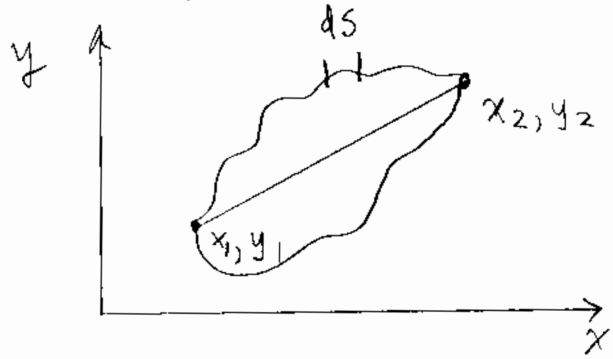
CALCULUS OF VARIATIONS

This classical subject forms the basis for modern formulations of field theory via the path integral technique (Feynman & Hibbs)

CLASSICAL PROBLEMS

[1] Shortest distance between 2 points



Formulating the problem: Given $y(x_1)=y_1$ $y(x_2)=y_2$ find the function $y=y(x)$ which minimizes (more generally which extremizes) the integral

$$I=\int_{(x_1,y_1)}^{(x_2,y_2)} ds \qquad (1)$$

infinitesimal arc length

In the 2-dimensional case: $ds^2 = dx^2+dy^2 \Rightarrow ds=\sqrt{dx^2+dy^2}$
$= dx\sqrt{1+[y'(x)]^2}$

We use this relation repeatedly:

$$\boxed{ds = dx\sqrt{1+\left(\frac{dy}{dx}\right)^2} = \sqrt{1+[y'(x)]^2}} \qquad (2)$$

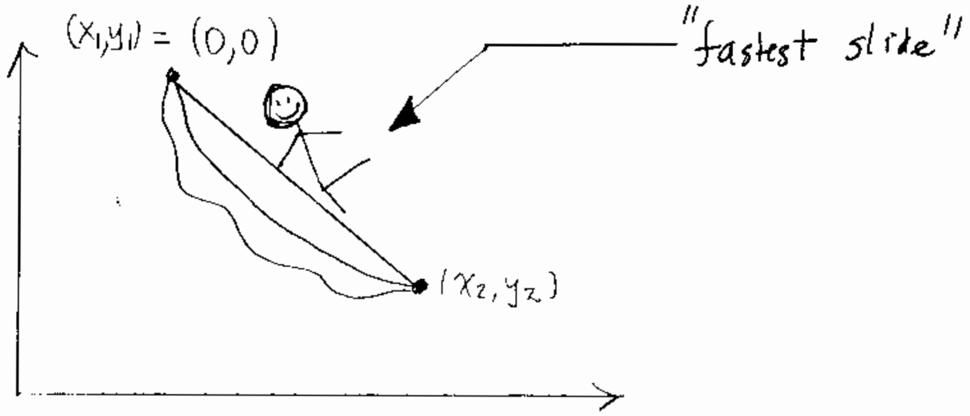
Hence the integral we seek to extremize is

$$I = \int_{x_1}^{x_2} dx \sqrt{1 + [y'(x)]^2}$$

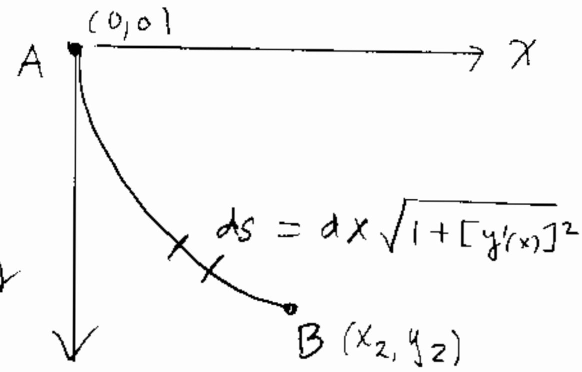
ANSWER is $y(x)$ (3)

GEODESIC = SHORTEST DISTANCE

[2] Brachistochrone = Shortest time (Johann Bernoulli - 1696)



Redraw this as:



To create the "fastest slide" we wish to minimize the time required to go from A to B:

$$\text{minimize } I = \int_{(0,0)}^{x_2, y_2} dt = \int_{0,0}^{x_2, y_2} \frac{ds}{v(x)} \leftarrow \text{speed} \qquad (4)$$

As in the previous example we wish to express everything in terms of x, so that we obtain an equation for $y(x)$.

To express everything in terms of x we use energy conservation:

$$\frac{1}{2} m v^2 - m g y = \text{const} = \frac{1}{2} m v_0^2 \quad (5)$$

g: 9.8 m/s²

v_0: speed at (0,0) ≡ 0

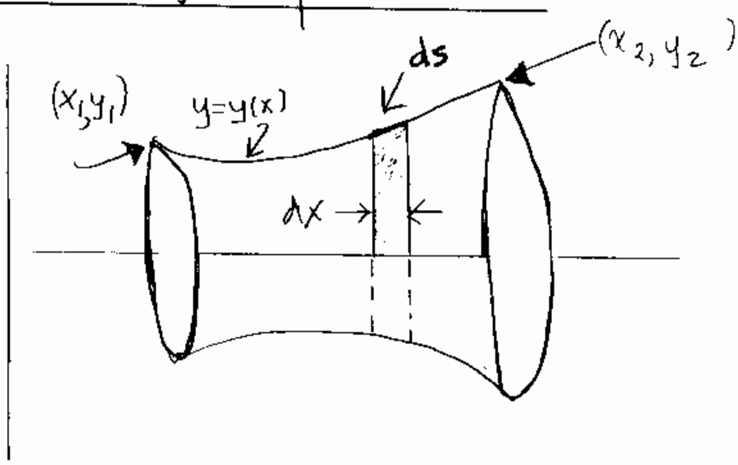
Hence $v^2(x) = 2 g y(x) \Rightarrow v = \sqrt{2 g y(x)}$ (6)

Thus the integral to be minimized is

$$\boxed{I = \int_{x=0}^{x=x_2} dx \, \frac{\sqrt{1+[y'(x)]^2}}{\sqrt{2 g y(x)}} = \int_0^{x_2} dx \sqrt{\frac{1+[y'(x)]^2}{2 g y(x)}}} \quad (7)$$

As in the previous example the "output" of this minimization is the function $y(x)$, which gives the shape of the "fastest slide."

[3] Minimum Surface of Revolution



The infinitesimal area dA of the shaded band is given by

$$dA = 2\pi y(x) \, ds = 2\pi y(x) \, dx \sqrt{1+[y'(x)]^2}$$

$$\boxed{\Rightarrow \text{Minimize } I = \int_{(x_1,y_1)}^{(x_2,y_2)} dA = 2\pi \int_{x_1}^{x_2} dx \, y(x) \sqrt{1+[y'(x)]^2}}$$

Comments:

(a) All of the 3 problems ask to minimize a function having the general form

the integral I

$$I = \int_{x_1}^{x_2} dx\, f(x, y, y'(x))$$

The "input" is $f(x, y, y')$ determined by the specific problem

The "output" = answer is the function $y(x)$.

(b) Although all these examples seek to minimize some integral I, in other situations we may want to maximize an integral. More generally we use the language extremize

$$\boxed{\text{extremize} = \text{minimize or maximize}}$$

SOLVING FOR $y(x)$: THE EULER-LAGRANGE EQUATIONS

Philosophical Point: As noted by FEYNMAN and others, even though we are seeking to minimize a global quantity such as the total distance between A & B, the only way this can be achieved using continuous (differentiable) functions is if these functions have the appropriate local behavior. This eventually leads to a differential equation for $y(x)$ via the EULER-LAGRANGE equations

Formulation of E-L Equations:

$$I = \int_{x_1}^{x_2} dx\, f(x, y, y') \tag{1}$$

Input: $f(x, y, y')$; $y(x_1) = y_1$; $y(x_2) = y_2$; x_1, x_2

Output: $y(x)$

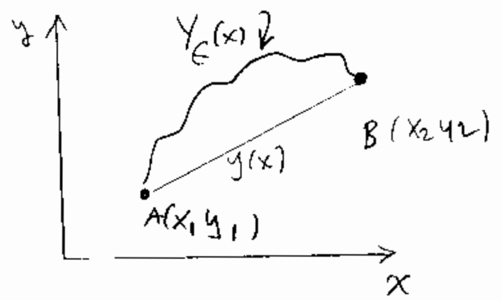
Method: Form a one-parameter family of comparison functions $Y_\epsilon(x)$

$$Y_\epsilon(x) = y(x) + \epsilon\,\eta(x) \tag{2}$$

small number

arbitrary function

correct answer



In order to compare the eventual "winner" $y(x)$ with all possible [differentiable] alternatives $\eta(x)$ must be arbitrary and satisfy,

$$\eta(x_1)=\eta(x_2)=0 \Rightarrow Y_\epsilon(x_1)=y(x_1)=y_1\;;\; Y_\epsilon(x_2)=y(x_2)=y_2 \qquad (3)$$

Next form the comparison integral

$$I(\epsilon)=\int_{x_1}^{x_2} dx\, f(x, Y_\epsilon(x), Y_\epsilon'(x)) \quad ; \quad Y_\epsilon'(x)=\frac{dY_\epsilon(x)}{dx}=y'(x)+\epsilon\,\underbrace{\eta'(x)}_{\partial\eta/\partial x} \qquad (4)$$

The point of introducing $I(\epsilon)$ is that <u>we know</u> (by <u>construction!</u>) <u>that the minimum value of</u> $I(\epsilon)$ [as a function of ϵ] <u>occurs at $\epsilon=0$</u>.

The condition for this minimum is

$$\frac{dI(\epsilon)}{d\epsilon}=0\equiv I'(\epsilon) \Rightarrow \text{minimum when } \boxed{I'(\epsilon)\Big|_{\epsilon=0}=0} \qquad (5)$$

To evaluate $I'(\epsilon)$ from (4) we note that Y_ϵ and Y_ϵ' are functions of ϵ, but x is not. Hence:

$$\frac{dI(\epsilon)}{d\epsilon}=\int_{x_1}^{x_2} dx\left\{\frac{\partial f}{\partial Y_\epsilon}\frac{\partial Y_\epsilon}{\partial \epsilon}+\frac{\partial f}{\partial Y_\epsilon'}\frac{\partial Y_\epsilon'}{\partial\epsilon}\right\}=\int_{x_1}^{x_2} dx\left\{\frac{\partial f}{\partial Y_\epsilon}\eta(x)+\frac{\partial f}{\partial Y_\epsilon'}\eta'(x)\right\} \qquad (6)$$

Next we impose the condition that the minimum occurs for $\epsilon=0 \Rightarrow Y_\epsilon(x)\to y(x)$

$$\therefore\; \frac{dI(\epsilon)}{d\epsilon}=0=\int_{x_1}^{x_2} dx\left\{\frac{\partial f}{\partial y}\eta(x)+\frac{\partial f}{\partial y'}\eta'(x)\right\} \qquad (7)$$

The second term in $\{\ldots\}$ can be partially integrated as follows:

Consider $\int_{x_1}^{x_2} dx \left\{ \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\eta\right]\right\} = \left[\frac{\partial f}{\partial y'}\eta\right]_{x_1}^{x_2} \equiv 0$ (8)

