united nations educational, scientific and cultural organization

international atomic energy agency

the
abdus salam
international centre for theoretical physics

Winter College on Optics and Photonics
7 - 25 February 2000

1218-18

"Polarization of Light"

F. GORI
Dipartimento di Fisica
Università di Roma Tre
Italy

ICTP/ICO/OSA Winter College

on Optics and Photonics

Polarization of light

Franco Gori

Dipartimento di Fisica

Università di Roma Tre

Miramare - Trieste, Italy, 7-25 February 2000

The monochromatic case

Strictly monochromatic light is always fully polarized

$$\begin{cases} E_x = A_x \cos(\varphi_x - \omega t) \\ E_y = A_y \cos(\varphi_y - \omega t) \end{cases} \rightarrow \begin{cases} E_x = A_x e^{i(\varphi_x - \omega t)} \\ E_y = A_y e^{i(\varphi_y - \omega t)} \end{cases}$$

$$\rightarrow \begin{cases} E_x = A_x e^{i\varphi_x} \\ E_y = A_y e^{i\varphi_y} \end{cases}$$

Jones vectors

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix} = \begin{pmatrix} A_x e^{i\varphi_x} \\ A_y e^{i\varphi_y} \end{pmatrix}$$

Examples

Linearly polarized light

$$\begin{pmatrix} E_x \\ 0 \end{pmatrix} \quad \text{along } x-\text{axis}$$

$$\begin{pmatrix} 0 \\ E_y \end{pmatrix} \quad \text{along } y-\text{axis}$$

Linearly polarized light

$$\begin{pmatrix} A \\ A \end{pmatrix} \quad \pi/4 \text{ with respect to the } x-\text{axis}$$

$$\begin{pmatrix} A\sqrt{3} \\ A \end{pmatrix} \quad \pi/6 \text{ with respect to the } y-\text{axis}$$

Circularly polarized light

$$\begin{pmatrix} A \\ iA \end{pmatrix} \quad \text{Right}-\text{handed}$$

$$\begin{pmatrix} A \\ -iA \end{pmatrix} \quad \text{Left}-\text{handed}$$

Basis vectors

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix} = E_x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + E_y \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Scalar product between $\begin{pmatrix} E_x \\ E_y \end{pmatrix}$ and $\begin{pmatrix} E_x' \\ E_y' \end{pmatrix}$

$$\begin{pmatrix} E_x^* & E_y^* \end{pmatrix} \begin{pmatrix} E_x' \\ E_y' \end{pmatrix} = E_x^* E_x' + E_y^* E_y'$$

Change of basis

$$\begin{pmatrix} p \\ q \end{pmatrix}, \begin{pmatrix} u \\ v \end{pmatrix} \qquad \{pv - qu \neq 0\}$$

$$\begin{pmatrix} p \\ q \end{pmatrix} = p\begin{pmatrix} 1 \\ 0 \end{pmatrix} + q\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} u \\ v \end{pmatrix} = u\begin{pmatrix} 1 \\ 0 \end{pmatrix} + v\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{pv - qu}\left[v\begin{pmatrix} p \\ q \end{pmatrix} - q\begin{pmatrix} u \\ v \end{pmatrix}\right]$$

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{pv - qu}\left[-u\begin{pmatrix} p \\ q \end{pmatrix} + p\begin{pmatrix} u \\ v \end{pmatrix}\right]$$

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix} = \frac{1}{pv - qu}\left\{\left(vE_x - uE_y\right)\begin{pmatrix} p \\ q \end{pmatrix} + \left(-qE_x + pE_y\right)\begin{pmatrix} u \\ v \end{pmatrix}\right\}$$

$$E_{pq} = \frac{1}{pv - qu}\left(vE_x - uE_y\right);\ E_{uv} = \frac{1}{pv - qu}\left(-qE_x + pE_y\right)$$

A first use of matrices

$$\begin{pmatrix} E_{pq} \\ E_{uv} \end{pmatrix} = T \begin{pmatrix} E_x \\ E_y \end{pmatrix}$$

$$T = \frac{1}{pv - qu} \begin{pmatrix} v & -u \\ -q & p \end{pmatrix}$$

Example: from $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ to $\begin{pmatrix} 1 \\ i \end{pmatrix}, \begin{pmatrix} 1 \\ -i \end{pmatrix}$

$$T = \frac{1}{2} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}$$

Further uses of matrices

Rotation of the x,y frame by an angle α

$$\begin{pmatrix} E_x' \\ E_y' \end{pmatrix} = \begin{pmatrix} C & -S \\ S & C \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix} \qquad \begin{Bmatrix} C = \cos\alpha \\ S = \sin\alpha \end{Bmatrix}$$

$$R = \begin{pmatrix} C & S \\ -S & C \end{pmatrix}$$

A rotator

$$\begin{pmatrix} E_x' \\ E_y' \end{pmatrix} = \begin{pmatrix} C & -S \\ S & C \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix} \qquad \begin{Bmatrix} C = \cos\alpha \\ S = \sin\alpha \end{Bmatrix}$$

$$M_R = \begin{pmatrix} C & -S \\ S & C \end{pmatrix}$$

A linear polarizer set along the x-axis

$$\begin{pmatrix} E_x' \\ E_y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix}$$

$$P_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

A linear polarizer set at an angle α with respect to the x-axis

$$P_\alpha = R^{-1} P_0 R = \begin{pmatrix} C^2 & CS \\ CS & S^2 \end{pmatrix}$$

Wave plates

$$\begin{pmatrix} E_x' \\ E_y' \end{pmatrix} = \begin{pmatrix} e^{i\delta_x} & 0 \\ 0 & e^{i\delta_y} \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix}$$

$$M_0 = \begin{pmatrix} e^{i\delta_x} & 0 \\ 0 & e^{i\delta_y} \end{pmatrix}$$

Examples

Quarter wave plate

$$M_Q = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$$

Half wave plate

$$M_H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Rotated wave plate

$$M_\alpha = R^{-1} M R$$

Matrices associated to rotators and wave plates form a group

Simon-Mukunda gadget

"Any wave plate or rotator can be synthesized by using two QWPs and a HWP with suitable orientations"

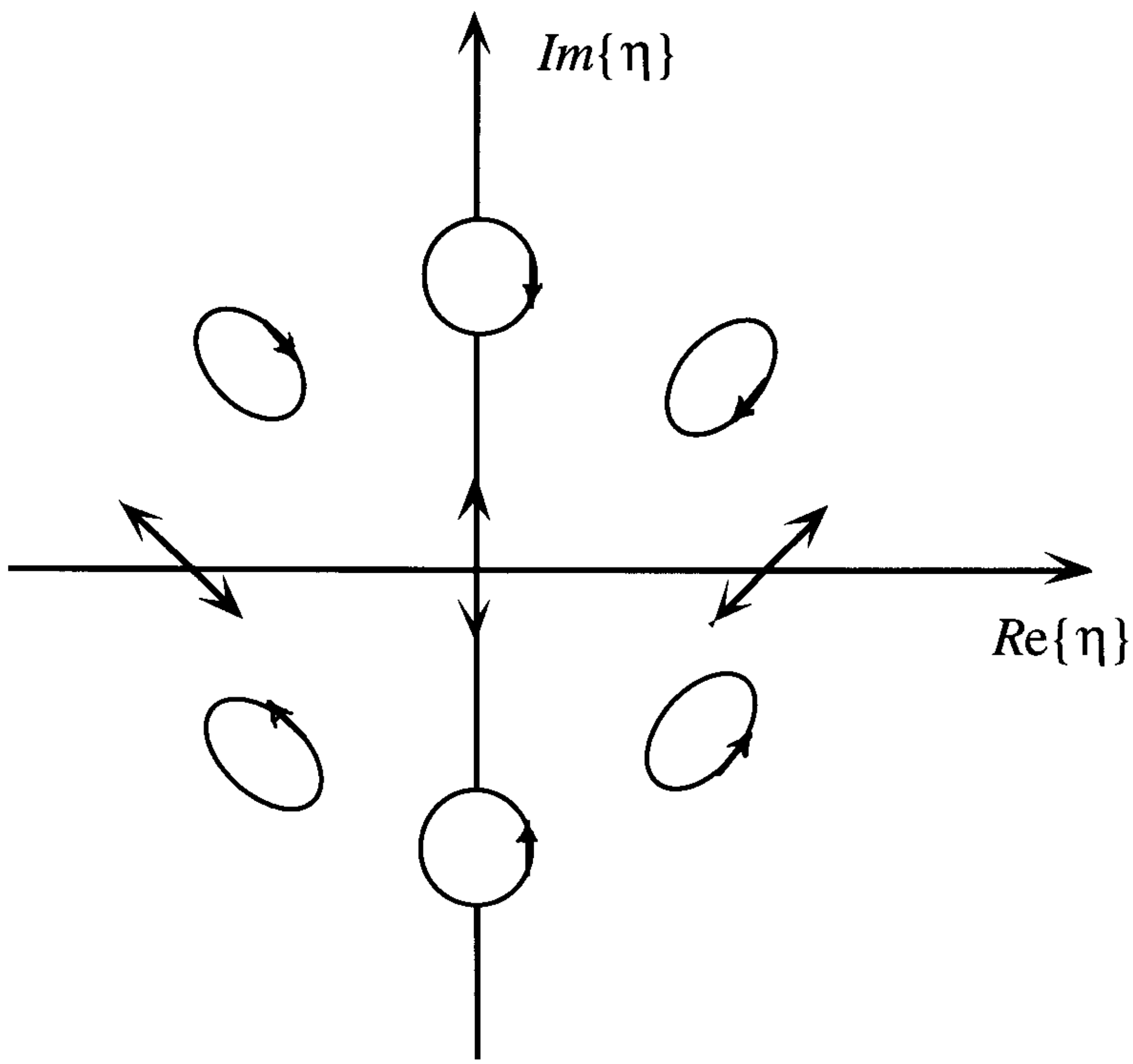
R. Simon and N. Mukunda, Phys. Lett. **140A** (1990) 165

V. Bagini, R. Borghi, F. Gori, M. Santarsiero, F. Frezza, G. Schettini, G. Schirripa Spagnolo, "The Simon Mukunda polarization gadget", European Journal of Physics, **17** (1996) 279

