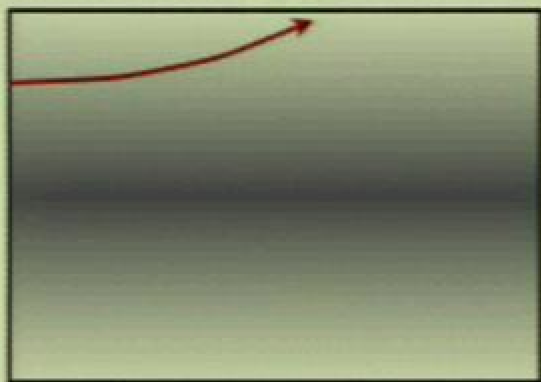
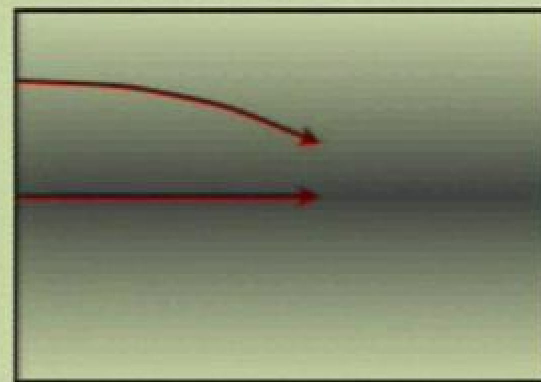
Why are the rays attracted towards the higher index?

Why is the ray trajectory not bending like this ... ... but rather like this?

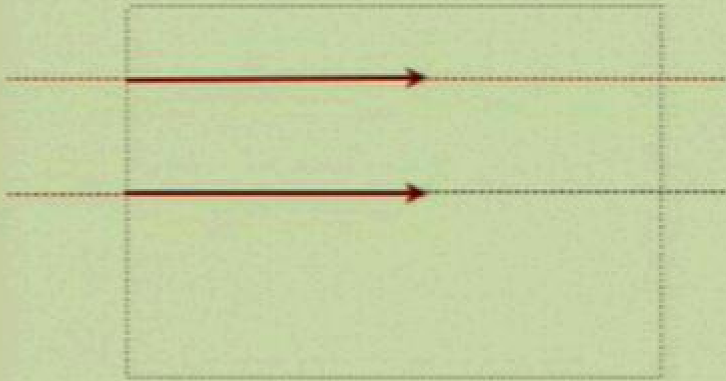




(after all, the ray is supposed to minimize its optical path: in the 1st option it does look like the ray is going through a lower index portion of the medium, and therefore shorter optical path)

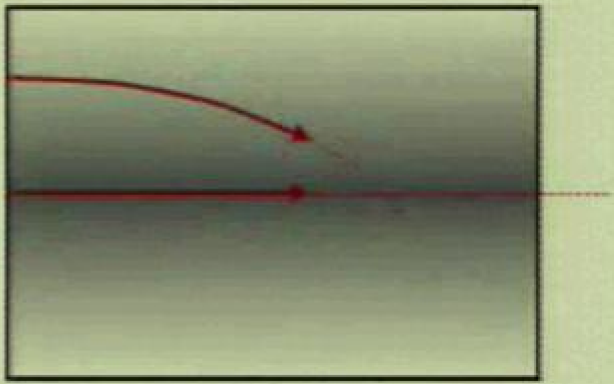
To properly contemplate an answer, we must consider not just a single ray, but a *pair of rays following different paths yet sharing a common beginning and a common end*

A pair of rays with equal paths



In free space, these two rays share a common beginning (infinity to the left) and common end (infinity to the right)

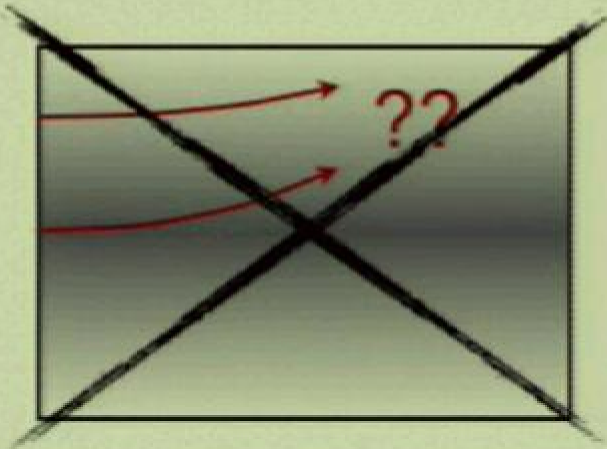
Clearly, the straight path is minimal for both, so Fermat's principle is satisfied. Moreover, clearly the two paths are equal.

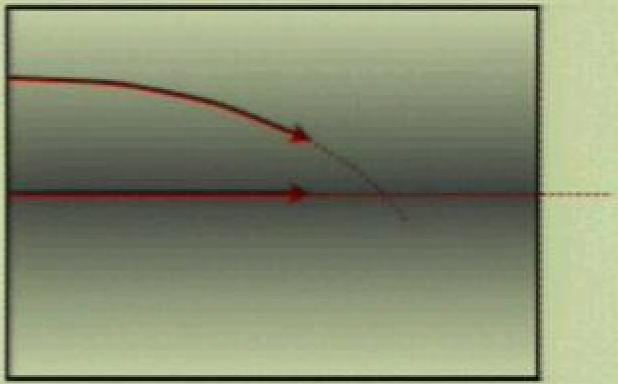


In the GRIN, the paths must be again equalized.

However, the on-axis ray undergoes a longer optical path due to the higher index. Therefore, the off-axis ray must bend its path to match it. As result, the two rays now meet at a finite distance (but still have traveled equal paths to get there!)

But why does the on-axis ray have to go straight into the higher index (and longer optical path)?



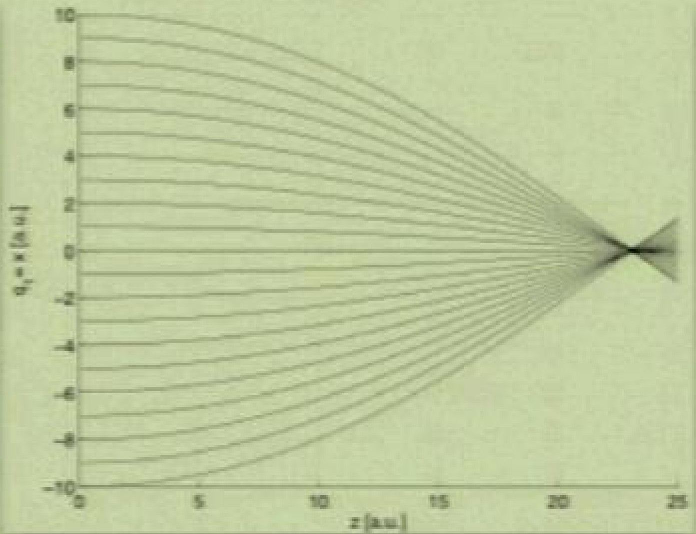


For example, what's wrong with this scenario: both rays are escaping the "undesirable" high index area by bending their paths away from it.

Answer: some thought will convince you that if that were the case, should the rays ever meet again (and they will, eventually, at infinity at the latest) it is impossible for them to have equal optical paths because the ray that spends most of its trajectory in the low index area also has the shortest geometrical length.

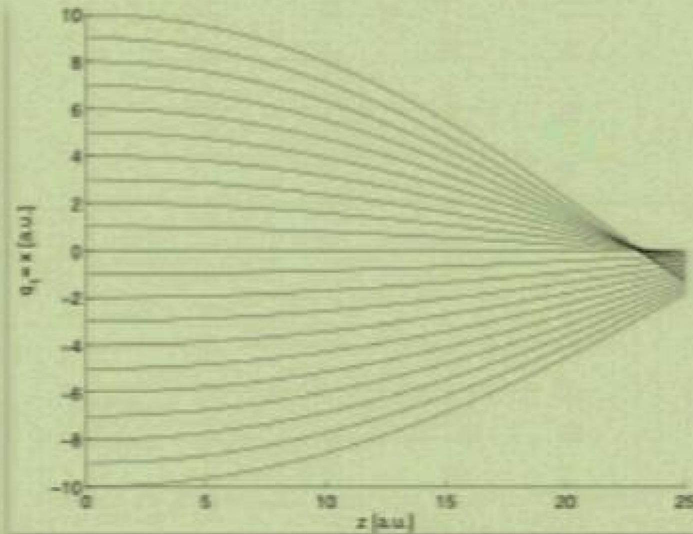
Therefore, the only possibility to meet the minimum path requirement is the one that predicts focusing of the ray from the low index area toward the high index area.