$\eta(x_2) = \eta(x_1) = 0$

But $\int_{x_1}^{x_2} dx \left\{ \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\eta\right]\right\} = \int \left\{ \frac{\partial f}{\partial y'}\eta'(x) + \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\cdot \eta \right\} dx = 0$ (9)

Hence in (7) we can write: $\int_{x_1}^{x_2} dx \frac{\partial f}{\partial y'}\eta'(x) = -\int_{x_1}^{x_2} dx \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\eta(x)$ (10)

Combining (7) & (10) we see that now both terms have a common factor of $\eta(x)$, so that

$$\frac{dI(\epsilon)}{d\epsilon} = 0 = \int_{x_1}^{x_2} dx \; \eta(x) \left\{ \frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right\} \quad (11)$$

Since $\eta(x)$ is an ARBITRARY function of x, the only way that the integral in (11) can vanish is if the expression in $\{\ldots\}$ vanishes and this gives the E-L equation

$$\boxed{\frac{\partial f(x,y,y')}{\partial y(x)} - \frac{d}{dx}\left(\frac{\partial f(x,y,y')}{\partial y'(x)}\right) = 0} \quad (12)$$

EULER-LAGRANGE EQUATION

Note that the solution to this equation is $y(x)$.
The derivative d/dx in (12) can be expanded as:

$$\frac{d}{dx} = \frac{\partial}{\partial x} + \frac{\partial y}{\partial x}\frac{\partial}{\partial y} + \frac{\partial y'}{\partial x}\frac{\partial}{\partial y'} = \frac{\partial}{\partial y} + y'\frac{\partial}{\partial y} + y''\frac{\partial}{\partial y'} \quad (13)$$

Hence E-L $\Rightarrow$ $\boxed{\frac{\partial f}{\partial y} - \frac{\partial^2 f}{\partial x \partial y'} - y'\frac{\partial^2 f}{\partial y \partial y'} - y''\frac{\partial^2 f}{\partial y'^2} = 0}$ (14)

COMMENTS:

[1] Depending on the specific problem the E-L equations in (12) or (14) may be the more useful form.

[2] The condition $I'(\epsilon)\big|_{\epsilon=0}=0$ is a necessary condition for an extremum, but not a sufficient condition, since this could also describe an inflection point

However, in practice it is usually obvious (at least in simple cases!) whether this gives a maximum, minimum or an inflection point.

Solutions of our 3-Problems

[1] Shortest distance

$$f(x,y,y') = \sqrt{1+[y'(x)]^2} \tag{1}$$

Note that this is the simplest case in which $f(x,y,y')$ does not actually depend on either x or y. This leads to some simplifications:

$$\underbrace{\frac{\partial f}{\partial y}}_{0} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0 \Rightarrow \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0 \Rightarrow \frac{\partial f}{\partial y'} = \text{const}(x) \equiv c_1 \tag{2}$$

$$\frac{\partial f}{\partial y'} = \frac{y'}{\sqrt{1+[y'(x)]^2}} = c_1 \Rightarrow [y']^2 = \frac{c_1^2}{1-c_1^2} \tag{3}$$

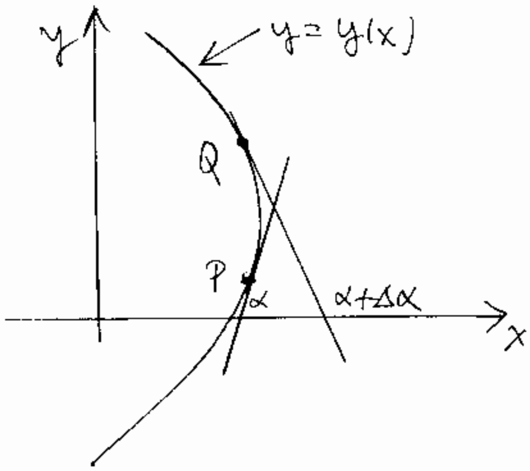
Hence $$y' = \frac{dy}{dx} = \frac{c_1}{\sqrt{1-c_1^2}} \equiv c_2 \Rightarrow \boxed{y(x) = c_2 x + c_3} \tag{4}$$

This is the expected answer, where the constants c_2 & c_3 can be easily chosen to pass through the points (x_1,y_1) and (x_2,y_2).

Digression on Curvature

There is another way to use the E-L equations to solve this problem, and this is to describe the straight line in terms of curvature. We introduce this concept here since we will need it later when we discuss the isoperimetric problem. We proceed to show that the E-L equations can be cast in a form which leads to the conclusion that the shortest distance between two points in a plane is the curve with zero curvature.

Curvature:



Intuitively the fact that the curve $y=y(x)$ has nonzero curvature is described by the fact that the angle α between the tangent to the curve and the x-axis changes as one moves along the curve. Hence we define the <u>Scalar curvature</u> K_p at a point P as

$$\boxed{K_p = \left|\frac{d\alpha}{ds}\right|_P} \quad (5)$$

From the figure $\tan\alpha = y' = dy/dx \Rightarrow \alpha = \tan^{-1} y'(x)$ (6)

$$\therefore\ K = \left|\frac{d\alpha}{ds}\right| = \left|\left(\frac{d\alpha}{dx}\right)\left(\frac{dx}{ds}\right)\right| \quad (7)$$

$$\text{Since } ds = dx\sqrt{1+[y'(x)]^2} \Rightarrow \frac{dx}{ds} = \frac{1}{\sqrt{1+[y']^2}} \quad (8)$$

Next we need $\dfrac{d\alpha}{dx} = \dfrac{d}{dx}\tan^{-1} y'(x)$

<u>Side Comment on Differentiating $\tan^{-1}(x)$</u>

Recall from last semester that $\ln z = \underbrace{\ln(x^2+y^2)^{1/2}}_{u(x,y)} + i\underbrace{\tan^{-1}\frac{y}{x}}_{v(x,y)}$ (9)

Since $\ln z$ is analytic u & v satisfy the CAUCHY-RIEMANN Conditions:

$$\partial u/\partial x = \partial v/\partial y\ ;\ \partial v/\partial x = -\partial u/\partial y$$

$$\therefore \quad \frac{\partial u}{\partial x} = \underbrace{\frac{\partial}{\partial x}\left[\ln(x^2+y^2)^{1/2}\right]}_{} = \frac{\partial}{\partial y}\tan^{-1}\frac{y}{x} \qquad (10)$$

$$\frac{1}{2}\,\frac{1}{x^2+y^2}\cdot 2x = \frac{x}{x^2+y^2} = \frac{(1/x)}{1+(y^2/x^2)} \qquad (11)$$

let $y/x \equiv w \Rightarrow \frac{\partial}{\partial y} = \frac{1}{x}\frac{\partial}{\partial w} \Rightarrow \frac{\partial}{\partial y}\tan^{-1}\frac{y}{x} = \frac{1}{x}\frac{\partial}{\partial w}\tan^{-1}w = \frac{1}{x}\frac{1}{1+w^2}$ (12)

Hence altogether (9)–(12) $\Rightarrow$

$$\boxed{\frac{\partial}{\partial w}\tan^{-1}w = \frac{1}{1+w^2}} \qquad (13)$$

Returning to our problem: $\frac{d\alpha}{dx} = \frac{d}{dx}\tan^{-1}y'(x) = \frac{1}{1+[y']^2}\cdot\frac{d}{dx}y'(x)$ (14)

$$\therefore \quad \boxed{\frac{d\alpha}{dx} = \frac{y''(x)}{1+[y']^2}} \qquad (15)$$

Combining (7), (8), & (15) we then find:

$$K = \left|\frac{d\alpha}{ds}\right| = \left|\frac{d\alpha}{dx}\cdot\frac{dx}{ds}\right| = \left|\frac{y''}{1+[y']^2}\cdot\frac{1}{\sqrt{1+[y']^2}}\right| \qquad (16)$$

Hence altogether:

$$\boxed{K = K(x) = \left|\frac{y''(x)}{\left[1+[y'(x)]^2\right]^{3/2}}\right|} \qquad (17)$$

Checks: (a) For a straight line $y(x) = c_2x + c_3 \Rightarrow y'' = 0 \Rightarrow K \equiv 0$ ✓ (18)

(b) For a circle $x^2+y^2 = a^2 \Rightarrow y(x) = (a^2-x^2)^{1/2}$ (19)

$$y'(x) = -x(a^2-x^2)^{-1/2} \quad ; \quad y'' = -a^2(a^2-x^2)^{-3/2} \qquad (20)$$

Combining Eqs. (17)-(20) we find:

$$K = \left| \frac{-a^2(a^2-x^2)^{-3/2}}{[1+x^2(a^2-x^2)^{-1}]^{3/2}} \right| = \left| \frac{-a^2(a^2-x^2)^{-3/2}}{\left[\frac{a^2}{a^2-x^2}\right]^{3/2}} \right| = \frac{1}{a} \checkmark \qquad (21)$$

This is intuitively reasonable since the curvature of a circle should be a constant, inversely proportional to the radius. ✓

Application to our Problem (Shortest Distance Between 2 Points)

Using the explicit form of the E-L equation as in 7,8 (14) we write

$$\frac{\partial f}{\partial y} - \frac{\partial^2 f}{\partial x \partial y'} - y' \frac{\partial^2 f}{\partial y \partial y'} - y'' \frac{\partial^2 f}{\partial y'^2} = 0 \qquad (22)$$

Since $f = \sqrt{1+[y']^2}$ does not explicitly depend on x or y, (22) reduces to

$$y'' \frac{\partial^2 f}{\partial y'^2} = 0 \qquad (23)$$

$$\frac{\partial f}{\partial y'} = y'\left[1+[y']^2\right]^{-1/2} \;; \quad \frac{\partial^2 f}{\partial y'^2} = \cdots = \left[1+[y']^2\right]^{-3/2} \qquad (24)$$

Combining (23) & (24) the E-L equations give

$$\boxed{0 = y'' \frac{\partial^2 f}{\partial y'^2} = \frac{y''}{[1+[y']^2]^{3/2}} \equiv K(x) = \text{curvature}} \qquad (25)$$

↳ see 11,12 (17)

Hence the E-L equations lead directly to the result that the shortest distance between 2 points is the "curve" with zero curvature, which is again a straight line.

Brachistochrone Problem [fastest slide]

$$f(x,y,y') = \left\{ \frac{1+[y'(x)]^2}{2gy(x)} \right\}^{1/2} \tag{1}$$

For simplicity we temporarily neglect the constant factor $1/\sqrt{2g}$.

Since $f(x,y,y')$ does not actually depend on x we can simplify the application of the E-L equations, as follows: Consider

$$\frac{d}{dx}\left[y'\frac{\partial f}{\partial y'} - f\right] = y''\frac{\partial f}{\partial y'} + y'\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{df}{dx} \tag{2}$$

$$= \cancel{y''\frac{\partial f}{\partial y'}} + y'\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \left(\frac{\partial f}{\partial x} + y'\frac{\partial f}{\partial y} + \cancel{y''\frac{\partial f}{\partial y'}}\right) \leftarrow \text{see 7,8(14)} \tag{3}$$

$$= -y'\left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right] - \frac{\partial f}{\partial x} \tag{4}$$

Hence
$$\boxed{\frac{d}{dx}\left[y'\frac{\partial f}{\partial y'} - f\right] = -y'\underbrace{\left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right]}_{=0 \text{ using E-L equations}} - \frac{\partial f}{\partial x}} \tag{5}$$

Thus the final identity is:

$$\boxed{\frac{d}{dx}\left[y'\frac{\partial f}{\partial y'} - f\right] = -\frac{\partial f}{\partial x}} \tag{6}$$

It follows from (6) that when $\partial f/\partial x = 0$ that

$$\frac{d}{dx}\left[y'\frac{\partial f}{\partial y'} - f\right] = 0 \Rightarrow \boxed{\left[y'\frac{\partial f}{\partial y'} - f\right] = \text{const}(x) = c_1} \tag{7}$$

This gives a first integral of motion: It has as a consequence that in (7) at most 1st derivatives appear in the equation for $y(x)$,

The result in (7) is familiar from classical mechanics:
There the E-L equations are:

$$\sum_i \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} \right) = 0 \qquad (8) \qquad L = \text{Lagrangian}$$

The analog of Eq. (7) is: $\left(\sum_i \dot{q}_i \underbrace{\frac{\partial L}{\partial \dot{q}_i}}_{p_i} - L \right) = \text{constant}$ (9)

We note that in (9) $\sum_i (\dot{q}_i p_i) - L = H$ = Hamiltonian, so that Eq. (9) is the statement of <u>energy conservation</u>.

Hence <u>if $L = L(q_i, \dot{q}_i, t)$ is independent of $t \Rightarrow$ energy conservation</u>

Returning to the brachistochrone problem,

$$f = \left\{ \frac{1+[y']^2}{y} \right\}^{1/2} \qquad (10)$$

Then:

$$y'\frac{\partial f}{\partial y'} - f = \frac{(y')^2}{y}\left\{ \frac{1+(y')^2}{y} \right\}^{-1/2} - \left\{ \frac{1+(y')^2}{y} \right\}^{1/2} = c_1 \qquad (11)$$

$$= \ldots (\text{some algebra}) = -\frac{1}{[y(1+(y')^2)]^{1/2}} \qquad (12)$$

This can be solved for $(y')^2$:

$$(y')^2 = \frac{1/c_1^2 - y}{y} \Rightarrow \boxed{y' = \frac{dy}{dx} = \sqrt{\frac{1/c_1^2 - y}{y}}} \qquad (13)$$

To find $y(x)$ let $1/c_1^2 \equiv 2a$ and write:

$$x = \int dy \frac{\sqrt{y}}{\sqrt{2a-y}} \qquad (14)$$

Define $y = 2a\sin^2\theta/2 \Rightarrow dy = 2a\sin\frac{\theta}{2}\cos\frac{\theta}{2}\,d\theta$ (15)

Hence $$x = \int \left(2a\sin\frac{\theta}{2}\cos\frac{\theta}{2}\,d\theta\right)\frac{\sqrt{\cancel{2a}\sin^2\theta/2}}{\sqrt{\cancel{2a}-\cancel{2a}\sin^2\theta/2}} \quad (16)$$

$$\therefore x = 2a\int d\theta\,\frac{\sin\theta/2}{\cancel{\cos}\,\theta/2}\cdot\left(\sin\theta/2\,\cancel{\cos\theta/2}\right) = 2a\int d\theta\,\underbrace{\sin^2\theta/2}_{\frac{1}{2}(1-\cos\theta)} \quad (17)$$

$$\therefore x = 2a\cdot\frac{1}{2}\int d\theta\,(1-\cos\theta) = a(\theta-\sin\theta)+x_0 \quad (18)$$

Hence combining the previous results the complete solution to the brachistochrone problem is

CYCLOID ⊛
$$\boxed{\begin{aligned} x &= a(\theta-\sin\theta)+x_0 \\ y &= 2a\sin^2\theta/2 = a(1-\cos\theta)\end{aligned}} \quad (19)$$

Note that the solution contains 2 integration constants a, x_0 which can be used to guarantee that the curve (cycloid) passes through the points (x_1,y_1) and (x_2,y_2). [Recall that the solution for the straight line also had 2 integration constants, reflecting the fact that the E-L equations give a 2ND order differential equation.

⊛ Curve swept out by a fixed point on the perimeter of a rolling circle.

Huygens: A mass point oscillating without friction on a vertical cycloid under the influence of gravity has a period which is strictly independent of amplitude. Hence the cycloid is also a TAUTOCHRONE.

[3] Minimum Surface of Revolution

$$f = f(x,y,y') = y\sqrt{1+[y']^2} \qquad (20)$$

As in the previous case $\partial f/\partial x = 0 \Rightarrow y'\frac{\partial f}{\partial y'} - f = \text{const} \equiv c_1$ (21)

From (20) $$\frac{\partial f}{\partial y'} = \frac{y y'}{\sqrt{1+[y']^2}} \qquad (22)$$

Hence E-L $\Rightarrow$ $$\frac{y(y')^2}{\sqrt{1+(y')^2}} - y\sqrt{1+(y')^2} = c_1 \qquad (23)$$

$$\therefore\ -c_1 = \frac{y}{\sqrt{1+[y']^2}} \Rightarrow \boxed{y' = \frac{dy}{dx} = \sqrt{\frac{y^2}{c_1^2} - 1}} \qquad (24)$$

$$\therefore x = \int \frac{dy}{\sqrt{y^2/c_1^2 - 1}} = c_1 \cosh^{-1}(y/c_1) + c_2$$

Inverting this: $$\boxed{y(x) = b\cosh\frac{x-a}{b}}$$ $a = c_2$, $b = c_1$

CATENARY (*)

(*) Latin for "chain"; This is the curve which describes the shape of a hanging chain (under the influence of gravity).
[English: concatenate = to link together as in a chain]

This gives the minimum surface of revolution assuming that $y''(x)$ exists. There is, however, a solution which is not differentiable called the GOLDSCHMIDT SOLUTION

ISOPERIMETRIC PROBLEMS

This is a class of problems where it is given to extremize one quantity subject to the constraint that another quantity remain fixed.

For example: A farmer with a fixed amount of fence material wants to enclose the maximum possible area for his horse to graze.

Formulation: We are given to extremize the integral I

$$I=\int_{x_1}^{x_2} dx\, f(x,y,y') \qquad y(x_1)=y_1 \quad y(x_2)=y_2 \qquad (1)$$

Subject to the constraint that some other integral J remains fixed:

$$J=\int_{x_1}^{x_2} dx\; g(x,y,y') = \text{constant} \qquad (2)$$

The solution to this problem requires LAGRANGE MULTIPLIERS which we review now.

Review of Lagrange Multipliers: [ARFKEN]

Consider the function $f(x,y,z)$ and evaluate

$$df=\left(\frac{\partial f}{\partial x}\right)dx+\left(\frac{\partial f}{\partial y}\right)dy+\left(\frac{\partial f}{\partial z}\right)dz \qquad (3)$$

To find an extremum of f we set $df=0$. Since the variations dx, dy, and dz are arbitrary, the only way that $df=0$ can hold is if

$$\boxed{\partial f/\partial x=\partial f/\partial y=\partial f/\partial z=0} \quad (4)$$

Suppose now that we find that there is a constraint in the problem which can be expressed by some equation of the form

$$g(x,y,z)=0 \qquad (5)$$

Because of this constraint, the variations dx, dy, dz are no longer independent, which was the assumption needed to derive the condition in (4). Specifically

$$g(x,y,z)=0 \Rightarrow 0=(\partial g/\partial x)dx + (\partial g/\partial y)dy + (\partial g/\partial z)dz \qquad (6)$$

Since $\partial g/\partial x$, $\partial g/\partial y$, and $\partial g/\partial z$ are known, one can solve (6) explicitly for dz, for example, in terms of dx and dy:

$$dz = -(\partial g/dz)^{-1}\left[(\partial g/\partial x)dx + (\partial g/\partial y)dy\right] \qquad (7)$$

Because of this equation, dz is dependent on dx and dy and the previous arguments to find the extremum are not valid.

One can of course eliminate dz simply by using (7) to replace dz everywhere. This can be done but is tedious.

There is another way to eliminate dz using Lagrange multipliers: Using (3) & (6) form the function $f(x,y,z)+\lambda g(x,y,z)$. Then the extremum

$$df = 0 \qquad (8)$$

can be rewritten as $df + \lambda dg = 0$, since $g(x,y,z)=0 \Rightarrow dg=0$

This gives the following equation:

$$df(x,y,z)+\lambda dg(x,y,z) = 0 = \left(\frac{\partial f}{\partial x}+\lambda\frac{\partial g}{\partial x}\right)dx + \left(\frac{\partial f}{\partial y}+\lambda\frac{\partial g}{\partial y}\right)dy + \left(\frac{\partial f}{\partial z}+\lambda\frac{\partial g}{\partial z}\right)dz \qquad (9)$$

Since dz (for example) is not linearly independent it should not appear in (9), and one way of ensuring this is to choose λ to make the coefficient of dz vanish:

$$\frac{\partial f}{\partial z}+\lambda\frac{\partial g}{\partial z}=0 \qquad (10)$$

Having eliminated dz, the expressions which give the extremum are now: $\frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x} = 0$; $\frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y} = 0$ (11)

When these equations are solved, $df = 0$ and $f(x,y,z)$ is an extremum subject to the constraint $g(x,y,z) = 0$.

SUMMARY:

- We want to find the extremum of $f(x,y,z)$ subject to the constraint $g(x,y,z) = 0$. Finding the extremum means finding x_0, y_0, z_0.
- Once we introduce the Lagrange multiplier λ, we then have 4 unknowns to solve for: x_0, y_0, z_0, λ
- These 4 quantities are then determined by the following 4 equations

$$\left.\begin{aligned} &\frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x} = 0 \qquad (12a) \\ &\frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y} = 0 \qquad (12b) \end{aligned}\right\} \text{Eqs. (11) above}$$

$$\left.\frac{\partial f}{\partial z} + \lambda \frac{\partial g}{\partial z} = 0 \right\} \text{Eq. (10)} \qquad (12c)$$

$$\left. g(x,y,z) = 0 \right\} \text{Eq. (5)} \qquad (12d)$$

Example 1: extremize f(x,y) = 3x+5y, given that x^2+y^2=136

Example: Application of Lagrange Multipliers in QM

[ARFKEN] The ground state energy of a particle in rectangular QM box whose sides are a, b, c is given by $(E \equiv (4\pi^2)^{-1}\mathrm{E})$ actual energy

$$E = E(a,b,c) = \frac{\hbar^2}{8m}\left(\frac{1}{a^2} + \frac{1}{b^2} + \frac{1}{c^2}\right) \qquad (13)$$

We wish to find the shape of the box (i.e., a, b, c) such that E is a minimum for a fixed volume

$$V = V(a,b,c) = abc = \text{constant} \equiv k \qquad (14)$$

Solution: In our previous notation let $f(a,b,c) = E(a,b,c)$ and

$$g(a,b,c) = V(a,b,c) - k = 0 = abc - k \qquad (15)$$

We then solve: $$\underbrace{\frac{\partial E}{\partial a} + \lambda\frac{\partial V}{\partial a} = 0}_{\text{this may be viewed as eliminating } da} \;;\; \frac{\partial E}{\partial b} + \lambda\frac{\partial V}{\partial b} = 0 \;;\; \frac{\partial E}{\partial c} + \lambda\frac{\partial V}{\partial c} = 0 \qquad (16)$$

$$\frac{\partial E}{\partial a} + \lambda\frac{\partial V}{\partial a} = -\frac{\hbar^2}{4ma^3} + \lambda bc = 0 \qquad (17a)$$

Similarly:

$$\frac{\partial E}{\partial b} + \lambda\frac{\partial V}{\partial b} = 0 = \frac{-\hbar^2}{4mb^3} + \lambda ac = 0 \qquad (17b)$$

$$\frac{\partial E}{\partial c} + \lambda\frac{\partial V}{\partial c} = 0 = -\frac{\hbar^2}{4mc^3} + \lambda ab = 0 \qquad (17c)$$

Multiplying this equations in turn by a, b, c then gives:

$$\lambda abc = \frac{\hbar^2}{4ma^2} \;;\; \lambda abc = \frac{\hbar^2}{4mb^2} \;;\; \lambda abc = \frac{\hbar^2}{4mc^2} \qquad (18)$$

The solution to these equations is obviously $\boxed{a = b = c}$ (19)

$\Rightarrow$ rectangular box $\rightarrow$ cube

Note that we have solved the problem without having to actually determine λ. However, if we wish to solve for λ to give it a physical interpretation we can write:

$$\lambda abc = \frac{\hbar^2}{4ma^2} \xrightarrow{a=b=c} \lambda a^3 = \frac{\hbar^2}{4ma^2} \Rightarrow \boxed{\lambda = \frac{\hbar^2}{4ma^5}} \quad (20)$$

To interpret λ we note from (13) & (19) that

$$E = \frac{\hbar^2}{8m}\left(\frac{1}{a^2} + \frac{1}{b^2} + \frac{1}{c^2}\right) \rightarrow \frac{3\hbar^2}{8ma^2}$$

Hence the energy density is given by $\dfrac{E}{V} = \dfrac{(3/8)\hbar^2/ma^2}{a^3} = \dfrac{3}{8}\dfrac{\hbar^2}{ma^5}$

If we convert this to physical energy units, $E = 4\pi^2 E$ etc then

$$\boxed{\lambda = \frac{3\pi^2}{2}\frac{E}{V}}$$

So λ is a measure of the energy density

Return to the Isoperimetric Problem:

Minimize or Maximize $\quad I = \int_{x_1}^{x_2} dx\, f(x, y, y') \qquad (1)$

Subject to the constraint $\quad J = \int_{x_1}^{x_2} dx\, g(x, y, y') = \text{constant} \qquad (2)$

We proceed in analogy to the previous case (no constraints):

Form a 2-parameter comparison function

$$Y_{\epsilon_1 \epsilon_2}(x) = \underbrace{y(x)}_{\text{true answer}} + \epsilon_1 \eta_1(x) + \epsilon_2 \eta_2(x) \qquad (3)$$

One can think of the extra function as ensuring that the constraint (2) remains enforced while $\eta_1(x)$ is varying to sweep out the space of comparison functions. Because of this <u>ϵ_1 and ϵ_2 are not really independent</u>; but by introducing a Lagrange multiplier λ below, we can treat them as if they were independent. As before we impose the conditions:

$$\eta_1(x_1) = \eta_1(x_2) = \eta_2(x_1) = \eta_2(x_2) = 0 \qquad (4)$$

Generalizing the previous approach we form the integrals

$$I(\epsilon_1, \epsilon_2) = \int_{x_1}^{x_2} dx\, f(x, Y_{\epsilon_1,\epsilon_2}, Y'_{\epsilon_1,\epsilon_2}) \qquad (5)$$

$$J(\epsilon_1, \epsilon_2) = \int_{x_1}^{x_2} dx\, g(x, Y_{\epsilon_1,\epsilon_2}, Y'_{\epsilon_1,\epsilon_2}) \qquad (6)$$

Then form the expression

$$K(\epsilon_1,\epsilon_2) = I(\epsilon_1,\epsilon_2) + \lambda J(\epsilon_1,\epsilon_2) = \int_{x_1}^{x_2} dx \underbrace{\left\{ f(x,Y,Y') + \lambda g(x,Y,Y') \right\}}_{h(x,Y,Y')} \quad (7)$$

Lagrange multiplier

Note that although we really wish to differentiate $I(\epsilon_1,\epsilon_2)$ to find the extremum, we can in fact differentiate $K = I + \lambda J$ since both λ and J are to be treated as constants. Thus the condition for an extremum in $K(\epsilon_1,\epsilon_2)$ [which is really going to be an extremum in I] is

$$0 = \left.\frac{\partial K(\epsilon_1,\epsilon_2)}{\partial \epsilon_1}\right|_{\epsilon_1=\epsilon_2=0} = \left.\frac{\partial K(\epsilon_1,\epsilon_2)}{\partial \epsilon_2}\right|_{\epsilon_1=\epsilon_2=0} \quad (8)$$

As noted above, there are no terms like $\partial K/\partial\lambda$ or $\partial K/\partial J$.

Proceeding as in the previous case we find [for $j = 1$ or 2]

$$\frac{\partial K}{\partial \epsilon_j} = \int dx \left\{ \frac{\partial h}{\partial Y}\frac{\partial Y}{\partial \epsilon_j} + \frac{\partial h}{\partial Y'}\frac{\partial Y'}{\partial \epsilon_j} \right\} = \int dx \left\{ \frac{\partial h}{\partial Y}\eta_j(x) + \frac{\partial h}{\partial Y'}\eta_j'(x) \right\} \quad (9)$$

As before, <u>we know by construction</u> that the minimum occurs at $\epsilon_1 = \epsilon_2 = 0$ in which case $Y(x) \to y(x)$. This gives

$$0 = \left.\frac{\partial K(\epsilon_1,\epsilon_2)}{\partial \epsilon_j}\right|_{\epsilon_j=0} = \int_{x_1}^{x_2} dx \left\{ \frac{\partial h}{\partial y}\eta_j(x) + \frac{\partial h}{\partial y'}\eta_j'(x) \right\} \quad j = 1,2 \quad (10)$$

We can integrate the term $\sim \eta_j'(x)$ by parts so that

$$\int dx \frac{\partial h}{\partial y'}\eta_j'(x) = -\int \frac{d}{dx}\left(\frac{\partial h}{\partial y'}\right)\eta_j(x)\, dx \quad (11)$$

Using the previous results, the equation for the extremum then becomes

$$0 = \int_{x_1}^{x_2} dx \left\{ \frac{\partial h}{\partial y} - \frac{d}{dx}\left(\frac{\partial h}{\partial y'}\right)\right\} \eta_j(x) \qquad (12)$$

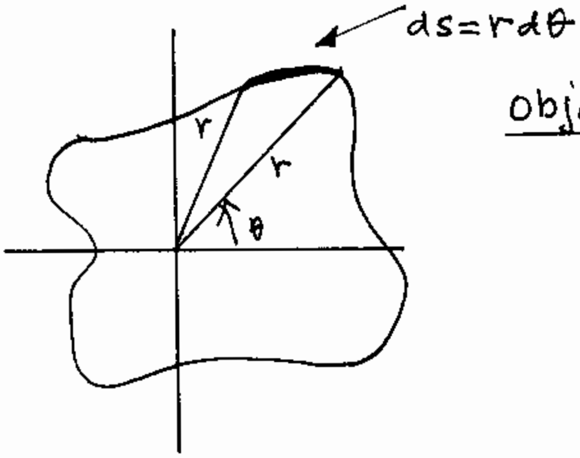
Since the $\eta_j(x)$ are arbitrary functions we have finally,

$$\boxed{\frac{\partial h(x,y,y')}{\partial y} - \frac{d}{dx}\left(\frac{\partial h(x,y,y')}{\partial y'}\right) = 0\,;\; h = f + \sum_{i=1}^{N} \lambda_i g_i} \qquad (13)$$

Note that in the last step we have generalized from one constraint to an arbitrary number N of constraints: For each such constraint $J_i = \int dx\, g_i(x,y,y')$ there is a corresponding Lagrange multiplier λ_i.

APPLICATIONS:

[1] ORIGINAL ISOPERIMETRIC PROBLEM



Object: maximize the area A enclosed by the curve while keeping perimeter = const $= P$

$$dA = \frac{1}{2} r \cdot r\,d\theta = \frac{1}{2} r^2 d\theta \Rightarrow A = \frac{1}{2}\int_0^{2\pi} d\theta\, r^2 \text{ [redacted]} \qquad (1)$$

$$\equiv r'$$

$$P = \text{perimeter} = \int_0^{2\pi} ds \qquad ds^2 = dr^2 + r^2 d\theta^2 \Rightarrow ds = d\theta\left[\left(\frac{dr}{d\theta}\right)^2 + r^2\right]^{1/2} \qquad (2)$$

Hence the problem is: <u>maximize</u>: $A = \frac{1}{2}\int_0^{2\pi} d\theta \cdot r^2$ (3)

<u>hold constant</u> $P = \int_0^{2\pi} d\theta\left[r^2 + r'^2\right]^{1/2}$ (4)

<u>Solution</u>: Let $f(\theta, r, r') = \frac{1}{2} r^2$; $g(\theta, r, r') = \left[r^2 + r'^2\right]^{1/2}$ (5)

$$h = f + \lambda g \Rightarrow \boxed{\frac{\partial h}{\partial r} - \frac{d}{d\theta}\left(\frac{\partial h}{\partial r'}\right) = 0} \qquad (6)$$

We find: $\frac{\partial h}{\partial r} = r + \lambda r\left[r^2 + r'^2\right]^{-1/2}$ (7)

To evaluate $\frac{d}{d\theta}(\ldots)$ write:

$$\boxed{\frac{d}{d\theta} = \frac{\partial}{\partial\theta} + r'\frac{\partial}{\partial r} + r''\frac{\partial}{\partial r'}} \qquad (8)$$

analog of previous result.

Hence $\frac{d}{d\theta}\left(\frac{\partial h}{\partial r'}\right) = \left(\cancel{\frac{\partial}{\partial \theta}} + r'\frac{\partial}{\partial r} + r''\frac{\partial}{\partial r'}\right)\left(\frac{\partial h}{\partial r'}\right)$ (9)

$$\frac{\partial h}{\partial r'} = \left\{\lambda r' [r^2+r'^2]^{-1/2}\right\} \; ; \; r'\frac{\partial}{\partial r}\{\cdots\} = -r r'^2 \lambda [r^2+r'^2]^{-3/2} \qquad (10)$$

$$r''\frac{\partial}{\partial r'}\{\cdots\} = r''\lambda\left\{[r^2+r'^2]^{-1/2} - \tfrac{1}{2} r' [r^2+r'^2]^{-3/2} \cdot 2r'\right\}$$

$$= \ldots \text{(algebra)} \ldots = \lambda r'' r^2 [r^2+r'^2]^{-3/2} \qquad (11)$$

Combining the previous results we have: E-L $\Rightarrow$

$$0 = r + \lambda r [r^2+r'^2]^{-1/2} + \lambda r r'^2 [r^2+r'^2]^{-3/2} - \lambda r'' r^2 [r^2+r'^2]^{-3/2} \qquad (12)$$

cancel factor of r:

$$\Rightarrow \; 0 = 1 + \lambda [r^2+r'^2]^{-3/2}\left\{[r^2+r'^2] + r'^2 - r'' r\right\} \qquad (13)$$

$$0 = 1 + \lambda [r^2+r'^2]^{-3/2}\left\{r^2 + 2r'^2 - r'' r\right\} \qquad (14)$$

Finally:

$$\boxed{\frac{r^2 + 2r'^2 - r'' r}{[r^2+r'^2]^{3/2}} \equiv \text{CURVATURE} = \left|\frac{1}{\lambda}\right|} \qquad (15)$$

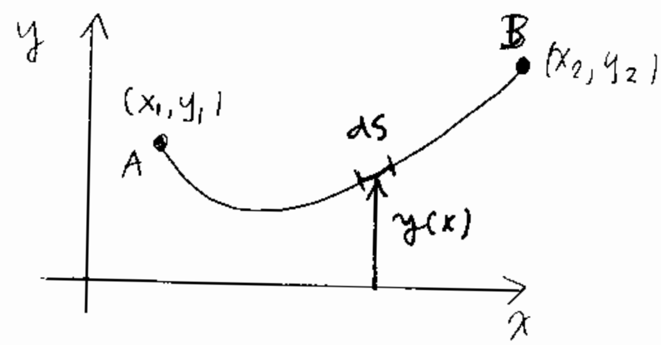
This shows that the curve which maximizes the area for a fixed perimeter is the curve with constant curvature $= 1/\lambda$

$\Rightarrow$ CIRCLE! In this case the Lagrange multiplier λ has a physical interpretation: λ = RADIUS OF CIRCLE $= R$

$P = 2\pi R =$ [redacted] $=$ LENGTH OF FENCE AVAILABLE $= 2\pi\lambda$

[2] Hanging Massive Chain

The problem here is to find the shape of a massive chain hanging between two points (x_1, y_1) & (x_2, y_2).



This is an isoperimetric problem because we are seeking to minimize the gravitational potential energy, subject to the constraint that the chain have a fixed length (and pass through the points A & B).

Solution: Let σ be the (constant) mass/unit length of the chain: The gravitational potential of the mass element shown is

$$dV = dm\,|\vec{g}|\,y(x) = (\sigma\, \underbrace{ds}_{dx\sqrt{1+(y')^2}})\,\underbrace{|\vec{g}|}_{9.8\,m/s^2}\,y(x) \qquad (1)$$

$$V = 0 \;@\; y = 0$$

Then: minimize: $$V = \sigma|\vec{g}|\int_{x_1}^{x_2} dx\,\sqrt{1+(y')^2}\cdot y \qquad (2)$$

constraint: $$L = \int_{x_1}^{x_2} dx\,\sqrt{1+(y')^2} = \text{constant} \qquad (3)$$

Hence
$$\left.\begin{aligned} f(x,y,y') &= \sigma|\vec{g}|\sqrt{1+(y')^2}\cdot y \\ g(x,y,y') &= \sqrt{1+(y')^2} \end{aligned}\right\} h = f + \lambda g \qquad (4)$$

E-L $\Rightarrow \dfrac{\partial h}{\partial y} - \dfrac{d}{dx}\left(\dfrac{\partial h}{\partial y'}\right) = 0$; $h = (\sigma|\vec{g}|y+\lambda)\sqrt{1+(y')^2}$

Since $h(x,y,y')$ does not depend explicitly on x it follows from our previous work that

$$y'\frac{\partial h}{\partial y'} - h = \text{Const} = -c_1 \; ; \; \frac{\partial h}{\partial y'} = (\sigma|\vec{g}|y+\lambda)\frac{y'}{\sqrt{1+(y')^2}} \qquad (5)$$

Hence E-L $\Rightarrow$ $\boxed{c_1 = \cdots = \dfrac{(\sigma|\vec{g}|y+\lambda)}{\sqrt{1+(y')^2}}}$ (6)

This is similar to the result obtained previously when studying the minimum surface of revolution. Hence using that result we can make the substitution $y \to (\sigma|\vec{g}|y+\lambda)$ to obtain immediately:

$$x = \int \frac{dy}{\sqrt{\dfrac{(\sigma|\vec{g}|y+\lambda)^2}{c_1^2} - 1}} \quad ; \text{ let } \sigma|\vec{g}|y+\lambda \equiv t \qquad (7)$$
$$\sigma|\vec{g}|dy = dt$$

$$\therefore x = \frac{1}{\sigma|\vec{g}|}\int \frac{dt}{\sqrt{(t/c_1)^2-1}} = \frac{1}{\sigma|\vec{g}|} c_1 \cosh^{-1}(t/c_1) + c_2 \qquad (8)$$

$$\therefore \cosh\left\{\left(\frac{x-c_2}{c_1}\right)\sigma|\vec{g}|\right\} = \frac{\sigma|\vec{g}|y+\lambda}{c_1} \qquad (9)$$

$$\boxed{\therefore y(x) = \frac{1}{\sigma|\vec{g}|}\left\{c_1 \cosh\left[\left(\frac{x-c_2}{c_1}\right)\sigma|\vec{g}|\right] - \lambda\right\}} \qquad (10)$$

This is the equation of a CATENARY: In fact this curve is so named because in Latin CATENA $\equiv$ CHAIN.

Note: $y(x)$ has 3 constants to be determined. These are fixed by having $y(x)$ pass through $A = (x_1, y_1)$ $B = (x_2, y_2)$, while having a length L.

VARIATIONAL METHODS IN CLASSICAL MECHANICS:

SEVERAL DEPENDENT VARIABLES

We wish to extremize the integral

$$I = \int_{t_1}^{t_2} dt\, f(q_i, \dot{q}_i, t) \qquad \dot{q}_i = dq_i/dt \qquad i = 1, \ldots, N \qquad (1)$$

$q_i(t_1)$ and $q_i(t_2)$ are specified

This is similar to the problem we have already solved, and so we can write

$$\boxed{\frac{\partial f}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{q}_i}\right) = 0} \quad (2)$$

As before, if f is explicitly independent of t, then we can immediately write down a first integral of the motion:

$$\boxed{\sum_i \left(\dot{q}_i \frac{\partial f}{\partial \dot{q}_i}\right) - f = \text{constant} = C_1} \quad (3)$$

f is independent of t

In addition, if f is independent of a coordinate q_i then we can find another first integral of the motion:

$$\frac{d}{dt}\left(\frac{\partial f}{\partial \dot{q}_i}\right) = 0 \Rightarrow \boxed{\frac{\partial f}{\partial \dot{q}_i} = \text{Constant} = C_2^i} \quad (4)$$

CONNECTION TO MECHANICS: If a classical mechanical system is Conservative (conserves energy) then it can be described by a potential energy function $V(q_i)$, where q_i are the generalized coordinates of the system. The Lagrangian of the system then has the form (T = kinetic energy)

$$L = T(\dot{q}_i) - V(q_i) \qquad (5)$$

The motion of the system is determined by HAMILTON'S PRINCIPLE

$$\left\| \; I = \int_{t_1}^{t_2} dt\, L(q_i, \dot{q}_i, t) = \text{EXTREMUM} \; \right\| \quad (6)$$

Carrying out the same analysis as before leads to the usual Lagrange equations:

$$\boxed{\frac{\partial L}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = 0}$$ LAGRAGE EQUATIONS (7)

It follows immediately that if L is not an explicit function of t then a first integral of the motion is:

$$\sum_i\left(\dot{q}_i \frac{\partial L}{\partial \dot{q}_i}\right) - L = c_1 = \text{constant} \qquad (8)$$

Since V is assumed to depend only on q_i and not on $\dot{q}_i \Rightarrow \frac{\partial L}{\partial \dot{q}_i} = \frac{\partial T}{\partial \dot{q}_i}$ (9)

Then (8) $\Rightarrow$

$$\sum_i \dot{q}_i \frac{\partial T}{\partial \dot{q}_i} - L = c_1 \qquad (10)$$

Typically T is a homogeneous function of order 2 in the variables $\dot{q}_i$: For a collect of free particles, for example, all having a common mass m

$$T = \frac{1}{2}\sum_i m(\dot{q}_i)^2 \Rightarrow \frac{\partial T}{\partial \dot{q}_i} = m\dot{q}_i = p_i \Rightarrow \sum_i \dot{q}_i \frac{\partial T}{\partial \dot{q}_i} = \sum_i m\dot{q}_i^2 = 2T \qquad (11)$$

Hence (10) $\Rightarrow$

$$\boxed{c_1 = \sum_i \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L = 2T - (T - V) = T + V = \underbrace{H}_{\text{TOTAL ENERGY}} = \text{const}} \qquad (12)$$

Thus when $L = L(t)$ this first integral of the motion expresses energy conservation

If $L \neq L(q_i)$ for some q_i then from (4)

$$\boxed{p_i \equiv \frac{\partial L}{\partial \dot{q}_i} = \text{Constant} \equiv c_2} \qquad (13)$$

Hence the invariance of L with respect to a particular coordinate q_i implies the conservation of the canonically conjugate p_i = momentum.

As a trivial example, for a free particle

$$L = T - V = \tfrac{1}{2} m \dot{q}_i^2$$

The energy integral gives $c_1 = T = \frac{1}{2} m \dot{q}_i^2 = \text{const}$ and the momentum integral gives

$$\frac{\partial L}{\partial \dot{q}_i} = m\dot{q}_i = p_i = \text{const.} \tag{14}$$

This is an example of a more general result NOETHER'S THEOREM:

If L is independent of this variable	Then the following quantity is conserved
t $\longleftrightarrow$	E = TOTAL ENERGY
q_i = coordinate $\longleftrightarrow$	p_i = MOMENTUM CONJUGATE TO q_i
θ = angular coordinate $\longleftrightarrow$	$L = mr^2\dot{\theta}$ = angular momentum

Note that the product of each pair of variables (connected by arrows $\longleftrightarrow$ above) is of the same dimensions as PLANCK'S CONSTANT $\hbar$. Such variables are said to be CANONICALLY CONJUGATE. The Heisenberg uncertainty relations connect such variable pairs:

$$\Delta q\, \Delta p \approx \hbar$$
$$\Delta t\, \Delta E \approx \hbar \tag{15}$$
$$\Delta \theta\, \Delta L \approx \hbar$$

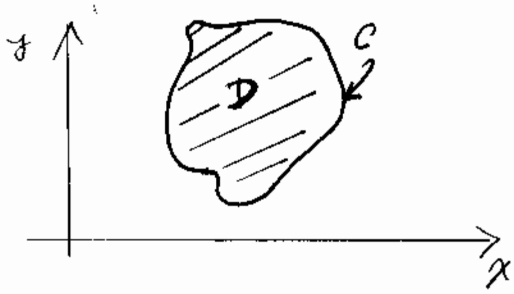
EXTREMIZATION OF MULTIPLE INTEGRALS & LAGRANGE DENSITIES

Suppose we have a multidimensional problem with more than one independent variable (e.g. x and y), and we wish to extremize the quantity

THE ANSWER

$$I = \iint_{D} dx\,dy\; f(\underbrace{x, y}_{\text{independent variables}},\, w(x,y),\, \underbrace{w_x}_{\frac{\partial w}{\partial x}},\, \underbrace{w_y}_{\frac{\partial w}{\partial y}}) \qquad (1)$$

D: Some bounded region analogous to the end points x_1, x_2 in the 1-dim case.



Note that now y denotes one of the independent variables, rather than the answer as before.

A physical example might be finding the equation of the surface that a chain mesh would form under gravity in analogy to the hanging chain.

By analogy to the one dimensional case where the extremization was carried out subject to the end points being fixed [so that $f(\ldots)$ was fixed] here we carry out the extremization subject to the constraint that $f(x, y, \ldots)$ have some fixed value along the boundary C of the domain D.

To find the differential equation for $w(x,y)$ we proceed in analogy to the 1-dimensional case: Form the family of 2-dim comparison functions

$$W(x,y) = w(x,y) + \epsilon\,\eta(x,y) \qquad (2)$$

$\eta(x,y)$ is then constrained to vanish along the surface. As before we form the integral

$$I(\epsilon) = \iint_{D} dx\,dy\; f(x, y, W, W_x, W_y) \qquad (3)$$

By construction, the minimum is given (as before) by

$$I'(\epsilon)\Big|_{\epsilon=0}=0 \equiv I'(0)=\iint_D dx\,dy\left\{\frac{\partial f}{\partial w}\eta+\frac{\partial f}{\partial w_x}\eta_x+\frac{\partial f}{\partial w_y}\eta_y\right\}=0 \qquad (4)$$

We now wish to carry out a <u>2-dimensional partial integration</u>, in order to replace $\eta_x=\partial\eta/\partial x$ and $\eta_y=\partial\eta/\partial y$ by $\eta(x)$. The 2-dim analog of partial integration follows from <u>GREEN'S THEOREM</u> in a plane [See, e.g., the book by Sokolnikoff & Redheffer, p. 392]:

$$\boxed{\iint_D dx\,dy\left(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}\right)=\underbrace{\iint_C}_{\substack{\text{bounding}\\ \text{contour}}}(P\,dy-Q\,dx) \qquad (5)}$$

This is valid if P, Q and their first derivatives are continuous in D

Let $P=\eta G \qquad Q=\eta F$; then (5) $\Rightarrow$

$$\iint_D dx\,dy\left(\eta\frac{\partial G}{\partial x}+G\underbrace{\frac{\partial \eta}{\partial x}}_{\eta_x}+\eta\frac{\partial F}{\partial y}+F\underbrace{\frac{\partial \eta}{\partial y}}_{\eta_y}\right)=\iint_C \eta\,(G\,dy-F\,dx) \qquad (6)$$

Eq. (6) $\Rightarrow$

$$\boxed{\iint_D dx\,dy\left\{G\eta_x+F\eta_y\right\}=\iint_C \eta\,(G\,dy-F\,dx)-\iint_D dx\,dy\,\eta\left(\frac{\partial G}{\partial x}+\frac{\partial F}{\partial y}\right)} \quad (7)$$

Comparing (4) & (7) set $\quad G=\partial f/\partial w_x \quad ; \quad F=\partial f/\partial w_y \quad \Rightarrow$

$$\iint_D dx\,dy\left\{\frac{\partial f}{\partial w_x}\eta_x+\frac{\partial f}{\partial w_y}\eta_y\right\}=\iint_C \eta(x)\left\{\frac{\partial f}{\partial w_x}dy-\frac{\partial f}{\partial w_y}dx\right\} \qquad (8)$$

this vanishes on C

$$-\iint_D dx\,dy\,\eta\left[\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial w_x}\right)+\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial w_y}\right)\right]$$

Then finally: (9)

$$\boxed{\iint_D dx\,dy\left\{\frac{\partial f}{\partial w_x}\eta_x+\frac{\partial f}{\partial w_y}\eta_y\right\}=-\iint_D dx\,dy\,\eta(x)\left[\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial w_x}\right)+\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial w_y}\right)\right]}$$

Combining Eqs (4) & (9) we have

$$0 = I'(\epsilon)\Big|_{\epsilon=0} = \iint_D dx\,dy \left\{ \frac{\partial f}{\partial W} - \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial W_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial W_y}\right) \right\} \eta(x,y) \qquad (10)$$

Since $\eta(x,y)$ is an arbitrary function of x, y it follows (as before!) that $\{\cdots\} = 0 \Rightarrow$

$$\boxed{\frac{\partial f}{\partial W} - \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial W_x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial W_y}\right) = 0} \qquad (11)$$

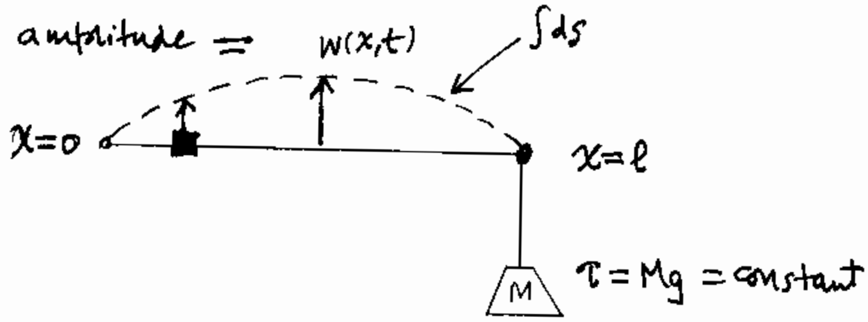
Note the presence of the partial derivatives $\partial/\partial x$ and $\partial/\partial y$ rather than the total derivatives as in the case of a single independent variable.

This can be generalized to the case of an arbitrary number N of independent variables:

$$\boxed{\frac{\partial f}{\partial W} - \sum_{i=1}^{N} \frac{\partial}{\partial x_i} \frac{\partial f}{\partial(\partial W/\partial x_i)} = 0} \qquad (12)$$

ARBITRARY NUMBER OF INDEPENDENT VARIABLES

Example: The vibrating string



From the figure the displacement of any point in the vertical direction (from the equilibrium position) is $W(x,t)$, and hence the vertical velocity at point is $W_t(x,t) \equiv (\partial/\partial t)\, W(x,t)$. If we consider a small segment as shown then its kinetic energy is given by

$$dT = \tfrac{1}{2}\, dm\, W_t^2(x,t) \quad ; \quad dm = \rho\, ds \cong \rho\, dx \ \ (\text{valid for KE}) \qquad (13)$$

$$\boxed{\therefore \quad T = \frac{1}{2}\int_0^{\ell} dx\, \rho(x) W_t^2(x,t)} \qquad (14)$$

need not be constant

Consider next the expression for the potential energy V of the string: If the length of the string is increased by an amount Δl while the string is subject to the tension τ, then the change in the potential energy (ΔV) is given by

$$V - V_0 = \Delta V = \tau \Delta l \qquad [\text{recall: } dW = F dx] \tag{15}$$

To calculate Δl we note that Δl is the difference between $\int ds$ — which gives the length of the stretched string — and the equilibrium value l. If we define $V_0 = 0$ as usual then

$$V = \tau\left\{\int ds - l\right\} = \tau\left\{\int_{x=0}^{x=l} dx\sqrt{1+\underbrace{(y')^2}_{\equiv w_x}} - l\right\} \tag{16}$$

Hence $$\boxed{V = \tau\left\{\int_{x=0}^{x=l} dx\sqrt{1+w_x^2} - l\right\}} \tag{17}$$

Note that in the case of V the entire effect arises from the difference between $\int ds$ and $\int dx$. However, for the kinetic energy T we can write $ds \cong dx$. Also in (17) we can write $\sqrt{1+w_x^2} \cong 1 + \frac{1}{2}w_x^2$. Hence

$$\boxed{V \cong \tau\left\{\int_0^l dx\left(\cancel{1}+\tfrac{1}{2}w_x^2\right) - \cancel{l}\right\} = \frac{1}{2}\tau\int_0^l dx\, w_x^2} \tag{18}$$

Hence (14) & (18) give the following expression for the Lagrangian $L = T - V$:

$$\boxed{L = T - V = \int_0^l dx\left\{\frac{1}{2}\rho(x)\, w_t^2(x,t) - \frac{1}{2}\tau\, w_x^2(x,t)\right\}} \tag{19}$$

From Hamilton's Principle the quantity that we wish to extremize is

$$I = \int_{t_1}^{t_2} dt\, L = \int_{t_1}^{t_2} dt \int_0^l dx \underbrace{\left\{\frac{1}{2}\rho(x) w_t^2(x,t) - \frac{1}{2}\tau w_x^2(x,t)\right\}}_{\mathcal{L}(x,t)} \tag{20}$$

<u>Lagrangean Density</u>

The Lagrangian density $\mathcal{L}(x,t)$ is the analog of the function $f(x,y,W,W_x,W_y)$ that we started with. In field theory it is the usual starting point:

$$f(x,y,W,W_x,W_y) \longrightarrow \mathcal{L}(x,t,W,W_x,W_t) \qquad (21)$$

Hence using the E-L equation in (11) or (12) we have:

$$\frac{\partial \mathcal{L}}{\partial W} - \frac{\partial}{\partial x}\left(\frac{\partial \mathcal{L}}{\partial W_x}\right) - \frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial W_t}\right) = 0 \qquad (22)$$

In this case $\mathcal{L}$ is not a function of W, but only of W_x and W_t. Thus

$$\frac{\partial \mathcal{L}}{\partial W_x} = -\tau W_x \Rightarrow \frac{\partial}{\partial x}\left(\frac{\partial \mathcal{L}}{\partial W_x}\right) = -\tau W_{xx} = -\tau \frac{\partial^2 W}{\partial x^2} \qquad (23)$$

$$\frac{\partial \mathcal{L}}{\partial W_t} = +\rho W_t \Rightarrow \frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial W_t}\right) = \text{[redacted]} = +\rho \frac{\partial^2 W}{\partial t^2} \qquad (24)$$

Then E-L $\Rightarrow$ $0 = -\tau \dfrac{\partial^2 W}{\partial x^2} + \rho \dfrac{\partial^2 W}{\partial t^2} \Rightarrow$ $\boxed{\dfrac{\partial^2 W}{\partial x^2} - \left(\dfrac{\rho}{\tau}\right)\dfrac{\partial^2 W}{\partial t^2} = 0}$ (25)

WAVE EQUATION

Note that this is the classical wave equation for propagation on a string. On dimensional grounds we see that

$$\boxed{v^2 = \tau/\rho} \qquad (26)$$

Smaller $\rho \Rightarrow$ faster v for fixed τ etc.

In a 3-dimensional relativistic theory $dx\,dt \to d^3x\,dt \equiv d^4x$ which is a <u>relativistic invariant</u>. Since I is also a relativistic invariant $\Rightarrow \mathcal{L}(\vec{x},t)$ is also a relativistic invariant. Knowing this significantly constraints possible forms for $\mathcal{L}(\vec{x},t)$.

SUMMARY RECIPE

In the general case we wish to extremize the quantity I

$$I = \underbrace{\int}_{n\text{-fold}} d^n x_i \; f\left(x_i, w_j, \frac{\partial w_j}{\partial x_i}\right) \qquad \begin{matrix} i = (1, \ldots, n) \\ j = (1, \ldots, m) \end{matrix} \tag{1}$$

independent variables (x_i); answers (w_j)

$$d^n x_i \equiv dx_1\, dx_2 \ldots dx_n \tag{2}$$

this can arise in 3-dimensions when many particles exist.

The constraints take the form

$$J_k = \underbrace{\int}_{n\text{-fold}} d^n x_i \; g_k\left(x_i, w_j, \frac{\partial w_j}{\partial x_i}\right) \qquad k = (1, \ldots, p) \tag{3}$$

The solution can be expressed as before: First form the function h,

$$h\left(x_i, w_j, \frac{\partial w_j}{\partial x_i}\right) = f\left(x_i, w_j, \frac{\partial w_j}{\partial x_i}\right) + \sum_{k=1}^{p} \lambda_k g_k\left(x_i, w_j, \frac{\partial w_j}{\partial x_i}\right) \tag{4}$$

The solution to the extremization problem is then obtained from

$$\boxed{\frac{\partial h}{\partial w_j} - \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left(\frac{\partial h}{\partial(\partial w_j/\partial x_i)}\right) = 0} \tag{5}$$

STURM-LIOUVILLE THEORY

Many of the special functions that we studied last semester are solutions of differential equations which can be viewed as special cases of the STURM-LIOUVILLE differential equation. These special functions include Bessel, Legendre, Laguerre, Hermite, and hypergeometric functions. Since the S-L differential equation is the solution of a particular extremization problem, this subject forms a natural bridge between CALCULUS OF VARIATIONS and DIFF. EQUATIONS.

The Origin of the STURM-LIOUVILLE (S-L) Equation:

We are given to extremize the integral

$$I = \int_{x_1}^{x_2} dx \left[\underbrace{p(x)(y')^2}_{\sim KE} + \underbrace{q(x)\,y^2(x)}_{\sim PE} \right] \quad ; \quad [\cdots] \sim f(x,y,y') \qquad (1)$$

Subject to the constraint

$$J = \int_{x_1}^{x_2} dx \, \underbrace{w(x)\, y^2(x)}_{\sim g(x,y,y')} = \text{constant} \qquad (2)$$

Here $p(x)$, $q(x)$, and $w(x) \geq 0$ are given functions of x in the interval $x_1 \leq x \leq x_2$. The problem can be solved in the usual way by forming $h = f - \lambda g$, where we choose the Lagrange multiplier to be $-\lambda$ by convention. Then

$$h = f - \lambda g = p(x)(y')^2 + q(x)\,y^2 - \lambda w(x)\,y^2 \qquad (3)$$
$$= p(y')^2 + (q - \lambda w)y^2$$

The E-L equation is, as before,

$$\frac{\partial h}{\partial y} - \frac{d}{dx}\left(\frac{\partial h}{\partial y'}\right) = 0 \qquad y' = dy/dx \qquad (4)$$

$$\frac{\partial h}{\partial y} = 2(q-\lambda w)y \;;\; \frac{\partial h}{\partial y'} = 2py' \Rightarrow \frac{d}{dx}\left(\frac{\partial h}{\partial y'}\right) = 2\frac{d}{dx}(py') = 2(p'y' + py'') \qquad (5)$$

We can then write from (4): E-L $\Rightarrow$ $\boxed{\frac{d}{dx}(py') - (q-\lambda w)y = 0}$ (6)

STURM–LIOUVILLE DIFF. EQN.

We have set up this variational problem in analogy with previous cases where the end points are fixed: $y(x_1) = y_1$ $y(x_2) = y_2$. One can also find solutions of the S-L problem with other boundary conditions

$$[pyy']_{x_1} = [pyy']_{x_2} = 0 \qquad (7)$$

which corresponds to the same variational problem with free end points. [We use this below!]

As we will see, many of the equations of interest to us can be cast into the S-L form (6) with appropriate choices of p, q, w. The advantage of the S-L formalism is that one can establish certain properties of the solutions (such as orthogonality) by studying the "generic" S-L equation, without having to consider each case individually.

Connection to Quantum Mechanics:

From the preceding discussion form $(I-\lambda J) = \int_{x_1}^{x_2} dx \left\{ p(y')^2 + (q-\lambda w)y^2 \right\}$ (1)

Integrate the term proportional to p by parts as follows:

$$\int_{x_1}^{x_2} dx\, p(y')^2 = \underbrace{p y y' \Big]_{x_1}^{x_2}}_{=0 \text{ (boundary condition)}} - \int_{x_1}^{x_2} dx \left\{ y \frac{d}{dx}(py') \right\} \quad (2)$$

Then
$$(I-\lambda J) = \int_{x_1}^{x_2} dx \left\{ -y \frac{d}{dx}(py') + (q-\lambda w)y^2 \right\} = \int_{x_1}^{x_2} dx \cdot y(x) \left[\underbrace{-\frac{d}{dx}(py') + (q-\lambda w)y} \right] \quad (3)$$

When I is an extremum subject to the constraint $J =$ constant, then y is a solution to the S-L equation in 44(6) so that $\sim\sim\sim = 0$ in (3) above.

It follows that $I - \lambda J = 0 \Rightarrow \boxed{\lambda = I/J}$ (4)

This looks vaguely like an <u>eigenvalue equation for λ</u>. To clarify this, when $\lambda = 0$ we can write I in the form

$$I = \int_{x_1}^{x_2} dx\,(py'^2 + qy^2) \overset{(2)}{=} -\int_{x_1}^{x_2} dx\, y\left\{ \frac{d}{dx}\left(p\frac{dy}{dx}\right) - qy \right\} \quad (5)$$

$$= -\int_{x_1}^{x_2} dx\, y \underbrace{\left\{ \frac{d}{dx}\left(p\frac{d}{dx}\right) - q \right\}}_{H} y$$

In terms of H the S-L equation is:

$$Hy = -\lambda w y \quad (6)$$

and

$$I = -\int dx\,(yHy) \quad (7)$$

$$J = \int dx\,(y w y) \quad (8)$$

From (4): $$\lambda = \lambda(y) = \frac{I}{J} = \frac{-\int_{x_1}^{x_2} dx\; yHy}{\int_{x_1}^{x_2} dx\; ywy} = \frac{\int_{x_1}^{x_2} dx\,(py'^2+qy^2)}{\int_{x_1}^{x_2} dx\; ywy} \qquad (9)$$

In the special case $w=1$ Eq(9) gives:

$$Hy = -\lambda y \qquad \lambda = \frac{\int_{x_1}^{x_2} dx\; y(x)(-H)y(x)}{\int_{x_1}^{x_2} dx\; y^2(x)} \equiv \frac{-\langle y|H|y\rangle}{\langle y|y\rangle} \qquad (10)$$

<u>SUMMARY</u>: (a) If I is an extremum subject to $J=$ constant, then y is a solution of

$$Hy \equiv \left[\frac{d}{dx}\left(p\frac{d}{dx}\right) - q\right] y = -\lambda w y \qquad (11)$$

and λ is given by $\lambda = I/J$, where I (and J) are evaluated at the value $y(x)$ which solves the S-L differential equation

(b) Conversely if we evaluate $\lambda(y) = I(y)/J(y)$ and vary whatever parameters appear in $\lambda(y)$ to find a stationary (extremum) value then λ must be a solution of the differential equation

$$Hy = -\lambda w y \qquad (12)$$

which for $w=1 \Rightarrow$ <u>λ is an eigenvalue of H</u> (<u>RAYLEIGH'S PRINCIPLE</u>)

<u>Application</u>: <u>Rayleigh-Ritz Variational Method</u>

One can solve for $y(x)$ using the above formalism as follows: Choose a <u>trial function</u> $y(x, \alpha_1, \alpha_2, \dots, \alpha_n)$ which depends on some auxiliary set of parameters $(\alpha_1, \dots, \alpha_n)$ but which is chosen to satisfy the relevant boundary conditions: $y(x_1, \alpha_1 \dots \alpha_n) = y_1$; $y(x_2, \alpha_1, \dots \alpha_n) = y_2$ OR//

$$pyy'|_{x_1} = pyy'|_{x_2} = 0 \qquad (13)$$

One then substitutes this trial function into the expression

$$\lambda(y, \alpha_1, \ldots, \alpha_n) = \frac{I(y)}{J(y)} = \frac{\int_{x_1}^{x_2} dx\,(py'^2 + q y^2)}{\int_{x_1}^{x_2} dx\; w y^2} \qquad (14)$$

λ now depends explicitly on $\alpha_1, \ldots \alpha_n$, so that if λ is an extremum* with respect to these parameters this means that

$$\boxed{0 = \frac{\partial \lambda(\alpha_1, \ldots, \alpha_n)}{\partial \alpha_1} = \frac{\partial \lambda(\alpha_1, \ldots, \alpha_n)}{\partial \alpha_2} = \cdots = \frac{\partial \lambda(\alpha_1, \ldots, \alpha_n)}{\partial \alpha_n}} \qquad (15)$$

This gives a set of coupled equations in the parameters α_n which can then in principle be solved for the α_n. These parameters then determine the eigenvalue λ and the eigenfunction $y(x, \alpha_1, \ldots \alpha_n)$. **

* Since $\lambda = I/J$, where $J =$ constant, extremizing $\lambda(\alpha_1, \ldots \alpha_n)$ by varying the α_n, is equivalent to extremizing I, which is what we set out to do.

** We note that the "true" value of y [the "exact" solution] would (by definition) make $\int dx\,(yHy)$ a minimum, and hence would make the eigenvalue λ a minimum. It follows that any trial function (which represents a guess) must give a larger eigenvalue λ than the correct value. Better guesses at the trial function $\Rightarrow$ we approach the correct value more closely from above.

General Properties of STURM-LIOUVILLE Systems:

We return now to characterize the properties of differential equations which can be cast in the S-L form. This is actually a broad class including any equation of the form

$$\boxed{f y'' + g y' + (h_1 + \lambda h_2) y = 0} \quad (1)$$

$f = f(x)$ $g = g(x)$
$h_{1,2} = h_{1,2}(x)$

Any such equation can be written in the S-L form:

$$\frac{d}{dx}(py') + (\lambda w - q) y = 0 \qquad (2)$$

To show this define $p(x) = e^{\int^x (g/f)dx}$; $q = -p h_1/f$; $w = p h_2/f$ (3)

As before we define

$$H = \frac{d}{dx}\left(p\frac{d}{dx}\right) - q \quad ; \quad Hy = -\lambda w y \qquad (4)$$

$$\text{Then } Hy = p'y' + py'' - qy = \underbrace{e^{\int^x (g/f)dx}(g/f)}_{p'} y' + e^{\int^x (g/f)dx} y'' + e^{\cdots}\frac{h_1}{f} y \qquad (5)$$

$$\therefore Hy = \underset{\sim\sim}{e^{\int^x (g/f)dx}} \left\{ \frac{g}{f} y' + y'' + \frac{h_1}{f} y \right\} \qquad (6)$$

$$\text{Since } Hy = -\lambda w y = -\lambda \underset{\sim\sim}{e^{\int^x \cdots}} \frac{h_2}{f} y \Rightarrow \quad \sim\sim = \text{common factor which cancels} \qquad (7)$$

$$Hy = -\lambda w y \Rightarrow \left(\frac{g}{f} y' + y'' + \frac{h_1}{f} y\right) = -\lambda \frac{h_2}{f} y \qquad < f \qquad (8)$$

$$\therefore \quad f y'' + g y' + (h_1 + \lambda h_2) = 0 \qquad (9)$$

Hence any differential equation of this form can be put in S-L form.

