CHAPTER V

PARAXIAL or GAUSSIAN APPROXIMATION

We have seen in the previous chapter that the conditions for perfect stigmatism are fulfilled only for a limited number of optical systems, and for very specific pairs of points. A first version of approximate stigmatism consisted in allowing only small displacements in the vicinity of stigmatic points, while keeping large apertures: it is the approximate stigmatism of the Abbe or Herschel type.

For all the situations when these conditions of stigmatism are not fulfilled, the image will present aberrations. However, when we reduce the aperture of these optical systems and we observe small enough objects, we find out that images recover a good quality.

We will start this chapter using the example of the spherical refractive surface: the approximation of small angles will lead to a linear equation between the parameters defining the incident ray and those defining the emerging ray, and to a formula for the relative positions of the object and its image independent of the ray that we considered.

We will generalize this property to any homogeneous centered optical system, when we restrict to rays with a small angle with respect to the optical axis and to objects with a small size. This is the framework of the paraxial approximation of geometrical optics also called gaussian approximation. The rays that we will consider are called paraxial rays.

We will then see how the coefficients of these linear equations are connected to the characteristics of the optical system and to the positions and sizes of the object and the image. This will allow us to define a few characteristic points, cardinal points of a centered optical system, from which we can calculate or construct geometrically the image of any object.

I. Example of linear approximation in the case of a spherical refractive surface

1) Rigorous calculation

Using Snell-Descartes' law in its vectorial form, we can calculate the emerging ray starting from the incident ray without any approximation. We must:

- calculate the intersection I of the incident ray with the refractive surface;

- calculate the vector **N** normal to the surface at I (unit vector in the direction of **CI**);

- calculate the angle i between **u** and **N**;

- calculate the angle i' after refraction using Snell's law $n \sin i = n' \sin i'$;

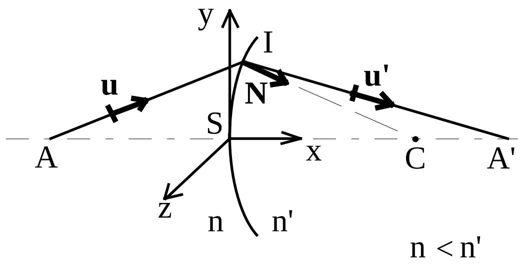
-calculate **u'** using the vectorial law $n\mathbf{u} - n'\mathbf{u}' = (n\cos i - n'\cos i')\mathbf{N}$.

The emerging ray is finally defined by I and **u'**.

To determine the image of a point with this method, we must repeat this calculation procedure for a large number of rays originating at the object point, and look for the intersection of the emerging rays with a plane orthogonal to the optical axis, in order to find the position of the image. This is a long and tedious calculation, and it would be useful to find at least approximately the final position of the image. In order to do that, we will consider first the rays that travel close to the optical axis.

2) Linear approximation

We use Sxyz as a reference frame, where S is a point on the spherical surface and Sx is the axis passing through C (optical axis of the refractive surface).



Since we consider only rays that travel close to the axis, we can write:

- the law of refraction in the linearized form $ni = n'i'$ and $n'\mathbf{u}' - n\mathbf{u} = (n'-n)\mathbf{N}$;

- the coordinates of vectors **u**, **u’** and **N** in the following form:

$$\mathbf{u}\begin{bmatrix}1\\ \beta\\ \gamma\end{bmatrix} \qquad \mathbf{u}'\begin{bmatrix}1\\ \beta'\\ \gamma'\end{bmatrix} \qquad \mathbf{N}\begin{bmatrix}1\\ -\dfrac{y_I}{R}\\ -\dfrac{z_I}{R}\end{bmatrix}$$

where y_I and z_I are the coordinates of point I and $R=\overline{SC}$ is the (algebraic) radius of curvature of the surface. The approximation implies in particular that we can consider I to be in the Syz plane, tangent to the spherical surface ($x_I=0$).

We get two equations of the same type that we can combine if we define new parameters that are complex numbers:

$$Z = y_I + jz_I \qquad \Omega = n\beta + jn\gamma \qquad \Omega' = n'\beta' + jn'\gamma'$$

$$\Omega' = \Omega + (n-n')\frac{Z}{R}$$

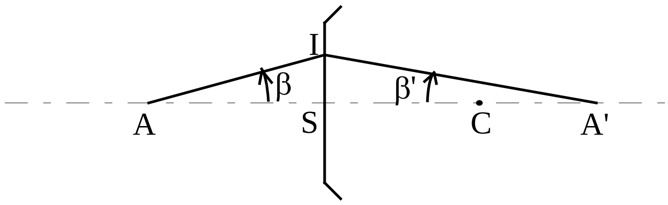
(Z, Ω) defines the incident ray and (Z, Ω’) the emerging ray. We finally obtain a linear relationship between those parameters that we can write in a matrix form:

$$\begin{bmatrix}Z\\ \Omega'\end{bmatrix} = \begin{bmatrix}1 & 0\\ -\dfrac{n'-n}{R} & 1\end{bmatrix}\begin{bmatrix}Z\\ \Omega\end{bmatrix}$$

Note that the fact that we include indices in the definition of parameters Ω leads to a matrix that always has a determinant of 1. For a plane refractive surface or a plane mirror ($R=\infty$), this matrix is the identity matrix.

