Reflecting on the Shortest Path between Two Points

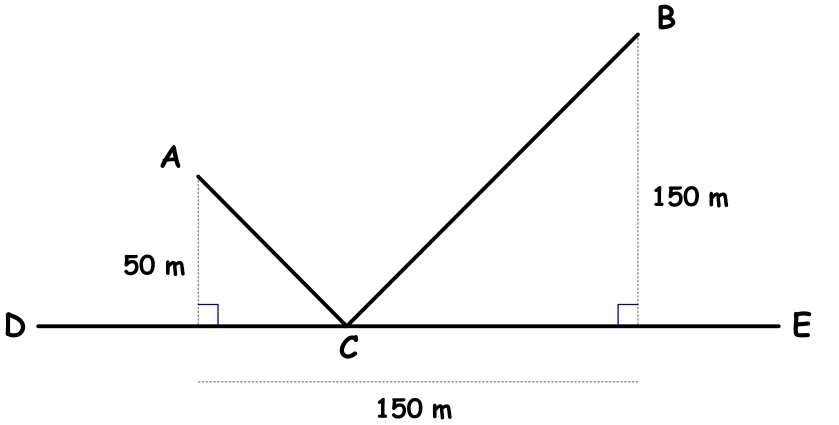
Duncan Samson
St Andrew's College, Grahamstown

Given two points on a surface, a geodesic is a curve on that surface representing the shortest path between the two points. Given two points in a plane, there is only one shortest curve through them, *viz.* the straight line joining them. On a curved surface it is possible for there to be more than one shortest curve between two given points. Imagine two diametrically opposite points on a sphere – in this scenario there are *infinitely* many shortest curves, since any great semicircle between the two points would represent a shortest path on the surface of the sphere.

The geometry of the sphere has important practical application in terms of navigating on or above the surface of the earth. Airline maps show curved flight paths between distant cities. This is because the shortest path between two points on a sphere is along an arc of a great circle (a circle whose centre coincides with the centre of the sphere).

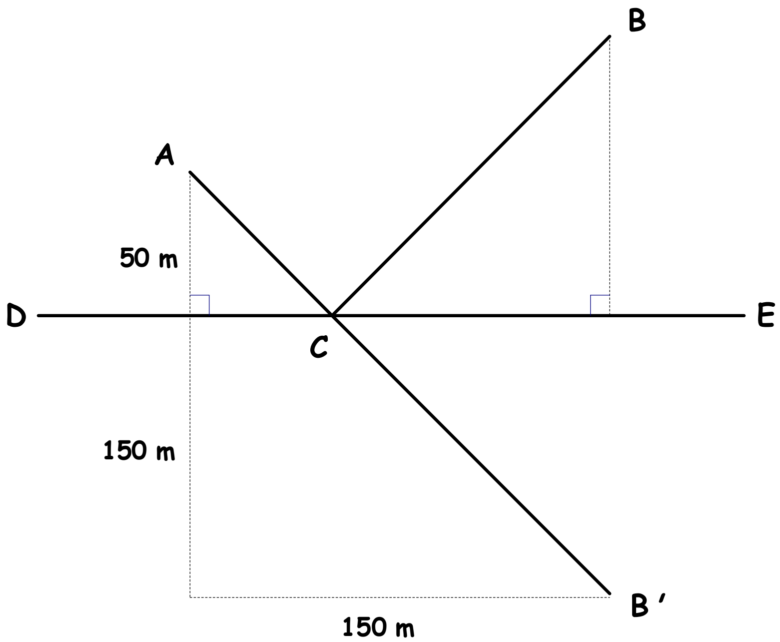
Geometrical problems connected with determining the shortest route from one point to another on a curved surface are often difficult, but geodesics on flat surfaces are in general readily determinable. For two points in the same plane, the shortest path between the two points is simply the straight line connecting them. Using this knowledge in conjunction with the principle of reflection allows for an elegant approach to solving a number of problems.

The diagram alongside shows two houses (A and B) and a straight pipeline DE. The two houses must be connected to the pipeline DE via junction C with two straight connecting pipes, AC and BC. What is the smallest length of piping required to complete the job?