One final comment: the on-axis ray does *not* stay straight because of symmetry



Symmetric index profile

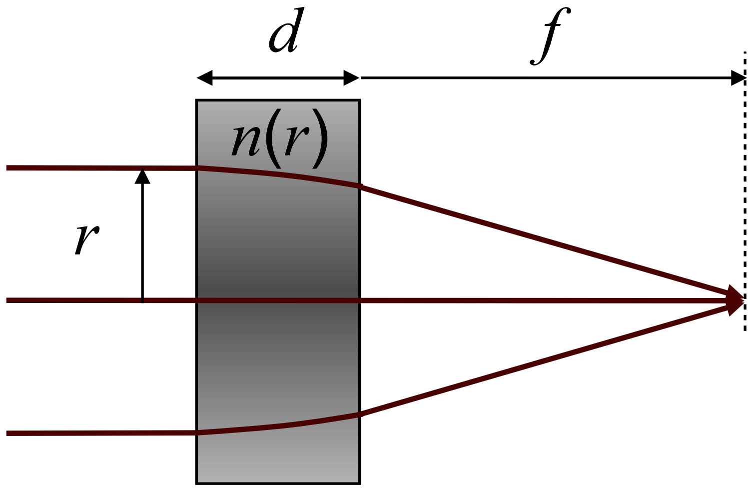
$$n(x) = 1.5 - 3.3 \times 10^{-3}\, x^2$$



Asymmetric index profile

$$n(x) = \begin{cases} 1.5 - 3.3 \times 10^{-3}\, x^2, & \text{if } x > 0; \\ 1.5 - 2.2 \times 10^{-3}\, x^2, & \text{otherwise.} \end{cases}$$

Paraxial focusing by a thin quadratic GRIN lens



Consider a ray from infinity entering the GRIN at elevation r. If focusing is to be achieved, this ray must meet the on–axis ray at a distance f from the exit face; therefore, their optical paths must be equal according to Fermat's principle.

For the on–axis ray,

$$\mathrm{OPL}\,(r=0) = n_{\max} d + f.$$

For the ray at elevation r,

$$\mathrm{OPL}\,(r) \approx n_{\max}\left(1 - \frac{\alpha r^2}{2}\right) d + \sqrt{r^2 + f^2},$$

where we have neglected the small elevation decline due to the bending of the ray inside the GRIN.

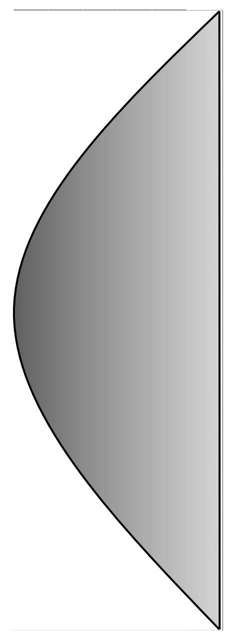
Applying Fermat's principle in the paraxial approximation,

$$\begin{aligned}
n_{\max} d + f &\approx n_{\max}\left(1 - \frac{\alpha r^2}{2}\right) d + \sqrt{r^2 + f^2} \Rightarrow \\
f + \frac{n_{\max}\alpha d}{2} r^2 &\approx \sqrt{r^2 + f^2} \approx f\left(1 + \frac{r^2}{2f^2}\right) \Rightarrow \\
\frac{n_{\max}\alpha d}{2} r^2 &\approx \frac{r^2}{2f} \Rightarrow \\
f &\approx \frac{1}{n_{\max}\alpha d}.
\end{aligned}$$

Gradient Index (GRIN) optics: axial

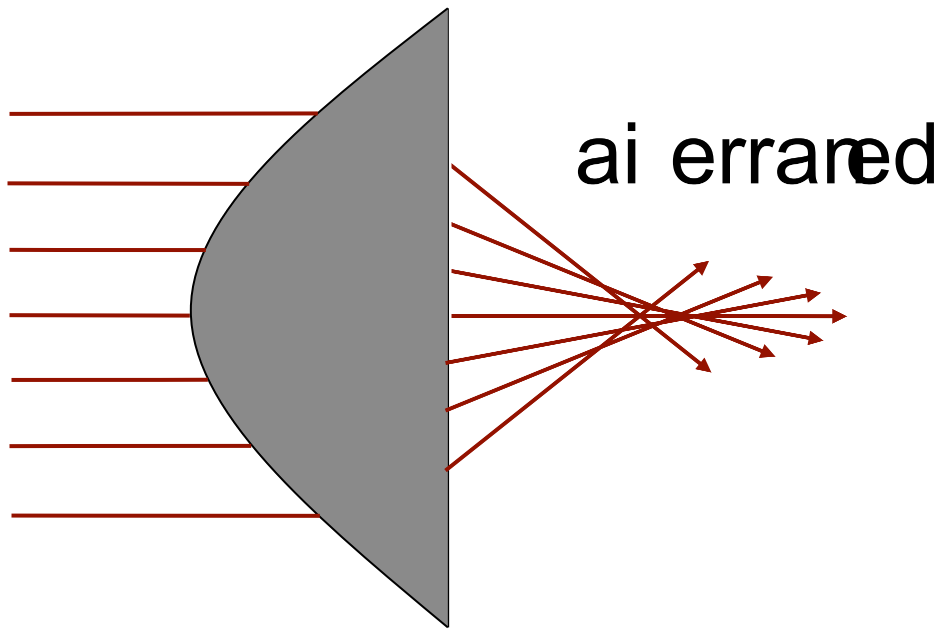
• ial inde profile:

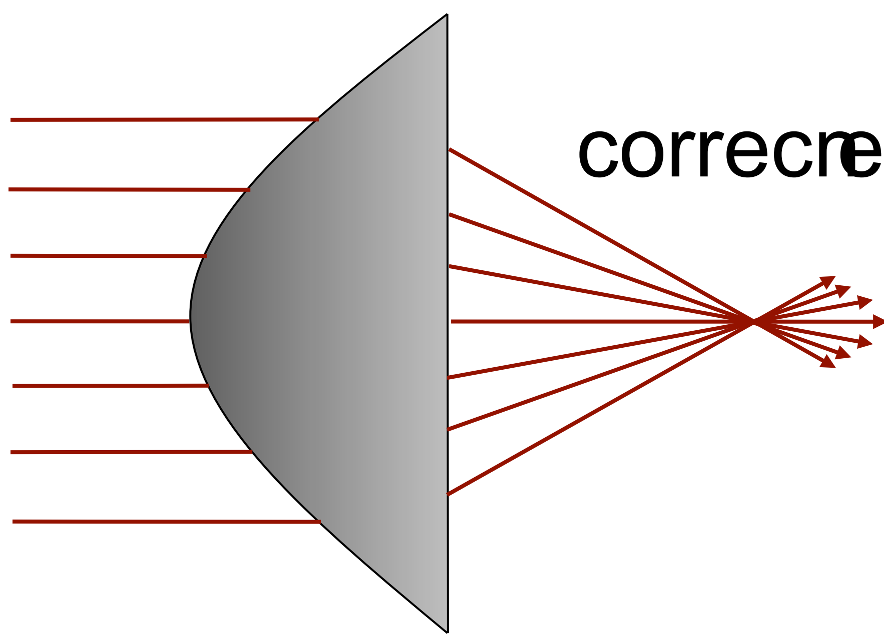
fai rication i r f elding & grinding



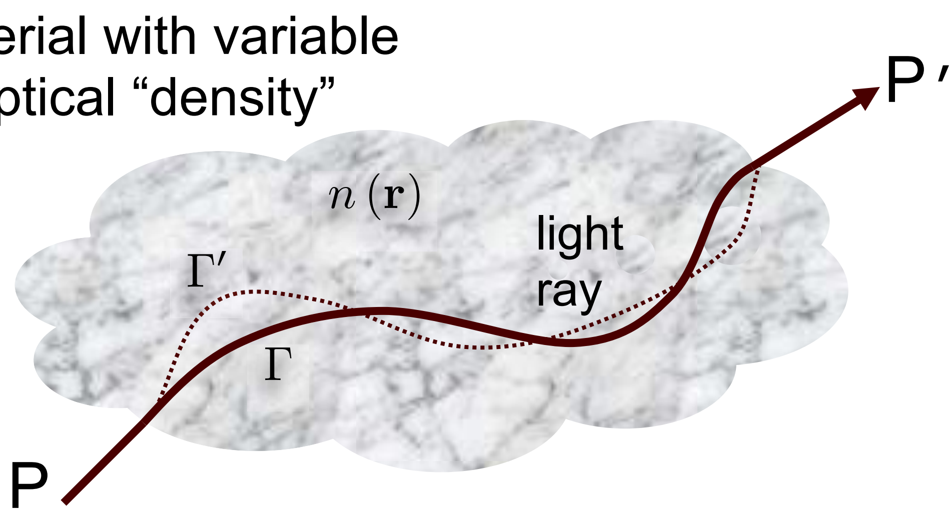
- Stack
- Meld
- Grind & polish to a sphere
- Result:
  Spherical refractive surface with axial index profile $n(z)$

Correction of spherical aberration by axial GRIN lenses





Generalized GRIN: what is the ray path through arbitrary *n*(*r*)?



