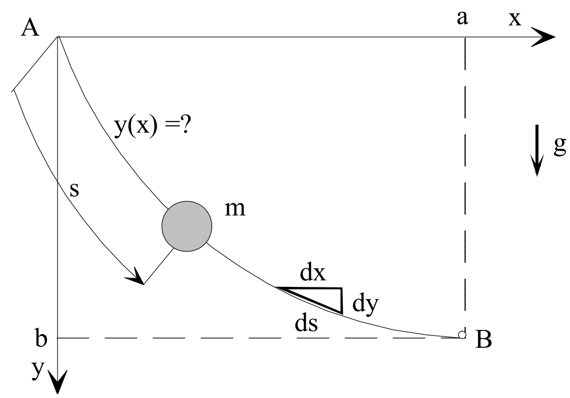
INTRODUCTION TO CALCULUS OF VARIATIONS

Brachistochrone problem

The problem was posed by John Bernoulli (1667-1748) in 1696. brachist = shortest, chronos = time (from the Greek)



<u>Problem formulation</u>

A bead slides under the gravity along a smooth wire joining two fixed points A and B (not in the same vertical line). The bead is released from rest at A and it should slide to B in minimum time. The question is: what shape of wire should be? [A(0,0),B(a,b)] – coordinates of points A and B

Velocity $\quad v = \dfrac{ds}{dt} \Rightarrow \quad dt = \dfrac{ds}{v} \quad \Rightarrow \quad$ time $T = \int_A^B dt = \int_A^B \dfrac{ds}{v} \Rightarrow$ MIN.

The total energy is conserved, so

$$\frac{mv^2}{2} = mgy \qquad \Rightarrow \qquad v = (2gy)^{1/2}$$

For $\quad ds^2 = dx^2 + dy^2 \quad \Rightarrow \quad$ hence $\quad ds = \left(dx^2 + dy^2\right)^{1/2} = dx\left(1 + y'^2\right)^{1/2}$

Substituting v and ds we have

$$\boxed{T = \frac{1}{(2g)^{1/2}} \int_0^a \left(\frac{1 + y'^2}{y}\right)^{1/2} dx} \Rightarrow \qquad \text{is minimized.}$$

The simplest problem of calculus of variations

Find the smooth curve y = y(x), which joins the fixed points A(0,0) and B(a, b) that gives the functional

$$I(y) = \int_0^a f(x, y, y')dx$$

its minimum value. The boundary conditions are y(0) = 0, y(a) = b.

Another formulation

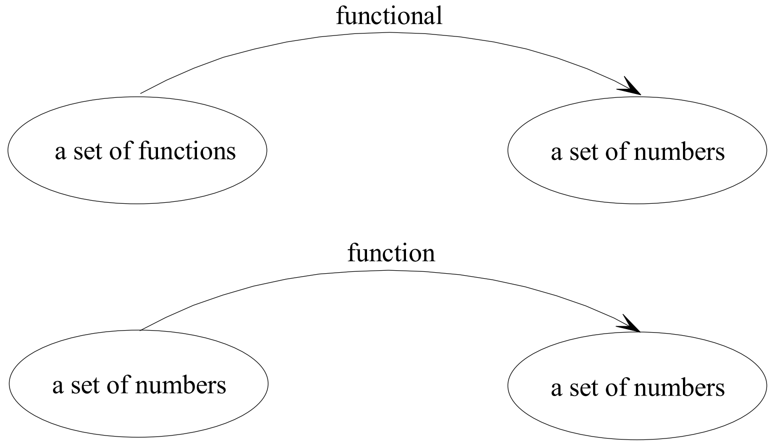
Find the smooth curve x = x(t) that gives the functional

$$I(x) = \int_{t_0}^{t_1} f(t, x, \dot{x})dt$$

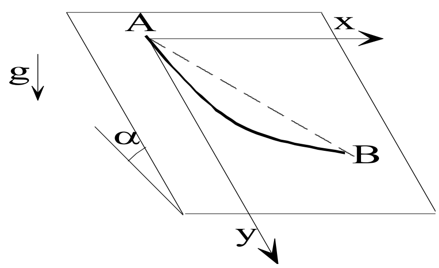
its minimum value and satisfies the boundary conditions

$$x(t_0) = x_0, \qquad x(t_1) = x_1.$$

$I(x)$ – it is functional. For $x(t)$ given the integral $I(x)$ leads to a specific numerical value. Integral $I(x)$ acts on a set of functions to produce a corresponding set of numbers.



Sample problem (G. Twardokens, 1990)
The skier is on the slope. He/she should cover the distance from A to B in minimum time. Find the shape of his/her path.



The fixed end – point problem
Minimize the functional

$$I(x) = \int_{t_0}^{t_1} f(t, x, \dot{x})dt \quad \Rightarrow \quad \text{MIN}$$

with boundary conditions

$$x(t_0) = x_0,\ x(t_1) = x_1.$$

Definition: The curve $x = x^*(t)$ minimizes functional, when

$$(*) \qquad I(x) \ge I(x^*)$$

for all curves $x = x(t)$ satisfying boundary conditions. Equality condition is satisfied only when $x(t)$ and $x^*(t)$ coincide, $x(t) = x^*(t)$.

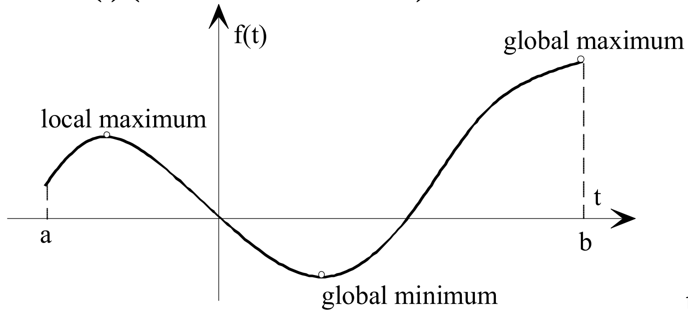
This definition is not particularly useful. If gives no algorithm how $x^*(t)$ might be found. It requires that each candidate $x(t)$ be tested using (*). The number of admissible curves $x(t)$ to be considered goes to infinity.

