IMAGING PROCESSES AND COHERENCE IN OPTICS

A. MARECHAL
Institut d'Optique - Orsay
France

Optics is the domain of physics where the fundamental concepts of image formation developped : geometrical optics was used to interpret the behaviour of mirrors, lenses, of the eye etc... and allowed the realization of conventional optical instruments.

Physical optics opened new perspectives not only for the representation of the propagation of light but also for the formation of optical images and the existence of fundamental limitations imposed by diffraction. Phase contrast and interferential microscopy illustrate the role of physical optics in the development of potentialities of optical images ; they pointed out the importance of coherence and stimulated the study of the transition between coherence and incoherence i.e. the domain of partial coherence where the leaders have been Zernike, Hopkins, Wolf and Blanc-Lapierre in the period 1948-1960. Those basic concepts proved to be very important for understanding the formation of optical images in instruments like microscopes, optical projectors etc...

What is then the meaning of the word coherence ? According to tradition it expresses the ability to produce interferences when beams are superimposed ; we will use this term in this precise meaning, hoping that distorsions that could appear in other fields would not lead to appreciable "incoherences" with this original meaning.

We will first analyse the problem of coherence and then the problem of image formation where coherence is one of the basic factors.

I Coherence, incoherence, partial coherence, space and time problems

A) Partial coherence

Two examples will easily demonstrate the existence of intermediate situations between coherence and incoherence

1- Young fringes with an extended source : Let us imagine that we perform the classical experiment by Th. Young with a very small monochromatic source So illuminating two small holes T_1 T_2 (fig.1) :

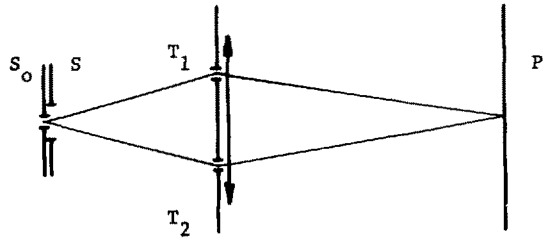


Fig. 1

fringes are observable in the plane P, and we conclude that vibrations from T_1 and T_2 are perfectly coherent. Let us now suppose that we use a larger source S : the fringe systems generated by the various elements of S (emitting incoherent vibrations) will be superimposed, the minima of illumination will no longer be zero, the contrast of fringes will decrease and it is very easy to increase the size of S and suppress interference phenomena : according to classical interpretation we can consider that this is the result of a blurring of interference phenomena having various positions in plane P. We also can declare that vibrations from holes T_1 and T_2 are no longer coherent : if the source S is large no interference phenomena can be observed and the vibrations from T_1 and T_2 behave like incoherent vibrations : the increase of dimensions of S opens the possibility of studying and intermediate domain between perfect coherence and perfect incoherence : the domain of partial coherence.

2-The Michelson interferometer with increasing OPD : Let us now consider a conventional Michelson interferometer illuminated with a very small monochromatic source (fig. 2) and let us vary the optical path difference (OPD) from zero to an importante value : it is well known that with conventional sources fringes will progressively vanish : their contrast will decrease from unity for a small OPD to zero when the OPD reaches values that depend on the spectral purity of the light : for conventional sources it is often expressed in fractions of a millimeter, reaching centimeters for good monochromatic sources and even one meter for sources used for the optical definition of standard lengths. Here the vibrations can be discomposed into spectral components contributing various interference fringe systems that are blurred if the OPD is important. We also can consider that the two vibrations coming from the two arms of the apparatus become incoherent if the OPD is important. In that case also a domain of partial coherence appears between perfect coherence and perfect incoherence but it is of completely different nature of the preceding case : the incoherence is closely related to the optical path difference Δ .