"optical path length"

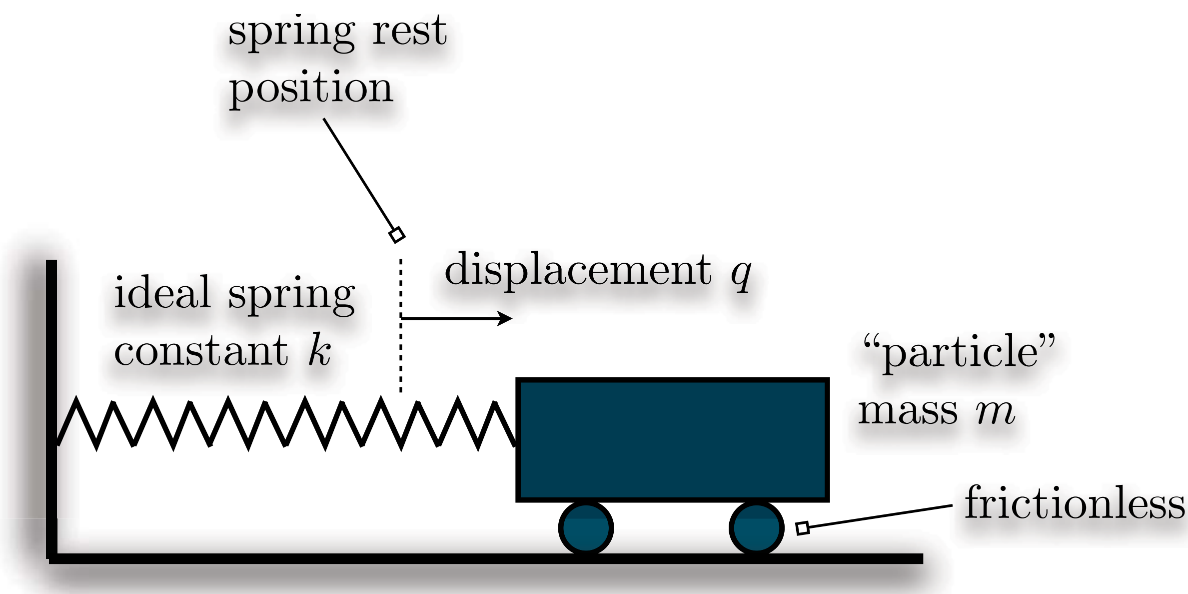
$$\int_\Gamma n(\mathbf{r})\,\mathrm{d}l$$

Fermat's principle:

The path Γ that the ray follows is such that the value of the path integral of refractive index $n(\mathbf{r})$ along Γ is smaller than all other possible paths Γ'.

Let's take a break from optics ...

Mechanical oscillator



Potential energy: $V = \frac{1}{2}kq^2.$

Kinetic energy: $T = \frac{1}{2}m\dot{q}^2 = \frac{1}{2}\frac{p^2}{m},$

where $p \equiv m\dot{q}$ is the momentum.

Since there is no dissipation,
the total energy

$$H = T + V$$

must be conserved.

Introduction to the Hamiltonian formulation of dynamics

The Hamiltonian formulation is a set of differential equations describing the trajectories of particles that are subject to a potential (force.) The trajectory is described in terms of the particle position $\mathbf{q}(t)$ and momentum $\mathbf{p}(t)$. The Hamiltonian is the total energy, *i.e.* the sum of kinetic and potential energies, and it is conserved if there is no dissipation in the system. For example, for a harmonic oscillator the Hamiltonian is expressed as

$$H(\mathbf{q},\mathbf{p}) = \frac{\mathbf{p}^2}{2m} + \frac{k\mathbf{q}^2}{2}. \tag{1}$$

The first term is the kinetic energy for a particle of mass m, and the second term is the potential energy for linear spring constant k.

The Hamiltonian equations in general are

$$\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t} = \frac{\partial H}{\partial \mathbf{p}}, \tag{2}$$

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = -\frac{\partial H}{\partial \mathbf{q}}. \tag{3}$$

The expressions on the right–hand side are the gradients of the Hamiltonian with respect to the vectors $\mathbf{p}$ and $\mathbf{q}$, respectively.

Let us consider the simplest case of a one–dimensional harmonic oscillator. In this case the position and momentum are scalars q, p. The Hamiltonian equations become

$$\left.\begin{aligned}\frac{\mathrm{d}q}{\mathrm{d}t} &= \frac{p}{m}\\ \frac{\mathrm{d}p}{\mathrm{d}t} &= -kq.\end{aligned}\right\} \Rightarrow \frac{\mathrm{d}^2 q}{\mathrm{d}t^2} = \frac{1}{m}\frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{k}{m}q \Rightarrow \frac{\mathrm{d}^2 q}{\mathrm{d}t^2} + \frac{k}{m}q = 0. \tag{4}$$

We have arrived at the familiar 2$^{\text{nd}}$–order harmonic differential equation. For example, assuming a particle that is initially at position $q(t=0) = q_0$ and at rest, $p(t=0) = 0$, the solution to the Hamiltonian equations is

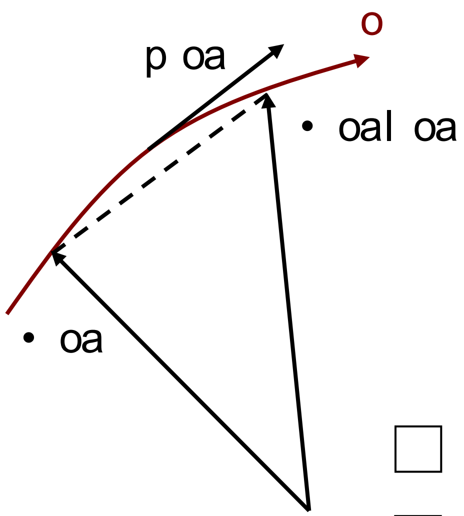
$$q(t) = q_0 \cos\left(\sqrt{\frac{k}{m}}t\right), \tag{5}$$

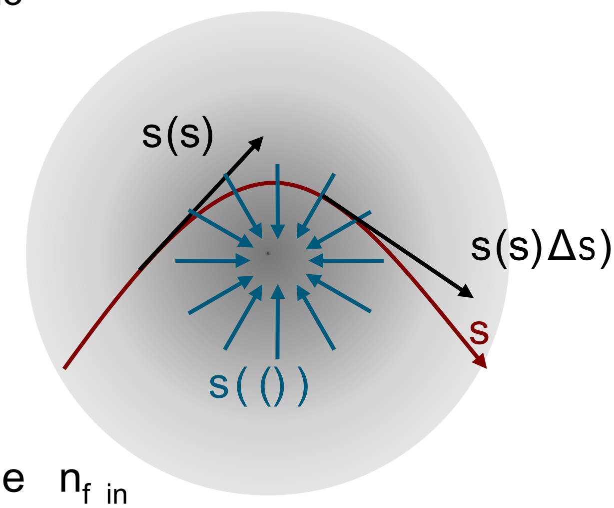
$$p(t) = -q_0\sqrt{km}\sin\left(\sqrt{\frac{k}{m}}t\right). \tag{6}$$

The solution set $\left\{q(t), p(t)\right\}$ is the trajectory of the particle. The motion represented by the trajectory that we found is clearly a harmonic oscillation.