3) Calculation of the image of an object point on axis

The object is the point A on axis, at which originates the rays in the direction **u**. We are looking for the position of A', intersection of the emerging ray with the optical axis. Since the surface is symmetrical around the optical axis, we can simplify by choosing a point I in the Sxy plane, so that $z_I=0$ and $\gamma=\gamma'=0$. In addition, the previous hypotheses lead us to choose I in the tangential plane Syz. Finally we can determine A' from a figure where the spherical surface is replaced by its tangential plane:



$$\beta = \frac{y_I}{\overline{AS}} \qquad \beta' = \frac{y_I}{\overline{A'S}}$$

The previous linear equation between the rays leads to:

$$\frac{n'}{\overline{A'S}} = \frac{n}{\overline{AS}} + \frac{(n-n')}{R}$$

The dependence with y_I has disappeared. The position of the image A' does not depend any more on the ray that was considered: we satisfy the stigmatism condition.

4) Validity of the approximation

Let us look for the conditions on the rays such that the gaussian approximation is valid.

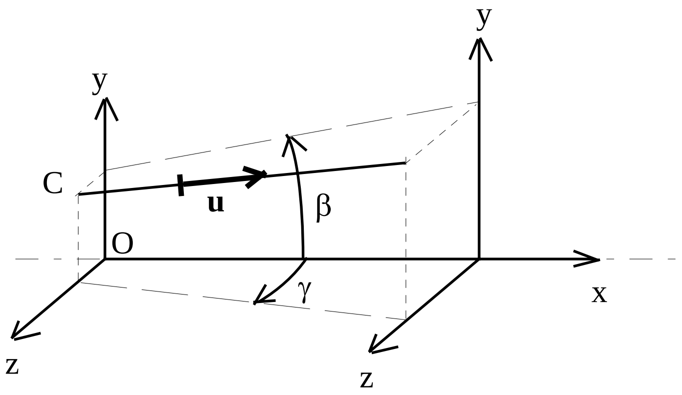
We have seen that the angles of the rays with respect to the normal to the surface or to the optical axis must be sufficiently small so that sini ≈ tani ≈ i and cosi ≈ 1: this is true to within 10% for angles smaller than approximately 25° and to within 1% for angles smaller than 8°. It is common to translate these numbers in terms of aperture number of the optical system $N=1/2n\sin\alpha$, which is approximately equal to 3.5 for an angle α of 8° and a medium with index n=1: we say that such a system is open at f/3.5. The Gaussian approximation will thus be very close to reality for optical systems open at f/N with N larger than 4.

II. Linear approximation and stigmatism

We will now show in the general case the properties that we just saw for the spherical refractive surface.

1) Relationship between incident and emerging ray

In all the following, we will consider only centered optical systems, with a symmetry of revolution around the optical axis O*x*, constituted of homogeneous media separated with refractive surfaces: the rays will thus appear as straight lines changing direction at the surfaces. The incident and emerging rays will be straight lines, that we can characterize using four parameters.



We will define a ray by its intersection C (y_1, z_1) with the O*yz* plane, and by its unit vector **u** (β, γ), which has coordinates in the O*xyz* reference frame that are:

$$C\begin{bmatrix}0\\ y_1\\ z_1\end{bmatrix} \qquad \mathbf{u}=\frac{1}{1+tg^2\beta+tg^2\gamma}\begin{bmatrix}1\\ tg\beta\\ tg\gamma\end{bmatrix}$$

We will also include the information relative to the index of refraction of the medium where this ray is traveling by choosing parameters *n*β and *n*γ, instead of β and γ.

In the most general case, each parameter of the emerging ray (y'_1, z'_1, $n'\beta'$, $n'\gamma'$) can be written as a Taylor series as a function of the parameters (y_1, z_1, $n\beta$, $n\gamma$) of the incident ray, such as:

$$y'_1=a_0+a_1y_1+a_2z_1+a_3n\beta+a_4n\gamma+\text{higher order terms}$$

Ox being the axis of symmetry of the system, the ray along the Ox axis with parameters (0,0,0,0) emerges from the optical system still along the axis ($a_0 = 0$).

The linear approximation will consist in neglecting terms of the second or higher orders. This amounts to saying that the rays that we consider are close to the optical axis. In particular the angles β and γ are small, so that the unit vector **u** along the ray has the following coordinates:

$$\mathbf{u}\begin{bmatrix}1\\ \beta\\ \gamma\end{bmatrix}$$

The relationship between the parameters of the emerging ray and those of the incident ray can then write in a matrix form:

$$\begin{bmatrix} y'_1 \\ z'_1 \\ n'\beta' \\ n'\gamma' \end{bmatrix} = \mathbf{M} \begin{bmatrix} y_1 \\ z_1 \\ n\beta \\ n\gamma \end{bmatrix}$$

M is a 4x4 matrix, with real coefficients a_{ij}, that depends on the optical system.

