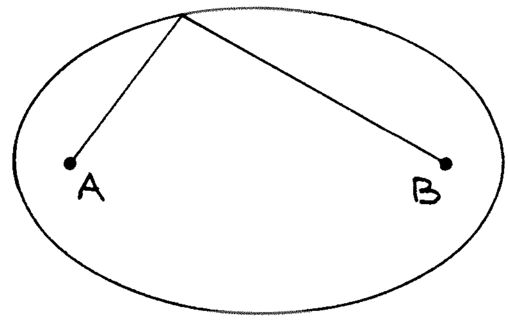
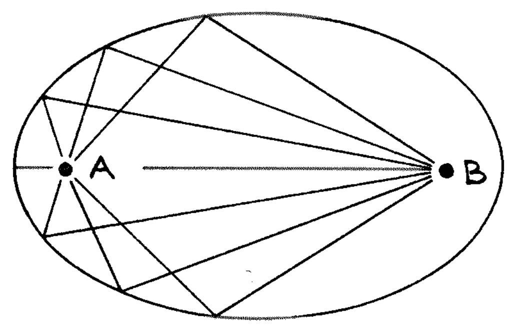
A NAIVE APPROACH TO GEOMETRICAL OPTICS

by A. Naeve

If you pin down two nails
on a piece of paper
at points A and B
and tie a piece of string
between them
without pulling it tight,
then every student knows that if he pushes the string tight
with his pen, and then moves it in this fashion,
the tip of the pen will trace out an ellipse
with A and B as focalpoints.



Most students also know
that this curve
has the property
of reflecting
a point source wave field
diverging from A
into a point field
converging to B



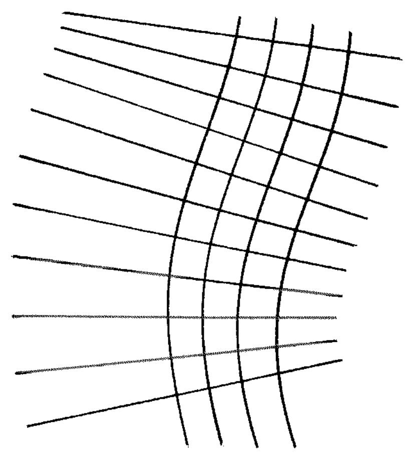
This is a particular case of a more general problem:

Suppose we are given two (two-dimensional) wave fields $\mathcal{A}$ and $\mathcal{B}$ in a homogeneous, isotropic medium. How can we construct the set of curves that will reflect them into each other?

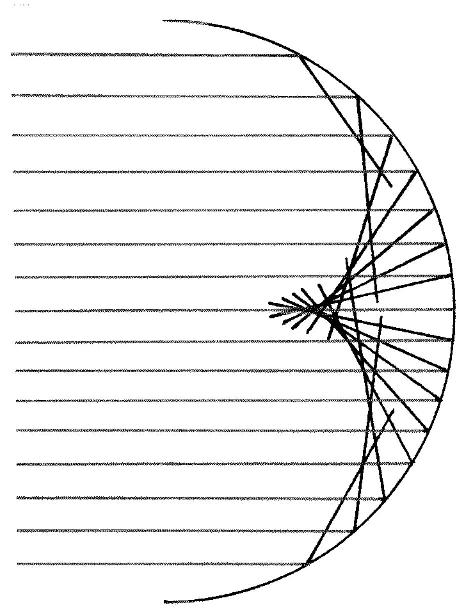
For the sake of convenience, let us think of the fields $\mathcal{A}$ and $\mathcal{B}$ as fields of light.

By the assumptions on the medium, the wavefronts of each field will form a so called <u>normal congruence</u>.
This means that through each point of any particular wavefront of the field, there is a straight line which is orthogonal to all the wavefronts of the field.
Such a line is called <u>a ray</u> of the field.

Thus the normal congruence of the wavefronts of our lightfields φ and ψ simply means that they can be thought of as aggregates of two complementary families of objects:
<u>wavefronts</u> and <u>lightrays</u>
each member of one family being orthogonal to all members of the other.



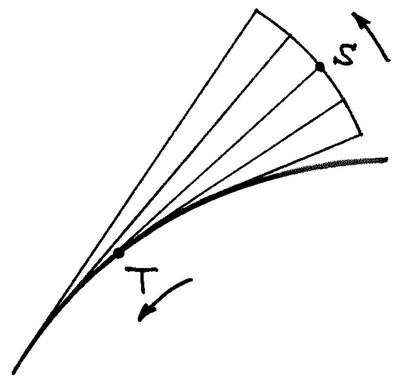
When the sun shines through a glass of water we can observe a bright, shining "curve of light" at the bottom of the glass.
This curve is called a <u>caustic</u> and it represents a local concentration of field energy.
If we think of the field as consisting of lightrays, then its caustic is of course identical to the <u>envelope</u> of the rays



If we think of the field as consisting of wavefronts, the following construction will show us how to describe its caustic.

Take a piece of string and stretch it tightly along the caustic on the convex side of it (we can think of the caustic as a ridge, extending out of the plane of the paper, so that we can actually wrap a piece of string around it.) Keeping one end fixed, start unwrapping the string by the other end (S) always keeping the unwrapped part of it stretched in a straight line.