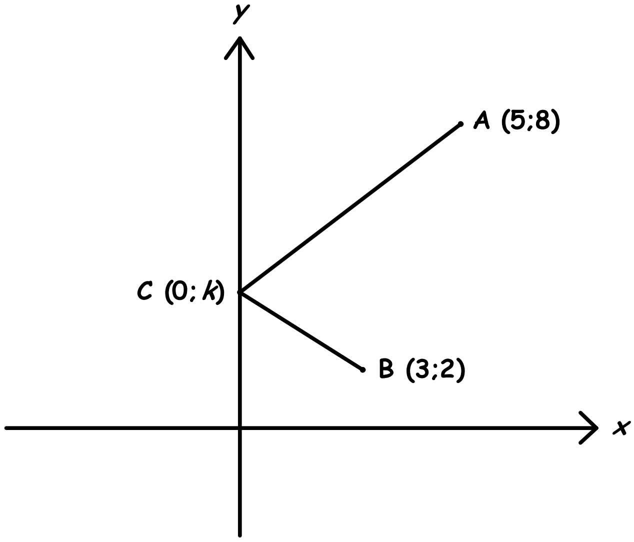


In essence we are being asked to calculate the smallest possible value of AC + CB. If we let B' be the reflection of B in the line DE then CB = CB'. AC + CB will be minimized when AC + CB' is a minimum. Since the shortest path between to points in a plane is the straight line connecting them, AC + CB' will be minimized when A, C and B' lie in a straight line.

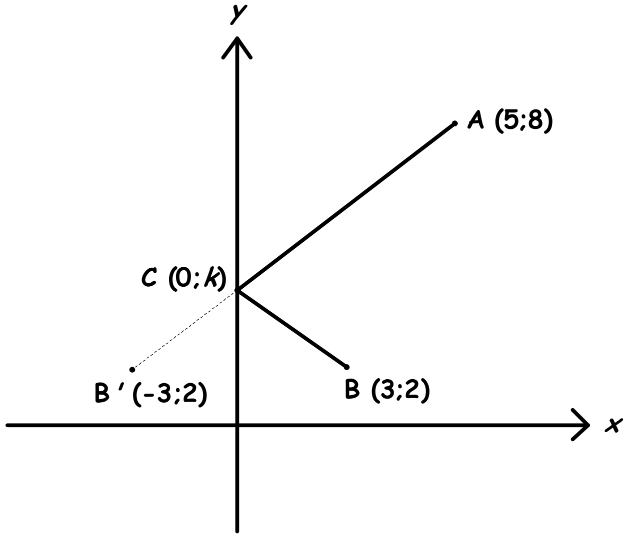
AB' is the hypotenuse of a right-angled triangle with sides 150 m, 200 m and therefore 250 m. The smallest possible value of AC + CB', and hence of AC + CB, is thus 250 m.



Let us now consider a related problem in the Cartesian Plane. The coordinates of A, B and C are (5;8), (3;2) and (0;k) respectively. What value of k will make AC + CB as small as possible? Clearly the y-coordinate of C must lie between the y-coordinates of A and B if we are to minimize AC + CB. We can thus represent the situation as follows:

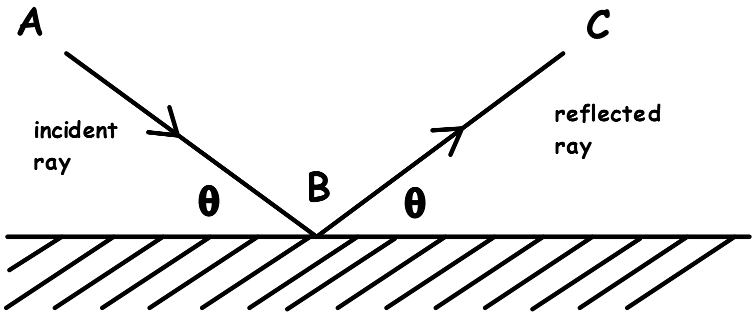


Let B' be the reflection of B in the y-axis. B' thus has coordinates (-3;2). Since CB = CB', AC + CB will be a minimum when AC + CB' is minimized. Since the shortest distance between two points is the straight line connecting them, AC + CB will be minimized when A, C and B' lie in a straight line:



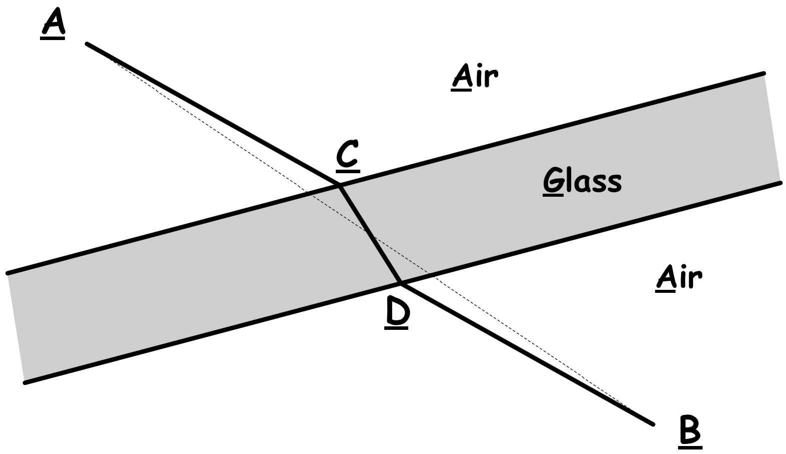
C is thus the y-intercept of the straight line from A to B'. The equation of straight line ACB' is $y = \frac{3}{4}x + 4\frac{1}{4}$ and the value of k is thus $4\frac{1}{4}$.

There is a wonderful connection between these reflection problems and the behaviour of light. If we consider a light ray reflecting off a flat, smooth surface then the angle of reflection will equal the angle of incidence. This is known as the Law of Reflection. Note that technically the angles of incidence and reflection are measured from the normal to the reflecting surface (i.e. from a line drawn perpendicular to the surface of reflection at the point where the light ray strikes it).



The path of the ray of light from A to C via B represents the shortest possible path, or more strictly the path that would take the least possible time to travel. This general principle, which can be used to determine the actual paths of light rays, was developed by Pierre de Fermat (1601 – 1665) and is known as Fermat's Principle of Least Time.

Imagine a light ray traveling from A to B through different refracting media: in this instance the light ray doesn't travel on the shortest path from A to B (dotted line) but rather on the path that will take the shortest time (solid line). The dotted line from A to B would require the light ray to travel further through the denser medium of glass (where its speed is less) and would not be the optimum path in terms of Fermat's Principle.



At first glance the actual path followed by the light ray from A to B may seem somewhat counter-intuitive. Why doesn't the light ray simply travel through the denser medium perpendicularly to its edges, thus minimizing the time spent traveling at a slower speed? Although this would minimize the time spent in the denser medium, it would increase the time spent traveling through the air on exiting the glass. The actual path followed from A to B (i.e. AC + CD + DB) represents a complex play-off in terms of minimizing the time for the journey as a whole.

This of course raises a somewhat perplexing question. If light rays always take the path of least time, how does light "choose" the path of least time? In other words, how does light "know" that other paths will take a longer time to travel? In a purely classical sense, this question is impossible to answer. However, we can approach the problem philosophically by imagining light being able to take all possible paths from point A to point B, and then "choosing" the path that takes the least time. Within the realm of quantum mechanics, this is precisely what light does.

The Law of Reflection, which states that the angle of reflection equals the angle of incidence, is a consequence of the Principle of Least Time. This topic can be worked up into a good investigation or extension exercise for Grade 12.

Firstly we will need to introduce the Chain Rule for differentiation:

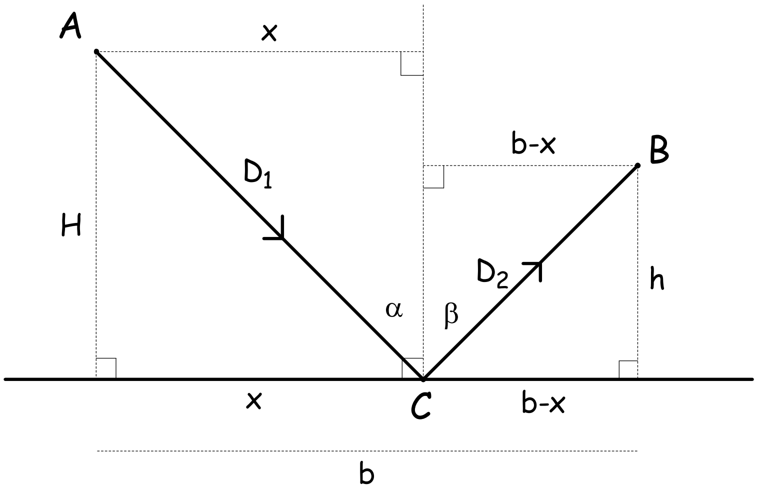
If $y = f(u)$ and $u = g(x)$ are both differentiable functions, then

