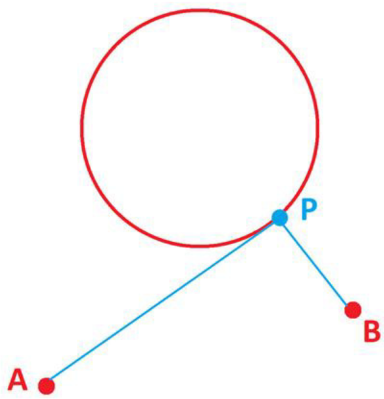
Minimum Path Via Circle

29 August 2022

Jim Stevenson



James Tanton provides another imaginative problem[1] on Twitter.

I am at point A and want to walk to point B via some point, any point, P on the circle. What point P should I choose so that my journey A → P → B is as short as possible?

Hint: I got ideas for a solution from two of my posts, "Square Root Minimum"[2] and "Maximum Product".[3]

My Solution

Just as in the "Square Root Minimum" problem, the sum of distances to a point from two fixed points reminded me of ellipses, where the sum of distances equals the major axis $2a$. Only in this case the ellipses are not allowed to shrink to as small as they can get. They must always intersect the circle. Figure 1 shows a set of ellipses with major axes approximately ranging from 11 to 23 in increasing increments (the figure is not entirely accurate). Clearly we can see the smallest sum of distances (smallest ellipse) is the one just tangent to the circle. So this is the qualitative answer to the problem.

Lagrange Mulitipliers. Figure 2 shows a quantitative approach based on Lagrange multipliers suggested by the solution to the "Maximum Product" problem. We need to establish a coordinate system to parameterize the problem. Assume the circle is centered at the origin (0, 0) and has radius 4. Then its equation is given by $g(x, y) = x^2 + y^2 - 16 = 0$. The point P is given the coordinates (x, y)

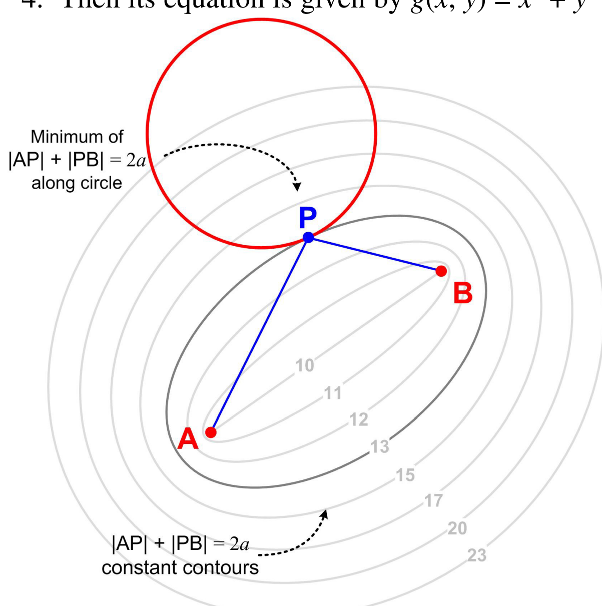


Figure 1

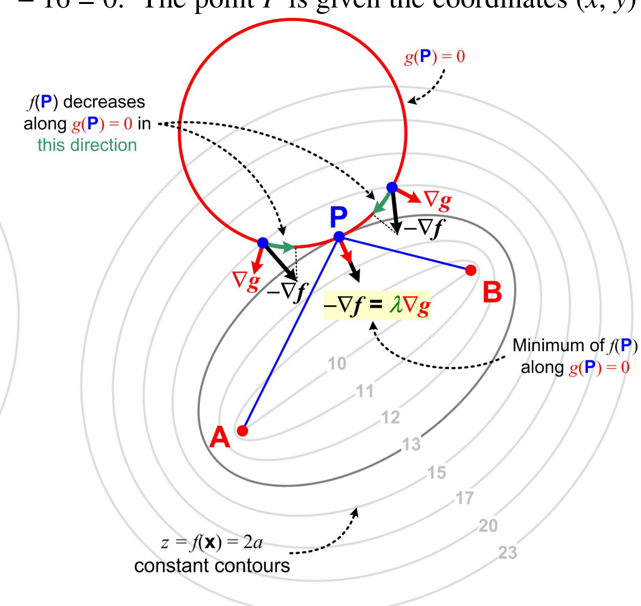


Figure 2 Lagrange Multipliers

[1] https://twitter.com/jamestanton/status/1562752143727251456, 25 August 2022
[2] https://josmfs.net/2022/09/03/square-root-minimum/
[3] http://josmfs.net/2019/08/10/maximum-product/