Local minimum
The functional attains a local minimum if the condition (*) is valid only for $x(t)$ from the ε – neighbourhood of $x^*(t)$

$$\left| x - x^* \right| < \varepsilon, \text{ where } \varepsilon \text{ is small.}$$

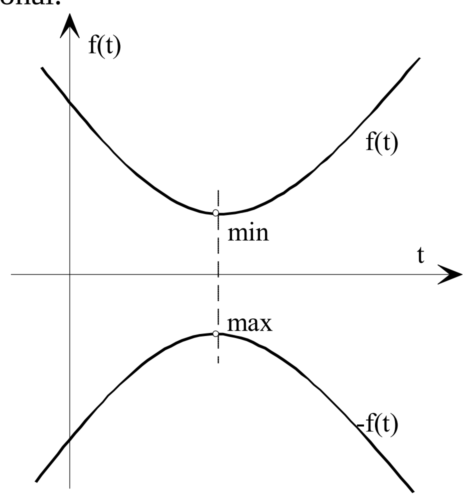
Global or absolute minimum

The functional attains a global minimum if the condition (*) is valid for any x(t) taken from an admissible domain of x(t) (ε is not small number).



$t \in \langle a, b \rangle$

Remark: Maximization of the functional may be converted into minimization of the functional changing the sign of the functional.



At that stage of considerations we consider the curves x(t) from a set of twice continuously differentiable functions – C^2 class. Now, we must clarify what we mean by saying that two curves are close to each other.

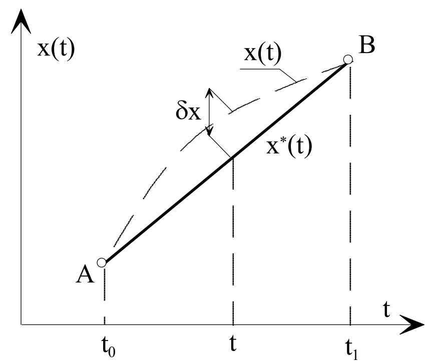
Definition: A weak variation

Let $x^*(t)$ be a minimizing curve and x(t) an admissible curve, i.e. the curve of C^2 class satisfying boundary conditions. Then, if there exist small numbers ε_1 and ε_2 such that

$$\left|x(t) - x^*(t)\right| < \varepsilon_1 \qquad \text{and} \qquad \left|\dot{x}(t) - \dot{x}^*(t)\right| < \varepsilon_2$$

for all t in $\langle t_0, t_1 \rangle$, than

$\delta x(t) = x(t) - x^*(t)$ is said to be a weak variation of $x^*(t)$.

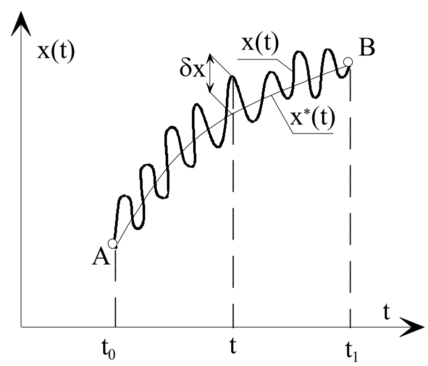


Definition: A strong variation

Let $x^*(t)$ be a minimizing curve and x(t) an admissible curve. If there exists a small number ε such that

$$\left|x(t) - x^*(t)\right| < \varepsilon \qquad \text{for all t in} < t_0, t_1 >, \text{ then}$$

$$\delta x(t) = x(t) - x^*(t) \text{ is said to be a strong variation of } x^*(t).$$



Remark: Weak variations are a sub–class of a class of strong variations. Some strong variations are weak variations. The inversed statement is not true.

The first necessary condition of local minimum

Let the neighbour function be in the form

$$x(t) = x^*(t) + \varepsilon\eta(t), \quad \text{where}$$

ε - a small quality, $\delta x(t) = \varepsilon\eta(t)$ -a weak admissible variation, η(t) – a function of the class C^2 satisfying boundary conditions

$$\eta(t_0) = \eta(t_1) = 0\,.$$

Definition: The increment of the functional I is

$$\Delta I \stackrel{\Delta}{=} I(x(t)) - I(x^*(t)).$$

Thus

$$\Delta I = I(x^* + \varepsilon\eta) - I(x^*) = \int_{t_0}^{t_1} \left[f(t, x^* + \varepsilon\eta, \dot{x}^* + \varepsilon\dot{\eta}) - f(t, x^*, \dot{x}^*) \right] dt\,.$$

At each t in $< t_0, t_1 >$ we expand $f(t, x^* + \varepsilon\eta, \dot{x}^* + \varepsilon\dot{\eta})$ in a Taylor series in the two variables $\varepsilon\eta$ and $\varepsilon\dot{\eta}$ about $(x^*(t), \dot{x}^*(t))$. All partial derivatives of f are to be evaluated on the minimizing curve.

$$\Delta I = \varepsilon \underbrace{\int_{t_0}^{t_1} \left(\frac{\partial f}{\partial x}\eta + \frac{\partial f}{\partial \dot{x}}\dot{\eta} \right) dt}_{V_1} + \frac{1}{2}\varepsilon^2 \underbrace{\int_{t_0}^{t_1} \left(\frac{\partial^2 f}{\partial x^2}\eta^2 + 2\frac{\partial^2 f}{\partial x \partial \dot{x}}\eta\dot{\eta} + \frac{\partial^2 f}{\partial \dot{x}^2}\dot{\eta}^2 \right) dt}_{V_2} + 0\left(\varepsilon^2\right).$$

$\delta I = \varepsilon V_1$ is called the first variation of the functional,

$\delta^2 I = \frac{1}{2}\varepsilon^2 V_2$ is called the second variation.

From definition of minimum $\Delta I \geq 0$

$$\Delta I = \varepsilon\left(V_1 + 0.5\varepsilon V_2 + 0\left(\varepsilon^2\right)\right) \geq 0 .$$

The small quality ε may be positive or negative, therefore

$$V_1 + 0.5\varepsilon V_2 + 0\left(\varepsilon^2\right) \geq 0, \quad \text{for } \varepsilon > 0 ,$$

$$V_1 + 0.5\varepsilon V_2 + 0\left(\varepsilon^2\right) \leq 0, \quad \text{for } \varepsilon < 0.$$

Now let $\varepsilon \to 0$ (positive or negative). We must have both $V_1 \geq 0$ and $V_1 \leq 0$. It is possible for $V_1 = 0$ for all $\eta(t)$.

The necessary condition is

$$V_1 = \int_{t_0}^{t_1} \left(\frac{\partial f}{\partial x}\eta + \underbrace{\frac{\partial f}{\partial \dot{x}}\dot{\eta}}_{\substack{\text{integrating} \\ \text{by parts}}} \right) dt = 0 \qquad \text{for all } \eta(t).$$

Integrating by parts

$$V_1 = \int_{t_0}^{t_1} \left(\frac{\partial f}{\partial x}\eta \right) dt + \left[\eta \frac{\partial f}{\partial \dot{x}} \right]_{t_0}^{t_1} - \int_{t_0}^{t_1} \eta \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}} \right) dt = 0 .$$

$$\left[\eta \frac{\partial f}{\partial \dot{x}} \right]_{t_0}^{t_1} = 0 \qquad \text{for} \qquad \eta(t_0) = \eta(t_1) = 0.$$

Because $\eta(t_0) = \eta(t_1) = 0$ we can re–write as

$$\int_{t_0}^{t_1} \underbrace{\eta}_{\text{free}} \underbrace{\left\{ \frac{\partial f}{\partial x} - \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}} \right) \right\}}_{\text{continuous}} dt = 0 .$$

It may be shown that, if $x = x^*(t)$ is of C^2 class (the term in the curly brackets is continuous function of time t) and $\eta(t)$ is unspecified, the term in the curly brackets should be equal zero. The proof is not trivial, however.

