Analytical expressions for optimizing the digging/filling volume for installing long straight arms

Technical report for the Cosmic Explorer collaboration (CE DCC number T2600009)

François Schiettekatte
Département de physique, Université de Montréal
francois.schiettekatte@montreal.ca

We develop analytical expression to compute:

i) The earth moving volume to create a trench and berms in order to install a straight arm at the surface of a spherical earth. We present
   a. The special case of an arm for the middle is at level (berm-only scenario);
   b. The special case of an arm for which each end is at level (trench-only scenario);
   c. The optimal case where the depth of the arm minimizes the volume of earth to move.

ii) The slopes and the (common) constant of two straight lines (i.e. the two arms of an L with an intersecting corner) given topographic data along such arms, and the resulting earth-moving volume, with an optimisation of the computing effort.

i) Volume to dig or fill to create a trench and berms to install a straight arm at the surface of a spherical earth

Let L be the length of a straight arm, and x the distance from the center of the arm, as illustrated in Fig. 1 a). Knowing the Earth radius R, the depth $z(x)$ to dig (positive) or fill (negative) is such that

$$R^2 = x^2 + (z - z_0 + R)^2$$

where z_0 is the sagitta, i.e. digging depth at $x = 0$, the middle of an arm. Hence,

$$z = R\sqrt{1 - (x/R)^2} - R + z_0.$$

Since $L \ll R$ and $|x| \leq L/2$, to an excellent approximation,

$$z \approx R - \frac{R}{2}\left(\frac{x}{R}\right)^2 - R + z_0 = z_0 - \frac{x^2}{2R}. \quad (1)$$

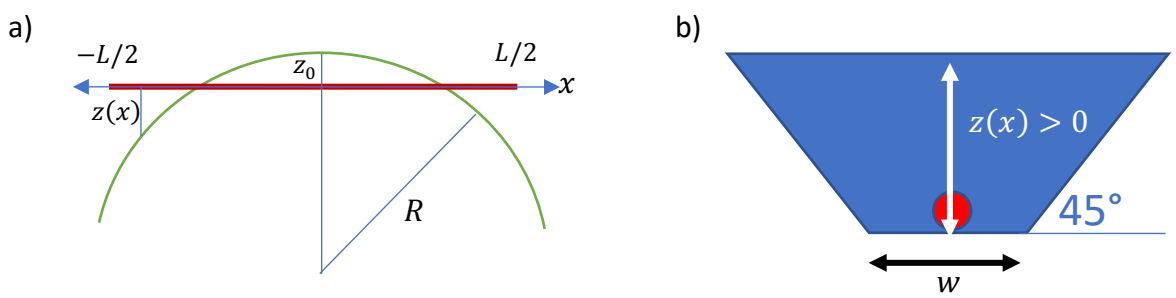

Fig. 1 a) Straight arm of length L at the surface of the Earth (radius R), the middle of which is inside a trench which depth is given by the sagitta z_0, and where $z(x)$ represents the depth of the trench ($z(x)$>0) or the height of the berm ($z(x) < 0$). b) Shape of a trench with a platform of width w at the bottom and 45° walls.

Assuming 45° walls allows us to compute the area represented by the two lateral triangles as illustrated in Fig. 1b) simply as $z^2(x)$. The digging/filling volume is therefore

$$\mathrm{V} = \int_{-\frac{L}{2}}^{\frac{L}{2}} z^2(x) + w|z(x)|\, dx = 2\int_0^{\frac{L}{2}} z^2(r) + w|z(x)|\, dx \tag{2}$$

if we assume that the digging/filling is symmetrical, as illustrated in Fig. 1a). To get rid of the absolute value in the integral, we find where $z(x)$ changes sign. This happens when $z(x) = 0$ in Eq. 1, i.e.