Complex plane representation

of polarized light

$$\eta = \frac{E_x}{E_y} = \frac{A_x}{A_y} e^{i(\varphi_x - \varphi_y)}$$



Normalized vectors

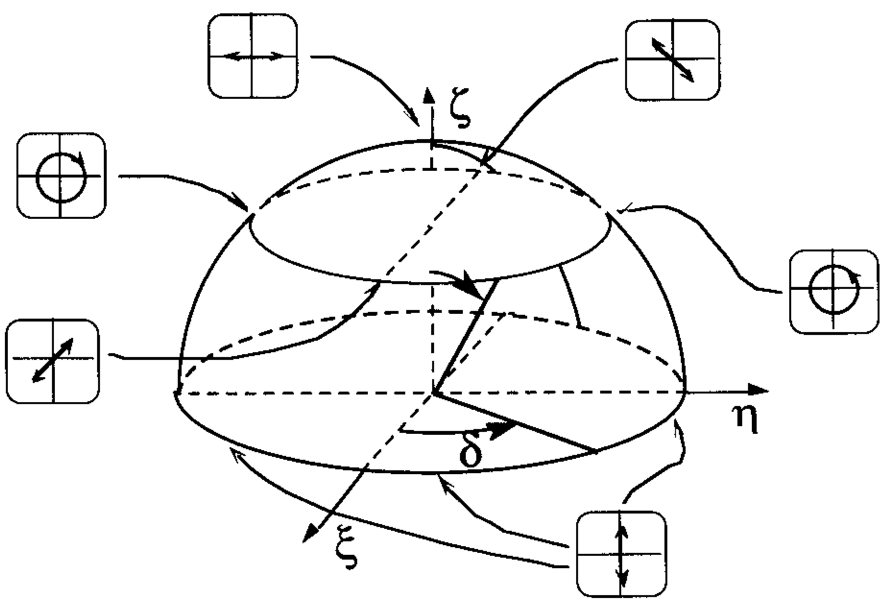
$$A_x^2 + A_y^2 = 1$$

We could set

$$A_x = \cos\gamma \qquad A_y = \sin\gamma$$

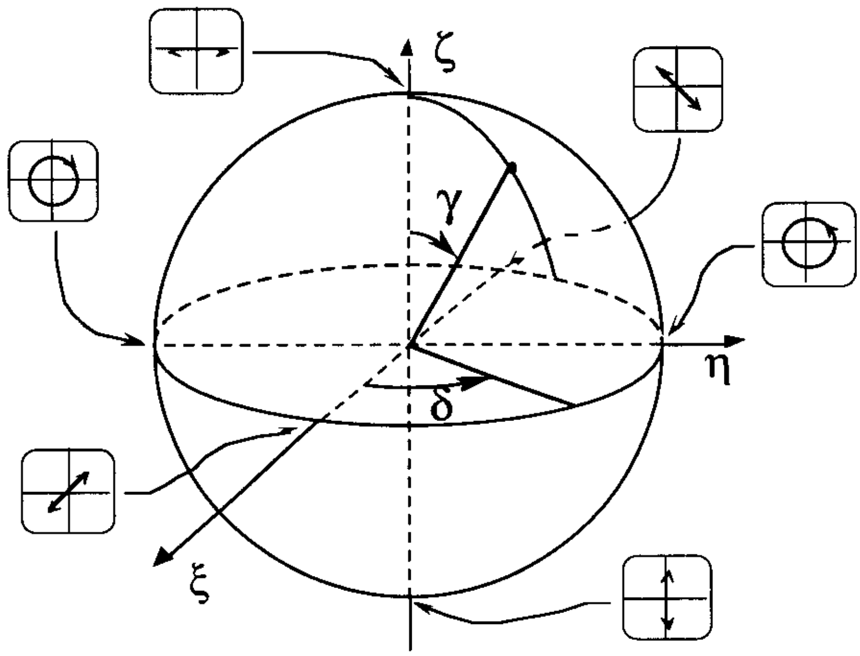
$$\begin{pmatrix} E_x \\ E_y \end{pmatrix} \rightarrow \begin{pmatrix} \cos\gamma \\ \sin\gamma \, e^{i\delta} \end{pmatrix} \qquad \{\delta = \varphi_y - \varphi_x\}$$

Polarization states mapped on a hemisphere



Polarization states mapped on a sphere

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix} \rightarrow \begin{pmatrix} \cos\gamma/2 \\ \sin\gamma/2 \; e^{i\delta} \end{pmatrix}$$



Note that antipodal points correspond to orthogonal states.

This gives a connection with spinor formalism

Stokes parameters

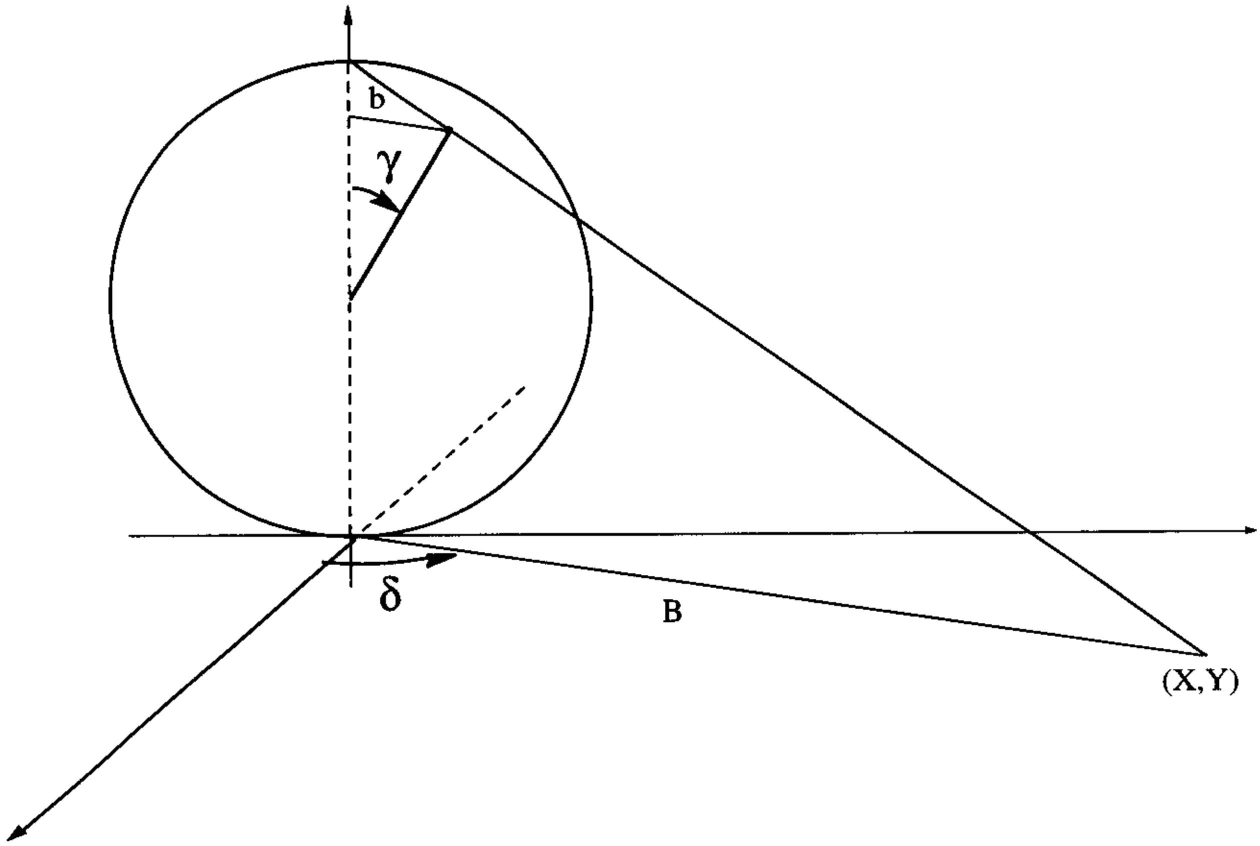
$$s_0 = A_x^2 + A_y^2$$
$$s_1 = A_x^2 - A_y^2$$
$$s_2 = 2A_xA_y \cos\delta$$
$$s_3 = -2A_xA_y \sin\delta$$

For the mapping on the sphere (s_0=1)

$$x = \sin\gamma\cos\delta = 2\sin\frac{\gamma}{2}\cos\frac{\gamma}{2}\cos\delta = s_2$$
$$y = \sin\gamma\sin\delta = 2\sin\frac{\gamma}{2}\cos\frac{\gamma}{2}\sin\delta = -s_3$$
$$x = \cos\gamma = \cos^2\frac{\gamma}{2} - \sin^2\frac{\gamma}{2} = s_1$$

For the original Poincaré sphere: $x = s_1$; $y = s_2$; $z = s_3$

Stereographic projection



$$b = \sin\gamma; \quad \frac{b}{B} = \frac{1-\cos\gamma}{2} = \sin^2\frac{\gamma}{2}$$

$$B = 2\frac{\cos\gamma/2}{\sin\gamma/2}$$

$$X = 2\frac{\cos\gamma/2}{\sin\gamma/2}\cos\delta = 2\,\mathrm{Re}\left\{\frac{E_x^*}{E_y^*}\right\}$$

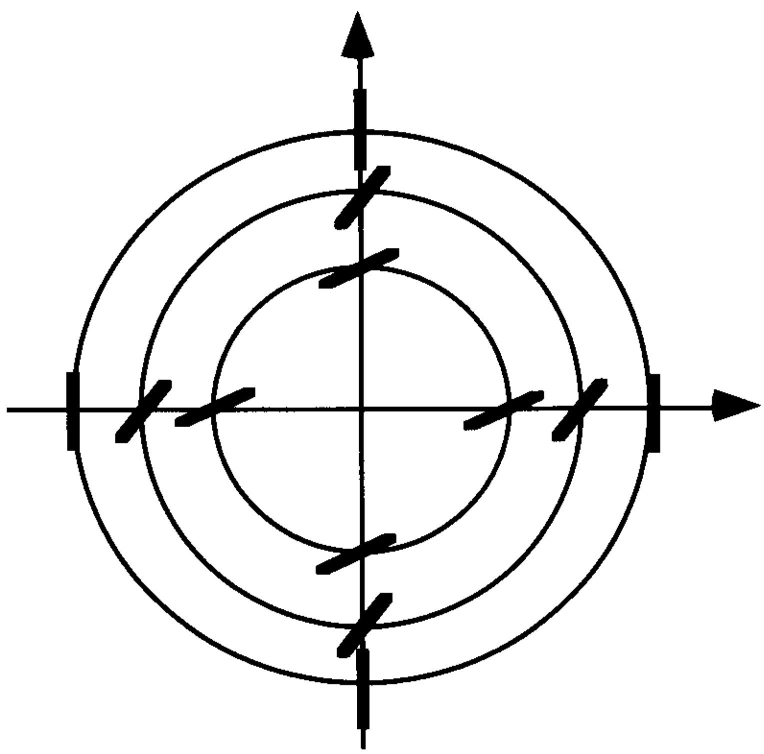
$$Y = 2\frac{\cos\gamma/2}{\sin\gamma/2}\sin\delta = 2\,\mathrm{Im}\left\{\frac{E_x^*}{E_y^*}\right\}$$

Non-uniform polarization

$$\begin{pmatrix}1\\ i\end{pmatrix} + e^{i\varphi}\begin{pmatrix}1\\ -i\end{pmatrix} = 2e^{i\varphi/2}\begin{pmatrix}\cos\varphi/2\\ \sin\varphi/2\end{pmatrix}$$