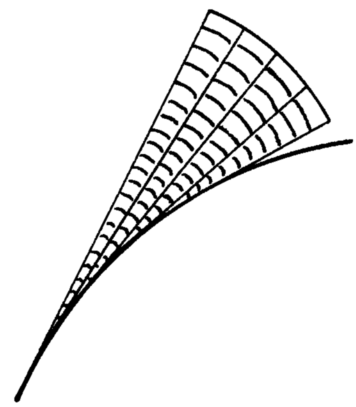
Then, as the string moves, the straight part of it (ST) will range through all the rays of the field, whereas the endpoint (S) of the string will trace out a single wavefront



By changing the initial position of S, we can trace out any wavefront of the field in this manner. Thus, all wavefronts are <u>involutes</u> of the caustic and therefore the caustic must be the <u>evolute</u> of the wavefronts.

This can also be seen directly by observing that the motion of S is "locally circular" with centre T, which means that, as T moves it must trace out the locus of the centre of curvature of any wavefront.

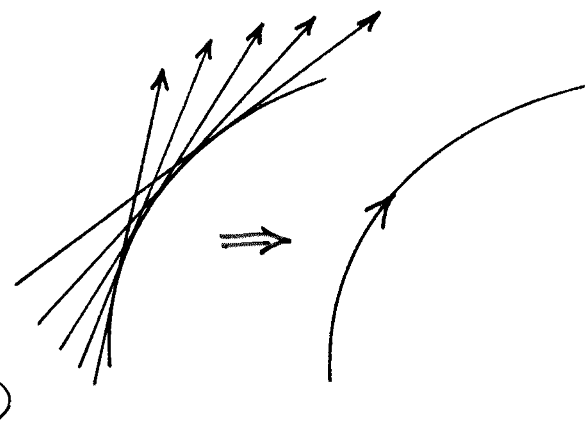
Therefore a point on the caustic can be regarded as a <u>local focal point</u> of the wavefronts in the corresponding direction.



After these preliminaries, let us now return to the problem posed at the beginning:

Given two wavefields φ and ψ, how can we construct the set of curves that will reflect them into each other?

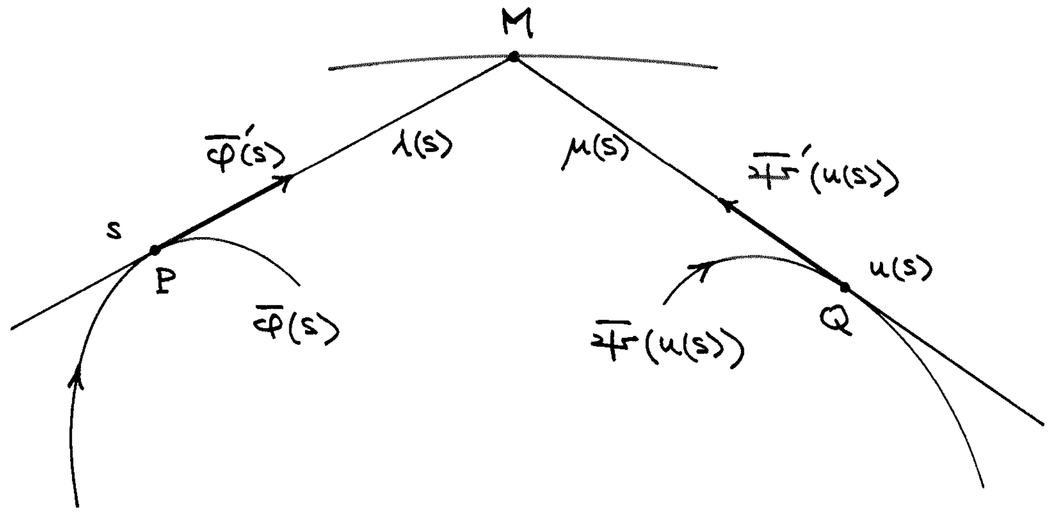
We will assume that both fields are stationary and that all rays involved carry an equal amount of energy. This means that each field is completely characterized by its caustic and its raydirections. These directions will induce a direction on the caustic in an obvious way. (Note that the points on the caustic have infinite field intensity)



Before tackling the problem, let us fix some notation.

Let φ be the wavefield that is incident on the mirror m, and let ψ be the reflection of φ in m.

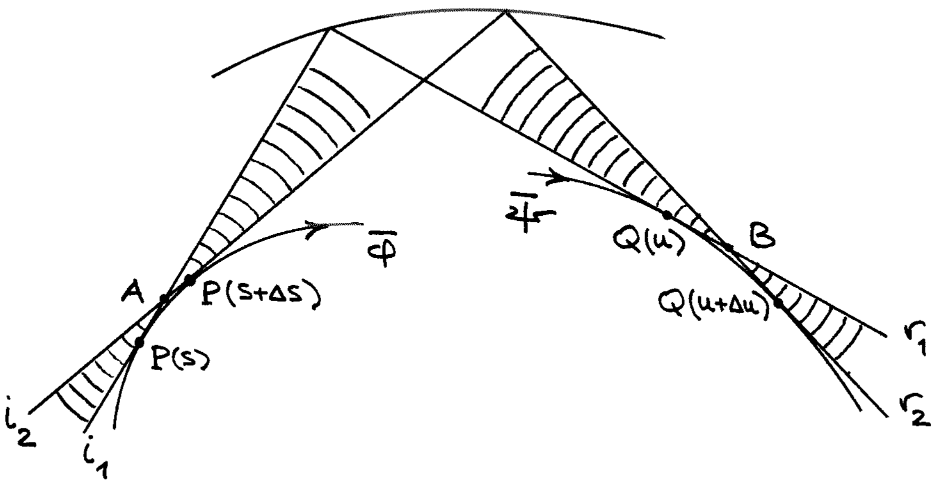
We will parametrize everything in terms of the arclength s of the caustic of φ.



