Ray Tracing

Reading

Required:

- Shirley, section 10.1-10.7 (handout)
- Triangle intersection handout

Further reading:

- Shirley errata on syllabus page, needed if you work from his book instead of the handout, which has already been corrected.
- T. Whitted. An improved illumination model for shaded display. Communications of the ACM 23(6), 343-349, 1980.
- A. Glassner. An Introduction to Ray Tracing. Academic Press, 1989.
- K. Turkowski, "Properties of Surface Normal Transformations," Graphics Gems, 1990, pp. 539-547.

Geometric optics

Modern theories of light treat it as both a wave and a particle.

We will take a combined and somewhat simpler view of light – the view of **geometric optics**.

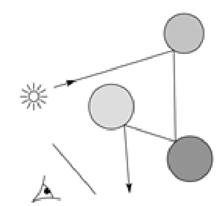
Here are the rules of geometric optics:

- Light is a flow of photons with wavelengths. We'll call these flows "light rays."
- Light rays travel in straight lines in free space.
- Light rays do not interfere with each other as they cross.
- Light rays obey the laws of reflection and refraction.
- Light rays travel from the light sources to the eye, but the physics is invariant under path reversal (reciprocity).

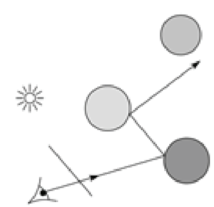
Eye vs. light ray tracing

Where does light begin?

At the light: light ray tracing (a.k.a., forward ray tracing or photon tracing)



At the eye: eye ray tracing (a.k.a., backward ray tracing)

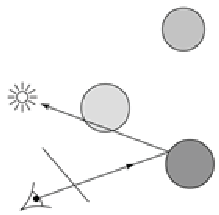


We will generally follow rays from the eye into the scene.

Precursors to ray tracing

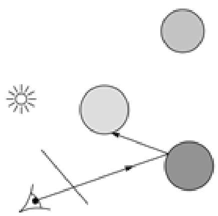
Local illumination

- Cast one eye ray, then shade according to light



Appel (1968)

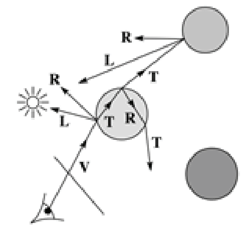
- Cast one eye ray + one ray to light



Whitted ray-tracing algorithm

In 1980, Turner Whitted introduced ray tracing to the graphics community.

- Combines eye ray tracing + rays to light
- Recursively traces rays



<u>Algorithm:</u>

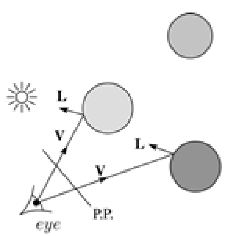
1. For each pixel, trace a **primary ray** in direction **V** to the first visible surface.

2. For each intersection, trace **secondary rays**:

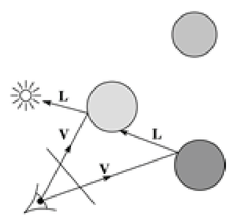
- **Shadow rays** in directions $\mathbf{L_i}$ to light sources
- **Reflected ray** in direction **R**.
- **Refracted ray** or **transmitted ray** in direction **T**.

Whitted algorithm (cont'd)

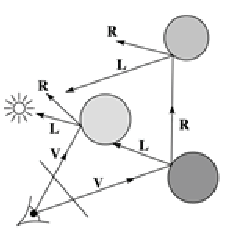
Let's look at this in stages:



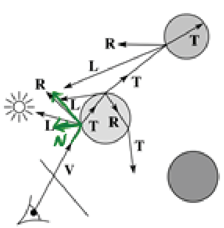
Primary rays



Shadow rays



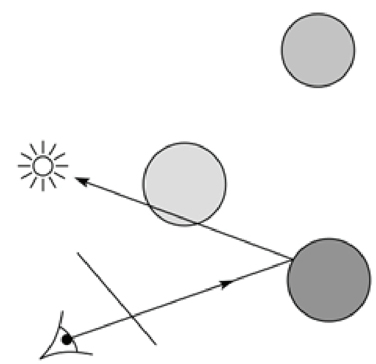
Reflection rays



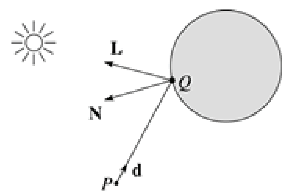
Refracted rays

Ray casting and local illumination

Now let's actually build the ray tracer in stages. We'll start with ray casting and local illumination:



Direct illumination



A ray is defined by an origin **P** and a unit direction **d** and is parameterized by $t > 0$:

$$\mathbf{r}(t) = P + t\mathbf{d}$$

Let $I(P, \mathbf{d})$ be the intensity seen along a ray. Then:

$$I(P, \mathbf{d}) = I_{\text{direct}}$$

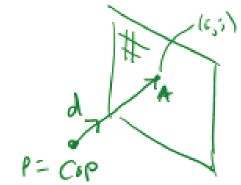
where

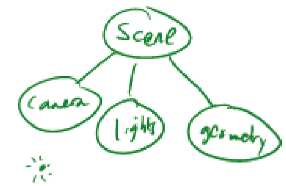
- I_{direct} is computed from the Phong model

Ray-tracing pseudocode

We build a ray traced image by casting rays through each of the pixels.

function *traceImage* (scene):
  for each pixel (i,j) in image
    A = *pixelToWorld*(i,j)
    P = **COP** ←
    $\mathbf{d} = (A - P)/\|A - P\|$ ⇐
    I(i,j) = *traceRay*(scene, P, **d**)
  end for
end function



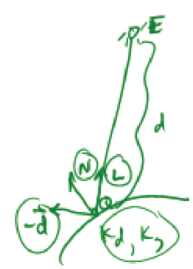


function *traceRay*(scene, P, **d**):
  (t, **N**, mtrl) ← scene.*intersect* (P, **d**)
  Q ← ray (P, **d**) evaluated at t
  I = *shade*(mtrl, N, scene, Q, -d)
  return I
end function

Shading pseudocode

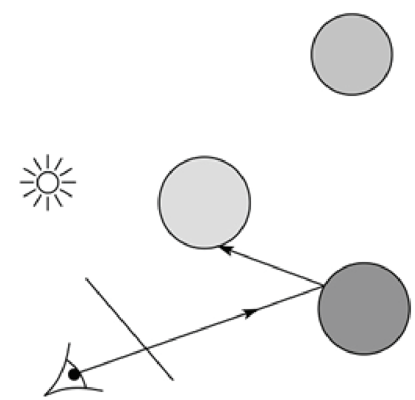
Next, we need to calculate the color returned by the *shade* function.

function *shade*(mtrl, scene, Q, **N**, **d**):
  I ← mtrl.k_e + mtrl. k_a * I_a
  for each light source L **do**:
    atten = L -> *distanceAttenuation*(Q)
    I ← I + atten*(diffuse term + specular term)
  end for
  return I
end function



Ray casting with shadows

Now we'll add shadows by casting shadow rays:



Shading with shadows

To include shadows, we need to modify the shade function:

function *shade*(mtrl, scene, *Q*, **N**, **d**):

  I ← mtrl.k_e + mtrl. k_a * I_a

  for each light source L **do**:

    atten = L -> *distanceAttenuation*(*Q*) *

      L -> *shadowAttenuation*(scene, Q)

    I ← I + atten*(diffuse term + specular term)

  end for

  return I

end function

Shadow attenuation

Computing a shadow can be as simple as checking to see if a ray makes it to the light source.

For a point light source:

function *PointLight*::*shadowAttenuation*(scene, *P*)

  d = (this.position - *P*).*normalize*()

  (t, **N**, mtrl) ← scene.*intersect*(*P*, **d**)

  Compute t_{light}

  if (t < t_{light}) **then**:

    atten = 0

  else

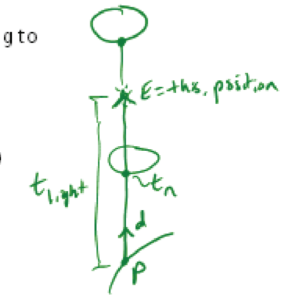
    atten = 1

  end if

  return atten

end function

For a directional light, $t_{light} = \infty$.



Shading in "Trace"

The Trace project uses a version of the Phong shading equation we derived in class, with two modifications:

- Distance attenuation is clamped to be at most 1:

$$A_j^{dist} = \min\left\{1, \frac{1}{a_j + b_j d_j + c_j d_j^2}\right\}$$

- Shadow attenuation A^{shadow} is included.