Example: two spherical waves with different curvatures

$$\varphi \propto \left(x^2 + y^2\right)$$



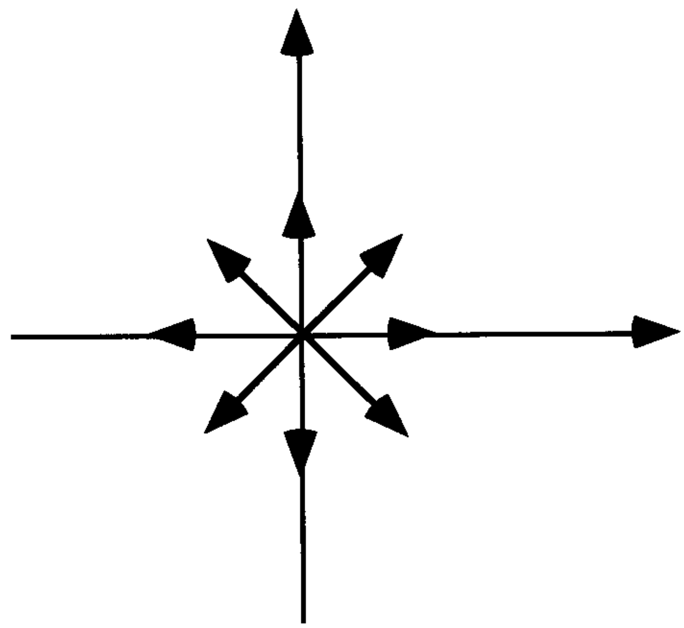
Vortex fields

Superposing two counterrotating circularly polarized beams possessing single-vortices with opposite charges.

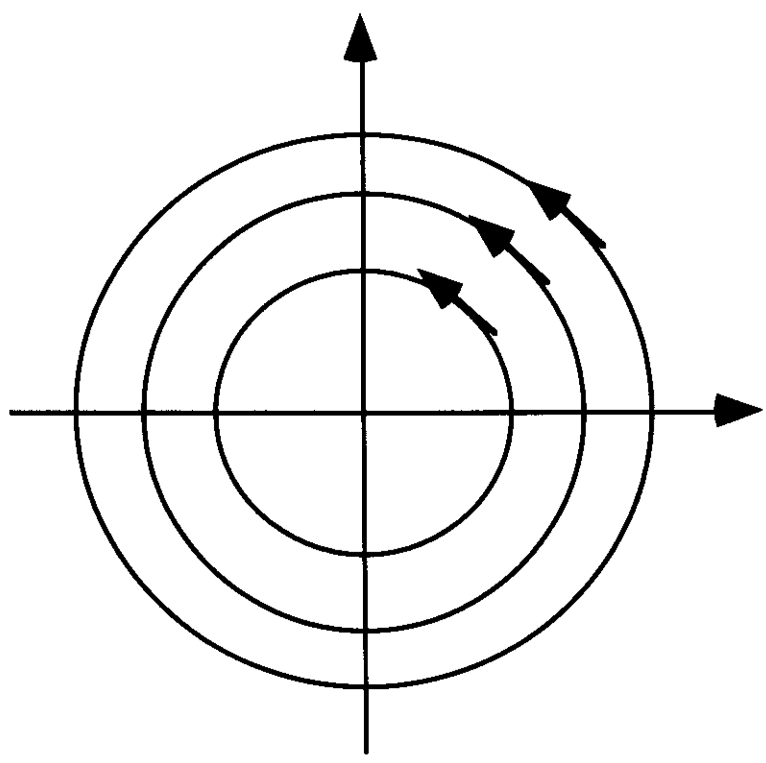
$$\begin{pmatrix} f(r)e^{-i(\vartheta+\alpha)} \\ if(r)e^{-i(\vartheta+\alpha)} \end{pmatrix} + \begin{pmatrix} f(r)e^{i(\vartheta+\alpha)} \\ -if(r)e^{i(\vartheta+\alpha)} \end{pmatrix} = 2f(r)\begin{pmatrix} \cos(\vartheta+\alpha) \\ \sin(\vartheta+\alpha) \end{pmatrix}$$

Note that no vortices are now present within the beam section.

$\alpha = 0$: radial polarization



$\alpha = \pi/2$: azimuthal polarization



Propagation of radially or azimuthally polarized fields

For a vortex field we have

$$z = 0: \quad f(r)e^{\pm i\vartheta}$$

$$z > 0: \quad g(r,\vartheta,z) = -\frac{k e^{\frac{ikr^2}{2z}} e^{\pm i\vartheta}}{z} \int_0^\infty f(\rho) e^{\frac{ik\rho^2}{2z}} J\left(\frac{k\rho r}{z}\right) \rho d\rho$$

Radially or azimuthally polarized fields, being the superposition of vortex fields, obey the same law of propagation.

Quasi-monochromatic case

$$E_x(t) = A_x(t)e^{i[\delta_x(t)-\omega t]}$$

$$E_y(t) = A_y(t)e^{i[\delta_y(t)-\omega t]}$$

Polarization matrix

$$J = \begin{bmatrix} J_{xx} & J_{xy} \\ J_{yx} & J_{yy} \end{bmatrix}$$

$$J_{\alpha\beta} = \left\langle E_\alpha^*(t)E_\beta(t) \right\rangle, \qquad (\alpha, \beta = x, y)$$

$$\mu_{xy} = \frac{J_{xy}}{\sqrt{J_{xx}J_{yy}}} = \left|\mu_{xy}\right|e^{i\varepsilon}$$

$$\left|J_{xy}\right| \le \sqrt{J_{xx}J_{yy}}$$

$$0 \le \left|\mu_{xy}\right| \le 1$$

Completely polarized field

$$J = \begin{pmatrix} A_x^2 & A_x A_y e^{i\delta} \\ A_x A_y e^{-i\delta} & A_y^2 \end{pmatrix}$$

Completely unpolarized field

$$J = \begin{pmatrix} I/2 & 0 \\ 0 & I/2 \end{pmatrix}$$

How to estimate polarization with a single number ($|\mu_{xy}|$ is not enough)

Note: On superposing two uncorrelated fields with matrix elements $J_{\alpha\beta}^{(1)}$, $J_{\alpha\beta}^{(2)}$ we have

$$J_{\alpha\beta} = J_{\alpha\beta}^{(1)} + J_{\alpha\beta}^{(2)}$$

Any polarization matrix can be decomposed as follows

$$\begin{pmatrix} J_{xx} & J_{xy} \\ J_{yx} & J_{yy} \end{pmatrix} = \begin{pmatrix} K & 0 \\ 0 & K \end{pmatrix} + \begin{pmatrix} A_x^2 & A_x A_y e^{-i\delta} \\ A_x A_y e^{-i\delta} & A_y^2 \end{pmatrix}$$

$$\begin{aligned} I_{pol} &= A_x^2 + A_y^2 \\ I_{unpol} &= 2K \\ I_{tot} &= I_{pol} + I_{unpol} \end{aligned}$$

Degree of polarization

$$P = \frac{I_{pol}}{I_{tot}} = \frac{A_x^2 + A_y^2}{A_x^2 + A_y^2 + 2K}$$

$$P = \frac{\sqrt{\left(J_{xx} - J_{yy}\right)^2 + 4\left|J_{xy}\right|^2}}{J_{xx} + J_{yy}}$$

$$P = \sqrt{1 - \frac{4\det(J)}{[\mathrm{Tr}(J)]^2}}$$

Stokes parameters

They can be defined for partially polarized light too

$$
\begin{aligned}
s_0 &= \left\langle |E_x|^2 \right\rangle + \left\langle |E_y|^2 \right\rangle = J_{xx} + J_{yy} \\
s_1 &= \left\langle |E_x|^2 \right\rangle + \left\langle |E_y|^2 \right\rangle = J_{xx} - J_{yy} \\
s_2 &= 2\,\mathrm{Re}\left\{ \left\langle E_x^* E_y \right\rangle \right\} = 2\left|J_{xy}\right| \cos\varepsilon \\
s_3 &= 2\,\mathrm{Im}\left\{ \left\langle E_x^* E_y \right\rangle \right\} = -2\left|J_{xy}\right| \sin\varepsilon
\end{aligned}
$$

A possible representation of partially polarized light is then through colums vectors of the form

$$
\begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \end{bmatrix}
$$

The effects of anisotropic optical elements could then be described by suitable 4 x 4 matrices known as Mueller matrices.

General references

M. Born and E. Wolf, Principles of Optics, 7-th edition, (Cambridge University Press, Cambridge, England, 1999)

L. Mandel and E. Wolf, Optical Cooherence and Quantum Optics, Cambridge University Press, Cambridge, England, 1995)

C. Brosseau, Fundamentals of polarized light (John Wiley & Sons, New York, 1998)

R. M. A. Azzam and N. M. Bashara, Ellipsometry and polarized light (North-Holland, Amsterdam, 1989)

Partially coherent beams with partial polarization.
A matrix treatment.

Suppose we have two beams. We are told that they are different and we have to find out where the difference lies. We are also told that both beams are quasi-monochromatic with the same spectrum and that the difference has not to do with photon statistics.

How shall we proceed?

Presumably, our first attempt could be to get information about the spatial distribution of the beam power. For example, we measure the M^2-factor of the beams, or in a more complete way, we measure the intensity distribution across as many transverse planes as possible. With some disappointment we will discover that distinguishing the two beams is not that simple. They produce the same intensity distribution everywhere.

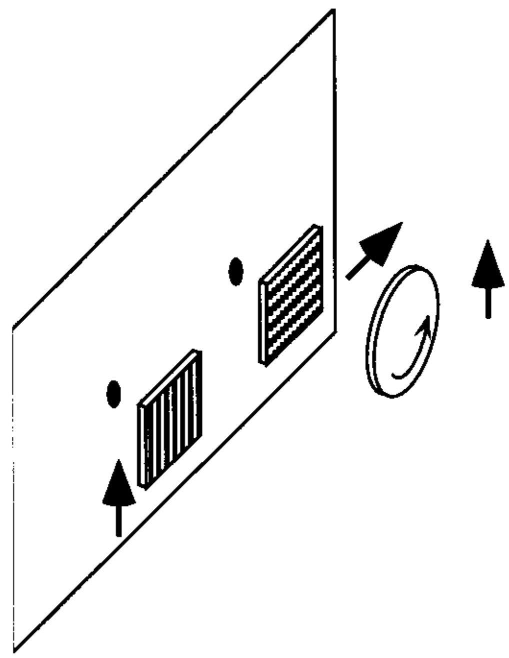
We could suspect that the difference is about polarization properties. Accordingly, equipped with a linear polarizer and a set of wave plates, we try to determine the polarization states of the beams. We will be disappointed again. Both beams are simply unpolarized.

