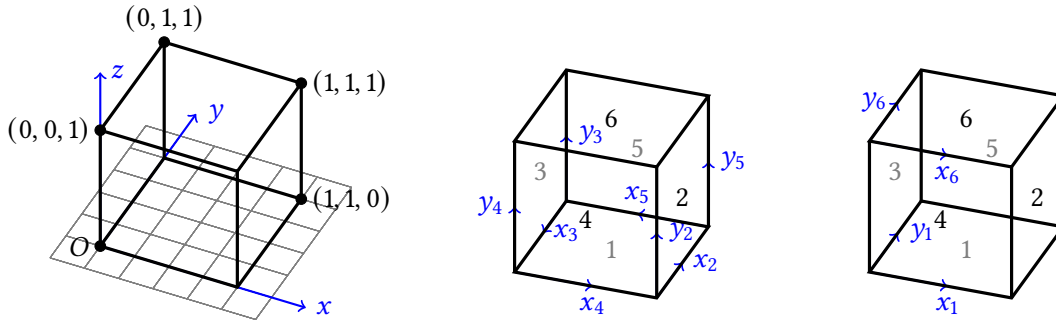# GEODESICS ON A CUBE AND A CAPPED CYLINDER

Heikki von Koch

1 GEODESICS ON A CUBE

We start our examples with a cube (surface) of side length 1 with the origin of $\mathbb{R}^3$ and labelling of its faces as indicated by the pictures. For the standard $\mathbb{R}^3$-coordinates, we always use the triple $(\cdot, \cdot, \cdot)$. In addition, each face F has their own fixed local two-dimensional coordinate system denoted by $(x_F, y_F)_F$, where $F \in \{1, 2, 3, 4, 5, 6\}$ and $x_F, y_F \in [0, 1]$ (we shall often just use the shorthand notation $(x, y)_F$). Their orientations are shown above.

As an example, we have that $(0, 0)_1 = (1, 0)_3 = (0, 0)_4$ and $(1, 0)_6 = (0, 1)_2 = (1, 1)_4$ and, in fact, we always have $(x, y)_1 = (x, y, 0)$ and $(x, y)_6 = (x, y, 1)$ (permanently fixing the vertices).

1.1 ANATOMY OF GEODESICS ON A CUBE

All geodesics on the cube are straight lines when the cube is unfolded in any manner but not all straight lines in the unfolding are geodesics as, for example, some go through a vertex and some appear to go through a face onto the same face. The important fact is that the globally shortest straight lines are always geodesics. There are effectively three different places where the geodesic might end with respect to the starting point:

1. the endpoint is on the same face;

2. the endpoint is on an adjacent face;

3. the endpoint is on the opposite face.

Geodesics of the 1. kind are obviously regular straight lines in the plane and need no further analysis.

Geodesics of the 2. kind have three options: going through the shared side or going through either of the faces that also share a side with the face of the endpoint. For example, let the starting point be on face 1 and the endpoint on face 2, then we have the three options: $1 - 2$, $1 - 5 - 2$, and $1 - 4 - 2$, where $a - b - c$ means that we start from face a, move to face b, and move to face c where we stop.

Geodesics of the 3. kind have 12 options: for each adjacent face, there are three options (as can be seen from the 2. kind). Note that not all of these will necessarily be actual geodesics with the same being true for geodesics of the 2. kind. However, at least one of them will be a geodesic as well as realise the globally shortest path between the start and endpoints.

Now all we need are the face-to-face transformations between local coordinate systems, which can then be composed together to make longer chains between the faces. A coordinate transformation from one face to another is made by rotating (with respect to the common edge) the target face onto the same plane as the starting face (keeping this face fixed) such that there is no overlap. For the unfoldings of the different coordinate grids, we essentially need two transformations of a point (x, y) on the plane:

- Rotation of (x, y) about a point (a, b) for $+90°$: $(x, y) \rightarrow (b + a - y, b - a + x)$;

- Rotation of (x, y) about a point (a, b) for $-90°$: $(x, y) \rightarrow (a - b + y, a + b - x)$.

The following matrix contains all of the coordinate transformations between the faces with $x, y \in [0, 1]$ and their standard three-dimensional coordinates.