on the circle. The foci A and B are given the coordinates (a_1, a_2) and (b_1, b_2), respectively. Then the elliptical constant contours are given by

$$f(x, y) = \sqrt{(x - a_1)^2 + (y - a_2)^2} + \sqrt{(x - b_1)^2 + (y - b_2)^2} = \text{constant} = 2a \tag{1}$$

where a is the semi-major axis for the given elliptical contour.

As shown in Figure 2, the minimum for f constrained to $g(x, y) = 0$ occurs when the gradients of f and g are parallel, that is,

$$-\nabla f = \lambda \nabla g \tag{2}$$

for some constant λ, called the Lagrange multiplier. Note that we are using the negative gradient for f, since we are looking for the minimum and the gradient points in the direction of maximum increase. So the negative gradient points in the direction of maximum decrease.

Alternative Representation. An alternative quantitative model for the problem is provided when the equation of the circle is given parametrically via a vector-valued function of a real number, in this case the angle. That is,

$$\mathbf{P}(\theta) = 4 \cos \theta \; \mathbf{i} + 4 \sin \theta \; \mathbf{j}$$

Then f restricted to the circle is given by the real-valued function of a real variable as $f(\mathbf{P}(\theta))$. So the minimum can be found by taking the (negative) real derivative and setting it to zero.

$$-df(\mathbf{P}(\theta))/d\theta = -\nabla f \cdot \mathbf{P}'(\theta) = 0.$$

As shown in Figure 3, $\mathbf{P}'(\theta)$ is a tangent vector to the circle. So the vanishing of the derivative occurs when the negative gradient of f is perpendicular to the tangent to the circle. This of course is equivalent to the Lagrange multiplier formulation. (Figure 3 shows $-df(\mathbf{P}(\theta))/d\theta = -\nabla f \cdot \mathbf{P}'(\theta)$ as the projection of $-\nabla f$ onto $\mathbf{P}'(\theta)$. Technically this is only true if $\mathbf{P}'(\theta)$ is a unit tangent vector, namely if we divide $\mathbf{P}'(\theta)$ by the radius 4.)

