stellar interferometry : an overview about basics

Yves Rabbia
Observatoire de la Côte d'Azur
UMR Lagrange, Nice, France

rabbia@oca.eu

Perso 06 24 33 84 96

stellar interferometry : an overview about basics

sections

- introduction : a problem raised
- science context and motivation
- few academic reminders
- basics for interferometry and aperture synthesis
- limitations and subsequent needs
- interferometers : principle, production, typology
- difficulties in real world (and some remedies)
- managing with data and some results
- quick-look at some alternative HAR methods
- nulling interferometry and coronagraphy

just talking (where are we ??)

time-signals can be described
either by amplitudes (signal) or by frequencies (spectrum)

similarly
there are two ways to describe brightness distributions :
the direct description, based on coordinates (image)
the other one, based on spatial frequencies (spatial spectrum).

For brightness distributions, when images are not at hand
a "spectrum analyser" must be found
Once available, coming back from the spectrum to the image
could be possible.

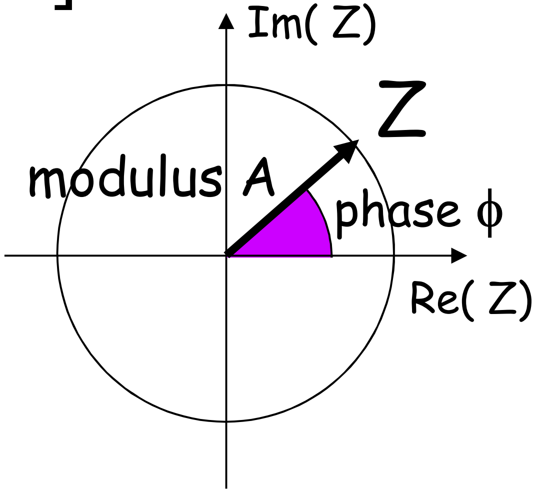
This is the goal of aperture synthesis

Thus we have now to conceive and built
such a spectrum analyser

interferometry : the very first tool, not a joke !

$$z_1 \in C \,,\; z_2 \in C$$

$$|z_1 + z_2|^2 = |z_1|^2 + |z_2|^2 + 2.Re[z_1 . z_2 *]$$



in other words

$$\left| A_1.e^{i.\phi_1} + A_2.e^{i.\phi_2} \right|^2 =$$

$$= A_1^2 + A_2^2 + 2.A_1.A_2. Re\left[e^{i.(\phi_1 - \phi_2)} \right]$$

$$= A_1^2 + A_2^2 + 2.A_1.A_2. cos(\phi_1 - \phi_2)$$

no comments, see later on

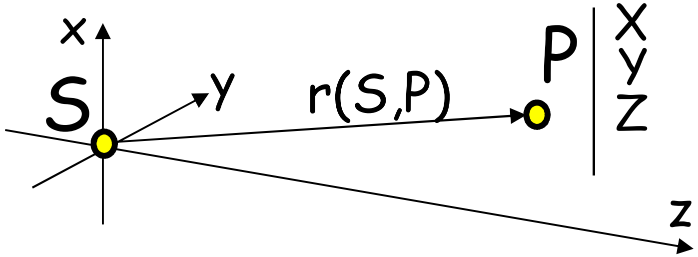
an important auxiliary : Fourier optics

Fourier optics

- key protagonists :
  complex amplitude and wavefront
- building the tool :
  - a very quick look
  - Huyghens Fresnel principle
  - here comes Fourier formalism
- illustrative example(s)

complex amplitude of fields

source S, observation point P



Wave at S $\quad V_S(t) = A.\exp(\, i.\, 2\pi.\nu.t)$

At P, same behaviour but time-delayed

$V_P(t,x,y,z) = A.\exp(\, i.2\pi.\nu.(t - r/c))$

$$V_P(t,x,y,z) = A.\exp(\, i.(\nu.\, t - r/\lambda)\,)$$

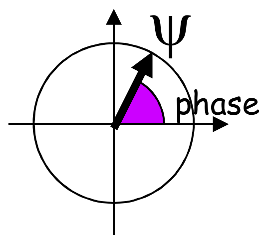
for imaging, relevant information is in the phase distribution $2\pi.r/\lambda$ at each observation point, so the useful description is given by the complex amplitude

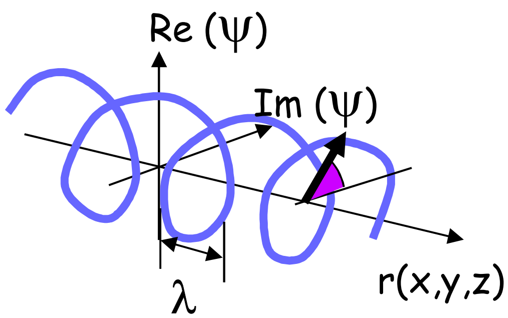
It describes the shape of the wavefront (equiphase surface)

$$\psi\,(x,y,z) = A.\exp(\, i.2\pi.\, r/\lambda)) = A.\exp\,[i.\phi(x,y,z)]$$

intensity
(power density)

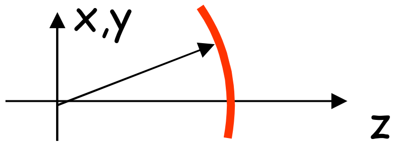
$$|\psi|^2 = A^2$$





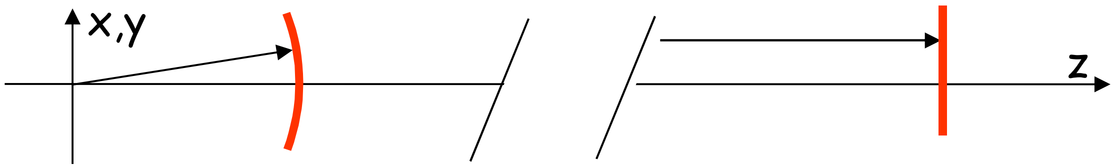
the meaning of "phase of the field" : wavefront

spherical wave



$$r^2(x,y,z) = x^2+y^2+z^2$$

plane wave (on-axis point-like source at infinity



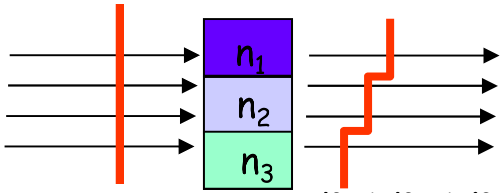
$$r\,(x,y,z) = z$$

the phase mathematically describes the shape of the wavefront

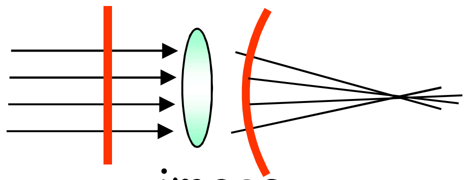
wavefronts are "equiphase surfaces"

propagation within material medium
("n" is refractive index)

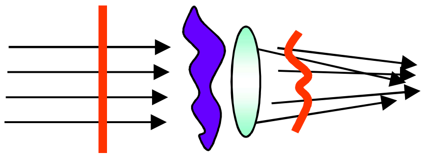
$$\phi = \frac{2\pi}{\lambda}.index.pathlength = \frac{2\pi}{\lambda}.n.r$$



$n_1 < n_2 < n_3$



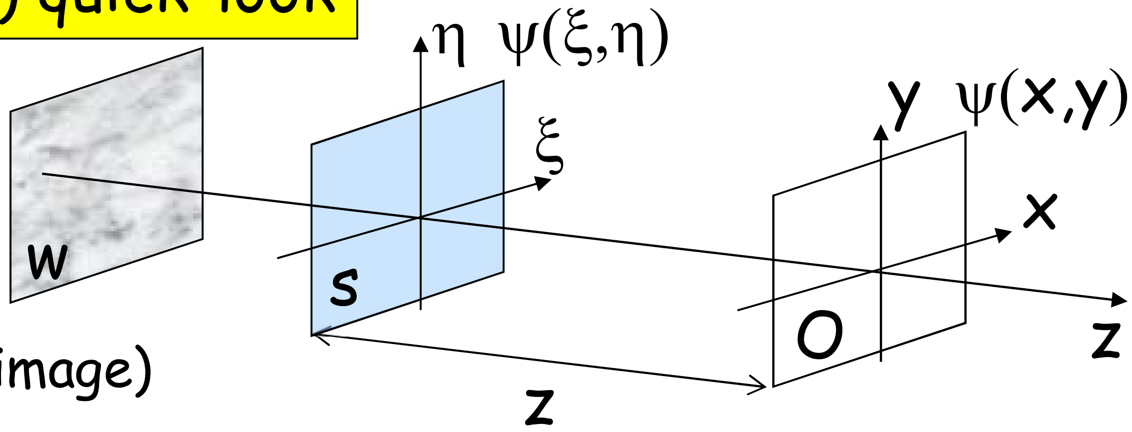
image



abberated image

Fourier Optics a (very) quick-look

frequent set-up



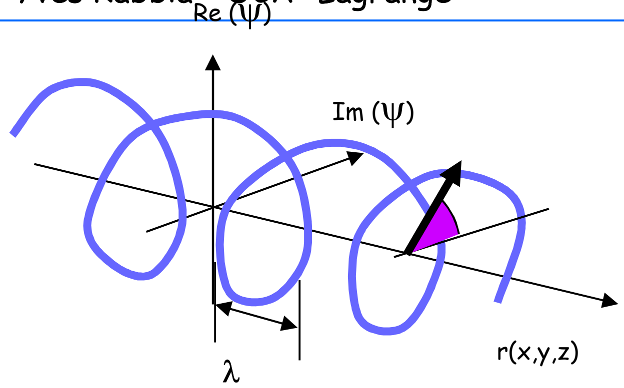
w : incoming wavefront
s : diffracting screen
O : observation screen (image)