$$z_0 = \frac{R}{2}\left(\frac{x}{R}\right)^2$$

or

$$x = \sqrt{2z_0 R}\,.$$

Hence, Eq. 2 can be re-written as

$$\mathrm{V} = 2\int_0^{\frac{L}{2}} z^2(x)\, dx + 2\int_0^{\sqrt{2z_0R}} w\, z(x)\, dx - 2\int_{R\sqrt{2z_0R}}^{\frac{L}{2}} w\, z(x) dx \tag{3}$$

where we've put a negative sign in front of the term where $z(x) < 0$ so the result is positive (and corresponds to the filling for berms, towards the ends of the arm). Solving the first integral in Eq. 3, we get

$$2\int_0^{\frac{L}{2}}\left(z_0 - \frac{x^2}{2R}\right)^2 dx = 2\int_0^{\frac{L}{2}} z_0^2 - z_0\frac{x^2}{R} + \frac{x^4}{4R^2}\, dx = 2z_0^2\frac{L}{2} - \frac{2z_0}{R}\frac{\left(\frac{L}{2}\right)^3}{3} + \frac{2}{4R^2}\frac{\left(\frac{L}{2}\right)^5}{5}$$
$$= 2z_0^2\frac{L}{2} - \frac{2z_0}{3R}\left(\frac{L}{2}\right)^3 + \frac{1}{10\,R^2}\left(\frac{L}{2}\right)^5.$$

The second integral gives

$$2\int_0^{\sqrt{2z_0R}} w\left(z_0 - \frac{x^2}{2R}\right) dx = 2w\, z_0\sqrt{2z_0R} - \frac{w}{3R}(2z_0R)^{\frac{3}{2}} = 2w\, z_0^{\frac{3}{2}}\sqrt{2R} - 2\frac{w}{3}z_0^{\frac{3}{2}}\sqrt{2R} = \frac{4w}{3}z_0^{\frac{3}{2}}\sqrt{2R}$$

while, for the third integral, we have

$$-2\int_{\sqrt{2z_0R}}^{\frac{L}{2}} w\left(z_0 - \frac{x^2}{2R}\right) dx = -2w\, z_0\left(\frac{L}{2} - \sqrt{2z_0R}\right) + \frac{w}{3R}\left(\left(\frac{L}{2}\right)^3 - (2z_0R)^{\frac{3}{2}}\right)$$
$$= -2w\, z_0\left(\frac{L}{2}\right) + 2w\, z_0\sqrt{2z_0R} + \frac{w}{3R}\left(\frac{L}{2}\right)^3 - \frac{w}{3R}(2z_0R)^{\frac{3}{2}}$$
$$= \frac{4w}{3}z_0^{\frac{3}{2}}\sqrt{2R} - 2w\, z_0\left(\frac{L}{2}\right) + \frac{w}{3R}\left(\frac{L}{2}\right)^3.$$

We note that the terms in red are the same as the result of the second integral.

Putting together the result for each part of Eq. 3, the volume of earth to displace is therefore given by

$$V = 2\, z_0(\, z_0 - w)\frac{L}{2} - \frac{2z_0 - w}{3R}\left(\frac{L}{2}\right)^3 + \frac{1}{10\, R^2}\left(\frac{L}{2}\right)^5 + \frac{8w}{3} z_0^{3/2}\sqrt{2R}\,, \tag{4}$$

a relatively complicated equation but one of the main results of this report. We see from this expression that for a z_0 assumed constant (which might not be the case, as we'll see in subsection c. below), the volume is given by a polynomial of with only odd powers of L, up to L^5. Such a high power of L means the volume of earth to displace may significantly depend on the arm length. The L^5 originates from the fact that the height of the walls scale as L^2 so their area goes as L^4, times length L for the volume. It is, however, relatively difficult to appreciate from Eq. 4 if all terms are important or not. Let's look at special cases that simplify the expression.

a. Special case of an arm for which the middle is at level (berm-only scenario);

In the special case where $z_0 = 0$, we see from Fig. 1a) that this corresponds to the case where berms would have to be built on each side and to both ends of the arm in order to support it. In that case, Eq. 4 reduces to

$$V = \frac{w}{3R}\left(\frac{L}{2}\right)^3 + \frac{1}{10\, R^2}\left(\frac{L}{2}\right)^5$$

or, factoring out an $(L/2)^3$ term,

$$V = \left(\frac{1}{40}\left(\frac{L}{R}\right)^2 + \frac{w}{3R}\right)\left(\frac{L}{2}\right)^3. \tag{5}$$