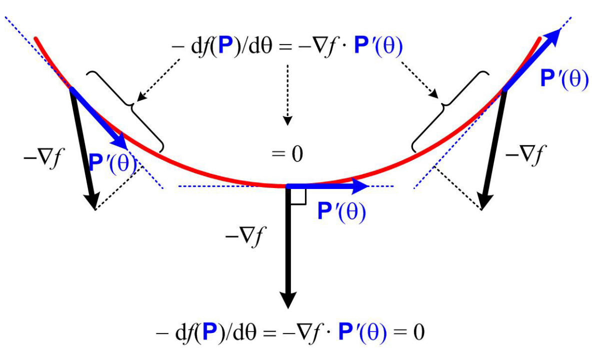


Figure 3 Vanishing Derivative at Extreme Point

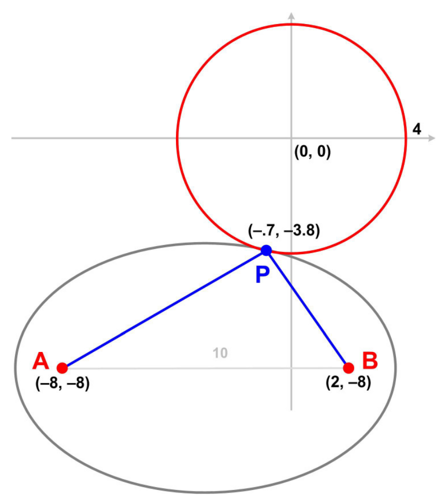


Figure 4

Computation. I thought I would try to compute the result for a given case. As shown in Figure 4, I rotated the problem around the unit circle until the line AB was parallel to the x-axis. I chose coordinates (–8, –8) for point A and (2, –8) for B. I filled these values into equation (1) and computed the gradients for equation (2). This was not hard, but trying to solve equation (2) involved some complicated radicals that I could not see how to simplify sufficiently to obtain the value for λ and use it to obtain the desired point $P(x, y)$. I tried the alternative formulation with the trig functions, but that only increased the nightmare.

I then carefully plotted the example in Visio and obtained an approximate solution of $P(-.7, -3.8)$. I estimated the major axis of $2a = 13.26$ from the resulting ellipse plot. When I plugged $P(-.7, -3.8)$ into equation (1), I got 13.41 as the major axis. This error was within my ability to estimate ± a couple of tenths.

Related Function. The typical way of calculating problems of distances involving radicals is to square the radicals. It is true that the minimum distance of P to A as P ranges over the circle will occur at the same point where the minimum of the square of the distance will occur. But this problem involves the sum of *two* distances. Taking the sum of the squares of the two distances is a different problem, reminiscent of a least squares problem. That type of problem weights the fitting toward minimizing the larger distances, in this case, the distance of P to A versus the distance of P to B. So we would expect the minimum point to move to the left along the circle. That problem is actually easily solvable.

Instead of equation (1), let f be defined by the sum of squares

$$f(x, y) = (x + 8)^2 + (y + 8)^2 + (x - 2)^2 + (y + 8)^2 \tag{3}$$

Then

$$f(x, y) = 2x^2 + 2y^2 + 12x + 32y + 196$$

The gradients are easily computed:

$$-\nabla f = -(2x + 12)\mathbf{i} - (2y + 32)\mathbf{j} = \lambda\nabla g = \lambda(2x\mathbf{i} + 2y\mathbf{j})$$

Therefore

$$x = -6/(\lambda + 2), \quad y = -16/(\lambda + 2)$$

Substituting these results into $g(x, y) = 0$ yields

$$16 = x^2 + y^2 = (36 + 256)/ (\lambda + 2)^2$$

Therefore,

$$\lambda + 2 = \pm \sqrt{(292/16)} = \pm 4.27$$

So

$$x = -6/4.27 = -1.41,$$

$$y = -16/4.27 = -3.75$$

As expected, this minimum solution is a little bit to the left of the solution to the original problem via Visio. –3.75 is within the approximation error for –3.8 from Visio, but –1.41 is not for –0.7.

Reflected Rays. Recall the ellipse property that the lines from the foci to the tangent point of the ellipse make equal angles with the tangent (and therefore with the normal)—"angle of incidence equals angle of reflection." This property is noted in the Twitter solutions below.