As the point P (with arclength s) varies along this caustic, its position vector is called $\overline{\varphi}(s)$. The lightray through P will have the direction $\overline{\varphi}'(s)$ and it will hit the mirror at the point M with position vector $\overline{m}(s)$ at a distance $\lambda(s)$ from P. The reflected ray MQ will touch the "reflected caustic" $\overline{\psi}(u(s))$ at the point Q with caustic arclength $u(s)$ at a distance $\mu(s)$ from M.

For technical reasons that will appear later, we will assume that both s and u are increasing in the direction towards the mirror $\overline{m}$. This means that the in-ray PM is directed along $\overline{\varphi}'(s)$, but the out-ray MQ is directed along $-\overline{\psi}'(u(s))$. Observe that both $\overline{\varphi}'(s)$ and $\overline{\psi}'(u(s))$ are unit vectors, since both s and u are arclength parameters.

Let us look at two neighboring in-rays i_1 and i_2 touching $\overline{\varphi}$ at the points $P(S)$ and $P(S+\Delta S)$ respectively, and intersecting each other at the point A.
The corresponding out-rays are r_1 and r_2, touching $\overline{\psi}$ at $Q(u)$ and $Q(u+\Delta u)$ and intersecting each other at the point B.
Study the bundle C_A of in-rays between i_1 and i_2 (let us call it "the in-cone from A"). The corresponding part of the wavefronts will have centres of curvature (local focal points) ranging from $P(S)$ to $P(S+\Delta S)$ along $\overline{\varphi}$.



When it hits the mirror, C_A will be reflected into C_B (the out-cone from B) whose corresponding wavefronts have local focalpoints ranging from $Q(u)$ to $Q(u+\Delta u)$ along $\overline{\psi}$.
If ΔS is small enough, the variation (ΔS) of the local focalpoints of the wavefronts of C_A and the corresponding variation (Δu) of C_B are small compared to the distances from A and B to the "place of reflection" at the mirror. Therefore, at the mirror, the wavefronts of C_A and C_B will appear to be approximately circular and concentric with centres P and Q respectively

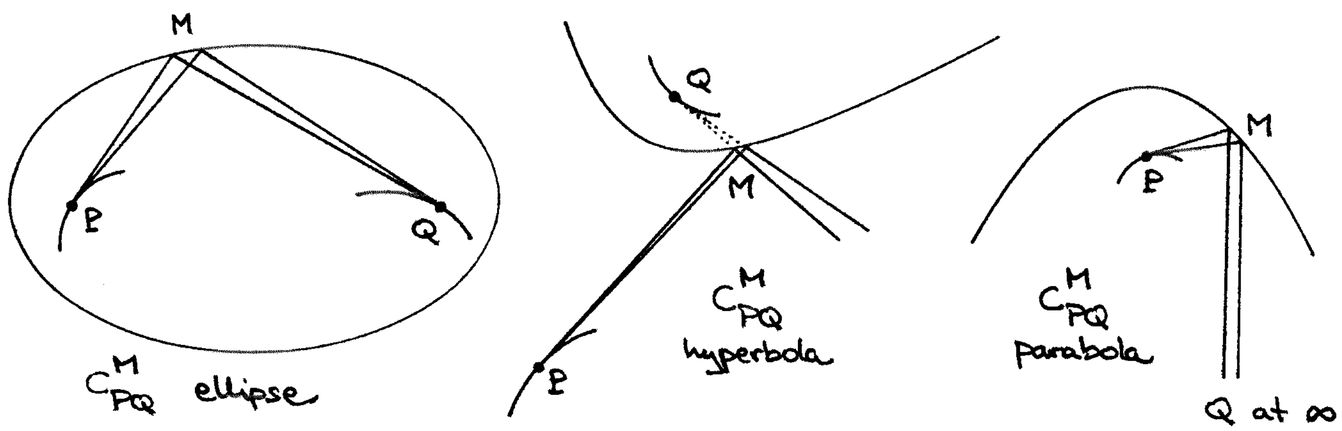
As $\Delta S \to 0$, we have $A \to P(S)$ and $P(S+\Delta S) \to P(S)$.
Therefore, as ΔS decreases, the approximation will be better and better, and in the limit it will be "perfect" ($C_A \to C_{P(S)}$ and $C_B \to C_{Q(u)}$).
We can therefore state, that the action of the mirror turns the (infinitesimal) in-cone $C_{P(S)}$ into the (infinitesimal) out-cone $C_{Q(u)}$.

But this is equivalent to the statement:

Proposition 1:

WHEN IT TRANSFORMS $C_{P(s)}$ INTO $C_{Q(u)}$
THE MIRROR ACTS AS A CONIC
WITH FOCALPOINTS $P(s)$ AND $Q(u)$
PASSING THROUGH THE POINT OF REFLECTION $M(s)$

This conic will be called the local conic at P, Q, M and it will be designated by C^M_{PQ}.
It is easy to see, that if P and Q lie on the same side of the tangent to the mirror at M, then C^M_{PQ} is an ellipse, if P and Q lie on opposite sides of this tangent, then C^M_{PQ} is a hyperbola (degenerated into a straight line if $|PM| = |MQ|$), if either P or Q lies at infinity C^M_{PQ} is a parabola and finally, if both P and Q lie at infinity C^M_{PQ} is a straight line.