We see that the term in L^5 is dominant only if $(L/R)^2$ is large compared to w/R. The ratio of the first term to the second gives $3L^2/40wR$. Assuming that L = 40 km, and that w = 18 m, and knowing that R = 6 371 km, we find that the ratio if close to 1, so neither term dominates in that case.

b. **Special case of an arm for which each end is at level (trench-only scenario);**

In the case where the end of each arm is at level (i.e. trench-only situation), we first need to find the corresponding z_0. It's given by Eq. 1 when $z = 0$ at $x = \pm L/2$:

$$0 = z_0 - \frac{(\pm L/2)^2}{2R}\,,$$

hence

$$z_0^{\text{to}} = \frac{L^2}{8R}. \tag{6}$$

Replacing z_0^{to} for z_0 in Eq. 4, we have

$$V = 2\,\frac{L^2}{8R}\left(\frac{L^2}{8R} - w\right)\frac{L}{2} - \frac{2(L^2/8R) - w}{3R}\left(\frac{L}{2}\right)^3 + \frac{1}{10\, R^2}\left(\frac{L}{2}\right)^5 + \frac{8w}{3}\left(\frac{L^2}{8R}\right)^{3/2}\sqrt{2R}\,.$$

Regrouping by powers of $(L/2)$, we get

$$V = \left(\frac{1}{2R^2} - \frac{1}{3R^2} + \frac{1}{10\,R^2}\right)\left(\frac{L}{2}\right)^5 - \left(\frac{w}{R} - \frac{w}{3R}\right)\left(\frac{L}{2}\right)^3 + \frac{8w}{3}L^3\frac{1}{8R}\sqrt{\frac{2R}{8R}}$$

$$= \left(\frac{4}{15R^2}\right)\left(\frac{L}{2}\right)^5 - \left(\frac{w}{R} - \frac{w}{3R}\right)\left(\frac{L}{2}\right)^3 + \frac{8w}{3}\frac{1}{2R}\left(\frac{L}{2}\right)^3$$

$$= \frac{4}{15R^2}\left(\frac{L}{2}\right)^5 + \frac{2w}{3R}\left(\frac{L}{2}\right)^3$$

or, finally,

$$V = \left(\frac{1}{15}\left(\frac{L}{R}\right)^2 + \frac{2w}{3R}\right)\left(\frac{L}{2}\right)^3 \tag{7}$$

which is the expression we published previously in Rev. Sci. Instrum. **96** (2025) 014502. So, Eq. 4 for the volume appears to be correct given that it works for these two special cases, and provides a general expression for any value of z_0, or at least for $0 < z_0 < z_0^{\text{to}}$, for which it has also been checked numerically.

c. Optimal case: z_0 that minimizes the volume of earth to move

We can use Eq. 4 to find the value of z_0 that minimizes V. This should be when $dV/dz_0 = 0$. For this derivative, we get

$$\frac{dV}{dz_0} = (2z_0 - w)L - \frac{2}{3R}\left(\frac{L}{2}\right)^3 + 4w\sqrt{2Rz_0} = 0$$

or

$$32w^2Rz_0 = \left(\frac{2}{3R}\left(\frac{L}{2}\right)^3 - (2z_0 - w)L\right)^2$$

which gives

$$(4z_0^2 - 2z_0w + w^2)L^2 - \frac{4}{3R}\left(\frac{L}{2}\right)^3(2z_0 - w)L + 32w^2Rz_0 + \frac{4}{9R^2}\left(\frac{L}{2}\right)^6 = 0$$

an ugly 2nd order polynomial in z_0 for which we have to find the roots, quite an error-prone exercise. Instead, we rely on an online symbolic calculator (Symbolab), which tells us that:

$$z_0 = \frac{L^4 + 12L^2wR + 96w^2R^2 \pm 8\sqrt{3}\,wR\sqrt{48w^2R^2 + 12wRL^2 + L^4}}{24L^2R}.$$

Dividing each term of the numerator by the denominator and inserting $\sqrt{3}/3 = 1/\sqrt{3}$ into the radical, we get

$$z_0 = \frac{L^2}{24R} + \frac{w}{2} + 4\frac{w^2R}{L^2} \pm \frac{w}{L^2}\sqrt{16w^2R^2 + 4wRL^2 + L^4/3}\,.$$