Twitter Solutions

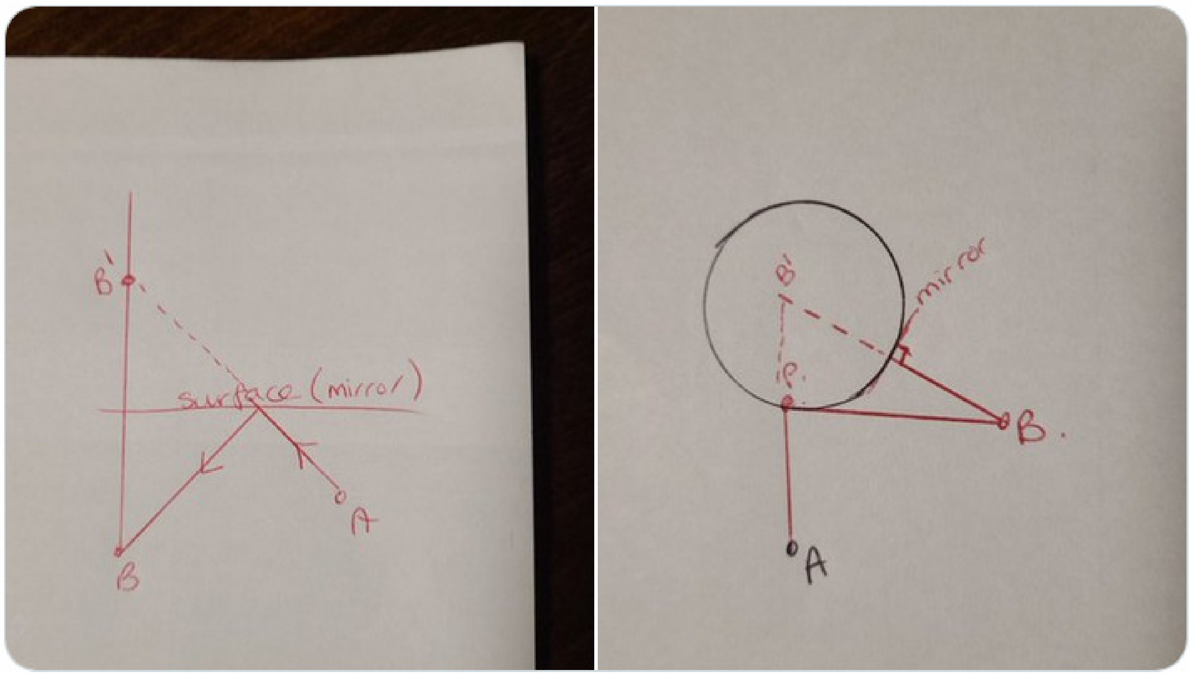
These were comments added to Tanton's Tweet (as of 8/29/2022). My ideas to various degrees are reflected in some of the remarks below.

(I have reformatted the comments and omitted some that were not germane.)

Christine Lenghaus @Lenghaus Aug 25

(https://twitter.com/Lenghaus/status/1562757948778381313)

Mirror (along tangent at P) B inside the circle. Draw line from A to B’ and where it hits circle is P.



7:05 AM · Aug 25, 2022·Twitter for Android

A Dog On The Internet @ADogOnTheNet Aug 25
(https://twitter.com/ADogOnTheNet/status/1562775264613912578)

This doesn’t look right. What if the image of B puts B’ outside of the circle entirely? My intuition is that you need to map B so that its distance from the center changes from kR to 1/k R.

Or maybe better still, find the conformal mapping from the circle to a straight line and solve the problem as for a flat mirror.

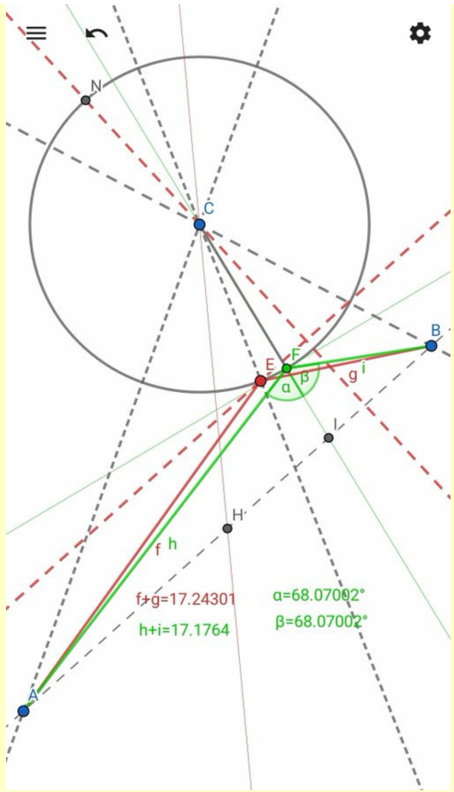
Colin V. Parker @ColinVParker Aug 25
(https://twitter.com/ColinVParker/status/1562777560483635200)

I believe that the ellipse through P having A and B as foci should be tangent to the circle. This leads to the conclusion that AP and PB might not be perpendicular.

Calling C the circle’s center, I would put P on the bisector of CA and CB?

