The Calculus of Variations

M. Bendersky *

Revised, December 29, 2008

*These notes are partly based on a course given by Jesse Douglas.

Contents

1 Introduction. Typical Problems

The Calculus of Variations is concerned with solving **Extremal Problems** for a **Functional**. That is to say Maximum and Minimum problems for functions whose domain contains functions, $Y(x)$ (or $Y(x_1, \cdots x_2)$, or n-tuples of functions). The range of the functional will be the real numbers, $\mathbb{R}$

Examples:

I. Given two points $P_1 = (x_1, y_1), P_2 = (x_2, y_2)$ in the plane, joined by a curve, $y = f(x)$. The Length Functional is given by $L_{1,2}(y) = \int_{x_1}^{x_2} \underbrace{\sqrt{1 + (y')^2}dx}_{ds}$. The domain is the set of all curves, $y(x) \in C^1$ such that $y(x_i) = y_i, i = 1, 2$. The minimum problem for $L[y]$ is solved by the straight line segment $\overline{P_1P_2}$.

II. (Generalizing I) The problem of Geodesics, (or the shortest curve between two given points) on a given surface. e.g. on the 2-sphere they are the shorter arcs of great circles (On the Ellipsoid Jacobi (1837) found geodesics using elliptical coördinates in terms of Hyperelliptic integrals, i.e.

$$\int_a f(\sqrt{a_0 + a_1 x + \cdots a_5 x^5} dx, f \text{ rational })$$

.

III. In the plane, given points, P_1, P_2 find a curve of given length ℓ ($> |P_1P_2|$) which together with segment $\overline{P_1P_2}$ bounds a maximum area. In other words, given $\ell = \int_{x_1}^{x_2} \sqrt{1 + (y')^2}dx$, maximize $\int_{x_1}^{x_2} y dx$

This is an example of a problem with given constraints (such problems are also called isoperimetric problems). Notice that the problem of geodesics from P_1 to P_2 on a given surface, $F(x, y, z) = 0$ can also be formulated as a variational problem with constraints:

Given $\quad F(x, y, z) = 0$

Find $\quad y(x), z(x)$ to minimize $\int_{x_1}^{x_2} \sqrt{1 + (\frac{dy}{dx})^2 + (\frac{dz}{dx})^2}dx$,

where $y(x_i) = y_i, z(x_i) = z_i$ for $i = 1, 2$.

IV. Given P_1, P_2 in the plane, find a curve, $y(x)$ from P_1 to P_2 such that the surface of revolution obtained by revolving the curve about the x-axis has minimum surface area. In other words minimize $2\pi \int_{P_1}^{P_2} yds$ with $y(x_i) = y_i$, $i = 1, 2$. If P_1 and P_2 are not too far apart, relative to $x_2 - x_1$ then the solution is a Catenary (the resulting surface is called a Catenoid).

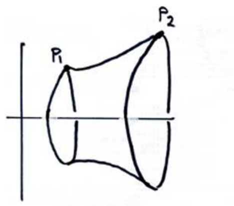


Figure 1: Catenoid

Otherwise the solution is Goldschmidt's discontinuous solution (discovered in 1831) obtained by revolving the curve which is the union of three lines: the vertical line from P_1 to the point $(x_1, 0)$, the vertical line from P_2 to $(x_2, 0)$ and the segment of the x-axis from $(x_1, 0)$ to $(x_2, 0)$.

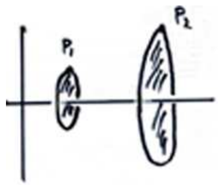


Figure 2: Goldscmidt Discontinuous solution

This example illustrates the importance of the role of the category of functions allowed. If we restrict to continuous curves then there is a solution only if the points are close. If the points are far apart there is a solution only if allow piecewise continuous curves (i.e. continuous except possibly for finitely many jump discontinuities. There are similar meanings for piecewise class C^n.)

The lesson is that the class of unknown functions must be precisely prescribed. If other curves are admitted into "competition" the problem may change. For example the only solutions to minimizing

$$L[y] \underset{\text{def}}{\equiv} \int_a^b (1-(y')^2)^2 dx, \quad y(a) = y(b) = 0.$$

are polygonal lines with $y' = \pm 1$.

V.The Brachistochrone problem. This is considered the oldest problem in the Calculus of Variations. Proposed by Johann Bernoulli in 1696: Given a point 1 higher than a point 2 in a vertical plane, determine a (smooth) curve from $1 \rightarrow 2$ along which a mass can slide along the curve in minimum time (ignoring friction) with the only external force acting on the particle being gravity.

Many physical principles may be formulated in terms of variational problems. Specifically the *least-action principle* is an assertion about the nature of motion that provides an alternative approach to mechanics completely independent of Newton's laws. Not only does the least-action principle offer a means of formulating classical mechanics that is more flexible and powerful than Newtonian mechanics, but also variations on the least-action principle have proved useful in general relativity theory, quantum field theory, and particle physics. As a result, this principle lies at the core of much of contemporary theoretical physics.

VI. Isoperimetric problem In the plane, find among all closed curves, C, of length ℓ the one(s) of greatest area (Dido's problem) i.e. representing the curve by $(x(t), y(t))$: given $\ell = \int\limits_C \sqrt{\dot{x}^2 + \dot{y}^2}dt$ maximize $A = \frac{1}{2}\int\limits_C (x\dot{y} - \dot{x}y)dt$ (recall Green's theorem).

VII. Minimal Surfaces Given a simple, closed curve, $\mathcal{C}$ in $\mathbb{R}^3$, find a surface, say of class C^2, bounded by $\mathcal{C}$ of smallest area (see figure 3).

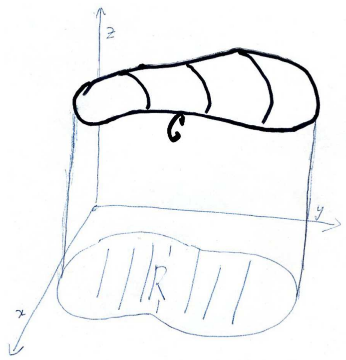


Figure 3:

Assuming a surface represented by $z = f(x, y)$, passes through $\mathcal{C}$ we wish to minimize

$$\iint_R \sqrt{1 + (\frac{\partial z}{\partial x})^2 + (\frac{\partial z}{\partial y})^2} dxdy$$

Proving the existence of minimal surface is Plateau's problem which was solved by Jesse Douglas in 1931.

2 Some Preliminary Results. Lemmas of the Calculus of Variations

Notation 1. *Denote the category of piecewise continuous functions on* $[x_1, x_2]$*. by* $\widetilde{C}[x_1, x_2]$

Lemma 2.1. *(Fundamental or Lagrange's Lemma) Let* $M(x) \in \widetilde{C}[x_1, x_2]$*. If* $\int_{x_1}^{x_2} M(x)\eta(x)dx = 0$ *for all* $\eta(x)$ *such that* $\eta(x_1) = \eta(x_2) = 0, \quad \eta(x) \in C^n, \quad 0 \leq n \leq \infty$ *on* $[x_1, x_2]$ *then* $M(x) = 0$ *at all points of continuity.*

Proof : . Assume the lemma is false, say $M(\overline{x}) > 0$, M continuous at $\overline{x}$. Then there exist a neighborhood, $N_{\overline{x}} = (\overline{x}_1, \overline{x}_2)$ such that $M(x) \geq p > 0$ for $x \in N_{\overline{x}}$. Now take

$$\eta_0(x) \underset{\text{def}}{\equiv} \begin{cases} 0, & \text{in } [x_1, x_2] \text{ outside } N_{\overline{x}} \\ (x - \overline{x}_1)^{n+1}(\overline{x}_2 - x)^{n+1}, & \text{in } N_{\overline{x}} \end{cases}$$

Then $\eta_0 \in C^n$ on $[x_1, x_2]$ and

$$\int_{x_1}^{x_2} M(x)\eta_0(x)dx = \int_{\overline{x}_1}^{\overline{x}_2} M(x)\eta_0(x)dx \geq p \int_{\overline{x}_1}^{\overline{x}_2} (x - \overline{x}_1)^{n+1}(\overline{x}_2 - x)^{n+1}dx \geq 0$$

For the case $n = \infty$ take

$$\eta_0 \equiv \begin{cases} 0, & \text{in } [x_1, x_2] \text{ outside } N_{\overline{x}} \\ e^{\frac{1}{x - \overline{x}_2}} e^{\frac{1}{\overline{x}_1 - x}}, & \text{in } N_{\overline{x}} \end{cases}$$

q.e.d.

Lemma 2.2. *Let* $M(x) \in \widetilde{C}[x_1, x_2]$*. If* $\int_{x_1}^{x_2} M(x)\eta'(x)dx = 0$ *for all* $\eta(x)$ *such that* $\eta \in C^\infty, \quad \eta(x_1) = \eta(x_2) = 0$ *then* $M(x) = c$ *on its set of continuity.*

Proof : (After Hilbert, 1899) Let a, a' be two points of continuity of M. Then for b, b' with $x_1 < a < b < a' < b' < x_2$ we construct a C^∞ function[1] $\eta_1(x)$ satisfying

$$\begin{cases} 0, & \text{on } [x_1, a] \text{ and } [b', x_2] \\ p \text{ (a constant } > 0), & \text{on } [b, a'] \\ \text{increasing on } [a, b], & \text{decreasing on } [a', b'] \end{cases}$$

Step 1: Let $\widehat{\eta}_0$ be as in lemma (2.1)

$$\widehat{\eta}_0 = \begin{cases} 0, & \text{in } [x_1, x_2] \text{ outside } [a, b] \\ e^{\frac{1}{x-\overline{x}_2}} e^{\frac{1}{\overline{x}_1 - x}}, & \text{in } [a, b] \end{cases}$$

Step 2: For some c such that $b < c < a'$ and $x_1 \leq x \leq c$ set

$$\eta_1(x) = \frac{p}{\int_a^b \widehat{\eta}_0(t)dt} \int_a^x \widehat{\eta}_0(t)dt$$

Similarly for $c \leq x \leq x_2$ define $\eta_1(x)$ by

$$\eta_1(x) = \frac{p}{\int_{a'}^{b'} \widehat{\widehat{\eta}}_0(t)dt} \int_x^{b'} \widehat{\widehat{\eta}}_0(t)dt$$

where $\widehat{\widehat{\eta}}_0(x)$ is defined similar to $\widehat{\eta}_0(t)$ with $[a', b']$ replacing $[a, b]$.

