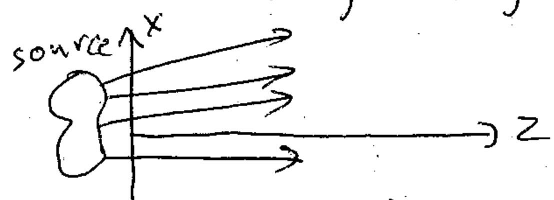
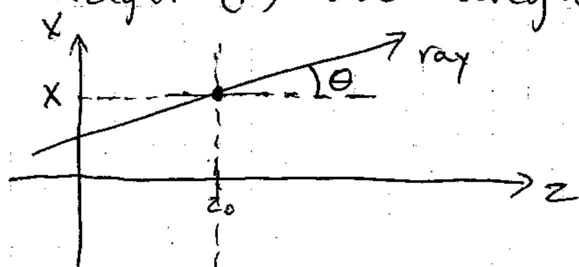
Phase space

Let us first work in 2D. Let z coincide with the optical axis, and x to the transverse axis. Assume all light is going towards larger z:



At a given z, each ray can be identified by its height (x) and angle (θ):



Let us define p (sometimes called the optical momentum) as

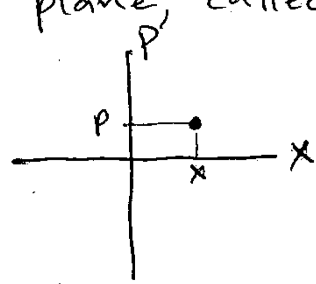
$$p = n \sin\theta$$

$\uparrow$ refractive index.

Then, the ray at z is fully characterized by x & p.

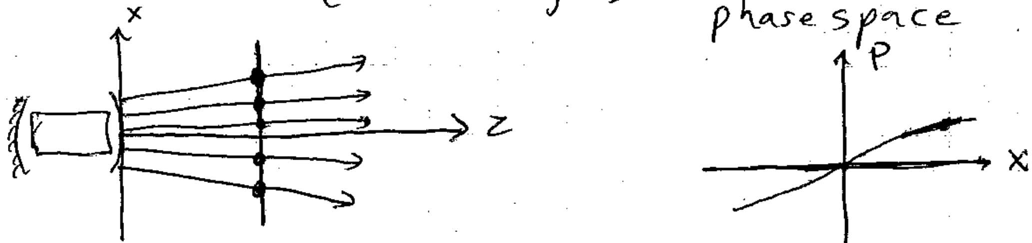
Let us define the vector $\underline{V} = \begin{pmatrix} x \\ p \end{pmatrix}$

The ray can then be represented by a point in the x vs p plane, called "phase space".

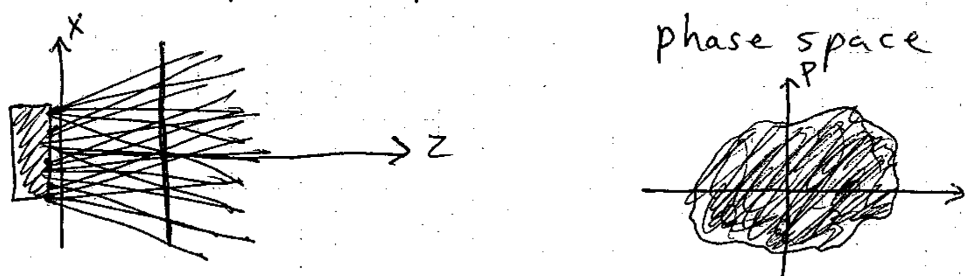


Of course, as the ray propagates in z, this point generally moves.

For a coherent source (in 2D), all rays form a 1-parameter family, so their points in phase space form a curve (can be straight!):

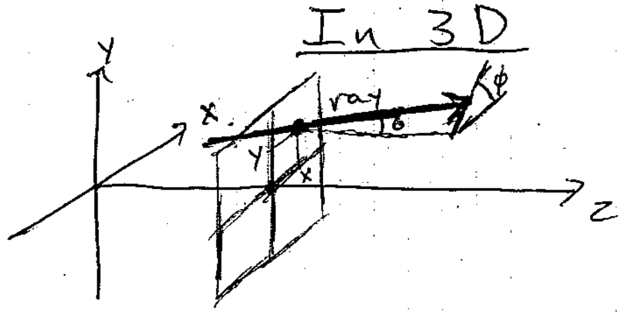


For an incoherent source (in 2D), the rays fill an area in phase space:



Propagation in free space or an optical system causes these curves and areas to change form.

<u>In 3D</u>



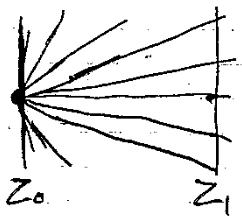
the transverse position is $\underline{x} = (x, y)$ instead of just x. The optical momentum is $\underline{p} = (p_x, p_y)$, where $p_x = n \sin\theta \cos\phi$, $p_y = n \sin\theta \sin\phi$.

Therefore $\underline{v} = \begin{pmatrix} x \\ y \\ p_x \\ p_y \end{pmatrix} = \begin{pmatrix} \underline{x} \\ \underline{p} \end{pmatrix}$, and phase space is four-dimensional! The rays of a coherent source form a surface (called the Lagrange manifold), while the rays of an incoherent source fill hypervolumes.

Exercise:

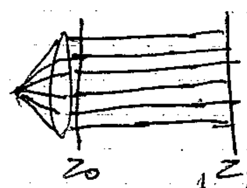
Represent in phasespace (2D) the rays generated by the following sources:

1) pointsource

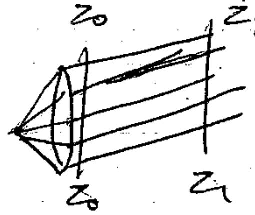


Draw it both for $z=z_0$, $z=z_1$

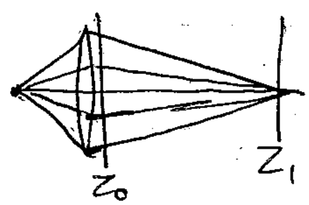
2) collimated beam



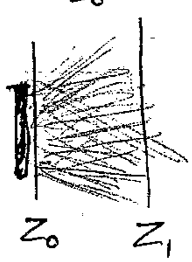
3) tilted collimated beam



4) focused beam

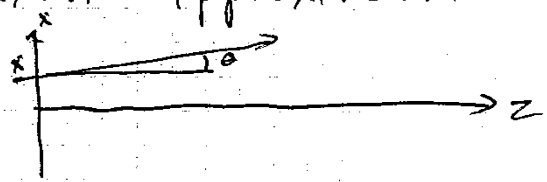


5) Extended incoherent source



Geometrical optics

- Paraxial approximation

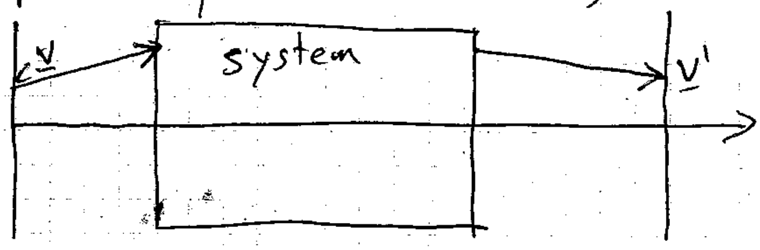


$\theta \ll \pi/2$, $\sin\theta \approx \theta \approx \tan\theta$

$\underline{p} \simeq (\theta\cos\phi, \theta\sin\phi)$

- Remember the ray vector $\underline{V} = \begin{pmatrix} \underline{x} \\ \underline{p} \end{pmatrix}$

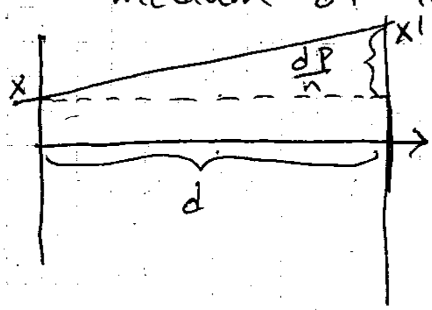
- Optical systems (1^{st} order):



$\underline{V}' = \mathbb{M}\,\underline{V}$

matrix describing the system

Examples: free propagation in a homogeneous medium of index n:



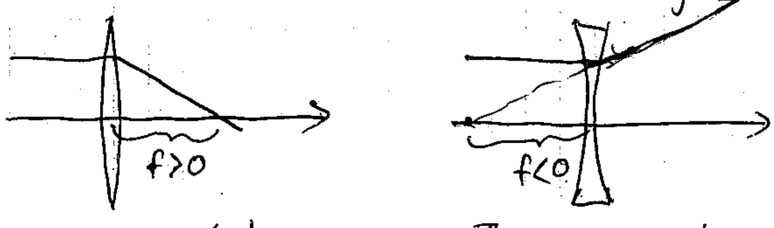
$\underline{x}' = \underline{x} + \frac{d}{n}\underline{p}$

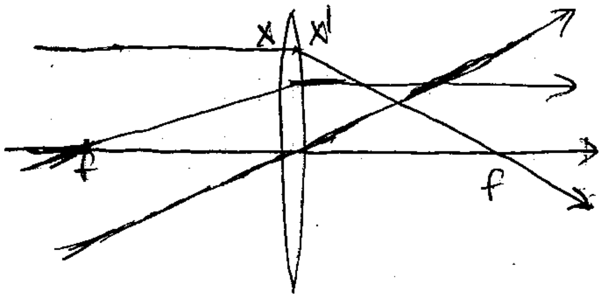
$\underline{p}' = \underline{p}$

s.o.