In all probability, our next step will be to inquire about the spatial coherence properties of the beams. So we pass to the use of a Young interferometer and measure the degree of coherence for a sufficiently high number of point pairs. The result we find is that the beams are partially coherent from the spatial standpoint. In polar coordinates the degree of coherence is described by a function of the form

$$\mu_{12} = \cos(\vartheta_1 - \vartheta_2)$$

However, we find that this formula holds equally good for both beams. Consequently, they are still indistinguishable.

There is not much left. After further thinking we decide to set up an experiment like this



Each aperture of the Young interferometer is covered by a linear polarizer, the two polarizers being crossed. In addition one of the polarizers is followed by a $\pi/2$ rotator so that the outcoming fields can interfere (if correlated).

Here, at last we catch the difference. For one of the beams we find that no fringes appear, irrespective of the position of the two holes. For the other one, fringes can be seen and the associated degree of correlation turns out to be

$$\mu_{12} = \sin(\vartheta_1 - \vartheta_2)$$

Incidentally, this is unity when the degree of coherence, as measured without anisotropic elements, is zero.

Let us summarize. We have two beams that appear identical in a scalar description as far as intensity distribution and coherence are concerned. Furthermore, they are both unpolarized. The only difference we found out is that orthogonal field components **at distinct points** can be correlated in one beam while are uncorrelated in the other.

Which theoretical tool should we use in order to account for this difference?

I realize that some of you could think: “All right, we found out a small difference between the two beams. So what? Why should we worry about finding a theoretical tool accounting for such details? Could not we simply ignore them?”

The answer is that our previous example in a particularly simple one. It merely aims at pointing out that the scalar description of a beam or the use of ordinary concepts drawn from the theory of polarized light can be insufficient for a complete description of the light field.

In a more general situation the need for a suitable tool can be more evident.

Just to give you an idea, let me add that one can easily build up examples of optical fields that upon propagation appear to be completely unpolarized at some planes and yet polarized at some others.

It may be useful to recall the customary approaches that can be used to describe polarized light beams.

a) Stokes parameters

b) Polarization (or coherence) matrix

Here we shall refer to the second choice.

We have seen that the polarization matrix is

$$\hat{J} = \begin{vmatrix} J_{xx} & J_{xy} \\ J_{yx} & J_{yy} \end{vmatrix} = \begin{vmatrix} \langle E_x^* E_x \rangle & \langle E_x^* E_y \rangle \\ \langle E_y^* E_x \rangle & \langle E_y^* E_y \rangle \end{vmatrix}$$

Field components are taken at a typical point in a cross-section and at equal times. Angular brackets denote times averages.

The important point to be made is that in the previous definition it is assumed that:

a) polarization state is the same at any point of a beam section

b) spatial coherence is complete

Neither of these hypotheses can be taken for granted in several cases of interest.

To deal with the most general case one should use the complete tensorial theory of the electromagnetic field developed by E. Wolf many years ago. This is not a simple task because, even in the quasi-monochromatic case, it involves the use of 3 different 3x3 matrices whose elements depend on 6 space variables.

Beam Coherence-Polarization matrix

(PCB matrix)

Many fields of interest propagate in the form of beams (multimode lasers, synthesized partially coherent sources). In this case:

a) longitudinal field components may be neglected

b) a single vector, e.g. the electric field, may be used (as for a plane wave)

One then introduces the matrix

$$\hat{\Gamma}^{(E)} = \begin{pmatrix} \Gamma_{xx}^{(E)}(\mathbf{r}_1, \mathbf{r}_2, z; \tau) & \Gamma_{xy}^{(E)}(\mathbf{r}_1, \mathbf{r}_2, z; \tau) \\ \Gamma_{yx}^{(E)}(\mathbf{r}_1, \mathbf{r}_2, z; \tau) & \Gamma_{yy}^{(E)}(\mathbf{r}_1, \mathbf{r}_2, z; \tau) \end{pmatrix}$$

where

$$\Gamma_{\alpha\beta}^{(E)}(\mathbf{r}_1, \mathbf{r}_2, z; \tau) = \left\langle E_\alpha^*(\mathbf{r}_1, z; t) E_\beta(\mathbf{r}_2, z; t + \tau) \right\rangle \qquad (\alpha, \beta = x, y)$$

In the quasi-monochromatic case we can limit ourselves to the BCP matrix

$$\hat{J}(\mathbf{r}_1,\mathbf{r}_2,z)=\begin{pmatrix} J_{xx}(\mathbf{r}_1,\mathbf{r}_2,z) & J_{xy}(\mathbf{r}_1,\mathbf{r}_2,z) \\ J_{yx}(\mathbf{r}_1,\mathbf{r}_2,z) & J_{yy}(\mathbf{r}_1,\mathbf{r}_2,z) \end{pmatrix}$$

where

$$J_{\alpha\beta}(\mathbf{r}_1,\mathbf{r}_2,z)=\left\langle E_\alpha^*(\mathbf{r}_1,z;t)E_\beta(\mathbf{r}_2,z;t)\right\rangle \qquad (\alpha,\beta=x,y)$$

The difference between the BCP matrix and the ordinary polarization matrix is that the corresponding elements are **numbers** (independent of z) in the latter while they are **correlation functions** in the former.

It will be noted that the elements of the BCP matrix have the same structure as the mutual intensity of the scalar coherence theory.

In particular, for $\mathbf{r}_1 = \mathbf{r}_2$, we have the **local** polarization matrix. At first, it might be thought that, disregarding coherence properties, such matrix could be enough in order to describe the polarization state of the beam. This may be considered as true across a certain plane of the beam. However, when it comes to evaluating the propagated quantities at a different plane knowledge of the local polarization matrix is insufficient to solve the problem. We need to know the BCP matrix at the starting plane because propagation formulas apply to the correlation functions. In other words, correlation and polarization properties are strictly tied.

Let us discuss a few properties of the BCP matrix. First, its elements can be normalized by introducing the degrees of correlation

$$j_{\alpha\beta}(\mathbf{r}_1,\mathbf{r}_2,z)=\frac{J_{\alpha\beta}(\mathbf{r}_1,\mathbf{r}_2,z)}{\sqrt{J_{\alpha\alpha}(\mathbf{r}_1,\mathbf{r}_1,z)J_{\beta\beta}(\mathbf{r}_2,\mathbf{r}_2,z)}}$$

In particular j_{xx} and j_{yy} are the complex degrees of coherence for the x- and y-components of the field. The anti-diagonal element j_{xy} is the degree of cross-coherence between the x-component at $\mathbf{r}_1$, z and the y-component at $\mathbf{r}_2$, z.

It is to be noted that $J_{xy}(\mathbf{r}_1, \mathbf{r}_2, z)$ is **not** locally connected to $J_{xx}(\mathbf{r}_1, \mathbf{r}_2, z)$ and $J_{yy}(\mathbf{r}_1, \mathbf{r}_2, z)$ [except for $\mathbf{r}_1 = \mathbf{r}_2$]. So, for certain triples ($\mathbf{r}_1$, $\mathbf{r}_2$, z), J_{xy} can be different from zero even if J_{xx} and J_{yy} are zero.

On the other hand integral inequalities are to be met. For example

$$\iint \Big[f_1^*(\mathbf{r}_1) f_1(\mathbf{r}_2) J_{xx}(\mathbf{r}_1,\mathbf{r}_2,z) + f_2^*(\mathbf{r}_1) f_2(\mathbf{r}_2) J_{yy}(\mathbf{r}_1,\mathbf{r}_2,z) + 2\,\mathrm{Re}\left\{ f_1^*(\mathbf{r}_1) f_2(\mathbf{r}_2) J_{xy}(\mathbf{r}_1,\mathbf{r}_2,z) \right\} \Big]\, d\mathbf{r}_1 d\mathbf{r}_2 \geq 0$$

Locally, the degree of polarization is computed through the usual formula

$$P(\mathbf{r},z) = \sqrt{1 - \frac{4\det\left\{\hat{J}(\mathbf{r},\mathbf{r},z)\right\}}{\left(Tr\left\{\hat{J}(\mathbf{r},\mathbf{r},z)\right\}\right)^2}}$$

When some anisotropic flat optical elements is inserted across the beam its effect can be described by a Jones matrix of the form

$$\hat{T}(\mathbf{r}) = \begin{pmatrix} a(\mathbf{r}) & b(\mathbf{r}) \\ c(\mathbf{r}) & d(\mathbf{r}) \end{pmatrix}$$

The new BCP matrix is given by

$$\hat{J}'(\mathbf{r}_1, \mathbf{r}_2, z) = \hat{T}^*(\mathbf{r}_1)\hat{J}(\mathbf{r}_1, \mathbf{r}_2, z)\hat{T}(\mathbf{r}_2)$$

As an example if we use a linear polarizer set at an angle φ the element is described by the matrix

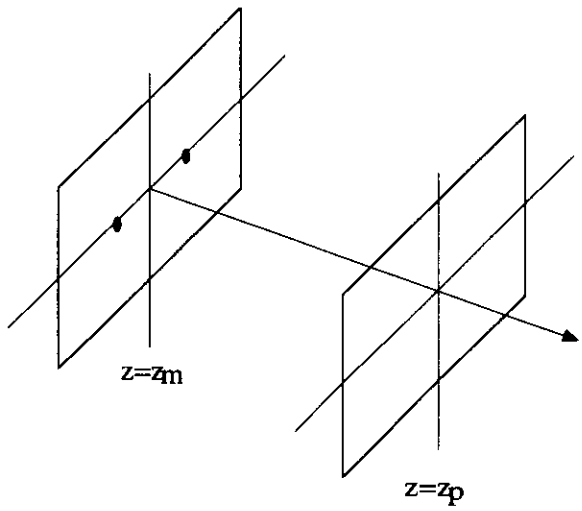
$$\hat{T} = \begin{pmatrix} C^2 & CS \\ CS & S^2 \end{pmatrix}$$

where C = cosφ and S = sinφ.

The intensity emerging from the polarizer is easily found to be

$$I_\varphi(\mathbf{r}, z) = J_{xx}(\mathbf{r}, \mathbf{r}, z)C^2 + J_{yy}(\mathbf{r}, \mathbf{r}, z)S^2 + 2\,\mathrm{Re}\left\{J_{xy}(\mathbf{r}, \mathbf{r}, z)\right\}CS$$

Let us see the role of the BCP matrix in the classical Young interference experiment.



