Variational Calculus - Week 8 - Lagrangian Multipliers & the Isoperimetric Problem

Antonio León Villares

November 2022

Contents

1 Lemma: The Method of Lagrange Multipliers in Several Variable Calculus

Let:

$$f, g : U \to \mathbb{R}, \qquad U \subset \mathbb{R}^n$$

Let $\underline{x}_0$ *be an interior point of* U*, such that* $\underline{x}_0$ *is an* ***extremum*** *of* f*, subject to* $g(\underline{x}) = 0$*.*
Then, if:

$$\nabla g(\underline{x}_0) \neq 0$$

there exists a ***Lagrange multiplier*** $\lambda \in \mathbb{R}$*, such that* $(\underline{x}_0, \lambda)$ *is a* ***critical point*** *of:*

$$F : U \times \mathbb{R} \to \mathbb{R}$$

$$F(\underline{x}, \lambda) = f(\underline{x}) - \lambda g(\underline{x})$$

In particular, we have a ***necessary*** *condition for a* ***critical point*** *of* f ***constrained*** *by* $g = 0$*: if* $\underline{x}_0$ *is a critical point, then* $\exists \lambda_0 \in \mathbb{R}$ *such that:*

$$\nabla F(\underline{x}_0, \lambda_0) = 0 \implies \begin{cases} \dfrac{\partial F}{\partial \lambda} = 0 \;\therefore\; g(\underline{x}_0) = 0 \\ \dfrac{\partial F}{\partial x^i} = 0 \;\therefore\; \dfrac{\partial f}{\partial x^i}(\underline{x}_0) = \lambda_0 \dfrac{\partial g}{\partial x^i}(\underline{x}_0) \end{cases}$$

(Lemma 9.1)

2 Constrained Optimisation in Functionals

2.1 Motivation: The Isoperimetric Problem

2.1.1 Theorem: Green's Theorem

Let C be a:

- ***positively oriented*** *(i.e can be traversed counterclockwise)*
- ***piecewise smooth***
- ***simple*** *(i.e not self-intersecting)*
- ***closed***

curve, enclosing a region D.
Moreover, let $F, G \in C^1$ be functions $\mathbb{R}^2 \to \mathbb{R}$.
Then, ***Green's Theorem*** *states:*

$$\oint_C F dx + G dy = \iint_D \left(\frac{\partial G}{\partial x} - \frac{\partial F}{\partial y} \right) dx dy$$

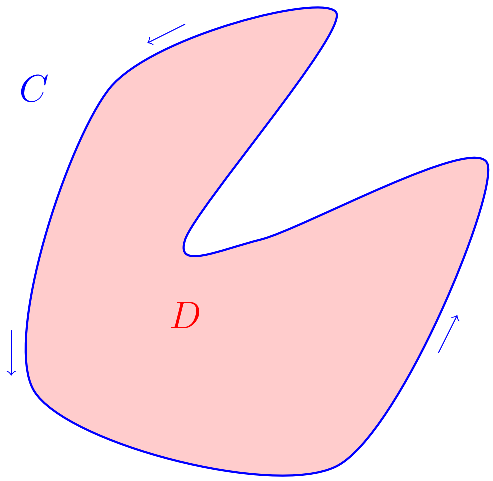
Moreover, if we parametrise C via:

$$\underline{x} = (x(t), y(t))$$

we have that:

$$\int_0^1 [F(\underline{x}(t))\dot{x}(t) + G(\underline{x}(t))\dot{y}(t)] \, dt = \iint_D \left(\frac{\partial G}{\partial x} - \frac{\partial F}{\partial y} \right) dx dy$$

Green's Theorem *can be thought of a two-dimensional fundamental theorem of calculus, relating a line integral to an area integral. For concreteness, an example for C, and D could be:*