$\psi(\xi,\eta)$: amplitude after "s"
$\psi(x,y)$: amplitude at O

summary :
if I know $\psi(\xi,\eta,0)$ amplitude transmitted by "s" (pupil plane)
I can calculate $\psi(x,y,z)$ amplitude over screen O (image plane)
Intensity at O = squared modulus of $\psi(x,y)$

tools to do that :
approximations + Huyghens-Fresnel principle

Huyghens-Fresnel principle

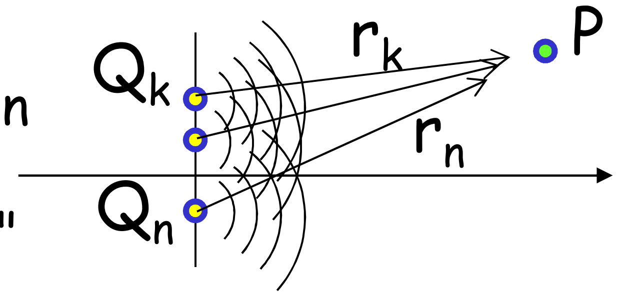


complex amplitude

already seen : as the field propagates the phase increases according to length of traveled path

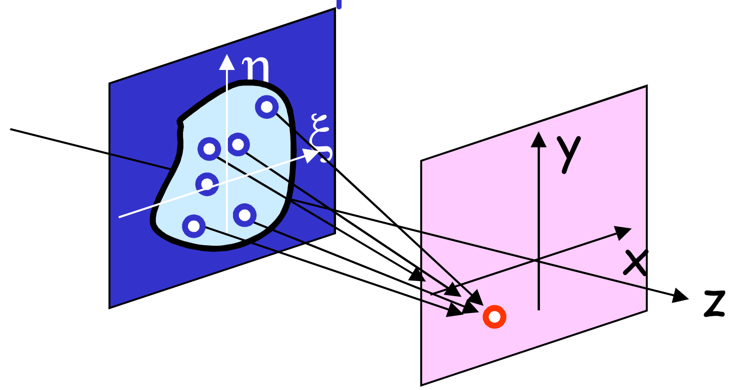
Huyghens Fresnel principle:

every point Q_n within an amplitude distribution emits a spherical wave, all waves are synchronous but not necessarily "in phase"



The field seen by P is the sum of the amplitudes of the spherical waves (complex numbers)

$$\psi(P) = \sum \psi(Q_n).\exp(i.\frac{2\pi}{\lambda}.path))$$



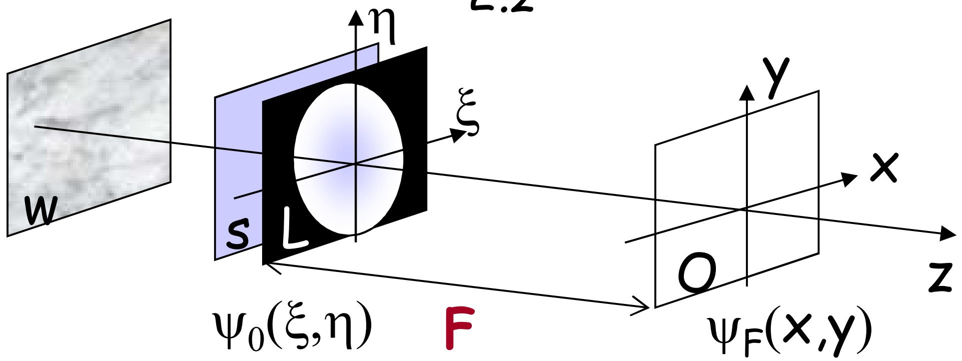
$$\psi_z(x,y) = \int \psi_0(\xi,\eta).\exp(i.\frac{2\pi}{\lambda}.r(\xi,\eta,x,y)).d\xi.d\eta$$

here comes Fourier formalism

approximations : scalar field $z \gg \xi,\eta,x,y$ $\lambda \ll \xi,\eta,x,y$

+ Pythagore : $r(\xi,\eta,x,y) \cong z + \dfrac{[x-\xi]^2+[x-\eta]^2}{2.z}$

most frequent set-up :
L lens, focal F
ψ_0 : transmitted amplitude



then (some algebra)

$$\psi_F(x,y) \propto \int [\psi_0(\xi,\eta)]\, e^{-i.\frac{2\pi}{\lambda.F}.(x.\xi+y.\eta)} .\ d\xi.d\eta$$

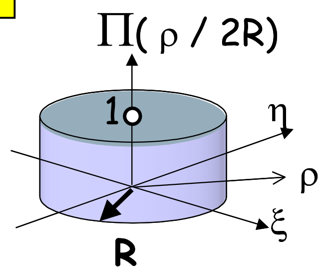
other writing : **WELCOME FOURIER !**

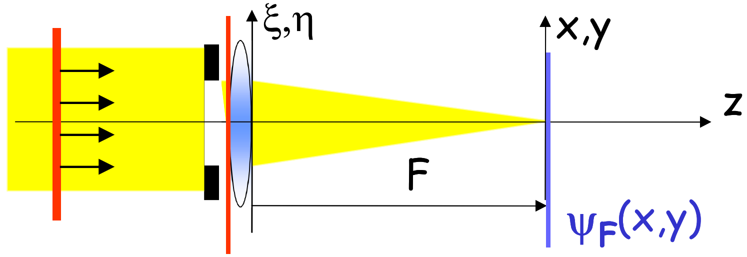
$$\psi_F(u,v) \propto \hat{\psi}_F(u,v) = \int [\psi_0(\xi,\eta)]\, e^{-i.2\pi.(u.\xi+v.\eta)} .\ d\xi.d\eta$$

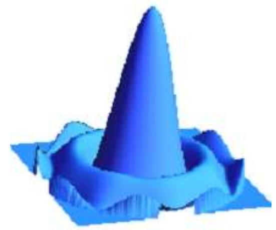
with u, v spatial frequencies : $u = \dfrac{x}{\lambda F}, v = \dfrac{y}{\lambda F}$