Now

$$\int_{x_1}^{x_2} M(x)\eta_1'(x)dx = \int_a^b M(x)\eta_1'(x)dx + \int_{a'}^{b'} M(x)\eta_1'(x)dx$$

where $M(x)$ is continuous on $[a, b], [a', b']$. By the mean value theorem there are $\alpha \in [a, b], \alpha' \in [a', b']$ such that the integral equals $M(\alpha) \int_a^b \eta_1'(x)dx + M(\alpha') \int_{a'}^{b'} \eta_1'(x)dx = p(M(\alpha) -$

[1] If M were differentiable the lemma would be an immediate consequence of integration by parts.

$M(\alpha'))$. By the hypothesis this is 0. Thus in any neighborhood of a and a' there exist α, α' such that $M(\alpha) = M(\alpha')$. It follows that $M(a) = M(a')$. **q.e.d.**

- η_1 in lemma (2.2) may be assumed to be in C^n. One uses the C^n function from lemma (2.1) in the proof instead of the C^∞ function.

- It is the fact that we imposed the endpoint condition on the test functions, η, that allows non-zero constants for M. In particular simply integrating the bump function from lemma (2.1) does not satisfy the condition.

- The lemma generalizes to:

  Lemma 2.3. *If $M(x)$ is a piecewise continuous function such that*

  $$\int_{x_1}^{x_2} M(x)\eta^{(n)}(x)dx = 0$$

  for every function that has a piecewise continuous derivative of order n and satisfies $\eta^{(k)}(x_i) = 0, i = 1, 2, \quad k < n$ then $M(x)$ is a polynomial of degree $n - 1$.

  (see [AK], page 197).

Definition 2.4. The normed linear space $\mathcal{D}^n(a, b)$ consist of all continuous functions, $y(x) \in \tilde{C}^n[a, b]$[2] with bounded norm $||y||_n = \sum_{i=0}^{n} \max_{x_1 \le x \le x_2} |y^{(i)}(x)|$.[3]

If a functional, $J : \mathcal{D}^n(a, b) \to \mathbb{R}$ is continuous we say J is continuous with respect to $\mathcal{D}^n$.

[2] i.e. have continuous derivatives to order n except perhaps at a finite number of points.

[3] $||y||_n$ is a norm because f is assumed to be continuous on $[a, b]$.

The first examples we will study are functionals of the form

$$J[y] = \int_a^b f(x, y, y')ds$$

e.g. the arc length functional. It is easy to see that such functionals will be continuous with respect to $\mathcal{D}^1$ but are not continuous as functionals from $\mathcal{C} \to \mathbb{R}$. In general functionals which depend on the n-th derivative are continuous with respect to $\mathcal{D}^n$, but not with respect to $\mathcal{D}^k, k < n$.

We assume we are given a function, $f(x, y, z)$ say of class C^3 for $x \in [x_1, x_2]$, and for y in some interval (or region, G, containing the point $\overline{y} = (y_1, y_2, \cdots, y_n)$) and for all real z (or all real vectors, $\overline{z} = (z_1, z_2, \cdots, z_n)$).

Consider functions, $y(x) \in \mathcal{D}^1[x_1, x_2]$ such that $y(x) \in G$. Let $\mathcal{M}$ be the set of all such $y(x)$. For any such $y \in \mathcal{M}$ the integral

$$J[y] \underset{\text{def}}{\equiv} \int_a^b f(x, y(x), y'(x))dx$$

defines a functional $J : \mathcal{M} \to \mathbb{R}$.

Problem: To find relative or absolute extrema of J.

Definition 2.5. Let $y = y_0(x), a \leq x \leq b$ be a curve in $\mathcal{M}$.

(a) A **strong ϵ neighborhood** of y_0 is the set of all $y \in \mathcal{M}$ in an ϵ ball centered at y_0 in $\mathcal{C}$.

(b) a **weak ϵ neighborhood** of y_0 is an ϵ ball in $\mathcal{D}^1$ centered at y_0. [4]

[4] For example let $y(x) = \frac{1}{n}\sin(nx)$. If $\frac{1}{n} < \epsilon < 1$, then y_0 lies in a strong ϵ neighborhood of $y_0 \equiv 0$ but not in a weak ϵ neighborhood.

A function $y_0(x) \in \mathcal{M}$ furnishes a *weak relative minimum* for $J[y]$ if and only if $J[y_0] < J[y]$ for all y in a weak ϵ neighborhood of y_0. It furnishes a *strong relative minimum* if $J[y_0] < J[y]$ for all y in a strong ϵ neighborhood of y_0. If the inequalities are true for all $y \in \mathcal{M}$ we say the minimum is *absolute.* If $<$ is replaced by $\leq$ the minimum becomes *improper.* There are similar notions for maxima instead of minima[5]. In light of the comments above regarding continuity of a functional, we are interested in finding weak minima and maxima.

Example(A Problem with no minimum)

Consider the problem to minimize

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

on $\mathcal{D}^1 = \{y \in C^1[0,1], y(0) = 0, y(1) = 1\}$ Observe that $J[y] > 1$. Now consider the sequence of functions in $\mathcal{D}^1$, $y_k(x) = x^k$. Then

$$J[y_k] = \int_0^1 x^{k-1}\sqrt{x^2 + k^2}dx \leq \int_0^1 x^{k-1}(x+k)dx = 1 + \frac{1}{k+1}.$$

So $inf(J[y]) = 1$ but there is no function, y, with $J[y] = 1$ since $J[y] > 1$.[6]

[5]Obviously the problem of finding a maximum for a functional $J[y]$ is the same as finding the minimum for $-J[y]$

[6]Notice that the family of functions $\{x^k\}$ is not closed, nor is it equicontinuous. In particular Ascoli's theorem is not violated.

3 A First Necessary Condition for a Weak Relative Minimum: The Euler-Lagrange Differential Equation

We derive Euler's equation (1744) for a function $y_0(x)$ furnishing a weak relative (improper) extremum for $\int_{x_1}^{x_2} f(x, y, y')dx$.

Definition 3.1 (The variation of a Functional, Gâteaux derivative, or the Directional derivative)**.** Let $J[y]$ be defined on a $\mathcal{D}^n$. Then the first variation of J at $y \in \mathcal{D}^n$ in the direction of $\eta \in \mathcal{D}^n$ (also called the Gâteaux derivative) in the direction of η at y is defined as

$$\lim_{\epsilon \to 0} \frac{J[y + \epsilon\eta] - J[y]}{\epsilon} = \frac{\partial}{\partial\epsilon} J[y + \epsilon\eta]\Big|_{\epsilon=0} \underset{def}{\equiv} \delta J$$

Assume $y_0(x) \in \tilde{C}^1$ furnishes such a minimum. Let $\eta(x) \in C^1$ such that $\eta(x_i) = 0,\ i = 1, 2$. Let $B > 0$ be a bound for $|\eta|, |\eta'|$ on $[x_1, x_2]$. Let $\epsilon_0 > 0$ be given. Imbed $y_0(x)$ in the family $y_\epsilon \underset{def}{\equiv} y_0(x) + \epsilon\eta(x)$. Then $y_\epsilon(x) \in \tilde{C}^1$ and if $\epsilon < \epsilon_0(\frac{1}{B})$, y_ϵ is in the *weak* ϵ_0-neighborhood of $y_0(x)$.

Now $J[y_\epsilon]$ is a real valued function of ϵ with domain $(-\epsilon_0, \epsilon_0)$, hence the fact that y_0 furnishes a weak relative extremum implies

$$\frac{\partial}{\partial\epsilon} J[y_\epsilon]\Big|_{\epsilon=0} = 0$$

We may apply Leibnitz's rule at points where

$f_y(x, y_0(x), y_0'(x))\eta(x) + f_{y'}(x, y_0(x), y_0'(x))\eta'(x)$ is continuous :

$$\frac{d}{d\epsilon} J[y_\epsilon] = \frac{d}{d\epsilon} \int_{x_1}^{x_2} f(x, y_0 + \epsilon\eta, y_0' + \epsilon\eta')dx =$$

$$\int_{x_1}^{x_2} f_y(x, y_0 + \epsilon\eta, y_0' + \epsilon\eta')\eta + f_{y'}(x, y_0 + \epsilon\eta, y_0' + \epsilon\eta')\eta' dx$$

Therefore

(2)

$$\frac{d}{d\epsilon} J[y_\epsilon]\Big|_{\epsilon=0} = \int_{x_1}^{x_2} f_y(x, y_0, y_0')\eta + f_{y'}(x, y_0, y_0')\eta' dx = 0$$

We now use integration by parts, to continue.

$$\int_{x_1}^{x_2} f_y \eta dx = \eta(x)(\int_{x_1}^{x} f_y dx)\Big|_{x_1}^{x_2} - \int_{x_1}^{x_2} (\int_{x_1}^{x} f_y dx)\eta'(x)dx = -\int_{x_1}^{x_2} (\int_{x_1}^{x} f_y dx)\eta'(x)dx$$

Hence

$$\delta J = \int_{x_1}^{x_2} [f_{y'} - \int_{x_1}^{x} f_y dx]\eta'(x)dx = 0$$

for any class C^1 function η such that $\eta(x_i) = 0, i = 1, 2$ and $f(x, y_0(x), y_0'(x))\eta(x)$ continuous. By lemma (2.2) we have proven the following

Theorem 3.2 (Euler-Lagrange, 1744). *If $y_0(x)$ provides an extremum for the functional $J[y] = \int_{x_1}^{x_2} f(x, y, y')dx$ then $f_{y'} - \int_{x_1}^{x} f_y dx = c$ at all points of continuity. i.e. everywhere $y_0(x), y_0'(x)$ are continuous on $[x_1, x_2]$.* [7]

[7] y_0' may have jump discontinuities. That is to say y_0 may be a broken extremal sometimes called a discontinuous solution. Also in order to use Leibnitz's law for differentiating under the integral sign we only needed f to be of class C^1. In fact no assumption about f_x was needed.