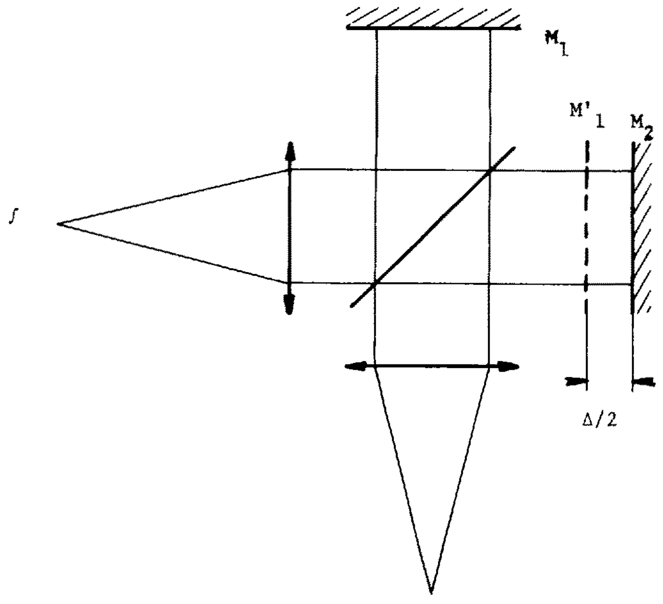


Fig. 2

The interferometer combines vibrations emitted by the same small source S at instants t and t + Θ, if Θ is the time difference Δ /c between the times of propagation through the two arms : this experiment raises the question of coherence of the vibration emitted by a small source at instants separated by a time shift Θ . This could be called a problem of temporal coherence, in opposition with the preceding case where the relative position of holes T_1 T_2 was the dominant factor : the problem of coherence of T_1 T_2 in the Th Young's experiment is a problem of spatial coherence.

B) The representation of vibrations by a fluctuating amplitude

The existence of interference phenomena establishes the wave nature of light and justifies the representation of the vibration by the expression R (a exp jω t) where a is a complex amplitude. In fact this representation is very convenient for a limited interval of time : if we take the example of a good monochromatic source able to produce interference fringes with an OPD of 1cm this representation will be valid over a time domain of about 3 10^{-11} sec. But if we increase the time delay Θ the experiment shows that fringes disappear : this representation is no longer valid for large OPDs or Θ s. In other words we have to consider that the amplitude a is no longer a constant but undergoes a relatively slow variation : during time interval of some periods T (T= 1.6 10^{-15} s for λ =0,5μm) a can be considered as a constant but it varies appreciably if the time interval is increased to approximately 3 10^{-11} s. In other words a becomes a function of time a (t) that does not vary in some periods but changes

appreciably during a time interval defined by the OPD that produces blurring of the fringes. The light emitted by a source being in fact the sum of contributions of a large number of atoms (or molecules) it obeys statistical laws and it is convenient to consider a (t) as a fluctuating quantity whose mean value $<a(t)>$is zero (otherwise there would be coherence for large values of Θ) the mean square$< a(t)\ a^{*}(t)>$ representing the energy.
On the other hand we have to notice that the time of observation is generally large when compared with the time delay θ that affects coherence : the observable quantity is in fact$< a\ (t)\ a^{*}\ (t)>$.

C)The degree of partial coherence

Let us now combine two amplitudes $a_1(t)$ and $a_2(t)$ in such a way that we could obtain fringes : allowing the waves W_1 W_2 to present an angle α between them (of between the wave vectors k_1 and k_2)we shall obtain inference fringes in the observation plane with a fringe separation $i = \lambda\ /\ \alpha$ If ϕ is the phase difference $2\ \pi\ \alpha y/\lambda$ at the point M (fig.3)