examples

circular aperture





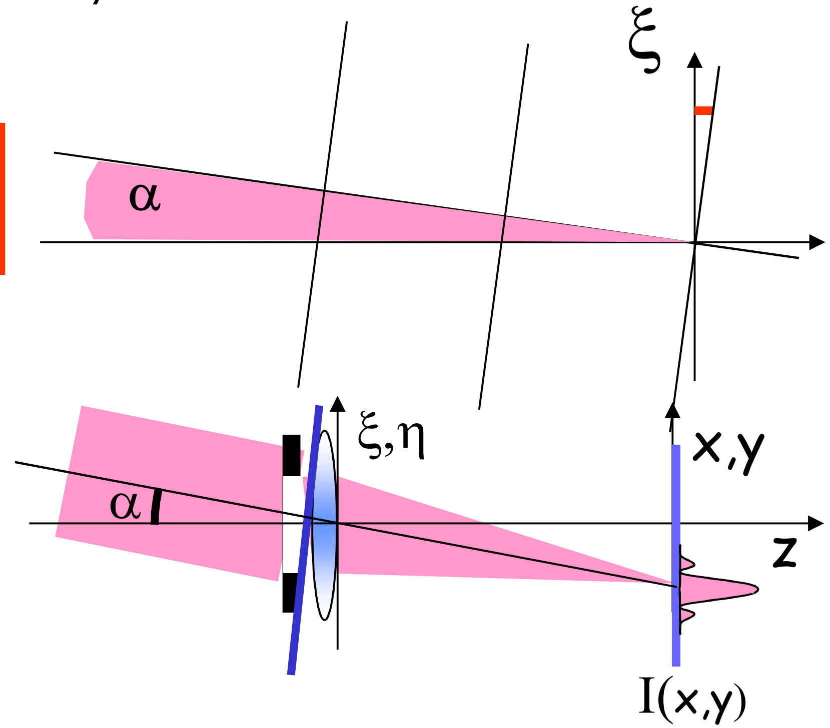


$$\psi_F(\rho) \propto 2.\frac{J1(Z)}{Z} \quad \text{with } Z = 2\pi.\rho.R$$

tilted incoming wavefront

$$\varphi(\xi) = -\frac{2\pi}{\lambda}.path = -\frac{2\pi}{\lambda}.\xi.\alpha$$

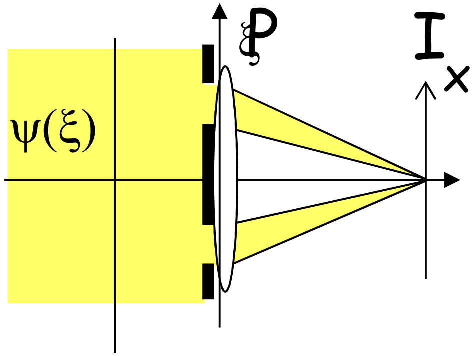
$$P(\xi).exp(-i.\frac{2\pi}{\lambda}.\xi.\alpha) \Leftrightarrow \hat{P}(u) * \delta(u-\alpha)$$

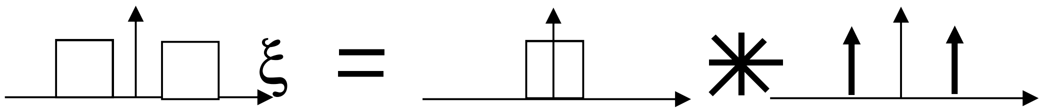


warning ! mixed and ill-used notations

another example

double slit as aperture

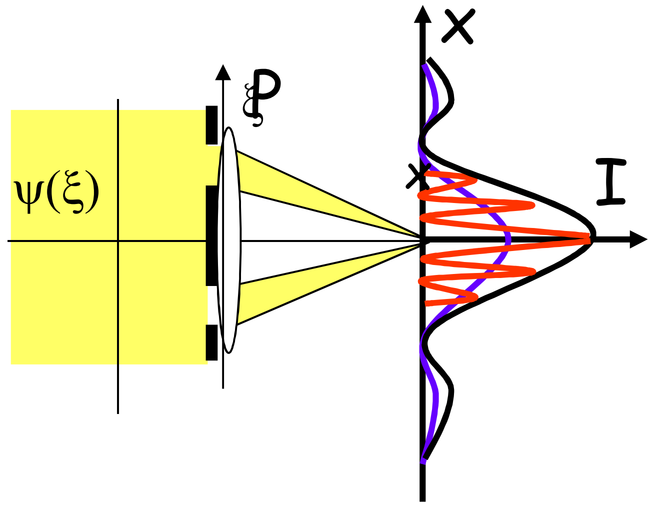




$$\psi_0 = P_+ + P_- = P * [\delta_+ + \delta_-]$$

$$\psi_F = \hat{P} . [e^{i\varphi} + e^{-i\varphi}]$$

$$I = |\psi_F|^2 = 2.I_0.[1 + cos(\ \ )] \quad with \quad I_0 = |\hat{P}|^2$$




ECUADOR
Malgusto.com

other useful auxiliaries :

linear filtering
and
transfer function

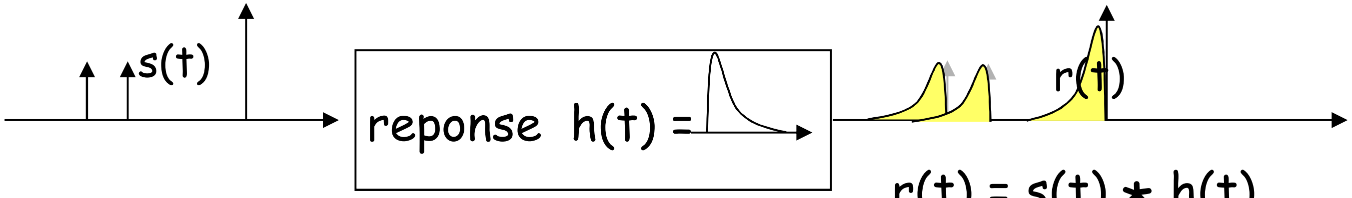
overview

signal → physical system → response

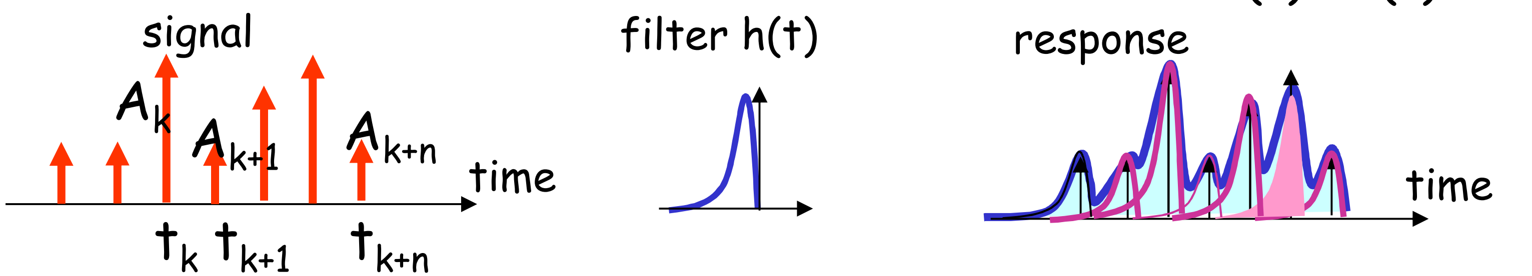
dirac-like input → physical system → response to dirac (impulse)

syst. phys. = linear filter ⟺ input-output relationship is a convolution

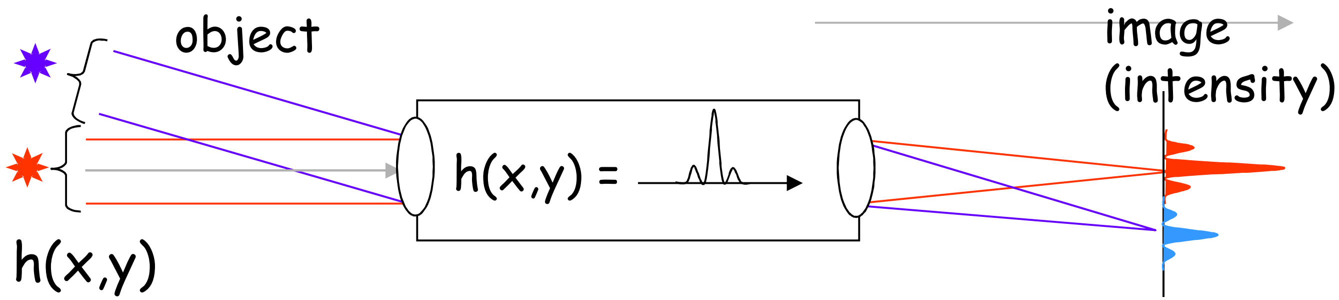
time signals



$r(t) = s(t) \ast h(t)$



2D-objects



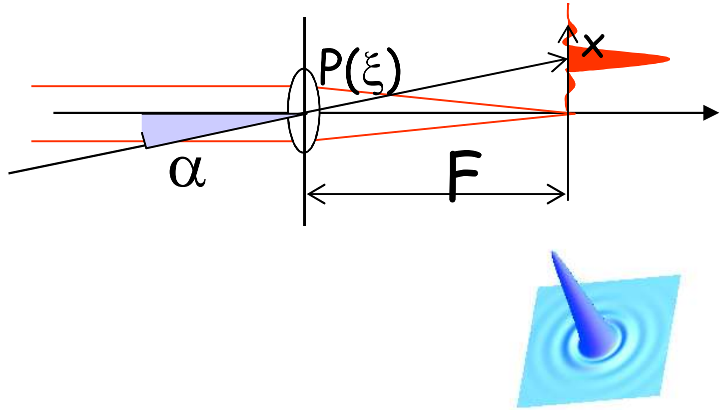
$I(x,y) = O(x,y) \ast h(x,y)$

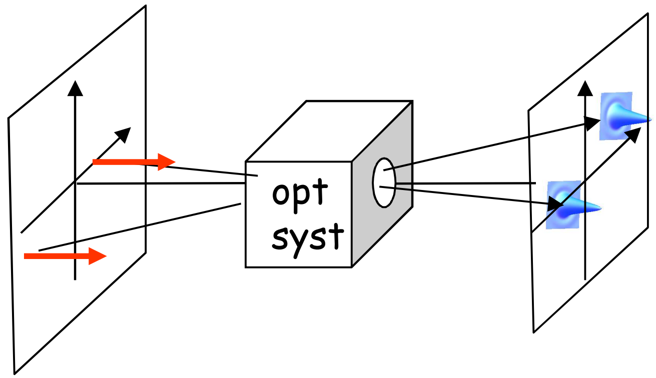
bi-dimensional situation

object-image relationship in the direct space (coordinates space)