$\mathbb{M} = \mathbb{T}(d/n) = \begin{pmatrix} \mathbb{I} & d/n\,\mathbb{I} \\ \mathbb{O} & \mathbb{I} \end{pmatrix}$

Thin lenses of focal length f



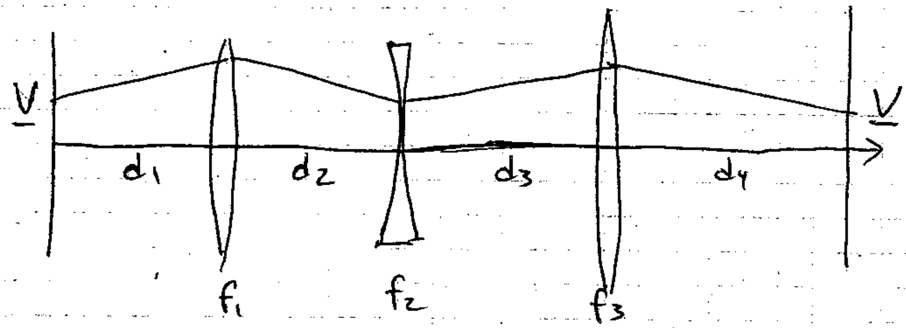


$\underline{x}' = \underline{x}$

$\underline{p}' = \underline{p} - \underline{x}/f$

$\mathbb{M} = \mathbb{L}(f) = \begin{pmatrix} \mathbb{I} & \mathbb{O} \\ -\frac{\mathbb{I}}{f} & \mathbb{I} \end{pmatrix}$

For a full system



$$\underline{V}' = \underbrace{\mathbb{T}(d_{n+1})\mathbb{L}(f_n)\ldots\mathbb{L}(f_2)\mathbb{T}(d_2)\mathbb{L}(f_1)\mathbb{T}(d_1)}_{\mathbb{M}}\underline{V}$$

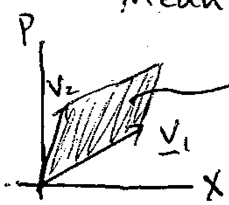
(ABCD systems because $\mathbb{M} = \begin{pmatrix} \mathbb{A} & \mathbb{B} \\ \mathbb{C} & \mathbb{D} \end{pmatrix}$)

Note: $\text{Det}\{\mathbb{T}(d)\} = \text{Det}\{\mathbb{L}(f)\} = \underline{1}$

therefore

$$\text{Det}\{\mathbb{M}\} = 1$$

Meaning? Think of 2D:



Think of the phase space area in a rhomboid defined by the rays $\underline{V}_1$, $\underline{V}_2$.

This area, which is an étendue in 2D, is given by

$$\mathcal{E} = |X_1 P_2 - X_2 P_1| = |\underline{V}_1^{+} \sigma \underline{V}_2|, \text{ where}$$

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

After propagation through the system,

$$\underline{V}_{\substack{1\\2}} \to \underline{V}'_{\substack{1\\2}} = \mathbb{M}\,\underline{V}_{\substack{1\\2}}, \text{ so}$$

$$\mathcal{E}' = |\underline{V}_1'^{+} \sigma \underline{V}_2'| = |\underline{V}_1^{+} \mathbb{M}^{+} \sigma \mathbb{M} \underline{V}_2|$$

but $\mathbb{M}^{+}\sigma\mathbb{M} = \begin{pmatrix} m_{11} & m_{21} \\ m_{12} & m_{22} \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} = \begin{pmatrix} m_{11} & m_{21} \\ m_{12} & m_{22} \end{pmatrix}\begin{pmatrix} m_{21} & m_{22} \\ -m_{11} & -m_{12} \end{pmatrix}$

$$= \begin{pmatrix} m_{11}m_{21} - m_{21}m_{11}, & m_{11}m_{22} - m_{12}m_{21} \\ m_{12}m_{21} - m_{11}m_{22}, & m_{12}m_{22} - m_{22}m_{12} \end{pmatrix} = \text{Det}\{\mathbb{M}\}\,\sigma$$

therefore:

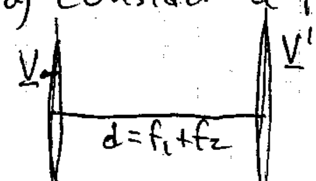
$$\mathcal{E}' = |\mathrm{Det}\{\mathbb{M}\}|\, |\underline{V}_1^{+} \sigma \underline{V}_2| = |\mathrm{Det}\{\mathbb{M}\}|\, \mathcal{E}$$

because $\mathrm{Det}\{\mathbb{M}\} = 1$, $\underline{\mathcal{E}' = \mathcal{E}}$

therefore $\mathrm{Det}\{\mathbb{M}\} = 1$ is analogous to Liouville's theorem.

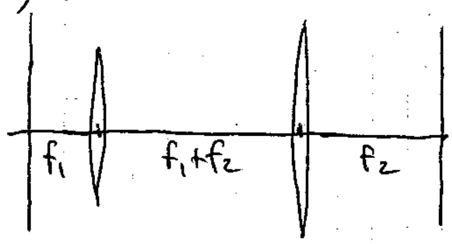
In 3D, if the system has rotational symmetry around the z axis, there are other invariants besides the étendue. More on this later...

Exercise: a) consider a telescope setup



find $\mathbb{M}$. What does the system do to both x & p?

b) Now surround this system with some free space as

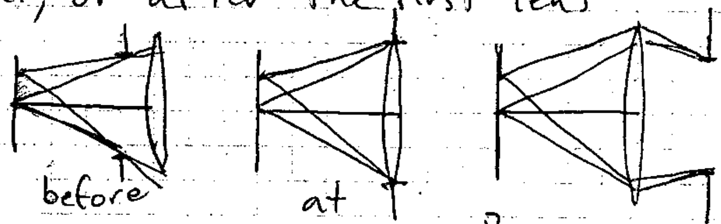


What is the new matrix? what does it do to the étendue occupied by a bundle of rays in phase space?

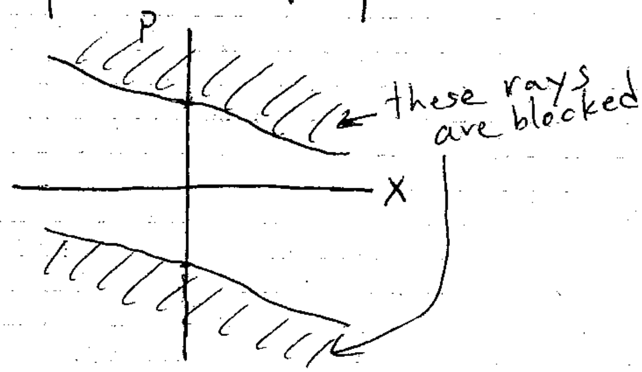
Stops and pupils

They limit the total étendue (or throughput) of the system.

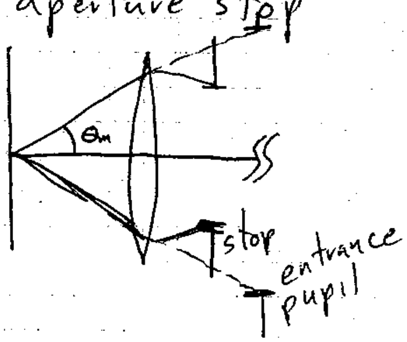
- Aperture stop: puts an angular limit to the rays entering the system. It can be before, at, or after the first lens



In phase space: (at the plane of the object)



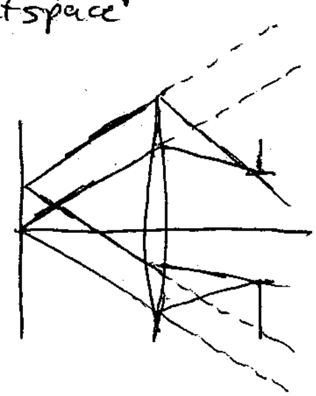
- Entrance pupil: image, viewed from the object, of the aperture stop



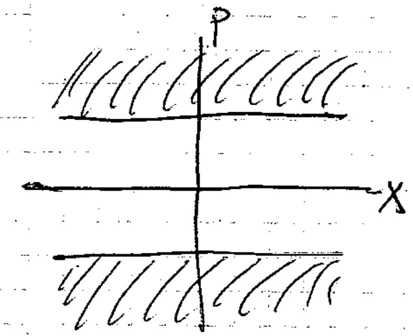
Note: if the stop is before or at the first lens, the stop itself is the entrance pupil.