The observed intensity distribution is

$$I(\mathbf{r}, z_p) = |K_1|^2\left[J_{xx}(\mathbf{r}_1, \mathbf{r}_1, z_m) + J_{yy}(\mathbf{r}_1, \mathbf{r}_1, z_m)\right]$$
$$+|K_2|^2\left[J_{xx}(\mathbf{r}_2, \mathbf{r}_2, z_m) + J_{yy}(\mathbf{r}_2, \mathbf{r}_2, z_m)\right]$$
$$+2\,\mathrm{Re}\left\{K_1^* K_2\left[J_{xx}(\mathbf{r}_1, \mathbf{r}_2, z_m) + J_{yy}(\mathbf{r}_1, \mathbf{r}_2, z_m)\right]\exp(-i2\pi\bar{\nu}\tau)\right\}$$

where

$$\tau = \frac{1}{c}\left\{\sqrt{(\mathbf{r}-\mathbf{r}_1)^2 + (z_p - z_m)^2} - \sqrt{(\mathbf{r}-\mathbf{r}_2)^2 + (z_p - z_m)^2}\right\}$$

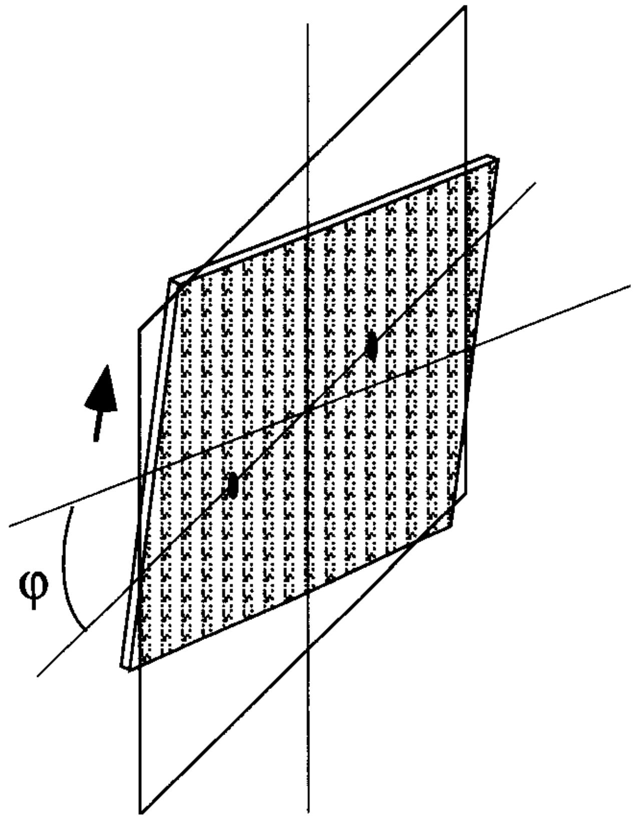
Alternatively , we can write such intensity as

$$I(\mathbf{r}, z_p) = |K_1|^2 I(\mathbf{r}_1, z_m) + |K_2|^2 I(\mathbf{r}_2, z_m)$$
$$+2\,\mathrm{Re}\left\{K_1^* K_2 J_{eq}(\mathbf{r}_1, \mathbf{r}_2, z_m)\exp(-i2\pi\bar{\nu}\tau)\right\}$$

where we defined the equivalent mutual intensity

$$J_{eq}(\mathbf{r}_1, \mathbf{r}_2, z_m) = J_{xx}(\mathbf{r}_1, \mathbf{r}_2, z_m) + J_{yy}(\mathbf{r}_1, \mathbf{r}_2, z_m)$$

If no anisotropic elements are present the equivalent mutual intensity is the sum of those pertaining to the x- and y-components.


φ

If the Young mask is covered with a linear polarizer the intensity across the observation plane is

$$I_{\varphi}(\mathbf{r}, z_p) = |K_1|^2 I_{\varphi}(\mathbf{r}_1, z_m) + |K_2|^2 I_{\varphi}(\mathbf{r}_2, z_m)$$
$$+2\,\mathrm{Re}\left\{K_1^* K_2 J_{\varphi eq}(\mathbf{r}_1, \mathbf{r}_2, z_m)\exp(-i2\pi\bar{\nu}\tau)\right\}$$

where

$$J_{\varphi eq}(\mathbf{r}_1, \mathbf{r}_2, z_m) =$$
$$= J_{xx}(\mathbf{r}_1, \mathbf{r}_2, z_m)C^2 + J_{yy}(\mathbf{r}_1, \mathbf{r}_2, z_m)S^2 + 2\,\mathrm{Re}\left\{J_{xy}(\mathbf{r}_1, \mathbf{r}_2, z_m)CS\right\}$$

This is a new type of interference law. In fact, it has to deal with mutual intensities rather than mere optical intensities.

Let us come back to the example that we treated in qualitative terms at the beginning.

Consider the incoherent superposition of two Laguerre-Gauss modes of the form

$$\begin{cases} E_x^{(1)}(\mathbf{r},0;\mathrm{t}) = a(t)\dfrac{r}{\upsilon\sqrt{2}}\exp\left(-\dfrac{r^2}{\upsilon^2}\right)\exp(i\vartheta) \\ E_y^{(1)}(\mathbf{r},0;\mathrm{t}) = b(t)\dfrac{r}{\upsilon\sqrt{2}}\exp\left(-\dfrac{r^2}{\upsilon^2}\right)\exp(-i\vartheta) \end{cases}$$

These are vortex modes with opposite charges. The first one is polarized along the x-axis, the other along the y-axis. We assume that

$$\langle a^*(t)b(t)\rangle = 0\,; \qquad \left\langle |a(t)|^2 \right\rangle = \left\langle |b(t)|^2 \right\rangle$$

Consider now a different incoherent superposition.

$$\begin{cases} E_x^{(2)}(\mathbf{r},0;\mathrm{t}) = [a(t)\exp(i\vartheta) + b(t)\exp(-i\vartheta)]\dfrac{r}{2\upsilon}\exp\left(-\dfrac{r^2}{\upsilon^2}\right) \\ E_y^{(2)}(\mathbf{r},0;\mathrm{t}) = i[a(t)\exp(i\vartheta) - b(t)\exp(-i\vartheta)]\dfrac{r}{2\upsilon}\exp\left(-\dfrac{r^2}{\upsilon^2}\right) \end{cases}$$

This time the modes are circularly polarized with opposite helicity.

It is not difficult to show that for both cases the equivalent mutual intensity is

$$J_{eq}^{(j)}(\mathbf{r}_1,\mathbf{r}_2,0) = I_0\frac{\mathbf{r}_1\mathbf{r}_2}{\upsilon^2}\exp\left(-\frac{\mathbf{r}_1^2+\mathbf{r}_2^2}{\upsilon^2}\right)\cos(\vartheta_2-\vartheta_1) \qquad (j=1,2)$$

In particular the optical intensity is

$$I^{(j)}(\mathbf{r},0) = I_0\frac{\mathbf{r}_2}{\upsilon^2}\exp\left(-\frac{2\mathbf{r}^2}{\upsilon^2}\right) \qquad (j=1,2)$$

This proves that the two beams appear identical in a scalar description.

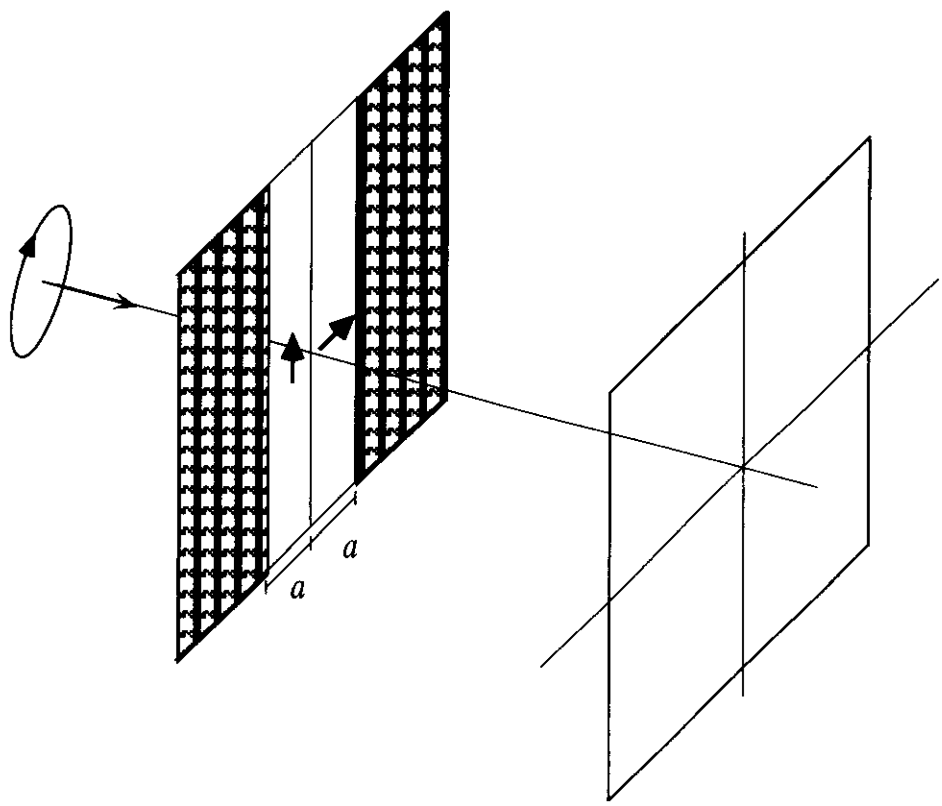
However the beams are physically different. This is revealed by the BCP matrix whose elements turn out to be

$$\begin{cases} J_{xx}^{(1)}(\mathbf{r}_1,\mathbf{r}_2,0) = I_0 \dfrac{\mathbf{r}_1\mathbf{r}_2}{2\upsilon^2} \exp\left[-\dfrac{\mathbf{r}_1^2+\mathbf{r}_2^2}{\upsilon^2} + i(\vartheta_2-\vartheta_1)\right] \\ J_{yy}^{(1)}(\mathbf{r}_1,\mathbf{r}_2,0) = \left[J_{xx}^{(1)}(\mathbf{r}_1,\mathbf{r}_2,0)\right]^* \\ J_{xy}^{(1)}(\mathbf{r}_1,\mathbf{r}_2,0) = 0 \end{cases}$$

$$\begin{cases} J_{xx}^{(2)}(\mathbf{r}_1,\mathbf{r}_2,0) = I_0 \dfrac{\mathbf{r}_1\mathbf{r}_2}{2\upsilon^2} \exp\left(-\dfrac{\mathbf{r}_1^2+\mathbf{r}_2^2}{\upsilon^2}\right) \cos(\vartheta_2-\vartheta_1) \\ J_{yy}^{(2)}(\mathbf{r}_1,\mathbf{r}_2,0) = J_{xx}^{(2)}(\mathbf{r}_1,\mathbf{r}_2,0) \\ J_{xy}^{(2)}(\mathbf{r}_1,\mathbf{r}_2,0) = -I_0 \dfrac{\mathbf{r}_1\mathbf{r}_2}{2\upsilon^2} \exp\left(-\dfrac{\mathbf{r}_1^2+\mathbf{r}_2^2}{\upsilon^2}\right) \sin(\vartheta_2-\vartheta_1) \end{cases}$$

Note in particular the anti-diagonal elements.