Let's factor out the first term in the radical, and factor out a $w/6$ from the first term of the expression, we find:

$$z_0 = \frac{w}{6}\frac{L^2}{4wR} + \frac{w}{2} + \frac{4w^2R}{L^2} \pm \frac{4w^2R}{L^2}\sqrt{1 + \frac{L^2}{4wR} + \frac{1}{3}\left(\frac{L^2}{4wR}\right)^2}$$

or

$$z_0^{\text{opt}} = \frac{w}{2}\left[1 + \frac{\alpha}{3} + \frac{2}{\alpha}\left(1 \pm \sqrt{1 + \alpha + \frac{\alpha^2}{3}}\right)\right], \qquad \alpha = \frac{L^2}{4wR}, \tag{8a}$$

a second important result of this report.

Equation 8a could be inserted in Eq. 4 in order to get the full expression of the optimal $V(L)$, but this would give a quite intractable expression, so it's better to compute them separately. Yet, we see that Eq. 8a brings an extra dependence of the volume on L, this one being far from just polynomial terms. Let's check how z_0^{opt} behaves to understand how it affects V. And yet, we have not determined which of the positive or negative roots is physically meaningful.

First, assuming again L = 40 km, w = 18 m and R = 6 371 km, we find that $\alpha \approx 3.488$, so no term can be neglected in Eq. 8a. In that case, we get z_0^{opt} = 39.1 m for the positive root, and z_0^{opt} = 9.54 m for the negative one, the latter being the solution sought for. This and other checks confirm that the negative root is the correct one, so the equation to use for z_0^{opt} is actually

$$z_0^{\text{opt}} = \frac{w}{2}\left[1 + \frac{\alpha}{3} + \frac{2}{\alpha}\left(1 - \sqrt{1 + \alpha + \frac{\alpha^2}{3}}\right)\right], \qquad \alpha = \frac{L^2}{4wR}. \tag{8b}$$

Before plotting z_0^{opt} to better understand it, let's find how it behaves when α is small or large. In the limit where $\alpha \ll 1$ (e.g. α = 0.035 if L = 4 km and w = 18 m), then to 2nd order, $\sqrt{1+y} \approx 1 + y/2 - y^2/8$, so

$$z_0^{\text{opt}} \approx \frac{w}{2}\left[1 + \frac{\alpha}{3} + \frac{2}{\alpha}\left(1 - \left(1 + \frac{\alpha}{2} + \frac{\alpha^2}{6} - \frac{1}{8}\left(\alpha + \frac{\alpha^2}{3}\right)^2\right)\right)\right].$$

Let's keep only the terms to order α^2, we get:

$$z_0 \approx \frac{w}{2}\left[1 + \frac{\alpha}{3} + \frac{2}{\alpha}\left(-\frac{\alpha}{2} - \frac{\alpha^2}{6} + \frac{\alpha^2}{8}\right)\right]$$
$$\approx \frac{w}{2}\left[1 + \frac{\alpha}{3} - 1 - \frac{\alpha}{3} + \frac{\alpha}{4}\right].$$

We see from this that all the terms from the 1st order expansion cancel out with the rest of the expression and we needed at least part or the 2nd order term. We therefore get

$$z_0^{\text{opt}} \approx \frac{\alpha w}{8} = \frac{L^2}{32R} = \frac{1}{4}z_0^{\text{to}}$$

which is z_0 = 0.0049 for the conditions above, resulting in V = 5688 m^3. We note in passing that the only negligible term in Eq. (4) for V in such conditions is the term in L^5.