2.1.2 Corollary: Area Formulae from Green's Theorem

Notice, if:

$$\frac{\partial G}{\partial x} - \frac{\partial P}{\partial y} = 1$$

*Then **Green's Theorem** tells us that:*

$$Area(D) = \oint_C F dx + G dy$$

Sensible choices for F, G then give:

$$F(x,y) = -y \quad \& \quad G(x,y) = 0 \implies Area(D) = \oint_C -y dx$$

$$F(x,y) = 0 \quad \& \quad G(x,y) = x \implies Area(D) = \oint_C x dy$$

$$F(x,y) = -\frac{y}{2} \quad \& \quad G(x,y) = \frac{x}{2} \implies Area(D) = \frac{1}{2}\oint x dy - y dx$$

2.1.3 The Isoperimetric Problem

*The **Isoperimetric Problem** is the following:*

Find the closed plane curve of a given length that encloses the largest area.

*This can be reformulated as a **variational problem** in terms of **functionals**:*

Let:

$$\underline{x} : [0,1] \to \mathbb{R}^2$$

be a ***continuously differentiable*** *(C^1)* ***closed*** *curve:*

$$\underline{x}(0) = \underline{x}(1)$$

The ***area*** *enclosed by $\underline{x}$ is a* ***functional****:*

$$A[\underline{x}] = \frac{1}{2}\int_0^1 (x^1\dot{x}^2 - x^2\dot{x}^1)dt$$

(by using ***Green's Theorem****)*
Its ***perimeter*** *is our well known* ***arclength*** *functional:*

$$S[\underline{x}] = \int_0^1 \|\dot{\underline{x}}(t)\|dt = \int_0^1 \sqrt{(\dot{x}^1)^2 + (\dot{x}^2)^2}dt$$

The ***isoperimetric problem*** *requires that we* ***extremise*** *$A[\underline{x}]$, subject to the* ***functional constraint*** *$S[\underline{x}] = \ell$.*

Notice, this is reminiscent of ***constrained optimisation*** *in* ***SVC****, but this time we* ***extremise*** *functionals, and use functional constraints. Hence, we will have to develop* ***Lagrange Multipliers*** *for* ***variational calculus*** *to solve the* ***Isoperimetric Problem****.*

2.2 Theorem: The Method of Lagrange Multipliers for Functionals

Let:

$$J[\underline{y}] = \int_0^1 L(\underline{y}, \underline{y}', x)$$

$$I[\underline{y}] = \int_0^1 K(\underline{y}, \underline{y}', x)$$

*be **functionals** for a C^1 function:*

$$\underline{y} : [0,1] \to \mathbb{R}^n$$

*subject to the **boundary conditions**:*

$$\underline{y}(0) = \underline{y}_0 \qquad \underline{y}(1) = \underline{y}_1$$

*Now, suppose that $\underline{y}(x)$ **extremises** J, subject to the **isoperimetric constraint**:*

$$I[\underline{y}] = 0$$

*Then, if $\underline{y}$ is **not** an extremal of $I[\underline{y}]$, there is a **Lagrange multiplier** $\lambda \in \mathbb{R}$, such that $\underline{y}$ **extremises** the functional:*

$$P[\underline{y}] = J[\underline{y}] - \lambda I[\underline{y}]$$

(Theorem 9.2)

---

*Note, an **isoperimetric problem** is a **general** problem were a functional is extremised subject to a functional constraint.*
*Moreover, note the **similarities** with **Lagrange Multipliers** in SVC:*

$$g(\underline{x}) = 0 \iff I[\underline{y}] = 0$$

$$\nabla g(\underline{x}) \neq 0 \iff \underline{y} \textit{ doesn't extremise } I[\underline{y}]$$

$$F = f - \lambda g, \nabla F(\underline{x}, \lambda) = 0 \iff \underline{y}, \lambda \textit{ extremise } P = J - \lambda I$$