2) Simplification due to the radial symmetry of the system

This leads to a significant simplification of the previous 4x4 matrix. We can show that many terms in that matrix are zero or identical using for example the following properties:

a) an incident ray in the O*xy* plane leads to an emerging ray in the O*xy* plane (because O*xy* includes the axis of symmetry): if $z_1 = \gamma = 0$, we always have $z'_1 = \gamma' = 0$; which means that coefficients a_{21}, a_{23}, a_{41}, a_{43} of the matrix are zero.

b) same thing for an incident ray in the O*xz* plane: a_{12}, a_{14}, a_{32}, a_{34} are zero.

c) the system is invariant with respect to a rotation by 90° around O*x*.
We choose an incident ray R_1 (y_1, 0, $n\beta$, *0*), in the O*xy* plane, which emerges as a ray R'_1 (y'_1, 0, $n'\beta'$, 0), in the O*xy* plane. The ray R_2 (0, y_1, 0, $n\beta$), which results from a 90° rotation of R_1, will give an emerging ray R'_2 (0, y'_1, 0, $n'\beta'$), which results from the same 90° rotation of R'_2. This proves that the coefficients of the matrix are identical by pairs: $a_{11} = a_{22}$, $a_{13} = a_{24}$, $a_{31} = a_{42}$, $a_{33} = a_{44}$.

Finally the matrix M has the form:

$$\mathbf{M} = \begin{bmatrix} a_1 & 0 & a_2 & 0 \\ 0 & a_1 & 0 & a_2 \\ a_3 & 0 & a_4 & 0 \\ 0 & a_3 & 0 & a_4 \end{bmatrix}$$

We can simplify the expression of the linear relationship using complex parameters:

$$Z = y_1 + jz_1 \qquad \Omega = n\beta + jn\gamma$$

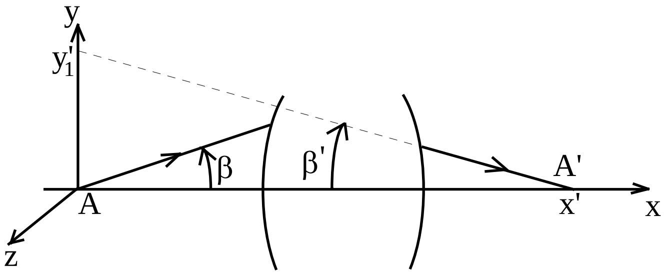
$$\begin{bmatrix} Z' \\ \Omega' \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} Z \\ \Omega \end{bmatrix}$$

The transformation between the rays is thus characterized by a 2x2 matrix with real coefficients.

Note: the radial symmetry of the optical system implies in particular that the relationship between incident and emerging rays is invariant by symmetry around the axis. If we change sign to all the parameters of the incident ray, all the parameters of the emerging ray must change sign also. This implies that in the expression of (y'_1, z'_1, $n'\beta'$, $n'\gamma'$) *as a* *function of the* (y_1, z_1, $n\beta$, $n\gamma$)*, there are no even order terms. The first non-zero order higher than the linear approximation is a third order term in the parameters of the incident ray.*

3) Stigmatism for an object on axis

We will now study what is implied by the linear approximation above when we consider image formation.



To determine the image of an object point A on the axis, we will consider an arbitrary ray originating at that point and look for the intersection of the emerging ray with the optical axis. Let us choose A*xyz* as a reference frame, where A*x* is the optical axis and A*xy* contains the incident ray. The coordinates of a ray originating at A are $Z = 0$ and $\Omega = n\beta$. The corresponding emerging ray has the coordinates $Z' = a_2 n\beta = y'_1$ and $\Omega' = a_4 n\beta = n'\beta'$. The position of the intersection A' with the axis is:

$$x' = \overline{AA'} = -\frac{y'_1}{\beta'} = -\frac{a_2 n'}{a_4}$$

It is thus independent of the ray that we chose in the O*xy* plane. Due to the radial symmetry, the position of the image is independent of the ray no matter how we chose it (however it does depend on A via the coefficients a_2 and a_4 of the matrix expressed in the reference frame attached to A).

4) Effect of a change of reference frame

Let us consider the expression (Z_1, Ω_1) of a ray in the reference frame O_1xyz and find out how it is modified if we change to the reference frame O_2xyz, where the origin has been translated by a quantity O_1O_2 along O_1x. We easily show that the new coordinates (Z_2, Ω_2) write:

$$\begin{bmatrix} Z_2 \\ \Omega_2 \end{bmatrix} = \begin{bmatrix} 1 & \dfrac{\overline{O_1O_2}}{n} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} Z_1 \\ \Omega_1 \end{bmatrix}$$

We will call <u>translation matrix</u> the matrix of the translation of the reference frame. We can thus use the most convenient reference frame to write the coordinates of a ray or the matrix of an optical system, then multiply by displacement matrices to change reference frame. Note that the determinant of this matrix is equal to 1, so that the determinant of the matrix will not change with the reference frame.