$\mathbb{R}^3$	$(x, y, 0)$	$(1, x, y)$	$(0, 1 - x, y)$	$(x, 0, y)$	$(1 - x, 1, y)$	$(x, y, 1)$
	1	2	3	4	5	6
1	$(x, y)_1$	$(y + 1, x)_1$	$(-y, 1 - x)_1$	$(x, -y)_1$	$(1 - x, y + 1)_1$	
2	$(y, x - 1)_2$	$(x, y)_2$		$(x - 1, y)_2$	$(x + 1, y)_2$	$(y, 2 - x)_2$
3	$(1 - y, -x)_3$		$(x, y)_3$	$(x + 1, y)_3$	$(x - 1, y)_3$	$(1 - y, x + 1)_3$
4	$(x, -y)_4$	$(x + 1, y)_4$	$(x - 1, y)_4$	$(x, y)_4$		$(x, y + 1)_4$
5	$(1 - x, y - 1)_5$	$(x - 1, y)_5$	$(x + 1, y)_5$		$(x, y)_5$	$(1 - x, 2 - y)_5$
6		$(2 - y, x)_6$	$(y - 1, 1 - x)_6$	$(x, y - 1)_6$	$(1 - x, 2 - y)_6$	$(x, y)_6$

Note the abundance of regularity in the table because of the way the local coordinate systems were chosen. The side faces all share the y-coordinate, and the x-coordinate always has the same orientation, that is, the coordinate transformations from one side face to the next when going counterclockwise (or clockwise, for that matter) are the same.

1.2 COORDINATE CALCULATIONS

The distances can now be easily calculated with the Pythagorean theorem. For example, let our starting point $(x_1, y_1)_1$ be on face 1 and endpoint $(x_2, y_2)_2$ on face 2. First we need to figure out the three different possible coordinates of $(x_2, y_2)_2$ in the coordinate system of face 1 by composing the transformations together using Table 1:
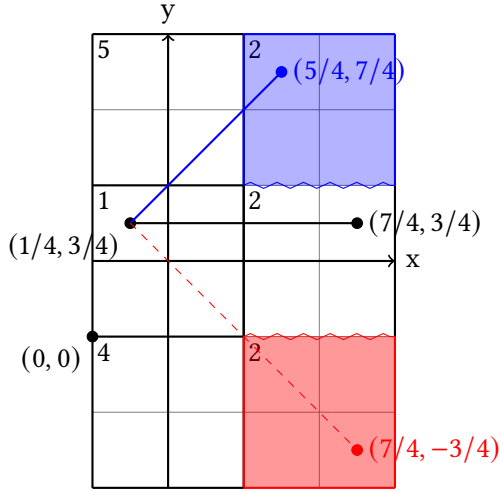
1. Path $1 - 2$: $(x, y)_2 \rightarrow (y + 1, x)_1$;

2. Path $1 - 5 - 2$: $(x, y)_2 \rightarrow (x - 1, y)_5 \rightarrow (2 - x, y + 1)_1$;

3. Path $1 - 4 - 2$: $(x, y)_2 \rightarrow (x + 1, y)_4 \rightarrow (x + 1, -y)_1$.

This gives us three possible shortest paths between them with the distances

1. Path $1 - 2$: $d((x_1, y_1)_1, (x_2, y_2)_2) = \sqrt{(x_1 - (y_2 + 1))^2 + (y_1 - x_2)^2}$;

2. Path $1 - 5 - 2$: $d((x_1, y_1)_1, (x_2, y_2)_2) = \sqrt{(x_1 - (2 - x_2))^2 + (y_1 - (y_2 + 1))^2}$;

3. Path $1 - 4 - 2$: $d((x_1, y_1)_1, (x_2, y_2)_2) = \sqrt{(x_1 - (x_2 + 1))^2 + (y_1 + y_2)^2}$.

Let $(x_1, y_1)_1 = (1/2, 1/2)_1$ and the endpoint similarly be the middle point of face 2, that is, $(1/2, 1/2)_2$. Clearly the shortest distance between the start and endpoint is 1. The distance formulas give us 1, $\sqrt{2}$, $\sqrt{2}$, respectively, that is, the shortest path is rather unsurprisingly through the common side (in fact, the other paths aren't geodesics at all since they go through vertices as can be seen from drawing a simple picture).

Lets change our starting point to $(1/4, 3/4)_1$ and our endpoint to $(3/4, 3/4)_2$. Now we have the distances $3/2$, $\sqrt{2}$, $\sqrt{18}/2$, respectively, that is, the shortest path is through face 5.