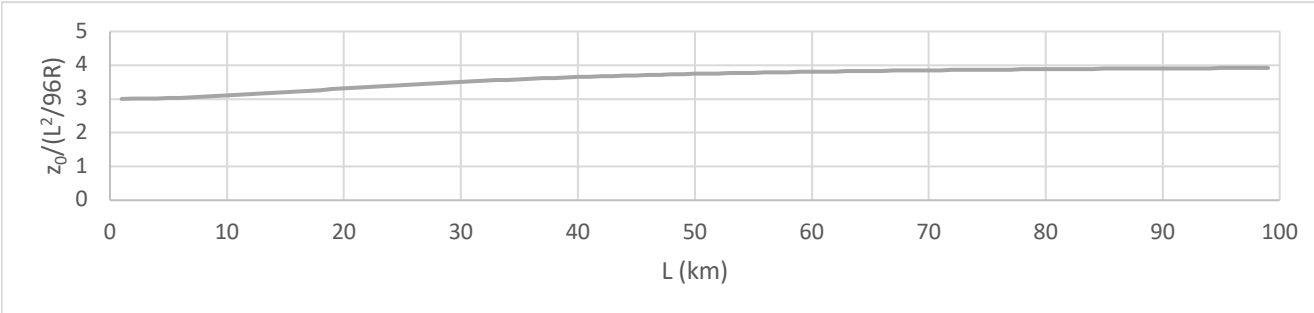

Fig. 2: $z_0^{\mathrm{opt}}/12z_0^{\mathrm{to}}$ (i.e Eq. 8b divided by 12 times Eq. 6) as a function of L assuming w = 18 m.

If $\alpha \gg 1$ (e.g. α = 21.8 if L = 100 km and w = 18 m), then $\alpha^2 \gg \alpha$, and

$$z_0 \approx \frac{w}{2}\left[1+\frac{\alpha}{3}+\frac{2}{\alpha}\left(1-\frac{\alpha}{\sqrt{3}}\right)\right] = \frac{w}{2}\left[\left(1-\frac{2}{\sqrt{3}}\right)+\frac{\alpha}{3}\right].$$

Further considering that $\alpha \gg 1$, we find in that case that

$$z_0^{\mathrm{opt}} \approx \frac{\alpha w}{6} = \frac{L^2}{24R} = \frac{1}{3}z_0^{\mathrm{to}}.$$

So, despite its complicated shape, save a factor between 1/4 and 1/3, Eq. 8b remains of order z_0^{to}, which was the depth of the trench in the middle of a "trench-only" scenario, with both ends of the arm at level. Figure 2 plots $z_0^{\mathrm{opt}}/12z_0^{\mathrm{to}}$ (i.e Eq. 8b divided by 12 times Eq. 6) as a function of L, assuming w = 18 m. We see that the value of z_0^{opt} smoothly changes between the two limits we found and is somewhere in the middle for a 40 km arm. The midpoint is when $\alpha \approx 2$ or $L \approx \sqrt{8wR}$.

Figure 3 shows Eq. 4 as a function of z_0 as a dashed curve, with each term of Eq. 4 plotted separately. No term is negligible although the 2nd term is small near the minimum. But this is simply because it becomes 0 when $2z_0 = w$, which happens at $z_0 = 9$ m when assuming w = 18 m. Also, the negative root of Eq. 8 appears to give the right location of the minimum in Eq. 4, but it's not clear what the positive root corresponds to, as Eq. 4 continues to increase monotonically.