ORTHOGONALITY OF SOLUTIONS OF THE S-L EQUATION

Let u & v be 2 solutions of the S-L equation

$$Hy = \left[\frac{d}{dx}\left(p\frac{d}{dx}\right) - q\right]y = -\lambda w y \tag{1}$$

Satisfying the boundary condition: $p v u'\big|_{x_1} = p v u'\big|_{x_2}$. (2)

This is a somewhat more general form of

$$p y y'\big|_{x_1} = p y y'\big|_{x_2} = 0 \tag{3}$$

Let λ_m and λ_n be the eigenvalues corresponding to the eigenfunctions $u \equiv y_m$ and $v = y_n$. We assume for the moment that $\lambda_m \neq \lambda_n$. Then

$$\frac{d}{dx}(p y_m') + (\lambda_m w - q) y_m = 0 \quad ; \quad \frac{d}{dx}(p y_n') + (\lambda_n w - q) y_n = 0 \tag{4}$$

Multiplying these equations by y_n and y_m respectively gives:

$$y_n \frac{d}{dx}(p y_m') + \lambda_m w y_m y_n - q y_m y_n = 0 \tag{5a}$$

$$y_m \frac{d}{dx}(p y_n') + \lambda_n w y_n y_m - q y_n y_m = 0 \tag{5b}$$

Subtracting (5b) from (5a) gives:

$$y_n \frac{d}{dx}(p y_m') - y_m \frac{d}{dx}(p y_n') + (\lambda_m - \lambda_n) w y_m y_n = 0 \tag{6}$$

Next we integrate all terms in (6) between x_1 and x_2:

$$\int_{x_1}^{x_2} dx \left\{ y_n \frac{d}{dx}(p y_m') - y_m \frac{d}{dx}(p y_n') \right\} + (\lambda_m - \lambda_n)\int_{x_1}^{x_2} dx\, w(x) y_m(x) y_n(x) = 0 \tag{7}$$

We can partially integrate the terms in $\{\ldots\}$ by noting that

$$\int_{x_1}^{x_2} dx \, \frac{d}{dx}(y_n p y_m') = y_n p y_m' \Big|_{x_1}^{x_2} = 0 \tag{8}$$

$$= \int_{x_1}^{x_2} dx \left[y_n \frac{d}{dx}(p y_m') + p y_m' \frac{d}{dx} y_n \right] = \int_{x_1}^{x_2} dx \left[y_n \frac{d}{dx}(p y_m') + p y_m' y_n' \right] \tag{9}$$

Combining the results in (7) & (9) we find:

$$\int_{x_1}^{x_2} dx \left[y_n \frac{d}{dx}(p y_m') \right] = -\int_{x_1}^{x_2} dx \,(p y_m' y_n') \qquad (10)$$

Inserting this equation into Eq. (7), along with the analogous result for $n \leftrightarrow m$ we find:

$$\boxed{\int_{x_1}^{x_2} dx \left\{ \cancel{-p y_m' y_n'} \;\; \cancel{+ p y_n' y_m'} \right\} + (\lambda_m - \lambda_n) \int_{x_1}^{x_2} dx \; w(x) y_m(x) y_n(x) = 0 \qquad (11)}$$

Hence if $\lambda_m \neq \lambda_n$ then (11) expresses the orthogonality of the solutions $y_m(x)$ and $y_n(x)$ with respect to the non-negative weight function $w(x)$, on the interval $[x_1, x_2]$.

Having shown that the eigenfunctions of the S-L equation are orthogonal, we can normalize them appropriately by demanding that

$$J = \int_{x_1}^{x_2} dx \; w(x) \, y_m^2(x) = 1$$

Thus in the end it is <u>the normalization condition which makes the S-L system an isoperimetric problem.</u>

COMPLETENESS OF THE SOLUTIONS OF THE S-L EQUATION

We start by invoking a theorem (not proved here!) that if H is a Hermitian operator ($H = H^\dagger$) and

$$H\phi_n = \lambda_n \phi_n \qquad (1)$$

then the ϕ_n form a complete (orthogonal) set if the λ_n satisfy

$$\begin{aligned} &\text{a) there exists a minimum (finite) } \lambda_n \\ &\text{b) } \lim_{n\to\infty} \lambda_n = \infty \end{aligned} \qquad (2)$$

Hence completeness follows if we can show that

$$H = \frac{d}{dx}\left(p(x)\frac{d}{dx}\right) - q = H^\dagger \qquad (3)$$

Recall the properties of Hermitian operators:

$$\langle f|H|g\rangle \equiv \langle f|Hg\rangle \underset{\substack{\uparrow \\ \text{def. of } H^\dagger}}{=} \langle H^\dagger f|g\rangle \underset{\substack{\uparrow \\ \text{def. of scalar product}}}{=} \langle g|H^\dagger f\rangle^* \overset{\substack{H=H^\dagger \\ \downarrow}}{=} \langle g|H|f\rangle^* \qquad (4)$$

We can use the above to write the Hermiticity condition as

$$\langle f|H|g\rangle \equiv \langle f|Hg\rangle = \langle H^\dagger f|g\rangle \underset{\substack{\uparrow \\ H=H^\dagger}}{=} \langle Hf|g\rangle \Rightarrow \boxed{\langle f|Hg\rangle = \langle Hf|g\rangle} \qquad (5)$$

In the present case this translates into

$$\boxed{\int_{x_1}^{x_2} dx\, f^* H g \overset{?}{=} \int_{x_1}^{x_2} dx\, (Hf)^* g} \qquad (6)$$

$$Hg = \left[\frac{d}{dx}\left(p\frac{d}{dx}\right) - q\right]g = \frac{d}{dx}(pg') - qg = (pg')' - qg \qquad (7)$$

We want now to integrate by parts to remove the derivatives from g so that they act on f as required by (6).

Integrating by parts (for the term $\sim p$)

$$\int_{x_1}^{x_2} dx\, f^*(pg')' = \int_{x_1}^{x_2} dx\, f^* \frac{d}{dx}(pg') = \underbrace{f^* p g'\Big|_{x_1}^{x_2}}_{=0} - \int_{x_1}^{x_2} dx\, f^{*\prime} p g' \qquad (8)$$

we assume this boundary condition

Integrating by parts a second time $\Rightarrow$

EQ.(8)

$$\int_{x_1}^{x_2} dx\, f^*(pg')' = -\int_{x_1}^{x_2} dx\, f^{*\prime} p g' = -\int_{x_1}^{x_2} dx\, \frac{dg}{dx}(pf^{*\prime}) = \underbrace{-g p f^{*\prime}\Big|_{x_2}^{x_1}}_{=0} + \int_{x_1}^{x_2} dx\, g \frac{d}{dx}(pf^{*\prime}) \qquad (9)$$

Hence altogether:

$$\boxed{\int_{x_1}^{x_2} dx\, f^*(pg')' = \int_{x_1}^{x_2} dx \left[\frac{d}{dx}(pf')^*\right] g} \qquad (10)$$

$\leftarrow$ assumes $p = p^*$

Consider next the contribution in $H \sim q$

$$\int_{x_1}^{x_2} dx\, f^* \cdot q \cdot g = \int_{x_1}^{x_2} dx\, (qf)^* g \leftarrow \text{trivially} \qquad (11)$$

Hence combining (10) & (11):

$$\boxed{\int_{x_1}^{x_2} dx\, f^* \left\{\frac{d}{dx}\left(p\frac{d}{dx}\right) - q\right\} g = \int_{x_1}^{x_2} dx \left\{\left[\frac{d}{dx}\left(p\frac{d}{dx}\right) - q\right] f\right\}^* g}$$

$$\langle f | Hg \rangle = \langle Hf | g \rangle \checkmark \qquad (12)$$

This completes the proof of <u>COMPLETENESS</u>, provided the λ_n have the correct properties (see below).

We note in passing that from (12) we can characterize differential equations in the S-L form as <u>SELF-ADJOINT EQUATIONS</u>.

PROPERTIES OF THE EIGENVALUES:

We summarize below the properties of the eigenvalues of some special functions obtained from S-L equations:

LEGENDRE: Eigenvalues are $\ell(\ell+1)$; $\ell = 0, 1, 2, \ldots$ (ℓ = orbital angular momentum)

HERMITE: " " $2n$; $n = 0, 1, 2, \ldots$ (number of quanta $\hbar\omega_0$)

LAGUERRE: " " n ; $n = 0, 1, 2 \ldots$ (~Bohr level)

BESSEL " " n^2 ; $n = 0, 1, 2 \ldots$

Some properties of these equations in S-L form are given in the accompanying Table.

Seccion ARFKEN - p. 446

Table of Orthogonal Functions Arising from Sturm - Liouville Systems

Name and Physical Application	Rodrique's Formula	Generating Function	Differential Equation	S -L Form of D. E.	Orthonormality
Legendre Polynomials 1) Multiple Expansion 2) ∇^2 in sph. coor.	$P_n(x)=\frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n$	$\sum_{n=0}^{\infty} P_n(x)t^n=\frac{1}{\sqrt{1-2xt+t^2}}$ $(0<t<1)$	$(x^2-1)P_n''+2xP_n'-n(n+1)P_n=0$	$\frac{d}{dx}\left((1-x^2)P_n'\right)+n(n+1)P_n=0$	$\int_{-1}^{1}P_nP_m\,dx=\delta_{nm}\frac{2}{2n+1}$
Hermite Polynomials Quantum Oscillator	$H_n(x)=(-1)^n e^{x^2}\frac{d^n e^{-x^2}}{dx^n}$	$\sum_{n=0}^{\infty}\frac{H_n(x)}{n!}t^n=e^{-t^2+2tx}$ $(t>0)$	$H_n''-2xH_n'+2nH_n=0$	$\frac{d}{dx}(e^{-x^2}H_n')+2ne^{-x^2}H_n=0$	$\int_{-\infty}^{\infty}H_nH_me^{-x^2}dx=\delta_{nm}\sqrt{\pi}\,2^n n!$
Laguerre Polynomials H - atom	$L_n(x)=e^x\frac{d^n}{dx^n}(x^ne^{-x})$	$\sum_{n=0}^{\infty}\frac{L_n(x)}{n!}t^n=\frac{e^{-\frac{xt}{1-t}}}{1-t}$ $(0<t<1)$	$xL_n''+(1-x)L_n'+nL_n=0$	$\frac{d}{dx}(xe^{-x}L_n')+ne^{-x}L_n=0$	$\int_0^{\infty}L_nL_me^{-x}dx=\delta_{nm}(n!)^2$
	Series Presentation				
Bessel's Function (of integral order) ∇^2 in cylindrical coordinates	$J_n(x)=\sum_{m=0}^{\infty}\frac{(-1)^m\left(\frac{x}{2}\right)^{2m+n}}{(n+m)!\,m!}$	$\sum_{n=-\infty}^{\infty}J_n(x)t^n=e^{\frac{1}{2}x\left(t-\frac{1}{t}\right)}$ $(t>0)$	$x^2J_n''+xJ_n'+(x^2-n^2)J_n=0$	$\frac{d}{dx}(xJ_n')+\left(x-\frac{n^2}{x}\right)J_n=0$	
Trigonometric Functions Classical Oscillator	$f_n=\sin nx=\sum_{m=0}^{\infty}(-1)^m\frac{(nx)^{2m+1}}{(2m+1)!}$ $g_n=\cos nx=\sum_{m=0}^{\infty}(-1)^m\frac{(nx)^{2m}}{(2m)!}$		$f_n''+n^2f_n=0$ $g_n''+n^2g_n=0$	$\frac{d}{dx}(f_n')+n^2f_n=0$ $\frac{d}{dx}(g_n')+n^2g_n=0$	$\int_{-\pi}^{\pi}f_nf_m\,dx=\delta_{nm}\pi$ $\int_{-\pi}^{\pi}g_ng_m\,dx=\delta_{nm}\pi$ $\int_{-\pi}^{\pi}f_ng_m\,dx=0$

161