Let us note here, that in terms of its mirror action on a wavefield from a real or virtual pointsource at one focalpoint, we have:

Ellipse $= \begin{pmatrix} \text{real} \to \text{real} \\ \text{virtual} \to \text{virtual} \end{pmatrix}$ pointfieldtransformer

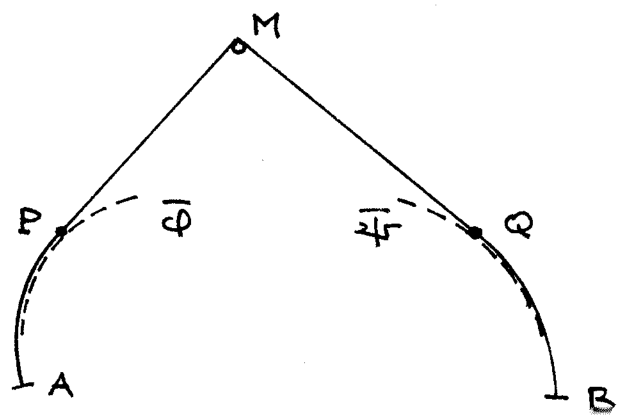
Hyperbola $= \begin{pmatrix} \text{real} \to \text{virtual} \\ \text{virtual} \to \text{real} \end{pmatrix}$ pointfieldtransformer

Parabola $= \begin{pmatrix} \binom{\text{real}}{\text{virtual}} \to \text{parallell} \\ \text{parallell} \to \binom{\text{real}}{\text{virtual}} \end{pmatrix}$ pointfieldtransformer

We will not attempt a rigorous proof of proposition 1 at this point. Rather we will use it to conceive of a construction, that will vindicate it in retrospect.

Starting with two caustics $\overline{\varphi}$ and $\overline{\psi}$, let us wrap a piece of string around them and attach its endpoints rigidly to the points A and B.

If we push the string tight with our pen M, and then move it in this fashion, (just as in the elementary elliptic string construction) what kind of curve will be traced out by the tip of the pen?



By observing the motion locally around the points of contact P and Q between the string and the two caustics, and comparing it to the elliptic string construction, we see that in the immediate vicinity of P and Q the pen will trace out the local conic C^M_{PQ}.

Now, if this curve represents the local action of the mirror (as stated by proposition 1), then it is reasonable to expect that:

<u>Proposition 2</u>:

AS THE PEN MOVES, IT WILL TRACE OUT THE CURVE OF THE MIRROR THAT REFLECTS $\overline{\varphi}$ INTO $\overline{\psi}$ AND PASSES THROUGH THE STARTING POINT OF THE PEN

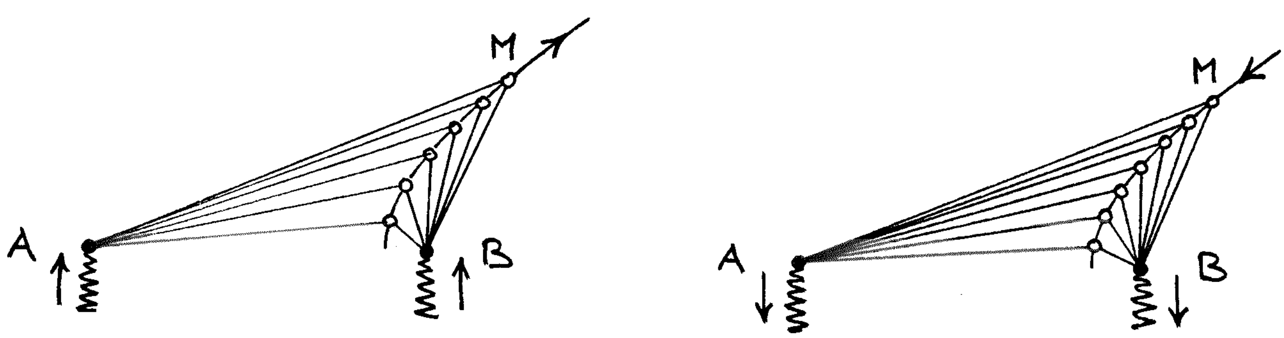
Just as before, we will postpone the proof until we have developed a more general construction, the proof of which will give us the proofs of all the preceding propositions as special cases.

In our aim to generalize, let us observe to begin with, that proposition 2 has a limited range of applicability. It only works when all pairs of local focalpoints P, Q lie on the same side of the corresponding tangent to the mirrorcurve at the point of reflection M, i.e. when all the local conics C^{M}_{PQ} are ellipses. It is easy to see that this happens precisely when our caustics $\overline{\varphi}$ and $\overline{\mathcal{Z}}$ are either both real or both virtual.

How can we generalize our string construction to make it work for arbitrary local conics?

Before we can do this, we must of course avail ourselves of a string construction for the hyperbola. Remembering that the hyperbola is the locus of points whose distance from two fixed points (the focalpoints) have a constant <u>difference</u>, it is obvious that a hyperbolic string construction cannot be obtained by using a string with constant length.

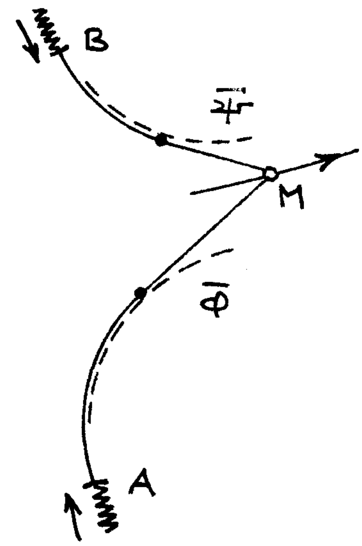
However, if we modify the elliptic string construction as shown in the figure, by attaching the pen rigidly to the string at M and by pulling in (or letting out) string with equal speed from the points A and B, then it is clear that the pen will move towards (or away from) A and B with equal speed, and thus it will maintain a constant difference of distance to A and B. Hence it will trace out a hyperbola with A and B as focalpoints.



