Introduction to optical design: I. Matrix techniques and imaging

David Kleinfeld and Philbert Tsai – UCSD

Background

Light and the interaction of light and matter is governed by the rules of quantum electrodynamics. When the photon densities are so too high that electrons are accelerated to relativistic speeds (as occurs at the highest laboratory laser fields), the rules of quantum mechanics can be used. We start here.

QM description: $E = h\nu$[1]. One- and multi-photon transition cross-sections

$\delta E\ \delta t \approx h$. Uncertainty and lifetimes ($h \equiv$ Planks constant).

Wave description: $\nu = c/\lambda$. Diffraction and interference effects.

$\delta\lambda/\lambda \approx \lambda/L_{coh}$ or $L_{coh} \approx \lambda^2/(\delta\lambda)^2$. Spatial coherence effects.

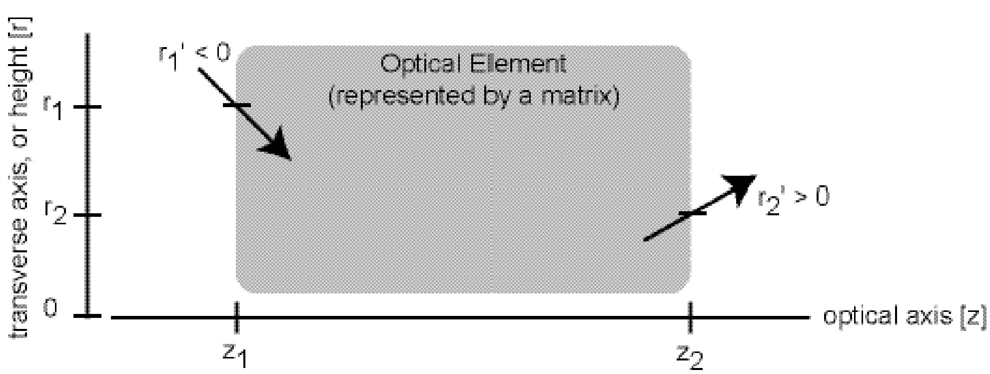
$\delta\theta\ \delta a \approx \lambda$. Uncertainty between angular spread and size of beam.

Geometric optics: $\lambda \to 0$, so that diffraction is ignored and the resolution angle, λ/(size of system) goes to zero as well. Light is described by rays.

Paraxial optics: $\partial[\cdot]/\partial z << \partial[\cdot]/\partial r$. Only small angles, so that imaging design is in terms of straight lines and two parameters, the height of a ray and the slope of the ray.

Matrix method for rays: We start with a notation for specifying optics systems.

- z axial location of a ray.
- r_1 height of a ray at axial location z_1.
- r_1' slope of a ray at axial location z_1, *i.e.*, $r' = \partial r/\partial z$.



---

[1] $E = hc/\lambda = 1.24/\lambda$ eV with λ in micrometers, so that blue light is 3 eV, red light is 2 eV.

[2] Noting that $\delta\lambda \approx c\ \delta t$, $L_{coh} = \lambda^2/(c\ \delta t) \approx (0.8\ \mu m)^2/(3x10^{14}\ \mu m/s \bullet 1x10^{-13}\ s) = 2\ \mu m$ for 2-photon excitation.

n_1 optical index at axial location z_1.

d separation distances along axial axis.

R radius of curvature of a spherical lens.

f focal length (function of n and R at each interface);
we will show that f = R/(n-1) for a thin planoconvex lens.

Formalism to connect rays at two locations. General idea:

- Two unknowns requires two equations.