Hamiltonian Optics postulates

o: paraf ererir arion of ri e rar rajecrorr
• oa pooirion f ecror for ri e rar rajecrorr ano;
p oa rangenn f ecror ro ri e rar rajecrorr ano





• low inde n$_f$ in

• i igi inde n$_f$ a

• eof errical poorulare:

• ar o are conrinuouo and piecewioe differenriai le

$$\Delta\mathbf{q}(s) \approx \frac{\mathbf{p}(s')}{|\mathbf{p}(s')|}\Delta s$$

$$\Rightarrow \frac{\mathrm{d}\mathbf{q}}{\mathrm{d}s} = \frac{\mathbf{p}(s)}{|\mathbf{p}(s)|}$$

• r naf ical poorulare:

• of ennf ci angeo along rajecrorr arc lengri in proporrion ro ri e local refracrif e inde gradienn

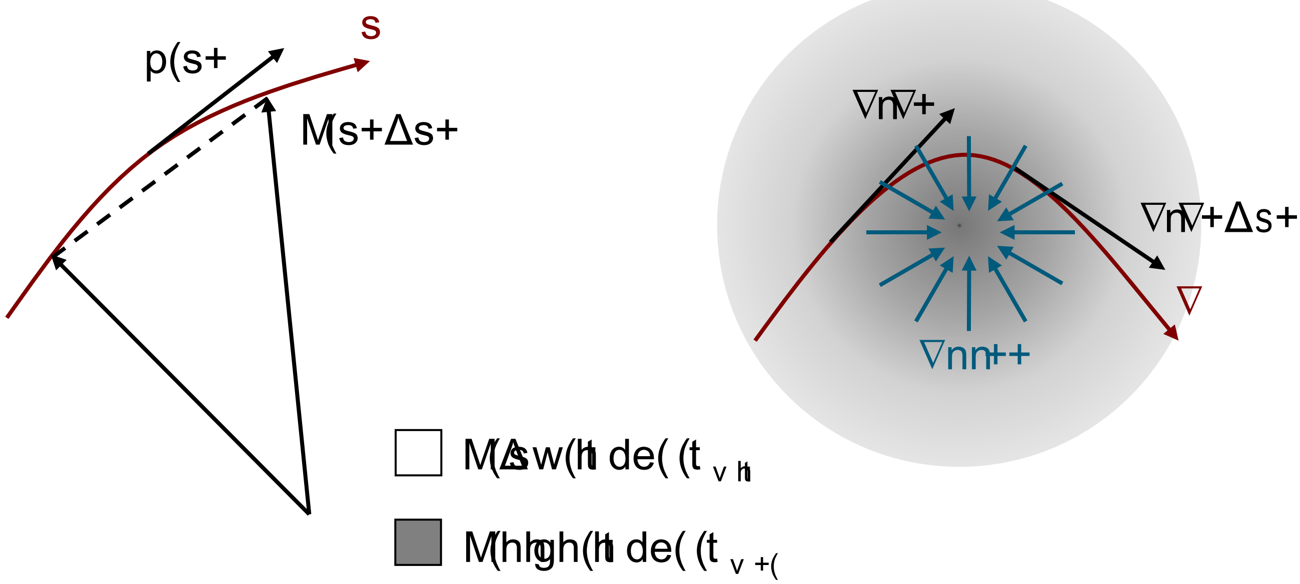
$$\Delta\mathbf{p}(s) \approx \nabla n\Big(\mathbf{q}(s')\Big)\Delta s$$

$$\Rightarrow \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}s} = \nabla n\Big(\mathbf{q}(s)\Big)$$

These are the “equations of motion,”
i.e. they yield the ray trajectories.

The ray Hamiltonian

s:(p+z+v etezhz+thst (sv(the(z+z(tz+jectszz
M(s+:(psshthst (vectsz(vsz(the(z+z(tz+jectszz(+t(s;(
p(s+:(t+t get t(vectsz(ts(the(z+z(tz+jectszz(+t(s



The choice $H = |\mathbf{p}| - n(\mathbf{q})$ yields

$$\nabla_{\mathbf{q}} H \equiv \frac{\partial H}{\partial \mathbf{q}} = -\nabla_{\mathbf{q}} n(\mathbf{q}) \qquad \nabla_{\mathbf{p}} H \equiv \frac{\partial H}{\partial \mathbf{p}} = \frac{\mathbf{p}(s)}{|\mathbf{p}(s)|}$$

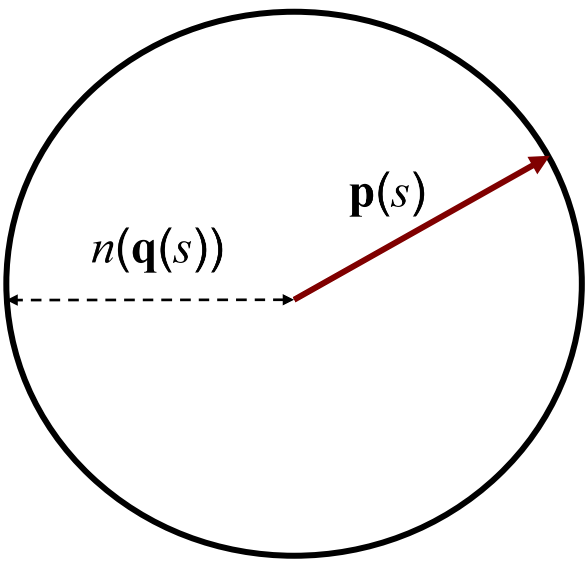
Therefore, the equations of motion become

$$\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}s} = \frac{\partial H}{\partial \mathbf{p}} \qquad \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}s} = -\frac{\partial H}{\partial \mathbf{q}}$$

Since the ray trajectory satisfies a set of Hamiltonian equations on the quantity *H*, it follows that *H* is conserved.

The actual value of *H*=const. is <u>arbitrary</u>.

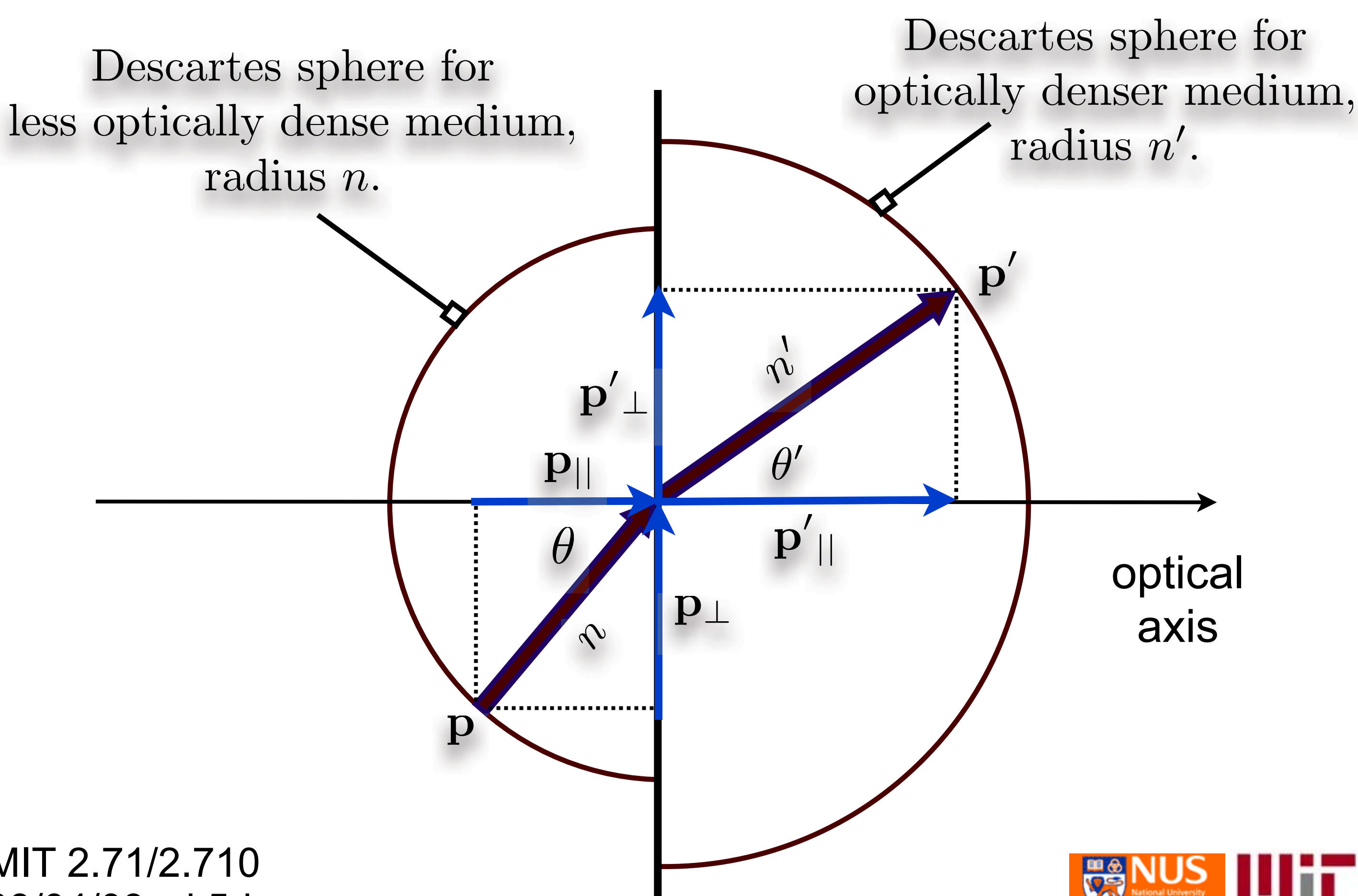
The ray Hamiltonian and the Descartes sphere