Here's what it should look like:

$$I = k_e + k_a L_a + \sum_j A_j^{shadow} A_j^{dist} L_j \left[k_d (\mathbf{N} \cdot \mathbf{L}_j)_+ + k_s B_j (\mathbf{V} \cdot \mathbf{R}_j)_+^{n_s} \right]$$

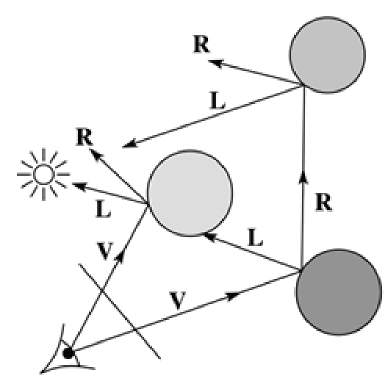
I.e., we are not using the OpenGL shading equation, which is somewhat different.

Note: the "**R**" here is the reflection of the *light* about the surface normal.

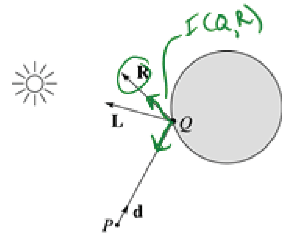
You must use the shading equation on this slide, not the one in Shirley's textbook.

Recursive ray tracing with reflection

Now we'll add reflection:



Shading with reflection



Let $I(P, \mathbf{d})$ be the intensity seen along a ray. Then:

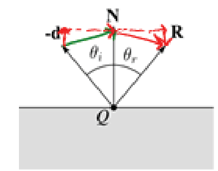
$$I(P,\mathbf{d}) = I_{\text{direct}} + I_{\text{reflected}}$$

where

- I_{direct} is computed from the Phong model, plus shadow attenuation
- $I_{\text{reflected}} = k_{\mathbf{r}} I\,(Q, \mathbf{R})$

Typically, we set $k_{\mathbf{r}} = k_s$.

Reflection



Law of reflection:

$$\theta_i = \theta_r$$

R is co-planar with **d** and **N**.

Ray-tracing pseudocode, revisited

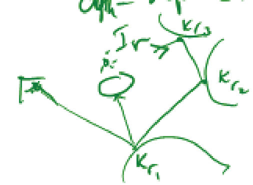
function *traceRay*(scene, *P*, **d**):
  (t, **N**, mtrl) ← scene.*intersect*(*P*, **d**)
  Q ← ray (*P*, **d**) evaluated at t
  I = *shade*(scene, mtrl, *Q*, **N**, **-d**)
  R = *reflectDirection*(-d, N)
  I ← I + mtrl.k_r ∗ *traceRay*(scene, *Q*, **R**)
  return I
end function

Terminating recursion

Q: How do you bottom out of recursive ray tracing?

Possibilities:

depth ≤ depth_Threshold



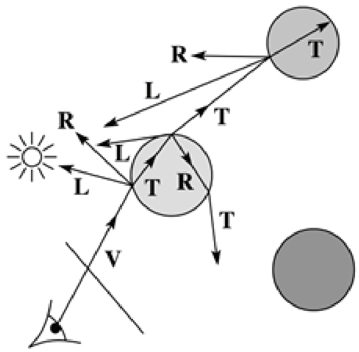
$K_{r_1}, K_{r_2} K_{r_3}\ I_r$

$\prod_j K_{r_j} <$ thresh.

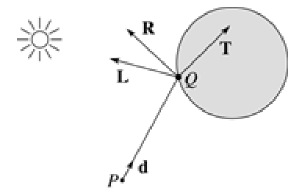
early / adaptive ray termination

Whitted ray tracing

Finally, we'll add refraction, giving us the Whitted ray tracing model:



Shading with reflection and refraction



Let $I(P, \mathbf{d})$ be the intensity seen along a ray. Then:

$$I(P, \mathbf{d}) = I_{\text{direct}} + I_{\text{reflected}} + I_{\text{transmitted}}$$

where

- I_{direct} is computed from the Phong model, plus shadow attenuation
- $I_{\text{reflected}} = k_r I\,(Q, \mathbf{R})$
- $I_{\text{transmitted}} = k_t I\,(Q, \mathbf{T})$

Typically, we set $k_r = k_s$ and $k_t = 1 - k_s$ (or 0, if opaque).

[Generally, k_r and k_t are determined by "Fresnel reflection," which depends on angle of incidence and changes the polarization of the light. This is discussed in Shirley's textbook and can be implemented for extra credit.]