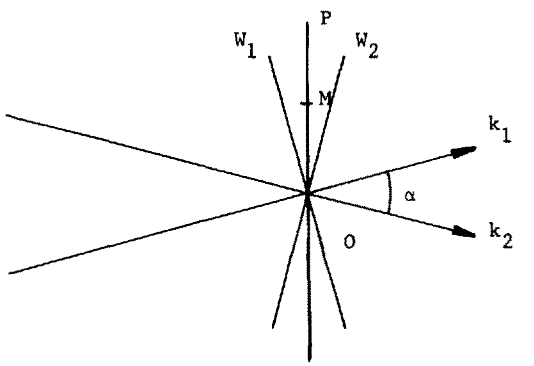


Fig. 3

the amplitude in M can be written, a_1 and a_2 being the amplitudes in O :

$$a_1(t)\ \exp\ (j\ \phi/2)\ +\ a_2\ (t)\ \exp\ (-j\ \phi\ /2)$$

The expression of the energy received in a point M where the phase diffêrence is ϕ will be :

$$<a_1\ a_1^{*}>+\ <a_2\ a_2^{*}>\ +\ 2\ R[\ <\ a_1\ a_2^{*}>\ e^{j\ \phi}]$$

where the importance of the variation of illumination with ϕ depends upon the quantity $<a_1\ (t)\ a_2^{*}\ (t)>$

This quantity determines the quality of the interference fringes

- its modulus determines the contrast of fringes

- its argument determines the contrast of fringes

We will write :

$$\Gamma 1.2 = \langle a_1 (t)\ a_2^{*} (t)\rangle$$

and call this complex quantity the <u>degree of partial coherence</u>.

It is possible to normalise this quantity by writing

$$\Gamma_N = \frac{\langle a_1\ a_2 \rangle}{(E_1\ E_2)^{1/2}}$$

E_1 and E_2 being the energies corresponding of the two beams.

It is easy to check that

- if $\Gamma_N = 0$ (two independant beams) no fringes appear ; this is the case of incoherence.
- if $|\Gamma_N| = 1$ the beams are completely coherent.

In fact the coherence of two light beams is expressed by the correlation function of the two fluctuating amplitudes a_1, a_2 normalised if necessary.

D) Spatial coherence - The Zernike-Van Cittert theorem.

Let us now show how this degree of partial coherence can be expressed, in the schematic case where the source is far from the reference plane where we intend to express the coherence between two points : one is located in the origin O and the other has coordinates y. z. (fig. 4)

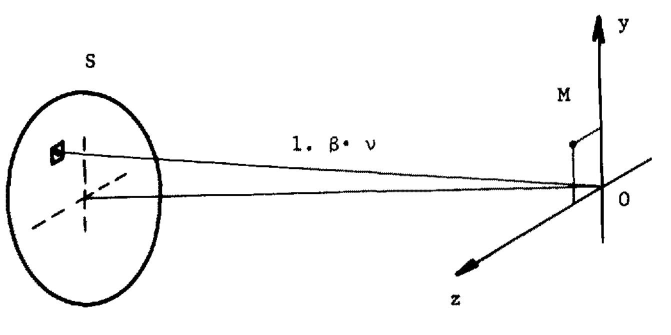


Fig. 4

If 1, β , γ are the components of a unit vector(β and γ being small)the density of energy on the source is a function $\varepsilon(\beta,\gamma)$ of the direction . Assuming that the various elements of the source are completely incoherent and supposing that the m^{th} element produces an amplitude $a_m(t)$ in point O it produces an amplitude $a_m(t) \exp(jk(\beta_m y+\gamma_m z)$ in point M. If now we compute the degree of coherence between O and M

$$\Gamma\ O.M = \langle \Sigma\ a_m(t) \quad \Sigma\ a_m^{*}(t) \exp(-j\ k\ (\beta_m\ y + \nu_m z))\rangle$$

taking into account the incoherence of the various elements we obtain as a result expressed in terms of an integral :

$$\Gamma = \iint \varepsilon(\beta,\gamma) \exp(-j\,k\,(\beta y+\gamma z))\; d\beta\, d\gamma$$