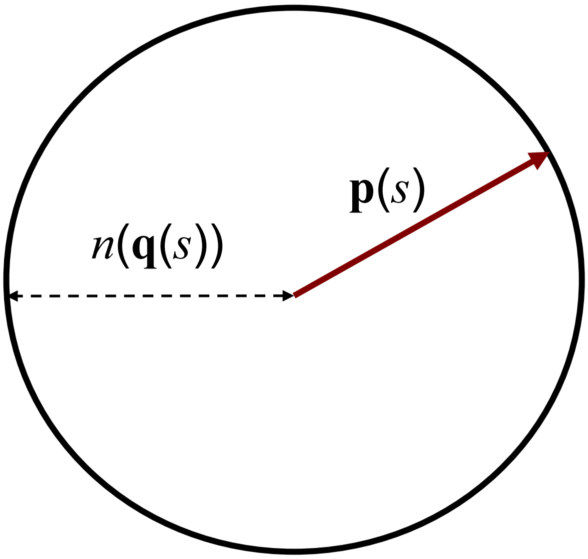
$H = |\mathbf{p}| - n(\mathbf{q}) = 0$

The ray momentum $\mathbf{p}$ is constrained to lie on a sphere of radius n at any ray location $\mathbf{q}$ along the trajectory s

<u>Application</u>:
Snell's law of refraction



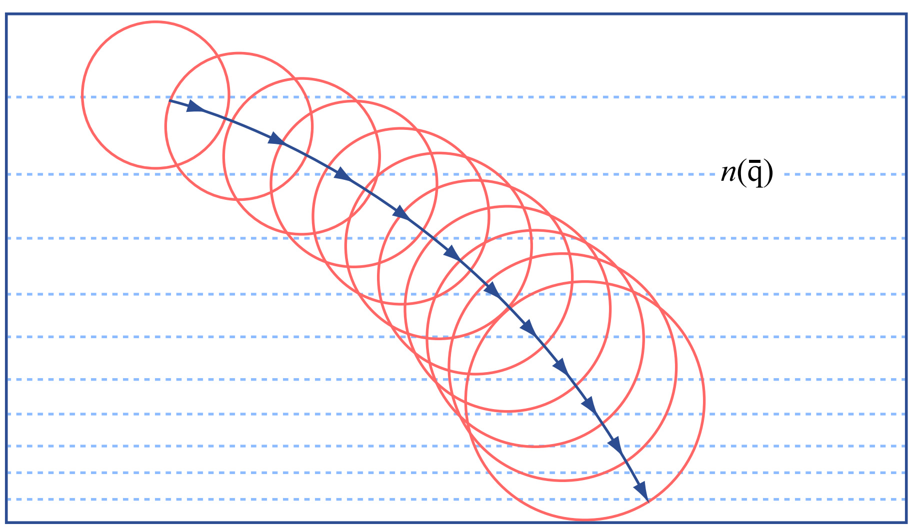
The ray Hamiltonian and the Descartes sphere



$$H = |\mathbf{p}| - n(\mathbf{q}) = 0$$

The ray momentum $\mathbf{p}$ is constrained to lie on a sphere of radius n at any ray location $\mathbf{q}$ along the trajectory s

<u>Application</u>:
propagation in a GRIN medium



The Descartes sphere radius is proportional to n($\mathbf{q}$); as the rays propagate, the lateral momentum is preserved by gradually changing the ray orientation to match the Descartes spheres.

Figure by MIT OpenCourseWare. Adapted from Fig. 1.5 in Wolf, Kurt B. Geometric Optics in Phase Space. New York, NY: Springer, 2004.

Hamiltonian analogies: optics *vs* mechanics

Hamiltonian of
mami animal y tam

$$H_{\mathrm{m}} = \frac{|\mathbf{p}|^2}{2m} + V(\mathbf{q}) \qquad E = H_{\mathrm{m}} \quad \text{(conserved)}$$

Energy

$$\text{Momentum} \quad \mathbf{p} = m\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t} \Rightarrow \frac{\mathrm{d}s}{\mathrm{d}t} = \frac{|\mathbf{p}|}{m} \quad \text{Velocity}$$

Mechanical Hamiltonian equations

$$\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}s} = \frac{\mathbf{p}}{|\mathbf{p}|} \qquad \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}s} = -\frac{m}{|\mathbf{p}|}\frac{\partial V(\mathbf{q})}{\partial \mathbf{q}}$$

Optical Hamiltonian equations

$$\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}s} = \frac{\mathbf{p}}{|\mathbf{p}|} \qquad \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}s} = \frac{\partial n(\mathbf{q})}{\partial \mathbf{q}}$$

$$\text{Analogous if: } \frac{\partial V(\mathbf{q})}{\partial \mathbf{q}} \quad \leftrightarrow \quad -\frac{|\mathbf{p}|}{m}\frac{\partial n(\mathbf{q})}{\partial \mathbf{q}} = -\frac{n(\mathbf{q})}{m}\frac{\partial n(\mathbf{q})}{\partial \mathbf{q}} =$$

$$= -\frac{1}{2m}\frac{\partial n^2(\mathbf{q})}{\partial \mathbf{q}}.$$

$$\Rightarrow V(\mathbf{q}) \quad \leftrightarrow \quad -\frac{n^2(\mathbf{q})}{2m} + \text{const.}$$

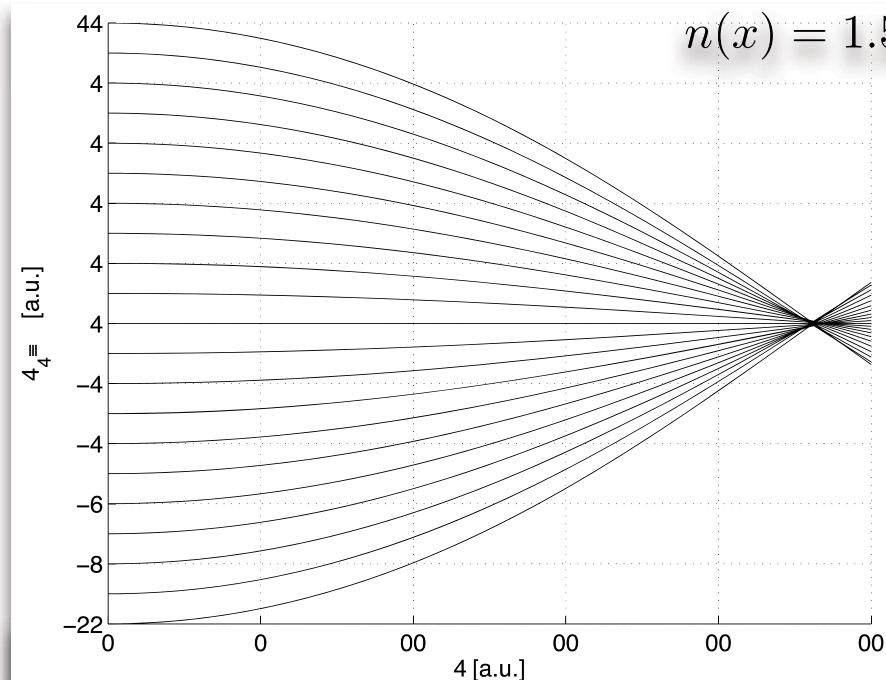
$$\text{Choose const.} = -\frac{E+1}{2m} \Rightarrow 2m\left[E - V(\mathbf{q})\right] \quad \leftrightarrow \quad n^2(\mathbf{q}) - 1$$

$$\text{and} \qquad E_{\text{kinetic}} \propto E - V(\mathbf{q}) > 0 \quad \leftrightarrow \quad n(\mathbf{q}) > 1.$$