Proof. We prove this for a one-dimensional curve:

$$y : [0,1] \to \mathbb{R}$$

Suppose that y extremises J. Then, $J[y + s\varepsilon]$, where ε is a variation, has an extrema at $s = 0$. In other words:

$$\frac{d}{ds} J[y + s\varepsilon]\bigg|_{s=0} = 0$$

The problem is that we require that the variation ε also satisfy the isoperimetric constraint; however, we have no guarantee that such (non-zero) variations even exist. As such, we introduct a **correction term** via $r\eta$, such that:

$$\hat{y}(x) = y(x) + s\varepsilon(x) + r\eta(x)$$

where $r\eta(x)$ ensures that for fixed, arbitrary ε, and small s in the neighbourhood of 0, $\hat{y}$ satisfies the isoperimetric condition:

$$I[\hat{y}] = 0$$

Now, notice our functionals now become functions of 2 variables, r, s:

$$f(r,s) = J[\hat{y}] = J[y + s\varepsilon + r\eta]$$
$$g(r,s) = I[\hat{y}] = I[y + s\varepsilon + r\eta]$$

These function will be differentiable, assuming that L, K are differentiable. Now, we are in a setting where Lagrange Multipliers can be applied: y extremises J subject to I, so in particular it follows that $(0,0)$ is a critical point of f, and y obeys the isoperimetric constraint:

$$I[y] = 0 \implies g(0,0) = 0$$

Moreover, if we assume that $\nabla g(0,0) \neq 0$, by the Lemma on Lagrange multipliers, $\exists \lambda \in \mathbb{R}$, such that $(0,0,\lambda)$ is a critical point of:

$$F(r,s,\lambda) = f(r,s) - \lambda g(r,s)$$

In other words, we must have that $\nabla F(0,0,\lambda) = 0$, so we get the following equations:

$$\left.\frac{\partial F}{\partial s}\right|_{(0,0,\lambda)} = 0 \qquad \left.\frac{\partial F}{\partial r}\right|_{(0,0,\lambda)} = 0 \qquad \left.\frac{\partial F}{\partial \lambda}\right|_{(0,0,\lambda)} = 0$$

The last equation just tells us that:

$$g(0,0) = 0$$

which is just our constraint satisfaction equation $I[y] = 0$.

The remaining equations can be rewritten as:

$$\left.\frac{\partial}{\partial s}\left(J[\hat{y}] - \lambda I[\hat{y}]\right)\right|_{s=r=0} = 0 \qquad \left.\frac{\partial}{\partial r}\left(J[\hat{y}] - \lambda I[\hat{y}]\right)\right|_{s=r=0} = 0$$

Using:

$$\hat{y} = y + s\varepsilon + r\eta$$
$$\hat{y}' = y' + s\varepsilon' + r\eta'$$

we compute:

$$\begin{aligned}
\left.\frac{\partial}{\partial s}\left(J[\hat{y}]\right)\right|_{s=r=0} &= \int_0^1 \left.\frac{\partial}{\partial s} L(\hat{y}, \hat{y}', x)\right|_{s=r=0} dx \\
&= \int_0^1 \left.\frac{\partial L}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial s} + \frac{\partial L}{\partial \hat{y}'}\frac{\partial \hat{y}'}{\partial s}\right|_{s=r=0} dx \\
&= \int_0^1 \left.\frac{\partial L}{\partial \hat{y}}\varepsilon + \frac{\partial L}{\partial \hat{y}'}\varepsilon'\right|_{s=r=0} dx \\
&= \int_0^1 \left(\frac{\partial L}{\partial y}\varepsilon + \frac{\partial L}{\partial y'}\varepsilon'\right) dx
\end{aligned}$$

Now, if we apply integration by parts, we know that:

$$u = \frac{\partial L}{\partial y'} \qquad du = \frac{d}{dx}\frac{\partial L}{\partial y'}$$

$$dv = \varepsilon' \qquad v = \varepsilon$$

so:

$$\int_0^1 \frac{\partial L}{\partial y'}\varepsilon' dx = \left[\varepsilon \frac{\partial L}{\partial y'}\right]_0^1 - \varepsilon \int_0^1 \frac{d}{dx}\frac{\partial L}{\partial y'} dx$$