For example we will be easier to write the coordinates (Z,) of the incident ray in a frame that has its origin at the object point A and to write the emerging ray coordinates (Z', Ω') in a frame centered at A' image of A. We then get:

$$\begin{bmatrix} Z' \\ \Omega' \end{bmatrix}_{(A'xyz)} = \begin{bmatrix} 1 & \frac{x'}{n'} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} Z \\ \Omega \end{bmatrix}_{(Axyz)}$$

Since $x' = -n' a_2 / a_4$, the transformation matrix can be written in the form:

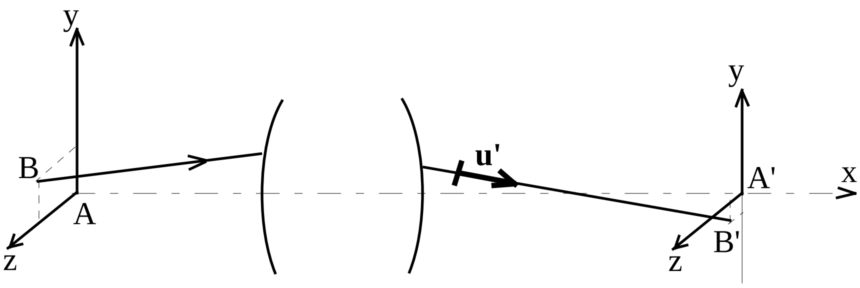
$$\begin{bmatrix} Z' \\ \Omega' \end{bmatrix}_{(A'xyz)} = \begin{bmatrix} a'_1 & 0 \\ a'_3 & a'_4 \end{bmatrix} \begin{bmatrix} Z \\ \Omega \end{bmatrix}_{(Axyz)}$$

The matrix of an optical system between two conjugate points is thus simplified.

5) Stigmatism for an object point close to the axis

Let us consider now an object point B, located in the Ayz plane and close to A. An arbitrary ray originating at B (0, y, z) will have parameters in the reference frame Axyz that are: $Z = y+jz$ and $\Omega = n\beta+jn\gamma$. The fact that B is close to the axis is necessary so that we can use the linear approximation with y and z being small.

Note: in this case the incident ray (B, **u**) *is NOT in a plane containing the axis; we cannot say anything special about the corresponding emerging ray*.



To determine the emerging ray, we use the matrix of the system between points A and A’: we immediately see that, due to $a'_2=0$, the parameter Z’ of the emerging ray does not depend on Ω, thus on the direction of the incident ray. But Z’ also represents the coordinates of the point of intersection of the emerging ray with the A’yz plane: we thus find that all the incident rays originating at B intersect at the same point B’ located in the A’yz plane. The stigmatic condition is thus fulfilled, and the image of a plane orthogonal to the axis is a plane orthogonal to the axis. In addition if B is along Ay (z=0), its image is along A'y', in other words A, A', B and B' are in the same plane.

As a conclusion, we have shown that for a centered optical system and rays that satisfy the linear approximation (small angle and close to the axis), the stigmatism is fulfilled for any object point on axis, and the image of a portion of a plane close to the axis is also a plane.

6) Other applications of this formalism

The matrix formalism that we developed here allows us to emphasize the linear nature of the gaussian approximation.

It is also well suited to a computed calculation, which becomes useful in the paraxial domain when the number of refractive surface becomes very large. We can even extend this type of matrix formalism to include higher order of calculation (larger matrices).

You will also find the same 2x2 matrices when you study the laser cavities, when they will allow you to determine the stability conditions and the parameters of the emitted beam in the fundamental mode (gaussian beams).

Finally, this kind of formalism is also used for electron beam optics, or for any kind of particle beam that requires a linear approximation.

It is however not advisable to use matrix formalism for simple systems where the graphical methods and the analytical formulas, that we will see further down in this course, is much faster and allow you to visualize the coherence of the results.

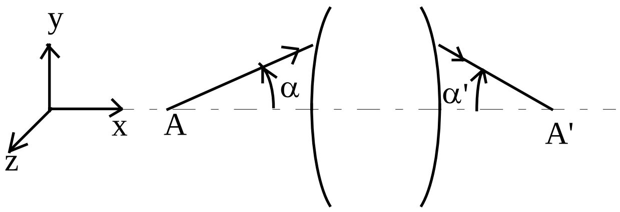
III. Characterization of optical systems in the linear approximation

In this paragraph, we will connect the matrices, in particular their coefficients, to the properties of the optical system that they represent, as well as the relationship between the position of the object and the position of the image (gaussian formula).

Following on the remark of §**II.4)**, we will choose Axyz as a reference frame for the incident ray and A'xyz for the emerging ray, so that the coefficient a'_2 of the matrix of the optical system is zero. To simplify the notations, we will change back to the letters without prime a_1, a_2, a_3, a_4 for the coefficients of the matrix, thus we will have $a_2 = 0$.

1) Magnifications. Power.

*Let us consider a ray originating at A and making an angle α with the axis. Let us choose Axy as the plane containing this ray.



The relationship between the rays is given by:

$$\begin{bmatrix} 0 \\ n'\alpha' \end{bmatrix} = \begin{bmatrix} a_1 & 0 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} 0 \\ n\alpha \end{bmatrix}$$

The coefficient a_4 of the matrix is thus connected to the <u>angular magnification</u> $g_\alpha = \alpha'/\alpha$ through:

$$a_4 = \frac{n'}{n} g_\alpha$$

*Let us consider a point B close to A with coordinate y. The transformation of an arbitrary ray originating at B is given by:

$$\begin{bmatrix} y' \\ \Omega' \end{bmatrix} = \begin{bmatrix} a_1 & 0 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} y \\ \Omega \end{bmatrix}$$

The coefficient a_1 is thus equal to the transverse magnification gy = y'/y.