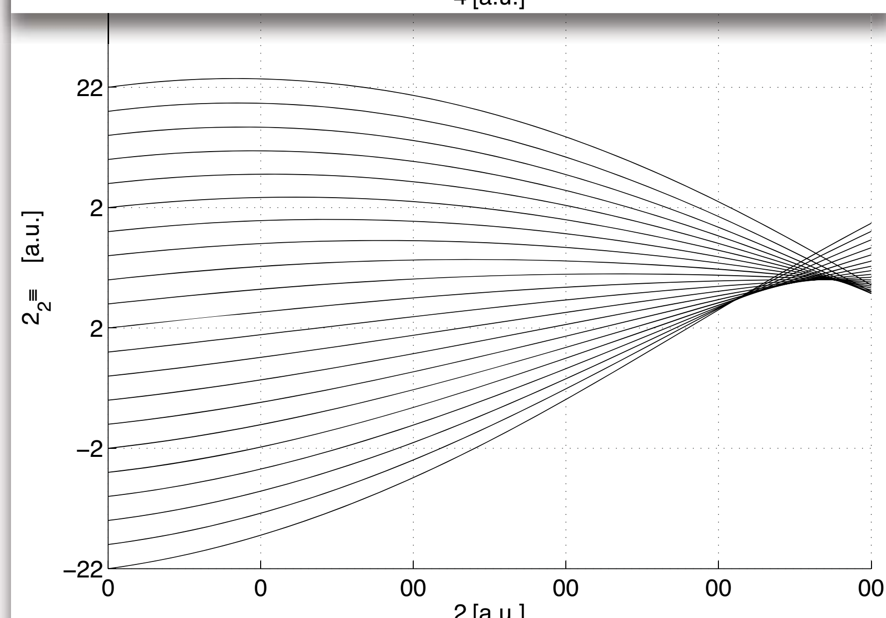
physically allowable kinetic energy — physically allowable refractive index

Example: Hamiltonian ray tracing of quadratic GRIN

$$n(x) = 1.5 - 0.0033\,x^2$$



on-axis incidence



off-axis incidence, $p_x(0)=0.2$

Further reading:

- M. Born and E. Wolf, *Principles of Optics*, Cambrige University Press, 7th edition, sections 4.1-4.2
- K. B. Wolf, *Geometrical Optics on Phase Space,* Springer, chapters 1, 2
- K. Tian, *Three-dimensional (3D) optical information processing,* PhD dissertation, MIT 2006.

So far

- Geometrical Optics
  - Light propagation is described in terms of rays and their refraction / reflection
  - Image formation is described in terms of point sources and point images
- Region of validity of geometrical optics: features of interest should be much bigger than the wavelength λ
  - Problem: geometrical point objects/images are infinitesimally small, definitely *smaller* than λ!!!
  - So light focusing at a single point is an artifact of the geometric approximations
  - Moreover, especially in microscopy, we are interested in resolving object features at distances comparable to λ
  - To understand light behavior at scales ~ λ we need to graduate from geometric to the **wave description** of light propagation, i.e. **Wave Optics**.

What is a wave?

Water waves: Vouliagmeni Bay, Greece (photo by G. B.)

Image removed due to copyright restrictions.
Please see:
http://www.reuters.com/news/pictures/searchpopup?picId=8505659

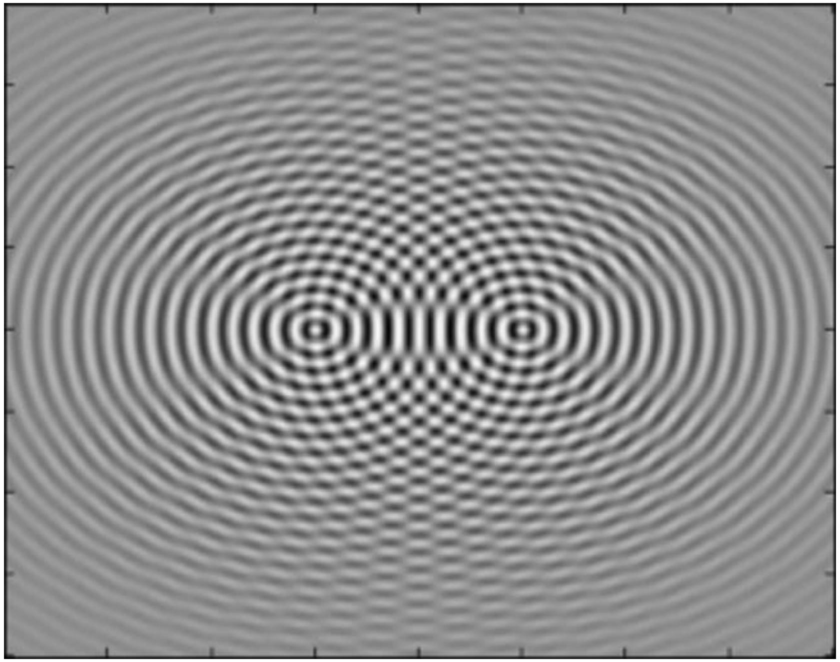
Photo of a Yagi-Uda antenna removed due to copyright restrictions.

Radio waves: Yagi-Uda television antenna

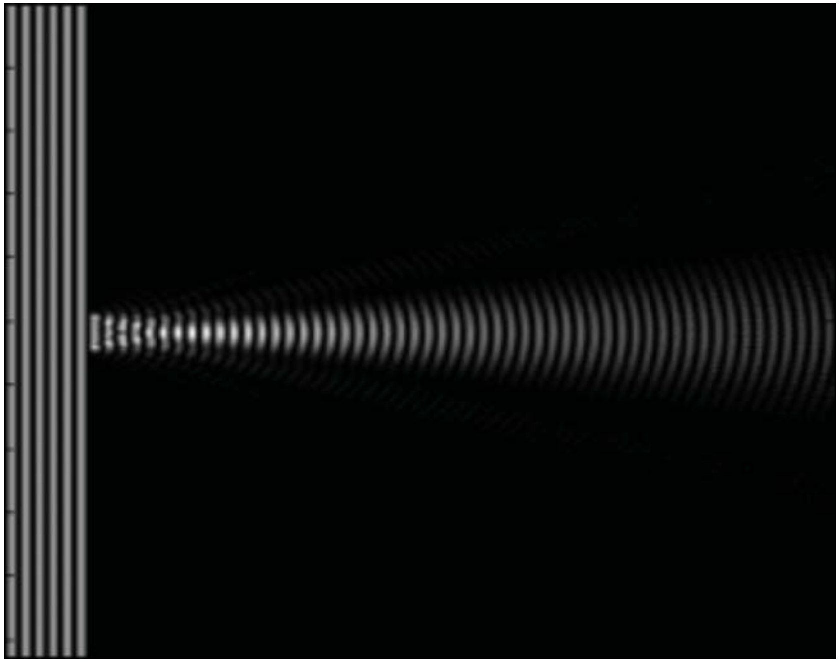
Ground surface waves: Chocolate Hills, Philippines
http://www.bohol.ph/picture48.html

What is a wave?

- A wave is a traveling disturbance
- Evidence of wave behavior:

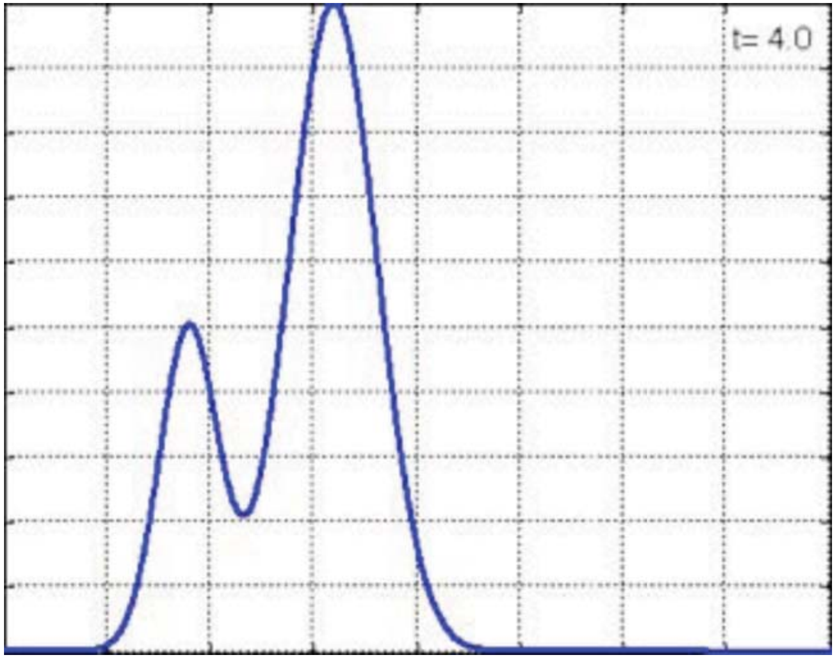


Interference

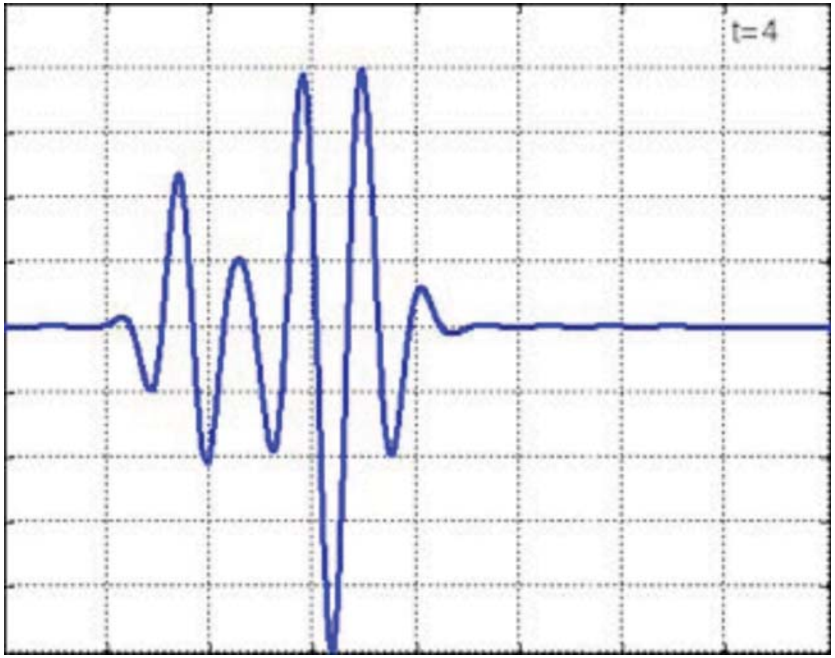


Diffraction

Traveling waves

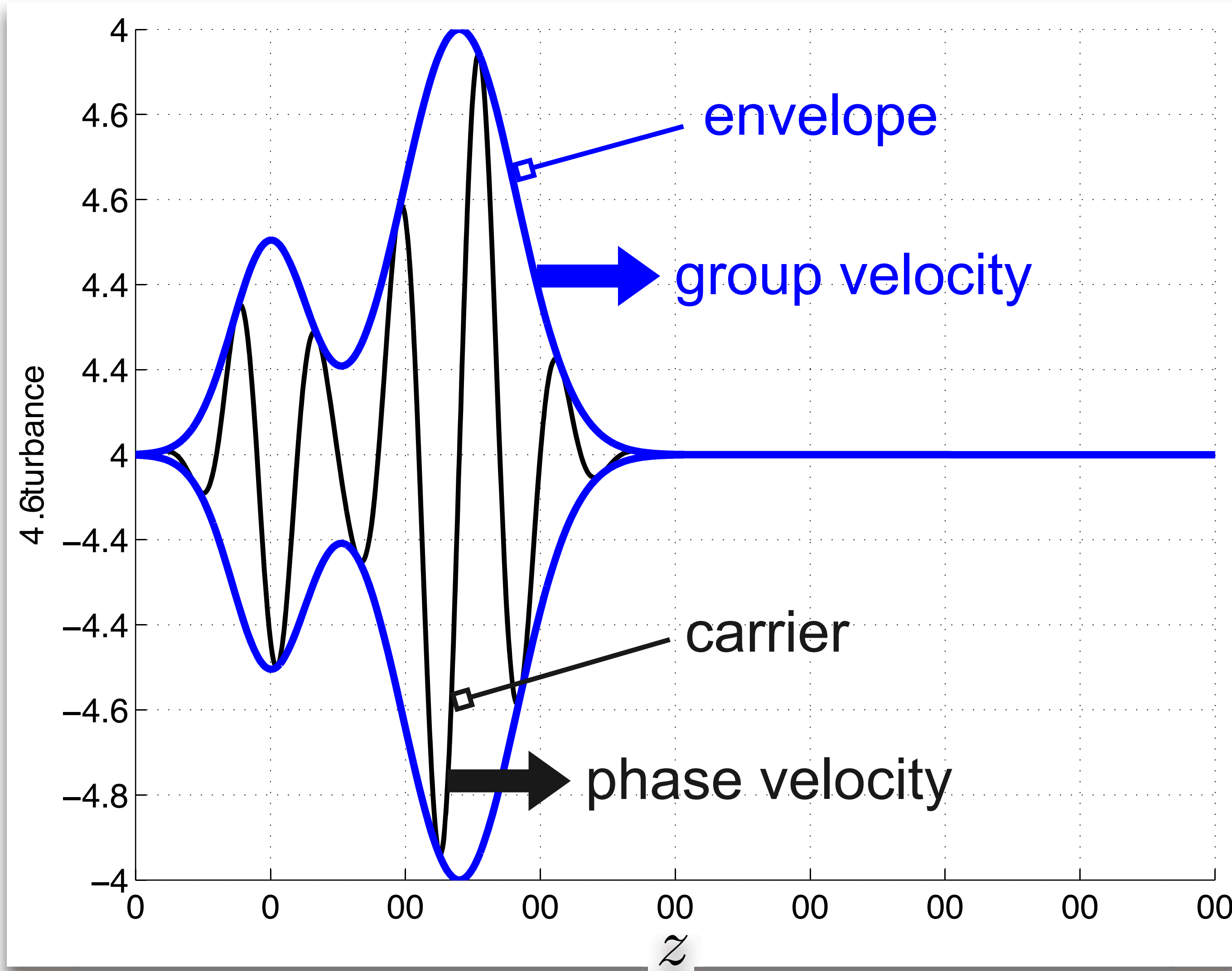


Traveling disturbance
(envelope)

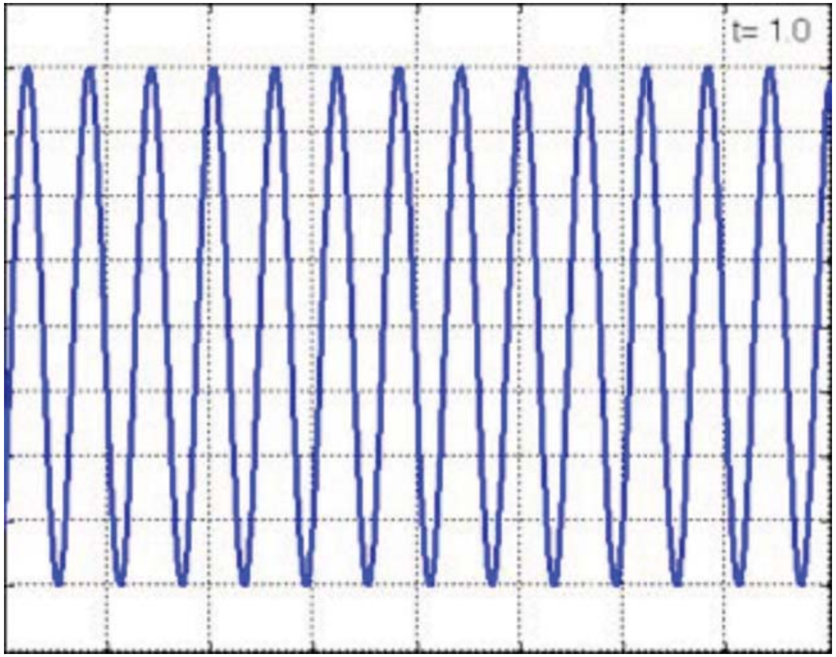


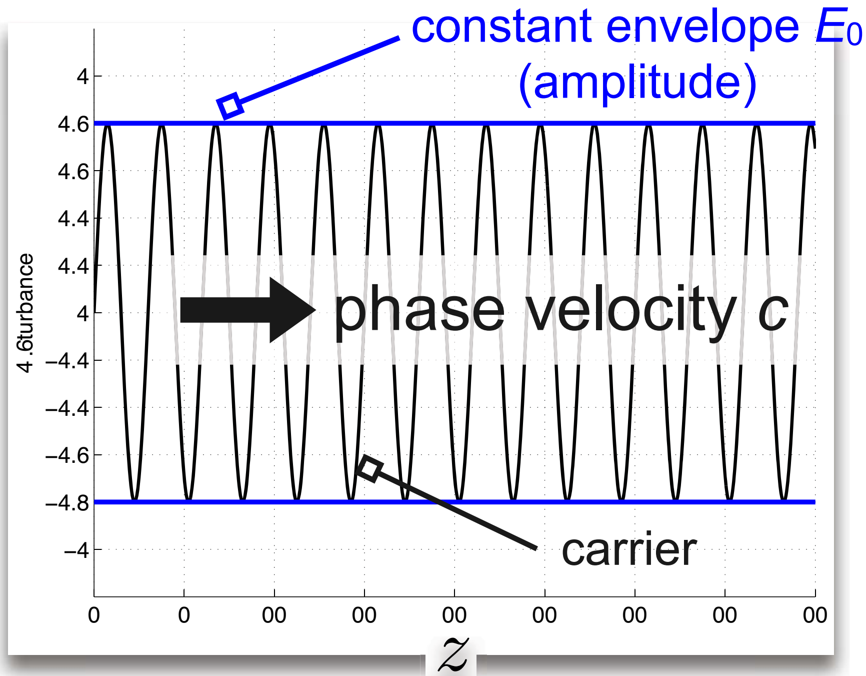
Traveling disturbance
with sinusoidal modulation
(envelope on a carrier)