An unfolding of the cube with the labels of the faces marked on their upper left corners. Included are the three colour-coded paths from face 1 to face 2 with starting point $(1/4, 3/4)_1$ and endpoint $(3/4, 3/4)_2$, and their endpoints in the coordinate system of face 1. The dashed red path is not a geodesic as it goes through the vertex $(1, 0)_1$. Note that the two zigzag edges between faces 2 are not real but merely artifacts of the drawing.

2 GEODESICS ON A CAPPED CYLINDER

A cylinder (more specifically, a right circular cylinder) is a surface defined as the set of points $x, y, z \in \mathbb{R}$ such that $(x, y, z) = (r \cos \theta, r \sin \theta, z) = (\theta, z)$, where $r > 0$ is the radius of the circle defined by x and y. Furthermore, we will consider the capped cylinder by adding circular discs of the same radius r to the top and bottom of the cylinder.

2.1 GEODESICS ON EITHER THE SIDE OR THE TOP AND BOTTOM DISCS

Geodesics between points on the same discs are clearly straight lines and one can check (using the geodesic equation) that geodesics on the side are helices (in the degenerate cases, these include straight lines and arcs of circles). A helix is a space curve parametrised by

$$x(\varphi) = r \cos \varphi, \quad y(\varphi) = r \sin \varphi, \quad z(\varphi) = c\varphi, \quad \text{for } \varphi \in [0, d],$$

where c is a constant such that $2\pi c$ gives the vertical separation of the helix's loops and d a parameter depending on c, such that we stay on the side of the cylinder (so not in the interior of top or bottom disc). Instead of vertical separation, one could also say that the helix has a slope of r/c. In fact, this is its defining feature: a curve for which the tangent makes a constant angle with a fixed line (center axis for our cylinder). The helix between two points $p_1 = (z_1, \varphi_1)$ and $p_2 = (z_2, \varphi_2)$ is given by

$$\gamma(\varphi) = \left(r \cos \varphi, r \sin \varphi, \frac{z_1 - z_2}{\varphi_1 - \varphi_2} \varphi + \frac{\varphi_1 z_2 - \varphi_2 z_1}{\varphi_1 - \varphi_2} \right)$$

and its arc length L can be calculated with $L = d\sqrt{r^2 + c^2}$.

Example 2.1 (No vertical separation between loops). Since there is no separation between loops, $2\pi c = 0 \implies c = 0$, meaning that $z(\varphi) = 0$ for all φ, i.e., the z-coordinate does not change and thus the parametrisation is that of an arc of a circle of radius r. The arc length is naturally $L = dr$.

Example 2.2 (Zero angle with central axis). Since the angle is zero, we have that $\tan(0) = r/c = 0 \implies r = 0$, meaning that $x(\varphi) = y(\varphi) = 0$ for all φ, i.e., the parametrisation is that of a straight line, where the points only differ by their z-coordinates. The arc length is now $L = dc$.

Example 2.3 (Existence of a shortest path and arbitrary long paths). Between any two points on the cylinder that, w.l.o.g., only differ in their z-coordinate, there exist infinitely many geodesics. Furthermore, there is no upper bound for their length but there does exist at least one shortest one. By controlling the angle φ (or c), we control how many rotations we take around the cylinder. For $n \geq 0$ rotations we have by Pythagoras (by unfolding the cylinder n times and excluding the top and bottom), that $\tan(\varphi) = nr/c \iff \varphi = \arctan(nr/c)$. The shortest path is now with zero rotations while there is no longest path as we can do as many rotations as we which: $n \to \infty \implies \varphi \to \pi/2 \implies L \to \infty$.

For now, we have only discussed geodesics lying strictly on the curved side of the cylinder or on the flat top and bottom discs. What about when we go on two or more of them?

2.2 SIDE -> DISC GEODESICS

Let p_1 be a point on the side and p_2 a point on one of the discs. Denote the point where the geodesic crosses the edge by x; we now have a picture where the side of the cylinder is unfolded and the disc is touching this at the point x. The point x needs to be chosen such that the line $p_1 x p_2$ is straight (in the unfolded picture), i.e., we have

$$d(p_1, p_2) = \min_x \left\{ d_{\text{cyl}}(p_1, x) + d_{\text{disc}}(x, p_2) \right\}.$$