$n \sin\theta_m = NA$
numerical aperture in object space

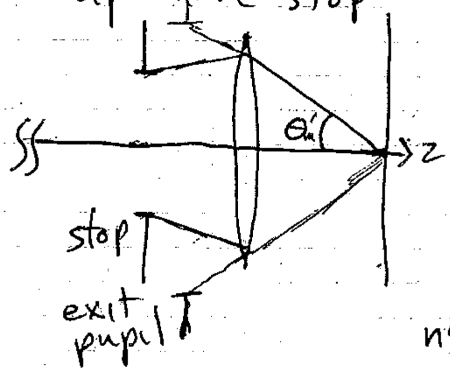
If the exit pupil is at infinity, e.g. if the stop is at the back focal plane of the 1^{st} lens, the system is called "entrance telecentric".



For entrance telecentric systems, the angular limit is independent of object position



• <u>Exit pupil</u>: image, viewed from the image space, of the aperture stop

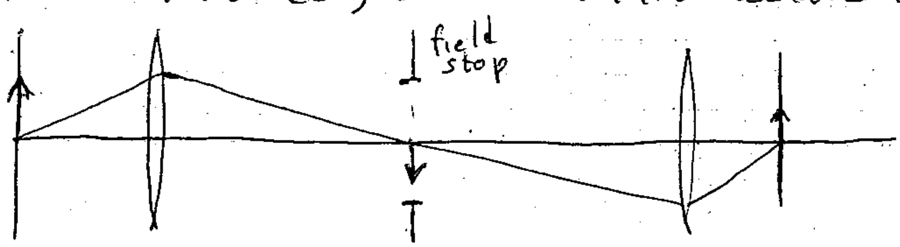


If the stop is at or after the last lens, the stop itself is the exit pupil

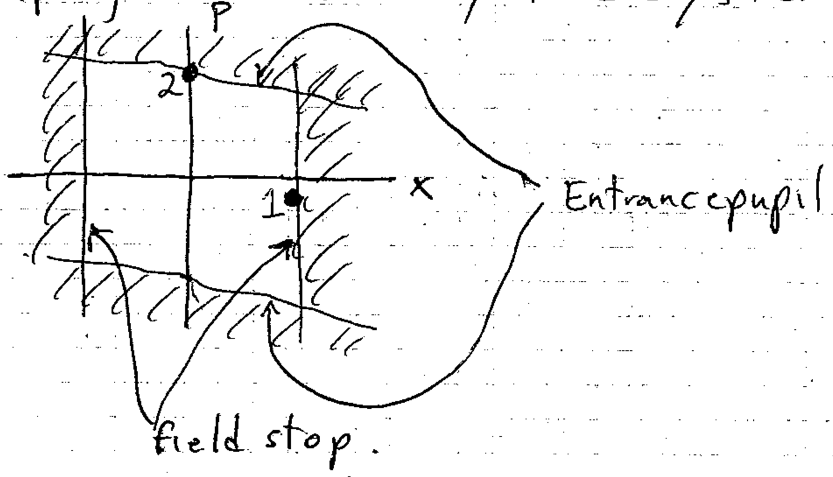
$n' \sin \theta_m' = NA'$, <u>numerical aperture in image space</u>

If the exit pupil is at infinity, then the system is called "exit telecentric". The image size is then insensitive to defocus.

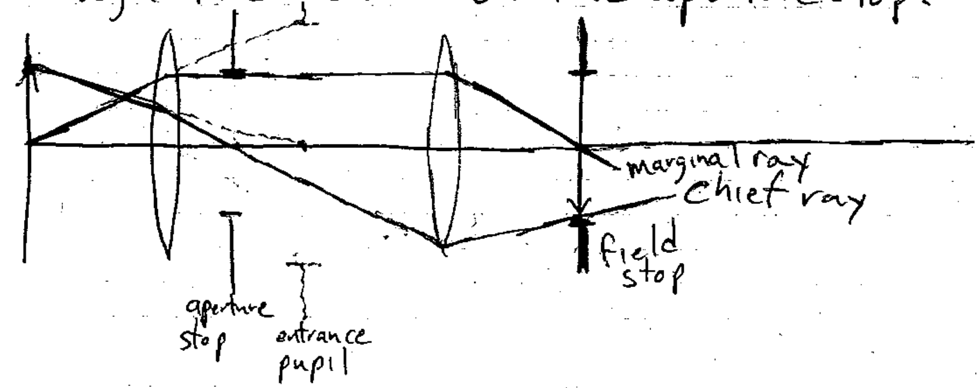
• <u>Field stop</u>: determines the size of the object that can be imaged. It can be at the object, the final image (e.g. at the film or ccd) or at an intermediate image:



The field and aperture stops (entrance pupil) determine the total étendue (also called throughput) admitted by the system.



1. Chief ray: A ray coming from the edge of the object (field stop) headed to the center of the entrance pupil, which then passes through the center of the aperture stop.



2. Marginal ray: A ray coming from the center of the object, headed for the edge of the exit pupil, and touching the edge of the aperture stop.

These are "extreme" rays in the system.

Lagrange invariant

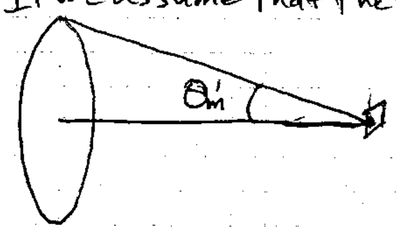
$$H = X_m P_c - X_c P_m = \underline{V}_c^{+} \sigma \underline{V}_m$$

constant along the system.
In 2D this is the total $\mathcal{E}$.
In 3D, the total $\mathcal{E}$ is given by

$$\mathcal{E} = \pi^2 H^2$$

Irradiance of an image

Suppose we are imaging a Lambertian object of radiance L_o.
If we assume that the transmissivity of the medium is T, then



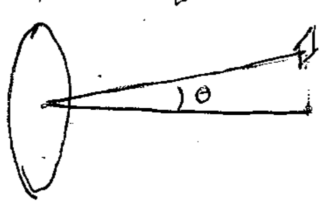
the radiance at the exit pupil is TL_o.
The irradiance at the image plane, on axis, is given by
(see Problem 1)

$$E_{image} = \pi L_o T \sin^2\theta_m' \quad \text{(assuming uniform T)}.$$

Since $NA' = n \sin\theta_m'$,

$$E_{image} = \pi T \frac{L_o}{n'^2} NA'^2$$

Off axis, we get a factor of $\cos^4\theta$.

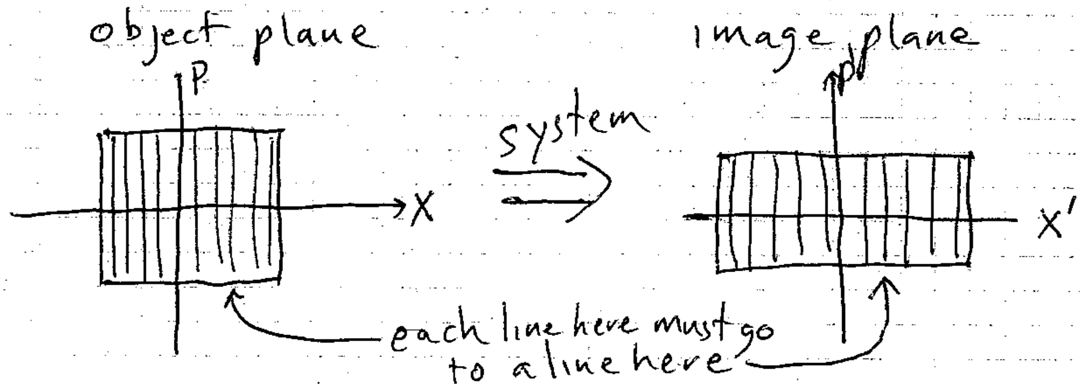


(see prob. 3)

Imaging vs. nonimaging (illumination) optics

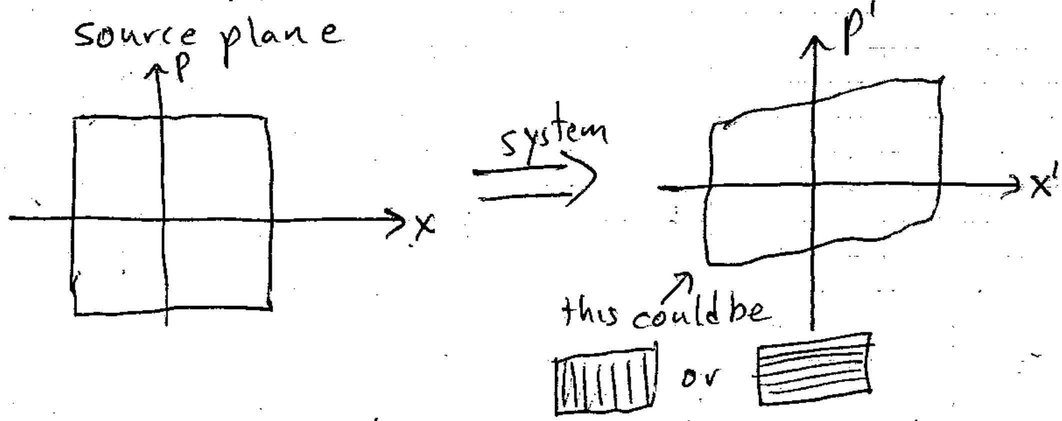
Imaging optics

The goal is to form good images of an object, i.e. to make all rays from the object go (at least approximately) to a corresponding point at the image plane. In phase space this means



Nonimaging optics

The goal is to take the light from a source and use it to illuminate an object in a given way. In phase space, this means



It is not important what ray goes to what point. The "imaging" solution is not forbidden, though.