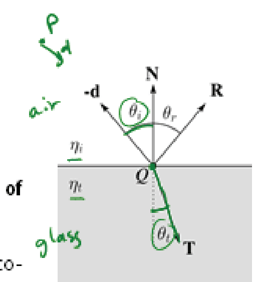
Refraction

Snell's law of refraction:

$$\eta_i \sin\theta_i = \eta_t \sin\theta_t$$

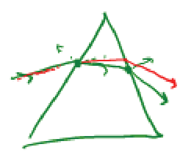
where η_i, η_t are **indices of refraction**.

In all cases, **R** and **T** are co-planar with **d** and **N**.



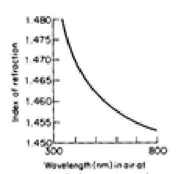
$\sin\theta_t = \frac{\eta_i}{\eta_t} \sin\theta_i$

The index of refraction is material dependent.



It can also vary with wavelength, an effect called **dispersion** that explains the colorful light rainbows from prisms. (We will generally assume no dispersion.)

Medium	Index of refraction
Vaccum	1
Air	1.0003
Water	1.33
Fused quartz	1.46
Glass, crown	1.52
Glass, dense flint	1.66
Diamond	2.42



Index of refraction variation for fused quartz

Total Internal Reflection

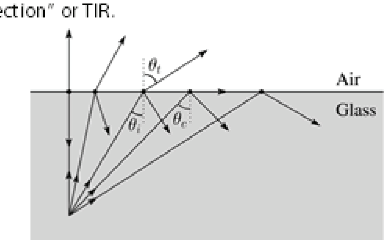
The equation for the angle of refraction can be computed from Snell's law:

$$\theta_t = \sin^{-1}\left(\frac{\eta_i}{\eta_t} \sin\theta_i\right)$$

What happens when $\eta_i > \eta_t$?

When θ_t is exactly 90°, we say that θ_i has achieved the "critical angle" θ_c.

For $\theta_i > \theta_c$, *no rays are transmitted*, and only reflection occurs, a phenomenon known as "total internal reflection" or TIR.



Shirley handout

Shirley uses different symbols. Here is the translation between them:

$\mathbf{r} = \mathbf{R}$

$\mathbf{t} = \mathbf{T}$

$\phi = \theta_t$

$\theta = \theta_r = \theta_i$

$n = \eta_i$

$n_t = \eta_t$

Also, Shirley has two important errors that have already been corrected in the handout.

But, if you're consulting the original text, be sure to refer to the errata posted on the syllabus and on the project page for corrections.

Ray-tracing pseudocode, revisited

function *traceRay*(scene, *P*, **d**):

(t, **N**, mtrl) ← scene.*intersect* (*P*, **d**)

Q ← ray (*P*, **d**) evaluated at t

I = *shade*(scene, mtrl, *Q*, **N**, -**d**)

R = *reflectDirection*(**N**, -**d**)

I ← I + mtrl.k_r * *traceRay*(scene, *Q*, **R**)

if ray is entering object **then**

n_i = index_of_air

n_t = mtrl.index

else

n_i = mtrl.index

n_t = index_of_air

if (*notTIR* (N, d, n_i, n_t)) **then**

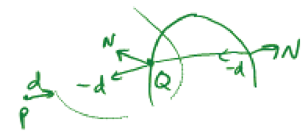
T = *refractDirection* ( )

I ← I + mtrl.k_t * *traceRay*(scene, *Q*, **T**)

end if

return I

end function

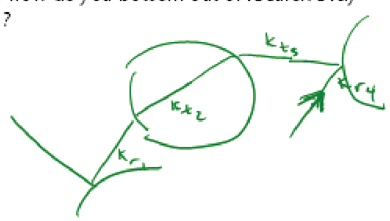


$(-d \cdot N) > 0 \Rightarrow$ entering

Q: How do we decide if a ray is entering the object?

Terminating recursion, incl. refraction

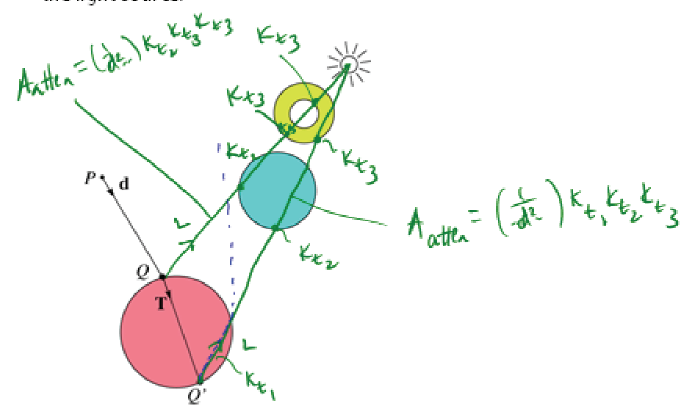
Q: *Now* how do you bottom out of recursive ray tracing?