$I(x,y) = O(x,y) * h(x,y)$

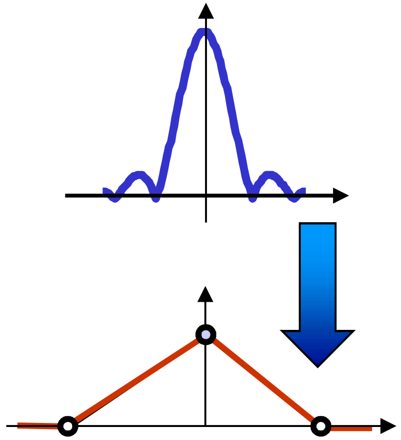
h(x,y) = response to dirac : Point Spread Function (typical : Airy pattern)





description in the frequencies space (Fourier space)

$$\hat{I}(u,v) = \hat{O}(u,v).\hat{h}(u,v) = \hat{O}(u,v).T(u,v)$$

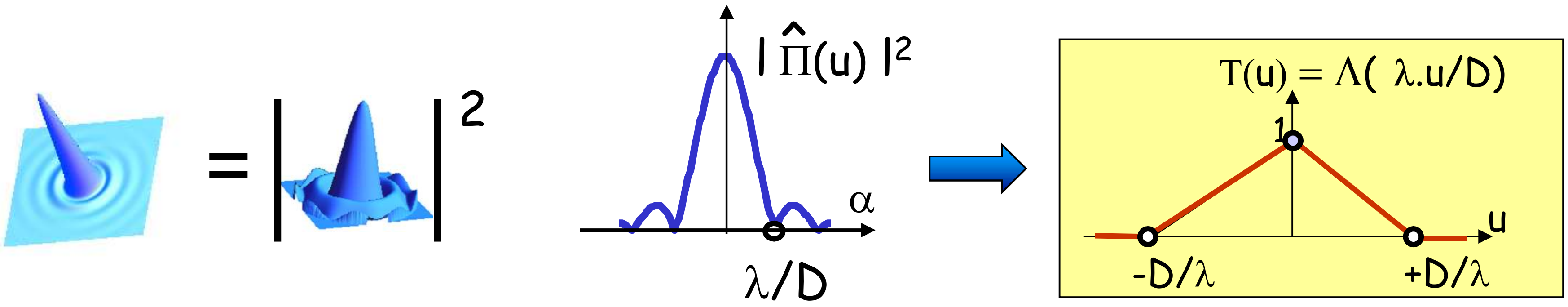


T(u,v) = transfert function = FT of impulse response
shows how frequencies are transmitted : gain (complex)

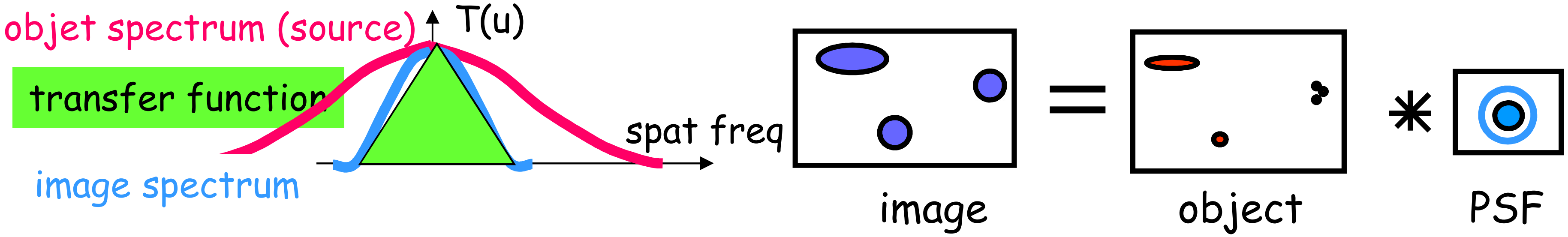
Also : T(u,v) shaped like autocorrelation of pupilla (Rayleigh theorem)