Notice that the Gâteaux derivative can be formed without reference to any norm. Hence if it not zero it precludes $y_0(x)$ being a local minimum with respect to any norm.

Corollary 3.3. *At any x where $f_y(x, y_0(x), y_0'(x))$ is continuous*[8] *y_0 satisfies the Euler-Lagrange differential equation*

$$\frac{d}{dx}(f_{y'}) - f_y = 0$$

Proof : At points of continuity of $f_y(x, y_0(x), y_0'(x))$ we have
$f_{y'} = \int_{x_1}^{x} f_y dx + c$. Differentiating both sides proves the corollary.

Solutions of the Euler-Lagrange equation are called *extremals.* It is important to note that the derivative $\frac{d}{dx}f_{y'}$ is not assumed to exist, but is a consequence of theorem (3.2). In fact if $y_0'(x)$ is continuous the Euler-Lagrange equation is a second order differential equation[9] even though y'' may not exist.

Example 3.4. Consider the functional $J[y] = \int_{-1}^{1} y^2(2x - y')^2 dx$ where $y(-1) = 0, \quad y(1) = 1$. The Euler-Lagrange equation is

$$2y(2x - y')^2 = \frac{d}{dx}[-2y^2(2x - y')]$$

The minimum of $J[y]$ is zero and is achieved by the function

$$y_0(x) = \begin{cases} 0, & \text{for } -1 \le x \le 0 \\ x^2, & \text{for } 0 < x \le 1. \end{cases}$$

y_0' is continuous, y_0 satisfies the Euler-Lagrange equation yet $y_0''(0)$ does not exist.

[8]Hence wherever $y_0'(x)$ is continuous.

[9]In the simple case of a functional defined by $\int_{x_1}^{x_2} f(x, y, y')dx$.

A condition guaranteeing that an extremal, $y_0(x)$, is of class C^2 will be given in (4.4).

A curve furnishing a weak relative extremum may not provide an extremum if the collection of allowable curves is enlarged. In particular if one finds an extremum to a variational problem among smooth curves it may be the case that it is not an extremum when compared to all piecewise-smooth curves. The following useful proposition shows that this cannot happen.

Proposition 3.5. *Suppose a smooth curve, $y = y_0(x)$ gives the functional $J[y]$ an extremum in the class of all admissible smooth curves in some weak neighborhood of $y_0(x)$. Then y provides an extremum for $J[y]$ in the class of all admissible piecewise-smooth curves in the same neighborhood.*

Proof : We prove (3.5) for curves in $\mathbb{R}^1$. The case of a curve in $\mathbb{R}^n$ follows by applying the following construction coordinate wise. Assume y_0 is not an extremum among piecewise-smooth curves. We show that curves with 1 corner offer no competition (the case of k corners is similar). Suppose $\tilde{y}$ has a corner at $x = c$ with $\tilde{y}(x_i) = y_0(x_i) \quad i = 1, 2$ and provides a smaller value for the functional than $y_0(x)$ i.e.

$$J[\tilde{y}] = J[y_0] - h, \quad h > 0$$

in a weak ϵ-neighborhood of y_0. We show that there exist $\widehat{y}(x) \in C^1$ in the weak ϵ-neighborhood of $y_0(x)$ with $|J[\widehat{y}] - J[\tilde{y}]| \leq \frac{h}{2}$. Then $\hat{y}$ is a smooth curve with $J[\hat{y}] < J[y_0]$ which is a contradiction. Let $y = \tilde{z}$ be the curve $\tilde{z} = \frac{d\tilde{y}}{dx}$. Then $\tilde{z}$ lies in a 2ϵ neighborhood of the curve $z_0 = \frac{dy_0}{dx}$ (in the sup norm). For a small $\delta > 0$ construct a curve, $\overline{\mathbf{z}}$ from

$(c-\delta, \tilde{z}(c-\delta))$ to $(c+\delta, \tilde{z}(c+\delta))$ in the strip of width 2ϵ about z_0 such that

$$\int_{c-\delta}^{c+\delta} \overline{\mathbf{z}} dx = \int_{c-\delta}^{c+\delta} \tilde{z} dx.$$

Outside $[c-\delta, c+\delta]$ set $\overline{\mathbf{z}} = \frac{d\tilde{y}}{dx}$. Now define $\hat{y}(x) = \tilde{y}(x_1) + \int_{x_1}^{x} \overline{\mathbf{z}} dx$ (see figure 6). If δ is sufficiently small $\hat{y}$ lies in a weak ϵ neighborhood of y_0 and $|J[\tilde{y}] - J[\hat{y}]| \leq \int_{c-\delta}^{c+\delta} |f(x, \tilde{y}.\tilde{y}') - f(x, \hat{y}, \hat{y}')| dx \leq \frac{1}{2} h$.

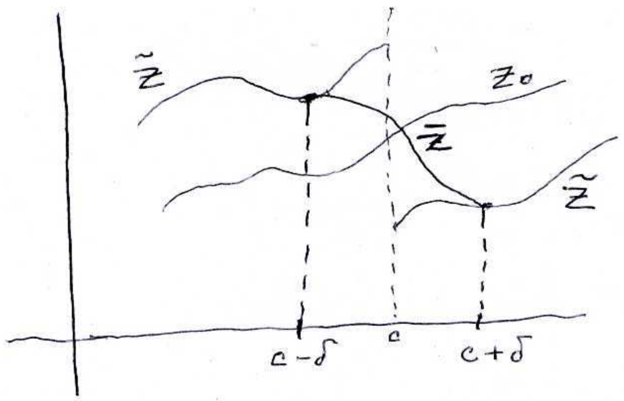


Figure 4: Construction of $\overline{\mathbf{z}}$.

4 Some Consequences of the Euler-Lagrange Equation. The Weierstrass-Erdmann Corner Conditions.

Notice that the Euler-Lagrange equation does not involve f_x. There is a useful formulation which does.

Theorem 4.1. *Let $y_0(x)$ be a $\widetilde{C}^1$ extremal of $\int_{x_1}^{x_2} f(x,y,y')dx$. Then y_0 satisfies*

$$f - y'f_{y'} - \int_{x_1}^{x} f_x dx = c.$$

In particular at points of continuity of y_0' we have

(3)

$$\frac{d}{dx}(f - y'f_{y'}) - f_x = 0$$

- Similar to the remark after (3.3) we are not assuming $\frac{d}{dx}(f - y'f_{y'})$ exist. The derivative exist as a consequence of the theorem.
- If y_0'' exist in $[x_1, x_2]$ then (3) becomes a second order differential equation for y_0.
- If y_0'' exist then (3) is a consequence of Corollary 3.3 as can be seen by expanding (3).

Proof : (of Theorem 4.1) With a a constant, introduce new coordinates (u,v) in the plane by $u = x - ay, v = y$. Hence the extremal curve, C has a parametric representation in the plane with parameter x

$$u = x - ay(x), v = y(x).$$

We have $x = u + av, y = v$. If a is sufficiently small the first of these equations may be solved for x in terms of u.[10]. So for a sufficiently small the extremal may be described in (v, u) coordinates as $v = y(x(u)) = v(u)$. In other words if the new (u, v)-axes make a sufficiently small angle with the (x, y)-axes then equation $y = y_0(x), x_1 \le x \le x_2$ becomes $v = v_0(u), x_1 - ay_1 = u_1 \le u \le u_2 = x_2 - ay_2$ and every curve $v = v(u)$ in a sufficiently small neighborhood of v_0 is the transform of a curve $y(x)$ that lies in a weak neighborhood of $y_0(x)$. The functional $J[y(x)] = \int_{x_1}^{x_2} f(x, y, y')dx$ becomes $J[v(u)] = \int_{u_1}^{u_2} F(u, v(u), \dot{v}(u))du$ where $\dot{v}$ denotes differentiation with respect to u and $F(u, v(u), \dot{v}(u)) = f(u+av(u), v(u), \frac{\dot{v}(u)}{1+a\dot{v}(u)})(1+a\dot{v}(u))$.[11] Therefore the Euler-Lagrange equation applied to $F(u, v_0, \dot{v}_0)$ becomes

$$F_{\dot{v}} - \int_{x_1-ay_1}^{x-ay} F_v du = c$$

but $F_{\dot{v}} = af + \frac{f_{y'}}{1+a\dot{v}}; F_v = (af_x + f_y)(1 + a\dot{v})$. Therefore

$$af + \frac{f_{y'}}{1 + a\dot{v}} - \int_{x_1-ay_1}^{x-ay} (af_x + f_y)(1 + a\dot{v})du = c$$

Substituting back to (x, y) coordinates, noting that $\frac{1}{1+a\dot{v}} = 1 - ay'$ yields

$$af + (1 - ay')f_{y'} - \int_{x_1}^{x} af_x + f_y dx = c.$$

Subtract the Euler-Lagrange equation (3.2) and dividing by a proves the theorem. **q.e.d.**

[10]since $\frac{\partial}{\partial x}[x - ay(x) - u] = 1 - ay'(x)$ and $y \in \mathcal{D}^1$ implies y' bounded on $[x_1, x_2]$.

[11] $\frac{dy}{dx} = \frac{dv}{du} \cdot \frac{du}{dx}$

Applications:

- Suppose $f(x, y, y')$ is independent of y. Then the Euler-Lagrange equation reduces to $\frac{d}{dx} f_{y'} = 0$ or $f_{y'} = c$.

- Suppose $f(x, y, y')$ is independent of x. Then (3) reduces to $f - y' f_{y'} = c$.

