A New Derivation of Snell's Law Without Calculus

John A. Quintanilla

John Quintanilla (John.Quintanilla@unt.edu), known as "Dr. Q" to his students, is a University Distinguished Teaching Professor of Mathematics at the University of North Texas. He is currently Associate Dean for Undergraduate Studies for the College of Science. He was a founding co-director of Teach North Texas, a UTeach replication and UNT's program for preparing secondary teachers of mathematics and science. He earned his BS and MS from Stanford and his MA and PhD from Princeton. His blog is www.meangreenmath.com. At home, he enjoys time with his wife and daughter.

Snell's Law describes the refraction of light through a linear interface between two regions (like air and water) in which light travels at different speeds. In Figure 1, we depict a path of light traveling from A to D, two points in different regions which are distances a and d, respectively, from the horizontal interface. We also let c be the horizontal distance between A and D.

Due to diffraction, light traveling from A to D crosses the interface at a point P which is not on $\overline{AD}$; we let x be the horizontal distance between P and D. Let v and w be the speed of light along $\overline{AP}$ and $\overline{PD}$, respectively; without loss of generality, we will assume that $v > w > 0$. Then the travel time along APD is

$$T_1(x) = \frac{\sqrt{a^2 + (c-x)^2}}{v} + \frac{\sqrt{x^2 + d^2}}{w}.$$

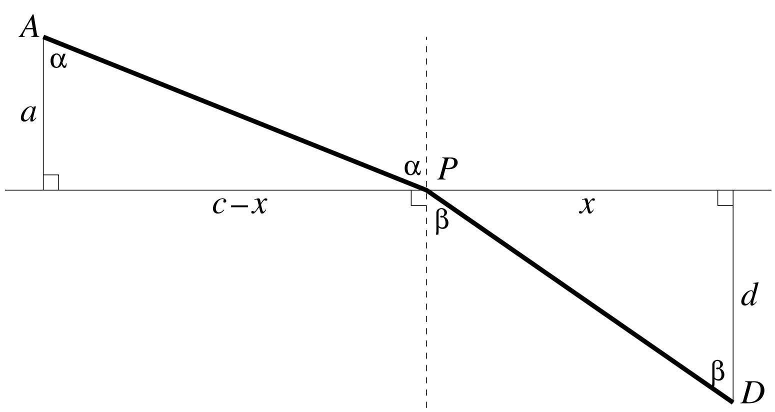

Figure 1. Snell's Law arises from minimizing $T_1(x)$, the travel time along APD, where the speeds on $\overline{AP}$ and $\overline{PD}$ are different.

Fermat's Principle dictates that light travels from A to D along the path APD that requires the least amount of time. Identifying the point P, or equivalently the distance x, is an optimization problem that is often presented in calculus textbooks. Indeed, as we discuss later, solving the equation $T_1'(x) = 0$ produces Snell's Law,

$$\frac{\sin\alpha}{v} = \frac{\sin\beta}{w},$$

where the angle of incidence α and the angle of refraction β are depicted in Figure 1. Alternatively, Snell's Law can be derived without calculus by using Huygens's Principle [**1, 2, 3**] or Ptolemy's Theorem [**4, 5**].

We now present a new way of deriving Snell's Law without calculus. We will see that the key equation (2) in this new approach, obtained by setting the ratio of two lengths equal to the ratio of the two speeds, is equivalent to the equation $T_1'(x) = 0$.

The Shortest Distance Between Two Points

In Figure 2, we depict a related but much simpler problem: if light has the *same* speed v in both regions, find the point Q so that light travels along path AQB in least time. In this figure, the distances a and c are the same as in Figure 1, but the distance b is (for now) arbitrary.