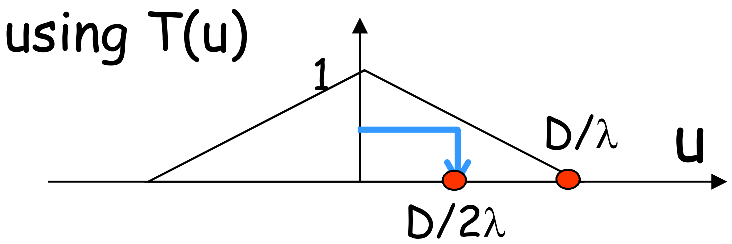
transfer function, mostly pictorial

T(u) = FT of PSF (Airy) ➡ $T(u) = \Lambda(\frac{\lambda . u}{D})$

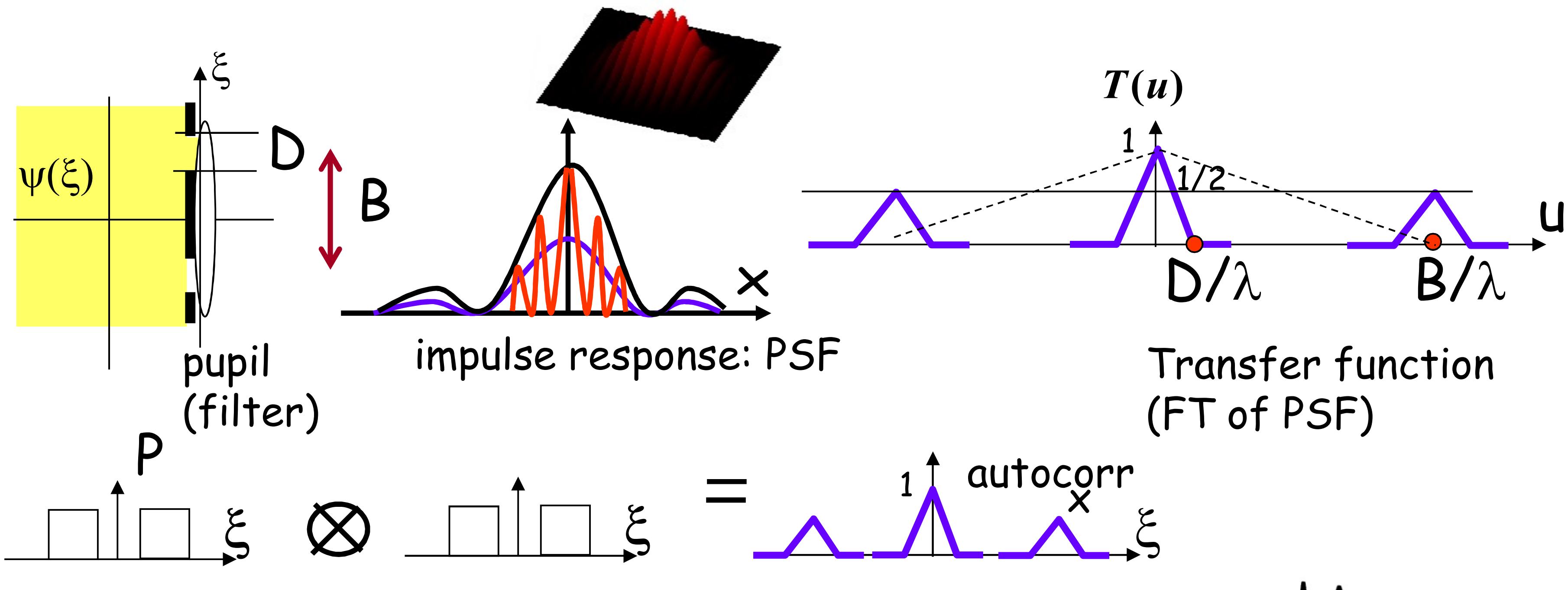


D/λ : cut-off frequency. Consequence : higher frequencies are lost

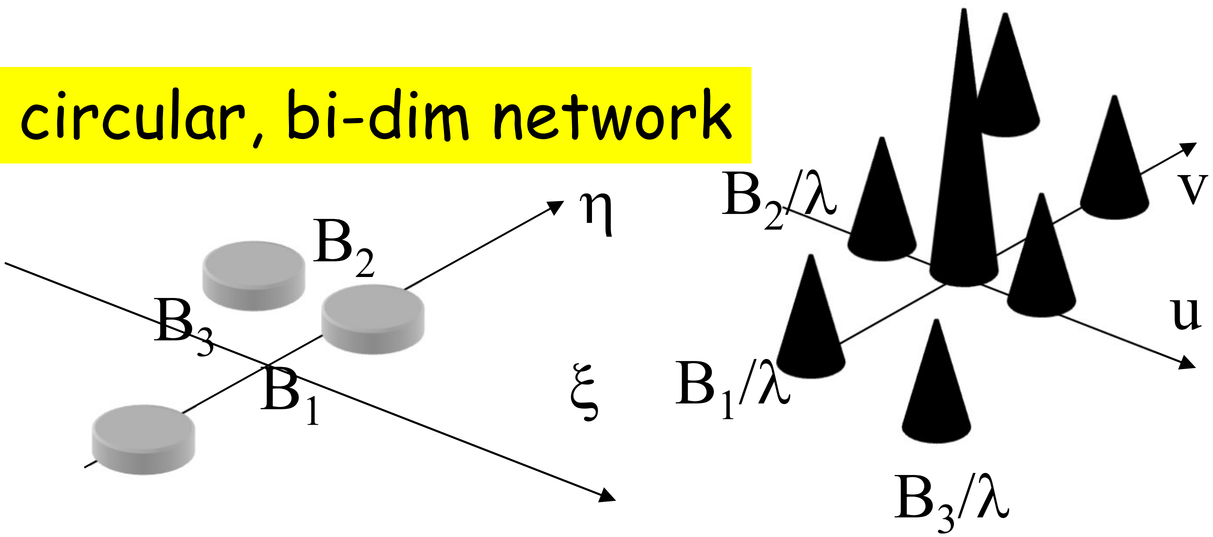




a special one (anticipation) : double slit aperture



further ? several circular, bi-dim network



fundamental principles :

Coherence
and
the second key : VCZ theorem

defining coherence : a mutual notion

coherence of fields is the ability they might have
to produce observable interferences when mixed
(duration equal or larger than the detector integration time say 10^{-6} s)

coming back to our first tool

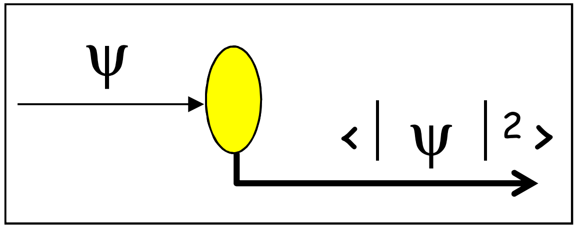
quadratic detection, energy detection (visible and infrared)

incoming field ψ

detection process : output = $<|\psi|^2>_\tau$

τ = integration time,

notation $<>_\tau$ means "averaging over τ "



"mixing fields" means ψ_1 and ψ_2 arriving together on the detector

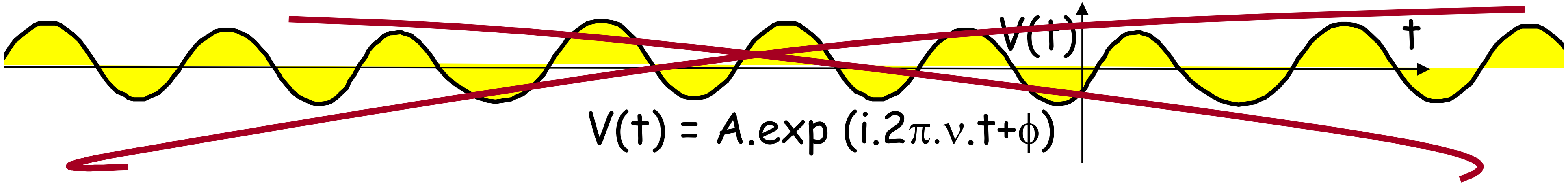
$$<|\psi_1+\psi_2|^2> \;=<|\psi_1|^2>+<|\psi_2|^2>\;+\;\mathbf{2.Re}(<\psi_1\,.\,\psi^*{}_2>)$$

$$<|\psi_1+\psi_2|^2> \;=\; \boxed{I_1+I_2} \;+\; \boxed{\mathbf{2.Re}(<\psi_1\,.\,\psi^*{}_2>)}$$

energy terms — interference term

how to quantify this coherence ability ? _ 1

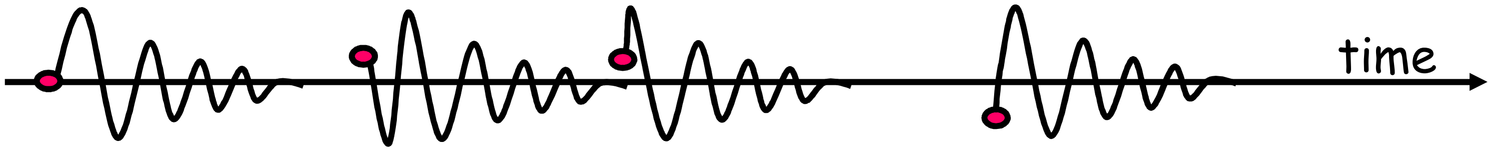
the single cosine model is not compatible with observation of interferences



need another model for wave of light, a model with random features

a relevant model : train of damped oscillations with random emission time and phase at origin

length of a wagon : $\tau_c \approx 1/$ spectral $\Delta\nu$



$$\psi(t,\phi) = \sum_k \psi_k$$

$$\text{with } \psi_k = A(t - t_k).\exp(i 2\pi\nu(t - t_k) + \phi_k)$$

how to quantify this coherence ability ? _2

$$< |\psi_1 + \psi_2|^2 > \; = \boxed{I_1 + I_2} + \boxed{2.\mathrm{Re}(<\psi_1 . \psi^*_2>)}$$

$<\psi_1.\psi_2^*>$ is something like a covariance
it could serve but depends on incoming amplitudes (leading to intensities)
better to define
a dimensionless quantity free from differences in intensity
A complex number which modulus varies between 0 and 1 is appropriate

complex degree of coherence

with $I_k = < |\psi_k|^2 >$

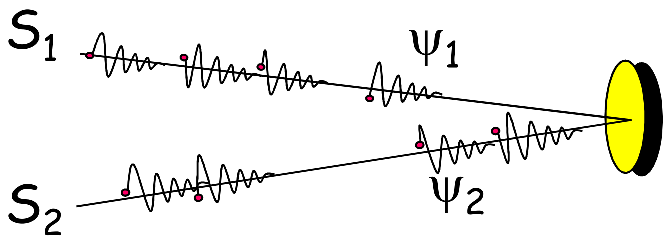
$$\gamma_{12} = \frac{<\psi_1 . \psi_2^* >}{\sqrt{I_1 . I_2}}$$

γ_{12} can be seen as a complex normalized covariance

about addition of fields : generic situations _ 1

incoherent addition

the interference term is destroyed
by averaging over the random phases



$$I = \langle |\psi_1 + \psi_2|^2 \rangle = \left\langle \left| \sum_k \psi_{1k} + \sum_n \psi_{2n} \right|^2 \right\rangle$$

$$I = \left\langle \left| \sum_k \psi_{1k} \right|^2 \right\rangle + \left\langle \left| \sum_n \psi_{2n} \right|^2 \right\rangle + 2.Re\left[\left\langle \sum_k \sum_n \psi_{1k}.\psi^*_{2n} \right\rangle \right]$$

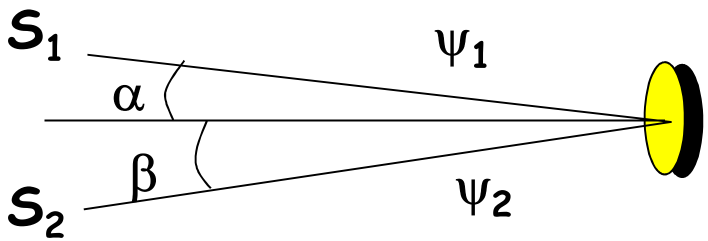
all $\langle \psi_{1k}.\psi^*_{2n} \rangle$ *convey* $\langle exp[i(\phi_k - \phi_n)] \rangle$ *and averaging results in zero*

$I = I_1 + I_2$ + nothing !