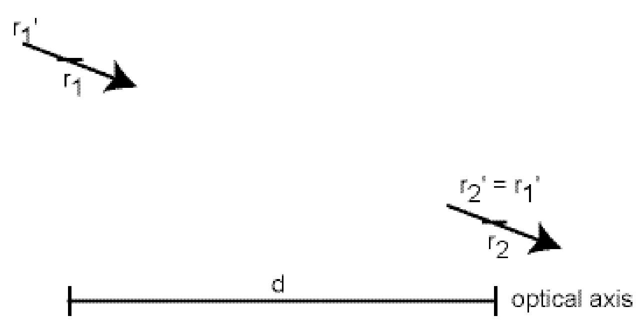
$$\begin{pmatrix} r_2 \\ r_2' \end{pmatrix} = \begin{pmatrix} Ar_1 + Br_1' \\ Cr_1 + Dr_1' \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \bullet \begin{pmatrix} r_1 \\ r_1' \end{pmatrix} \equiv \mathbf{T} \bullet \begin{pmatrix} r_1 \\ r_1' \end{pmatrix}$$

- Represent each interface or transition by an appropriate ABCD, or transfer matrix. Then whole optical systems are reduced to a single matrix by multiplying through matrices for each component. We recall that matrix algebra involves
$\mathbf{T} \bullet \mathbf{t} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \bullet \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} Aa + Bc & Ab + Bd \\ Ca + Dc & Cb + Dd \end{pmatrix}$.
- Ignore reflections – consider only "forward" propagation, so that amplitudes are constant.

Just to complete the story, an extension of this formalism also holds for beams with a Gaussian profile, in which r is replaced by a complex variable that describes the width and curvature of the bean (see Yariv).

Propagation through space. We start with a ray that propagates a distance d in space with constant index.

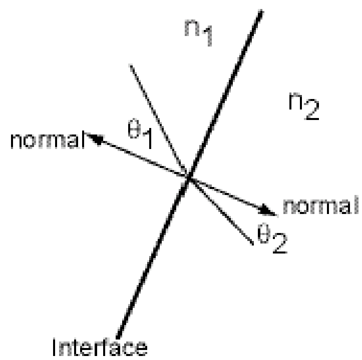
- The slope will remain unchanged.
- The height will change by the slope times the distance.



We have $\begin{pmatrix} r_2 \\ r_2' \end{pmatrix} = \begin{pmatrix} r_1 + d r_1' \\ 0 r_1 + r_1' \end{pmatrix} = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} \bullet \begin{pmatrix} r_1 \\ r_1' \end{pmatrix}$, so the matrix for propagating a distance d is

$$\mathbf{T}_{prop} = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix}.$$

Snells' Law. Before we can move on to rays that impinge on an interface, we need to recall Snell's law – which describes the angle of the transmitted ray.. A derivation in terms of minimizing the time to travel between two points is given in appendix A. We consider a ray that propagates from a material with index n_1 to one with index n_2. The ray is incident at an angle θ_1 relative to the normal to the interface. Snells' law relates the exit angle, θ_2, to the indices and entrance angle by $n_1 \sin\theta_1 = n_2 \sin\theta_2$.



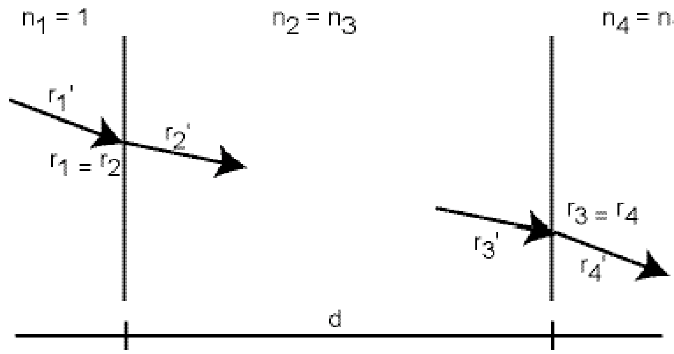
Flat interface. We move on to an interface with index n_1 to the left and n_2 to the right.