Theorem: The necessary condition of the local minimum of the functional I on the curve $x = x^*(t)$ of C^2 class is that

$$\boxed{\frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}} \right) - \frac{\partial f}{\partial x} = 0}.$$

This differential equation is called the **Euler–Lagrange equation**.

Example: Find the extremal of the functional

$$I = \int_1^2 \underbrace{\dot{x}^2 t^3}_{f(t,x,\dot{x})} dt$$

for the boundary conditions $x(1) = 0$, $x(2) = 3$.

Solution:

$$f(t,x,\dot{x}) = \dot{x}^2 t^3; \qquad \frac{\partial f}{\partial x} = 0; \qquad \frac{\partial f}{\partial \dot{x}} = 2\dot{x}t^3$$

The Euler – Lagrange equation is

$$\frac{d}{dt}\left(2\dot{x}t^3\right) = 0 \ .$$

The integral is $\dot{x}t^3 = C_1 = \text{const}$, or

$$\frac{dx}{dt} = \frac{C_1}{t^3} \ .$$

Separating variables $\quad x = \int \frac{C_1}{t^3}dt + C_2 \quad$ and $\quad x = -\frac{C_1}{2t^2} + C_2$.

When we apply the end conditions

$$\left.\begin{aligned} 0 &= -\frac{C_1}{2} + C_2 \\ 3 &= -\frac{C_1}{8} + C_2 \end{aligned}\right\}$$

we find

$$\boxed{x = -\frac{4}{t^2} + 4} \ .$$

The fixed–end points problem for n unknown functions

Minimize the functional

$$I(\mathbf{x}) = \int_{t_0}^{t_1} f(t, x_i, \dot{x}_i dt) \ \rightarrow \ \text{MIN} \qquad \text{for } i = 1,\ldots, n$$

and boundary conditions: $\qquad x_i(t_0) = x_i^0, \qquad x_i(t_1) = x_i^1.$

The increment of the functional is

$$\Delta I \overset{\Delta}{=} \int_{t_0}^{t_1}\left[f\left(t, x_i^* + \underbrace{\varepsilon}_{\substack{\text{the same}\\ \text{for all } \eta_i}} \eta_i, \dot{x}_i^* + \varepsilon\dot{\eta}_i \right) - f\left(t, x_i^*, \dot{x}_i^*\right)\right]dt \underbrace{=}_{\substack{\text{expanding in}\\ \text{Taylor series}}} \varepsilon \int_{t_0}^{t_1} \sum_{i=1}^{n}\left(\frac{\partial f}{\partial x_i}\eta_i + \underbrace{\frac{\partial f}{\partial \dot{x}_i}\dot{\eta}_i}_{\substack{\text{integrating}\\ \text{by parts}}}\right)dt + 0\left(\varepsilon^2\right),$$

$$\delta I = \left\{ \left[\sum_{i=1}^{n} \frac{\partial f}{\partial \dot{x}_i} \underbrace{\eta_i}_{\substack{=0_i \text{ employing}\\ \text{bonduary}\\ \text{conditions}}} \right]_{t_0}^{t_0} + \int_{t_0}^{t_1} \sum_{i=1}^{n} \underbrace{\left(\frac{\partial f}{\partial x_i} - \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}_i}\right)\right)}_{\substack{=0_i \text{ because } \eta_i \text{ are independent}\\ \text{each other}}} \eta_i dt \right\} = 0 .$$

The terms in the brackets are equal to zero, $(\cdot) = 0_i$, because η_i are independent each other. We can select them identically equal to zero $\eta_i \equiv 0$ except one. Finally we can obtain a system of Euler-Lagrange equations

$$\boxed{\frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}_i}\right) - \frac{\partial f}{\partial x_i} = 0} \qquad\qquad i = 1,\ldots,n.$$

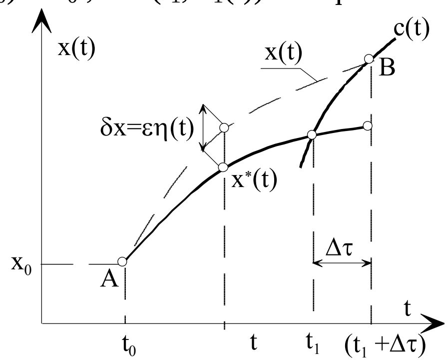
There are ordinary differential equations of the second order. Their solutions contain 2n arbitrary constants which may be computed employing boundary conditions:

$$x_i(t_0) = x_i^{\circ}, \qquad x_i(t_1) = x_i^1 .$$

Problem in which the end point is not fixed
Minimize the functional

$$I = \int_{t_0}^{t_1} f(t, x, \dot{x})dt \quad \rightarrow \quad MIN$$

when (t_0, x_0) is fixed, i.e. $x(t_0) = x_0$, but $(t_1, x_1(t))$ is required to lie on a given curve $\boxed{x = c(t).}$



Let $x^*(t)$ be the minimizing curve, c(t) – the target curve,

$\delta x = x(t) - x^*(t) = \varepsilon\eta(t)$ be a weak variation,

$\Delta\tau$ - is small, it is $0(\varepsilon)$ – it is small of degree ε.