Returning to *broken extremals*, i.e. extremals $y_0(x)$ of class $\tilde{C}^1$ where $y_0(x)$ has only a finite number of jump discontinuities in $[x_1, x_2]$. As functions of x

$$\int_{x_1}^{x} f_y(\overline{x}, y_0(\overline{x}), y_0'(\overline{x}) d\overline{x} \text{ and } \int_{x_1}^{x} f_{\overline{x}}(\overline{x}, y_0(\overline{x}), y_0'(\overline{x})) d\overline{x}$$

occurring in (3.2) and (4.1) respectively are continuous in x. Therefore $f_{y'}(x, y_0(x), y_0'(x))$ and $f - y' f_{y'}$ are continuous on $[x_1, x_2]$. In particular at a corner, $c \in [x_1, x_2]$ of $y_0(x)$ this means

Theorem 4.2 (Weierstrass-Erdmann Corner Conditions). [12]

$$f_{y'}(c, y_0(c), y_0'(c^-)) = f_{y'}(c, y_0(c), y_0'(c^+))$$

$$(f - y' f_{y'})\Big|_{c^-} = (f - y' f_{y'})\Big|_{c^+}$$

Thus if there is a corner at c then $f_{y'}(c, y_0(c), z)$ as a function of z assumes the same value for two different values of z. As a consequence of the theorem of the mean there must be a solution to $f_{y'y'}(c, y_0(c), z) = 0$. Therefore we have:

[12] There will be applications in later sections.

Corollary 4.3. *If $f_{y'y'} \neq 0$ then an extremal must be smooth (i.e. cannot have corners).*

The problem of minimizing $J[y] = \int_{x_1}^{x_2} f(x, y, y')dx$ is *non-singular* if $f_{y'y'}$ is > 0 or < 0 throughout a region. Otherwise it is singular. A point (x_0, y_0, y_0') of an extremal $y_0(x)$ of $\int_{x_1}^{x_2} f(x, y, y')dx$ is *non-singular* if $f_{y'y'}(x_0, y_0(x), y_0') \neq 0$. We will show in (4.4) that an extremal must have a continuous second derivative near a non-singular point.

Suppose a variational problem is singular in a region, i.e. $f_{y'y'}(x, y, y') \equiv 0$ in some (x, y, y') region. Then $f_{y'} = N(x, y)$ and $f = M(x, y) + N(x, y)y'$. Hence the Euler-Lagrange equation becomes $N_x - M_y = 0$ for extremals. Note that $J[y] = \int_{x_1}^{x_2} f(x, y, y')dx = \int_{x_1}^{x_2} Mdx + Ndy$. There are two cases:

1. $M_y \equiv N_x$ in a region. Then $J[y]$ is independent of path from $P_1 to P_2$. Therefore the problem with fixed end points has no relevance.

2. $M_y = N_x$ only along a curve $y = y_0(x)$ (i.e. $M_y \neq N_x$ identically) then $y_0(x)$ is the unique extremum for $J[y]$, provided this curve contains P_1, P_2 since it is the only solution of the Euler-Lagrange equation.

Theorem 4.4 (Hilbert)**.** *If $y = y_0(x)$ is an extremal of $\int_{x_1}^{x_2} f(x, y, y')dx$ and (x_0, y_0, y_0') is non-singular (i.e. $f_{y'y'}(x_0, y_0(x), y_0') \neq 0$) then in some neighborhood of x_0 $y_0(x)$ is of class C^2.*

Proof : Consider the implicit equation in x and z

$$f_{y'}(x, y_0(x), z) - \int_{x_1}^{x} f_y(\overline{x}, y_0(\overline{x}), y_0'(\overline{x}))d\overline{x} - c = 0$$

Where c is as in the (3.2). Its locus contains the point $(x = x_0, z = y_0'(x_0))$. By hypothesis the partial derivative with respect to z of the left hand side is not equal to 0 at this point. Hence by the implicit function theorem we can solve for $z(= y')$ in terms of x for x near x_0 and obtain y' of class C^1 near x_0. Hence y'' exists and is continuous. **q.e.d.**

Near a non-singular point, (x_0, y_0, y_0') we can expand the Euler-Lagrange equation and solve for y''

(4)

$$y'' = \frac{f_y - f_{y'x} - y' f_{y'y}}{f_{y'y'}}$$

Near a non-singular point the Euler-Lagrange equation becomes (4). If f is of class C^3 then the right hand side of (4) is a C^1 function of y'. Hence we can apply the existence and uniqueness theorem of differential equations to deduce that there is one and only one class-C^2 extremal satisfying given initial conditions, $(x_0, y_0(x), y_0')$.

5 Some Examples

Shortest Length[Problem I of §1]

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

This is an example of a functional with integrand independent of y (see the first application in §4), i.e. the Euler-Lagrange equation reduces to $f_{y'} = c$ or $\frac{y'}{\sqrt{1+(y')^2}} = c$. Therefore $y' = a$ which is the equation of a straight line.

Notice that any integrand of the form $f(x, y, y') = g(y')$ similarly leads to straight line solutions.

Minimal Surface of Revolution[Problem IV of §1]

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

This is an example of a functional with integrand of the type discussed in the second application in §4. So we have the differential equation:

$$f - y' f_{y'} = y\sqrt{1 + (y')^2} - \frac{y(y')^2}{\sqrt{1 + (y')^2}} = c$$

or

$$\frac{y}{\sqrt{1 + (y')^2}} = c$$

So $\frac{dx}{dy} = \frac{c}{\sqrt{y^2 - c^2}}$. Set $y = c \cosh t$. Then

$$x = \int_{x_0}^{x} \frac{dx}{dy} dy = c \int_{t_0}^{t} \frac{\sinh \bar{t}}{\sinh \bar{t}} d\bar{t} = ct + b.$$

So the solution to the Euler-Lagrange equation associated to the minimal surface of revolution problem is

$$y = c\cosh(\frac{x-b}{c}) \quad \text{satisfying } y_i = c\cosh(\frac{x_i-b}{c}), i = 1, 2$$

This is the graph of a *catenary*. It is possible to find a catenary connecting $P_1 = (x_1, y_1)$ to $P_2 = (x_2, y_2)$ if y_1, y_2 are not too small compared to $x_2 - x_1$ (otherwise we obtain the Goldschmidt discontinuous solution). Note that $f_{y'y'} = \frac{y}{(1+(y')^2)^{\frac{3}{2}}}$. So corners are possible only if $y = 0$ (see corollary 4.3).

The following may help explain the Goldschmidt solution. Consider the problem of minimizing the cost of a taxi ride from point P_1 to point P_2 in the x, y plane. Say the cost of the ride given by the y-coordinate. In particular it is free to ride along the x axis. The functional that minimizes the cost of the ride is exactly the same as the minimum surface of revolution functional. If x_1 is reasonably close to x_2 then the cab will take a path that is given by a catenary. However if the points are far apart in the x direction it seem reasonable the the shortest path is given by driving directly to the x-axis, travel for free until you are are at x_2 then drive up to P_2. This is the Goldschmidt solution.

The Brachistochrone problem[Problem V of §1]

For simplicity, we assume the original point coincides with the origin. Since the velocity of motion along the curve is given by

$$v = \frac{ds}{dt} = \sqrt{1 + (y\prime)^2}\frac{dx}{dt}$$

we have

$$dt = \frac{\sqrt{1+(y')^2}}{v}dx$$

Using the formula for kinetic energy we know that

$$\frac{1}{2}mv^2 = mgy.$$

Substituting into the formula for dt we obtain

$$dt = \frac{\sqrt{1+(y')^2}}{\sqrt{2gy}}dx.$$

So the time it takes to travel along the curve, y, is given by the the functional $J[y] = \int_{x_1}^{x_2} \sqrt{\frac{1+(y')^2}{y}}dx$. $J[y]$ does not involve x so we have to solve the first order differential equation (where we write the constant as $\frac{1}{\sqrt{2b}}$)

(5)

$$f - y'f_{y'} = \frac{1}{\sqrt{y}\sqrt{1+(y')^2}} = \frac{1}{\sqrt{2b}}$$

We also have

$$f_{y'y'} = \frac{1}{\sqrt{y}(1+(y')^2)^{\frac{3}{2}}}.$$

So for the brachistochrone problem a minimizing arc can have no corners i.e. it is of class C^2. Equation (5) may be solved by separating the variables, but to quote [B62] "it is easier if we profit by the experience of others" and introduce a new variable u defined by the equation

$$y' = -\tan\frac{u}{2} = -\frac{\sin u}{1+\cos u}$$

Then $y = b(1 + \cos u)$. Now

$$\frac{dx}{du} = \frac{dx}{dy}\frac{dy}{du} = (-\frac{1 + \cos u}{\sin u})(-b \sin u) = b(1 + \cos u).$$

Therefore we have a parametric solution of the Euler-Lagrange equation for the brachistochrone problem:

$$x - a = b(u + \sin u)$$

$$y = b(1 + \cos u)$$

for some constants a and b[13]. This is the equation of a *cycloid*. Recall this is the locus of a point fixed on the circumference of a circle of radius b as the circle rolls on the lower side of the x-axis. One may show that there is one and only one cycloid passing through P_1 and P_2. An (approximate) quotation from Bliss:

> The fact that the curve of quickest descent must be a cycloid is the famous result discovered by James and John Bernoulli in 1697. The cycloid has a number of remarkable properties and was much studied in the 17th century. One interesting fact: If the final position of descent is at the lowest point on the cycloid, then the time of descent of a particle starting at rest is the same no matter what the position of the starting point on the cycloid may be.

That the cycloid should be the solution of the brachistochrone problem was regarded with wonder and admiration by the Bernoulli brothers.

[13] u is an angle of rotation, not time.

