\mathcal{E}_1^2 + \mathcal{E}_2^2\big]\cos(\vartheta+\alpha)\cos(\vartheta+\beta)
\end{aligned}$$

One occasionally encounters the claim that "the angular momentum transported by a laser beam lives at the fringes of the beam," but in support of that claim authors who possess only a *scalar* theory of beams must argue rather vaguely that

i) beam angular momentum must arise from momentum circulation

ii) there can be no circulation *at the axis* of an axially-symmetric beam

iii) all $\boldsymbol{\mathcal{P}}$-circulation must therefore occur between the axis and the remote regions where the $\boldsymbol{E}$ and $\boldsymbol{B}$ fields have fallen off to zero—in short: "at the fringes" of the beam.

My effort has been to carry a *vector* theory of beams far enough to illuminate the details of the matter. Having achieved that objective, I must be content now to abandon my little "electromagnetic theory of beams"... but feel an obligation to list some of the respects in which the theory remains incomplete:

- It should be feasible (by the method sketched on page 321) to turn such beams on and off; *i.e.*, to construct laterally confined quasi-monochromatic Gaussian wavepackets—"classical photons," if you will.
- It should be feasible, moreover, to construct trains of such wavepackets, and to describe the coherence/polarization properties of such trains.
- One would like to be in position to describe the energy, momentum and angular momentum transported by such a "classical photon," and to identify conditions under which they stand in the quantum relationships

$$E = cP = \omega L$$

- To that end one would need to clarify certain salient properties of and interrelationships among the complicated functions a, b and c.
- Identical values of $\mathcal{E}_1$, $\mathcal{E}_2$ and δ were assigned to each of the plane waves from which our Gaussian beams were assembled. Do the Stokes parameters implicit (by (399)) in $\{\mathcal{E}_1, \mathcal{E}_2, \delta\}$ speak usefully about the polarization properties of the *assembled* beam?

It should be borne always in mind that the theory sketched above proceeds from an **inoffensive approximation** (page 322) and—within the bounds of that approximation—from a **convenient specification** (page 333) of the manner in which orthonormal vectors will be attached to $\hat{\boldsymbol{k}}$-vectors (and weighted). The theory is rich enough to support easily the notion of "higher beam modes," but this question remains open: *Is the theory—as I suspect—rich enough to account for the observed properties of the optical beams encountered in laboratories*?