Returning to our caustics $\bar{\Phi}$ and $\bar{\Psi}$, let us now assume that one (e.g. $\bar{\Phi}$) is real and the other ($\bar{\Psi}$) is virtual. This means of course, that the local conics C^M_{PQ} will all be hyperbolas (see figure)

Now it is obvious how to generalize the hyperbolic string construction to a locally hyperbolic mirror construction:

We wrap the string around the caustics as before, but instead of attaching its endpoints rigidly to A and B, we let out (or pull in) the string from A and B with equal speed, and attach the pen rigidly to the string at M



Let us state explicitely that:

Proposition 3:

AS THE PEN MOVES, IT WILL TRACE OUT THE LOCALLY HYPERBOLIC MIRROR THAT REFLECTS THE REAL (VIRTUAL) $\bar{\Phi}$ INTO THE VIRTUAL (REAL) $\bar{\Psi}$ AND PASSES THROUGH THE STARTING POINT OF THE PEN

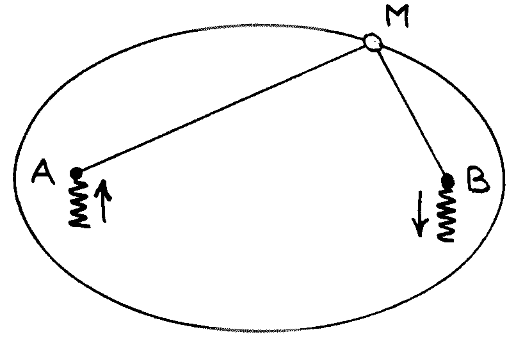
Postponing the proof (as usual!) we now have constructions that take care of both locally elliptic and locally hyperbolic reflexions. But we still cannot handle a reflexion that has a mixed character, i.e. where the family of local conics contains both ellipses and hyperbolas.

How can we connect the constructions of propositions 2 and 3 so as to be able to carry them both out at one stroke?

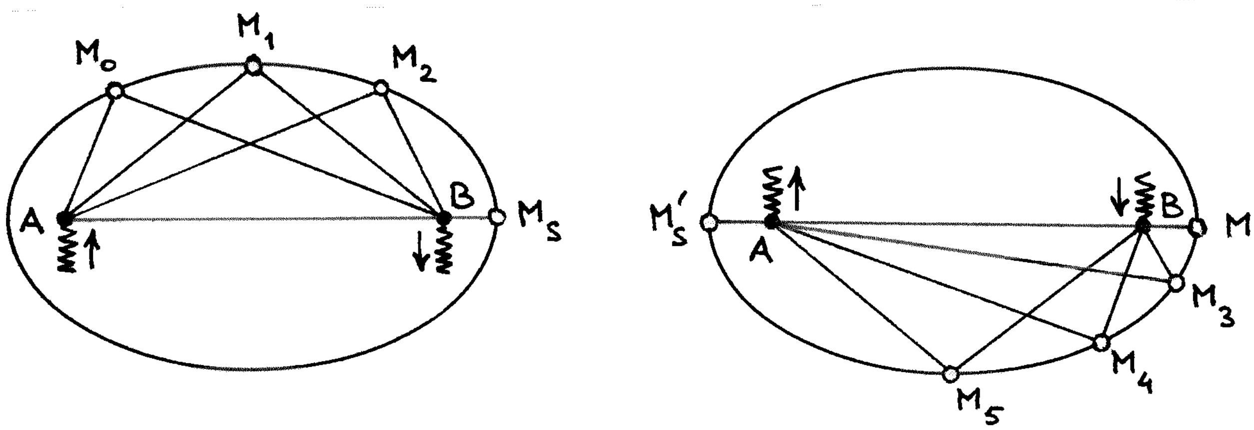
The key to this problem is a reformulation of the elementary elliptic string construction, that makes it more adaptable to the hyperbolic situation.

Imagine an ellipse being constructed as in figure
We can of course do it the usual way with fixed endpoints of the string at A and B

BUT WE CAN ALSO DO IT BY LETTING OUT STRING FROM A AND PULLING IT IN AT B WITH EQUAL SPEED



This is of course equivalent to the classical construction since the stringlength between A and B remains constant. Notice that in this new way of constructing the ellipse we can allow ourselves to attach the pen rigidly to the string at M, thus moving the pen by pulling the string, as long as we introduce the following convention:



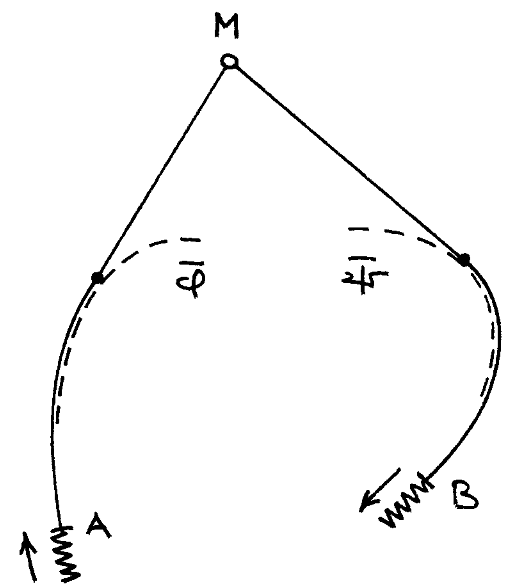
Starting at M_0, letting out string from A and pulling it in at B, the pen will move through $M_1, M_2, \ldots$ tracing out the desired ellipse until it reaches the vertex-point M_s at the end of the major axis.
However, if we try to continue beyond M_s, the pen will no longer follow the desired curve, because it will be pulled in too close to B and the part AM of the string will loose its tension.