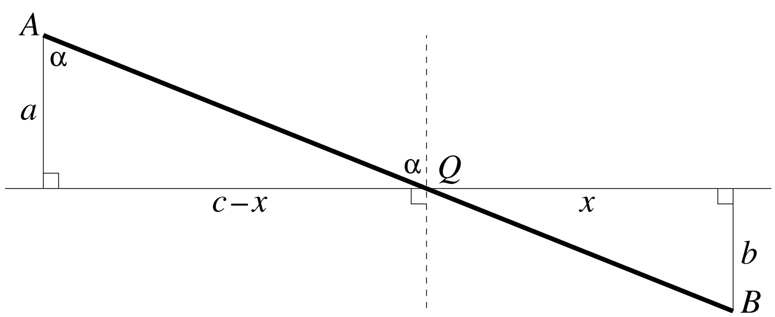

Figure 2. A simpler problem: Find the x_0 that minimizes $T_0(x)$, the travel time along AQB, if the speed is constant.

While calculus could be used to minimize the travel time along AQB, which is

$$T_0(x) = \frac{\sqrt{a^2 + (c-x)^2}}{v} + \frac{\sqrt{x^2 + b^2}}{v},$$

the constant speed v in both regions implies that T_0 is minimized when $AQ + QB$, the total distance traveled, is minimized. Geometrically, it is clear that this occurs when Q is on $\overline{AB}$. The two right triangles in Figure 2 are thus similar at this optimal $x = x_0$, so that

$$\frac{a}{c - x_0} = \frac{b}{x_0}, \qquad \text{or} \qquad x_0 = \frac{bc}{a+b}. \tag{1}$$

Connecting the Two Optimization Problems

We now relate Figures 1 and 2 by taking $d > 0$ as given, defining $k = v/w > 1$ to be the ratio of the two speeds, and selecting b to be the unique positive number so that $PD = k(QB)$, or

$$\sqrt{x_0^2 + d^2} = k\sqrt{x_0^2 + b^2}. \tag{2}$$

To see that b is unique, we use (1) to rewrite (2) as

$$d^2 = k^2 b^2 + (k^2 - 1)c^2 \left(\frac{b}{a+b}\right)^2. \tag{3}$$

For $b > 0$, the right-hand side of (3) is an increasing function of b with range $(0, \infty)$ since the constants k^2 and $(k^2 - 1)c^2$ are positive while

$$b \mapsto b^2 \qquad \text{and} \qquad b \mapsto \left(\frac{b}{a+b}\right)^2 = \left(1 - \frac{a}{a+b}\right)^2$$

are both increasing functions of b. We conclude that if $d > 0$ is given, there is a unique positive b that solves (3) and thus (2). We remark that we have demonstrated the existence and uniqueness of this distance b without using calculus.

Our main result is that, for given values of the distances a, c, and d, the distance x_0 in (1) minimizes $T_2(x) = T_1(x) - T_0(x)$ if b is the positive solution of (2) or (3). Since x_0 minimizes both $T_0(x)$ and $T_2(x)$, it also minimizes $T_1(x) = T_0(x) + T_2(x)$. Therefore, when T_0 and T_1 are minimized, the point P in Figure 1 is the same as the point Q in Figure 2, so that the angles α in the two figures are the same. Snell's Law follows since

$$\frac{\sin\beta}{w} = \frac{x_0}{w\sqrt{x_0^2 + d^2}} = \frac{x_0}{kw\sqrt{x_0^2 + b^2}} = \frac{x_0}{v\sqrt{x_0^2 + b^2}} = \frac{\sin\alpha}{v};$$

the last equality holds since the right triangles in Figure 2 are similar when Q is on $\overline{AB}$.

Thus, to prove Snell's Law, it suffices to show that the x_2 that minimizes

$$vT_2(x) = vT_1(x) - vT_0(x) = k\sqrt{x^2 + d^2} - \sqrt{x^2 + b^2},$$

where b and d are related by (2) or (3), is the same as the x_0 that minimizes $T_0(x)$. While minimizing T_2 is a straightforward exercise with calculus, we instead demonstrate that $x_2 = x_0$ without calculus by rotating a certain hyperbola.