Edge ray principle: it is enough to map the rays that define the edge of the region to solve the nonimaging problem.

3rd order theory (aberrations)

Let us assume that the system has axial symmetry. It is convenient to use, instead of $\underline{x}, \underline{\rho}$, the following normalized coordinates

$$\underline{h} = \frac{\underline{x}}{x_{max}} \quad \text{at the object plane}$$

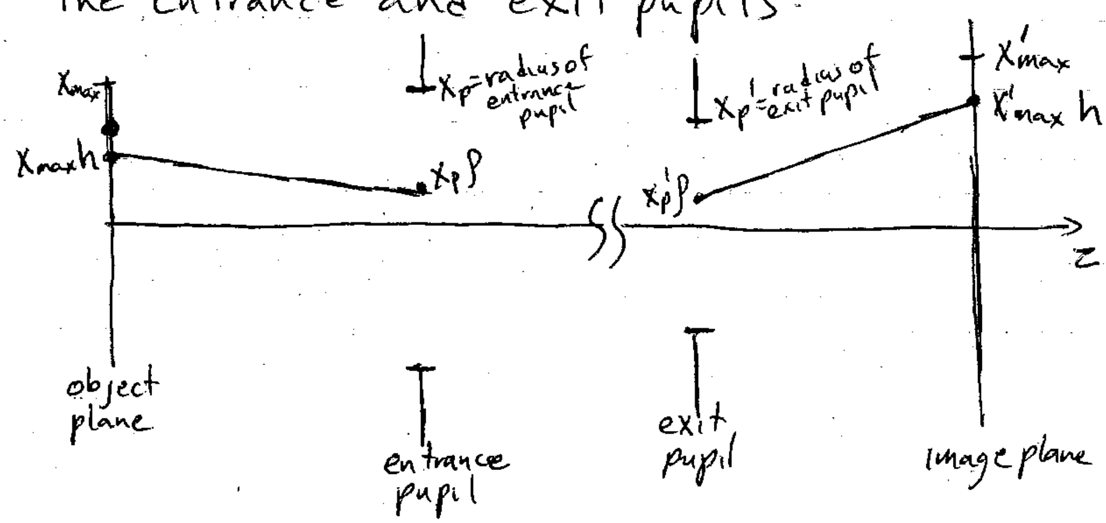
$$\underline{\rho} = \frac{\underline{x}}{x_{max}} \quad \text{at the aperture stop.}$$

Therefore, due to axial (rotational) symmetry, the three meaningful parameters are

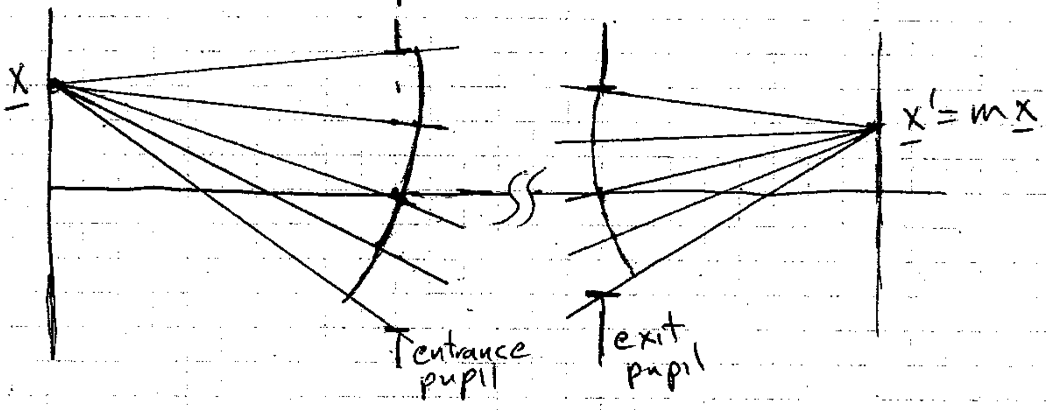
$$h = |\underline{h}|, \quad \rho = |\underline{\rho}|, \quad \phi = \arccos\left(\frac{\underline{h}\cdot\underline{\rho}}{h\rho}\right).$$

(Note that, for telecentric systems, $\underline{\rho} \propto \underline{\rho}$.)

We can think of an ideal system in terms of the entrance and exit pupils:



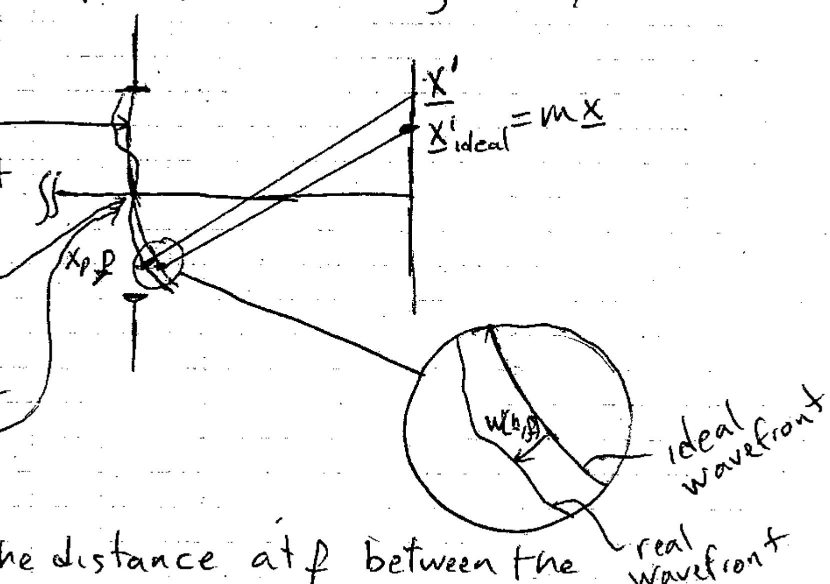
Ideally, all rays from an object point $\underline{x}$ go to the corresponding image point $\underline{x}' = m\underline{x}$ (where m is the magnification), so the phase fronts in both spaces are spherical:



However, in practice the phase front exiting the system is imperfect

We choose as the reference ideal wavefront the one that crosses the axis at the pupil plane.

We also choose the real one to do this



Let $W(\underline{h},\underline{\rho})$ be the distance at $\underline{\rho}$ between the real and the ideal wavefront (with $W>0$ if the real wavefront is to the right of the ideal one).

W is called the "wave aberration function".

The transverse error can be found:

$$\underline{\varepsilon}(\underline{h},\underline{\rho}) = \underline{x}' - \underline{x}'_{ideal} = -\frac{R}{n' x'_{max}} \frac{\partial W}{\partial \underline{\rho}}$$

refractive index

To study the system, we expand $W(\underline{h},\underline{\rho})$ in a Taylor series in $\underline{h}$ and $\underline{\rho}$ around $\underline{0}$:

$$
\begin{aligned}
W(\underline{h},\underline{\rho}) = {} & W_{000} && \text{constant term} \\
& + W_{020}\,\rho^2 + W_{111}\,\underline{\rho}\cdot\underline{h} + W_{200}\,h^2 && \text{quadratic terms} \\
& \left.\begin{aligned} & + W_{040}\,\rho^4 + W_{131}\,\rho^2\,\underline{\rho}\cdot\underline{h} + W_{220}\,\rho^2 h^2 \\ & + W_{222}\,(\underline{\rho}\cdot\underline{h})^2 + W_{311}\,\underline{\rho}\cdot\underline{h}\,h^2 + W_{400}\,h^4 \end{aligned}\right\} && 4^{\text{th}}\text{ order terms.} \ * \\
& + \ldots
\end{aligned}
$$

Notice that, when calculating $\underline{\varepsilon}$, because of the derivative, the order is reduced by one. For this reason the quadratic terms describe "linear" behavior, and the 4^{th} order ones (also called aberrations) lead to the "3^{rd} order theory".

Recall that we chose the reference and real wavefronts to cross at the axis ($\underline{\rho}=\underline{0}$), therefore $W(\underline{h},\underline{0})=0$, which implies $W_{000}=0$, $W_{200}=0$, $W_{400}=0$. Even if we did not set these constants to zero, they do not affect the image because these terms are independent of $\underline{\rho}$.