The way to remedy this deficiency is clear. When we reach the point M_s, we simply switch the in- and out directions of the string and start pulling it in at A and letting it out from B instead.

In this way we will trace the ellipse through $M_3, M_4, M_5, \ldots$ until we reach the other major vertex M_s'. At this point we again reverse the motion of the string and continue as before.
Calling a ray <u>singular</u> if it passes through both focalpoints without being reflected inbetween, we can name this procedure a <u>singularity switch</u>.

Observe that the singularity switch also occurs in the hyperbolic string construction (figure).
If we start by pulling in string at A and B, the pen M will move towards these points, tracing a hyperbolic arc, but as it reaches the singular ray AB, we have to switch the motion of the string and start letting it out from both A and B, if we want to trace the hyperbola further.

The new (sliding string) ellipse construction generalizes of course immediately to a (sliding string) locally elliptic mirror construction:
In performing the construction of proposition 2 (cfr. figure) we attach the pen rigidly to the string at M, and start letting out string from A and pulling it in at B (or vice versa) with equal speed, always observing the convention of the singularity switch whenever we come upon a singular ray.

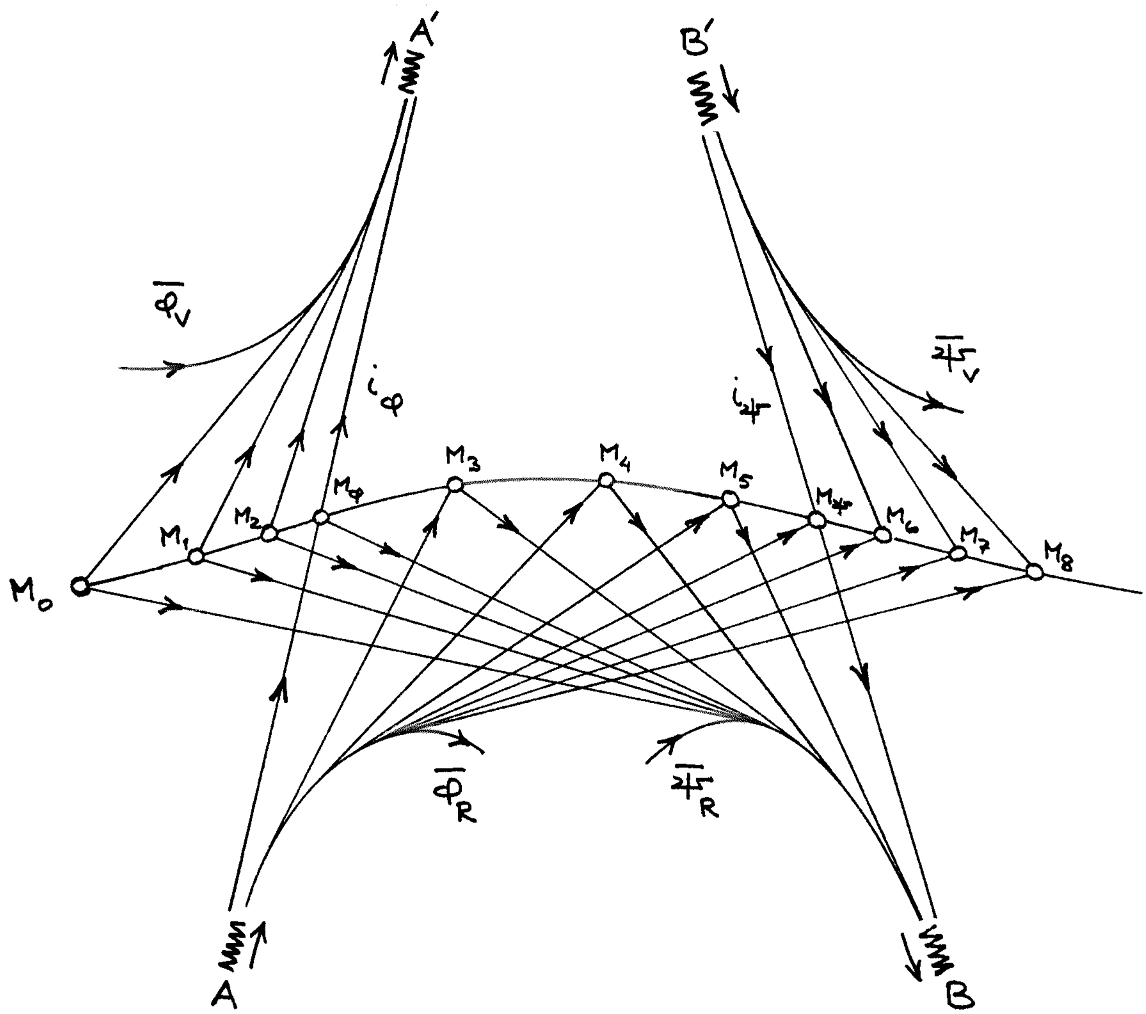


Let us summarize and reformulate proposition 2 as:

Proposition 4:

AS IT MOVES IN THIS WAY, THE PEN WILL TRACE THE LOCALLY ELLIPTIC MIRROR, THAT REFLECTS $\overline{\varphi}$ INTO $\overline{\psi}$ AND PASSES THROUGH THE STARTINGPOINT OF THE PEN

We are now in a position to connect the constructions of propositions 3 and 4:



Suppose we are given two caustics $\bar{\varphi}$ and $\bar{\psi}$ as in figure , each consisting of two arcs with a common asymptotic ray i_φ and i_ψ respectively. Localizing the mirror as in the figure (choosing M_0 and finding the rays of $\bar{\varphi}$ and $\bar{\psi}$ that intersect at M_0) gives us a real and a virtual arc of each caustic ($\bar{\varphi}_R$ and $\bar{\psi}_R$ real, $\bar{\varphi}_V$ and $\bar{\psi}_V$ virtual).