Denote $p_1 = (z_1, \varphi_1)$, $\angle \tilde{p}_1 O x = \alpha_1$, where $\tilde{p}_1$ is the point on the boundary of the disc corresponding to p_1 (same phase), and $p_2 = (r_2, \varphi_2)$, $\angle p_2 O x = \alpha_2$, where $d_{\text{disc}}(O, p_2) = r_2$. Assume further that $\varphi_2 > \varphi_1$ (mod 2π). We have

$$d_{\text{cyl}}(p_1, x) = \sqrt{z_1^2 + r^2 \alpha_1^2} \quad \text{and} \quad d_{\text{disc}}(x, p_2) = \sqrt{r^2 + r_2^2 - 2rr_2 \cos \alpha_2},$$

where the first is by Pythagoras (the non z-coordinate side of the triangle is simply the arc length $r\alpha_1$ of the circle) and the second by the cosine rule. Since $\alpha_2 = \varphi_2 - \varphi_1 - \alpha_1$, we get

$$d(p_1, p_2) = \min_{\alpha_1} \left\{ \sqrt{z_1^2 + r^2 \alpha_1^2} + \sqrt{r^2 + r_2^2 - 2rr_2 \cos(\varphi_2 - \varphi_1 - \alpha_1)} \right\}.$$

To calculate its minima, set $f(\alpha_1; p_1, p_2, r) = \sqrt{z_1^2 + r^2 \alpha_1^2} + \sqrt{r^2 + r_2^2 - 2rr_2 \cos(\varphi_2 - \varphi_1 - \alpha_1)}$ and calculate its derivatives as

$$D_{\alpha_1} f = \frac{r^2 \alpha_1}{\sqrt{z_1^2 + r^2 \alpha_1^2}} - \frac{rr_2 \sin(\varphi_2 - \varphi_1 - \alpha_1)}{\sqrt{r^2 + r_2^2 - 2rr_2 \cos(\varphi_2 - \varphi_1 - \alpha_1)}}$$

and

$$D_{\alpha_1}^2 f = \frac{r^2 z_1^2}{(z_1^2 + r^2 \alpha_1^2)^{3/2}} + \frac{rr_2 (r \cos(\varphi_2 - \varphi_1 - \alpha_1) - r_2)(r - r_2 \cos(\varphi_2 - \varphi_1 - \alpha_1))}{(r^2 + r_2^2 - 2rr_2 \cos(\varphi_2 - \varphi_1 - \alpha_1))^{3/2}}.$$

2.3 DISC -> SIDE -> DISC GEODESICS

Without loss of generality, let p_1 be a point on the bottom disc and p_2 on the top disc. Denote the points where the geodesic crosses the edges by x_1 and x_2, respectively; we now have a picture where the cylinder side is unfolded and the discs are touching the side at the points x_1 and x_2. These points need to be chosen such that the line $p_1x_1x_2p_2$ is straight (in the unfolded picture), i.e., we have

$$d(p_1, p_2) = \min_{(x_1, x_2)} \left\{ d_{\text{disc}}(p_1, x_1) + d_{\text{cyl}}(x_1, x_2) + d_{\text{disc}}(x_2, p_2) \right\}.$$

Denote $p_i = (r_i, \varphi_i)$, $\angle p_i O x_i = \alpha_i$, for $i = 1, 2$, where $d_{\text{disc}}(O_i, p_i) = r_i$. Since the line $p_1x_1x_2p_2$ is straight, we can define a rectangle $x_1\tilde{x}_1x_2\tilde{x}_2$ such that we have the correspondence of phases: $\varphi_{\tilde{x}_1} = \varphi_{x_2}$ and $\varphi_{\tilde{x}_2} = \varphi_{x_1}$, i.e., $\angle x_1O_1\tilde{x}_1 := \tilde{\alpha}_1 = \tilde{\alpha}_2 =: \angle x_2O_2\tilde{x}_2$. Assume further that $\varphi_2 > \varphi_1 \pmod{2\pi}$. We have

$$d_{\text{disc}}(p_1, x_1) = \sqrt{r^2 + r_1^2 - 2rr_1\cos\alpha_1}, \quad d_{\text{cyl}}(x_1, x_2) = \sqrt{h^2 + (\tilde{\alpha}_1 r)^2}, \quad \text{and} \quad d_{\text{disc}}(x_2, p_2) = \sqrt{r^2 + r_2^2 - 2rr_2\cos\alpha_2},$$

where the middle is by Pythagoras (the non z-coordinate side of the triangle is simply the arc length $r\tilde{\alpha}_i$ of a circle and h is the height of the cylinder) and the others by the cosine rule. We have