* Note that the notation for the coefficients W_{ijk} is such that:

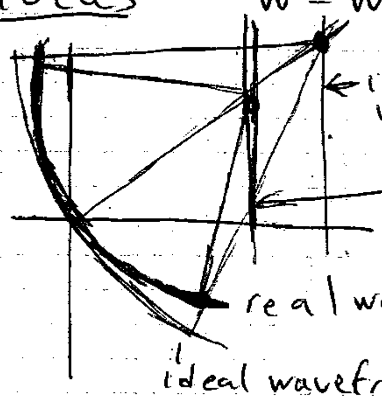
i = power of $|\underline{h}|$

j = power of $|\underline{\rho}|$

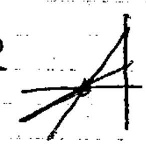
k = power of $\cos(\phi_h - \phi_\rho) = \dfrac{\underline{h}\cdot\underline{\rho}}{|\underline{h}||\underline{\rho}|}$.

The meaning of the aberrations is clarified if we first understand the two linear terms.

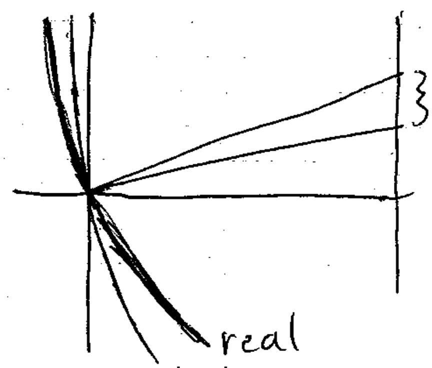
- <u>Defocus</u> $W = W_{020}\,\underline{\rho}^2$ ← this changes the curvature so the image plane is closer or further away.



$$\underline{\varepsilon} = \frac{-2R\,W_{020}\,\underline{\rho}}{n' X'_{max}}$$



- <u>Magnification error</u> $W = W_{111}\,\underline{\rho}\cdot\underline{h}$ ← this is linear, so it tilts the wavefront



$$\underline{\varepsilon} = -\frac{R}{n' X'_{max}}\, W_{111}\, \underline{h}$$

$$\underline{X}' = m\,\underline{X} + \underline{\varepsilon} = \underbrace{\left(m - \frac{R W_{111}}{n' X'_{max}}\right)}_{\text{correction in magnification}} \underline{X}$$

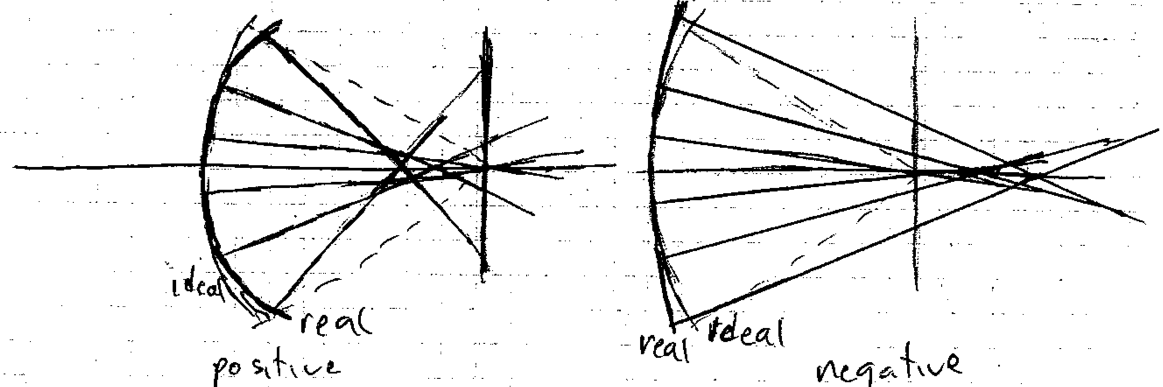
Aberrations

- **Spherical** $W = W_{040}\rho^4$

  Only one that is independent of $\underline{h}$, therefore only one that degrades the image on-axis; can think of it as ρ-dependent defocus:

$$W = \underbrace{(W_{040}\rho^2)}_{\text{like } W_{020}}\rho^2$$

so rays at the edge of the pupil are more defocused than those at the center:



- <u>Coma</u> $W = W_{131}\, \underline{h}\cdot\underline{\rho}\, \rho^2$

This is the only aberration linear in $\underline{h}$. If a system has no spherical aberration ($W_{040}=0$) or coma ($W_{131}=0$), then the image quality is good at and near the axis, since all other aberrations go as h^2 or h^3. Such a system is called "aplanatic".

Notice that coma looks both like a defocus and a magnification error:

$$W = \underbrace{\left(W_{131}(\underline{h}\cdot\underline{\rho})\right)}_{\text{like } W_{020}} \rho^2 = \underbrace{\left(W_{131}\,\rho^2\right)}_{\text{like } W_{111}} \underline{h}\cdot\underline{\rho}$$

Transverse error:

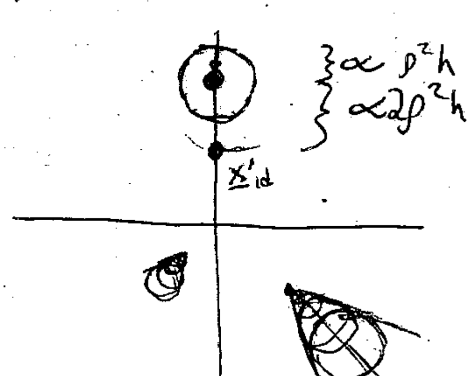
$$\underline{\varepsilon} = -\frac{R}{n'X'_{max}} \frac{\partial W}{\partial \underline{\rho}} = -\frac{2R W_{131}(\underline{h}\cdot\underline{\rho})}{n'X'_{max}}\, \underline{\rho} - \frac{R}{n'X'_{max}} W_{131}\, \rho^2\, \underline{h}$$

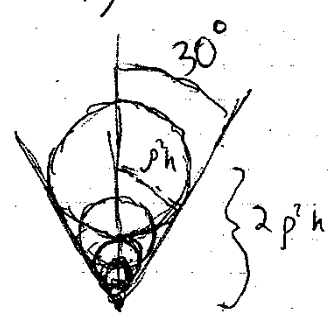
Consider an object point at the y axis: $\underline{h} = h(0,1)$

Let $\underline{\rho} = \rho(\cos\phi, \sin\phi)$. Then $\underline{h}\cdot\underline{\rho} = h\rho\sin\phi$ and

$$\underline{\varepsilon} = \frac{-R}{n'X'_{max}} W_{131}\, \rho^2 h \left[(2\sin\phi\cos\phi, 2\sin^2\phi) + (0,1)\right]$$

$$= \frac{-R}{n'X'_{max}} W_{131}\, \rho^2 h \left(\sin 2\phi,\ 2 - \cos 2\phi\right)$$

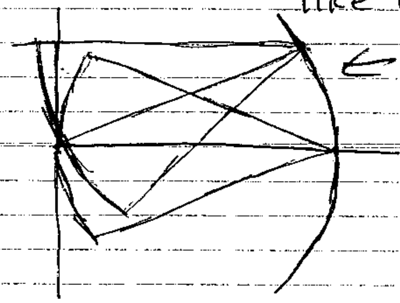




• Field curvature $W = W_{220}\, h^2 \rho^2$

Like a defocus proportional to h^2:

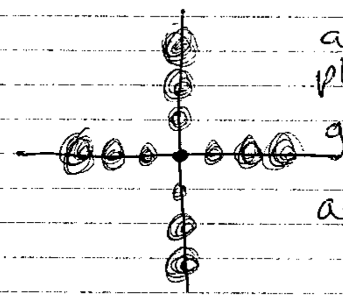
$$W = \underbrace{(W_{220}\, h^2)}_{\text{like } W_{020}} \rho^2$$



← Makes the ideal image plane be not a "plane" but a curved surface.

Can be >0 or <0.

$$\varepsilon = -\frac{2R}{n' x'_{max}} W_{220}\, h^2 \rho$$



at ideal image plane, images get out of focus away from axis.