- In the paraxial approximation, $r' = \tan\theta = \sin\theta$, so that Snells' law becomes $n_1 r'_1 = n_2 r'_2$.
- Since $r_1 = r_2$ at the interface, $\begin{pmatrix} r_2 \\ r_2' \end{pmatrix} = \begin{pmatrix} r_1 + 0 r_1' \\ 0 r_1 + \frac{n_1}{n_2} r_1' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{n_1}{n_2} \end{pmatrix} \bullet \begin{pmatrix} r_1 \\ r_1' \end{pmatrix}$

Thus the matrix for propagating across the interface is

$$\mathbf{T}_{flat} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{n_1}{n_2} \end{pmatrix}.$$

Example: Propagation through a slab of thickness d. This gives us a first nontrivial example to compute the propagation of a ray. We start on the right and consider the interface from n_1 to n_2 and then back to n_1 again.

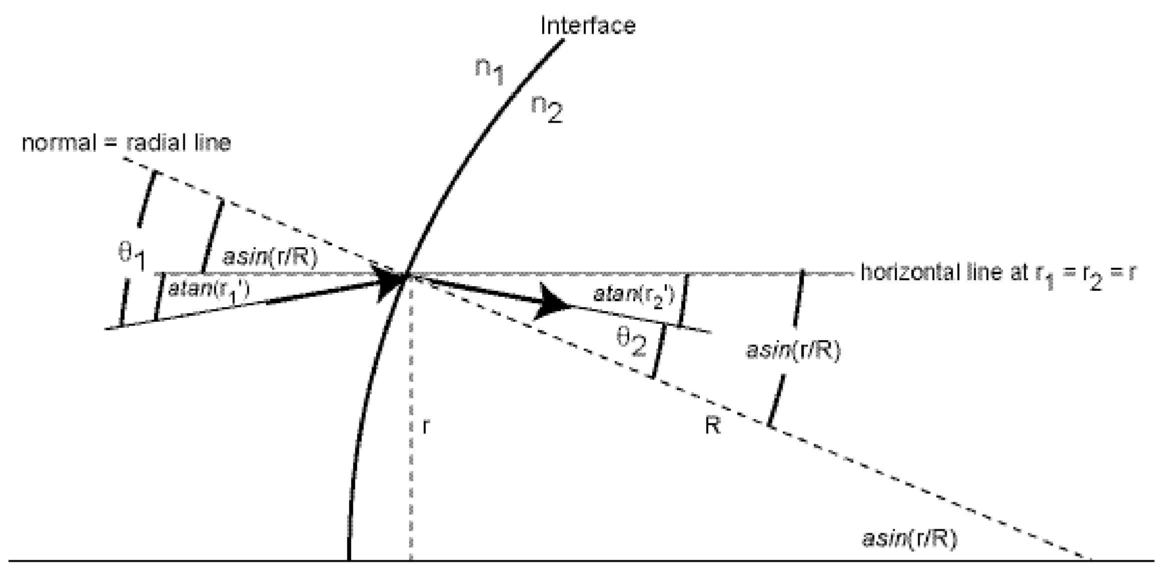


The relationship between the input and the output is:

$$\begin{pmatrix} r_4 \\ r_4' \end{pmatrix} = \mathbf{T_{flat}} \bullet \mathbf{T_{prop}} \bullet \mathbf{T_{flat}} \bullet \begin{pmatrix} r_1 \\ r_1' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{n_2}{n_1} \end{pmatrix} \bullet \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} \bullet \begin{pmatrix} 1 & 0 \\ 0 & \frac{n_1}{n_2} \end{pmatrix} \bullet \begin{pmatrix} r_1 \\ r_1' \end{pmatrix} = \begin{pmatrix} 1 & d\frac{n_1}{n_2} \\ 0 & 1 \end{pmatrix} \bullet \begin{pmatrix} r_1 \\ r_1' \end{pmatrix}$$

Note the ordering of the transfer matrices – which is right to left! The final matrix, $\begin{pmatrix} 1 & d\frac{n_1}{n_2} \\ 0 & 1 \end{pmatrix}$, looks just like the free propagator, except that the distance is normalized by n_1/n_2.[3] The position is displaced, but the entrance and exit angles remain unchanged.

Spherical Interface. We now come to propagation through a curved interface, which is the first step in understanding a lens. We derive the transfer matrix in a general way.