Assuming a fixed-point variation $\varepsilon(0) = \varepsilon(1) = 0$ (since $y(0), y(1)$ are fixed constants for any s), this gives us:

$$\frac{\partial}{\partial s} J[\hat{y}]\bigg|_{s=r=0} = \int_0^1 \varepsilon \left(\frac{\partial L}{\partial y} - \frac{d}{dx}\frac{\partial L}{\partial y}\right) dx$$

Similarly:

$$\frac{\partial}{\partial s} I[\hat{y}]\bigg|_{s=r=0} = \int_0^1 \varepsilon \left(\frac{\partial K}{\partial y} - \frac{d}{dx}\frac{\partial K}{\partial y}\right) dx$$

so it follows that:

$$\frac{\partial}{\partial s}\left(J[\hat{y}] - \lambda I[\hat{y}]\right)\bigg|_{s=r=0} = 0 \implies \int_0^1 \varepsilon \left[\left(\frac{\partial L}{\partial y} - \frac{d}{dx}\frac{\partial L}{\partial y}\right) - \lambda \left(\frac{\partial K}{\partial y} - \frac{d}{dx}\frac{\partial K}{\partial y}\right)\right] dx = 0$$

ε was arbitrary, so by the **Fundamental Lemma**:

$$\left(\frac{\partial L}{\partial y} - \frac{d}{dx}\frac{\partial L}{\partial y}\right) - \lambda \left(\frac{\partial K}{\partial y} - \frac{d}{dx}\frac{\partial K}{\partial y}\right) = 0$$

Defining $M = L - \lambda K$, this is nothing by the Euler-Lagrange Equation for M:

$$\frac{\partial M}{\partial y} - \frac{d}{dx}\frac{\partial M}{\partial y'} = 0$$

Now, we don't even need to compute:

$$\frac{\partial}{\partial r}\left(J[\hat{y}] - \lambda I[\hat{y}]\right)\bigg|_{s=r=0} = 0$$

Since we are assuming $\nabla g(0,0) \neq 0$, that means that g isn't singular at $(0,0)$, so by the implicit function theorem, the constraint:

$$g(r,s) = 0$$

tells us that for small values of s, $r = r(s)$ can be thought of as a function of s (alternatively, s could be a function of r). Because of this, η can be thought of as being part of a smaller variation class than ε. If alternatively we have $s = s(r)$, then η would've been our variation, and we would've reached the same conclusion, albeit by taking the partial derivative with respect to r.

The last step is to consider the effect of assuming $\nabla g(0,0) \neq 0$. If $\nabla g(0,0) = 0$, then this is equivalent to y extremising $I[y]$. Hence, for our Lagrange multipliers to apply, we must ensure that the constraint is satisfied ($I[y] = 0$) **and** that y doesn't extremise y. $\square$

2.2.1 Remarks

1. If y is an extremal of I ($\nabla g(0,0) = 0$), the problem is **abnormal**. In this case, $\exists \lambda, \mu \in \mathbb{R}$, not both 0, such that y is an extremal of:

$$\mu J - \lambda I$$

2. If we have many constraints, then we have to determine the extremal of the Lagrangian:

$$M = L - \lambda_1 K_1 - \lambda_2 K_2 - \ldots - \lambda_n K_n$$

3 Worked Examples

3.1 Solving Constrained Optimisation Problems with Functionals

The following are the steps to following when extremising $J[y]$, subject to $I[y] = 0$:

1. *Determine the* ***extremals*** *of $I[y]$, and ensure none of them satisfies:*

$$I[y] = 0$$

2. *Solve the* ***Euler-Lagrange Equation*** *for:*

$$M = L - \lambda K$$

3. *Determine* ***constants of integrations*** *by using the* ***boundary conditions****:*

$$y(0) = y_0 \qquad y(1) = y_1$$

4. *Determine λ by using the constraint $I[y] = 0$*

3.2 A Variant of the Isoperimetric Problem

Consider a curve:

$$y(x) : [0,1] \to \mathbb{R}, \qquad \forall x \in [0,1], y(x) \geq 0$$

where $y(0) = y(1) = 0$ are the only points were $y = 0$.
Determine y, such that it maximises the area under the curve, provided that he arclength of y is $\ell > 1$.

This corresponds to maximising the area functional:

$$A[x] = \int_0^1 y dx$$

subject to:

$$S[y] = \int_0^1 \sqrt{1 + y'(x)^2}dx = \ell$$

(to adhere to the formulation we have used, we define:

$$I[y] = S[y] - \ell$$

so that we have the constraint $I[y] = 0$)

---

(1)

We know that the arclength functional S has an extremum given by a straight line. Hence, $S[y] = 1$ in such a case. However, we are looking for y with arclength $\ell > 1$, so any y satisfying the constraint $I[y] = 0$ won't be an extremum of I, and thus, we can use Lagrange Multipliers.

(2)

We compute the Euler-Lagrange Equation for:

$$M = y - \lambda(\sqrt{1 + (y'(x))^2} - \ell)$$

Indeed:

$$\frac{\partial M}{\partial y} = 1$$

$$\frac{\partial M}{\partial y'} = -\frac{\lambda y'}{\sqrt{1 + (y')^2}}$$

which gives us the Euler-Lagrange Equation:

$$1 + \frac{d}{dx}\left(\frac{\lambda y'}{\sqrt{1 + (y')^2}}\right) = 0$$

Integrating once:

$$x + \frac{\lambda y'}{\sqrt{1 + (y')^2}} = c_1$$

We simplify:

$$\begin{aligned}
& x + \frac{\lambda y'}{\sqrt{1 + (y')^2}} = c_1 \\
\implies & \frac{\lambda^2 (y')^2}{1 + (y')^2} = (c_1 - x)^2 \\
\implies & \lambda^2 (y')^2 = (c_1 - x)^2 (1 + (y')^2) \\
\implies & (y')^2 (\lambda^2 - (c_1 - x)^2) = (c_1 - x)^2 \\
\implies & y' = \pm\sqrt{\frac{(c_1 - x)^2}{\lambda^2 - (c_1 - x)^2}} \\
\implies & y = \pm\int_0^1 \sqrt{\frac{(c_1 - x)^2}{\lambda^2 - (c_1 - x)^2}}dx
\end{aligned}$$

Now, using:

$$u = \lambda^2 - (c_1 - x)^2 \implies \frac{du}{dx} = -2(c_1 - x)$$

so:

$$\begin{aligned} y &= \pm \int \sqrt{\frac{(c_1 - x)^2}{u}} \times -\frac{du}{2(c_1 - x)} \\ &= \frac{1}{2} \int u^{-\frac{1}{2}} du \\ &= \sqrt{u} + c_2 \\ &= \sqrt{\lambda^2 - (c_1 - x)^2} + c_2 \\ \implies \lambda^2 &= (y - c_2)^2 + (x - c_1)^2 \end{aligned}$$

where we have taken y to be positive. Notice, y will trace out a segment of a circle of radius λ, centered at (c_1, c_2).

---

③

Now, we require that $y(0) = y(1) = 0$, so:

$$y(0) = 1 \implies \lambda^2 = c_2^2 + c_1^2$$

$$y(1) = 1 \implies \lambda^2 = c_2^2 + (1 - c_1)^2$$

so:

$$c_1^2 = (1 - c_1)^2 \implies 1 - 2c_1 + c_1^2 = c_1^2 \implies c_1 = \frac{1}{2}$$

Hence:

$$c_2 = \sqrt{\lambda^2 - \frac{1}{4}}$$

Thus, y traces out a segment of a circle, centered at:

$$\left(\frac{1}{2}, \sqrt{\lambda^2 - \frac{1}{4}}\right)$$

---

④

The final step is determining λ. For this we have to use our constraint:

$$I[y] = 0$$

which gives the arclength/perimeter of our segment about the x-axis. Indeed:

$$
\begin{aligned}
& S[y] - \ell = 0 \\
\implies & \int_0^1 \sqrt{1+(y')^2}dx = \ell \\
\implies & \int_0^1 \sqrt{1+\frac{(0.5-x)^2}{\lambda^2-(0.5-x)^2}}dx = \ell \\
\implies & \int_0^1 \sqrt{\frac{\lambda^2}{\lambda^2-(0.5-x)^2}}dx = \ell \\
\implies & \lambda\int_0^1 \sqrt{\frac{1}{\lambda^2-(0.5-x)^2}}dx = \ell
\end{aligned}
$$

Now define:

$$0.5 - x = \lambda\sin\theta \quad\implies\quad \frac{dx}{d\theta} = -\lambda\cos\theta$$

When $x=0$:

$$\theta = \arcsin\frac{1}{2\lambda}$$

and wehn $x=1$:

$$\theta = -\arcsin\frac{1}{2\lambda}$$

so that:

$$
\begin{aligned}
& \lambda\int_0^1 \sqrt{\frac{1}{\lambda^2-(0.5-x)^2}}dx = \ell \\
\implies & -\lambda^2\int_{\arcsin\frac{1}{2\lambda}}^{-\arcsin\frac{1}{2\lambda}} \sqrt{\frac{1}{\lambda^2-\lambda^2\sin^2\theta}}\cos\theta d\theta = \ell \\
\implies & -\lambda\int_{\arcsin\frac{1}{2\lambda}}^{-\arcsin\frac{1}{2\lambda}} d\theta = \ell \\
\implies & 2\lambda\arcsin\frac{1}{2\lambda} = \ell \\
\implies & 2\lambda\sin\frac{\ell}{2\lambda} = 1
\end{aligned}
$$

This is a transcendental equation, but it does have real solutions. For example, if $\ell = 2$, we get that $\lambda \approx \pm 0.52757$. We can plot this:

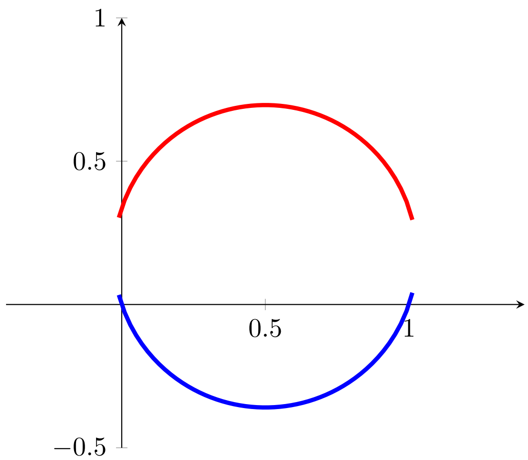


Figure 1: In red, $y(x) = \sqrt{\lambda^2 - (c_1 - x)^2} + c_2$ using $\lambda = 0.52757$. In blue the corresponding negative solution.

3.3 Solving the Isoperimetric Problem

The isoperimatric problem corresponds to extremising the functional:

$$P[x] = \int_0^1 \left[\frac{1}{2}(x^1 \dot{x}^2 - x^2 \dot{x}^1) - \lambda \sqrt{(\dot{x}^1)^2 + (\dot{x}^2)^2} \right] dt$$

3.4 The Catenary

Consider a uniform chain of length ℓ, hanging udner it sown weight from 2 poles a height h from the ground, and a distance $2\ell_0 < \ell$ apart.
Let $H(s)$ denote the height of the chaing, as a function of arclength s. The chain will try to minimise its ***gravitational potential energy****, which is given by:*

$$\int_0^\ell H(s)dx$$

If we parametrise the height of the chain by:

$$y(x) = H(s(x))$$

where $x \in [-\ell_0, \ell_0]$, we have that the potential energy will be:

$$\int_{-\ell_0}^{\ell_0} y(x)\sqrt{1 + y'(x)^2}dx$$

subject to the boundary conditions:

$$y(-\ell_0) = y(\ell_0) = h$$

and the isoperimetric constraint:

$$\int_{-\ell_0}^{\ell_0} \sqrt{1 + y'(x)^2}dx = \ell$$

---

(1)

Since we once have an arclength constraint, the curve which extremises it is a striaght line, whose length will be $2\ell_0$. Since $\ell > 2\ell_0$, any y satisfying the constraint won't be a straight line, so Lagrange Multipliers apply.

---

(2)

If we define:

$$L(y, y', x) = y\sqrt{1 + (y')^2} - \lambda\sqrt{1 + (y')^2} = (y - \lambda)\sqrt{1 + (y')^2}$$

then, and since the L doesn't depend explicitly on x, by Beltrami's Identity we get the the EL become:

$$\begin{aligned}
& y' \frac{\partial L}{\partial y'} - L = c \\
\implies & y' \frac{y'(y-\lambda)}{\sqrt{1+(y')^2}} - (y-\lambda)\sqrt{1+(y')^2} = c \\
\implies & \frac{y-\lambda}{\sqrt{1+(y')^2}} = c \\
\implies & \left(\frac{y-\lambda}{c}\right)^2 = 1 + (y')^2 \\
\implies & y' = \frac{1}{c}\sqrt{(y-\lambda)^2 - c^2} \\
\implies & x = c\int \frac{1}{\sqrt{(y-\lambda)^2 - c^2}} dy
\end{aligned}$$

Now recall:

$$\begin{aligned}
\cosh^2(x) - \sinh^2(x) = 1 \\
\frac{d}{dx}\cosh(x) = \sinh(x) \\
\frac{d}{dx}\sinh(x) = \cosh(x)
\end{aligned}$$

Thus, let:

$$y - \lambda = c\cosh(\theta) \quad \implies \quad \frac{dy}{d\theta} = c\sinh(\theta)$$

Our integral becomes:

$$\begin{aligned}
& x = c\int \frac{1}{\sqrt{(y-\lambda)^2 - c^2}} dy \\
\implies & x = c^2 \int \frac{1}{\sqrt{c^2\cosh^2(\theta) - c^2}} \sinh(\theta) d\theta \\
\implies & x = c\int d\theta \\
\implies & x = c\theta + d \\
\implies & y - \lambda = c\cosh\frac{x-d}{c}
\end{aligned}$$

---

(3)

Now, we apply the boundary conditions:

$$y(-\ell_0) = h \quad \implies \quad h - \lambda = c\cosh\frac{-\ell_0 - d}{c} = c\cosh\frac{\ell_0 + d}{c}$$

(since cosh is even)

$$y(\ell_0) = 0 \quad \implies \quad h - \lambda = c\cosh\frac{\ell_0 - d}{c}$$

Notice, this implies that:

$$\cosh\frac{\ell_0 + d}{c} = \cosh\frac{\ell_0 - d}{c} \quad \implies \quad d = 0$$

by using the fact that cosh is a **monotone** function.

Hence:

$$y = c \cosh \frac{x}{c} + \lambda$$

---

(4)

The work above already gives us λ:

$$h - \lambda = c \cosh \frac{\ell_0}{c} \implies \lambda = h - c \cosh \frac{\ell_0}{c}$$

so that:

$$y = c \cosh \frac{x}{c} + h - c \cosh \frac{\ell_0}{c}$$