$$\tilde{\alpha}_1 = \varphi_{\tilde{x}_1} - \alpha_1 - \varphi_1 \quad \text{and} \quad \tilde{\alpha}_2 = \varphi_2 - \alpha_2 - \varphi_{\tilde{x}_2} \implies \tilde{\alpha}_1 + \tilde{\alpha}_2 = \varphi_{\tilde{x}_1} - \varphi_{\tilde{x}_2} + \varphi_2 - \varphi_1 - \alpha_1 - \alpha_2,$$

and since $\varphi_{\tilde{x}_2} = \varphi_{x_1}$ with $\varphi_{\tilde{x}_1} - \varphi_{x_1} = \tilde{\alpha}_1$, we get

$$\tilde{\alpha}_2 = \varphi_2 - \varphi_1 - \alpha_1 - \alpha_2.$$

Putting everything together gives us

$$d(p_1, p_2) = \min_{(\alpha_1, \alpha_2)} \left\{ \sqrt{r^2 + r_1^2 - 2rr_1\cos\alpha_1} + \sqrt{h^2 + ((\varphi_2 - \varphi_1 - \alpha_1 - \alpha_2)r)^2} + \sqrt{r^2 + r_2^2 - 2rr_2\cos\alpha_2} \right\}.$$

Here it becomes clear that the previous disc -> side calculation is simply a special/modified case of this one; the first term is nonexistent, $\alpha_1 = 0$, and instead of h, we have a z-coordinate for the point on the side.

Finally, let us compute some derivatives. Set

$$f(\alpha_1, \alpha_2; p_1, p_2, r, h) = \sqrt{r^2 + r_1^2 - 2rr_1\cos\alpha_1} + \sqrt{h^2 + ((\varphi_2 - \varphi_1 - \alpha_1 - \alpha_2)r)^2} + \sqrt{r^2 + r_2^2 - 2rr_2\cos\alpha_2}.$$

$$D_{\alpha_i}f = \frac{rr_i\sin(\alpha_i)}{\sqrt{r^2 + r_i^2 - 2rr_i\cos\alpha_i}} - \frac{r^2(\varphi_2 - \varphi_1 - \alpha_1 - \alpha_2)}{\sqrt{h^2 + r^2(\varphi_2 - \varphi_1 - \alpha_1 - \alpha_2)^2}}.$$

$$D^2_{\alpha_1, \alpha_2}f = \frac{r^2h^2}{(h^2 + r^2(\varphi_2 - \varphi_1 - \alpha_1 - \alpha_2)^2)^{3/2}}.$$

$$D^2_{\alpha_i}f = \frac{rr_i(r\cos\alpha_i - r_i)(r - r_i\cos\alpha_i)}{(r^2 + r_i^2 - 2rr_i\cos\alpha_i)^{3/2}} + D^2_{\alpha_1, \alpha_2}f.$$

2.4 SIDE -> DISC -> SIDE GEODESICS

Do shortest paths ever cross a disc when both of the points are on the side of the cylinder? A simple drawing shows that the answer is yes: take two points opposite (same z-coordinate but phase differing by π) each other on a cylinder of radius r. Now their distance on the side is $r\pi$ while using the disc it is $2(r + h)$, where h is the z-coordinate of the points. By controlling h and r we can thus order these lengths as we wish.

It's clear that whether this happens or not depends on both the radius of the disc and the z-coordinates and phases of the points. Let us assume that this happens and calculate the distance. Let p_1 and p_2 be two points on the side of the cylinder and x_1 and x_2 be the points where the path crosses the sides of the disc, respectively. The length to minimise is thus

$$d(p_1, p_2) = \min_{(x_1, x_2)} \{ d_{\text{cyl}}(p_1, x_1) + d_{\text{disc}}(x_1, x_2) + d_{\text{cyl}}(x_2, p_2) \}.$$

To determine whether we stay on the side or the take a shortcut through the disc, one simply compares this distance to the one strictly on the side.

Remark 2.4. This problem is a somewhat generalisation of the first problem where we would've had $p_2 = x_2$ without the assumption that the point lies on the boundary of the disc. However, here we will deal with chords, whereas there we most likely did not.