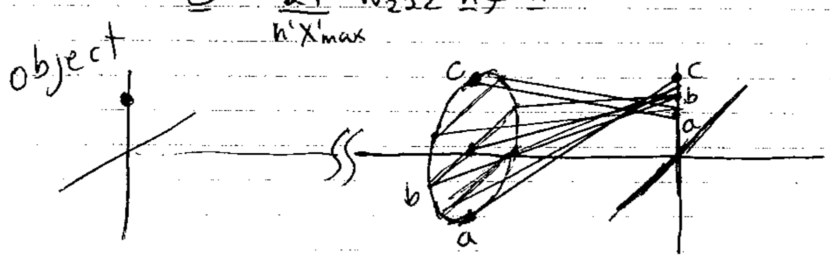
• Astigmatism $W = W_{222} (\rho \cdot h)^2$

Like a magnification error proportional to $\underline{h} \cdot \underline{\rho}$:

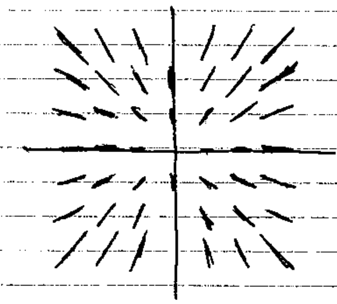
$$W = \underbrace{(W_{222}\, \underline{h} \cdot \underline{\rho})}_{\text{like } W_{111}}\, \underline{h} \cdot \underline{\rho}$$

Since this "change in magnification" is $\propto \underline{h} \cdot \underline{\rho}$, at a given point $\underline{h}$, the intersection of the rays moves from the ideal image an amount proportional to the component of ρ in the direction of $\underline{h}$

$$\varepsilon = -\frac{2R}{n' x'_{max}} W_{222}\, \underline{h}\underline{\rho}\, \underline{h}$$

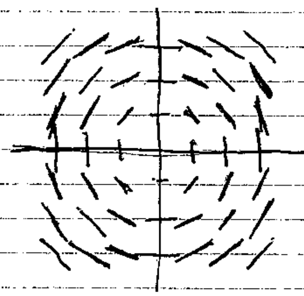


the image at the ideal plane then looks like

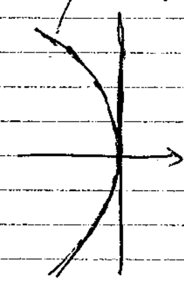


this plane is called the "sagittal image surface".

at a curved surface, the intersections are perpendicular

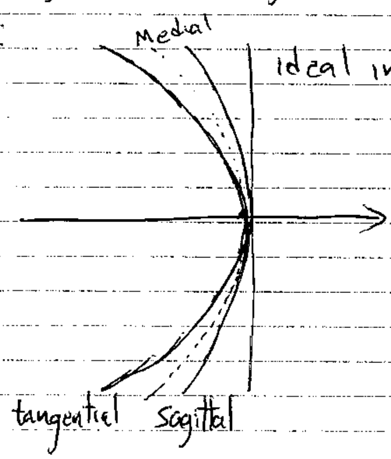


shape of this plane

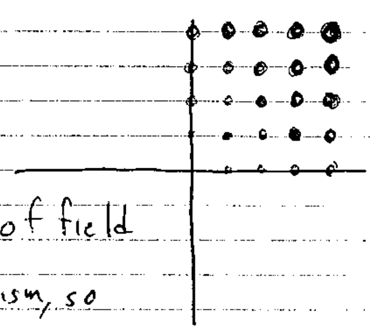


this is called the "tangential image surface".

If we combine field curvature and astigmatism the sagittal image surface gets curved, so we have:



Right between the sagittal and tangential surfaces, at the so-called "Medial image surface", the spread of each point is smaller and round:



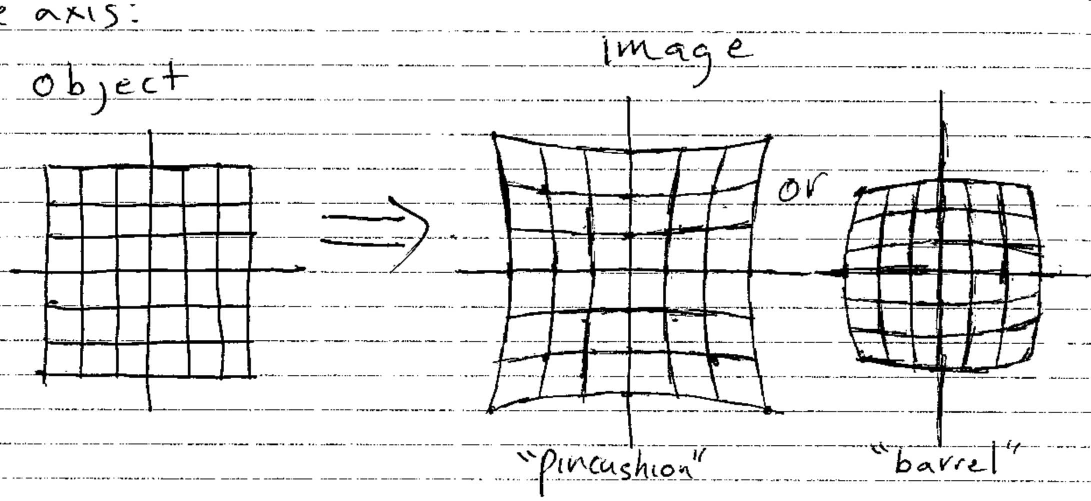
In some references, an amount of field curvature is grouped with astigmatism, so that, on its own, the medial image surface coincides with the flat image plane. The remainder of the field curvature is called "Petzval curvature".

- **Distorsion** $W = W_{311}\, h^2\, \underline{h}\cdot\underline{\rho}$

This looks like a magnification error proportional to h^2:

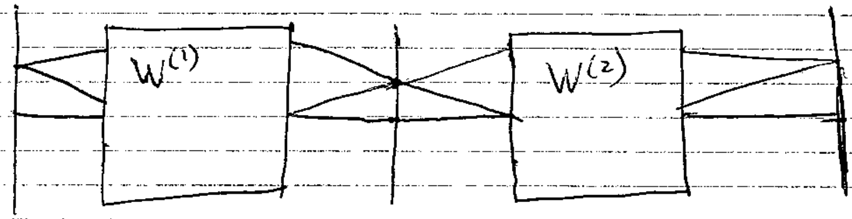
$$W = \underbrace{\left(W_{311}\, h^2\right)}_{\text{like } W_{111}} \underline{h}\cdot\underline{\rho}$$

That is, the magnification is different for object points away from the axis than for those near the axis:



Distorsion is the only aberration that does not affect the quality of the image at the image plane, but only its shape. Nowadays it can be corrected easily on the computer.

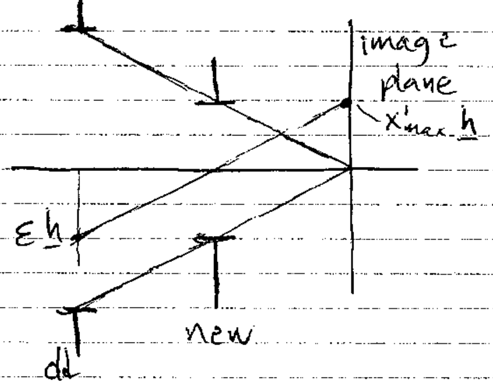
Concatenation of systems:



$W = W^{(1)} + W^{(2)}$, so each aberration is the sum of the aberrations.

Stop shifts

Shifting the stop changes some of the aberrations. Let us restrict ourselves to the case when the stop is also rescaled so that the output numerical aperture is preserved, i.e. so that the old and new exit pupil look like:

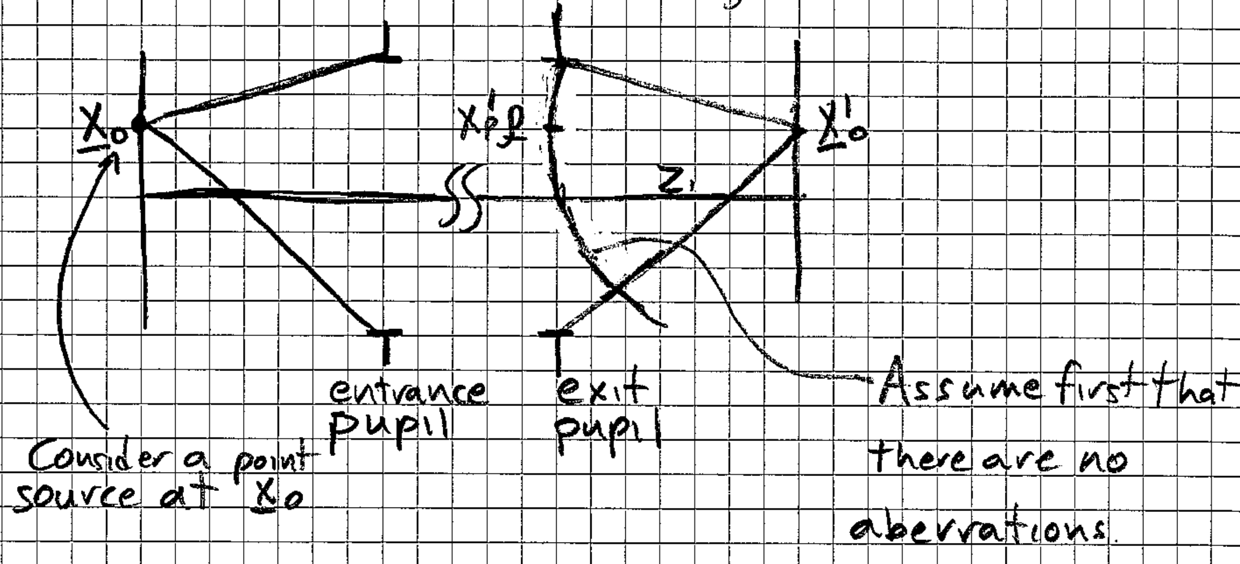


Notice that, then, the new pupil coordinate $\underline{\rho}'$ does not in general coincide with the old one, $\underline{\rho}$, unless $\underline{h} = 0$. more generally: $\underline{\rho} = \underline{\rho}' + \varepsilon\,\underline{h}$ (see figure above). Find the new aberrations W^{*}_{ijk} in terms of the old ones W_{ijk} by substituting $\underline{\rho} = \underline{\rho}' + \varepsilon\,\underline{h}$ in

$$W = W_{040}\rho^4 + W_{131}\rho^2 \underline{\rho}\cdot\underline{h} + W_{220}\rho^2 h^2 + W_{222}(\underline{\rho}\cdot\underline{h})^2 + W_{311}\,\underline{\rho}\cdot\underline{h}\,h^2.$$