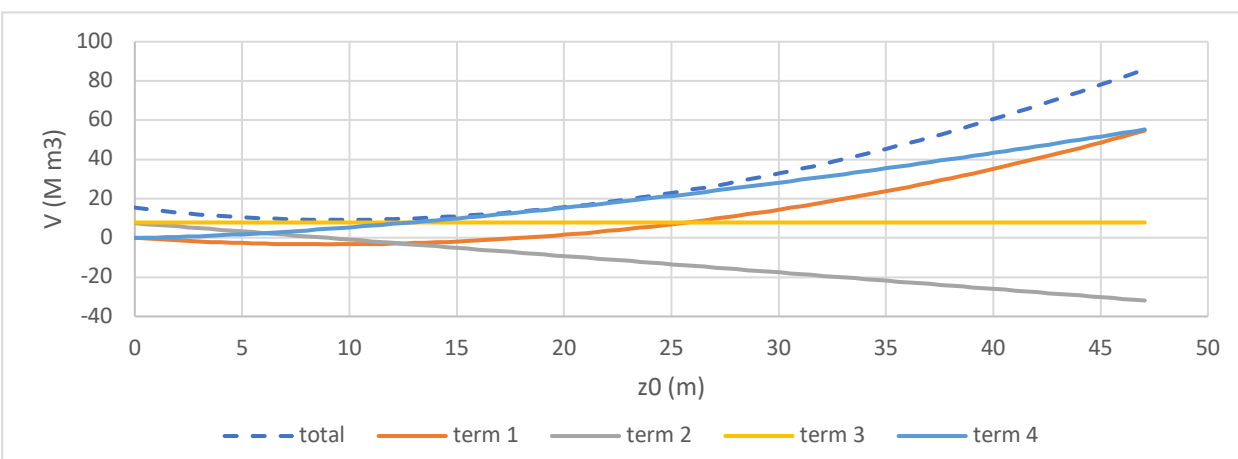

Fig. 3: Equation 4 as a function of z_0 assuming w = 18 m (dashed curve) and each of its term (solid curves).

Finally, Fig. 4 tries to capture in a single graph the behavior of Eq. 4 for the volume, as contours, as a function of z_0 and L, again assuming w = 18 m. $z_0^{\mathrm{opt}}(L)$, Eq. 8b, is plotted as a blue dashed line. It is seen that it crosses the contours at the maximum value of L each volume contour can have, validating that z_0^{opt} is indeed the value of $z_0$that provides the largest L for a given volume, or conversely, that it gives the smallest volume for a given L value. We observe that the volume changes relatively slowly with z_0 near the crossings, which means that the exact value of z_0 doesn't critically affects the volume near the optimum.

The orange dashed curve is Eq. 7 for the volume of the trench-only scenario, with both ends of the arm at ground level. We see that it crosses the contours at their maximum value (or more precisely, when their slope reaches 0). This means that this scenario gives the largest volume for a given arm length, hence the wort scenario. The contours continue left of the orange curve, but it is not clear if these solutions have any physical meaning. Finally, the intersection of the contours with the horizontal axis gives the length of an arm for a certain volume in the berm-only, arm-center-at-level scenario, as one would find with Eq. 5 (i.e. $z_0 = 0$). We see that these solutions always give a larger L then the "trench-only" scenario, since the volume of Eq. 5 is always smaller than that to Eq. 7.