The Bernoulli's did not have the Euler-Lagrange equation (which was discovered about 60 years later). Johann Bernoulli derived equation (5) by a very clever reduction of the problem to Snell's law in optics (which is derived in most calculus classes as a consequence of Fermat's principle that light travels a path of least time). Recall that if light travels from a point (x_1, y_1) in a homogenous medium M_1 to a point (x_2, y_2) $(x_1 < x_2)$ in a homogeneous medium M_2 which is separated from M_1 by the line $y = y_0$. Suppose the respective light velocities in the two media are u_1 and u_2. Then the point of intersection of the light with the line $y = y_0$ is characterized by $\frac{\sin a_1}{u_1} = \frac{\sin a_2}{u_2}$ (see the diagram below).

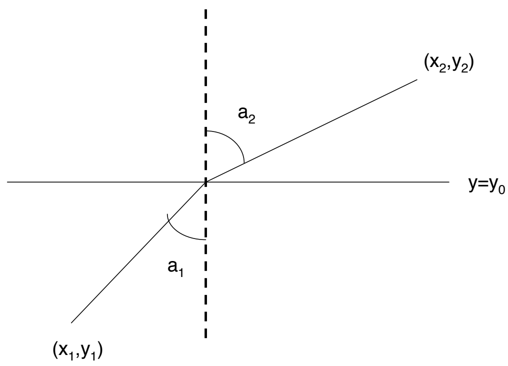


Figure 5: Snell's law

We now consider a single optically inhomogeneous medium M in which the light velocity is given by a function, $u(y)$. If we take $u(y) = \sqrt{2gy}$ then the path light will travel is the same as that taken by a particle that minimizes the travel time. We approximate M by a sequence of parallel-faced homogeneous media, $M_1, M_2, \cdots$ with M_i having constant velocity given by $u(y_i)$ for some y_i in the i-th band. Now by Snell's law we have

$$\frac{\sin\phi_1}{u_1} = \frac{\sin\phi_2}{u_3} = \cdots$$

i.e.

$$\frac{\sin\phi_i}{u_i} = C$$

where C is a constant.

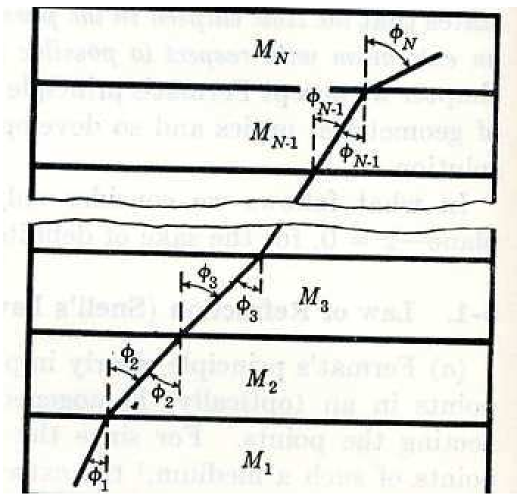


Figure 6: An Approximation to the Light Path

Now take a limit as the number of bands approaches infinity with the maximal width approaching zero. We have

$$\frac{\sin\phi}{u} = C$$

where ϕ is as in figure 7, below.

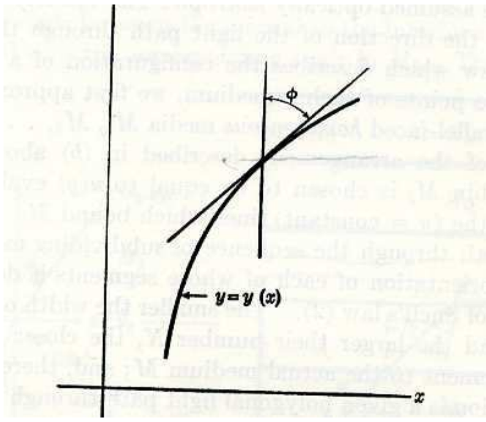


Figure 7:

In particular $y'(x) = \cot\phi$, so that

$$\sin\phi = \frac{1}{\sqrt{1+(y')^2}}.$$

So we have deduced that the light ray will follow a path that satisfies the differential equation

$$\frac{1}{u\sqrt{1+(y')^2}} = C.$$

If the velocity, u, is given by $\sqrt{2gy}$ we obtain equation (5).

6 Extension of the Euler-Lagrange Equation to a Vector Function, Y(x)

We extend the previous results to the case of

$$J[Y] = \int_{x_1}^{x_2} f(x, Y, Y')ds, \quad Y(x) = (y_1(x), \cdots, y_n(x)), \quad Y' = (y_1'(x), \cdots, y_n'(x))$$

For the most part the generalizations are straight forward.

Let $\mathcal{M}$ be continuous vector valued functions which have continuous derivatives on $[x_1, x_2]$ except, perhaps, at a finite number of points. A *strong* ϵ-neighborhood of $Y_0(x) \in \mathcal{M}$ is an ℓ^2 epsilon neighborhood. i.e.

$$\{Y(x) \in \mathcal{M} | \quad |Y(x) - Y_0(x)| = [\sum_{i=1}^{n} (y_i(x) - y_{i_0}(x))^2]^{\frac{1}{2}} < \epsilon\}$$

For a *weak* ϵ-neighborhood add the condition

$$|Y'(x) - Y_0'(x)| < \epsilon.$$

The first variation of $J[Y]$, or the directional derivative in the direction of H is given by

$$\delta J[Y_0] = \frac{dJ[Y_0 + \epsilon H]}{d\epsilon}\Big|_{\epsilon=0}$$

For H satisfying $H(x_i) = (0, \cdots, 0), i = 1, 2 \quad \delta J[Y_0] = 0$ is a necessary condition for $Y_0(x)$ to furnish a local minimum for $J[Y]$. By choosing $H = (0, \cdots, \eta(x), \cdots, 0)$ we may apply Lemma 2.2 to obtain n- Euler-Lagrange equations:

(6)

$$f_{y'_j} - \int_{x_1}^{x_2} f_{y_j} d\overline{x} = c_j$$

and whenever $y'_j \in C$ this yields:

(7)

$$\frac{df_{y'_j}}{dx} - f_{y_j} = 0.$$

To obtain the vector version of Theorem 4.1 we use the coordinate transformation

$$u = x - \alpha y_1, \quad v_i = y_i, \quad i = 1, \cdots, n$$

or

$$x = u + \alpha v_1, \quad , y_i = v_i \quad i = 1, \cdots, n.$$

For α small we may solve $u = x - \alpha y_1(x)$ for x in terms of u and the remaining n equations become $V = (v_1(u), \cdots, v_n(u))$ (e.g. $v_n = y_n(x(u)) = v_n(u)$)

For a curve $Y(x)$ the functional $J[Y]$ is transformed to

$$\widetilde{J}[V(u)] = \int_{x_1 - \alpha y_1(x_1)}^{x_2 - \alpha y_1(x_2)} f(u + \alpha v_1, V(u), \frac{\dot{V}}{1 + \alpha \dot{v}_1})(1 + \alpha \dot{v}_1) du$$

If Y provided a local extremum for J, $V(u)$ must provide a local extremum for $\widetilde{J}$ (for small α). We have the Euler-Lagrange equation for $j = 1$

$$\alpha f + f_{y'_1} - \frac{\alpha}{1 + \alpha \dot{v}_1} \sum_{i=1}^{n} f_{y'_i} \dot{v}_i - \int_{u_1 = x_1 - \alpha y_1(x_1)}^{u = x - \alpha x} (\alpha f_x + f_{y_1}) \underbrace{(1 + \alpha \dot{v}_1) du}_{d\overline{x}} = c.$$

Subtracting the case $j = 1$ of (6), dividing by α and using $\frac{\dot{v}_i}{1+\alpha\dot{v}_1} = \frac{dv_i}{du}\frac{du}{dx} = \frac{dy_i}{dx}$ yields the vector form of Theorem 4.1

Theorem 6.1. $f - \sum\limits_{i=1}^{n} y_i' f_{y_i'} - \int_{x_1}^{x} f_{\overline{x}} d\overline{x} = c$

The same argument as for the real valued function case, we have the $n+1$ *Weierstrass-Erdmann* corner conditions at any point x_0 of discontinuity of Y_0' ($Y_0(x)$ an extremal):

$$f_{y_i'}\Big|_{x_0^-} = f_{y_i'}\Big|_{x_0^+}, \quad i = 1, \cdots, n$$

and

$$f - \sum_{i=1}^{n} y_i' f_{y_i'}\Big|_{x_0^-} = f - \sum_{i=1}^{n} y_i' f_{y_i'}\Big|_{x_0^+}$$

Definition 6.2. $(x, Y(x), Y'(x))$ in *non-singular* for f iff

$$det\left[f_{y_i' y_j'}\right]_{n,n} \neq 0 \quad \text{at } (x, Y(x), Y'(x))$$

Theorem 6.3 (Hilbert)**.** *If $(x_0, Y(x_0), Y'(x_0))$ is a non-singular element of the extremal Y for the functional $J[Y] = \int_{x_1}^{x_2} f(x, Y, Y')dx$ then in some neighborhood of x_0 the extremal curve of of class C^2.*

Proof : Consider the system of implicit equations in x and $Z = (z_1, \cdots, z_n)$

$$f_{y_j'}(x, Y_0(x), Z) - \int_{x_1}^{x} f_{y_i}(\overline{x}, Y_0(\overline{x}), Y'(\overline{x}))d\overline{x} - c_j = 0, \quad j = 1, \cdots, n$$

where c_j are constants. Its locus, contains the point $(x_0, Z_0 = (y_{0_i}'(x), \cdots, y_{0_n}))$. By hypothesis the Jacobian of the system with respect to $(z_1, \cdots, z_n)$ is not 0. Hence we obtain a solution $(z_1(x), \cdots, z_n(x))$ for Z in terms of x. viz $Y_0'(x)$ which near x_0 is of class C^1 (as are $f_{y'}$ and $\int_{x_1}^{x} f_y d\overline{x}$ in x.) Hence Y_0'' is continuous near X_0. **q.e.d.**

Hence near a non-singular element $(x_0, Y_0(x_0), Y_0'(x_0)$ we can expand (7) to obtain a system of n linear equations in $y_{0_1}'', \cdots, y_{0_n}''$

$$f_{y_j' x} + \sum_{k=1}^{n} f_{y_j' y_k} \cdot y_{0_k}'(x) + \sum_{k=1}^{n} f_{y_j' y_k'} y_{0_k}'' - f_{y_j} = 0, \quad (j = 1, \cdots, n).$$

The system may be solved near x_0

$$y_{0_i}''(x) = \sum_{k=1}^{n} \frac{P_{ik}(x, Y_0, Y_0')}{det\left[f_{y_i' y_j'}\right]} y_{0_k}' + \sum_{k=1}^{n} \frac{Q_{ik}(x, Y_0, Y_0')}{det\left[f_{y_i' y_j'}\right]} (f_{y_k} - f_{y_k' x}) \quad i = 1, \cdots, n$$

where P_{ik}, Q_{ik} are polynomials in the second order partial derivatives of $f(x, Y_0, Y_0')$. From this we get the existence and uniqueness of class C^1 extremals near a non-singular element.