In the previous section the variation at the final point was zero, $\delta x(t_1) = \varepsilon\eta(t_1) = 0$.
Here such condition should be replaced by following considerations.

$$x(t_1 + \Delta\tau) = x^*(t_1 + \Delta\tau) + \varepsilon\eta(t_1 + \Delta\tau) \underbrace{=}_{\substack{\text{expanding}\\ \text{in Taylor}\\ \text{series}}} x^*(t_1) + \dot{x}^*(t_1)\Delta\tau + \varepsilon\eta(t_1) + 0(\varepsilon^2) .$$

But the final point should be on a target curve

$$x(t_1 + \Delta\tau) = c(t_1 + \Delta\tau) \underbrace{=}_{\text{expanding}} c(t_1) + \dot{c}(t_1)\Delta\tau + 0(\varepsilon^2) .$$

Right hand sides of above equations are equal each other, therefore

$$x^*(t_1) + \dot{x}^*(t_1)\Delta\tau + \varepsilon\eta(t_1) = c(t_1) + \dot{c}(t_1)\Delta\tau$$

because $x^*(t_1)$ should reach the target curve $c(t_1)$, and

$$(*) \qquad \boxed{\varepsilon\eta(t_1) = \left[\dot{c}(t_1) - \dot{x}^*(t_1)\right] \Delta\tau} .$$

The variation in functional I is given by

$$\delta I=\int_{t_0}^{t_1+\Delta\tau} f\left(t,x^*+\varepsilon\eta,\dot{x}^*+\varepsilon\dot{\eta}\right)dt-\int_{t_0}^{t_1} f\left(t,x^*,\dot{x}^*\right)dt=$$

$$=\int_{t_0}^{t_1}\underbrace{\left[f\left(t,x^*+\varepsilon\eta,\dot{x}^*+\varepsilon\dot{\eta}\right)-f\left(t,x^*,\dot{x}^*\right)\right]}_{\substack{\text{This integrand was considered ealier.}\\ \text{Expanding and integrating by parts.}}}dt+\underbrace{f\left(t_1,x^*+\varepsilon\eta,\dot{x}^*+\varepsilon\dot{\eta}\right)\Delta\tau}_{\text{Expanding in Tylor series}}=$$

$$=\int_{t_0}^{t_1}\varepsilon\underbrace{\eta}_{\text{it is arbitrary}}\underbrace{\left\{\frac{\partial f}{\partial x}-\frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}}\right)\right\}}_{=0,\ \text{E–L equation}}dt+\varepsilon\eta\left(t_1\right)\frac{\partial f}{\partial \dot{x}}\left(t_1\right)-\varepsilon\underbrace{\eta\left(t_0\right)}_{\substack{=0\text{ from}\\ \text{initial condition}}}\frac{\partial f}{\partial \dot{x}}\left(t_0\right)+f\left(t_1,x^*,\dot{x}^*\right)\Delta\tau+$$

$$+\left.\frac{\partial f}{\partial x}\right|_{\substack{x=x^*\\ t=t_1}}\underbrace{\varepsilon\eta\,\Delta\tau}_{=0\left(\varepsilon^2\right)}+\left.\frac{\partial f}{\partial \dot{x}}\right|_{\substack{x=x^*\\ t=t_1}}\varepsilon\dot{\eta}\Delta\tau\underbrace{=}_{\substack{\text{the necessary condition}\\ \text{of extremum }\delta I=0}}0.$$

We retain only terms of order ε. We use the fact that $x^*(t)$ is an extremal to obtain

$$\delta I=f\left(t_1\right)\Delta\tau+\underbrace{\varepsilon\eta\left(t_1\right)}_{\text{from}(*)}\frac{\partial f}{\partial \dot{x}}\left(t_1\right)+0\left(\varepsilon^2\right)=0\,.$$

Employing (∗)

$$\delta I=\underbrace{\left\{f\left(t_1\right)+\left[\dot{c}\left(t_1\right)-\dot{x}^*\left(t_1\right)\right]\frac{\partial f}{\partial \dot{x}}\left(t_1\right)\right\}}_{=0}\underbrace{\Delta\tau}_{\neq 0}+\underbrace{0\left(\varepsilon^2\right)}_{\text{small}}=0\,.$$

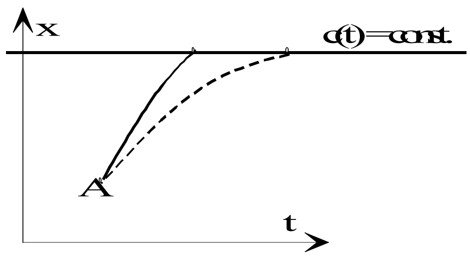
Finally

$$(\nabla\nabla)\qquad \boxed{f\left(t_1\right)+\left[\dot{c}\left(t_1\right)-\dot{x}^*\left(t_1\right)\right]\frac{\partial f}{\partial \dot{x}}\left(t_1\right)=0}\,.$$

This is the **transversality condition**. It replaces the condition $x^*(t_1) = x_1$, which is not valid. It is an algebraic equation relating the slope of the extremal and the slope of the target curve at the point at which they meet. We have derived it for the case in which the right-hand end–point lies on a given curve.

Special forms of the transversality condition

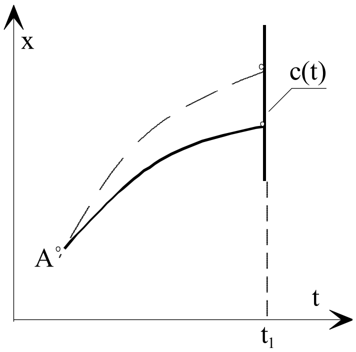
A) If $x(t_1)$ is fixed and t_1 is free.



Thus, $\dot{c}\left(t_1\right)=0$, and (∇∇) takes the simplified form

$$\boxed{f\left(t_1\right)-\dot{x}^*\left(t_1\right)\frac{\partial f}{\partial \dot{x}}\left(t_1\right)=0}.$$

B) If t_1 is fixed but $x(t_1)$ is free.



Dividing (∇∇) by $\dot{c}(t) \neq 0$

$$\underbrace{\frac{1}{\dot{c}(t_1)}\left\{f(t_1) - \dot{x}^*(t_1)\frac{\partial f}{\partial \dot{x}}(t_1)\right\}}_{\to 0} + \frac{\partial f}{\partial \dot{x}}(t_1) = 0 \ .$$

As $\dot{c}(t_1)$ goes to infinity the first term goes to zero.
Finally

$$\boxed{\frac{\partial f}{\partial \dot{x}}(t_1) = 0}.$$

Example: Find the extremal of the functional

$$I = \int_1^T \underbrace{\dot{x}^2 t^3}_{=f} dt$$

with the initial condition x(1) = 0 and with x(T) which lies on the curve

$$x(T) = c(T) = \frac{2}{T^2} - 3 .$$

Solution: From the previous example it follows

$$x = -\frac{C_1}{2t^2} + C_2 .$$

From the initial condition

$$0 = -\frac{C_1}{2} + C_2 \quad \text{therefore} \quad C_2 = \frac{C_1}{2}$$

Substituting into the solution of E-L equation containing arbitrary constants

$$x = \underbrace{-\frac{C_1}{2}}_{=k=const}\left(\frac{1}{t^2} - 1\right) = \frac{k}{t^2} - k$$

$$\dot{x} = \frac{-2k}{t^3} .$$

The target curve is $c(T) = \frac{2}{T^2} - 3$, so $\dot{c}(T) = \frac{-4}{T^3}$.