this is the Zernike-Van Cittert theorem :
The degree of coherence between two points receiving light from an extended source S is the Fourier transform of the energy luminance of the source.
This theorem applies very easily to the problem of the determination of stellar diameters by the measurement of fringe contrast as a function of the separation between two holes (or slits, according to the method of Michelson and Pease : assuming a uniform circular repartition for the star we obtain as degree of partial coherence $2J_1(Z)/Z$ with $Z = \frac{2\pi}{\lambda}\eta\, y$ where η is the angular radius of the star and y the separation between the holes. The first disparition of fringes will take place for $y = 1.2\ \lambda/2\eta$
This theorem can be generalised in the following way : S be ing an element of the source, (S A)and(SB) two optical paths, δ their difference the coherence between A and B can be expressed by the following relation

$$\Gamma = \iint \varepsilon(\beta,\gamma) \exp(-jk\,\delta(\beta,\gamma))\; d\beta\, d\gamma$$

where A and B can be non longer be located in the same plane.

E) Temporal coherence and spectral finesse.

Let us now look at the relation between the degree of partial coherence $\langle a(t)\, a^*(t-\theta)\rangle$ between vibrations emitted by the same source with a time delay θ and the spectral repartition of light ; it is well known that if we increase the "finesse" of the radiation i.e. if we decrease the width of the spectral band, we increase the possible O.P.D. in the interferometer and the possible delay θ . Let us now relate those various quantities :
If f (t) is the vibration entering a spectrometer : in the spectral plane of the spectrometer we will receive a repartition of amplitudes representing the Fourier analysis of f (t) ; if ν is the time frequency we will obtain a function g (ν) representing the spectral repartition of monochromatic amplitudes composing f (t) ; this simply results from the additivity of amplitudes of various frequencies ; if only one frequency is present one exit slit only receives light in the spectral plane and if the light is complex its amplitude is "Fourier analysed" by the spectrometer. In fact the energy received in the spectral plane is represented by $g(\nu)\, g^*(\nu) = S(\nu)$ which is the spectral repartition of light.
On the other hand it is well known that the interferogram I (Δ) turns out to be related to the F.T. of S :

$$I(\Delta) = \int_0^\infty S(\nu)\left(1+\cos 2\pi \frac{\Delta\nu}{c}\right) d\nu$$

or more precisely the variable part of I (Δ) is :

$$i(\Delta) = I(\Delta) - I_o = \int S(\nu) \cos 2\pi\Delta\nu/c \, d\nu$$

from the comparison of those relations we deduce (fig.5) that the spectral

$$f(t) \xleftrightarrow{\text{F.T.}} g(\nu)$$

$$\left.\begin{aligned} A(\theta) &= \text{autocorr.} f(t) \\ &= i(\Delta) \end{aligned}\right\} \xleftrightarrow{\text{F.T.}} S(\nu) = gg^{+}$$

Fig. 5

repartition of energy is the F.T. of the auto-correlation function of f(t) : $A(\theta) = \langle f(t) f(t-\theta)\rangle$ what means that $A(\theta)$ should be identified with $i(\Delta)$. In fact this is not astonishing : the interferometer acts on the amplitude f (t) as an autocorrelator, relating the amplitude at times t and $t - \theta$; a Michelson type interferometer could be called a time auto correlator for light amplitudes.

If now we develop the expression of $A(\theta)$ we obtain the following :

$$A(\theta) \propto R\left[\Gamma(\theta) \exp(j\omega\theta)\right]$$

which means that the two quantities $A(\theta)$ and $\Gamma(\theta)$ are very closely related : $A(\theta)$ is an oscillating quantity, the modulus being $|\Gamma(\theta)|$

Finally the energy spectrum $S(\nu)$ and the time partial coherence $\Gamma(\theta)$ are related as follows :

$$S(\nu) \propto \int_0^\infty A(\theta) \cos 2\pi\nu\theta \, d\theta = \int_0^\infty R|\Gamma(\theta) \exp j\omega\theta| \cos 2\pi\nu\theta d\theta$$

what shows that if $S(\nu)$ has a small extent, (very monochromatic light, high finesse) $\Gamma(\theta)$ has a large extent and the tolerable OPD on the interferometer is large.

F) The spatio temporal partial coherence.

In many cases the limitation of coherence results mainly from a predominant factor, (spatial or temporal) so that the relations established in D) or E) represent the situation. Nevertheless in some cases both geometry of the source and spectral finesse establish a limitation to coherence. In many of those cases it turns out that the O.P.D. does not vary appreciably when the point source moves on the source (otherwise the coherence would be very small) and if the source is homogeneous (the spectral repartition being everywhere the same) it is easy to dissociate spatial and temporal effects and the degree of partial coherence turns out to be the product of the two factors

$$\Gamma = \Gamma_S \; \Gamma_T$$

where Γ_S and Γ_T are spatial and temporal coherence expressed in sections D and E.

II Formation of optical images

We will mainly concentrate on the problem of the formation of optical images of extended objects : the basic mechanism is the superposition in the image plane of the images of the various elements of the objects : one little element of the object will contribute an element of the image represented by a diffusion function D (a diffraction image in the best cases, a more extended pattern in the presence of aberrations). In order to know how to combine the two images produced by two neighhouring elements of the object (fig. 6) we have to know whether vibrations coming from them are coherent, partially coherent or incoherent.

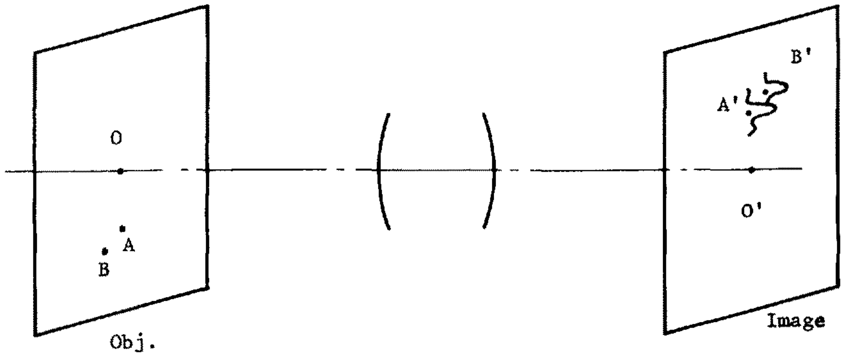


Fig.6

A) Incoherent illumination

The most common case is the incoherence : when taking an amateur photograph, an astronomical photograph, when separating various frequencies in a spectrograph as examples. In those conditions we have to add the energies coming from the various elements of the object, what leads to a convolution relation of the form $I = O \circledast D$: the image I is obtained by a convolution (in energy) between the object function (repartition of luminances) and the diffusion function D.

This can be expressed in the Fourier space : if i, o, and d are the Fourier transforms of I,O,D we can write $i = O \times d$ what means that the F.T. of the image is the product of transform of the object and of the transform of the diffusion pattern : the instrument can be characterised by an optical trnasfer function d that determines the way in which various spatial frequencies are transmitted by the instrument : this situation is very similar to that of the transmission of a temporal signal through an electronic or acoustic electronic device : the behaviour of the system can be known through a transfer function expressing the ability to transmit temporal frequencies. In optics we have to deal with spatial frequencies expressed in cycles per unit length and those components depend in fact on two parameters : the object or image are two

dimensional functions, as well as their transforms.
By application of the classical Fourier relations and taking into account the fact that the diffracted amplitude is the F.T. of the repartition of amplitudes on the pupil (pupil function) it turns out that the OTF is the autocorrelation of the pupil function: in the case of a circular pupil the OTF is represented by the common surface to two circles of radius α'(the angular radius of the pupil) shifted of a quantity proportionnal to the spatial frequency $1/p'$ if p' is the period of the component in the image plane (fig. 7). In those conditions the OTF for a perfect instruments is represented on fig.7 It has a strict cut off frequency $2\,\alpha'/\lambda$ what means that the instrument is unable to transmit any information outside of that bandwidth. one of the consequences is that it is possible to know completely the image by sampling the values at nodal points of a crossed grating having a period $\lambda/\ 4\ \alpha'$.
This means that for a perfect f/2 photo objective working for $\lambda=$ 0,5 μm the number of sampling points per square millemeter would be about $4\ 10^6$ and for a 24×36^{mm} image it reaches the tremendous number of $3.5\ 10^9$: this means that the quantity of information that can be recorded by optical means is very large and vice versa the necessary bandwidth for transmitting good T.V. image is very large.

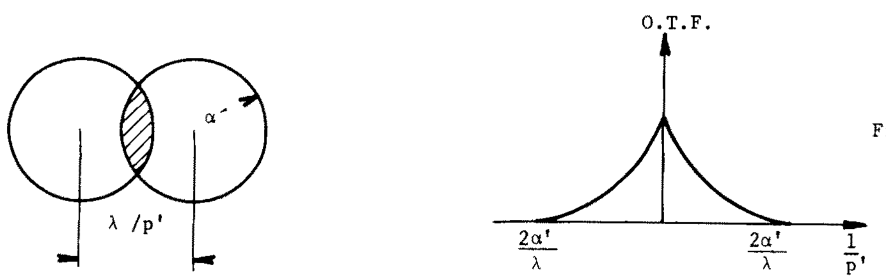


Fig.7

B) Coherent illumination

Fig.8 reprensents the typical mounting for performing completely coherent illumination: a point source S illuminates an object having an amplitude transmission Ω and the various elements f_o the object receive perfectly coherent vibrations. It becomes obvious that the image will be obtained by a convolution process applied to amplitudes :
The amplitude in the image is now the convolution of the amplitude in the object with the repartition of amplitudes in the diffraction image i.e. the F.T. of the pupillar repartition of amplitudes (the pupil function P). This process can also be established by studying the way in which light undergoes two successive diffractions :
The pupil receives the F.T. of Ω (first diffraction), transmits the product $P \times F\,T_{(\Omega)}$ and the second diffraction is responsible for producing in the image plane the

convolution F.T. (P) $\otimes$ Ω which is in complete agreement with the preceding conclusions. The optical transfer function turns out to be the pupil function P : according to the value of the spatial frequency the signal goes or not through the pupil : the OTF for amplitudes is a square function limited to the $\pm \alpha'/\lambda$; in coherent illumination the bandwidth is only one half of the value in incoherent illumination. Let us now point out one very interesting possibility of this coherent illumination : Abbe had already shown in his theory of the microscope that it is possible to transform the image of a periodic structure by eliminating various components for the diffraction phenomena in the pupil. We have shown how the attenuation of low spatial frequencies can help the perception of details and reinforce interesting informations : Fig. 9 is a typical example of those possibilities where a blurred text is treated in order to reinforce details and becomes readable by enhancement of details in the image. Another well known application is the technique of use of a matched filter suggested O'Neill and developped later by various authors : Fig. 10 is a typical example where the filter has been matched to the recognition of the letter e.

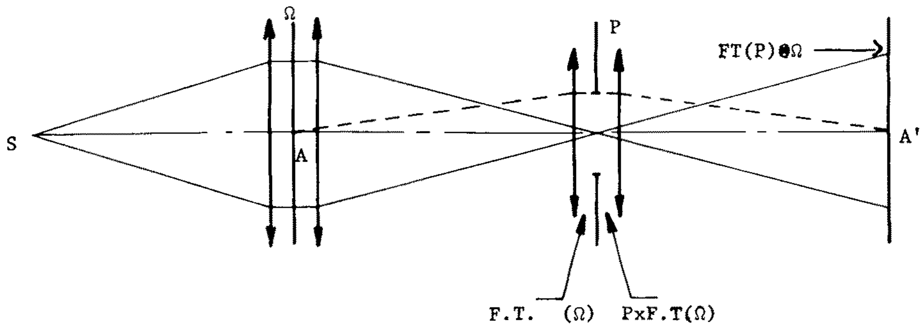


Fig.8

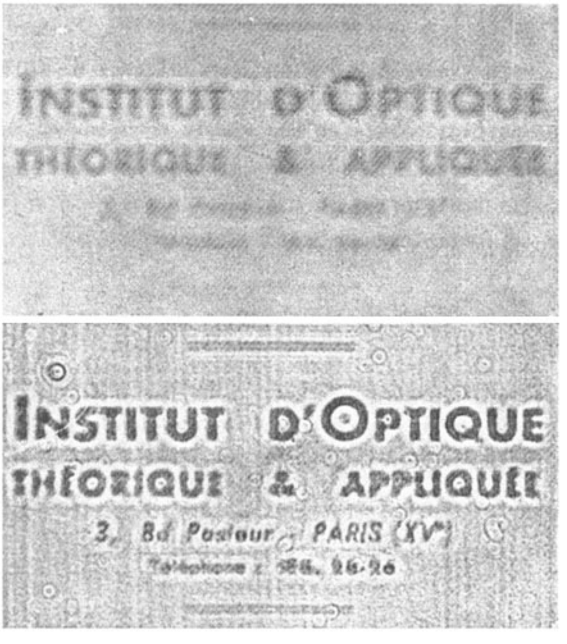


Fig. 9

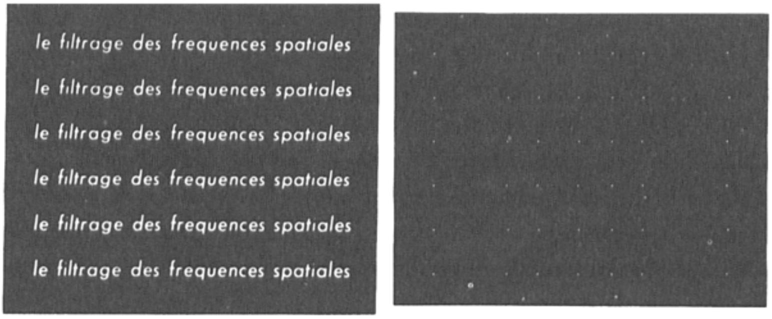


Fig. 10

Reconnaissance de la lettre e

C) Partially coherent illumination

The coherent illumination technique has the drawback that the illuminating point limits severely the total flux ; in many cases (microscopy, projection technique) it is necessary to use a source having an appreciable extension but this introduces a partially coherent illumination : if the source can be assimilated to a circular disc of angular radius α_c de degree of partial coherence is represented by a fonction $2\ J_1\ (Z)/Z$ with $Z = \frac{2\pi}{\lambda} \alpha_c y$; points that are very close are coherent ; then the coherence decreases and goes to zero for $y = 1.2\ \lambda / 2\ \alpha_c$ and oscillates and vanishes for distant points.

The relation between object and images becomes much more complicated : if Γ represents the degree of partial coherence between two points M_1 and M_2 according to the Zernike-Van Cittert theorem

$$\Gamma = \int \quad (S)\ e^{jkS\ (M_1 - M_2)}\ d\,S$$

where ζ (S)represents the energy repartition on the condenser.

The repartition of illumination in the image is expressed as follows :

$$I(M') = \iint \Omega(M_1)\ E\ (M' - M_1)\Omega^*\ (M_2)\ E^*\ (M'-M_2)\Gamma\ (M_2-M_1)\ dM_1 dM_2$$

In the general case of partial coherence the simple convolution relations are no longer valid and the relations are much more cumbersome. Nevertheless when the contrast of the object is small (what is often the case in optical microscopy) an approximation can be made and leads to a simple convolution relation

$$I(M') = I_o + 2R_e\left[\int \omega\ (M)\ E\ (M'-M)E_c^*\ (M'-M)\ d\ M\right]$$

where E and Ec are

- the repartition of the amplitude in the diffraction pattern produced by the full pupil
- the repartition in the pattern produced by the pupil limited to the aperture of the condenser.