$\underline{\text{Image formation}}$

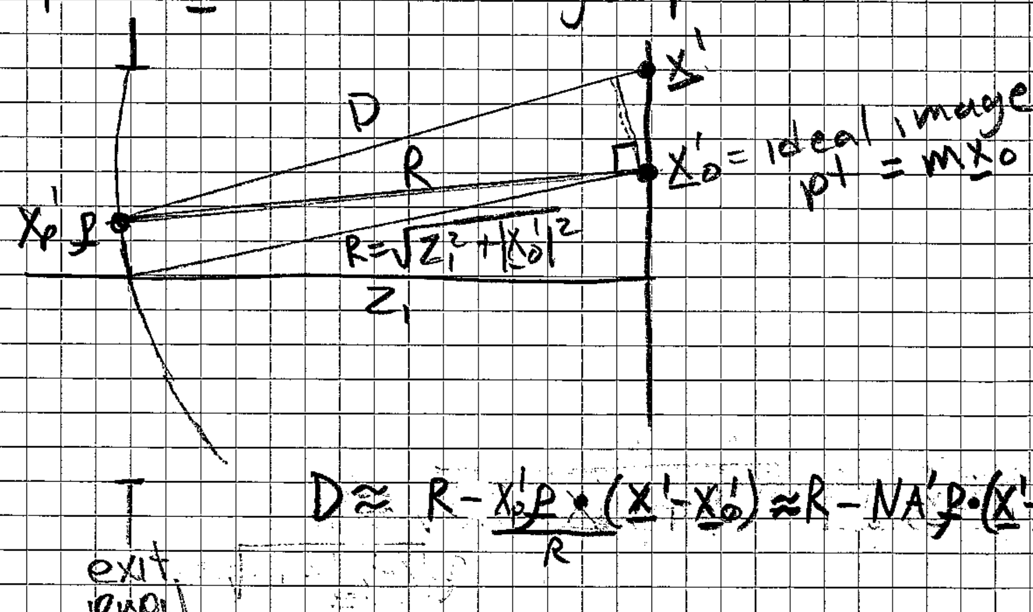
Assume incoherent objects.



The field in the image plane is roughly

$$U_{image}(\underline{x}') \propto \iint_{pupil} U_{pupil}(\underline{\rho})\, e^{ikD(\underline{\rho},\underline{x}')}\, d^2\rho$$

where $D(\underline{\rho},\underline{x}')$ is the distance from a point $(\underline{\rho})$ in the ideal reference wavefront and the point $\underline{x}'$ at the image plane.



$$D \approx R - \frac{x_p'\underline{\rho}\cdot(\underline{x}' - \underline{x}_o')}{R} \approx R - NA'\underline{\rho}\cdot(\underline{x}' - \underline{x}_o')$$

Therefore:

$$U_{image}(\underline{x}') \tilde{\propto} e^{ikR} \iint_{pupil} U_{pupil}(\underline{\rho})\, e^{-ikNA'\underline{\rho}\cdot(\underline{x}'-\underline{x}_0')} d^2\rho$$

$$\propto e^{ikR}\, \hat{\mathcal{F}} U_{pupil} \Big|_{kNA'(\underline{x}'-\underline{x}_0')}$$

$$= e^{ikR}\, \tilde{U}_{pupil}\left(kNA'(\underline{x}'-\underline{x}_0')\right)$$

Irradiance E_ν (what we usually call intensity)

$$E_\nu \tilde{\propto} \left| \tilde{U}_{pupil}\left(kNA'(\underline{x}'-\underline{x}_0')\right)\right|^2$$

For a perfect imaging system:

$$U_{pupil}(\underline{\rho}) = U_0 P(\underline{\rho}), \text{ where } P(\underline{\rho}) = \begin{cases} 1, & |\underline{\rho}| \le 1 \\ 0, & |\underline{\rho}| > 1. \end{cases}$$

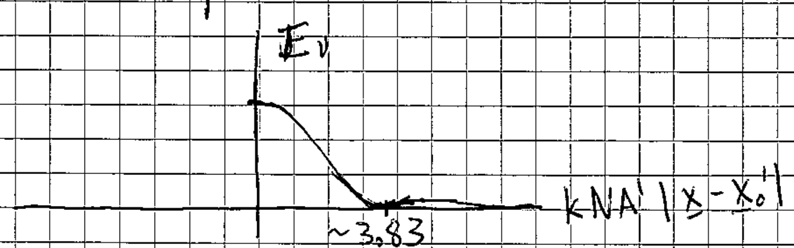
$$\tilde{U}_{pupil}(\underline{q}) = \frac{1}{2\pi} \iint_{|\underline{\rho}|\le 1} U_0\, e^{-i\underline{q}\cdot\underline{\rho}}\, d^2\rho$$

Bessel function

$$= \frac{U_0}{2\pi}\int_0^{2\pi}\int_0^1 e^{-iq\rho\cos(\phi_\rho-\phi_q)}\rho\, d\rho\, d\phi_\rho = U_0 \int_0^1 J_0(q\rho)\rho\, d\rho$$

$$= \frac{U_0}{q^2}\int_0^q J_0(u)\, u\, du = \frac{U_0}{q^2}\, q J_1(q) = U_0 \frac{J_1(q)}{q}$$

So $$E_\nu \tilde{\propto} \left| \frac{J_1(kNA'|\underline{x}-\underline{x}_0'|)}{kNA'|\underline{x}-\underline{x}_0'|} \right|^2$$ Airy Pattern



If the system has aberrations:

$$U_{pupil}(\underline{f}) = U_0 \, P(\underline{f}) \, e^{-ikW(\underline{f}, \underline{h}_0)}$$

$$E_V(\underline{x}') \tilde{\propto} \left| \iint_{|\underline{f}|<1} e^{-ikW(\underline{f}, \underline{h}_0) - ik\underline{f}\cdot(\underline{x}'-\underline{x}_0')NA'} \, d^2 f \right|^2$$

In general, no closed form solution, so must use numerical integration or fast Fourier transforms.

If instead of a point source object, we have an extended spatially incoherent source, then the irradiance images of all are superposed to give

$$E_{\nu,\text{image}}(\underline{X}') \propto \iint M_{\nu,\text{object}}(\underline{X}) \left|\hat{U}_{\text{pupil}}\left(kNA'(\underline{X}'-m\underline{X})\right)\right|^2 d^2X$$

$$\propto \iint M_{\nu,\text{object}}\left(\frac{\underline{X}_0'}{m}\right) \left|\tilde{U}_{\text{pupil}}\left(kNA'(\underline{X}'-\underline{X}_0')\right)\right|^2 d^2X_0'$$

$$= M_{\nu,\text{object}}\left(\frac{\underline{X}'}{m}\right) * \left|\tilde{U}_{\text{pupil}}\left(kNA'\,\underline{X}'\right)\right|^2$$

Now, Fourier-transform both sides and use the convolution theorem:

$$\tilde{E}_{\nu,\text{image}}(\underline{K}') \propto \tilde{M}_{\nu,\text{object}}(m\underline{K}')\ OTF(m\underline{K}')$$

or

$$\tilde{E}_{\nu,\text{image}}(\underline{K}/m) \propto \tilde{M}_{\nu,\text{object}}(\underline{K})\ OTF(\underline{K})$$

where OTF = optical transfer function.

$$OTF(\underline{K}) = \frac{1}{\underset{\text{normalization}}{N}} \iint \left|\tilde{U}_{\text{pupil}}\left(k\,\underbrace{NA'm}_{NA}\,\underline{X}\right)\right|^2 e^{-i\underline{K}\cdot\underline{X}}\, d^2X$$

$$= \frac{\iint U^*_{\text{pupil}}\left(\frac{\underline{K}_0-\underline{K}/2}{kNA}\right) U_{\text{pupil}}\left(\frac{\underline{K}_0+\underline{K}/2}{kNA}\right) d^2K_0}{\iint \left|U_{\text{pupil}}\left(\frac{\underline{K}_0}{kNA}\right)\right|^2 d^2K_0}$$

Let $U_{pupil}(\underline{\rho}) = P(\underline{\rho})\, e^{ikW(\underline{\rho})}$ ← aberrations.

$$P(\underline{\rho}) = \begin{cases} 1, & |\underline{\rho}| \le 1 \\ 0, & |\underline{\rho}| > 1. \end{cases}$$