$\prod_j K_{\{r,t\}_j}$

Shadow attenuation (cont'd)

Q: What if there are transparent objects along a path to the light source?

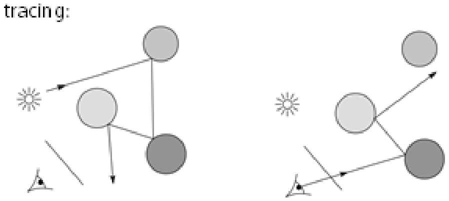


Suppose for simplicity that each object has a multiplicative transparency constant, k_t, which gets factored in every time an object is entered, possibly more than once for the same object.

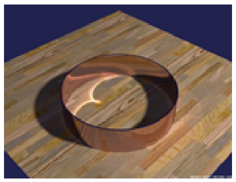
Shirley's textbook describes a better attenuation model based on Beer's Law, which you can implement for extra credit.

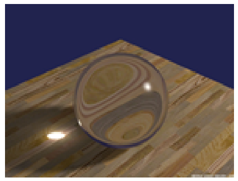
Photon mapping

Combine light ray tracing (photon tracing) and eye ray tracing:



...to get **photon mapping**.



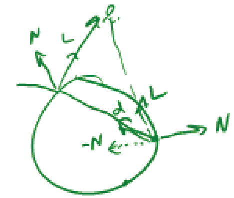


Renderings by Henrik Wann Jensen:
http://graphics.ucsd.edu/~henrik/images/caustics.html

Normals and shading when inside

When a ray is inside an object and intersects the object's surface on the way out, the normal will be pointing ***away*** from the ray (i.e., the normal always points to the outside by default).

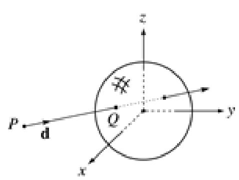
You must ***negate*** the normal before doing any of the shading, reflection, and refraction that follows.



Intersecting rays with spheres

Now we've done everything except figure out what that "scene.*intersect*(*P*, **d**)" function does.

Mostly, it calls each object to find out the t value at which the ray intersects the object. Let's start with intersecting spheres...



$r(t) = P + td$

Given:

- The coordinates of a point along a ray passing through *P* in the direction **d** are:

$$x = P_x + td_x$$
$$y = P_y + td_y$$
$$z = P_z + td_z$$

- A unit sphere *S* centered at the origin defined by the equation:

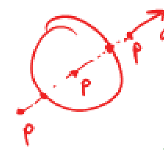
$x^2 + y^2 + z^2 = 1$

$g(x,y,z) = x^2 + y^2 + z^2 - 1$

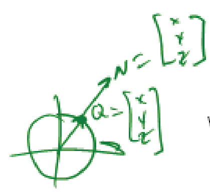
$N \sim \nabla g = (2x, 2y, 2z)$

Find: The *t* at which the ray intersects *S*.

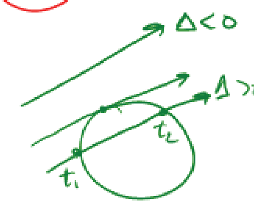
Intersecting rays with spheres



Solution by substitution:



$$x^2 + y^2 + z^2 - 1 = 0$$
$$(P_x + td_x)^2 + (P_y + td_y)^2 + (P_z + td_z)^2 - 1 = 0$$
$$at^2 + bt + c = 0$$



where

$$a = d_x^2 + d_y^2 + d_z^2$$
$$b = 2(P_x d_x + P_y d_y + P_z d_z)$$
$$c = P_x^2 + P_y^2 + P_z^2 - 1$$

Q: What are the solutions of the quadratic equation in *t* and what do they mean?

$$t = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$\Delta = b^2 - 4ac$ (discriminant)

$\Delta > 0 \Rightarrow$ 2 real roots $\Rightarrow 2 \cap$s

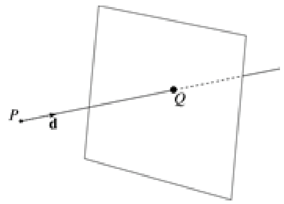
$\Delta < 0 \Rightarrow$ 2 complex roots $\Rightarrow$ no $\cap$s

$\Delta = 0 \Rightarrow$ 1 real double root $\Rightarrow$ tangent $\cap$ (1 $\cap$)

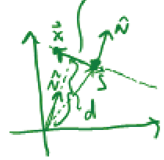
Q: What is the normal to the sphere at a point (*x,y,z*) on the sphere?