Since the two rays intersecting at M_0 have a real-virtual relationship to the desired mirror, we start the tracing of the curve by applying the locally hyperbolic construction, pulling the string towards A' and B. In this way the pen M will move through the points $M_1, M_2, \dots$ in a locally hyperbolic fashion, until it reaches the point M_φ, that corresponds to the asymptotic ray i_φ. At this point the motion becomes locally parabolic, since i_φ is tangent to $\bar{\varphi}_V$ at infinity (at A'). But i_φ is also tangent to $\bar{\varphi}_R$ at infinity (at A). Therefore, when it is at M_φ, THE STRING IS BEING PULLED TOWARDS A' AND LET OUT FROM A AT THE SAME TIME.

Hence, as it moves past M_φ, the string looses contact with $\bar{\varphi}_V$ and picks up contact with $\bar{\varphi}_R$, AND THE MOTION OF THE PEN CHANGES FROM LOCALLY HYPERBOLIC TO LOCALLY ELLIPTIC, the string now being let out from A but still pulled in at B.

The locally elliptic motion will continue through $M_3, M_4, \ldots$ until the pen reaches $M_{\mathcal{F}}$, corresponding to the other asymptotic ray $i_{\mathcal{F}}$. Here the motion again becomes locally parabolic, since the string is now touching $\overline{\mathcal{F}}$ at infinity, and at the same time the motion changes from locally elliptic to locally hyperbolic, since the string at $M_{\mathcal{F}}$ is being simultaneously pulled towards B and let out from B'. Thus loosing contact with $\overline{\mathcal{F}}_R$ and picking up contact with $\overline{\mathcal{F}}_V$, the motion continues in a locally hyperbolic way through $M_6, M_7, \ldots$

Granted the validity of propositions 3 and 4, we can describe the constructed curve by the following:

<u>Proposition 5</u>:

THE TRACE CURVE IN THE ABOVE CONSTRUCTION IS THE MIRROR THAT PASSES THROUGH M_0 AND REFLECTS $\overline{\Phi}$ INTO $\overline{\mathcal{F}}$

We now have a general construction procedure for mirrors, given the in- and out caustics $\overline{\Phi}$ and $\overline{\mathcal{F}}$. Observe that if we interpret a singularity switch as a reversal of all ray directions (which, by the reciprocity principle doesn't change the mirror), then THE MOTION OF THE STRING IS IDENTICAL TO THE PHYSICAL MOTION OF THE CORRESPONDING LIGHTRAYS.

This observation is the key to generalizing the construction of proposition 5 to incorporate lenses as well as mirrors.

Suppose that the caustics $\overline{\Phi}$ and $\overline{\mathcal{F}}$ are located in two different media A and B, each homogeneous and isotropic with index of refraction n_A and n_B respectively. This means that the corresponding lightvelocities are $v_A = \frac{c}{n_A}$ and $v_B = \frac{c}{n_B}$, where c is the velocity of light in vaccuum.

Let us modify the construction of proposition 5 in the following way:

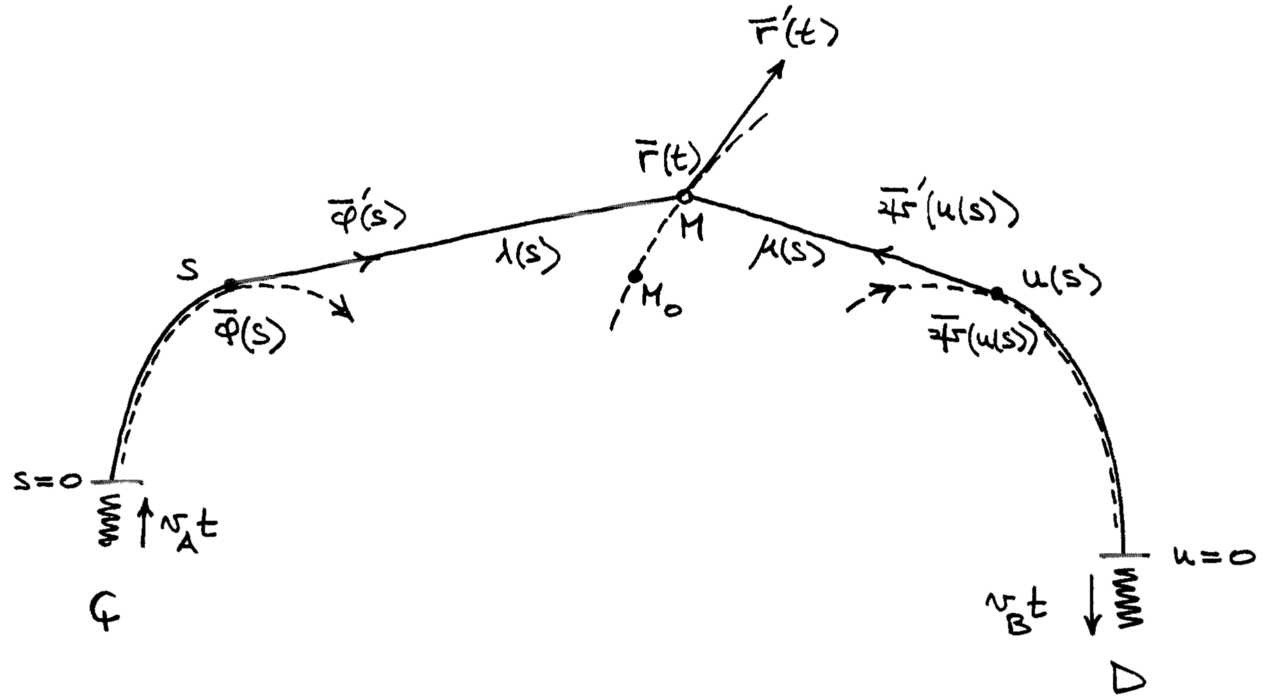
Keeping everything else as before, we change the let-out and pull-in speed of the string, so that the part of it that touches $\bar{\varphi}$ moves with the speed v_A, and the part that touches $\bar{\psi}$ moves with the speed v_B.