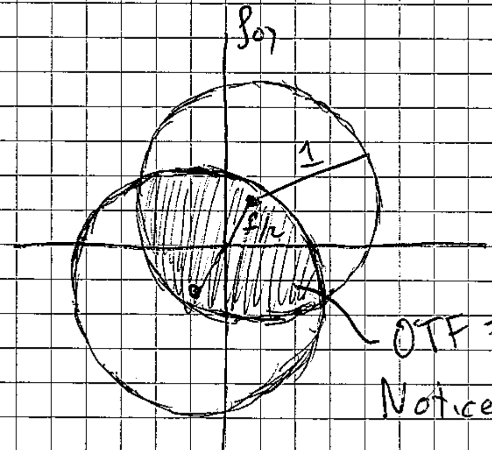
Then

$$OTF(\underline{K}) = \frac{\iint P\left(\underline{\rho}_0 - \frac{\underline{\rho}}{2}\right) P\left(\underline{\rho}_0 + \frac{\underline{\rho}}{2}\right) e^{ik\left[W\left(\underline{\rho}_0 + \frac{\underline{\rho}}{2}\right) - W\left(\underline{\rho}_0 - \frac{\underline{\rho}}{2}\right)\right]} d^2\rho_0}{\iint |P(\underline{\rho}_0)|^2 d^2\rho_0}$$

$$= \underbrace{\frac{1}{\pi}}_{\text{area of } P.} \iint P\left(\underline{\rho}_0 - \frac{\underline{\rho}}{2}\right) P\left(\underline{\rho}_0 + \frac{\underline{\rho}}{2}\right) e^{ik\left[W\left(\underline{\rho}_0 + \frac{\underline{\rho}}{2}\right) - W\left(\underline{\rho}_0 - \frac{\underline{\rho}}{2}\right)\right]} d^2\rho_0 \Bigg|_{\underline{\rho} = \frac{\underline{K}}{kNA}}$$

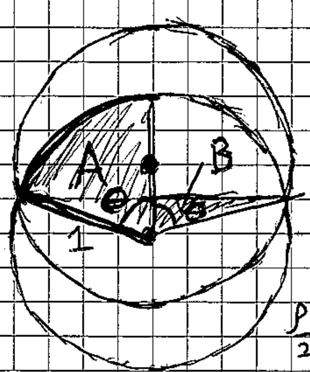
Example: if there are no aberrations ($W=0$)

$$OTF = \frac{1}{\pi} \iint P\left(\underline{\rho}_0 - \frac{\underline{\rho}}{2}\right) P\left(\underline{\rho}_0 + \frac{\underline{\rho}}{2}\right) d^2\rho_0 \Bigg|_{\underline{\rho} = \frac{\underline{K}}{kNA}}$$



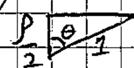
OTF = intersection area

Notice that it only depends on $|\underline{\rho}|$



$A = \pi\left(\frac{\theta}{2\pi}\right) = \frac{\theta}{2}$

π: area of unit circle; $\left(\frac{\theta}{2\pi}\right)$: fraction of unit circle

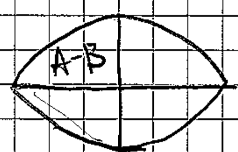


$$B = \frac{\sin\theta\cos\theta}{2}$$

But $\cos\theta = \frac{\rho}{2}$, $\sin\theta = \sqrt{1-\frac{\rho^2}{4}}$

so $A = \frac{1}{2}\arccos\left(\frac{\rho}{2}\right)$

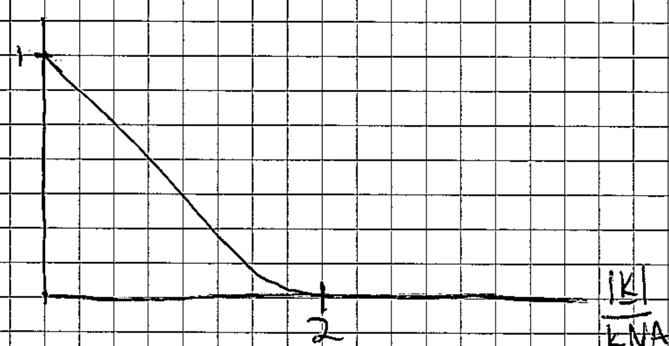
$$B = \frac{\rho}{4}\sqrt{1-\rho^2/4}$$



$$\text{area} = \frac{4(A-B)}{\pi} = \frac{2\arccos\left(\frac{\rho}{2}\right) - \rho\sqrt{1-\frac{\rho^2}{4}}}{\pi}$$

Test: area$(0) = 1$ ✓, area$(2) = 0$ ✓

$$OTF(\underline{k}) = \frac{2\arccos\left(\frac{|\underline{k}|}{2kNA}\right) - \frac{|\underline{k}|}{kNA}\sqrt{1-\frac{|\underline{k}|^2}{4k^2NA^2}}}{\pi}$$



If there are aberrations, the OTF cannot be calculated in closed form. Notice that, strictly speaking, we can only define the OTF if the aberrations are independent of $\underline{h}$ (for example defocus or spherical), since it is assumed in the use of the convolution theorem that the image of every point has the same shape. However, we can still compute the OTF for other aberrations to get an idea of what the system does to the image at different regions.

If there are aberrations, the OTF can become complex. Then we define

<u>Modulation transfer function</u>

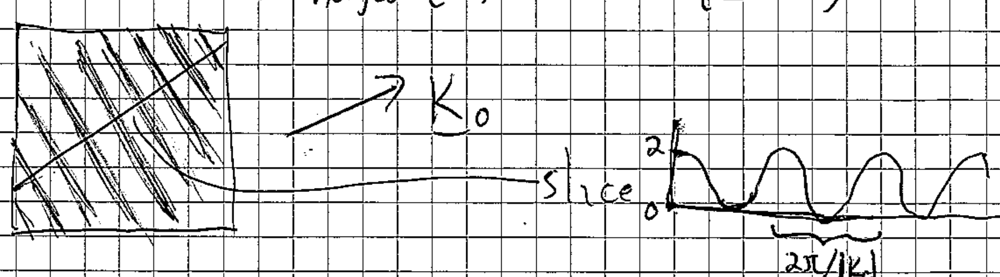
$$MTF(\underline{K}) = |OTF(\underline{K})|$$

<u>Phase transfer function</u>

$$PTF(\underline{K}) = \mathrm{Arg}\{OTF(\underline{K})\}.$$

- <u>Interpretation of MTF, PTF</u>

Supose the object is a sinusoidal intensity distribution of the form: $M_{\text{object}}(\underline{X}) = 1 + \cos(\underline{K}_0 \cdot \underline{X})$



The Fourier transform of this object is

$$\tilde{M}_{\nu,\text{object}}(\underline{K}) \propto \delta(\underline{K}) + \frac{\delta(\underline{K}-\underline{K}_0)}{2} + \frac{\delta(\underline{K}+\underline{K}_0)}{2}$$

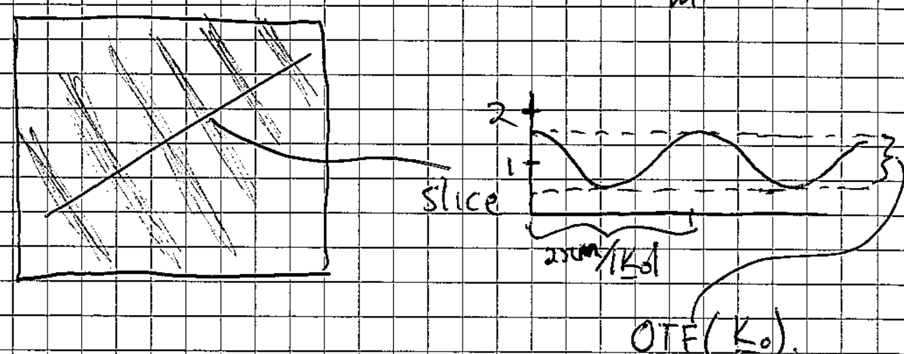
Therefore

$$\tilde{E}_{\nu,\text{image}}\left(\underline{K}/m\right) \propto OTF(\underline{K})\left[\delta(\underline{K}) + \frac{\delta(\underline{K}-\underline{K}_0)}{2} + \frac{\delta(\underline{K}+\underline{K}_0)}{2}\right]$$

$$= \cancelto{1}{OTF(\underline{0})}\,\delta(\underline{K}) + OTF(\underline{K}_0)\left[\frac{\delta(\underline{K}-\underline{K}_0)}{2} + \frac{\delta(\underline{K}+\underline{K}_0)}{2}\right]$$

assume OTF real, & therefore even

So $E_{\nu,\text{image}}(\underline{X}') \propto 1 + OTF(\underline{K}_0)\cos\left(\frac{\underline{K}_0\cdot\underline{X}'}{m}\right)$



So the visibility of the fringes is proportional to the MTF = |OTF| at the corresponding spatial frequency. If OTF < 0, then the fringes are reversed (bright ⇄ dark).

When there are aberrations, the MTF drops drastically. For example, for $W_{040} = \lambda$: