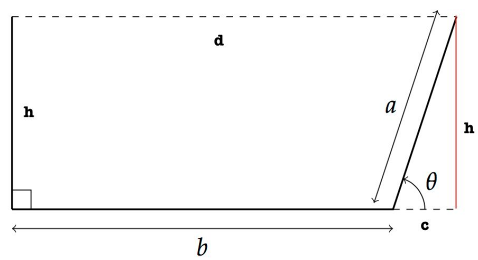
Minerva Schools At KGI
Assignment 1 - Project Design
Quang Tran
CS164 Spring 2020

All the angle measures are in radians, unless otherwise noted.

Part 1.

(a) **Visualize the surface plot. Estimate the maximum.**



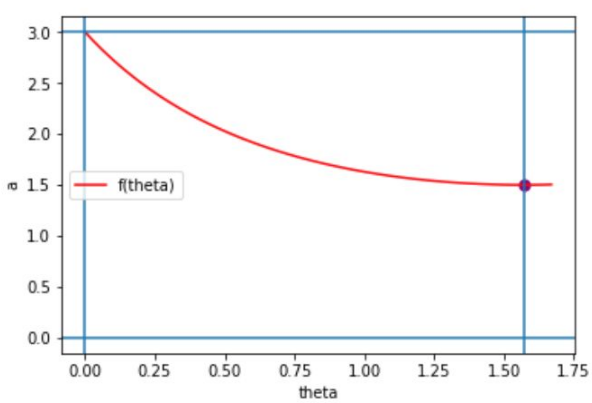
Here we will not consider configurations where $\theta > \pi/2$, because for any such configuration, we can bend the segment a so that the new segment is symmetric with the old segment w.r.t. the vertical line (the new θ is $\pi - \theta$) and this new configuration has a larger area.

$$A(a, \theta) = Wasin\theta - a^2sin^2\theta - a^2sin\theta + \tfrac{1}{2}a^2sin\theta cos\theta$$

The primary physically-realistic domain of a and θ is $a \in [0, W]$ (a can't exceed the perimeter) and $\theta \in [0, \pi/2]$ (Strictly speaking, a and θ cannot be 0, but we will later show that the maximum over this loose domain does not lie on the boundary.) We have one more constraint: h, b, a must be positive. This means that $a \leq W/(1 + sin\ \theta)$ (Again, b should be strictly positive, but as will be shown the maximum does not lie in this boundary.)

The physically-realistic domain is, therefore,

$$\begin{cases} a \in [0, W] \\ \theta \in [0, \pi/2] \\ a \leq W/(1 + \sin\theta) \end{cases}$$



The physically-realistic domain is the region bounded by the outer rectangular box ($a = 0,\ a = 3,\ \theta = 0, \theta = \pi/2$) and below the red curve (for $f(\theta) = W/(1 + sin\theta)$), including the boundaries. The red curve within the domain runs from left to right starting from $(0, f(0)) = (0, 3)$ and ends at $(\pi/2, f(\pi/2)) = (\pi/2, W/2) = (\pi/2, 1.5)$.

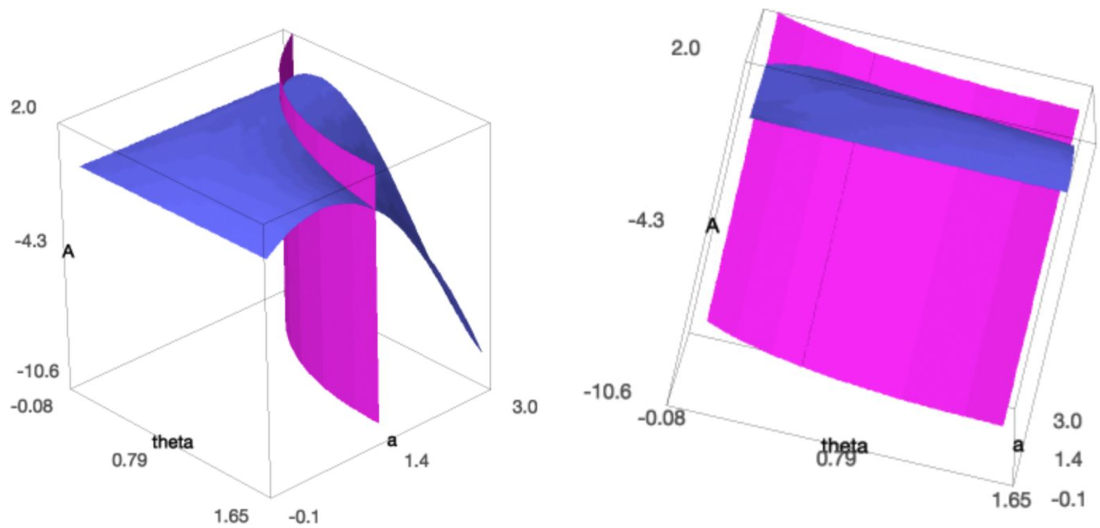


Figure 1. Surface plot of $A(a, \theta; W = 3)$ (purple) from two views. For purposes of visualization, I didn't include the $a = 0,\ a = W,\ \theta = 0, \theta = \pi$ surfaces. The part of A in the physically-realistic domain is the part of the purple curve in front of the magenta and green curves. We see that there is a bump with the highest point around $(a, \theta) = (1, 1)$ and it seems unique because there is no other bump within the considered domain.

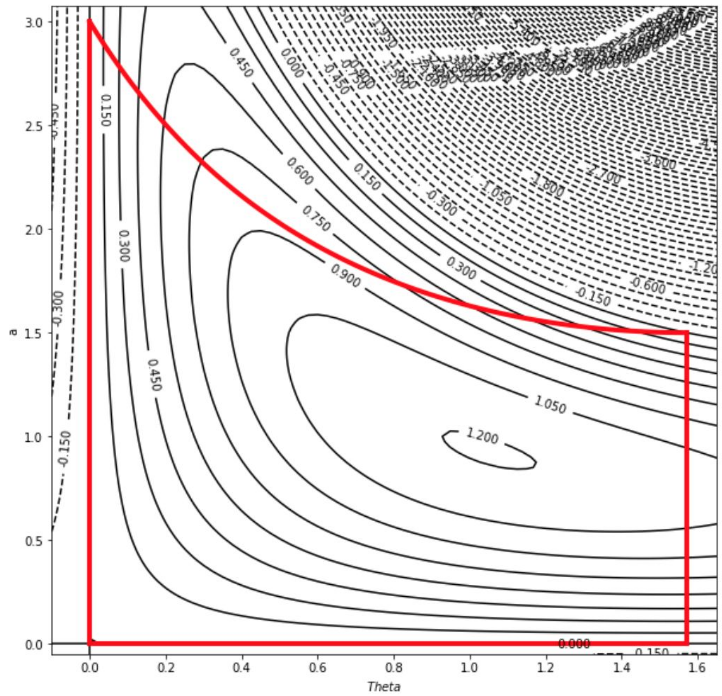


Figure. Contour plot for A. The domain of physically-realistic is bounded by the red curves. Negative lines are dashed.
The contour plot shows it more clearly that the maxima lies around (1,1).

(b) Is -A coercive?

$$-A(a,\theta) =-(Wa sin\theta - a^2 sin^2\theta - a^2 sin\theta + \frac{1}{2}a^2 sin\theta cos\theta)$$
$$=-\frac{1}{2}a\ sin\theta(2W - 2a\ sin\theta - 2a + a\ cos\theta)$$

For $-A$ to be coercive, we look at how $-A$ behaves as the norm of the input vector $\|(a,\theta)\| = \sqrt{a^2+\theta^2}$ tends to infinity. If we keep a constant, and let θ vary to infinity, i.e., looking at the θ direction (which then means $\|(a,\theta)\| \to \infty$), then
$B(\theta) =-A(a,\theta) =-Wa sin\theta + a^2 sin^2\theta + a^2 sin\theta - \frac{1}{2}a^2 sin\theta cos\theta$
We have: $-1 \le sin\theta,\ cos\theta, sin(2\theta) \le 1,$ then:

$$-Wa \le -Wa sin\theta \le Wa$$
$$0 \le a^2 sin^2\theta \le a^2$$
$$-a^2 \le a^2 sin\theta \le a^2$$
$$\frac{-1}{4}a^2 \le -\frac{1}{2}a^2 sin\theta cos\theta = \frac{-1}{4}a^2 sin(2\theta) \le \frac{1}{4}a^2$$

Summing up the LHS and the RHS of the above inequalities we get:
$-Wa - \frac{5}{4}a^2 \le B(\theta) =-A(a,\theta) \le Wa + \frac{9}{4}a^2$
which means $-A$ is then bounded above and below by $-Wa - \frac{5}{4}a^2$ and $Wa + \frac{9}{4}a^2$ respectively as we go along the said direction to the infinity. Because $-A$ does not get to infinity in this case, $-A$ is not coercive.
Another way to prove it is we keep $a = 0$ and then vary θ. In this case, despite θ tending to infinity, -A=0. Similarly, if we keep $\theta = 0$, then $sin\theta = 0$ and $-A = 0$.

Part 2.

$A(a,b,\theta) = \frac{1}{2}a\, sin\theta(2b + a\, cos\theta)$

Equality constraint: $g(a,b,\theta) = a\, sin\theta + b + a = 3$ (because h+b+a=W)

To apply Lagrange multipliers method, we need to solve the following two equations:

1. $\nabla A(a,b,\theta) = \lambda\nabla g(a,b,\theta)$
2. $a\, sin\theta + b + a = 3$ (the constraint function) **(a)**

We have $\nabla A(a,b,\theta) = (sin\theta(b + acos\theta),\ asin\theta,\ abcos\theta + \frac{1}{2}a^2cos(2\theta))$

And $\nabla g(a,b,\theta) = (sin\theta + 1, 1, acos\theta)$

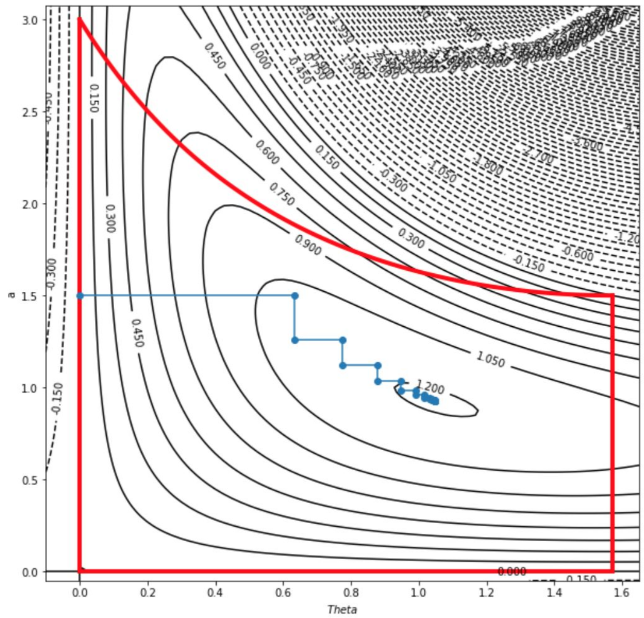
Equation 1 is equivalent to solving the following three equations:

a. $sin\theta(b + acos\theta) = \lambda(sin\theta + 1)$ **(b)**
b. $asin\theta = \lambda$ **(c)**
c. $abcos\theta + \frac{1}{2}a^2cos(2\theta) = \lambda acos\theta$ **(d)**

So we need to solve the system of equations (a), (b), (c), and (d). This will give two critical points: (a,b,θ) : $(0,3,0,0)$ (A=0) and $(4\sqrt{3} - 6\ , 3 - \sqrt{3}, \pi/3\ )$ ($A = 9 - \frac{9\sqrt{3}}{2}$) (See Appendix)

We first prove that these two points must be the maximum and the minimum points (and not an inflection point.) We have that our domain is closed ($a \in [0, W]$, $\theta \in [0, \pi/2]$, $a \le W/(1 + sin\ \theta)$), and bounded. The function A itself is continuous (it is a function composed of continuous functions by elementary operations). According to Weierstrasse extreme value theorem, there exists a minimum and a maximum points over this domain. Furthermore, we found only two critical points on this domain, so one of them must be a minimum and the other must be the maximum. Given the computed value of A at those points, the minimum is $(a,\theta) = (0,0)$ and the maximum is $(a,\theta) = (4\sqrt{3} - 6\ , \pi/3\ )$ ($A_{max} = 9 - \frac{9\sqrt{3}}{2}$)

Part 3.



Convergence using line exact search with bisection

We see that the algorithm does converge towards the "center" of the lowest point. At the last step, the point is $(\theta, a) = (1.04719666,\ 0.9282037)$, which agrees with the solution found above. The table that shows the convergence is included in the appendix.

Appendix

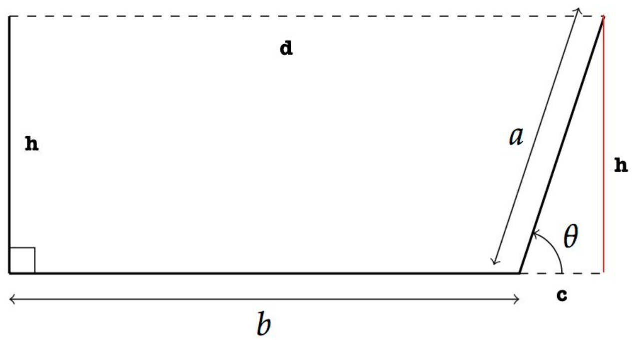
HC Application

#algorithm: Followed and implemented successfully the exact line search using bisection.
#optimization: Carefully constructed the constraints and the physically-realistic domain, applied correctly the Lagrange multipliers method and classified the critical points.

Code for surface plot in SageMath

https://gist.github.com/

Derivation for the expression for $A(a, \theta)$



$d = b + c$
$\sin\theta = h/a \Rightarrow h = a\sin\theta$
$\cos\theta = c/a \Rightarrow c = a\cos\theta$

$$\begin{aligned}A(a,\theta) &= \tfrac{1}{2}h(b+d) = \tfrac{1}{2}h(b+b+c)\\ &= \tfrac{1}{2}h(2b+c)\\ &= \tfrac{1}{2}h\,(2(W-h-a)+c) \text{ (because } W = h+b+a \text{ is the perimeter)}\\ &= \tfrac{1}{2}h(2W-2h-2a+c)\\ &= \tfrac{1}{2}a\sin\theta(2W-2a\sin\theta-2a+a\cos\theta)\\ &= Wa\sin\theta - a^2\sin^2\theta - a^2\sin\theta + \tfrac{1}{2}a^2\sin\theta\cos\theta\end{aligned}$$

Derivation for h, b, and a being positive means that $a < W/(1+\sin\theta)$.

- $h > 0 \Leftrightarrow a\sin\theta > 0$. This is always true for $a > 0$ and $\theta \in (0, \pi)$.
- $a > 0$. This is always true because in the primary constraint stated above, $a \in (0, W)$
- $b > 0 \Leftrightarrow W - a - h > 0 \Leftrightarrow W - a - a\sin\theta > 0 \Leftrightarrow a(1+\sin\theta) < W$ $\Leftrightarrow a < W/(1+\sin\theta)$ (we can divide both sides by $1+\sin\theta$ without changing the inequality sign because $1+\sin\theta > 0$ for $\theta \in [0, \pi/2]$)

In sum, the condition translates to $a < W/(1+\sin\theta)$.

Finding the critical points

$A(a,b,\theta) = \frac{1}{2}a\sin\theta(2b + a\cos\theta)$
$a\sin\theta + b + a = 3$ (the constraint function) **(a)**
$\sin\theta(b + a\cos\theta) = \lambda(\sin\theta + 1)$ **(b)**
$a\sin\theta = \lambda$ **(c)**

$abcos\theta + \frac{1}{2}a^2cos(2\theta) = \lambda acos\theta$ **(d)**

So we need to solve the system of equations (a), (b), (c), and (d).
We first consider these following cases:

- If $sin\theta = 0$. Then this means $\theta = 0$ (because $\theta \in [0, \pi/2]$), $cos\theta = cos2\theta = 0$, $\lambda = 0$ (from (c)). From (a), b+a=3. We further have $ab + a^2/2 = 0$ (from (d). These last two equations means $(a, b) = (0, 3)$ *or* $(6, -3)$. But we don't consider (6,-3) because it's outside our physically-realistic domain. In sum, when $sin\theta = 0$:
  - $(a, b, \theta, \lambda) = (0, 3, 0, 0)$ and $A = 0$
- If $cos\theta = 0$. This means $\theta = \pi/2$ (because $\theta \in [0, \pi/2]$), $sin\theta = 1$. From (d) we have $-a^2/2 = 0$ which means $a = 0$. From (c), $\lambda = 0$. Because $\lambda + b + a = 3$ (from (a) and (c)), $b = 3$. But these results contradict with (b). Therefore, $cos\theta \neq 0$.
- If $a = 0$, then $b = 3$ (from (a)), $\lambda = 0$ (from (c)), and $\theta = 0$ (from (b).) In sum, when $a = 0$:
  - $(a, b, \theta, \lambda) = (0, 3, 0, 0)$ and $A = 0$
- If $\lambda = 0$: From (c): either $a = 0$, which brings us back to the case $a = 0$, or $sin\theta = 0$, which brings us back to the case $sin\theta = 0$.
  - $(a, b, \theta, \lambda) = (0, 3, 0, 0)$ and $A = 0$

In the derivation below, we therefore assume that $sin\theta,\ cos\theta,\ a,\ \lambda \neq 0$.
Multiplying both sides of (d) by $sin^2\theta$ we get:
$absin^2\theta cos\theta + \frac{1}{2}a^2sin^2\theta cos(2\theta) = \lambda asin^2\theta cos\theta$
Substituting $asin\theta$ with λ (from (c)), we get:
$b\lambda sin\theta cos\theta + \frac{1}{2}a\,\lambda sin\theta cos(2\theta) = \lambda^2 cos\theta sin\theta$
$\Leftrightarrow bsin\theta cos\theta + \frac{1}{2}a\ sin\theta cos(2\theta) = \lambda\ sin\theta cos\theta$ (dividing both sides by λ)
$\Leftrightarrow bsin\theta cos\theta + \frac{1}{2}asin\theta\ (2cos^2\theta - 1) = \lambda sin\theta\ cos\theta$
$\Leftrightarrow bsin\theta cos\theta + asin\theta cos^2\theta - (asin\theta)/2) = \lambda\ sin\theta cos\theta$
$\Leftrightarrow sin\theta cos\theta(b + acos\theta) - (asin\theta)/2) = \lambda\ sin\theta cos\theta$
Substitute $sin\theta(b + acos\theta)$ with $\lambda(sin\theta + 1)$:
$\lambda cos\theta(sin\theta + 1) - (asin\theta)/2) = \lambda\ sin\theta cos\theta$
$\Leftrightarrow \lambda cos\theta - (asin\theta)/2) = 0$
Substitute $asin\theta$ with λ:
$\lambda cos\theta - \lambda/2) = 0$
$\Leftrightarrow cos\theta = 1/2$ (dividing both sides by λ)
$\Leftrightarrow \theta = \pi/3$
With this, (a), (c), and (d) become:

- $\left(\frac{\sqrt{3}}{2} + 1\right)a + b = 3$
- $\frac{\sqrt{3}}{2}a = \lambda$
- $b - \frac{1}{2}a = \lambda$

Solving the above system we get: $a = 4\sqrt{3} - 6,\ b = 3 - \sqrt{3}, \lambda = 6 - 3\sqrt{3}$. Therefore, we have two critical points (a, b, θ): $(0, 3, 0, 0)$ (A=0) and $(4\sqrt{3} - 6\ , 3 - \sqrt{3}, \pi/3\ )$ ($A = 9 - \frac{9\sqrt{3}}{2}$)

Code for Gradient Descent

https://gist.github.com/quangntran/ef2172d857926e6c01077316aacc80ec

Convergence Table

	k	norm_d	theta	a	A
0	1	3.375000e+00	0.000000	1.500000	0.000000
1	2	3.371201e-01	0.634118	1.500000	1.080211
2	3	4.858194e-01	0.634118	1.260854	1.120522
3	4	2.654409e-01	0.773031	1.260854	1.153355
4	5	2.702426e-01	0.773031	1.119061	1.172174
5	6	1.874790e-01	0.877242	1.119061	1.185901
6	7	1.491810e-01	0.877242	1.034953	1.193785
7	8	1.199713e-01	0.948098	1.034953	1.198968
8	9	8.097943e-02	0.948098	0.986392	1.201881
9	10	7.093561e-02	0.992033	0.986392	1.203637
10	11	4.324153e-02	0.992033	0.959297	1.204598
11	12	3.970159e-02	1.017415	0.959297	1.205143
12	13	2.281280e-02	1.017415	0.944603	1.205435
13	14	2.147751e-02	1.031409	0.944603	1.205594
14	15	1.194402e-02	1.031409	0.936787	1.205678
15	16	1.139426e-02	1.038913	0.936787	1.205722
16	17	6.226110e-03	1.038913	0.932676	1.205746

17	18	5.980592e-03	1.042874	0.932676	1.205758

click to scroll output; double click to hide

18	19	3.237707e-03	1.042874	0.930529	1.205764
19	20	3.121199e-03	1.044948	0.930529	1.205768
20	21	1.681510e-03	1.044948	0.929411	1.205769
21	22	1.624023e-03	1.046029	0.929411	1.205770
22	23	8.727047e-04	1.046029	0.928830	1.205771
23	24	8.436836e-04	1.046591	0.928830	1.205771
24	25	4.527738e-04	1.046591	0.928528	1.205771
25	26	4.379366e-04	1.046883	0.928528	1.205771
26	27	2.348633e-04	1.046883	0.928372	1.205771
27	28	2.272260e-04	1.047034	0.928372	1.205771
28	29	1.218168e-04	1.047034	0.928291	1.205771
29	30	1.178714e-04	1.047113	0.928291	1.205771
30	31	6.317977e-05	1.047113	0.928249	1.205771
31	32	6.113638e-05	1.047154	0.928249	1.205771
32	33	3.276815e-05	1.047154	0.928227	1.205771
33	34	3.171158e-05	1.047175	0.928227	1.205771
34	35	1.699320e-05	1.047175	0.928215	1.205771
35	36	1.644648e-05	1.047186	0.928215	1.205771
36	37	8.812736e-06	1.047186	0.928210	1.205771

37	38	8.539493e-06	1.047191	0.928210	1.205771
38	39	4.567852e-06	1.047191	0.928207	1.205771
39	40	4.439755e-06	1.047194	0.928207	1.205771
40	41	2.361956e-06	1.047194	0.928205	1.205771
41	42	2.263503e-06	1.047196	0.928205	1.205771
42	43	1.260840e-06	1.047196	0.928204	1.205771
43	44	1.125319e-06	1.047197	0.928204	1.205771
44	45	6.825287e-07	1.047197	0.928204	1.205771