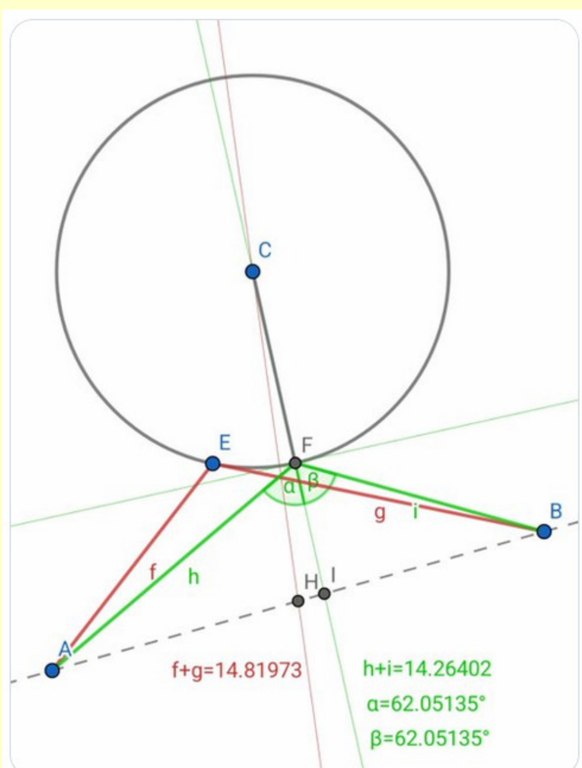
Carlos Cabral @cesscabral Aug 25
(https://twitter.com/cesscabral/status/1562875043482603521)

Not really.

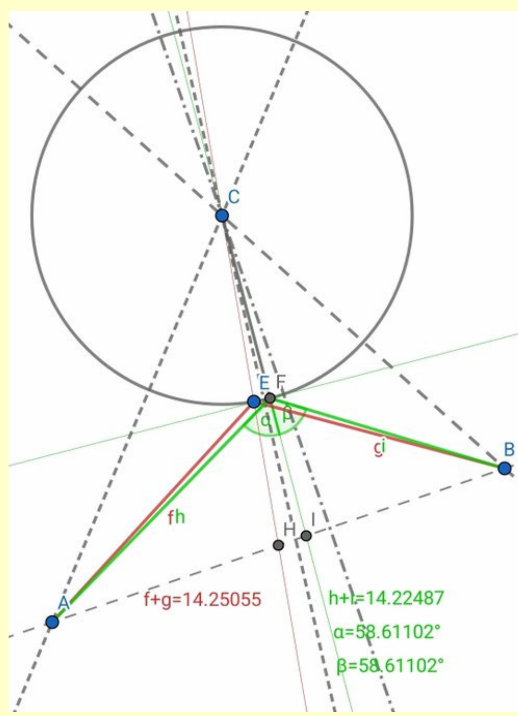


Carlos Cabral **Aug 25** (https://twitter.com/cesscabral/status/1562834553718206465)

It's the point F where we see A and B with the same angle against the radius



This point is better than any other point like the ones defined by the perpendicular to AB or the bisector of ABC or the median of AB from C



Alfred Simpson @alfredsimpson Aug 25
(https://twitter.com/alfredsimpson/status/1562841134891536384)

The angle at P should be 120 degrees.

Carlos Cabral @cesscabral Aug 25 (https://twitter.com/cesscabral/status/1562874580989263873)