*Let us consider another pair of arbitrary points A_1 and A'_1 (not necessarily conjugate points) and write the new matrix of the system. We can calculate it using the translation matrices defined in §**II.4**:

$$\begin{bmatrix} Z'_1 \\ \Omega'_1 \end{bmatrix}_{(A'_1 xyz)} = \begin{bmatrix} 1 & \frac{\overline{A'A'_1}}{n'} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 & 0 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} 1 & \frac{\overline{A_1 A}}{n} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} Z_1 \\ \Omega_1 \end{bmatrix}_{(A_1 xyz)} = \begin{bmatrix} a'_1 & a'_2 \\ a'_3 & a'_4 \end{bmatrix} \begin{bmatrix} Z_1 \\ \Omega_1 \end{bmatrix}_{(A_1 xyz)}$$

We note that the coefficient a'_3 of the new matrix is equal to a_3. This coefficient is thus independent on the choice of origins for the reference frames: it is a characteristic parameter of the optical system.

We will call the power P of an optical system the opposite of a_3. An optical system will be converging if P>0: such a system deviates the rays towards the axis. For example a spherical refractive surface with n<n' and $R=\overline{SC}>0$ (see §**I**) is converging: P = (n'-n)/R >0.

We shall call *focal systems* those with a non zero power P≠0 and *afocal systems* those with P=0.

2) Gaussian formulae and relationships between the magnifications

Let us go back to the matrix for the change of reference frame from (A, A') to (A_1, A'_1) written above, but considering now the case where A_1 and A'_1 are conjugate points. We will try to find the new coefficients a'_2, a'_3 and a'_4.

*Since A_1 and A'_1 are conjugate points, we know that the coefficient a'_2 must be zero. We thus get a relationship between the positions of A, A', A_1 and A'_1 that we can write in the form:

$$\frac{(g_y)_{AA'}}{\overline{A'A'_1}} - \frac{(g_\alpha)_{AA'}}{\overline{AA_1}} = \frac{P}{n'}$$

If we know the position of a specific pair of conjugate points and the magnifications for these points, we can then determine the position of all the other pairs of conjugate points. We shall call this relationship the gaussian formula for the optical system with its origin at points (A, A'), or Descartes' conjugation relationship. One example is the gaussian formula for the spherical refractive surface that we demonstrated in §**I.3** with A=A'=S.

*Note: This formula concerns only in principle the object points on axis. In fact, we have shown in §**II.3b** that the image of an object point B in the plane of A must be in the plane of A': the gaussian formula gives the projection on the axis of the image B' and we just have to multiply the distance $\overline{AB}$ by the transverse magnification (coefficient a_1 of the matrix) to find $\overline{A'B'}$.*

*We can transform the above gaussian formula into:

$$\frac{\overline{AA_1}}{\overline{A'A'_1}}=\frac{g_\alpha}{g_y}+\frac{P}{n'g_y}\overline{AA_1}$$

Considering a point A_1 very close to A, such that the last term of the equation is negligible, we get the lateral magnification, ratio between the small displacement of the image $dx'=\overline{A'A'_1}$ corresponding to the small displacement $dx=\overline{AA_1}$ of the object:

$$g_x=\frac{\overline{A'A'_1}}{\overline{AA_1}}=\frac{dx'}{dx}=\frac{g_y}{g_\alpha}$$

We get in that way a relationship between the three magnifications corresponding to the same pair of conjugate points (A,A'):

$$\boxed{g_y=g_x g_\alpha}$$

*The determinant of the new matrix is equal to the product of the determinants of the three matrices. Since the translation matrices have a determinant equal to 1, we thus have *a'd'=ad*. The quantity *ad* is thus the same for any pair of conjugate points. It can be written as:

$$\frac{n'}{n}g_y g_\alpha=\text{cst}$$

We can determine this constant noting that the matrix of the spherical refractive surface has a determinant of 1 (see §I). In the linear approximation, any system can be described as a succession of spherical refractive surfaces, so that its matrix will always have a determinant of 1. We thus get the following relationship:

$$\boxed{ny\alpha=n'y'\alpha'}$$

which is called Lagrange-Helmholtz relationship**.** We can notice that this is the equivalent of Abbe sine condition (see chap IV) for small angles.

Finally we see that two relationships connect the three magnifications, transverse, angular and lateral, so that knowing only one is enough.

*From the two relationships between the magnifications:

$$g_x=\frac{n'}{n}g_y{}^2$$

The lateral magnification is thus always positive, which means that object and image always move in the same direction. On the other hand, when the image is magnified by a factor of 2 along y, it is expanded by a factor of 4 along x: the proportions along x and y are not conserved.

*Let us calculate the coefficients a'_1 and a'_4 as a function of a_1 and a_4. This allows us to connect the magnifications for (A_1, A'_1) to the magnification for (A, A'):

$$(g_y)_{A_1A'_1}=(g_y)_{AA'}-\frac{P}{n'}\overline{A'A'_1}$$

$$(g_\alpha)_{A_1A'_1}=(g_\alpha)_{AA'}+\frac{P}{n'}\overline{AA_1}$$

What is interesting in these relationships is that they show that for a focal system ($P\neq 0$), we can reach an arbitrary transverse magnification or an arbitrary angular magnification (but the two are connected) if we choose the right pair of conjugate points.