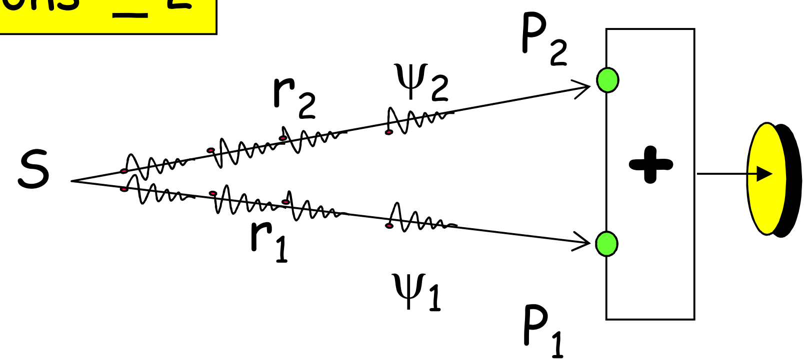
two distinct sources cannot produce
an observable interference term

consequence : introducing
a non-coherence relationship

$$\langle \psi(\alpha).\psi^*(\beta) \rangle = \langle |\psi(\alpha)|^2 \rangle.\delta(\alpha - \beta)$$



about addition of fields : generic situations _ 2



coherent addition

let us consider a machine performing addition of field collected at P1 and P2 from a unique point-like source and nearly equal paths r_1 and r_2

each "wagon" is like : $A(t-t_k).exp\left[i.\frac{2\pi}{\lambda}r_{1\,or\,2}+i.\phi_k\right]$

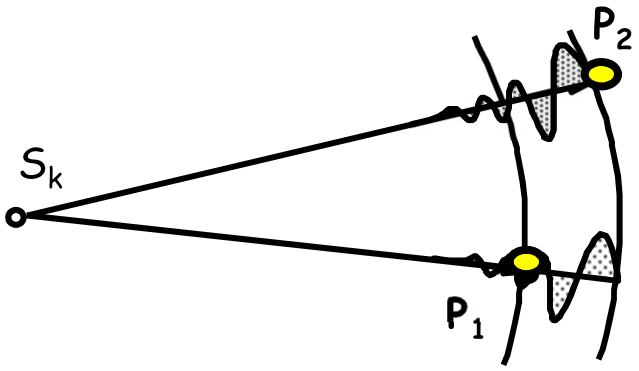
and interference term conveys components like :

$$< A(t-t_k-\frac{r_1}{c}).A(t-t_k-\frac{r_2}{c}).exp\left[i.\frac{2\pi}{\lambda}.(r_1-r_2)+i(\phi_k-\phi_k)\right] >$$

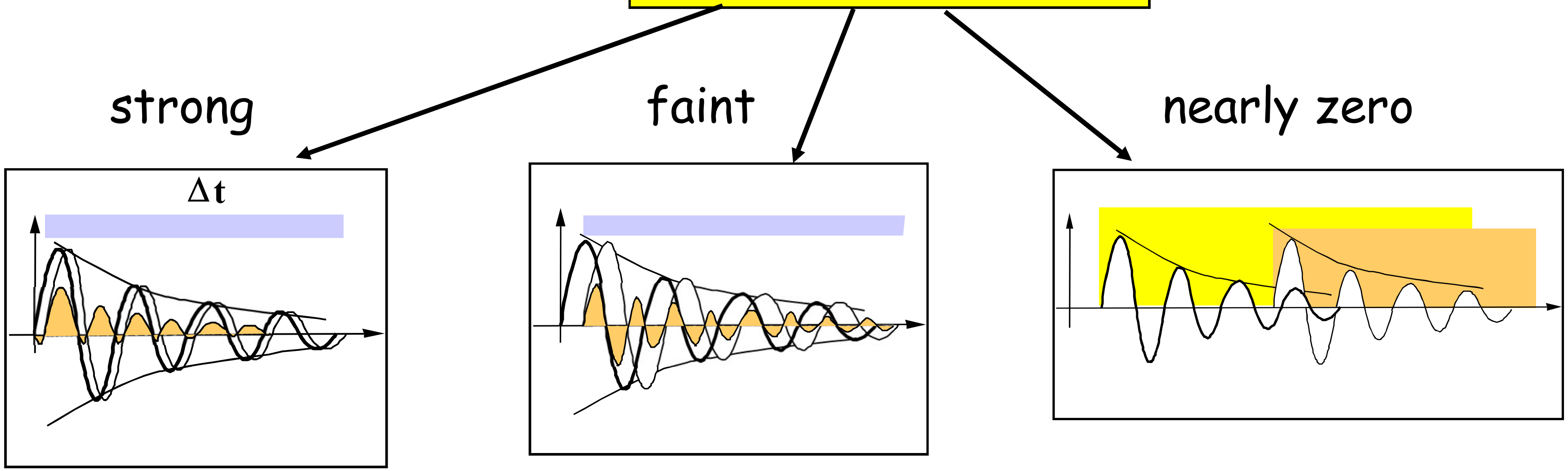
now averaging over ϕ has no effect, and interference term remains, in spite of the effect of statistics. zero

However it is non-zero only if "wagons" are overlaping what requires that (r1-r2) must be smaller than the length of a wagon (subsequently named : coherence length)

path difference must be small enough



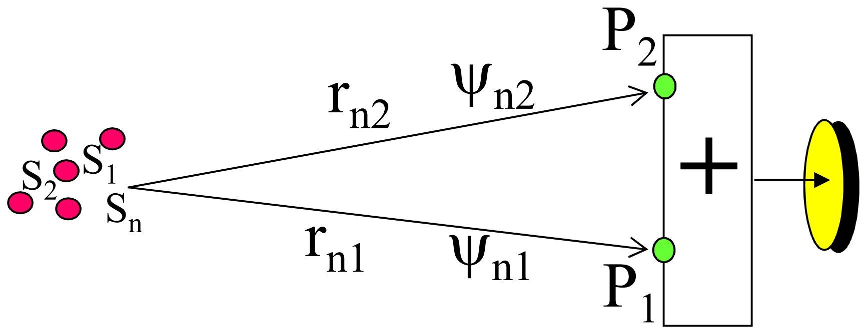
$$< \psi_1 . \psi_2^* >$$



a third situation : synthesis of the two previous ones _1

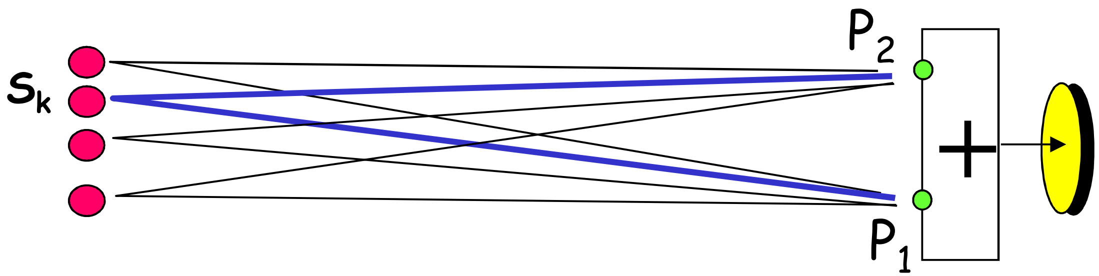
addition partially coherent

several sources $S_1, S_2, \ldots S_n$
and the addition machine



phenomenology

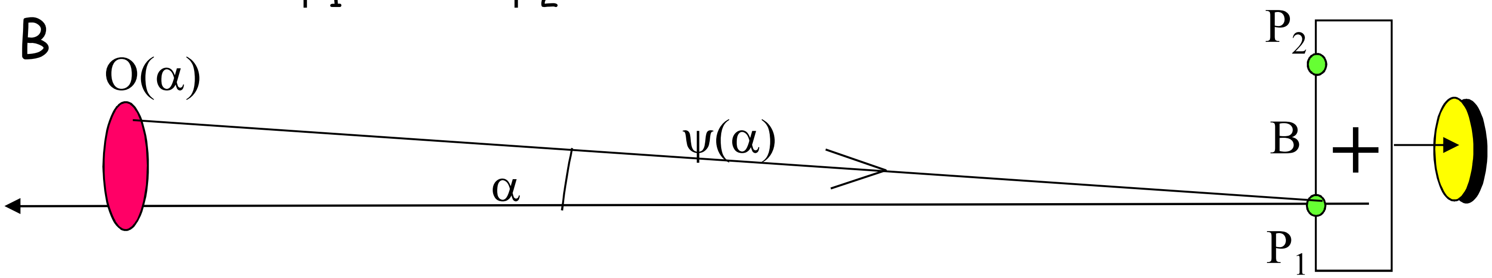
all sources Sk yield an observable interférence term (situation 2), independantly one to another, each term depends on (r_{k1} - r_{k2}) in other words depends on its location



the resulting value for the interference term depends on the (angular) extension of the distribution of point-like sources

a third situation : synthesis of the two previous ones_2

$O(\alpha)$ = continuous distribution of point-like sources
field from direction "α" noted $\psi(\alpha)$, such as $\langle|\psi(\alpha)|^2\rangle = O(\alpha)$
field incoming on P1 noted ψ_1 and ψ_2 for P2
distance P_1P_2 is B



On P_1 and P_2 comes the sum (over α) of fields from directions (α)
respective paths are $r_1(\alpha)$ and $r_2(\alpha)$

$$\psi_1 = \int \psi(\alpha).\exp\left(-i.\frac{2\pi}{\lambda}.r_1(\alpha)\right).d\alpha \qquad \psi_2 = \int \psi(\alpha).\exp\left(-i.\frac{2\pi}{\lambda}.r_2(\alpha)\right).d\alpha$$

then we have (classical expression for a product of integrals)

$$<\psi_1.\psi_2^*> = \int\int <\psi(\alpha).\psi^*(\beta)>.\exp\left(-i.\frac{2\pi}{\lambda}.[r_1(\alpha)-r_2(\beta)]\right).d\alpha.d\beta$$

look !

so what ?

$$<\psi_1.\psi_2^*> = \int\int <\psi(\alpha).\psi^*(\beta)>.exp(-i.\frac{2\pi}{\lambda}.[r_1(\alpha)-r_2(\beta)]).d\alpha.d\beta$$

and thanks to non-coherence relation

$$<\psi_1.\psi_2^*> = \int\int O(\alpha).\delta(\alpha-\beta).exp(-i.\frac{2\pi}{\lambda}.[r_1(\alpha)-r_2(\beta)]).d\alpha.d\beta$$

and thanks to dirac properties

$$<\psi_1.\psi_2^*> = \int O(\alpha).exp(-i.\frac{2\pi}{\lambda}.[r_1(\alpha)-r_2(\alpha)]).d\alpha \int \delta(\alpha-\beta).d\beta$$

= 1

and so and so ?

$$<\psi_1.\psi_2^*> = \int O(\alpha).exp(-i.\frac{2\pi}{\lambda}.[B.\alpha]).d\alpha$$

isn't that beautiful ? Fourier strikes back

$$<\psi_1.\psi_2^*> = \hat{O}(\frac{B}{\lambda})$$

last step towards complete happiness

a by-product:

$$<\psi_1 . \psi^*_1> \; = \; < |\psi_1|^2 > \; = \; I_1 \; = \; \int O(\alpha) \, . d\alpha \; = \hat{O}\,(u=0)$$

same for ψ_2

let us then recast the degree of coherence

$$\gamma_{12} = \frac{<\psi_1 . \psi_2^* >}{\sqrt{I_1 . I_2}} = \frac{\hat{O}\left(\frac{B}{\lambda}\right)}{\hat{O}(0)}$$

in other words

with $P_1P_2 = B$ and with λ

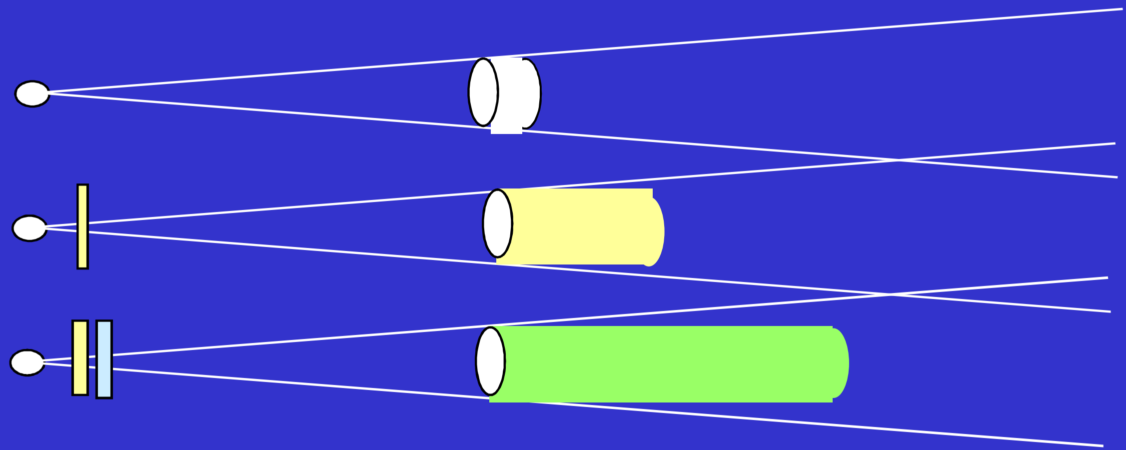
the degree of coherence is the FT at frequency B / λ of the angular brightness distribution of the source normalized on its FT at origin

this simply is the Van Cittert & Zernike theorem

coherence volume: in the guts !

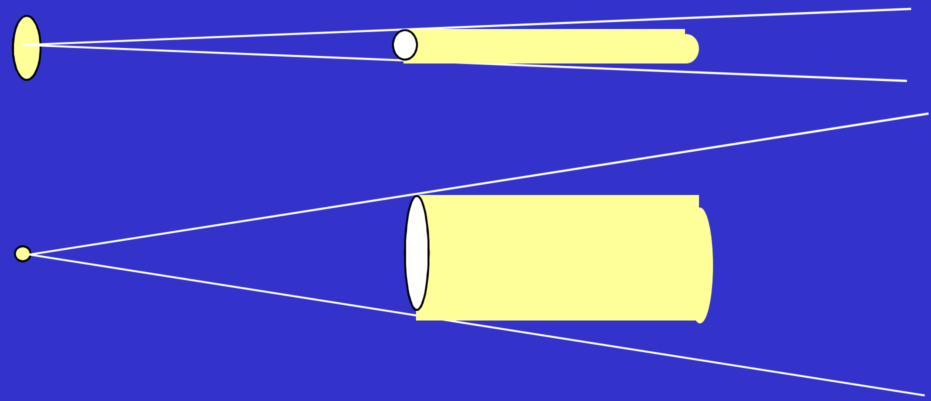
$\Delta\lambda$ decreased

t_c coherence time increased



apparent dimension decreased

coherence area increased



soufflons un peu, et prenons l'air

avant la suite qui est pire que tout !

Long Baseline Interferometry

functional features

interferometry : measuring by means of interferences

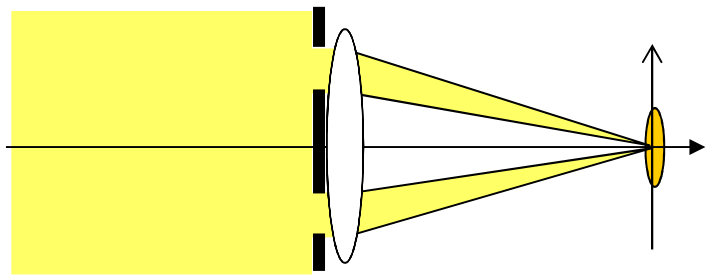
among several compact descriptions
the two main features here to handle are :

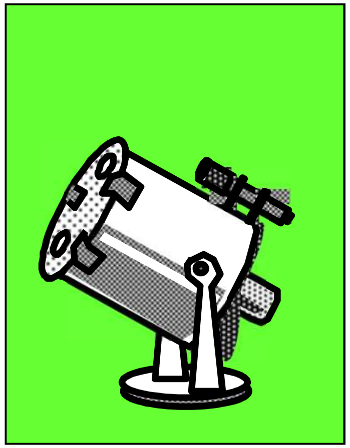
- ❑ interferometer = machine performing coherent addition
- ❑ interferometer = filter for spatial frequencies

but the main feature for science concerns is :
(simply a consequence of coherent addition)

interferometer = instrument making observable the degree of coherence

here we look at
the academic case
(Fizeau configuration)