The beam enters and leaves the interface at height

- $r_1 = r_2 \equiv r$.

All angles are all taken to be small

[3] This rescaling is also a statement that the wavelength of light changes with the index.

- The normal to the surface makes an angle a*sin*(r/R) ≈ r/R.
- The slopes are small, so that *atan*(r_1') ≈ r_1' and *atan*(r_2') ≈ r_2'.
- From Snells' law at small angles, $n_1\theta_1 \approx n_2\,\theta_2$ or $n_1\,(r'_1 + r/R) = n_2\,[r/R - (-\,r'_2)]$.
- Thus $r'_2 = (n_1/n_2 - 1)(r/R) + (n_1/n_2)r'_1$ so that $\begin{pmatrix} r_2 \\ r_2' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{n_1-n_2}{n_2}\frac{1}{R} & \frac{n_1}{n_2} \end{pmatrix} \bullet \begin{pmatrix} r_1 \\ r_1' \end{pmatrix}$

$$\mathbf{T}_{\text{curved}} = \begin{pmatrix} 1 & 0 \\ \frac{n_1-n_2}{n_2}\frac{1}{R} & \frac{n_1}{n_2} \end{pmatrix}.$$

In the limit that $R \to \infty$ we recover our result for the flat surface.

Lens. The most general lens involves a change in index through a curved surface of radius R_1, propagation through the lens material, and exit through a second curved surface of radius R_2. We save this general case for Appendix B, and consider the special but useful case of a thin lens that is curved on the left side with radius of curvature R_1 and on the right side with radius of curvature R_2. From the above result,

- The indices at the entrance side and exit side are identical, *i.e.*, n_1, while the index in the lens is n_2.
- $\begin{pmatrix} r_3 \\ r_3' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{n_2-n_1}{n_1}\frac{1}{R_2} & \frac{n_2}{n_1} \end{pmatrix} \bullet \begin{pmatrix} 1 & 0 \\ \frac{n_1-n_2}{n_2}\frac{1}{R_1} & \frac{n_1}{n2} \end{pmatrix} \bullet \begin{pmatrix} r_1 \\ r_1' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{n_2-n_1}{n_1}\left(\frac{1}{R_1} - \frac{1}{R_2}\right) & 1 \end{pmatrix} \bullet \begin{pmatrix} r_1 \\ r_1' \end{pmatrix}.$
- We define the focal length, denoted f, through

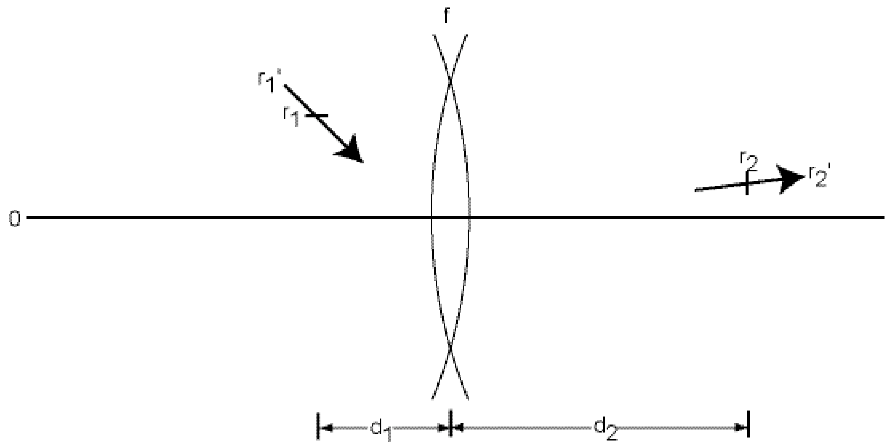
$$\frac{1}{f} = \frac{n_2-n_1}{n_1}\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

where R is positive for a convex surface (of particular relevance in microscopy) and negative for a concave surface. For a symmetric biconvex lens, $R_2 = -\,R_1$. For a planoconvex lens, $R_2 \to \infty$.

$$\mathbf{T}_{\text{thin\_lens}} = \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix}.$$

Imaging (at last!). We now consider three classic configurations of lenses.