Rotating a Hyperbola

Define $t(x) = k\sqrt{x^2 + d^2}$ and $s(x) = \sqrt{x^2 + b^2}$, so that the goal becomes minimizing $t - s$. Subtracting the equations $t^2/k^2 - x^2 = d^2$ and $s^2 - x^2 = b^2$ yields

$$\frac{t^2}{k^2(k^2 - 1)(x_0^2 + b^2)} - \frac{s^2}{(k^2 - 1)(x_0^2 + b^2)} = 1 \tag{4}$$

after using the relationship $d^2 - b^2 = (k^2-1)(x_0^2 + b^2)$ from (2). In other words, the points $(s(x), t(x))$ lie on a branch of a hyperbola in the st-plane, as shown in Figure 3. To find the x_2 that minimizes $vT_2(x) = t(x) - s(x)$, we need to find the point $R(s(x_2), t(x_2))$ on the upper branch of the hyperbola that is closest to the line $s = t$.

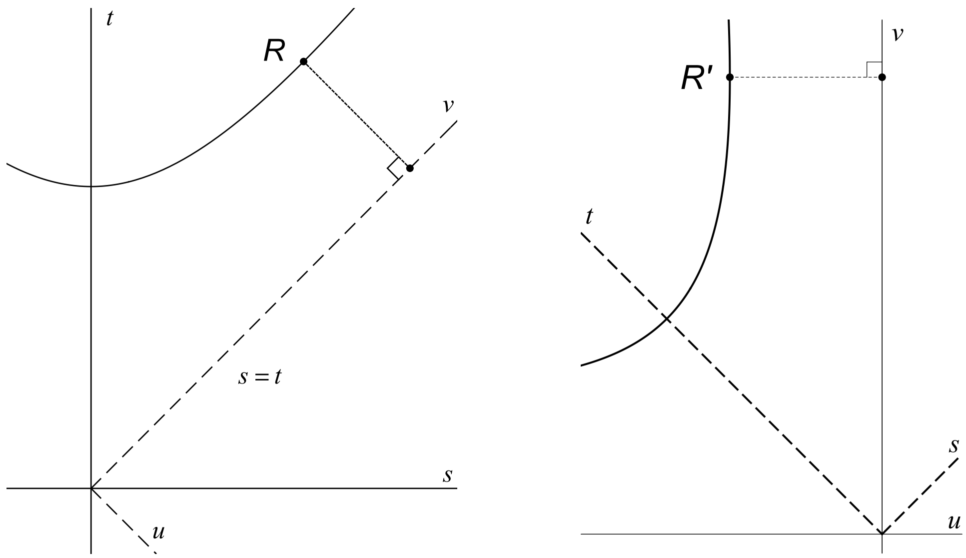

Figure 3. To find the point R on the hyperbola closest to the line $s = t$ and thus minimize T_2, we find the point R' on the rotated hyperbola with the least negative $u-$coordinate.

The coordinates of R may be found without calculus by first rotating the hyperbola $45°$ counterclockwise by substituting

$$s = \frac{u+v}{\sqrt{2}} \qquad \text{and} \qquad t = \frac{v-u}{\sqrt{2}}$$

into (4). With these substitutions, the line $s = t$ in the st-plane rotates to the v-axis in the uv-plane. We also define R' to be the image of R after the rotation. After some simplification, we find that the equation of the rotated hyperbola is

$$(k^2-1)v^2 + 2(k^2+1)uv + (k^2-1)u^2 + 2k^2(k^2-1)(x_0^2+b^2) = 0.$$

Using the quadratic formula to solve for v, we find

$$v = \frac{-(k^2+1)u \pm k\sqrt{4u^2 - 2(k^2-1)^2(x_0^2+b^2)}}{k^2-1}. \tag{5}$$

In other words, the graph of the rotated hyperbola branch is not given by a single function of u. Instead, (5) defines two functions of u whose graphs combine to form the rotated hyperbola branch.