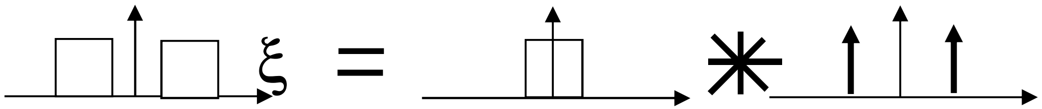


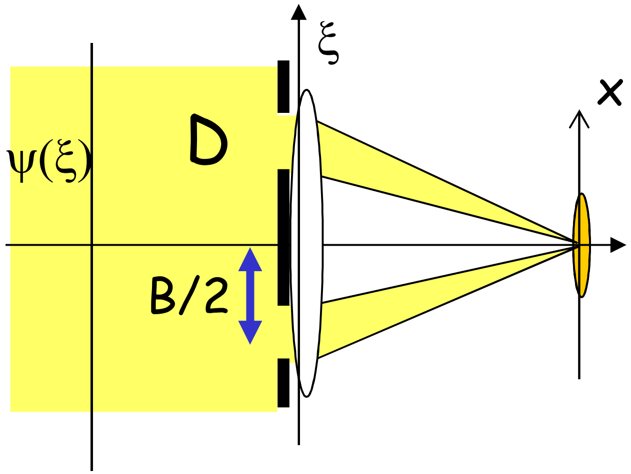


coherent addition of fields : observed intensity

incoming field $\psi(\xi)$

collecting pupil : $\Pi(\xi - B/2) + \Pi(\xi + B/2)$





collected field:

$Q(\xi) = \psi_1 \,.\, \Pi(\xi - B/2) + \psi_2 \,.\, \Pi(\xi + B/2)$

amplitude at focal plane (FT) :

$$\hat{Q}(u) = \hat{\Pi}(u)\left[\psi_1 \,.\, e^{-i.2\pi.\frac{B}{2}.u} + \psi_2 \,.\, e^{+i.2\pi.\frac{B}{2}.u}\right]$$

intensity at focal plane (squared modulus)

with $I_0 = \langle |\psi_1|^2 \rangle = \langle |\psi_2|^2 \rangle$

and $Airy = |\hat{\Pi}|^2$

$$I(x) = 2.Airy(x).I_0.\left\{1 + Re\left[\frac{\langle \psi_1.\psi^*_2 \rangle}{I_0}.\exp\left(i.2\pi.\frac{B.x}{\lambda.F}\right)\right]\right\}$$

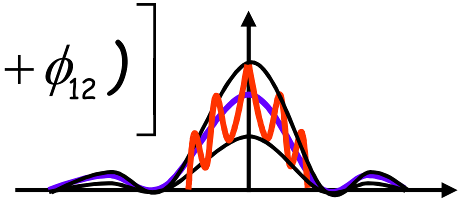
interferometer makes observable the degree of coherence

$$I(x) = 2.Airy.I_0.\left\{1 + Re\left[\frac{<\psi_1.\psi^*_2>}{I_0}.exp(i.2\pi.\frac{B.x}{\lambda.F})\right]\right\}$$

γ_{12} *might be complex*   here is $\gamma_{12}(B/\lambda)$

we write : $\gamma_{12} = \frac{|<\psi_1.\psi^*_2>|}{I_0}.exp(i\phi_{12}) = V.exp(i\phi_{12})$

then : $$I(x) = 2.Airy.I_0.\left[1 + V.cos(2\pi.\frac{B.x}{\lambda.F} + \phi_{12})\right]$$



warning :

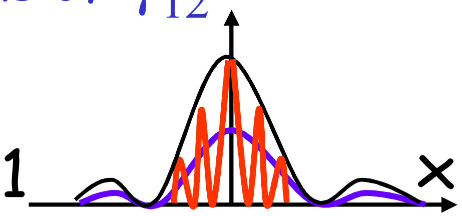
generally with two apertures, only V can be extracted from data

ϕ_{12} is mixed with random spurious phases and is not available

So the observation here, only yields V = modulus of γ_{12}

also called : amplitude of fringes

note : with point-like source V = 1



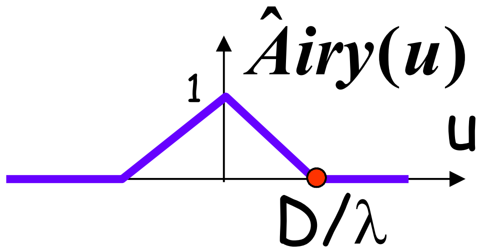
interferometer as filter for spatial frequencies

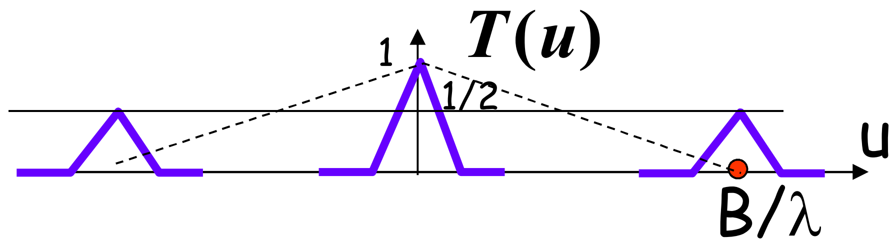
point-like source yields impulse response

$$I(x) = 2.I_0.Airy.\left[1 + cos\left(2.\pi.\frac{B}{\lambda}.\frac{x}{F}\right)\right] = 2.I_0.h(x)$$

transfer function (FT of h(x) normalized to unity)

$$T(u) = \hat{A}iry(u) * \left\{\delta(u) + \frac{1}{2}.\left[\delta(u - \frac{B}{\lambda}) + \delta(u + \frac{B}{\lambda})\right]\right\}$$





any source yields : $I(x) = O(x) * h(x)$ and $\hat{I}(u) = \hat{O}(u).T(u)$

interferometer allows sensing the source spectrum
at frequencies as high as B/λ

note : two telescopes, one baseline, one spatial frequency

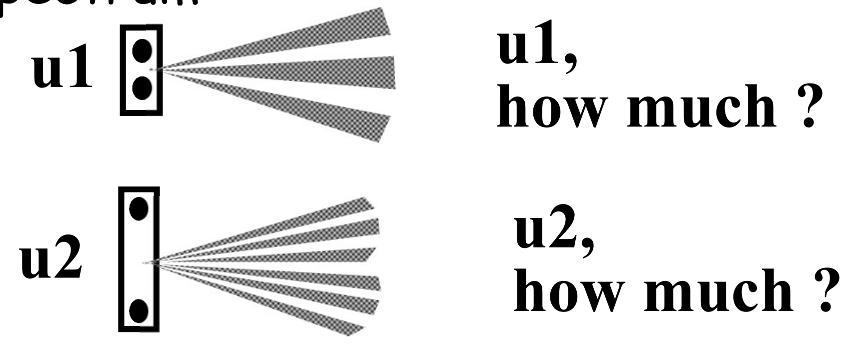
back to our "fringed lobe" : what is it doing ?

the lobe questions the source about the presence of a particular spatial frequency (the one born by the fringes : B/λ)

it makes a measure of this "presence" ,

in extracting info from the spatial spectrum

conceptually the source answer is the so-named "visibility" of the source at frequency B/λ



The visibility is given par by the modulation rate of the observed fringes on the camera

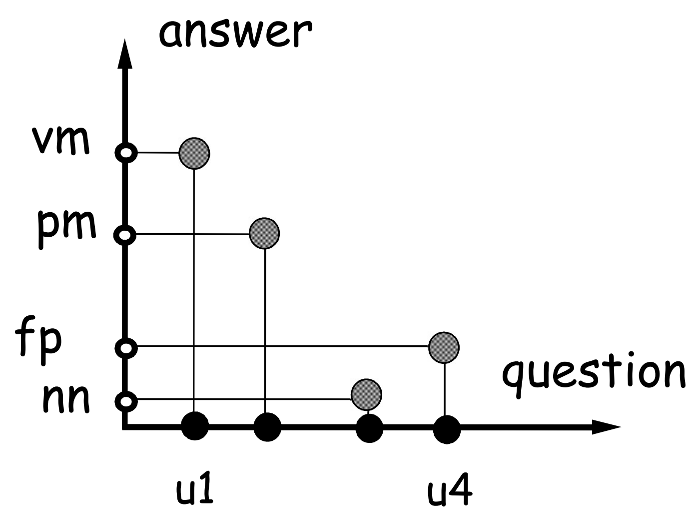
thanks to VanCittert & Zernike, we now know that this visibility is the degree of coherence of the source at frequency B/λ

note : one baseline, one frequency, each time a little piece of information (a component of the spatial spectrum)

listen to the source

base	freq	answer
B1	u1	very much
B2	u2	pretty much
B3	u3	nearly nothing
B4	u4	faint presence

answers are compiled
on a graph



so, baseline after baseline
is progressively built a so-called
visibility curve
which is assumed to reproduce
a cut in the spatial spectrum of the source

As announced , we have picked up information in the frequency space