Space-lens-space. A ray starts a distance d_1 from a thin lens and is observed a distance d_2 away.



Thus $\begin{pmatrix} r_2 \\ r_2' \end{pmatrix} = \begin{pmatrix} 1 & d_2 \\ 0 & 1 \end{pmatrix} \bullet \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \bullet \begin{pmatrix} 1 & d_1 \\ 0 & 1 \end{pmatrix} \bullet \begin{pmatrix} r_1 \\ r_1' \end{pmatrix} = \begin{pmatrix} 1-\frac{d_2}{f} & d_1 + d_2 - \frac{d_1 d_2}{f} \\ -\frac{1}{f} & 1-\frac{d_1}{f} \end{pmatrix} \bullet \begin{pmatrix} r_1 \\ r_1' \end{pmatrix}$

$$\mathbf{T}_{s\text{-}l\text{-}s} = \begin{pmatrix} 1-\frac{d_2}{f} & d_1 + d_2 - \frac{d_1 d_2}{f} \\ -\frac{1}{f} & 1-\frac{d_1}{f} \end{pmatrix}.$$

- **Imaging condition**. This corresponds to a final height, r_2, that is independent of initial slope, *i.e.*, r'_1, so that all rays that leave r_1 converge at r_2. Thus the "B" term of the transfer matrix must be zero, *i.e.*, $d_1 + d_2 - (d_1 d_2)/f = 0$. This leads to the famous lens formula

$$\boxed{\frac{1}{d_1} + \frac{1}{d_2} = \frac{1}{f}} \quad \text{or} \quad \boxed{(d_1 - f)(d_2 - f) = f^2}$$

- **Magnification.** When the imaging condition is satisfied, the "A" term corresponds to magnification, *i.e.*, $r_2 = [1-(d_2/f)]r_1 = -(d_2/d_1)r_1$ where the minus sign means the image in inverted.

$$\boxed{\frac{r_2}{r_1} = -\frac{d_2}{d_1}}$$

- **Real vs. Virtual.** Note that for the case of d, < f, the lens formula implies that $d_2 < 0$ and thus the sign of the magnification is now positive. Therefore, an upright image is formed to the left of the lens. There is no image to project on a screen. However, the upright image can form the object for another lens.
- **Ray tracing rules.** The imaging condition leads to a transfer matrix that can be written in three equivalent forms:

$$\mathbf{T}_{s\text{-}l\text{-}s} \rightarrow \begin{pmatrix} -\frac{f}{d_1-f} & 0 \\ -\frac{1}{f} & -\frac{d_1-f}{f} \end{pmatrix} = \begin{pmatrix} -\frac{d_2-f}{f} & 0 \\ -\frac{1}{f} & -\frac{f}{d_2-f} \end{pmatrix} = \begin{pmatrix} -\frac{d_2}{d_1} & 0 \\ -\left(\frac{1}{d_1} + \frac{1}{d_2}\right) & -\frac{d_1}{d_2} \end{pmatrix}.$$

Each of these forms leads to one of three ray-tracing rules:

1. Consider a horizontal ray from the "object", *i.e.*, $r'_1 = 0$:

$$\begin{pmatrix} r_2 \\ r'_2 \end{pmatrix} = \begin{pmatrix} -\frac{d_2-f}{f} & 0 \\ -\frac{1}{f} & -\frac{f}{d_2-f} \end{pmatrix} \bullet \begin{pmatrix} r_1 \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{d_2-f}{f} r_1 \\ -\frac{1}{f} r_1 \end{pmatrix}.$$

   The output ray will cross the axis at $d_2 = f$, the right focal point of the lens, and is imaged at

$$\begin{pmatrix} r_2 \\ r'_2 \end{pmatrix} = \begin{pmatrix} -\frac{d_2-f}{f} r_1 \\ -\frac{1}{f} r_1 \end{pmatrix} = \begin{pmatrix} -\frac{d_2}{d_1} r_1 \\ -\frac{1}{f} r_1 \end{pmatrix}.$$