$$\frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx}$$

By way of example, if $y = \sqrt{x^2+1}$ then using the above notation we have $y = \sqrt{u}$ and $u = x^2 + 1$. Thus $\frac{dy}{du} = \frac{1}{2}u^{-\frac{1}{2}} = \frac{1}{2\sqrt{u}}$ and $\frac{du}{dx} = 2x$. Thus:

$$\begin{aligned}\frac{dy}{dx} &= \frac{dy}{du} \times \frac{du}{dx} \\ &= \frac{1}{2\sqrt{u}}.2x \\ &= \frac{1}{2\sqrt{x^2+1}}.2x \\ &= \frac{x}{\sqrt{x^2+1}}\end{aligned}$$

Let us now consider a light ray reflecting off a flat, smooth surface. The light ray begins at point A, strikes the surface at C and is reflected to point B. The angles of incidence and reflection are α and β respectively. Note that the angles of incidence and reflection have been measured from the normal.



If the light travels at a constant speed v, then the total time for the journey from A to B via C is given by:

$$t = \frac{D_1 + D_2}{v}$$

Using Pythagoras, $D_1 = \sqrt{H^2 + x^2}$ and $D_2 = \sqrt{h^2 + (b-x)^2}$. Thus:

$$\begin{aligned} t &= \frac{D_1}{v} + \frac{D_2}{v} \\ &= \frac{\sqrt{H^2+x^2}}{v} + \frac{\sqrt{h^2+(b-x)^2}}{v} \\ &= \frac{1}{v}.\left(H^2+x^2\right)^{\frac{1}{2}} + \frac{1}{v}.\left(h^2+b^2-2bx+x^2\right)^{\frac{1}{2}} \end{aligned}$$

Now, bearing in mind that H, h, b and v are all constants, we can use the Chain Rule to determine $\frac{dt}{dx}$.

$$\begin{aligned} \frac{dt}{dx} &= \frac{1}{2v}.\left(H^2+x^2\right)^{-\frac{1}{2}}.(2x) + \frac{1}{2v}.\left(h^2+b^2-2bx+x^2\right)^{-\frac{1}{2}}.(-2b+2x) \\ &= \frac{x}{v.\sqrt{H^2+x^2}} + \frac{x-b}{v.\sqrt{h^2+b^2-2bx+x^2}} \end{aligned}$$

At this point we should note that $\frac{dt}{dx}$ is not defined when $\sqrt{H^2+x^2} = 0$ or when $\sqrt{h^2+b^2-2bx+x^2} = 0$. However, since $\sqrt{H^2+x^2} = D_1$ and $\sqrt{h^2+b^2-2bx+x^2} = D_2$, and both D_1 and D_2 are non-zero for our scenario, we can continue.

In order to minimize time, $\frac{dt}{dx} = 0$. Thus:

$$\frac{x}{v.\sqrt{H^2+x^2}} = -\frac{x-b}{v.\sqrt{h^2+b^2-2bx+x^2}}$$

i.e. $$\frac{x}{v.\sqrt{H^2+x^2}} = \frac{b-x}{v.\sqrt{h^2+b^2-2bx+x^2}} \quad \ldots (1)$$

From the previous diagram: $\sin\alpha = \dfrac{x}{D_1} = \dfrac{x}{\sqrt{H^2+x^2}}$

And: $\sin\beta = \dfrac{b-x}{D_2} = \dfrac{b-x}{\sqrt{h^2+(b-x)^2}} = \dfrac{b-x}{\sqrt{h^2+b^2-2bx+x^2}}$

It thus follows that $\dfrac{\sin\alpha}{v} = \dfrac{\sin\beta}{v}$ and hence $\sin\alpha = \sin\beta$. Since both α and β are acute angles, it follows that $\alpha = \beta$. We have thus shown that the angle of incidence equals the angle of reflection.