$N = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

Ray-plane intersection



$(\vec{x}-\vec{s})\cdot\hat{N}=0$



We can write the equation of a plane as:

$$ax + by + cz = d$$

$\hat{N}\cdot(\vec{x}-\vec{s})=0$

$\hat{N}\cdot\vec{x}-\hat{N}\cdot\vec{s}=0$

$\hat{N}\cdot\vec{x}=\hat{N}\cdot\vec{s}$

$\hat{N}\cdot\vec{x}=d$

$n_x x+n_y y+n_z z=d$

The coefficients a, b, and c form a vector that is normal to the plane, $\mathbf{n} = [a\ b\ c]^T$. Thus, we can re-write the plane equation as:

We can solve for the intersection parameter (and thus the point):

$\hat{N}\cdot r(t)=d$

$\hat{N}\cdot(P+t\vec{d})=d$

$\hat{N}\cdot P+\hat{N}\cdot(t\vec{d})=d$

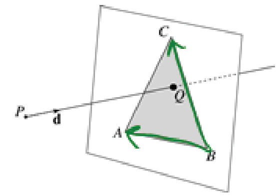
$\hat{N}\cdot P+t\hat{N}\cdot\vec{d}=d$

$t=\dfrac{d-\hat{N}\cdot P}{\hat{N}\cdot\vec{d}}$

if $(N\cdot\vec{d}=0) \Rightarrow$ no $\cap$

// to plane

Ray-triangle intersection



right-handed (CCW) ordering of vertices to indicate normal orientation

To intersect with a triangle, we first solve for the equation of its supporting plane.

How might we compute the (un-normalized) normal?

$N=(C-B)\times(A-B)$ $\qquad \hat{N}=\dfrac{N}{\|N\|}$

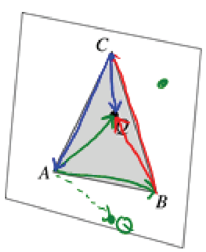
Given this normal, how would we compute d?

$\hat{N}\cdot A=d=\hat{N}\cdot B=\hat{N}\cdot C$

Using these coefficients, we can solve for Q. Now, we need to decide if Q is inside or outside of the triangle.

3D inside-outside test

One way to do this "inside-outside test," is to see if Q lies on the left side of each edge as we move counterclockwise around the triangle.



How might we use cross products to do this?

$[(B-A)\times(Q-A)]\cdot\hat{N}\geq 0$ and

$[(C-B)\times(Q-B)]\cdot\hat{N}\geq 0$ and

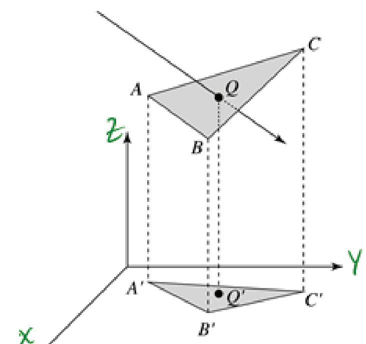
$[(A-C)\times(Q-C)]\cdot\hat{N}\geq 0$

$\Downarrow$

inside

2D inside-outside test

Without loss of generality, we can perform this same test after projecting down a dimension:



If Q' is inside of $A'B'C'$, then Q is inside of ABC.