The transversality condition (∇∇) is

$$\dot{x}(T)^2 \cdot T^3 + [\dot{c}(T) - \dot{x}(T)] \cdot [2\dot{x}(T) \cdot T^3] = 0$$

which becomes after substitution of $\dot{x}(T)$ and $\dot{c}(T)$

$$\left(\frac{-2k}{T^3}\right)^2 T^3+\left[-\frac{4}{T^3}+\frac{2k}{T^3}\right]\cdot\left(\frac{-4k}{T^3}\right)\cdot T^3=0\,.$$

T stands for time, therefore T > 0 and $k \neq 0$ (otherwise the extremal would be $x(t) \equiv 0$), so the transversality condition gives us one root $\boxed{k=4}$.

The required extremal is $\boxed{x=\frac{4}{t^2}-4}$ which meets the target curve at $T=\sqrt{2}$ and x(T) = -2.

Isoperimetric problem

Consider the problem of finding a curve which minimizes a given functional while giving another functional an assigned value. The name **isoperimetric** means the closed curves of the same length, that is of equal **perimeter**. These curves maximizes the enclosed area. Such a problem is historically oldest one.

The problem is formulated as follows. Minimize

$$I(x)=\int_{t_0}^{t_1} f(t,x,\dot{x})dt\,,$$

with the boundary conditions $x(t_0) = x_0,\ x(t_1) = x_1$ subject to the integral constraint

$$\tilde{I}(x)=\int_{t_0}^{t_1} g(t,x,\dot{x})dt=c\,,$$

where c is a given constant.

The legend about Dido: Legendary daughter of a king of Tyre. She founded Carthage and became its queen.

Theorem: In order that $x = x^*(t)$ be solution of the isoperimetric problem is necessary that, for a certain constant λ, $x = x^*(t)$ is the extremal of the augmented functional

$$\int_{t_0}^{t_1}\left[f(t,x,\dot{x})+\lambda\, g(t,x,\dot{x})\right]dt\,.$$

Proof (rather tricky): We will take admissible curves to be C^2 class, and the weak variation of the form

$$\delta x = x(t)-x^*(t)=\varepsilon\sigma(t)\,,$$

where ε is a small number, therefore

$$x(t)=x^*(t)+\varepsilon\,\sigma(t),\qquad\qquad \sigma(t_0)=\sigma(t_1)=0\,.$$

Varied curves pass through the end-points.

We rewrite $\sigma(t)$ in the form $\varepsilon\sigma(t)=\alpha\eta(t)+\beta\zeta(t)$, where α, β are constant and $\eta(t)$, $\zeta(t)$ vanish at the end-points. The functions $\eta(t)$ and $\zeta(t)$ are arbitrary and independent. There is no such constant k that $\eta(t) = k\,\zeta(t)$ for all t.

The increment of constraint is

$$\Delta\tilde{I}=\int_{t_0}^{t_1}\left[g(t,x^*+\varepsilon\sigma,\dot{x}^*+\varepsilon\dot{\sigma})-g(t,x^*,\dot{x}^*)\right]dt=c-c=0\,.$$

We expand in Taylor series around $x^*(t)$

$$\delta\tilde{I} = \int_{t_0}^{t_1}\left[\frac{\partial g}{\partial x}(\alpha\eta+\beta\zeta)+\underbrace{\frac{\partial g}{\partial \dot{x}}(\alpha\dot{\eta}+\beta\dot{\zeta})}_{\substack{\text{by parts +}\\ \text{end conditions}}}\right]dt = 0 \ .$$

Integrating by parts gives

$$\delta\tilde{I} = \int_{t_0}^{t_1}(\alpha\eta+\beta\zeta)\underbrace{\left[\frac{\partial g}{\partial x}-\frac{d}{dt}\left(\frac{\partial g}{\partial \dot{x}}\right)\right]}_{\substack{L(g)-\\ \text{linear operator}}}dt = 0 .$$

Let L(h) denote the linear operator

$$L(h)\overset{\Delta}{=}\frac{\partial h}{\partial x}-\frac{d}{dt}\left(\frac{\partial h}{\partial \dot{x}}\right),$$

so that we have

$$\delta\tilde{I} = \int_{t_0}^{t_1}(\alpha\eta+\beta\zeta)\ \underbrace{L(g)}_{\substack{\neq 0 \text{ since } x^*\\ \text{is not an}\\ \text{extremal of } \ddot{I}}}\ dt = 0 \ .$$

Since $x^*(t)$ is minimizing curve for I, $\delta I = 0$

$$\delta I = \int_{t_0}^{t_1}(\alpha\eta+\beta\zeta)\ \underbrace{L(f)}_{\substack{\neq 0 \text{ since}\\ (\alpha\eta+\beta\zeta)\\ \text{is dependent}}}\ dt = 0 .$$

The above conditions should be satisfied simultaneously. The conditions may be regarded as an algebraic equations relative to α and β. Eliminating α and β between the above equations gives.

$$\frac{\int_{t_0}^{t_1}L(f)\eta dt}{\int_{t_0}^{t_1}L(g)\eta dt}=\frac{\int_{t_0}^{t_1}L(f)\zeta dt}{\int_{t_0}^{t_1}L(g)\zeta dt}\underbrace{=}_{\substack{\text{since } \eta \text{ and } \zeta\\ \text{independent}\\ \text{functions of } C^2\\ \text{class}}} -\lambda = \text{const.}$$

Since $\eta(t)$ and $\zeta(t)$ are independent functions of C^2 class this can only be true if both sides are equal to constant $-\lambda$, say. Hence $x^*(t)$ must be such that

$$\int_{t_0}^{t_1}\underbrace{L(f+\lambda g)}_{\substack{=0 \text{ since } \eta(t)\\ \text{is arbitrary}}}\eta\, dt = 0 \qquad \text{for all admissible } \eta(t).$$

Then the necessary condition of optimality is

$$\frac{\partial}{\partial x}(f+\lambda g)-\frac{d}{dt}\left[\frac{\partial}{\partial \dot{x}}(f+\lambda g)\right]=0 \qquad \text{on } x = x^*(t).$$

Constant λ is termed Lagrange multiplier.

Algorithm:

1. Construct the functional $\int_{t_0}^{t_1}(f+\lambda g)dt$.

2. For this functional proceed Euler-Lagrange equation.
3. Find the solution of this equation containing constants C_1 and C_2.
4. Two constants C_1, C_2 and the multiplier λ obtain employing the end conditions and the isoperimetric constraint $\tilde{I} = c$.

Example: Find the extremal

$$I = \int_0^1 \dot{x}^2 dt$$

with the boundary conditions: x(0) = 2, x(1) = 4, and isoperimetric constraint

$$\tilde{I} = \int_0^1 x \, dt = 1.$$

Solution: The augmented functional takes the form

$$\int_0^1 \underbrace{\left(\dot{x}^2 + \lambda x\right)}_{f(t,x,\dot{x})} dt$$

and Euler-Lagrange equation $\dfrac{\partial f}{\partial x} - \dfrac{d}{dt}\left(\dfrac{\partial f}{\partial \dot{x}}\right) = 0$, the form

$$\lambda - \frac{d}{dt}\left[2\dot{x}\right] = 0.$$

It has the solution

$$x = \frac{\lambda t^2}{4} + C_1 t + C_2.$$

The end conditions give: $C_2 = 2$; $C_1 = 2 - \dfrac{\lambda}{4}$.

We find λ by applying the isoperimetric constraint

$$\int_0^1 \left\{\frac{\lambda}{4} t^2 + \left(2 - \frac{\lambda}{4}\right) t + 2\right\} dt = 1.$$

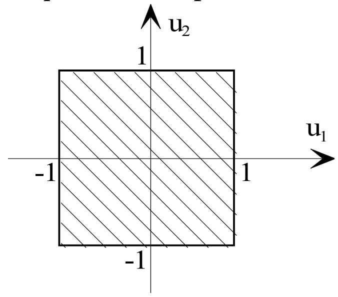
This gives $\lambda = 48$. The required extremal is the curve

$$x = 12 t^2 - 10t + 2.$$

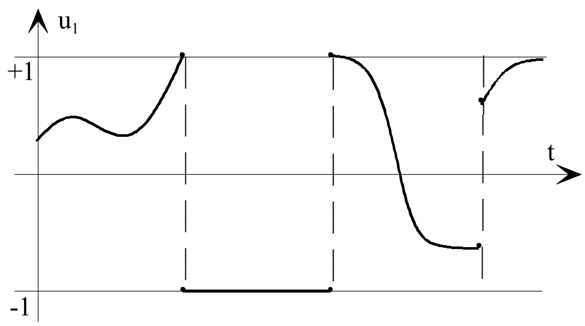
The optimal control problem

The state of the system at time t is described by a vector **x**(t) in n–dimensional Euclidean space X. This space is called **the state space**. The behaviour of the system (the state of the system) is represented by the point in this abstract space, for a given instant of time t. As the system evolves in time **x** trace out a continuous path in its state space. The control we model as an r–dimensional vector function of time, **u**. The components of **u** are allowed to be piecewise continuous. The values we can take are bounded, so that at any time t, $\mathbf{u} \in U$, it lies in some bounded region U of the control space. The control such that $\mathbf{u} \in U$ is called **admissible**.

Example: For r = 2, $\mathbf{u} = \lfloor u_1(t), u_2(t) \rfloor^T$, we impose the restriction $|u_i| \le 1$, i = 1,2. The admissible control belongs to the unit square in the plane.



Example: The piecewise continuous function has a finite number of discontinuity of the first kind.



Behaviour of the system is described by a set of n ordinary differential equations

$$\dot{x}_i = f_i(\mathbf{x}, \mathbf{u}), \qquad i = 1,\dots,n$$

or in vector form

$$\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{u}). \tag{1}$$

These are **the state equations**. We a assume that the function f_i are defined, continuous and continuously differentiable with respect to $\mathbf{x} = \lfloor x_1,\dots,x_n \rfloor^T$ for all $\mathbf{x} \in X$ and for all admissible controls $\mathbf{u} = \lfloor u_1,\dots,u_r \rfloor^T$.

The state equations should be supplemented by the boundary conditions

$$\mathbf{x}(t) = \mathbf{x}^0 \qquad \text{and} \qquad \mathbf{x}(t_1) = \mathbf{x}^1. \tag{2}$$

The state vector x(t) is continuous even though the control is piecewise continuous.

The state equations

$$\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{u}, t)$$

may be transformed to the equations of the form (1) substituting

$$x_{n+1} = t,$$

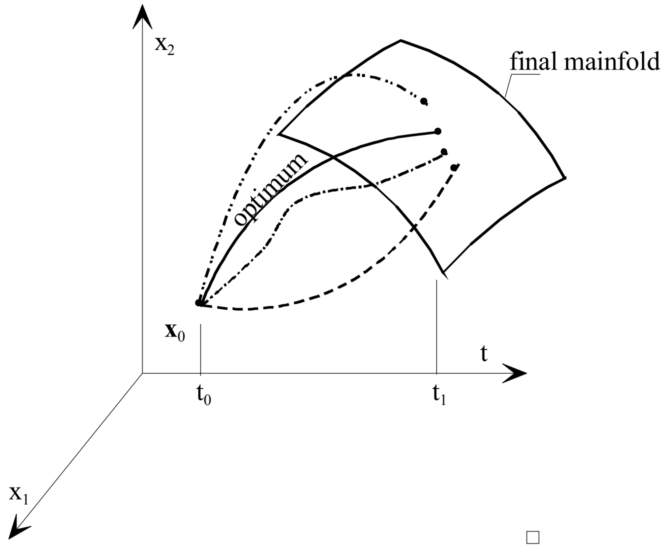
and adding the differential equation

$$\frac{dx_{n+1}}{dt} = 1.$$

We wish control the system from $\mathbf{x}^0$ at t_0 to $\mathbf{x}^1$ at $t = t_1$ in such a way that the performance index (cost functional)

$$I = \int_{t_0}^{t_1} f_0(\mathbf{x}, \mathbf{u})\,dt \quad \to \quad MIN$$

is minimized.



□

The final time t_1 may be specified or unspecified. We assume that the optimal control $\mathbf{u}^*(t)$ exists.

Proof of Pontryagin's maximum principle (1956)

Additional assumptions:

1. the control is unbounded , i.e. U is the whole control space,
2. the dimension of the state space is 2, n = 2, i.e. $\mathbf{x} = \lfloor x_1, x_2 \rfloor^T$,
3. the dimension of the control space is 1, r = 1, i.e. $\mathbf{u} = u$.