For the special case where $H = h$, i.e. where points A and B are at the same perpendicular distance from the reflecting surface, equation (1) becomes:

$$\frac{x}{v.\sqrt{H^2+x^2}} = \frac{b-x}{v.\sqrt{H^2+b^2-2bx+x^2}}$$

Multiplying through by v and squaring both sides gives:

$$\frac{x^2}{H^2+x^2} = \frac{b^2-2bx+x^2}{H^2+b^2-2bx+x^2}$$

The same argument holds as before regarding non-zero denominators, so we can multiply through by the LCD, thus giving:

$$x^2\left(H^2+b^2-2bx+x^2\right) = \left(H^2+x^2\right)\left(b^2-2bx+x^2\right)$$

Multiplying out gives:

$$H^2x^2+b^2x^2-2bx^3+x^4 = b^2H^2-2bH^2x+H^2x^2+b^2x^2-2bx^3+x^4$$

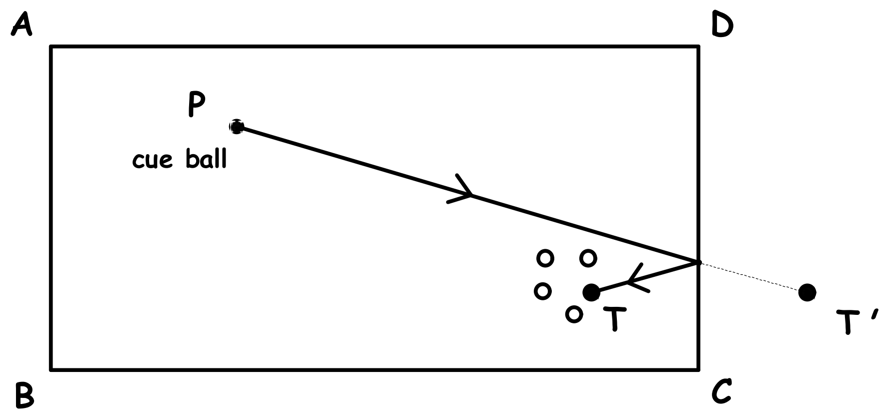
Canceling equivalent terms leaves $2bH^2x = b^2H^2$ which simplifies to $x = \dfrac{b}{2}$.

For the specific case where points A and B are at equal perpendicular distances from the reflecting surface, the path which takes the shortest time is the one for which $x = \dfrac{b}{2}$, i.e. where point C is the midpoint of horizontal length b. From symmetry/congruency considerations it follows that once again $\alpha = \beta$.

The Law of Reflection has further practical application in the game of snooker. If we ignore the effects of spin and friction, then a ball will rebound off the side of a snooker table at the same angle that it hit the side. This allows for a useful strategy for getting out of tight situations.

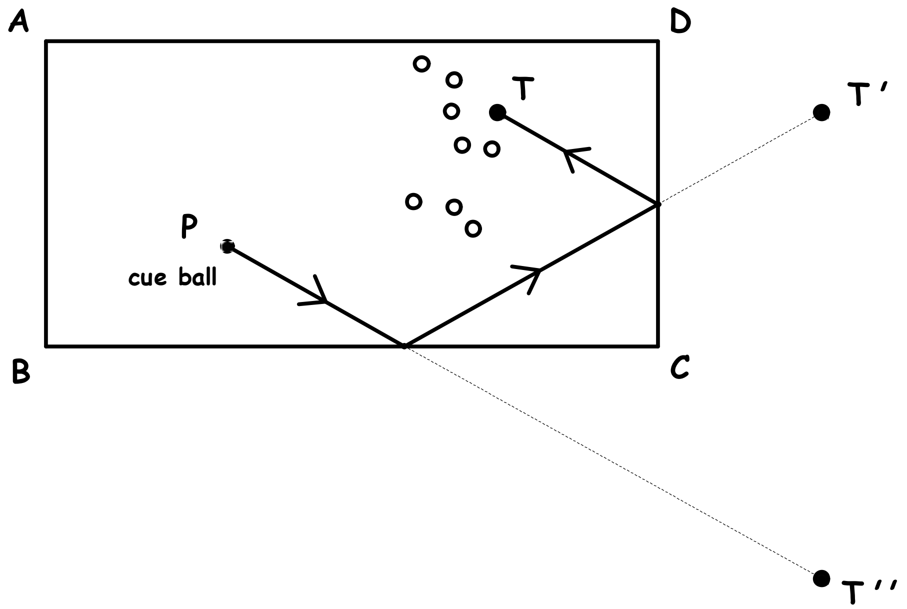
Imagine the following setup where P is the cue ball and T is the target ball which the cue ball needs to strike:

By reflecting T in the line DC we create T', the mirror image of T. If we hit the cue ball in the direction of T' it should bounce off the end cushion DC straight towards the target ball T.