On comparing scalar and vectorial treatments possible pitfalls should be avoided. Let us discuss an elementary example. A slit of width 2a is orthogonally illuminated by a monochromatic circularly polarized plane wave. Half the slit is covered by a linear polarizer with vertical axis and the other half by a horizontally oriented linear polarizer.



We want the intensity distribution in the far-zone.

It may be argued that in a scalar treatment, in which polarization is ignored, we deal with a uniformly illuminated slit of width 2a. Accordingly, the predicted intensity will be that of the diffraction pattern produced by such a slit.

On the other hand, taking polarization into account we realize that the intensity in the far-zone is due to the superposition of two (identical) patterns corresponding to a slit of width a.

Shall we conclude that the scalar theory leads to an incorrect prediction? One could simply say that the scalar treatment is unsuitable because we cannot neglect the vectorial features of the problem. There is, however, a subtler answer. We can use a scalar theory but we have to consider the slit as a **partially coherent** source. Indeed, fields emitted from the two halves do not interfere with one another.

The equivalence of the scalar and vectorial treatments as far as only the intensity is of interest is easily proved.

We have already seen that across a certain plane the connection between scalar and vectorial approaches is expressed by the formula

$$J_{eq}(\mathbf{r}_1, \mathbf{r}_2, z_m) = J_{xx}(\mathbf{r}_1, \mathbf{r}_2, z_m) + J_{yy}(\mathbf{r}_1, \mathbf{r}_2, z_m)$$

Since propagation formulas are linear this connection is at once extended at any plane.

In summary, we presented the basic elements for describing both **coherence and polarization** properties when the light field is beam-like. A single 2x2 matrix, the BCP matrix, is sufficient for this. Its elements have the structure of mutual intensities. Removal of the quasi-monochromaticity hypothesis is easily accomplished.

Scalar coherence theory has a very large number of applications to beam-like fields. It is likely that the present extension, aimed at dealing in a simple way with vectorial properties, can be of some usefulness.

References

1) L. Mandel and E. Wolf, Optical Coherence and Quantum Optics, Cambridge University Press, 1995
2) Martínez-Herrero, P.M. Mejías, and J. M. Movilla, Opt. Lett., **11** (1997) 1818
3) D. F. V. James, J. Opt. Soc. Am. A, **11** (1994) 1641
4) F. Gori, Optics Lett., **23** (1998) 241
5) F. Gori, M. Santarsiero, S. Vicalvi, R. Borghi and G. Guattari, J. E. O. S., Pure and Applied Optics, **7** (1998) 941

Selected topics

a) **Pancharatnam's theorem**

b) **Angular momentum of light**

c) **Polarization gratings**

a) **Pancharatnam's theorem**

Preliminaries

We shall denote Jones vectors by a bold letter

$$\mathbf{v} = \begin{pmatrix} \cos\dfrac{\vartheta}{2} \\ \sin\dfrac{\vartheta}{2}\, e^{i\varphi} \end{pmatrix}$$

The optical intensity associated to **v** is taken to be

$$I = |E_x|^2 + |E_y|^2 = \cos^2\left(\frac{\vartheta}{2}\right) + \sin^2\left(\frac{\vartheta}{2}\right)$$

and therefore equals one.

Letting, for the sake of brevity

$$\mathbf{v} = \begin{pmatrix} a \\ b \end{pmatrix}$$

the intensity associated with the superposition of two waves specified by certain vectors

$$\mathbf{v}_1 = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}; \qquad \mathbf{v}_2 = \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}$$

will be obtained by squaring the sum of the x-components and adding the square of the sum of the y-components.

In symbols

$$I = |a_1 + a_2|^2 + |b_1 + b_2|^2$$

This can also be written

$$I = |\mathbf{v}_1 + \mathbf{v}_2|^2 = 2[1 + \mathrm{Re}\{\mathbf{v}_1 \cdot \mathbf{v}_2\}]$$

using the inner product

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = a_1^* a_2 + b_1^* b_2$$

where the asterisk stands for complex conjugate.

Pancharatnam's connection

Pancharatnam started from the following problem. We have two waves with different states of polarization. When can they be said to be in phase? Many opticists could have argued that there was no answer to such question. Consider in fact the x- and y-components of the electric fields of the two waves. It may well happen that the x-components are in phase while the y-components are in antiphase (i. e. with a phase difference equal to π). Therefore, the two waves are neither in phase nor in antiphase. The question seems to be meaningless. The first great merit of Pancharatnam was to insist that the question could find an answer, at least in a conventional or operational sense.

To understand how, let us suppose that we are looking at the superposition of two waves with equal polarization and that we can change at will the phase of one of them. We progressively turn some knob and we have to decide when the two waves are in phase. How do we realise this is occurring? No doubt, we say the two waves are in phase when the intensity of their superposition reaches its maximum. Pancharatnam suggested that the same rule could be adopted for any pair of waves, regardless of their state of polarization. The rule breaks down only for orthogonal states of polarization, a case we shall exclude. Berry called this rule the Pancharatnam's connection.

We now take a further step by defining a sort of phase difference between two waves in different states of polarization. If the two waves are represented by Jones vectors $\mathbf{v}_1$ and $\mathbf{v}_2$, we multiply the second vector by exp(-iα) and maximise the resulting intensity

$$I = \left|\mathbf{v}_1 + \mathbf{v}_2 e^{-i\alpha}\right|^2$$

with respect to α. Assuming the vectors to be normalized, I can be written as

$$I = 2\left[1 + \mathrm{Re}\left\{(\mathbf{v}_1 \cdot \mathbf{v}_2)e^{-i\alpha}\right\}\right] = 2\left\{1 + |\mathbf{v}_1 \cdot \mathbf{v}_2| \cos\left[\arg(\mathbf{v}_1 \cdot \mathbf{v}_2) - \alpha\right]\right\}$$

We see that the value of α that maximises the intensity equals $\arg(\mathbf{v}_1 \cdot \mathbf{v}_2)$. We shall call it the phase mismatch between $\mathbf{v}_2$ and $\mathbf{v}_1$ and we shall denote it by ψ_{21}. More explicitly

$$\psi_{21} = \arg(\mathbf{v}_1 \cdot \mathbf{v}_2)$$

In order to put $\mathbf{v}_1$ and $\mathbf{v}_2$ in phase we must change the initial phase of $\mathbf{v}_2$ by $-\psi_{21}$. The scalar product can be written

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = \cos\frac{\vartheta_1}{2}\cos\frac{\vartheta_2}{2} + \sin\frac{\vartheta_1}{2}\sin\frac{\vartheta_2}{2}\cos(\varphi_2 - \varphi_1) + \\ + i\sin\frac{\vartheta_1}{2}\sin\frac{\vartheta_2}{2}\sin(\varphi_2 - \varphi_1)$$

from which the phase mismatch can be evaluated.

It is worthwhile to remark that ψ_{21} should not be confused with the "angle", say β_{21}, between the complex vectors $\mathbf{v}_1$ and $\mathbf{v}_2$, which is customarily defined through the relation

$$\cos^2 \beta_{21} = \frac{|\mathbf{v}_1 \cdot \mathbf{v}_2|^2}{|\mathbf{v}_1|^2 |\mathbf{v}_2|^2}$$

which in the present case ($|\mathbf{v}_1| = |\mathbf{v}_2| = 1$) gives

$$\cos^2 \beta_{21} = |\mathbf{v}_1 \cdot \mathbf{v}_2|^2$$

We now explore a most important consequence of the Pancharatnam's rule by working out a simple example. Let us consider the following three column vectors

$$\mathbf{v}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}; \quad \mathbf{v}_2 = \begin{pmatrix}1\\0\end{pmatrix}; \quad \mathbf{v}_3 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix}$$

The corresponding polarizations will be recognised as linear along a line oriented at $\pi/4$ between the x- and the y-axis ($\mathbf{v}_1$), linear along the x-axis ($\mathbf{v}_2$), and right-handed circular ($\mathbf{v}_3$).

Let us evaluate the phase mismatch between $\mathbf{v}_2$ and $\mathbf{v}_1$. On computing $\arg(\mathbf{v}_1 \cdot \mathbf{v}_2)$ we find at once $\psi_{21}=0$ and we conclude that $\mathbf{v}_1$ and $\mathbf{v}_2$ are in phase. Next, we evaluate the phase mismatch between $\mathbf{v}_3$ and $\mathbf{v}_2$. Proceeding as before, we find $\psi_{32}=0$ and we deduce that $\mathbf{v}_2$ and $\mathbf{v}_3$ are also in phase. We could be tempted to conclude that $\mathbf{v}_1$ and $\mathbf{v}_3$ are in phase too. This, however, is not true. As a matter of fact, when we insert the expressions of $\mathbf{v}_1$ and $\mathbf{v}_3$ into $\arg(\mathbf{v}_1 \cdot \mathbf{v}_2)$ we obtain $\psi_{31}= \pi/4$.

The present example shows that "to be in phase" in the sense of the Pancharatnam's connection is not a transitive property. For this reason, the quantity we called phase mismatch should not be thought of as an ordinary phase difference. Indeed, denoting by γ_1, γ_2 and γ_3 some arbitrary phases to be taken in the usual sense, the identity: $(\gamma_3 - \gamma_2) + (\gamma_2 - \gamma_1) = \gamma_3 - \gamma_1$ would hold, whereas in the example that we have just seen $\psi_{32} + \psi_{21} \neq \psi_{31}$.