Traveling wave terminology



Harmonic wave





$$\text{Disturbance } E(z,t) = E_0 \cos\left(\frac{2\pi}{\lambda}(z - ct) + \phi\right)$$

E_0 = amplitude

λ = wavelength (spatial period)

ν = frequency (temporal period)

ϕ = phase delay

One period $T \equiv 1/\nu$ later, the wave repeats; therefore,

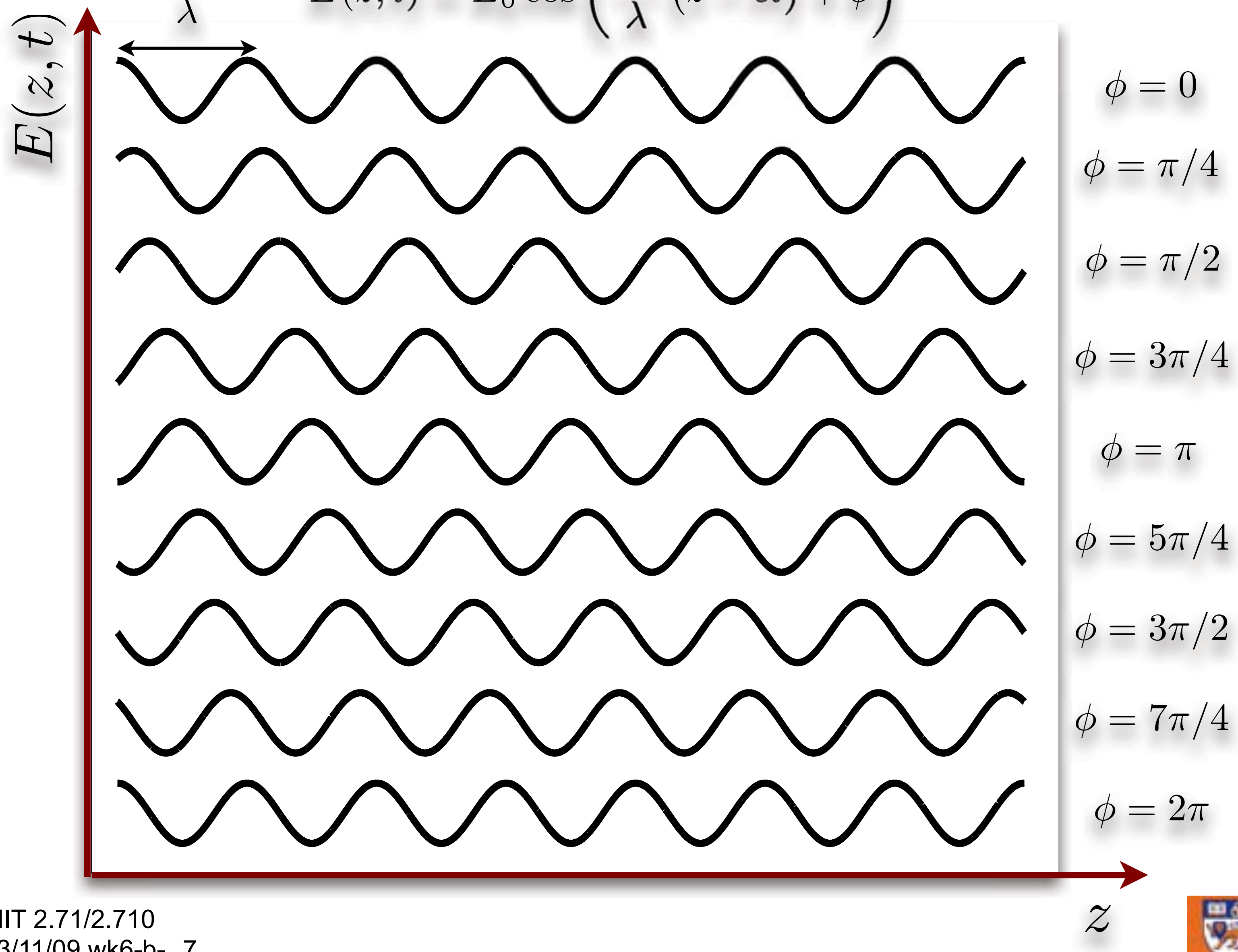
$$\frac{2\pi}{\lambda} cT = 2\pi \Rightarrow \frac{c}{\lambda\nu} = 1 \Rightarrow \boxed{c = \lambda\nu}$$

Dispersion relation

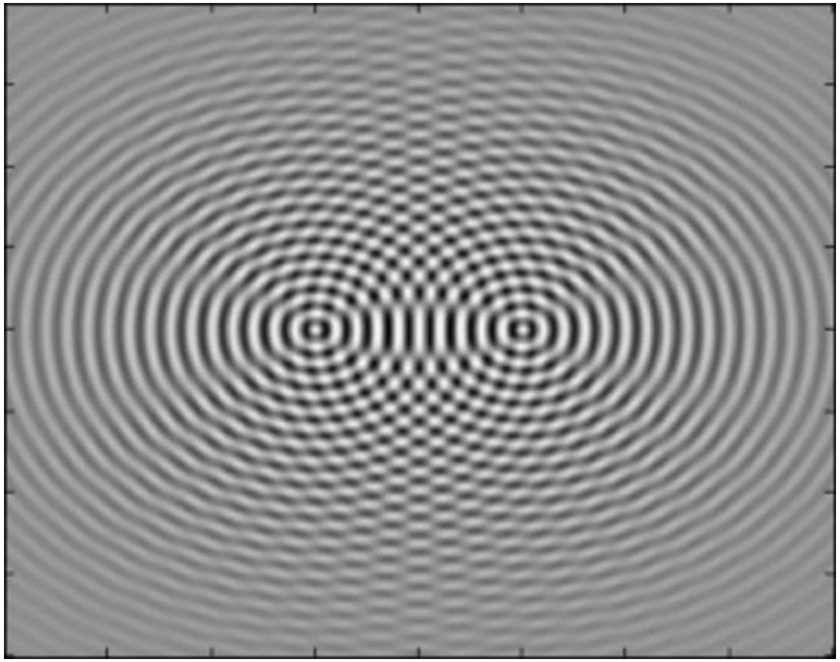
The significance of phase delays

$t = 0$

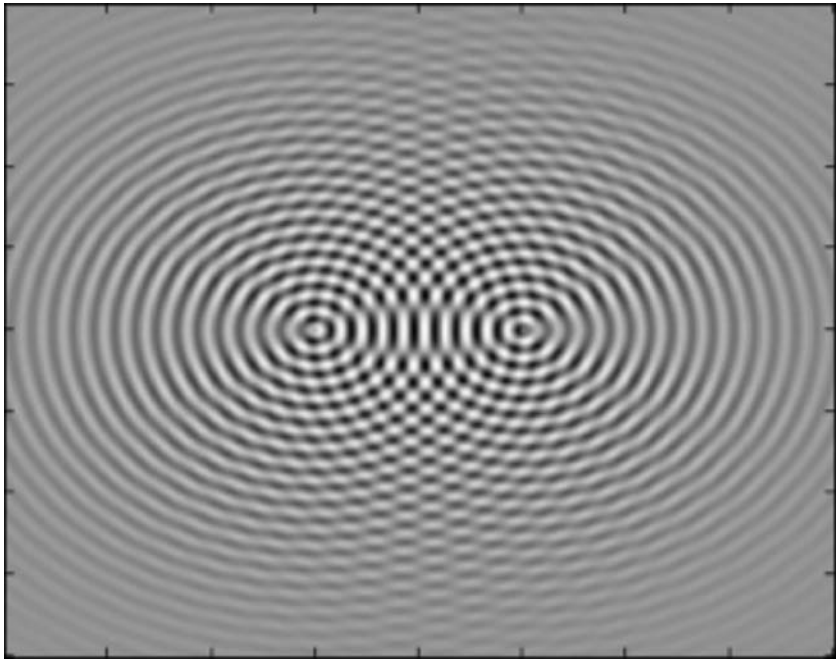
$$E(z,t) = E_0 \cos\left(\frac{2\pi}{\lambda}(z - ct) + \phi\right)$$



Phase delays and interference



in phase



out of phase