On the contrary, for afocal systems (P=0), the magnifications are independent of the pair of points. The gaussian formula written earlier thus takes the simple form:

$$\frac{x'}{x}=\frac{\overline{\mathrm{A'A'_1}}}{\overline{\mathrm{AA_1}}}=cst=\frac{g_y}{g_\alpha}=\frac{n'}{n}g_y^{\ 2}$$

and that, for any arbitrary pair of conjugate points chosen as origins for the reference frames.

3) Focal systems

In the case of focal systems, we can define an ensemble of specific points called cardinal points, with which we can construct the image of any object.

a) cardinal points

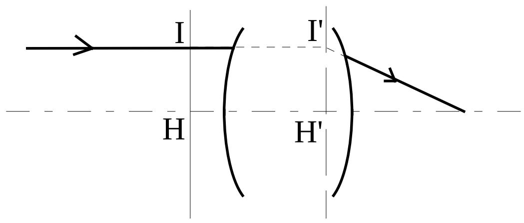
**principal points*

These points H and H' are conjugate points with a transverse magnification of 1. According to the Lagrange-Helmholtz relationship, the angular magnification is thus equal to n/n'.

We just saw that for a focal system, we can always find two such points. Taking these points as a reference, the matrix has a coefficient $a_1=1$ and still $a_2=0$ since the two reference points are conjugate. The Gaussian formula with an origin at the principal points is given by:

$$\boxed{\frac{n'}{x'}-\frac{n}{x}=C} \qquad \text{where } x'=\overline{\mathrm{H'A'}} \text{ et } x=\overline{\mathrm{HA}}$$

The planes perpendicular to the optical axis containing the principal points are called principal planes. They have a property that will be very useful for geometrical construction of rays. Consider an incident ray parallel to the axis of the system. It intersects the first principal plane in a point I, so the corresponding emerging ray must pass through the image I’ of I through the system.



This image of I is very easy to find since we know first that I’ belongs to the second principal plane, second that the transverse magnification is 1. I’ is thus at the intersection of the extended incident ray and of the second principal plane.

We can do this construction for any incident ray parallel to the axis: thus <u>the second principal plane is the intersection of the incident rays parallel to the axis and the corresponding emerging rays</u>.

As an exercise, you can do an analog reasoning to show that <u>the first principal plane is the intersection of the emerging rays parallel to the axis and the corresponding incident rays</u>.

**foci*

The second focus is the image of a point at infinity on axis. Its position, referenced with respect to the second principal point H’, is obtained when A is at infinity in the previous Gaussian formula:

$$\overline{\mathrm{H'F'}} = \frac{n'}{C} = f'$$

This quantity is called the focal length of the system. It is positive for a converging system.

The first focus has an image that is located at infinity on axis. Its position, referenced with respect to the first principal point, is given by:

$$\overline{\mathrm{HF}} = -\frac{n}{C} = f$$

f is the first focal length of the system. It is equal to the opposite of the second focal length when the incident and emerging media have the same index (n=n').

BE CAREFUL the first and second foci are not conjugate points.

**nodal points*

The nodal points N and N' are conjugate points with an angular magnification of 1. According to Lagrange-Helmholtz relationship, their transverse magnification is thus equal to n/n'. These points are different from the principal points only if the two incident and emerging media do not have the same index.

To determine the position of the first nodal point, we go back to the relationships between the angular magnifications when we change from (H,H') to (N,N') as reference conjugate points:

$$(g_\alpha)_{\mathrm{NN'}} = 1 = (g_\alpha)_{\mathrm{HH'}} + \frac{C}{n'}\overline{\mathrm{HN}} = \frac{n}{n'} + \frac{C}{n'}\overline{\mathrm{HN}} \text{ so that } \overline{\mathrm{HN}} = \frac{n'-n}{C} = f' + f$$

We do find that H and N are identical if *n*=*n'*, and we have:

$$\overline{\mathrm{FN}} = f'$$

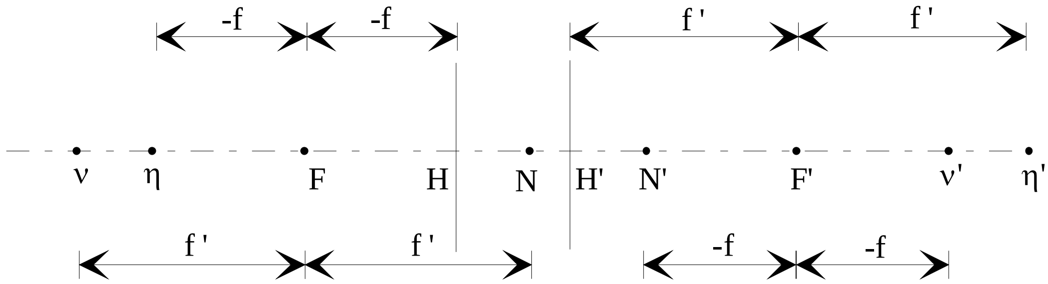
We can, in the same spirit, from the relationship between transverse magnifications, determine the position of the second nodal point, and we get:

$$\overline{\mathrm{N'F'}} = -f$$

** antiprincipal and antinodal points*

The antiprincipal points η and η' are conjugate points corresponding to a transverse magnification of -1 and the antinodal points ν and ν' those corresponding to an angular magnification of -1. We can show that these points are symmetric to the principal and nodal points with respect to the corresponding focal points.