2. Consider a ray from the "object" to the center of the lens, *i.e.*, $r'_1 = -r_1/d_1$:

$$\begin{pmatrix} r_2 \\ r'_2 \end{pmatrix} = \begin{pmatrix} -\frac{d_2}{d_1} & 0 \\ -\left(\frac{1}{d_1} + \frac{1}{d_2}\right) & -\frac{d_1}{d_2} \end{pmatrix} \bullet \begin{pmatrix} r_1 \\ -\frac{r_1}{d_1} \end{pmatrix} = \begin{pmatrix} -\frac{d_2}{d_1} r_1 \\ -\frac{r_1}{d_1} \end{pmatrix}.$$

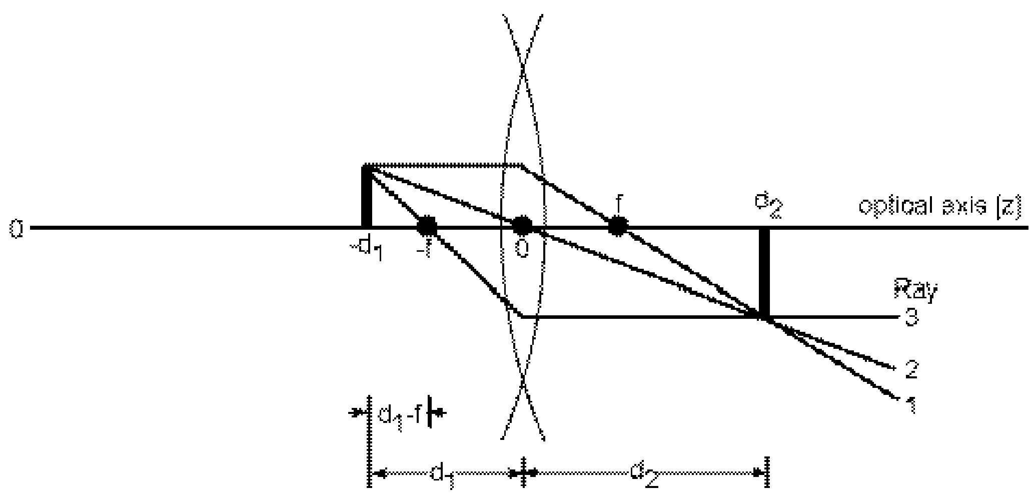
   The output ray has a slope that is unchanged.

3. Consider a ray from the "object" that passes through the left focal point, *i.e.*, $r'_1 = -r_1/(d_1-f)$

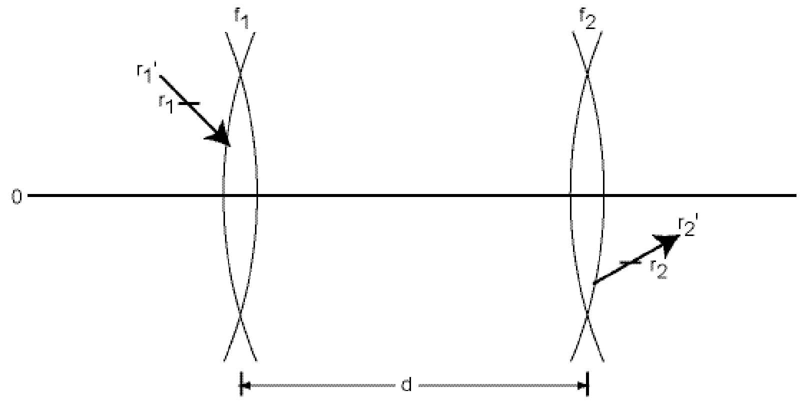
$$\begin{pmatrix} r_2 \\ r'_2 \end{pmatrix} = \begin{pmatrix} -\frac{f}{d_1-f} & 0 \\ -\frac{1}{f} & -\frac{d_1-f}{f} \end{pmatrix} \bullet \begin{pmatrix} r_1 \\ -\frac{r_1}{d_1-f} \end{pmatrix} = \begin{pmatrix} -\frac{d_2}{d_1} r_1 \\ 0 \end{pmatrix}.$$

   The output ray has a slope that is now zero.