Now

$$\begin{cases} \dot x_1 = f_1(x_1, x_2, u) \\ \dot x_2 = f_2(x_1, x_2, u) \end{cases} .$$

Let $u^*(t)$ be an optimal control and $\mathbf{x}^*(t)$ the corresponding optimal path. Consider the perturbed control and $u = u^* + \delta u(t)$ and corresponding perturbed state vector: $x_1 = x_1^* + \delta x_1$; $x_2 = x_2^* + \delta x_2$. The end conditions are fixed

$$x_i(t_1) = x_i^1, \qquad i = 1, 2.$$

The perturbed values of x_i^1 at the time instant $t_1 + \delta t$ (time t_1 is not fixed) are

$$x_i^*(t_1 + \delta t) + \delta x_i(t_1 + \delta t) = x_i^1, \qquad i = 1, 2 \tag{4}$$

Expanding in Taylor series and saving only the first-order effects, we deduce that

$$x_i^*(t_1) + \dot x_i^*(t_1)\delta t + \delta[x_i(t_1) + \dot x_i(t_1)\delta t] = x_i^1$$

$$\cancel{x_i^*(t_1)} + \dot x_i^*(t_1)\delta t + \delta x_i(t_1) + \underbrace{\cancel{\delta \dot x_i(t_1)\delta t} + \cancel{\dot x_i(t_1)\delta^2 t}}_{\text{small quantities}} = \cancel{x_i^1}$$

therefore

$$\delta x_i(t_1) + \dot x_i^*(t_1)\,\delta t = 0. \tag{5}$$

If we now use the right-hand sides of the state equations we obtain

$$\boxed{\delta x_i(t_1) = -f_i(t_1)\delta t}. \tag{6}$$

Remark: We denote

$$f_i(t) = f_i\left(x_1^*(t), x_2^*(t), u^*(t)\right)$$

that is considered on the optimal path. We adopt the same convention for $\partial f_i / \partial x_i$ and $\partial f_i / \partial u$.

The consequent change ΔI in I is

$$\Delta I=\int_{t_0}^{t_1+\delta t} f_0\left(x_1^*+\delta x_1, x_2^*+\delta x_2, u^*+\delta u\right)dt-\int_{t_0}^{t_1} f_0\left(x_1^*, x_2^*, u^*\right)dt \cong$$

$$=\underbrace{\int_{t_0}^{t_1}\left\{\frac{\partial f_0}{\partial x_1}\delta x_1+\frac{\partial f_0}{\partial x_2}\delta x_2+\frac{\partial f_0}{\partial u}\delta u\right\}dt}_{=\delta I=0,\ \text{this is the necessary condition of optimality}}+f_0(t_1)\delta t+0(\circ). \tag{7}$$

The variations: δx_1, δx_2, δu are not independent. They are linked by the state equations. We need to introduce two Lagrange multipliers $\Psi_1(t)$ and $\Psi_2(t)$ which to be time–dependent.

We consider the pair of integrals

$$\Phi_i=\int_{t_0}^{t_1}\Psi_i(t)\underbrace{\left(\dot{x}_i-f_i(x_1,x_2,u)\right)}_{\substack{=0,\ \text{since it is}\\ \text{the state equation}}}dt=0, \qquad i=1, 2. \tag{8}$$

Now we calculate the first variation of Φ_i, $\delta\Phi_i=0$. We consider only the second term, because the first one is zero.

$$\delta\Phi_i=\int_{t_0}^{t_1}\Psi_i(t)\left\{-\frac{\partial f_i}{\partial x_1}\delta x_1-\frac{\partial f_i}{\partial x_2}\delta x_2-\frac{\partial f_i}{\partial u}\delta u+\frac{d}{dt}(\delta x_i)\right\}dt=0. \tag{9}$$

Now

$$\int_{t_0}^{t_1}\Psi_i(t)\frac{d}{dt}(\delta x_i)dt \underbrace{=}_{\text{by parts}} \left[\Psi_i(t)\delta x_i\right]_{t_0}^{t_1}-\int_{t_0}^{t_1}\dot{\Psi}_i\delta x_i dt=-f_i(t_1)\Psi_i(t_1)\delta t-\int_{t_0}^{t_1}\dot{\Psi}_i\,\delta x_i\,dt \tag{10}$$

since $\delta x_i(t_0)=0$ and $\delta x_i(t_1)=-f_i(t_1)\delta t$ from equation (6).
Now, we use equation (10) in equation (9)

$$\delta\Phi_i=-\int_{t_0}^{t_1}\Psi_i(t)\left\{\frac{\partial f_i}{\partial x_1}\delta x_1+\frac{\partial f_i}{\partial x_2}\delta x_2+\frac{\partial f_i}{\partial u}\delta u\right\}dt-\int_{t_0}^{t_1}\dot{\Psi}_i\delta x_i dt-f_i(t_1)\Psi_i(t_1)\delta t=0\ . \tag{11}$$

The condition that $\delta I=0$ can now be replaced by the condition that

$$\delta I+\underbrace{\delta\Phi_1}_{=0}+\underbrace{\delta\Phi_2}_{=0}=0. \tag{12}$$

On substituting for δI equation (7), and for $\delta\Phi_i$ equation (11), rearranging terms we obtain

$$\int_{t_0}^{t_1}\delta x_1\left\{\underbrace{\frac{\partial f_0}{\partial x_1}-\Psi_1\frac{\partial f_1}{\partial x_1}-\Psi_2\frac{\partial f_2}{\partial x_1}}_{-\frac{\partial H}{\partial x_1}}-\dot{\Psi}_1\right\}dt+\int_{t_0}^{t_1}\delta x_2\left\{\underbrace{\frac{\partial f_0}{\partial x_2}-\Psi_1\frac{\partial f_1}{\partial x_2}-\Psi_2\frac{\partial f_2}{\partial x_2}}_{-\frac{\partial H}{\partial x_2}}-\dot{\Psi}_2\right\}dt+$$
$$+\int_{t_0}^{t_1}\delta u\underbrace{\left\{\frac{\partial f_0}{\partial u}-\Psi_1\frac{\partial f_1}{\partial u}-\Psi_2\frac{\partial f_2}{\partial u}\right\}}_{-\frac{\partial H}{\partial u}}dt+\underbrace{\left\{f_0(t_1)-f_1(t_1)\Psi_1(t_1)-f_2(t_1)\Psi_2(t_1)\right\}}_{-H(t_1)}\delta t=0 \tag{13}$$

This can be written more compactly if we introduce the Hamiltonian

$$\boxed{H\overset{\Delta}{=}-f_0(x_1,x_2,u)+\Psi_1 f_1(x_1,x_2,u)+\Psi_2 f_2(x_1,x_2,u)}\ . \tag{14}$$

Then equation (13) takes the form

$$\int_{t_0}^{t_1}\delta x_1\underbrace{\left\{\frac{\partial H}{\partial x_1}+\dot{\Psi}_1\right\}}_{=0}dt+\int_{t_0}^{t_1}\delta x_2\underbrace{\left\{\frac{\partial H}{\partial x_2}+\dot{\Psi}_2\right\}}_{=0}dt+\int_{t_0}^{t_1}\delta u\underbrace{\frac{\partial H}{\partial u}}_{=0}dt-\underbrace{H(t_1)}_{=0}\underbrace{\delta t}_{\text{or }=0}=0\,. \qquad (15)$$

For admissible variations $\{\delta u, \delta x_1, \delta x_2\}$ only one variation is independent. Two of them are dependent via the state equations (1). The multipliers Ψ_1 and Ψ_2 are at our disposal. We choose them to satisfy equations

$$\boxed{\dot{\Psi}_i=-\frac{\partial H}{\partial x_i}}, \qquad i=1,\ 2. \qquad (16)$$

These equations are termed the **adjoint equations** (co-state equations). For fixed time t_1, the variation $\delta t = 0$. For free time t_1 the Hamiltonian $H(t_1) = 0$.