Fig.11 represents the evolution of the "response function EEc for a perfect instrument where the pupil and the condenser are circular ; it is then possible to fill the gap between perfect coherence and perfect incoherence and study the evolution of the image or of the OTF (fig.12) between the two extreme cases.

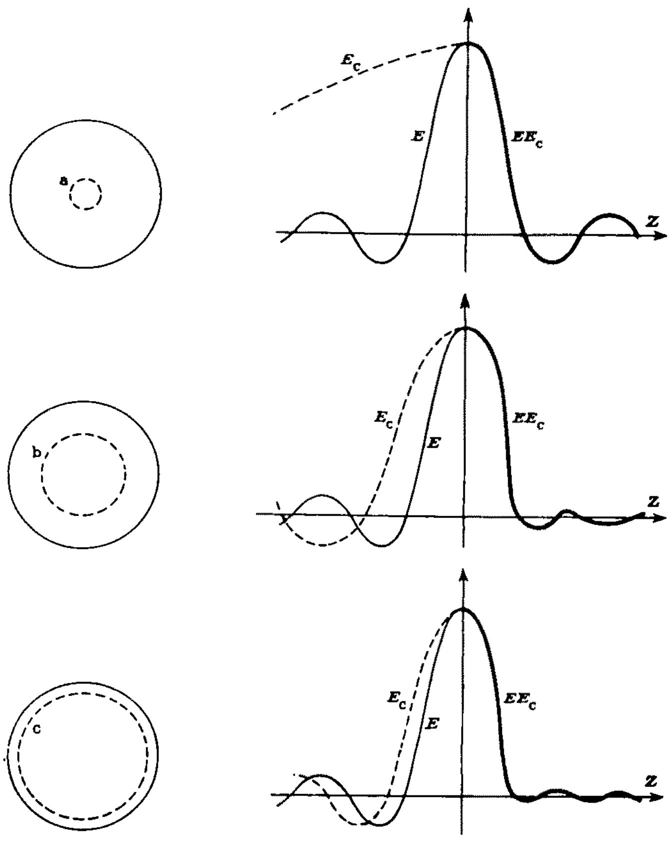
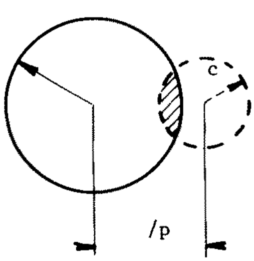


Fig. 11



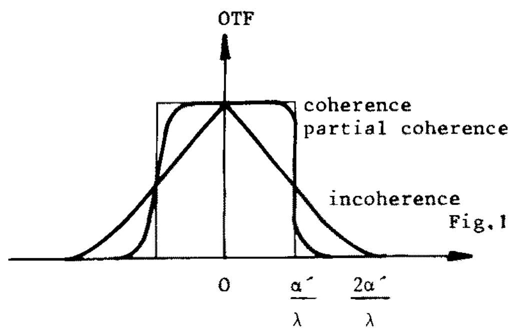


Fig.12

Bibliography

A.MARECHAL et M. FRANCON, Diffraction-Structure des Images Ed.Masson,Paris 1970
BORN M & WOLF E., Principles of optics. Perg.Press.
BLANC-LAPIERRE & DUMONTET P., Rev.Opt.,34, 1955,1.
HOPKINS H.H., Proc. Roy. Soc., 217, 1953, 408
MENZEL E., Optik, 15, 1958, 460
SLANSKY S., Opt.Acta, 2, 1955, 119 ;J.Phys.,16,1955,13 S.
WOLF E., Proc. Roy. Soc.,225, 1954, 96 ; Proc. Roy. Soc.,230,1955,246
ZERNIKE F., Proc.Phys.Soc., 61,1948,158.