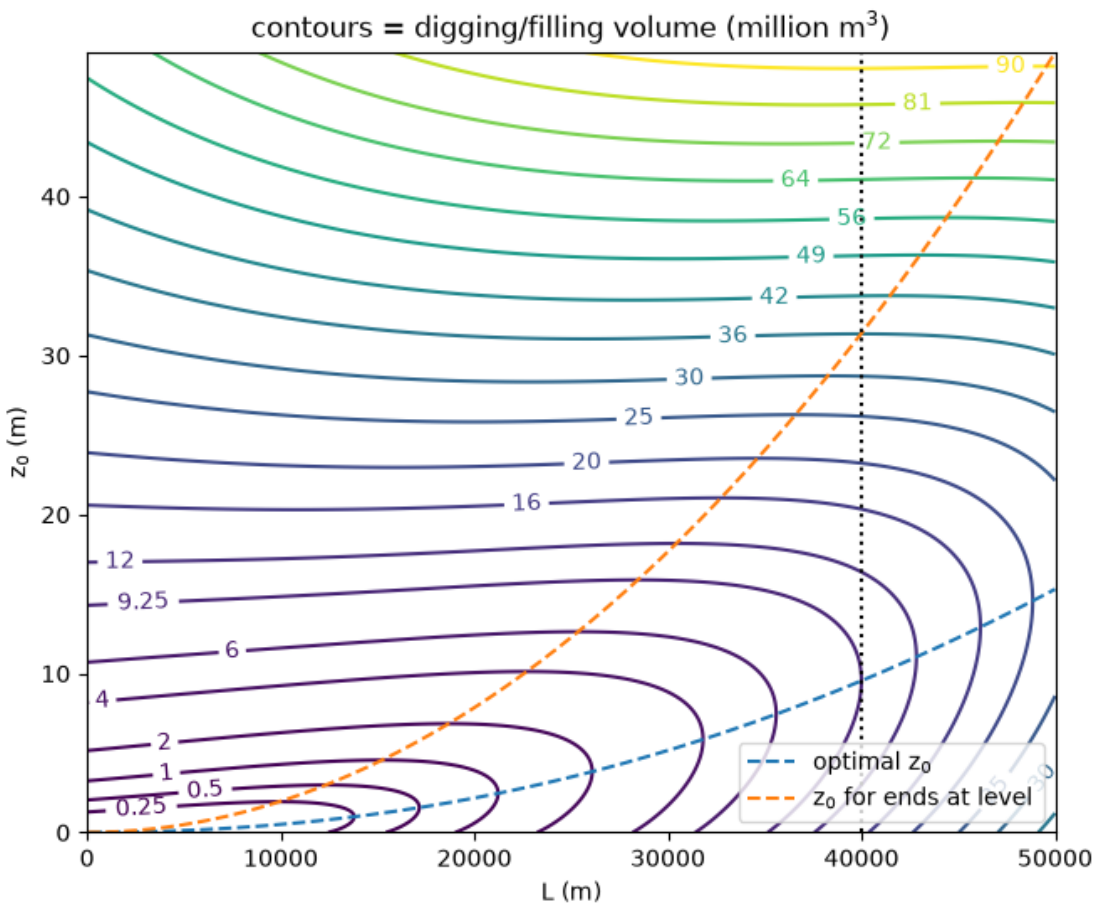

Fig. 4: Contour plot of the volume, Eq. 4, as a function of z_0 and L, assuming w = 18 m. Blue dashed curve: $z_0^{\mathrm{opt}}(L)$, Eq. 8b. Orange dashed curve: Eq. 7 for the volume of a trench-only, ends-at-level scenario.

ii) Equations for double linear regression with same constant

The equations presented in section i) only apply to a spherical Earth. They are useful to understand how, in general, the earth displacement volume is affected by the arm length and its depth in the middle, but finding the volume for actual locations on earth involve finding, at each of these location, how much earth needs to be moved.

This can be achieved rather simply, provided that reliable, precise and detailed topographic data are available in numerical format for each point. This is called Digital Elevation Maps (DEM) and are provided by different organizations. We note in passing that the accuracy of such data must be evaluated carefully as some of them can be inaccurate to a few meters, which may completely defeat the purpose of the exercise and may not provide reliable results.

To find the minimum digging volume for an arm at a given location, one only needs to retrieve the elevation data for a (not too small) number of points along the path of the arm, and carry out a linear regression. As we will see below, the χ^2 for each point gives the area formed by the two triangles in Fig. 1b), assuming 45° walls, while the difference between the DEM value and the fit gives the height of the berm or trench, which can be multiplied by the width w to get the other part of the section area of the berm or trench at that point. Integrating (or summing, times the distance between each point) gives the volume, for one arm.

An important point, however, is that we want to evaluate the digging volume for two arms intersecting at the same point, the L corner. So, we need to carry out two linear regressions but with the same constant. The other point is that we may need to do it for millions of locations and orientations of the Ls, as well as different values of L and angle between the arms, in order to find the best locations. So, the exercise below is not only to find how to do this "two-linear-regressions, same constant" fit, but how to do it fast. Hence, we derive the way to compute it, but also identify the terms that can be computed once vs. those that need to be computed at every location and orientation. Please note that equation numbering in this section restarts at 1.

We want to fit two sets of N data points, (r_i, y_i) and (r_i, z_i) with two linear regressions sharing the same constant c so the arms intersect at $r = 0$, i.e.,

$$y = ar + c,$$
$$z = br + c.$$

Here, we assume that the set of r_i is always the same, e.g. one point every 100 m or 1 km. For a linear regression based on least-square fit, we what to find a, b, c that minimize

$$f = \sum(y - y_i)^2 + \sum(z - z_i)^2 = \sum(ar_i + c - y_i)^2 + \sum(br_i + c - z_i)^2.$$