We now possess all the necessary ingredients to work out the Pancharatnam's theorem

The theorem

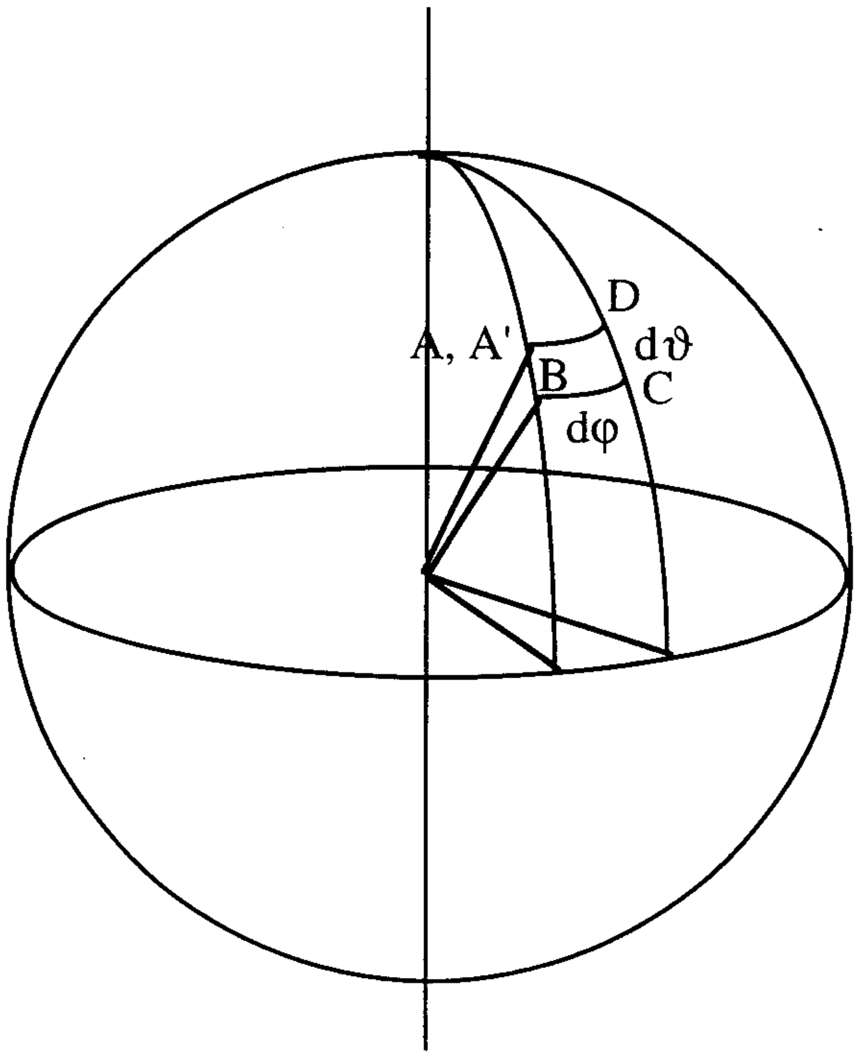
Let us suppose that we start from light with a certain state of polarization. Using anisotropic devices, like wave-plates, rotators and linear polarisers, we can change the state of polarization. We can imagine this is done in a continuous way, so that the state parameters are progressively changed by infinitesimal quantities. We can arrange things in such a way that, at the end of the transformation, we are back to the initial polarization state. In other words, the light wave performs a cyclic transformation, or a cycle. Suppose the form of the cycle has been fixed, in the sense that all intermediate states of polarizations are specified. Nonetheless, we can still choose at will the initial phase of any intermediate state.

Now, following Pancharatnam, we make the hypothesis that when we move along the transformation by infinitesimal steps, any new state is kept in phase with the previous one. This will be obtained by acting on the initial phase of the new state. Because of the non-transitivity of the phase matching, we understand that a typical state along the cycle need not be in phase with the starting one, although it is in phase with its immediate neighbours. Upon finishing the cycle, however, we would expect the final state to be in phase with the starting one. To our surprise, this is not true. In other words, we end up with an accumulated phase change that is generally different from zero. It can be said that the wave keeps a sort of memory of the transformation it passed through.

The magic is that Pancharatnam was able to find a very simple expression for the accumulated phase change. If we represent the cycle on the Poincaré sphere, the overall phase change is simply equal to half the solid angle subtended by the cycle at the origin. This is Pancharatnam's great theorem.

S. Pancharatnam, Proc. Ind. Acad. Sci. A, **44** (1956) 247

We shall limit ourselves to prove the theorem for an infinitesimal cycle, like the curvilinear quadrangle ABCD depicted below



We start from the polarization state corresponding to the point A specified by certain angular coordinates ϑ and φ. We move to B on increasing ϑ by $d\vartheta$ then to C on increasing φ by $d\varphi$. Next, we go to D on decreasing ϑ by $d\vartheta$ and then back again by decreasing φ. We denote the final state by A', meaning that although this state has the same polarization as A, it may have a different initial phase. Our goal indeed is to evaluate the phase mismatch between A' and A.

The essence of the theorem can be grasped by comparing the effects separately produced by infinitesimal variations of colatitude and longitude. We begin from the former, by inserting the angular coordinates corresponding to B and A into

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = \cos\frac{\vartheta_1}{2}\cos\frac{\vartheta_2}{2} + \sin\frac{\vartheta_1}{2}\sin\frac{\vartheta_2}{2}\cos(\varphi_2 - \varphi_1) + \\ + i\sin\frac{\vartheta_1}{2}\sin\frac{\vartheta_2}{2}\sin(\varphi_2 - \varphi_1)$$

More explicitly, we let $\vartheta_1 = \vartheta$, $\vartheta_2 = \vartheta + d\vartheta$ and $\varphi_1 = \varphi_2 = \varphi$. It is immediately seen that

$$\mathbf{v}_A \cdot \mathbf{v}_B = \cos\frac{d\vartheta}{2}$$

and this leads to the infinitesimal phase mismatch

$$d\psi_{BA} = 0$$

Therefore states A and B are in phase, so that we conclude that states along the same meridian line are in phase. Let us then see the effect of a change of longitude by moving from A to D. Letting now $\vartheta_1 = \vartheta_2 = \vartheta$ and $\varphi_1 = \varphi$; $\varphi_2 = \varphi + d\varphi$ we find

$$\mathbf{v}_A \cdot \mathbf{v}_D = \cos^2\frac{\vartheta}{2} + \sin^2\frac{\vartheta}{2}\cos d\varphi + i\sin^2\frac{\vartheta}{2}\sin d\varphi$$

Taking into account the infinitesimal nature of $d\varphi$, the phase mismatch $d\psi_{DA}$ is seen to be

$$d\psi_{DA} = \sin^2\frac{\vartheta}{2}\,d\varphi$$

This means that states at different longitudes are not in phase. In order to put D in phase with A we have to change its initial phase by $-d\psi_{DA}$. It is to be noted that, for a given value of $d\varphi$, the phase mismatch depends on ϑ. This is a crucial point. In particular, if we move from B to C we find the phase mismatch

$$d\psi_{CB} = \sin^2 \frac{\vartheta + d\vartheta}{2}\, d\varphi$$

Let us now move around the cycle. States A and B are in phase. The initial phase of C has to be changed by $-d\psi_{CB}$ in order to keep C in phase with B. States C and D are in phase. Finally, in order to bring A' in phase with D, we change the initial phase of A', by $-d\psi_{A'D} = d\psi_{DA}$. The overall phase change accumulated along the cycle is then

$$d\psi = -d\psi_{CB} + d\psi_{DA} = \left[-\sin^2\frac{\vartheta + d\vartheta}{2} + \sin^2\frac{\vartheta}{2}\right] d\varphi = -\frac{1}{2}\sin\vartheta \, d\vartheta \, d\varphi$$

where the infinitesimal nature of $d\vartheta$ has been exploited. It is seen that $d\psi$ is equal (disregarding the sign) to half the solid angle subtended by the line ABCD at the origin of the sphere. Hence, the theorem is proved for an infinitesimal cycle on the sphere.

b) Angular momentum of light

A photon of right-handed (left-handed) light propagating along the z-axis has an intrinsic (spin) angular momentum $\hbar\hat{z}(-\hbar\hat{z})$. If such photon passes through a half wave plate its angular momentum changes sign. As a consequence the plate experiences a mechanical torque. This effect, which is also predicted by classical electromagnetism was verified experimentally by Beth as early as 1936.

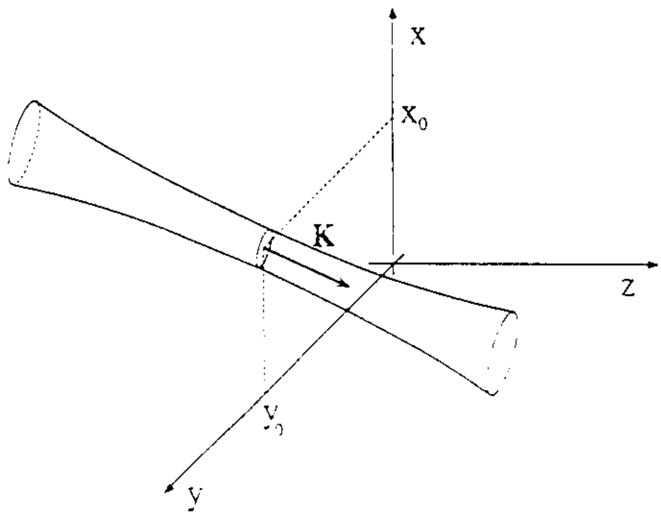
In addition, light can possess an orbital angular momentum.

The ordinary mathematical derivation of the orbital angular momentum is rather long and intricate. In recent times, several researches have been carried out on the orbital angular momentum of light beams, in which case the paraxial approximation can be used to a good approximation. These investigations have led to a lot of new results and have greatly clarified the subject. Furthermore significant applications of these results have been demonstrated. We shall now discuss a simple model for the orbital angular momentum of photons, which can be presented with minimal prerequisites. Essentially, the required concepts are that a photon possesses a linear momentum and that the optical intensity can be thought of as proportional to the spatial probability density of the photon.

Let us consider a certain plane, to be taken as z = 0, illuminated by a monochromatic light beam. As mentioned above the photon spin is connected with states of circular polarization. As a consequence, if we assume that the beam is linearly polarized, the expectation value of the spin angular momentum is zero and we can focus our attention on the orbital part only. In the scalar, complex representation of the light beam, we can describe the field distribution across z = 0 through a function V(x, y). There is no need to specify the exact meaning of V. For example, it could represent the complex electric field of the wave. The important point is its probabilistic meaning. Following the idea first put forward by Born, we assume that for a single photon, the squared modulus of V is proportional to the probability density that the photon crosses the plane z = 0 at point (x, y).

Examples

Suppose a TEM_{00} Gaussian mode impinges on the plane $z = 0$ centered at a point (x_0, y_0). Further, let us assume the axis of the beam to be inclined with respect to the z-axis.



Neglecting the ellipticity induced by such inclination (assumed to be small) we write the corresponding field distribution as

$$V(x,y) = A\, e^{-\frac{(x-x_0)^2+(y-y_0)^2}{v^2}}\, e^{i(K_x x+K_y y)}$$

where A is an amplitude term, v is the spot-size, and K_x and K_y are the x- and y-components of the mean wave-vector of the beam. In other words, **K** is directed along the beam axis.