Formulation of the Pontryagin maximum principle

Let $u^*(t)$ be an admissible control with corresponding path $\mathbf{x}^*=\lfloor x_1^*, x_2^*\rfloor^T$ that transfers the system from $\mathbf{x}^0$ at time $t = t_0$ to $\mathbf{x}^1$at some unspecified time t_1. Then in order that u^* and $\mathbf{x}^*$ be optimal (that is minimize I) it is necessary that there exist a non-trivial vector $\boldsymbol{\Psi}=\lfloor \Psi_0, \Psi_1, \Psi_2\rfloor^T$ satisfying adjoint equations (16) and a scalar function

$$H(\boldsymbol{\Psi},\mathbf{x},u)=\Psi_0 f_0(\mathbf{x},u)+\Psi_1 f_1(\mathbf{x},u)+\Psi_2 f_2(\mathbf{x},u)$$

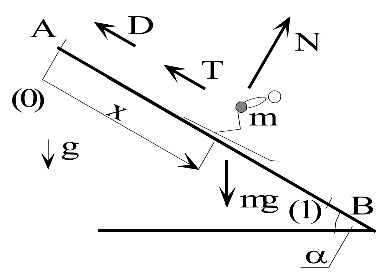
such that

(i) for every t in $t_0 \le t \le t_1$, H attains its maximum with respect to u at $u = u^*(t)$,

(ii) $H(\boldsymbol{\Psi}^*,\mathbf{x}^*,u^*) = 0$ and $\Psi_0 \le 0$ at $t=t_1$, where $\boldsymbol{\Psi}^*(t)$ is the solution of the system (16) for $u=u^*(t)$.

Furthermore it can be shown that $H(\boldsymbol{\Psi}^*(t), \mathbf{x}^*(t), u^*(t))$ = constant, so that $H=0$ and $\Psi_0\le 0$ at each point on an optimal trajectory.

Example:



The skier should cover the given distance from A to B in minimal time. At the point B his velocity v_B is given, and it is lower than the maximal one. Below the point B it is difficult segment of the slope. The velocity at the point A is also given, v_A. The skier uses his aerodynamic drag for braking changing his position (from dropped to upright position). Find the optimal solution of the formulated problem.

Equations of motion

$$\begin{cases} m\dfrac{dv}{dt}=-D-T+mg\sin\alpha, \\ \dfrac{dx}{dt}=v \end{cases} \qquad \begin{array}{l}\text{drag}\quad D=\dfrac{1}{2}\rho S C_x v^2 \\ \text{frictional force}\quad T=\mu N=\mu mg\cos\alpha\,. \end{array}$$

Substituting

$$\begin{cases} \dfrac{dv}{dt} = -\dfrac{1}{2}\dfrac{\rho}{m} SC_x v^2 - \mu \dfrac{mg}{m}\cos\alpha + \dfrac{mg\sin\alpha}{m} \\ \dfrac{dx}{dt} = v\,. \end{cases}$$

We change the independent variable: from t to x.

$$\frac{dv}{dt} = \frac{dv}{dx}\cdot\underbrace{\frac{dx}{dt}}_{=v} = -\frac{1}{2}\frac{\rho}{m}SC_x v^2 + g(\sin\alpha - \mu\cos\alpha)$$

$$\frac{dv}{dx} = -\underbrace{\frac{1}{2}\frac{\rho}{m}SC_x}_{=u\text{ – control variable}} v + \underbrace{\frac{g(\sin\alpha-\mu\cos\alpha)}{v}}_{=\frac{a}{v}},\ \text{a=constant},\quad u_{min} \le u \le u_{max}.$$

Finally

$$\frac{dv}{dx} = \underbrace{-uv + \frac{a}{v}}_{f_1}\ \text{- state equation,}$$

v – state variable
x – independent variable
u – control variable.
Boundary conditions: $v(x_0) = v_0$, $v(x_1) = v_1$, where x_0, x_1 are given.
Final time t_1 is minimized

$$t_1 = \int_{x_0}^{x_1} \underbrace{\frac{1}{v}}_{=f_0} dx \quad \Rightarrow \quad \text{MIN.}$$

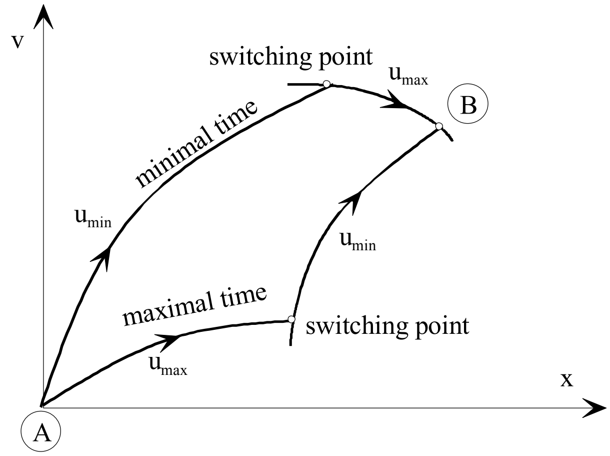
The Hamiltonian
$$\boxed{H = -f_0 + \Psi f_1 = -\frac{1}{v} + \Psi\left(-uv + \frac{a}{v}\right)}\,.$$

The adjoint equation
$$\frac{d\Psi}{dx} = -\frac{\partial H}{\partial v} = -\frac{1}{v^2} + \Psi\left(u + \frac{a}{v^2}\right).$$

$\dfrac{\partial H}{\partial u} = \dfrac{dH}{du} = -\Psi v \ne 0$. The function $\Psi(t) \equiv 0$ is not possible because for such function the adjoint equation cannot be satisfied.
The Hamiltonian takes the maximal value

$$H \to \text{MAX} \qquad \text{for} \quad \begin{cases} u = u_{min} & \text{if} \quad \Psi \ge 0 \\ u = u_{max} & \text{if} \quad \Psi < 0. \end{cases}$$

The control u is bang-bang type.


v
switching point
u_{max}
B
minimal time
u_{min}
u_{min}
maximal time
switching point
u_{max}
x
A