Denote $p_i = (z_i, \varphi_i)$ and $\angle \tilde{p}_i O x_i = \alpha_i$, for $i = 1, 2$, where $\tilde{p}_i$ is the point on the boundary of the disc corresponding to p_i (same phase) and assume again that $\varphi_2 > \varphi_1$. We have

$$\min_{(x_1, x_2)} \{ d_{\text{cyl}}(p_1, x_1) + d_{\text{disc}}(x_1, x_2) + d_{\text{cyl}}(x_2, p_2) \} = \min_{(\alpha_1, \alpha_2)} \left\{ \sqrt{z_1^2 + r^2 \alpha_1^2} + 2r \cos \beta + \sqrt{z_2^2 + r^2 \alpha_2^2} \right\},$$

where $\beta = 1/2(\alpha_1 + \alpha_2 + \varphi_1 - \varphi_2 + \pi)$ and $\varphi_2 > \varphi_1$. We could equivalently write $\cos \beta = \cos(1/2(\alpha_1 + \alpha_2 + \varphi_1 - \varphi_2 + \pi)) = -\sin(1/2(\alpha_1 + \alpha_2 + \varphi_1 - \varphi_2))$ and

$$\min_{(x_1, x_2)} \{ d_{\text{cyl}}(p_1, x_1) + d_{\text{disc}}(x_1, x_2) + d_{\text{cyl}}(x_2, p_2) \} = \min_{(\alpha_1, \alpha_2)} \left\{ \sqrt{z_1^2 + r^2 \alpha_1^2} - 2r \sin(1/2(\alpha_1 + \alpha_2 + \varphi_1 - \varphi_2)) + \sqrt{z_2^2 + r^2 \alpha_2^2} \right\}.$$

Set $f(\alpha_1, \alpha_2; p_1, p_2, r) = \sqrt{z_1^2 + r^2 \alpha_1^2} - 2r \sin(1/2(\alpha_1 + \alpha_2 + \varphi_1 - \varphi_2)) + \sqrt{z_2^2 + r^2 \alpha_2^2}$ and calculate its first derivatives and Hessian as

$$D_{\alpha_i} f = \frac{r^2 \alpha_i}{\sqrt{z_i^2 + r^2 \alpha_i^2}} - r \cos(\frac{1}{2}(\alpha_1 + \alpha_2 + \varphi_1 - \varphi_2))$$

and

$$\text{Hess}(f) = \begin{bmatrix} \frac{r^2 z_1^2}{(z_1^2 + \alpha_1^2 r^2)^{\frac{3}{2}}} + \frac{r}{2} \sin(\frac{1}{2}(\alpha_1 + \alpha_2 + \varphi_1 - \varphi_2)) & \frac{r}{2} \sin(\frac{1}{2}(\alpha_1 + \alpha_2 + \varphi_1 - \varphi_2)) \\ \frac{r}{2} \sin(\frac{1}{2}(\alpha_1 + \alpha_2 + \varphi_1 - \varphi_2)) & \frac{r^2 z_2^2}{(z_2^2 + \alpha_2^2 r^2)^{\frac{3}{2}}} + \frac{r}{2} \sin(\frac{1}{2}(\alpha_1 + \alpha_2 + \varphi_1 - \varphi_2)) \end{bmatrix}.$$

Example 2.5. Take opposite points on the side so that we have $\varphi_2 = \varphi_1 + \pi$ and $\alpha_1 = \alpha = \alpha_2$ and thus

$$\min_{(x_1, x_2)} \{ d_{\text{cyl}}(p_1, x_1) + d_{\text{disc}}(x_1, x_2) + d_{\text{cyl}}(x_2, p_2) \} = \min_{\alpha} \left\{ \sqrt{z_1^2 + r^2 \alpha^2} + 2r \cos \alpha + \sqrt{z_2^2 + r^2 \alpha^2} \right\}.$$

Differentiating this with respect to α and setting the derivative to zero gives us

$$\frac{r^2 \alpha}{\sqrt{z_1^2 + \alpha^2 r^2}} - 2r \sin \alpha + \frac{r^2 \alpha}{\sqrt{z_2^2 + \alpha^2 r^2}} = 0.$$

This holds, at least, when $\alpha = 0$. To see whether this is a maximum or a minimum, we take another derivative

$$\frac{r^2 z_1^2}{(z_1^2 + \alpha^2 r^2)^{\frac{3}{2}}} - 2r \cos \alpha + \frac{r^2 z_2^2}{(z_2^2 + \alpha^2 r^2)^{\frac{3}{2}}}$$

and set $\alpha = 0$ to get

$$r(\frac{r(z_1 + z_2)}{z_1 z_2} - 2).$$

We thus have a maximum when $(z_1 + z_2)/(z_1 z_2) < 2/r$ and a minimum when $(z_1 + z_2)/(z_1 z_2) > 2/r$. Furthermore, if the heights are equal ($z_1 = z = z_2$), then the minimum is only achieved when $z < r$.