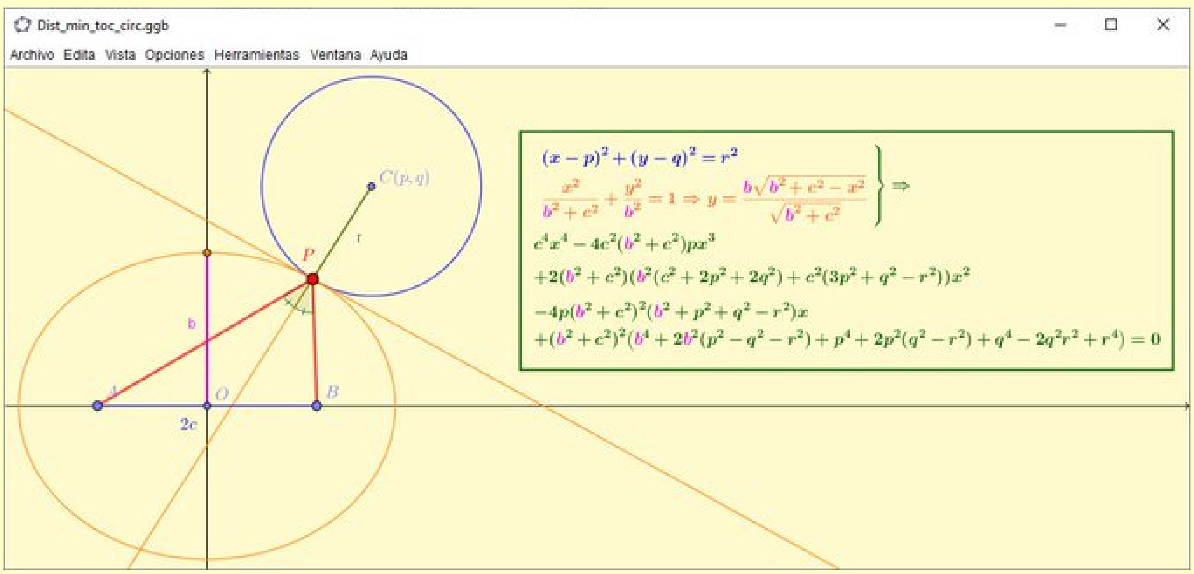
If points A and B are closer to each other than to the circle, α + β would be much less than 120°

Dr PD Prideaux @dr_prideaux Aug 25
(https://twitter.com/dr_prideaux/status/1562936760006594560)

The angle of incidence equals the angle of reflection?

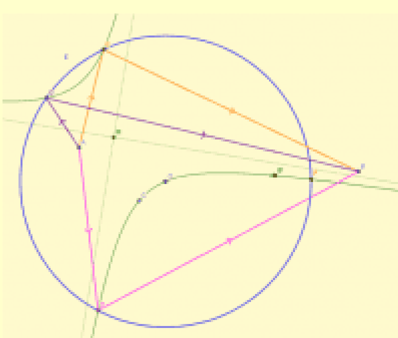
Ignacio Larrosa Cañestro @ilarrosac Aug 26
(https://twitter.com/ilarrosac/status/1563095565692715009)

P must be the point of contact of the ellipse with foci A and B externally tangent to the ◯. Using coordinates, we obtain a 4th degree equation in (x, y) that must have a unique solution. This leads to an equation again of degree 4 on the secondary semi-axis b of the ellipse.



Ignacio Larrosa Cañestro @ilarrosac Aug 26
(https://twitter.com/ilarrosac/status/1563099799875776513)

Only in simple cases I can get a 4th degree equation in b. In that cases, I get the correct solution by this method, coinciding with the solution foreseen with the naked eye. I think it is equivalent to Alhazen's problem, unsolvable with ruler and compass.



geogebra.org (https://t.co/KTn3efRteA)
(https://www.geogebra.org/m/qexheksq)

El problema de Alhazen

Problema de Alhazen, de reflexión en una circunferencia. En general no puede resolverse con regla y compás.

Almog Yalinewich @yalinewich Aug 27
(https://twitter.com/yalinewich/status/1563394146458431494)

wlog circle centre = origin and radius = 1

minimise AP + PB subject to P on circle using Lagrange multipliers

$$D = (P - A)^2 + (P - B)^2 + \lambda(P^2 - 1)$$

derive [differentiate] w.r.t P and equate to zero

$$P = (A + B)/(\lambda + 2)$$

determine λ from constraint $P = (A + B)/|A + B|$

Kate Rosengreen @rosengreenkate Aug 27
(https://twitter.com/rosengreenkate/status/1563480495463141379)

The point at which the angle APB is maximised?