The figure below shows an example of the positions of the cardinal points for a converging system with power P = 0.5 cm^{-1}, where the distance between principal planes is equal to 1.5cm, and for incident and emerging media with index n=1 and n'=1.5 respectively:



Note that if we set the relative positions of F, H, F' and H', all the other cardinal points can be deduced from them.

b) construction of the image of an object

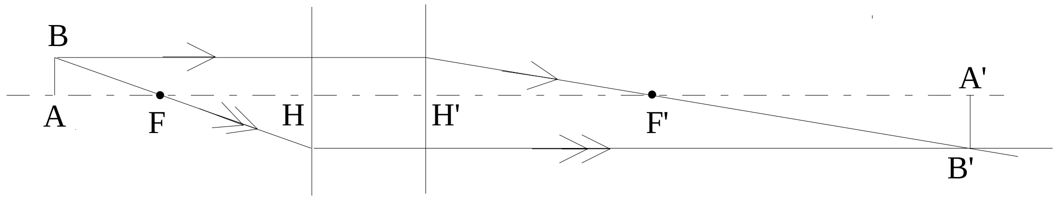
An optical system will be completely defined in the paraxial approximation by the positions of its principal planes and its foci.

Since we know that the system is stigmatic in the linear approximation, it is sufficient, to construct the image of an object B, to draw the path of two rays passing through this point B and to find the intersection B' of the two corresponding emerging rays. All the other rays will necessarily pass through B'.

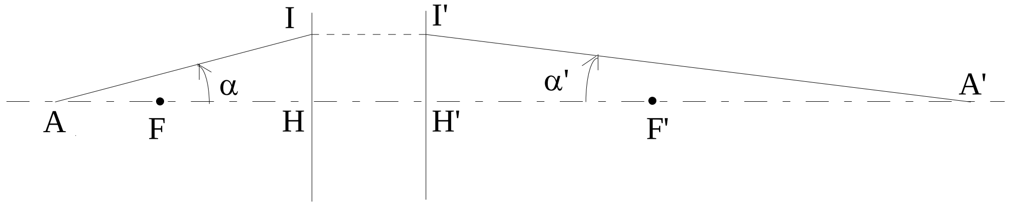
We thus choose two specific rays passing through B:

- an incident ray parallel to the optical axis: the corresponding emerging ray passes through F' and intersects the incident ray in the plane of H';

- an incident ray passing through F: the corresponding emerging ray comes out parallel to the optical axis and intersects the incident ray in the plane of H.



To construct the image of an object A on axis, the two suggested rays are identical: it is the optical axis of the system. To find the image A', we use the fact that the image of a small object AB perpendicular to the axis will be an image A'B' perpendicular to the axis also. We thus consider an arbitrary point object B in the plane of A, we construct its image B' and we obtain A' on the axis in the plane of B'.
We can also, using a ray passing through A, obtain a general expression for the angular magnification:

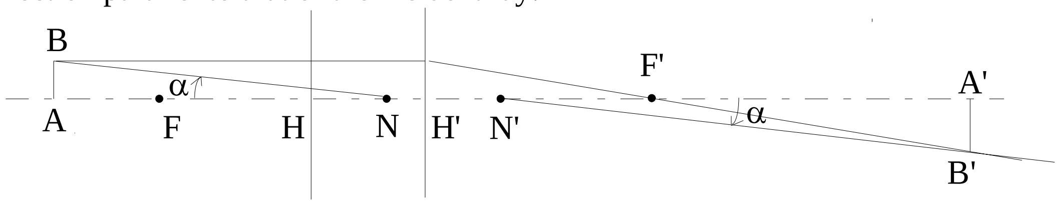


$$(g_\alpha)_{AA'} = \frac{\alpha'}{\alpha} = \frac{\overline{HA}}{\overline{H'A'}}$$

The incident ray passes through A and I, a point in the first principal plane: the emerging ray thus passes through A' image of A and through I', image of I, which is located in the second principal plane and corresponds to a transverse magnification of 1.

A construction of rays can prove very easily that the antiprincipal points (g_y = -1) are symmetric to the principal points with respect to the focal points.

Other specific rays can be used to construct an image, for example a ray passing through the first nodal point N will come out of the system through its conjugate point N' and in a direction parallel to that of the incident ray:

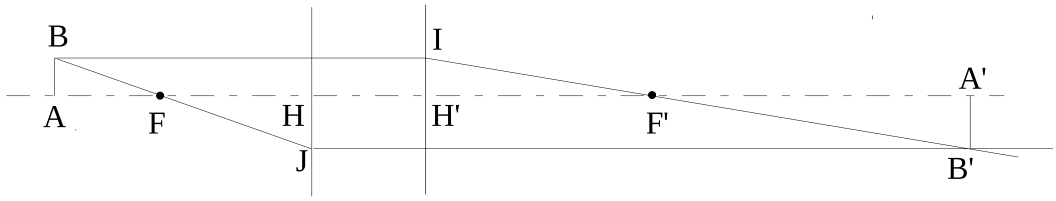


From this construction, we get a new expression of the transverse magnification:

$$(g_y)_{AA'} = \frac{\overline{A'B'}}{\overline{AB}} = \frac{\overline{N'A'}}{\overline{NA}}$$

c) Newton's formula for conjugate points

We want in this paragraph to connect the position of the image to the position of the object taking the focal points as references. Those formulae are different from the previous ones because focal points are not conjugate points. We can get those formulae either from the gaussian formula with H and H' as references writing $\overline{HA}$ as $\overline{HF}+\overline{FA}$ (same thing for $\overline{H'A'}$), or in a simpler way using a construction of rays. Let us go back to the construction of the image of a point B calling I and J the points of intersection of the incident rays with the principal planes:



From the figure, we get:

$$\frac{\overline{AB}}{\overline{FA}} = \frac{\overline{HJ}}{\overline{FH}} = -\frac{\overline{A'B'}}{f}$$

$$\frac{\overline{A'B'}}{\overline{F'A'}} = \frac{\overline{H'I}}{\overline{F'H'}} = -\frac{\overline{AB}}{f'}$$

We deduce from this Newton's formula for conjugate points A and A':

$$\boxed{\overline{FA}.\overline{F'A'} = ff'}$$

as well as two new expressions for the transverse magnification:

$$\boxed{g_y = \frac{\overline{A'B'}}{\overline{AB}} = -\frac{f}{\overline{FA}} = -\frac{\overline{F'A'}}{f'}}$$

We can see that in general, to find with a construction the formulae with references at specific points, we must draw rays passing through these points.

d) equivalence between a focal system and a single lens

If we review everything we just showed about focal systems, we can see that everything relative to the object space is positioned with respect to the first principal plane and to the first focal point, and everything relative to the image space is positioned with respect to the second principal plane and the second focal point. The distance between the principal planes appears as a translation between the image space and the object space. You can convince yourself of this going back to the construction that we just did in the previous paragraphs.

It is thus possible, in a first stage, to calculate everything in the image space reducing the distance between principal planes to zero, which amounts to replacing the optical system by a single thin lens of the same focal length (if necessary the first and second focal length can be different if the indices n and n' are different) located in the first principal plane of the system. In a second stage you will just have to translate the image space by a quantity HH' to find all the characteristics of the system in the paraxial approximation. Note that this method does not imply any other approximation than the paraxial one.

4) Afocal systems

Those are systems with zero power. Their matrices are diagonal, written as:

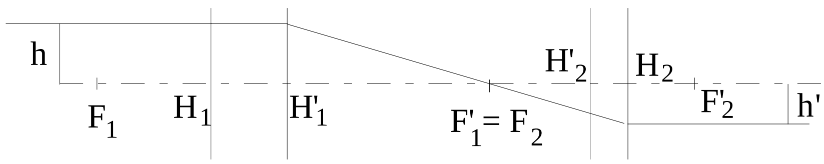
$$\begin{bmatrix} a_1 & 0 \\ 0 & a_4 \end{bmatrix}$$

There are called afocal because they do not have focal points (they are at infinity). Indeed if we consider an incident ray parallel to the axis, with coordinates (Z, Ω=0), the emerging ray is ($Z'=a_1Z$, Ω'=0), it is thus also parallel to the optical axis.

We have also seen that P=0 implies that the transverse and angular magnifications are independent of the pair of conjugate points that we choose, so the matrix of the system is also independent on those points. Since the magnifications are connected by the Lagrange-Helmholtz formula, the afocal system is characterized by the specification of one of its magnifications.

An afocal system can be made using two focal systems, with the second focal point of the first system in the same place as the first focal point of the second system.

Note that here we chose H'_2 before H_2



We can in this case find the transverse magnification using the construction:

$$g_y = \frac{h'}{h} = \frac{\overline{\mathrm{H_2F_2}}}{\overline{\mathrm{H'_1F'_1}}} = \frac{f_2}{f'_1}$$

We can also get the position of the image, using for example the relationship between magnifications that we saw in section III.2:

$$\frac{x'}{x} = \frac{n'}{n} g_y{}^2 = \frac{f_2 f'_2}{f_1 f'_1}$$

where *x* and *x*' give the positions of the object and its image with respect to any pair of other conjugate points (for example F_1 and F'_2).

CONCLUSION

The paraxial approximation allows you to study image formation in a first order approximation in terms of the ray parameters. It has the main advantage of giving very good results in many situations, even though it is very simple: the optical systems, even complex ones, are described by a few parameters (principal planes, focal points) so that you can study the whole system in terms of thin lenses, then you just have to translate the results by the distance between principal plane. The characteristics of images are obtained using linear formulae, or ray constructions, or linear algebra calculations with 2x2 matrices.

In the design of an optical system, the first step will then be to solve the problem in the linear approximation using only thin lenses (even mirrors can be replaced by lenses if you « unfold » the optical axis). Taking into account the role of each lens in the system (for example for which conjugate points it is used, for which aperture, etc), we can replace it by a more complex system that corrects the appropriate aberrations (see course on Optical Design). We will then have to take into account the translation HH' to find the relative positions of the elements in the final system.

In the following chapters, we will study one by one all the simple optical elements, single refractive surfaces, thin lenses and mirrors, for which it is often simpler to make a specific study rather than apply the general formulae seen in this chapter. The general formulae will reappear when we study thick lenses, as well as the combination of centered systems.