If t denotes time, what kind of curve will now be traced out by the pen? Regarding everything as functions of t, I claim that:

<u>Proposition 6</u>

IF THE ANGLE BETWEEN $\bar{\varphi}'(s)$ AND $\bar{r}'(t)$ LIES BETWEEN THE LIMITS OF "PHYSICAL INCIDENCE" THEN THE TRACE-CURVE $\bar{r}(t)$ WILL BE THE "LENSE-CURVE" PASSING THROUGH M_0 THAT REFRACTS $\bar{\varphi}$ INTO $\bar{\psi}$

The time for a proof has finally come.
We will assume that $\bar{\varphi}$ and $\bar{\psi}$ are both real and that they are located as in figure.
The other cases contain the same mathematics (with obvious modifications), and they can be easily worked out by the reader.



Since the string is being let out from C with speed v_A and pulled in at D with speed v_B, we have (referring back to page for notation)

(1) $\bar{r}(t) = \bar{\varphi}(s) + \lambda(s)\,\bar{\varphi}'(s)$; $s = s(t)$

(2) $\bar{r}(t) = \bar{\psi}(u(s)) + \mu(s)\,\bar{\psi}'(u(s))$

(3) $s + \lambda(s) = L_0 + v_A t$

(4) $u(s) + \mu(s) = L_1 - v_B t$

Here L_0 and L_1 denote the length of the φ- and ψ parts of the string corresponding to the startingpoint M_0. Differentiating (1) – (4) with respect to t gives us:

(1)' $\bar{r}'(t) = \bar{\varphi}'(s)s'(t) + \lambda'(s)s'(t)\bar{\varphi}'(s) + \lambda(s)\bar{\varphi}''(s)s'(t)$

(2)' $\bar{r}'(t) = \bar{\psi}'(u(s))u'(s)s'(t) + \mu'(s)s'(t)\bar{\psi}'(u(s)) + \mu(s)\bar{\psi}''(u(s))u'(s)s'(t)$

(3)' $s'(t) + \lambda'(s)s'(t) = v_A$

(4)' $u'(s)s'(t) + \mu'(s)s'(t) = -v_B$

Since both s and u are arclength parameters, we have

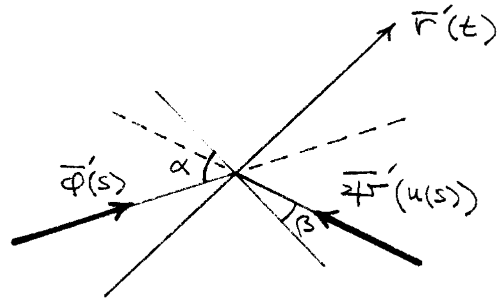
$$|\bar{\varphi}'(s)| = |\bar{\psi}'(u(s))| = 1 \qquad \text{and}$$

$$\langle \bar{\varphi}'(s), \bar{\varphi}''(s)\rangle = \langle \bar{\psi}'(u(s)), \bar{\psi}''(u(s))\rangle = 0$$

Utilizing these facts, we form the inner product of (1)′ with $\bar{\varphi}'(s)$ and (2)′ with $\bar{\psi}'(u(s))$. Taking account of (3)′ and (4)′ this gives us

$$(5) \begin{cases} \langle \bar{\varphi}'(s), \bar{r}'(t)\rangle = s'(t) + \lambda'(s)s'(t) = v_A \\ \langle \bar{\psi}'(u(s)), \bar{r}'(t)\rangle = u'(s)s'(t) + \mu'(s)s'(t) = -v_B \end{cases}$$

But (5) is nothing but another way of stating the familiar Snell's law of refraction, as shown by the following computation:



$$(6) \begin{cases} \langle \bar{\varphi}'(s), \bar{r}'(t)\rangle = |\bar{r}'(t)|\cos(\frac{\pi}{2}-\alpha) = |\bar{r}'(t)|\sin\alpha \\ \langle \bar{\psi}'(u(s)), \bar{r}'(t)\rangle = |\bar{r}'(t)|\cos(\frac{\pi}{2}+\beta) = -|\bar{r}'(t)|\sin\beta \end{cases}$$

$$(5) \text{ and } (6) \implies \frac{v_A}{v_B} = \frac{\sin\alpha}{\sin\beta} \implies n_A \sin\alpha = n_B \sin\beta$$

Observe that if $v_A = v_B$ we get the locally elliptic reflection (proposition 4), and if $v_A = -v_B$ we get the locally hyperbolic reflexion (proposition 3).