7 Euler's Condition for Problems in Parametric Form (Euler-Weierstrass Theory)

We now study *curve functionals* $J[C] = \int_{t_1}^{t_2} F(t, x(t), y(t), \dot{x}(t), \dot{y}(t))dt$.

Definition 7.1. A curve C is *regular* if if $x(t), y(t)$ are C^1 functions and $\dot{x}(t) + \dot{y}(t) \neq 0$.

For example there is the arc length functional, $\int_{t_1}^{t_2} \sqrt{\dot{x}(t) + \dot{y}(t)}dt$ where C is parametrized by $x(t), y(t)$ with $\dot{x}(t) + \dot{y}(t) \neq 0$.

To continue we require the following differentiability and invariance conditions on the base function F :

(8)

(a) F is of class C^3 for all t, all (x, y) in some region, R and all $(\dot{x}, \dot{y}) \neq (0, 0)$.

(b) Let $\mathfrak{M}$ be the set of all regular curves such that $(x(t), y(t)) \in$ R for all $t[t_1, t_2]$. Then if C is also parameterized by $\tau \in [\tau_1, \tau_2]$ with change of parameter function $t = \varphi(\tau), \quad \tau' > 0$[14], we require that

$$J[C] = \int_{t_1}^{t_2} F(t, x(t), y(t), \dot{x}(t), \dot{y}(t))dt = \int_{\tau_1}^{\tau_2} F(\tau, \tilde{x}(\tau), \tilde{y}(\tau), \dot{\tilde{x}}(\tau), \dot{\tilde{y}}(\tau))d\tau$$

Part (b) simply states that the functional we wish to minimize depends only on the curve, not on an admissible reparametrization of C.

[14]Such reparametrizations are called admissible.

Some consequences of 8 part (b)

$$\int_{t_1}^{t_2} F(t, x(t), y(t), \dot{x}(t), \dot{y}(t))dt =$$

$$\int_{\tau_1}^{\tau_2} F(\varphi(\tau), \tilde{x}(\tau), \tilde{y}(\tau), \frac{\dot{\tilde{x}}(\tau)}{\varphi'(\tau)}, \frac{\dot{\tilde{y}}(\tau)}{\varphi'(\tau)})\varphi'(\tau)d\tau = \int_{\tau_1}^{\tau_2} F(\tau, \tilde{x}(\tau), \tilde{y}(\tau), \dot{\tilde{x}}(\tau), \dot{\tilde{y}}(\tau))d\tau$$

The last equality may be considered an identity in τ_2. Take $\frac{d}{d\tau_2}$ of both sides. Therefore for any admissible $\varphi(\tau)$ we have

(9)

$$F(\varphi(\tau), \tilde{x}(\tau), \tilde{y}(\tau), \frac{\dot{\tilde{x}}(\tau)}{\varphi'(\tau)}, \frac{\dot{\tilde{y}}(\tau)}{\varphi'(\tau)})\varphi'(\tau) = F(\tau, \tilde{x}(\tau), \tilde{y}(\tau), \dot{\tilde{x}}(\tau), \dot{\tilde{y}}(\tau))$$

We now consider some special cases:

1. Let $\varphi(\tau) = \tau + c$: From 9 we have F depends only on $x, y, \dot{x}, \dot{y}$.

2. Let $\varphi\tau = k\tau, \quad k > 0$: From 9 we have

$$F(\tilde{x}, \tilde{y}, \frac{1}{k}\dot{\tilde{x}}, \frac{1}{k}\dot{\tilde{y}})k = F(\tilde{x}, \tilde{y}, \dot{\tilde{x}}, \dot{\tilde{y}})$$

   i.e. F is *homogeneous of degree 1* in $\dot{x}, \dot{y}$. One may show that these two properties suffice for the required invariance in (8) (b).

3. $F(x, y, k\dot{x}, k\dot{y}) \equiv kF(x, y, \dot{x}, \dot{y}), \quad k > 0$ implies

(10)

$$\dot{x}F_{\dot{x}} + \dot{y}F_{\dot{y}} \equiv F.$$

Differentiating this identity, first with respect to $\dot{x}$ then with respect to $\dot{y}$ yields;

$$\dot{x}F_{\dot{x}\dot{x}} + \dot{y}F_{\dot{x}\dot{y}} = 0, \quad \dot{x}F_{\dot{x}\dot{y}} + \dot{y}F_{\dot{y}\dot{y}} = 0.$$

Hence there exist a function $F_1(x, y, \dot{x}, \dot{y})$ such that

$$F_{\dot{x}\dot{x}} = \dot{y}^2 F_1, \quad F_{\dot{x}\dot{y}} = -\dot{x}\dot{y}F_1, \quad F_{\dot{y}\dot{y}} = \dot{x}^2 F_1.$$

Notice that F_1 is homogenous of degree -3 in $\dot{x}, \dot{y}$. For example for the arc length functional, $F = \sqrt{\dot{x}^2 + \dot{y}^2}$, $F_1 = \frac{1}{(\dot{x}^2+\dot{y}^2)^{\frac{3}{2}}}$

Definition 7.2. A *strong* ϵ-neighborhood of a curve, $C = (x(t), y(t)), \quad t \in [t_1, t_2]$ is given by

$$\{C = (\tilde{x}(t), \tilde{y}(t)) \quad | \quad (x_1(t) - x(t))^2 + (y_1(t) - y(t))^2 < \epsilon^2 \quad t \in [t_1, t_2]\}$$

The notion of a weak ϵ-neighborhood as defined in (2.5) depends on the parametrization, not on the curve. The following extra condition in the following definition remedies this:

Definition 7.3. $C_1 = (x_1(t), y_1(t))$ is in a *weak ϵ-neighborhood* of C if in addition to (7.2) we have $\sqrt{(x_1')^2 + (y_1')^2}$ is bounded by some fixed M[15] and

$$(x_1'(t) - x'(t))^2 + (y_1'(t) - y'(t))^2 < \epsilon^2 \sqrt{(x_1')^2 + (y_1')^2}\sqrt{(x')^2 + (y')^2}$$

[15]M depends on the parametrization, but the existence of an M does not.

(7.3) implies a weak neighborhood in the sense of (2.5). To see this first note that the set of functions $\{(x_1'(t), y_1'(t)\}$ in a weak ϵ-neighborhood of $(x(t), y(t))$ as defined in (2.5) are uniformly bounded on the interval $[t_1, t_2]$. Hence it is reasonable to have assumed that the set of functions in a weak ϵ-neighborhood, $\mathfrak{N}$ in the sense of (7.3) are uniformly bounded. It then follows that $\mathfrak{N}$ is a weak neighborhood in the sense of (2.5) (for a different ϵ).

The same proof of the Euler-Lagrange -equation in §6 leads to the necessary condition for a weak relative minimum:

(11)

$$F_{\dot{x}} - \int_{t_1}^{t} F_x d\bar{t} = a, \quad F_{\dot{y}} - \int_{t_1}^{t} F_y d\bar{t} = b$$

for constants a, b.

All we need to observe is that for $\tilde{\epsilon}$ sufficiently small, $(x(t), y(t)) + \tilde{\epsilon}\eta(t)$ lies in a weak ϵ-neighborhood of $(x(t), y(t))$.

Also (6.1) immediately generalizes to:

$$F - \dot{x}F_{\dot{x}} - \dot{y}F_{\dot{y}} - \int_{t_1}^{t} \underbrace{F_{\bar{t}}}_{=0} d\bar{t} = c$$

But this is jut (10). Hence the constant is 0.

Continuing, as before:

- If $\dot{x}(t)$ and $\dot{y}(t)$ are continuous we may differentiate (11) :

(12)

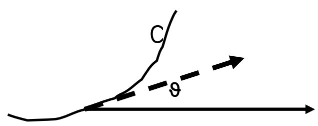
$$\frac{d}{dt}(F_{\dot{x}}) - F_x = 0, \quad \frac{d}{dt}(F_{\dot{y}}) - F_y = 0$$

- The Weirstrass-Erdmann corner condition is proven in the same way as in (4.2). At corner, t_0

$$F_{\dot{x}}\Big|_{t_0^-} = F_{\dot{x}}\Big|_{t_0^+}; \quad F_{\dot{y}}\Big|_{t_0^-} = F_{\dot{y}}\Big|_{t_0^+}$$

Note:

1. The term *extremal* is usually used only for smooth solutions of (11).
2. If t is arc length, then $\dot{x} = \cos\theta, \quad \dot{y} = \sin\theta$ where θ is the angle the tangent vector makes with the positive



A functional is *regular* at (x_0, y_0) if and only if $F_1(x_0, y_0, \cos\theta, \sin\theta) \neq 0$ for all θ; *quasi-regular* if $F_1(x_0, y_0, \cos\theta, \sin\theta) \geq 0$ (or ≤ 0) for all θ.