Why is this projection desirable? faster

Which axis should you "project away"?

axis corresponding to largest component of N

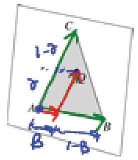
Barycentric coordinates

As we'll see in a moment, it is often useful to represent Q as an **affine combination** of A, B, and C:

$$Q = \alpha A + \beta B + \gamma C$$

where:

$$\alpha + \beta + \gamma = 1$$



We call α, β, and γ the **barycentric coordinates** of Q with respect to A, B, and C.

$$\alpha \begin{bmatrix} A_x \\ A_y \\ A_z \\ 1 \end{bmatrix} + \beta \begin{bmatrix} B_x \\ B_y \\ B_z \\ 1 \end{bmatrix} + \gamma \begin{bmatrix} C_x \\ C_y \\ C_z \\ 1 \end{bmatrix} = \begin{bmatrix} - \\ - \\ - \\ 1 \end{bmatrix}$$

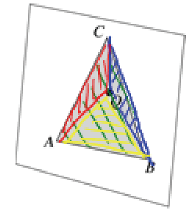
$$Q = A + \beta(B - A) + \gamma(C - A)$$
$$= A + \beta B - \beta A + \gamma C - \gamma A$$
$$= \underbrace{(1 - \beta - \gamma)}_{\alpha} A + \beta B + \gamma C$$

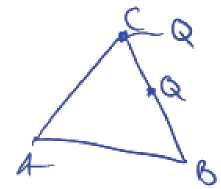
$$\alpha = 1 - \beta - \gamma \quad \Downarrow \quad \alpha + \beta + \gamma = 1$$

Computing barycentric coordinates

Given a point Q that is inside of triangle ABC, we can solve for Q's barycentric coordinates in a simple way:

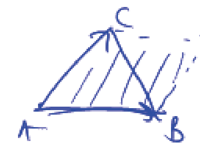
$$\alpha = \frac{\text{Area}(QBC)}{\text{Area}(ABC)} \quad \beta = \frac{\text{Area}(AQC)}{\text{Area}(ABC)} \quad \gamma = \frac{\text{Area}(ABQ)}{\text{Area}(ABC)}$$





How can cross products help here?

$$\text{Area}(ABC) = \frac{1}{2} \| (B - A) \times (C - A) \|$$



In the end, these calculations can be performed in the 2D projection as well!

Interpolating vertex properties

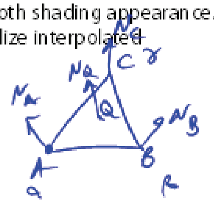
The barycentric coordinates can also be used to interpolate vertex properties such as:

- material properties
- texture coordinates
- normals

For example:

$$k_d(Q) = \alpha k_d(A) + \beta k_d(B) + \gamma k_d(C)$$

Interpolating normals, known as Phong interpolation, gives triangle meshes a smooth shading appearance. (Note: don't forget to normalize interpolated normals.)

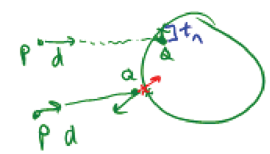


$$N_Q = \alpha N_A + \beta N_B + \gamma N_C$$

Epsilons

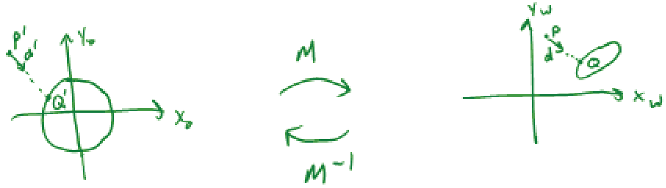
Due to finite precision arithmetic, we do not always get the exact intersection at a surface.

Q: What kinds of problems might this cause?



Q: How might we resolve this?

$t_n <$ RAY_EPSILON $\Rightarrow$ discard n

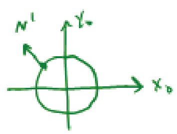


$$Q = P + t_n d$$

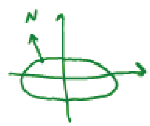
$$M^{-1}Q = M^{-1}(P + t_n d)$$

$$\underbrace{M^{-1}Q}_{Q'} = \underbrace{M^{-1}P}_{P'} + t_n \underbrace{M^{-1}d}_{d'}$$

$$Q' = P' + t_n d'$$



$$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



$$\Rightarrow N = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} N'$$

$$M = \left[\begin{array}{c|c} A_{3\times3} & t \\ \hline 0 & 1 \end{array}\right] \Rightarrow N = \left[\begin{array}{c|c} A_{3\times3}^{-T} & 0 \\ \hline 0 & 1 \end{array}\right] N'$$

Summary

What to take home from this lecture:

- The meanings of all the boldfaced terms.
- Enough to implement basic recursive ray tracing.
- How reflection and transmission directions are computed.
- How ray--object intersection tests are performed on spheres, planes, and triangles
- How barycentric coordinates within triangles are computed
- How ray epsilons are used.