After the rotation, R' is the point on the hyperbola that is closest to the $v-$axis and hence has the least negative u-coordinate. Since the graphs of the two functions of (5) also intersect at R', its coordinates (u_2, v_2) can be found by setting the discriminant in (5) equal to 0:

$$2u_2^2 = (k^2-1)^2(x_0^2+b^2) \qquad \text{and} \qquad v_2 = -\frac{(k^2+1)u_2}{k^2-1}.$$

The u-coordinate of R' must be negative, and so

$$\begin{aligned}(u_2, v_2) &= \left(\frac{(1-k^2)\sqrt{x_0^2+b^2}}{\sqrt{2}}, \frac{(k^2+1)\sqrt{x_0^2+b^2}}{\sqrt{2}}\right)\\ &= \left(\frac{s(x_0)-t(x_0)}{\sqrt{2}}, \frac{s(x_0)+t(x_0)}{\sqrt{2}}\right)\end{aligned}$$

since $k^2\sqrt{x_0^2+b^2} = k\sqrt{x_0^2+d^2} = t(x_0)$ after using (2). Therefore, the original point R in the st-plane has coordinates

$$(s(x_2), t(x_2)) = \left(\frac{u_2+v_2}{\sqrt{2}}, \frac{v_2-u_2}{\sqrt{2}}\right) = (s(x_0), t(x_0)).$$

We conclude that $x_2 = x_0$ since both $s(x)$ and $t(x)$ are bijections for non-negative x.

In summary, if the distances a, c, and d in Figure 1 are given and b is the unique positive solution of (3), then $x_0 = bc/(a+b)$ minimizes both $T_0(x)$ and $T_2(x)$ and thus also $T_1(x)$, from which Snell's Law follows.

Connection to Usual Derivation with Calculus

If we instead use calculus to minimize T_1, simplifying the equation $T_1'(x_0) = 0$ yields

$$\frac{c-x_0}{v\sqrt{a^2+(c-x_0)^2}} = \frac{x_0}{w\sqrt{x_0^2+d^2}}. \tag{6}$$

Attempting to solve for x_0 results in the nontrivial quartic equation

$$(c-x_0)^2(x_0^2+d^2) = k^2x_0^2[a^2+(c-x_0)^2]. \tag{7}$$

Indeed, this optimization problem is unusual among the examples normally considered in first-semester calculus in that textbooks typically do not attempt to find x_0. Instead, Snell's Law follows from (6) since

$$\sin\alpha = \frac{c-x_0}{\sqrt{a^2+(c-x_0)^2}} \qquad \text{and} \qquad \sin\beta = \frac{x_0}{\sqrt{x_0^2+d^2}}$$

by using the right triangles in Figure 1.

To gratifyingly connect these two different techniques for minimizing T_1, we isolate the d^2 term in (7) and then use (1):

$$\begin{aligned}d^2 &= \frac{k^2a^2x_0^2}{(c-x_0)^2} + (k^2-1)x_0^2\\ &= k^2b^2 + (k^2-1)c^2\left(\frac{b}{a+b}\right)^2,\end{aligned}$$

which is the same as (3). In other words, (6) is equivalent to (3), and thus (2), by using the solution for x_0 in (1). Therefore, the proportionality relationship (2) between the optimal PD and QB from Figures 1 and 2 is equivalent to the equation $T_1'(x_0) = 0$.

Summary. We present a novel method for deriving Snell's Law of refraction without using calculus by instead rotating a certain hyperbola. This approach includes a geometrically motivated equation that turns out to be equivalent to setting the derivative equal to zero.

This is an Accepted Manuscript of an article published by Taylor & Francis in **College Mathematics Journal** on January 28, 2022, available online:
`http://www.tandfonline.com/doi/full/10.1080/07468342.2022.2011544`.

References

1. Golumb, M. (1964), Proofs for the equivalence of Fermat's Principle and Snell's Law, *Amer. Math. Monthly* 71(5): 541–543. doi.org/10.1080/00029890.1964.11992274
2. Huygens, C. (1962). *Treatise on Light*. (Thompson, S. P., trans.) New York, NY: Dover.
3. Ling, S. J., Sanny, J., Moebs, W. (2016). *University Physics Volume 3*. Houston, TX: OpenStax. Available at https://openstax.org/books/university-physics-volume-3/pages/1-6-huygenss-principle
4. Niven, I. M. (1971). *Maxima and Minima Without Calculus*. Washington, DC: Mathematical Association of America.
5. Pedoe, D. (1964). A geometric proof of equivalence of Fermat's Principle and Snell's Law, *Amer. Math. Monthly* **71**(5): 543–544. doi.org/10.1080/00029890.1964.11992274