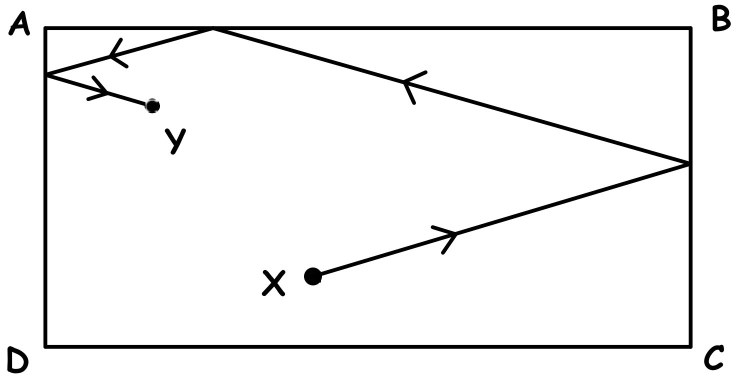


For more complex situations where the cue ball needs to bounce off two or more cushions, the principle of reflection can simply be extended. Consider the following setup where the cue ball P needs to strike the target ball T:

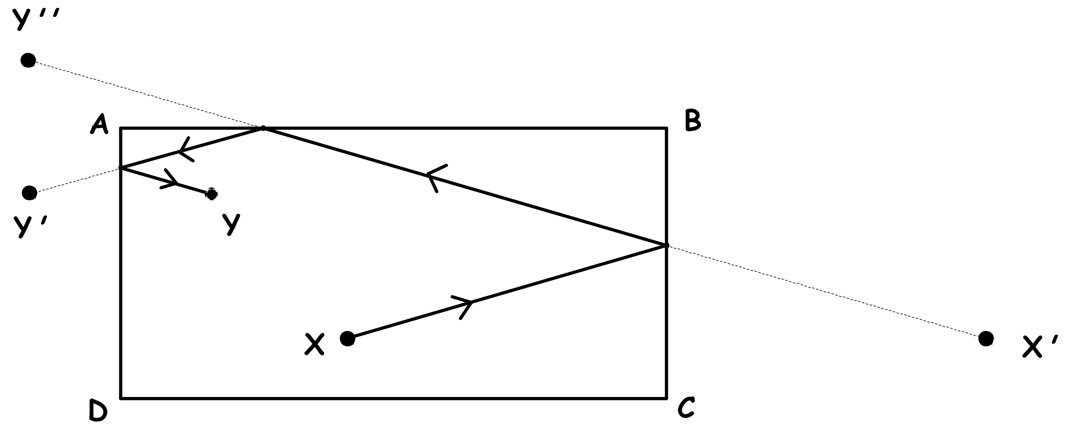
Working backwards, the cue ball needs to bounce off DC towards T. It must therefore travel towards DC in the direction of T', where T' is the mirror image of T in DC. In order to do this the cue ball must travel towards BC in the direction of T'', where T'' is the mirror image of T' in BC.



Here's a slight variation on the theme. Imagine a 2 m by 1 m rectangular pool table. A ball at X is hit and eventually comes to rest at Y. The path of the ball is shown in the following diagram:

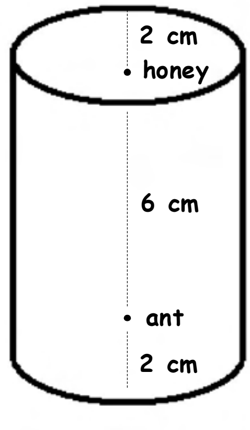


X is 0,6 m from AD and 0,2 m from DC. Y is 0,2 m from AD and 0,25 m from AB. If the ball always rebounds at the same angle as it hits the side, what is the total distance traveled by the ball from X to Y? If Y' is the reflection of Y in AD, Y'' is the reflection of Y' in AB, and X' is the reflection of X in BC, then the path of the ball from X' to Y'' is a straight line (the hypotenuse of a right-angled triangle).