Whenever $\ddot{x}(t), \ddot{y}(t)$ exist we may expand (12) which leads to:

(13)

$$F_{\dot{x}x}\dot{x} + F_{\dot{x}y}\dot{y} + F_{\dot{x}\dot{x}}\ddot{x} + F_{\dot{x}\dot{y}}\ddot{y} = F_x$$

$$F_{\dot{y}x}\dot{x} + F_{\dot{y}y}\dot{y} + F_{\dot{y}\dot{x}}\ddot{x} + F_{\dot{y}\dot{y}}\ddot{y} = F_y$$

In fact these are not independent. Using the equations after (10) which define F_1 and

$$F_{\dot{x}x}\dot{x} + F_{\dot{y}x}\dot{y} = F_x$$

which follows from (10) we can rewrite (13) as:

$$\dot{y}[F_{\dot{x}y} - F_{x\dot{y}} + F_1(\ddot{x}\dot{y} - \ddot{y}\dot{x})] = 0$$

and

$$\dot{x}[F_{\dot{x}y} - F_{x\dot{y}} + F_1(\ddot{x}\dot{y} - \ddot{y}\dot{x})] = 0.$$

Since $(\dot{x})^2 + (\dot{y})^2 \neq 0$ we have

Theorem 7.4. *[Euler-Weirstrass equation for an extremal in parametric form.]*

$$F_{\dot{x}y} - F_{x\dot{y}} + F_1(\ddot{x}\dot{y} - \ddot{y}\dot{x}) = 0$$

or

$$\frac{F_{\dot{x}y} - F_{x\dot{y}}}{F_1((\dot{x})^2 + (\dot{y})^2)^{\frac{3}{2}}} = \frac{\dot{x}\ddot{y} - \ddot{x}\dot{y}}{((\dot{x})^2 + (\dot{y})^2)^{\frac{3}{2}}} = \mathcal{K}(= curvature^{16})$$

[16] See, for example, [O] Theorem 4.3.

Notice that the left side of the equation is homogeneous of degree zero in $\dot{x}$ and $\dot{y}$. Hence the Euler-Weirstrass equation will not change if the parameter is changed.

We now consider curves Parametrized by arc length, s. An *element* for a functional $J[C] = \int_C F(x(s), y(s), \dot{x}(s), \dot{y}(x))ds$ is a point $X_0 = (x_0, y_0, \dot{x}_0, \dot{y}_0)$ with $||\frac{dC}{ds}_0||^2 = (\dot{x}_0)^2 + (\dot{y}_0)^2 = 1$. An element is *non-singular* if and only if $F_1(X_0) \neq 0$.

Theorem 7.5 (Hilbert)**.** *Let $(x_0(s), y_0(s))$ be an extremal for $J[C]$.*

Assume $(x_0(s_0), y_0(s_0), \dot{x}_0(s), \dot{y}_0(s_0))$ is non-singular. Then in some neighborhood of s_0 $(x_0(s), y_0(s))$ is of class C^2.

Proof : Since $(x_0(s), y_0(s))$ is an extremal, $\dot{x}$ and $\dot{y}$ are continuous. Hence the system

$$\begin{cases} F_{\dot{x}}(x_0(s), y_0(s), u, v) - \lambda u = \int_0^s F_x(x_0(\overline{s}), y_0(\overline{s}), \dot{x}_0(\overline{s}), \dot{y}_0(\overline{s}))d\overline{s} + a \\ F_{\dot{y}}(x_0(s), y_0(s), u, v) - \lambda v = \int_0^s F_y(x_0(\overline{s}), y_0(\overline{s}), \dot{x}_0(\overline{s}), \dot{y}_0(\overline{s}))d\overline{s} + b \\ u^2 + v^2 = 1 \end{cases}$$

(where a, b are as in (11)) define implicit functions of u, v, λ which have continuous partial derivatives with respect to s. Also for any $s \in (0, \ell)$ ($\ell =$ the length of C) the system has a solution

$$u = \dot{x}_0, v = \dot{y}_0, \lambda = 0$$

by (11). The Jacobian of the left side of the above system at $(\dot{x}_0(s), \dot{y}_0(s), 0, s) = 2F_1 \neq 0$. Hence by the implicit function theorem the solution $(\dot{x}_0(s), \dot{y}_0(s), 0)$ is of class C^1 near s_0. **q.e.d.**

Theorem 7.6. *Suppose an element,* (x_0, y_0, θ_0) *is non-singular*

(i.e. $F_1(x_0, y_0, \cos\theta_0, \sin\theta_0) \neq 0$*) and such that* $(x_0, y_0) \in \mathcal{R}$*, the region in which* F*is* C^3*. Then there exist a unique extremal* $(x_0(s), y_0(s))$ *through* (x_0, y_0, θ_0) *(i.e. such that* $x_0(s_0) = x_0, y_0(s_0) = y_0, \dot{x}_0(s_0) = \cos\theta_0, \dot{y}_0(s_0) = \sin\theta_0$*).*

Proof : The differential equations for the extremal are (7.4) and $\dot{x}^2 + \dot{y}^2 = 1$ ($\Rightarrow \dot{x}\ddot{x} + \dot{y}\ddot{y} = 0.$) Near $x_0, y_0, \dot{x}_0, \dot{y}_0$ this system can be solved for $\ddot{x}, \ddot{y}$ since the determinant of the left hand side of the system

$$\begin{cases} F_1(\ddot{x}\dot{y} - \ddot{y}\dot{x}) = F_{x\dot{y}} - F_{\dot{x}y} \\ \dot{x}\ddot{x} + \dot{y}\ddot{y} = 0 \end{cases}$$

is

$$\begin{vmatrix} F_1\dot{y} & -F_1\dot{x} \\ \dot{x} & \dot{y} \end{vmatrix} = (\dot{x}^2 + \dot{y}^2)F_1 = F_1 \neq 0.$$

The resulting system is of the form

$$\begin{cases} \ddot{x} = g(x, y, \dot{x}, \dot{y}) \\ \ddot{y} = h(x, y, \dot{x}, \dot{y}) \end{cases}$$

Now apply the Picard existence and uniqueness theorem.

8 Some More Examples

1. **<u>On Parametric and Non-Parametric Problems in the Plane</u>:**

(a) In a parametric problem, $J[C] = \int_C F(x, y, \dot{x}, \dot{y})dt$, we may convert J to a non-parametric form only for those curves that are representable by class $\widetilde{C}^1$ functions $y = y(x)$. For such a curve the available parametrization, $x(t) = t, y = y(t)$ converts $J[C]$ into

$\int_{x_1}^{x_2} \underbrace{F(x, y(x), 1, y'(x))}_{=f(x,y,y')} dx = \int_{x_1}^{x_2} f(x, y, y')dx.$

(b) In the reverse direction, parametrization of a non-parametric problem, $J[y] = \int_{x_1}^{x_2} f(x, y(x), y'(x))dx$, may enlarge the class of competing curves and thus change the problem. This is not always the case. For example if we change $J[y] = \int_{x_1}^{x_2} (y')^2 dx$, to a parametric problem by $x = \varphi(t)$ with $\dot{\varphi}(t) > 0$ the problem changes to maximizing $\int_{t_1}^{t_2} \frac{\dot{y}^2}{\dot{x}^2} dt$ among all curves. But this still excludes curves with $\dot{x}(t) = 0$, hence the curve cannot "double back".

Now consider the problem of maximizing $J[y] = \int_0^2 \frac{dx}{1+(y')^2}$ subject to $y(0) = y(2) = 0$. Clearly $y \equiv 0$ maximizes $J[y]$ with maximum $= 2$. But if we view this as a parametric problem we are led to $J[C] = \int_C \frac{\dot{x}^3 dt}{\dot{x}^2 + \dot{y}^2}$. Consider the following curve, C_h.

$J[C_h] = 2 + \frac{h}{2(2h^2+2h+1)} > 2$. This broken extremal (or discontinuous solution) can be approximated by a smooth curve, $\overline{C}_h$ for which $J[\overline{C}_h] > 2$.

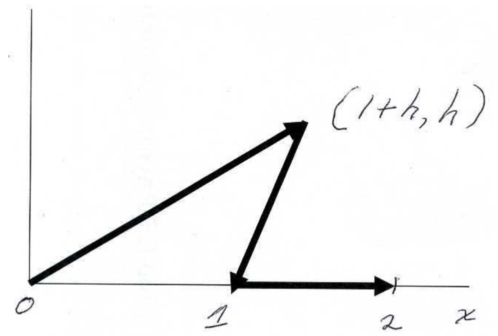


Figure 8: C_h

2. **The Arc Length Functional:** For a curve, ω, $J[\omega] = \int_\omega ||\dot{\omega}||dt$. $F_x = F_y = 0, F_{\dot{x}} = \frac{\dot{x}}{||\omega||}, F_{\dot{y}} = \frac{\dot{y}}{||\dot{\omega}||}, F_1 = \frac{1}{||\dot{\omega}||^3} \neq 0$. So this is a regular problem (i.e. only smooth extremals). Equations, (11), are $\frac{\dot{x}}{||\omega||} = a$, $\frac{\dot{y}}{||\omega||} = b$ or $\frac{\dot{x}}{\dot{y}}$ is a constant, i.e. a straight line!

Alternately the Euler-Weierstrass equation, (7.4) gives $\ddot{x}\dot{y} - \dot{x}\ddot{y} = 0$, again implying $\frac{\dot{x}}{\dot{y}}$ is a constant.

Yet a third, more direct method is to arrange the axes such that C has end points $(0,0)$ and $(\ell, 0)$. Then $J[C] = \int_{t_1}^{t_2} \sqrt{\dot{x}^2 + \dot{y}^2}dt \geq \int_{t_1}^{t_2} \sqrt{\dot{x}^2}dt \geq \int_{t_1}^{t_2} \dot{x}dt = x(t_2) - x(t_1) = \ell$. The first $\geq$ is $=$ if and only if $\dot{y} \equiv 0$, or $y(t) = 0$ (since $y(0) = 0$). The second $\geq$ is $=$ if and only if $\dot{x} > 0$ for all t (i.e. no doubling back). Hence only the straight line segment gives $J[C]_{\text{min}} = \ell$.