The solution satisfies

$$\frac{\partial}{\partial a} f = 0 \Rightarrow \sum 2(ar_i + c - y_i)r_i = 0$$
$$\Rightarrow a \underbrace{\sum r_i^2}_{\alpha} + c \underbrace{\sum r_i}_{\beta} = \underbrace{\sum y_i r_i}_{C},$$

$$\frac{\partial}{\partial b} f = 0 \Rightarrow \sum 2(br_i + c - z_i) r_i = 0$$
$$\Rightarrow b\sum r_i^2 + c\sum r_i = \underbrace{\sum z_i r_i}_{D},$$

and

$$\frac{\partial}{\partial c} f = 0 \Rightarrow \sum 2(ar_i + c - y_i) + \sum 2(br_i + c - z_i) = 0$$
$$\Rightarrow (a+b)\sum r_i + 2Nc = \underbrace{\sum y_i + \sum z_i}_{E},$$

where N is the number of elements in each sum, e.g. 41 if r_i ranges from 0 to 40 inclusively by steps of 1. Using the constants defined in the equations above, we get the following system of equations:

$$a\alpha + c\beta = C, \quad (1)$$
$$b\alpha + c\beta = D, \quad (2)$$
$$a\beta + b\beta + 2Nc = E. \quad (3)$$

We could solve it quite easily using linear algebra. However, since in our computation, the r_i are always the same, we can pre-compute α, β and other constant terms, and save computation time. Only C, D, E need to be computed for each of the millions or billions of fits. Isolating a and b in Eqs. (1) and (2), we get

$$a = (C - c\beta)\alpha^{-1}, \quad (4)$$
$$b = (D - c\beta)\alpha^{-1}. \quad (5)$$

Injecting them in (3), we get

$$\beta C - c\beta^2 + \beta D - c\beta^2 + 2\alpha Nc = \alpha E$$

where we multiplied both sides by α. Isolating c, we find

$$c\ \ 2(\alpha N - \beta^2) = \alpha E - \beta C - \beta D$$

i.e.

$$c = (\alpha E - \beta(C + D))\ \underbrace{(2(\alpha N - \beta^2))^{-1}}\ . \quad (6)$$

It is worth pointing out that the term with an underbrace in Eq. (6) is constant through the computation and can be computed beforehand to save computation time. We can also pre-compute α^{-1} to avoid computing millions of divisions in Eqs. 4 and 5 (which are much slower than multiplications).

The problem at stake is therefore solved by computing Eq. (6) and using the result to solve Eqs. (4) and (5), with α, β, N, α^{-1} and $(2(\alpha N - \beta^2))^{-1}$ known or computed beforehand.

As a result of the computations above, we want to know the volume of earth to move, which is

$$V = \sum[(ar_i + c - y_i)^2 + w|ar_i + c - y_i| + (br_i + c - z_i)^2 + w|br_i + c - z_i|] \quad (7)$$

It involves hundreds of computations at each of the millions of locations and orientations scanned. It's not obvious to reduce the computational time of this equation.

One way is to simplify the squared terms is expressed as follows:

$$\sum[(ar_i + c - y_i)^2 + (br_i + c - z_i)^2]$$

$$= \sum[a^2 r_i^2 + 2acr_i - 2ar_i y_i + c^2 - 2cy_i + y_i^2 + b^2 r_i^2 + 2bcr_i - 2br_i z_i + c^2 - 2cz_i + z_i^2]$$

$$= (a^2 + b^2) \underbrace{\sum r_i^2}_{\alpha} + 2\left((a+b)c \underbrace{\sum r_i}_{\beta} - a \underbrace{\sum r_i y_i}_{C} - b \underbrace{\sum r_i z_i}_{D} + c^2 - c \underbrace{\sum (y_i + z_i)}_{E} \right) + \sum(y_i^2 + z_i^2). \quad (8)$$

We see that only the last term involves a new loop on i and it would likely reduce the computation time.

On the other hand, the terms involving an absolute value on each point in Eq. (7) have to be computed individually. We note that the argument of the absolute value is the same as the terms that are squared. So once the arguments of the absolute values are computed, they can be squared and summed, which does not represent more computational effort than computing the last term of Eq. (8).

Hence, to compute Eq. (7) we first find the absolute value of the residues along each arm:

$$p_i = |ar_i + c - y_i|, \qquad q_i = |br_i + c - z_i|$$

and then compute the volume as

$$V = \sum[p_i^2 + w(p_i + q_i) + q_i^2].$$