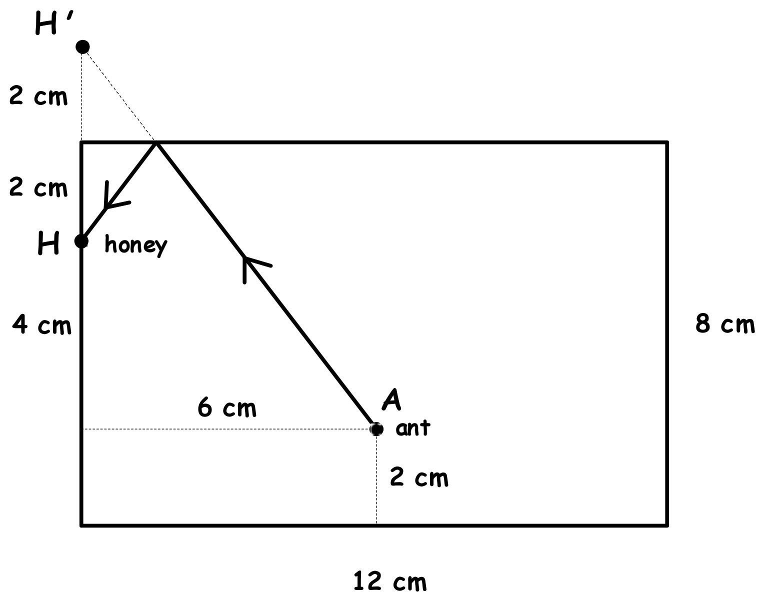
The total distance traveled from X' to Y'' parallel to AB is 1,4 m (to BC) plus 2 m (to AD) plus 0,2 m (to Y''), making 3,6 m. Similarly, the total distance traveled from X' to Y'' parallel to AD is 0,8 m (to AB) plus 0,25 m (to Y''), making 1,05 m. Using Pythagoras, the straight-line distance from X' to Y'' is 3,75 m. The ball thus travels 3,75 m on its path from X to Y.

Let us finish by looking at a classic question of its type. Imagine a cylindrical glass, 8 cm high and 12 cm in circumference. On the inside of the glass, 2 cm from the top, is a drop of honey. On the diametrically opposite outside surface, 2 cm from the bottom of the glass, is an ant. What is the shortest possible path by which the ant could walk to the honey?



As mentioned at the beginning, finding geodesics on a curved surface can often be difficult. However, the length of a path on the surface of a cylinder is not changed if the curved surface is flattened. By unrolling the cylinder and flattening it into a rectangle, a single reflection allows us to determine the ant's path.

The distance from A to H' is the hypotenuse of a 6, 8, 10 right-angled triangle. The shortest path from the ant to the honey is thus 10 cm.



References

Bolt, B. (1982). Mathematical Activities: a resource book for teachers. Cambridge: Cambridge University Press, pp. 75 – 76.

Chin, L. C. (2004). *Fermat's principle of least time*. Retrieved March 4, 2006 from http://www.geocities.com/theophoretos/fermat.html

Gardner, M. (1987). The 2nd Scientific American Book of Mathematical Puzzles & *Diversions*. Chicago, Ill.: University of Chicago Press, pp. 36 – 42.

Higgins, P. M. (1998). *Mathematics for the Curious*. Oxford: Oxford University Press, pp. 129 – 133.

Kinsey, L. C. and Moore, T. E. (2002). Symmetry, Shape, and Space: An Introduction to *Mathematics Through Geometry*. Emeryville, CA: Key College Pub., pp. 20 – 29.

Kraitchik, M. (1943). *Mathematical Recreations*. London: Allen & Unwin, pp. 17 – 21.

Serway, R. A. (1990). Physics for Scientists and Engineers with Modern Physics (3rd ed.). Philadelphia: Saunders College Publishing, pp. 988 – 1003.

Stewart, J. (1991). *Calculus* (2nd ed.). California: Brooks/Cole.

Wickelgren, W. W. (1974). *How to Solve Mathematical Problems*. New York: Dover, pp. 161 – 166.