These three rays will converge at a common point, $d_2 = (fd_1)/(d_1-f)$, as given by the imaging condition.



Lens-space-lens. A ray enters one thin lens, travels a distance d, and then exits through a second thin lens.



Thus $\begin{pmatrix} r_2 \\ r_2' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{1}{f_2} & 1 \end{pmatrix} \bullet \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} \bullet \begin{pmatrix} 1 & 0 \\ -\frac{1}{f_1} & 1 \end{pmatrix} \bullet \begin{pmatrix} r_1 \\ r_1' \end{pmatrix} = \begin{pmatrix} 1-\frac{d}{f_1} & d \\ \frac{d-f_1-f_2}{f_1 f_2} & 1-\frac{d}{f_2} \end{pmatrix} \bullet \begin{pmatrix} r_1 \\ r_1' \end{pmatrix}$

$$\mathbf{T}_{\text{l-s-l}} = \begin{pmatrix} 1-\frac{d}{f_1} & d \\ \frac{d-f_1-f_2}{f_1 f_2} & 1-\frac{d}{f_2} \end{pmatrix}.$$

- **Back-to-back lenses.** We let $d \to 0$ and the transfer matrix becomes

$$\mathbf{T}_{\text{l-s-l}} \to \begin{pmatrix} 1 & 0 \\ -\frac{f_1+f_2}{f_1 f_2} & 1 \end{pmatrix}$$

Thus the composite lens has an effective focal length equal to the geometric mean of the focal length of the two lenses, *i.e.*, $\frac{1}{f_{\text{effective}}} = \frac{1}{f_1} + \frac{1}{f_2}$.

- **Beam expander.** This corresponds to the case of collimated output, for which the slope r'_2 is independent of r_1. Thus the "C" term in $\mathbf{T}_{\text{l-s-l}}$ must be equal to zero, *i.e.*, $d - f_1 - f_2 = 0$. This leads to the famous telescope formula:

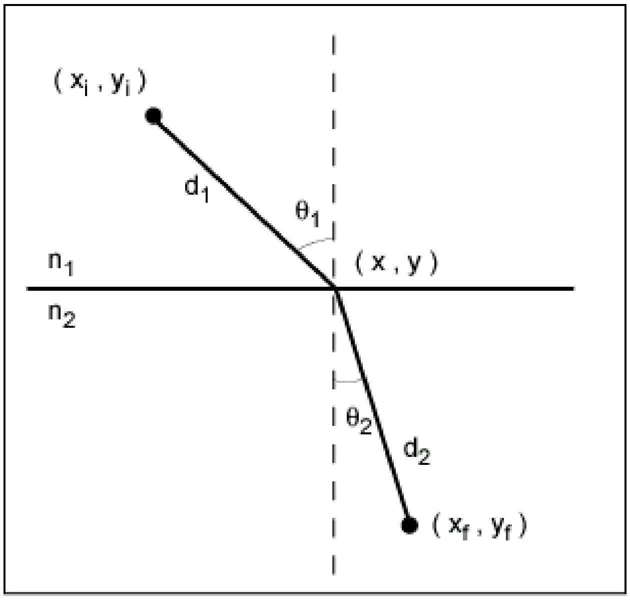
$$\boxed{d = f_1 + f_2}$$

which can be used to change the height of parallel light by $\frac{r_2}{r_1} = \frac{f_2}{f_1}$. The transfer matrix becomes:

$$\mathbf{T}_{\text{l-s-l}} \to \begin{pmatrix} -\frac{f_2}{f_1} & f_1 + f_2 \\ 0 & -\frac{f_1}{f_2} \end{pmatrix}.$$

Note that f_1 and f_2 are both be positive, or that one of f_1 and f_2 can be positive and the other negative.