Now, let k_x and k_y be the transverse components of $\mathbf{k}$. Then the z-component of the angular momentum of the photon is

$$m_z = \hbar\left(xk_y - yk_x\right)$$

Of course, we do not know where the photon hits the plane z=0. Therefore, we must content ourselves with the expectation value of m_z, which is to be computed through the probability density for the crossing point. In view of the above remark about the meaning of $|V(x, y)|^2$ such probability density, say p(x, y), can be written

$$p(x,y) = \frac{\left|V(x,y)\right|^2}{\iint \left|V(x,y)\right|^2 dxdy}$$

The expectation value for the angular momentum along the z-axis is

$$\langle m_z \rangle = \hbar \iint \left(xk_y - yk_x \right) p(x,y) dxdy$$

There remains to be seen how k_x and k_y can be derived from the knowledge of V(x, y). To this end, let us write V(x, y) in the form

$$V(x,y) = \left| V(x,y) \right| e^{i\phi(x,y)}$$

If the wavefront is sufficiently regular, as in the case of paraxial beams we can expand the phase in a neighbourhood of (x,y) and consider only the first order terms, i.e.

$$\phi(x+\xi, y+\eta) = \phi(x,y) + \frac{\partial\phi}{\partial x}\xi + \frac{\partial\phi}{\partial y}\eta$$

ξ and η represent small deviations along x and y, respectively. When this expression of the phase is compared to the phase distribution, say $\psi(x,y)$, produced by a plane wave across the plane z = 0, namely

$$\psi(\xi,\eta) = \alpha_0 + k_x(x+\xi) + k_y(y+\eta)$$

where α_0 is an initial phase term, we see that the following equations hold

$$k_x = \frac{\partial\phi}{\partial x}; \qquad k_y = \frac{\partial\phi}{\partial y}$$

On inserting these results into the expression for the expected angular momentum we obtain

$$\langle m_z \rangle = \hbar \frac{\iint \left(x\frac{\partial \phi}{\partial y} - y\frac{\partial \phi}{\partial x} \right) |V(x,y)|^2 \, dxdy}{\iint |V(x,y)|^2 \, dxdy}$$

This is the expected value for the z-component of the orbital angular momentum of a photon. It coincides with the expression derived from Maxwell's equations.

A few comments can be of help. We evaluated the expected value of the angular momentum along the z-axis. A different value can be expected if we refer the angular momentum to an axis parallel to the z-axis but passing through a typical point (x_a, y_a) at $z = 0$. In this case, the angular momentum becomes

$$m_z' = \hbar\left[(x - x_a)k_y - (y - y_a)k_x\right]$$

and its expected value is

$$\langle m_z' \rangle = \langle m_z \rangle - \hbar\left(x_a \langle k_y \rangle - y_a \langle k_x \rangle\right)$$

where

$$\langle k_\alpha \rangle = \iint k_\alpha p(x, y)dxdy; \qquad (\alpha = x, y)$$

are the expected values of the x- and y-components of the wave-vector. It may well happen that both $\langle k_x \rangle$ and $\langle k_y \rangle$ vanish (we shall see an example later). In this case the orbital angular momentum becomes an intrinsic feature in that it is independent of the coordinates of the point at which the chosen axis crosses the plane $z = 0$.

As a further remark, let us note that for certain types of field distributions polar rather than cartesian coordinates are used. Letting

$$x = r\cos\vartheta; \quad y = r\sin\vartheta$$

we can write

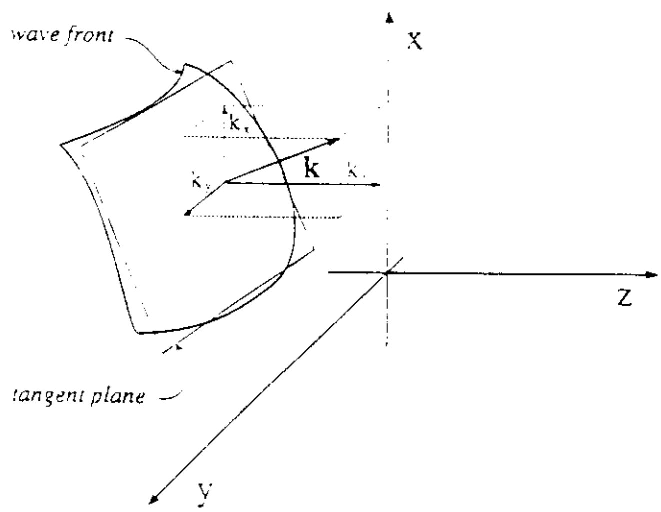
$$\frac{\partial}{\partial\vartheta} = \frac{\partial}{\partial x}\frac{\partial x}{\partial\vartheta} + \frac{\partial}{\partial y}\frac{\partial y}{\partial\vartheta} = x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}$$

so that the expression of the expected angular momentum becomes

$$\langle m_z \rangle = \hbar \frac{\int_0^{2\pi}\int_0^{\infty} \frac{\partial\phi}{\partial\vartheta} |U(r,\vartheta)|^2 \, r \, dr \, d\vartheta}{\int_0^{2\pi}\int_0^{\infty} |U(r,\vartheta)|^2 \, r \, dr \, d\vartheta}$$

where U specifies the field distribution at z = 0 in polar coordinates.

The other idea we need is that a monochromatic photon of frequency ν possesses a linear momentum with modulus hν/c, where h is Planck's constant and c is the speed of light. We must specify, however, the direction of such vector. To this aim, let us assume that in the neighbourhood of (x, y) the wavefront of the beam can be approximated by its tangent plane. In other words, we locally replace the wavefront by a plane wave.



Then the linear momentum of a photon passing at (x, y) can be thought of as directed along the wave-vector, say **k**, of such plane wave. Since $k = 2\pi\nu/c$, the linear momentum has a modulus $\hbar k$.

Using the previous expression for V(x,y) we find through simple passages

$$\langle m_z \rangle = \hbar\left(x_0 K_y - y_0 K_x\right)$$

It may be worthwhile to note that this is the same value that would pertain to a photon crossing $z = 0$ at x_0, y_0. This can be interpreted by saying that such point plays the role of a center of mass for the beam. It is seen that the present angular momentum is not an intrinsic feature of the beam. Indeed it can be made arbitrarily high (at least in principle) or vanishing through a suitable choice of (x_0, y_0) and (K_x, K_y).

As a second example, we shall consider a well known class of Laguerre-Gauss beams specified at their waist by the distributions

$$U_n(r,\vartheta) = A\, r^n e^{\pm in\vartheta}\, e^{-\frac{r^2}{v^2}}$$

where A is a constant and n is an integer number. On using this expression we obtain

$$\langle m_z \rangle = n\hbar$$

This beautiful result was first derived by Allen et al.

L. Allen, M. W. Beijersbergen, J. C. Spreeuw and J. P. Woerdman, Phys. Rev. A, **45** (1992) 8185

Writing $\phi=\tan^{1}(y/x)$ we deduce that

$$k_x = \frac{\partial\phi}{\partial x} = -\frac{y}{x^2+y^2} = -\frac{\sin\vartheta}{r}; \qquad k_y = \frac{\partial\phi}{\partial y} = \frac{x}{x^2+y^2} = \frac{\cos\vartheta}{r}$$

It can be easily seen that the expected values of k_X and k_y vanish. Therefore, the expected value of the angular momentum has an intrinsic meaning. This can be traced back to the form of the wavefronts. On moving around a circle of constant radius at the plane z = 0 the phase steadily increases by $\pm 2\pi n$. Therefore, the wavefronts, which constitute examples of optical vortices, are helicoidally shaped. This means that there is a local tilting of the wavefront with respect to the plane z = 0. It is this shape of the wavefront that determines the existence of an angular momentum along the z-axis.

F. Gori, M. Santarsiero, R. Borghi and G. Guattari, Eur. J. Phys **19** (1998) 439

c) **Polarization gratings**

Let us consider a set of strips cut from a linearly polarizing film. The strips are arranged in a plane, side by side, with the long sides oriented parallel to each other. A coordinate system is used in which the x and y axes are parallel to the short and the long sides of the strips, respectively. The orientation of the transmission axis can change when one passes from one strip to the other. Accordingly, we speak of a local transmission axis. The angle between the latter and the x axis is denoted by $\varphi(x)$. We assume φ to be periodic.

Suppose now that a monochromatic plane wave in any state of polarization impinges orthogonally (say, along the z axis) on the grating. We can specify such a wave through the Jones vector

$$E_i = \begin{pmatrix} A_x \\ A_y \, e^{i\delta} \end{pmatrix}$$

We have already seen that the Jones matrix describing the action of a linear polarizer is

$$P = \begin{pmatrix} C^2 & CS \\ CS & S^2 \end{pmatrix}$$

Now, however, φ depends on x.

The vector specifying the emerging field, say E_e, is

$$E_e = \begin{pmatrix} \dfrac{A_x}{2} + \dfrac{A_x - iA_y\, e^{i\delta}}{4} e^{2i\varphi} + \dfrac{A_x + iA_y\, e^{i\delta}}{4} e^{-2i\varphi} \\ \dfrac{A_x}{2} - i\dfrac{A_x - iA_y\, e^{i\delta}}{4} e^{2i\varphi} + i\dfrac{A_x + iA_y\, e^{i\delta}}{4} e^{-2i\varphi} \end{pmatrix}$$

In order to interpret this result we have to make a hypothesis about φ. For simplicity, we assume φ to be a linear function of x. The period of φ is the interval over which φ changes by π. The terms containing $\exp(2i\varphi)$ and $\exp(-2i\varphi)$ represent diffracted waves of the first order, while the remaining terms pertain to the undiffracted field.

The corresponding Jones vectors, to be denoted by E_1, E_{-1} and E_0 are

$$E_0 = \frac{1}{2}\begin{pmatrix} A_x \\ A_y \, e^{i\delta} \end{pmatrix}$$

$$E_1 = \frac{A_x - iA_y \, e^{i\delta}}{4}\begin{pmatrix} 1 \\ -i \end{pmatrix}$$

$$E_{-1} = \frac{A_x + iA_y \, e^{i\delta}}{4}\begin{pmatrix} 1 \\ i \end{pmatrix}$$

The zero-order wave is identical to the incident wave except for an amplitude factor. The order 1 is a left-circularly polarized wave (L-wave), the order -1 a right-circularly polarized wave (R-wave).

The states of polarization of the orders ±1 do not depend on the polarization of the incident wave. On the other hand, the amplitudes of these waves depend on the polarization of the incoming field. In particular, consider the case in which the incident wave is circularly polarized, say an L-wave. We easily find that the only first order wave whose amplitude is different from zero is the R-wave. So, disregarding the undiffracted field, an orthogonal incident L-wave is converted into an inclined R-wave (and viceversa fon an incoming R-wave). In general, it turns out that the amplitude of the L-wave (R-wave) associated with the order 1 (-1) is proportional to the R-polarized (L-polarized) component of the incident field.

It can be noted that there is a transfer of angular momentum between the light beam and the grating.

We considered monochromatic and hence completely polarized light. The treatment could be extended to partially polarized quasi-monochromatic radiation.

For broadband radiation, in which the polarization properties can depend on the wavelength, an angular dispersion effect will be present owing to the very nature of the grating.

It can be shown that the decomposition performed by the polarization grating can be used for evaluation of the Stokes parameters of a light beam.

We worked with a simple linear law of variation for φ, but other situations could be considered. For example if φ(x) has the form γx^2 (with real positive γ) one finds that a plane L-wave is converted into a converging cylindrical R-wave, whereas an incident plane R-wave gives rise to a diverging L-wave. The addition of a linear term to φ gives rise to a regime similar to that characterizing off-axis holography.

F. Gori, Opt. Lett. **24** (1999) 584.

F. Gori, M. Santarsiero, R. Borghi, G. Guattari, Atti Fondaz. G. Ronchi **54** (1999) 59.