Appendix A - Derivation of Snell's Law based on principle of minimum time



Distances

$$d_1 = \sqrt{(x - x_i)^2 + (y - y_i)^2}$$

$$d_2 = \sqrt{(x_f - x)^2 + (y_f - y)^2}$$

Speeds
(c ≡ speed of light in vacuum)

$$v_1 = \frac{c}{n_1}$$

$$v_2 = \frac{c}{n_2}$$

Calculate the total time of transit, T.

$$T = \frac{d_1}{v_1} + \frac{d_2}{v_2} = \frac{n_1}{c}\sqrt{(x - x_i)^2 + (y - y_i)^2} + \frac{n_2}{c}\sqrt{(x_f - x)^2 + (y_f - y)^2}$$

Take the derivative of the total transit time with respect to the position, x. As we are interested at the contact point at the plane of the interface, we set y = 0.

$$\left.\frac{dT}{dx}\right|_{y=0} = \frac{n_1}{c}\frac{2(x - x_i)}{\sqrt{(x - x_i)^2 + (y - y_i)^2}} - \frac{n_2}{c}\frac{2(x_f - x)}{\sqrt{(x_f - x)^2 + (y_f - y)^2}}$$

$$= \frac{n_1}{c}\, 2 \sin(\theta_1) - \frac{n_2}{c}\, 2 \sin(\theta_2)$$

The total transit time is minimized when $dT/dx|_{y=0} = 0$. This leads to the identity

$$\mathbf{n_1}\, \boldsymbol{sin\theta_1} = \mathbf{n_2}\, \boldsymbol{sin\theta_2}.$$

Appendix B – Thick Lens

The most general lens involves a change in index through a curved surface of radius R_1, propagation through the lens material, and exit through a second curved surface of radius R_2.

The indices at the entrance side and exit sides are identical, *i.e.*, n_1, while the index in the lens is n_2. The thickness is denoted s. Thus

$$\begin{pmatrix} r_4 \\ r_4' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{n_2-n_1}{n_1}\frac{1}{R_2} & \frac{n_2}{n_1} \end{pmatrix} \bullet \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} \bullet \begin{pmatrix} 1 & 0 \\ \frac{n_1-n_2}{n_2}\frac{1}{R_1} & \frac{n_1}{n2} \end{pmatrix} \bullet \begin{pmatrix} r_1 \\ r_1' \end{pmatrix}$$

$$= \begin{pmatrix} 1-\frac{n_2-n_1}{n_2}\frac{S}{R_1} & \frac{n_1}{n_2}s \\ -\frac{n_2-n_1}{n_1}\left(\frac{1}{R_1}-\frac{1}{R_2}+\frac{n_2-n_1}{n_2}\frac{S}{R_1R_2}\right) & 1+\frac{n_2-n_1}{n_2}\frac{S}{R_2} \end{pmatrix} \bullet \begin{pmatrix} r_1 \\ r_1' \end{pmatrix}.$$

This lens cannot be parameterized by a single number as for the case of a thin lens, *i.e.*, the focal length f. Nonetheless, for symmetric biconvex lens, so that $R_2 = - R_1$, we have

$$\mathbf{T} = \begin{pmatrix} 1-\frac{S}{2f} & \frac{n_1}{n_2}s \\ -\frac{1}{f}\left(1-\frac{S}{4f}\right) & 1-\frac{S}{2f} \end{pmatrix}$$

Where we used our previous definition of the focal length. For a planoconvex lens, so that $R_2 \rightarrow \infty$, we have

$$\mathrm{T} = \begin{pmatrix} 1-\frac{S}{f} & \frac{n_1}{n_2}s \\ -\frac{1}{f} & 1 \end{pmatrix}.$$

Thus the thickness of the lens is only an issue when the thickness approaches the focal length of the lens, which implies that the thickness approaches the radius of curvature. This can occur for objectives and high efficiency collector lenses.