3. **Geodesics on a Sphere:** Choose coordinates on the unit sphere, φ (latitude) and θ (longitude). Then for a curve $C : (\varphi(t), \theta(t))$, $ds = \sqrt{\dot{\varphi}^2 + (\cos^2(\varphi))\dot{\theta}^2}$. Therefore the arc length is given by $J[C] = \int_{t_1}^{t_2} \sqrt{\dot{\varphi}^2 + (\cos^2(\varphi))\dot{\theta}^2}dt$.

1. $F_\varphi = \frac{-\dot\theta^2 \sin\varphi\cos\varphi}{\sqrt{\dot\varphi^2+(\cos^2(\varphi))\dot\theta^2}}$

2. $F_\theta = 0$

3. $F_{\dot\varphi} = \frac{\dot\varphi}{\sqrt{\dot\varphi^2+(\cos^2(\varphi))\dot\theta^2}}$

4. $F_{\dot\theta} = \frac{\dot\theta\cos^2\varphi}{\sqrt{\dot\varphi^2+(\cos^2(\varphi))\dot\theta^2}}$

5. $F_1 = \frac{\cos^2\varphi}{(\dot\varphi^2+(\cos^2(\varphi))\dot\theta^2)^{\frac{3}{2}}}$

So $F_1 \neq 0$ except at the poles (and only due to the coordinate system chosen). Notice that the existence of many minimizing curves from pole to pole is consistent with (7.6).

If one develops the necessary conditions, (11) and (7.4) one is led to unmanageable differential equations[17]. The direct method used for arc length in the plane works for geodesics on the sphere. Namely choose axes such that the two end points of C have the same longitude, θ_0. Then $J[C] = \int_{t_1}^{t_2}\sqrt{\dot\varphi^2+(\cos^2(\varphi))\dot\theta^2}dt \geq \int_{t_1}^{t_2}\sqrt{\dot\varphi^2}dt \geq \int_{t_1}^{t_2}\dot\varphi dt = \varphi(t_2)-\varphi(t_1)$. Hence reasoning as above it follows that the shortest of the two arcs of the meridian $\theta=\theta_0$ joining the end points is the unique shortest curve.

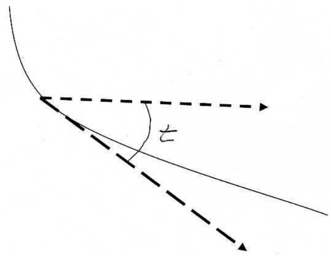
4. <u>Brachistochrone revisited</u> $J[C] = \int_{x_1}^{x_2}\sqrt{\frac{1+(y')^2}{y}}dx = \int_{t_1}^{t_2}\sqrt{\frac{\dot x^2+\dot y^2}{y}}dt$. Here $F_x = 0$, $F_{\dot x} = \frac{\dot x}{\sqrt{y(\dot x^2+\dot y^2)}}$, $F_1 = \frac{1}{\sqrt{y}(\dot x^2+\dot y^2)^{\frac{3}{2}}}$ (so it is a regular problem). (11) reduces to the equation

$$\frac{\dot x}{\sqrt{y(\dot x^2+\dot y^2)}} = a \neq 0^{18}$$

[17] See the analysis of geodesics on a surface, part (5) below.

[18] Unless the two endpoints are lined up vertically.

As parameter, t choose the angle between the vector $\overrightarrow{(1,0)}$ and the tangent to C (as in the following figure).



Therefore $\frac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2}} = \cos t$, and $\cos t = a\sqrt{y}$ or $y = \frac{1}{2a^2}(1+\cos 2t)$. Hence

$$\dot{x}^2 \sin^2 t = \dot{y}^2 \cos^2 t = \frac{4}{a^4}\sin^2 t \cos^4 t, \quad \dot{x} = \pm\frac{2}{a^2}\cos^2 t = \pm\frac{1}{a^2}(1+\cos 2t).$$

Finally we have

$$x - c = \pm\frac{1}{2a^2}(2t + \sin 2t)$$
$$y = \frac{1}{2a^2}(1+\cos 2t)$$

This is the equation of a cycloid with base line $y = 0$. The radius of the wheel is $\frac{1}{\sqrt{2}a}$ with two degrees of freedom (i.e. the constants a, c) to allow a solution between any two points not lined up vertically.

5. <u>Geodesics on a Surface</u> We use the Euler Weierstrass equations, (12), to show that if C is a curve of shortest length on a surface then $\kappa_g = 0$ along C where κ_g is the geodesic curvature. In [M] and [O] the vanishing of κ_g is taken to be the definition of a geodesic. It is then proven that there is a geodesic connecting any two points on a surface and the curve of shortest distance is given by a geodesic.[19] We will prove that the solutions to the Euler Weierstrass equations for the minimal distance on a surface functional have $\kappa_g = 0$. We have already shown that there exist solutions between any two points. Both (7.6) and Milnor [M] use the existence theorem from differential equations.

A quick trip through differential geometry

The reference for much of this is Milnor's book, [M, Section 8]. The surface, M (which for simplicity we view as sitting in $\mathbb{R}^3$), is given locally by coordinates $\overline{u} = (u^1, u^2)$ i.e. there is a diffeomorphism, $\overline{x} = (x^1(\overline{u}, x^2(\overline{u}), x^3(\overline{u})) : \mathbb{R}^2 \to M$. There is the notion of a covariant derivative, or connection on M (denoted $\nabla_X Y$, with X and Y smooth vector fields on M). $\nabla_X Y$ is a smooth vector field with the following properties:

1. $\nabla_X Y$ is bilinear as a function of X and of Y.

2. $\nabla_{fX} Y = f(\nabla_X Y)$.

3. $\nabla_X (fY) = (Xf)Y + f(\nabla_X Y),$ the derivation law .

In particular, since $\{\partial_i\}$ are a basis for T_{M_p}, the connection is determined by $\nabla_{\partial_i}\partial_j$. We define the Christoffel symbols by the identity:

$$\nabla_{\partial_i}\partial_j = \sum_k \Gamma^k_{ij}\partial_k (= \Gamma^k_{ij}\partial_k \text{ using the Einstein summation convention}).$$

[19]There may be many geodesics connecting two points. The shortest distance is achieved by at least one of them. e.g. the great circles on a sphere.

If C is a curve in M given by functions, $\{u^i(t)\}$, and V is any vector field along C, then we may define $\frac{DV}{dt}$, a vector field along C. If V is the restriction of a vector field, Y on M to C, the we simply define $\frac{DV}{dt}$ by $\frac{DV}{dt} = \nabla_{\frac{dC}{dt}} Y$. For a general vector field,

$$V = \Sigma v^j \partial_j$$

along C we define $\frac{DV}{dt}$ by a derivation law. i.e.

$$\frac{DV}{dt} = \Sigma_j (\frac{dv^j}{dt}\partial_j + v^j \nabla_{\frac{dc}{dt}} \partial_j) = \Sigma_k (\frac{dv^k}{dt} + \Sigma_{i,j} \frac{du^i}{dt} \Gamma^k_{ij} v^j)\partial_k$$

Now we suppose M has a Riemannian metric. A connection is Riemannian if for any curve, C in M, and vector fields, V, W along C we have:

$$\frac{d}{dt}\langle V, W\rangle = \langle \frac{Dv}{dt}, W\rangle + \langle V, \frac{DW}{dt}\rangle$$

A connection is *symmetric* if $\Gamma^k_{ij} = \Gamma^k_{ji}$. For the sequel we shall assume the connection is a symmetric Riemannian metric.

Now define functions, $g_{ij} \underset{def}{=} \langle \partial_i, \partial_j\rangle = \overline{x}_i \cdot \overline{x}_j = \frac{\partial \overline{x}}{\partial u^i}\frac{\partial \overline{x}}{\partial u^j}$ (we are using the summation convention). Note that some authors use E, F, G instead of g_{ij} on a surface. For example see [O, page 220, problem 7]. The arc length is given by

$$J[C] = \int_{t_1}^{t_2} \sqrt{g_{ij}\dot{u}^i \dot{u}^j}\, dt.$$

In particular if the curve is parametrized by arc length, then $g_{ij}\dot{u}^i\dot{u}^j = 1$. The matrix $[g_{ij}]$ is non singular. Define $[g^{ij}]$ by $[g^{ij}] \underset{def}{=} [g_{ij}]^{-1}$We have the following identities which connect the covariant derivative with the metric tensor.

$$\text{The first Christoffel identity:} \quad \sum_{\ell}\Gamma^{\ell}_{ij}g_{\ell k} = \tfrac{1}{2}\left(\frac{\partial g_{jk}}{\partial u_i}+\frac{\partial g_{ik}}{\partial u_j}-\frac{\partial g_{ij}}{\partial u_k}\right)$$

$$\text{The second Christoffel identity:} \quad \Gamma^{\ell}_{ij} = \sum_{k}\tfrac{1}{2}\left(\frac{\partial g_{jk}}{\partial u_i}+\frac{\partial g_{ik}}{\partial u_j}-\frac{\partial g_{ij}}{\partial u_k}\right)g^{k\ell}$$

Notice the symmetry in the first Christoffel identity: $\frac{\partial g_{jk}}{\partial u_i}$ and $\frac{\partial g_{ik}}{\partial u_j}$.

Now if C is given by $\overline{x} = \overline{x}(s)$, where s is arc length then the geodesic curvature is defined to be the curvature of the projection of the curve onto the tangent plane of the surface. In other words there is a decomposition $\underbrace{x''(s)}_{\substack{curvature\\ vector}} = \underbrace{\kappa_n}_{\substack{normal\\ curvature}} \underbrace{\overline{N}}_{\substack{surface\\ normal}} + \kappa_g B$ where B lies in the tangent plane of the surface and is perpendicular to the tangent vector of C. The geodesic curvature may also be defined as $\frac{D}{dt}(\frac{dC}{dt})$ (see [M, page 55]). Intuitively this is the acceleration along the curve in the tangent plane, which is the curvature of the projection of C onto the tangent plane. Recall also that $\kappa_g B = (\ddot{u}^i + \Gamma^i_{k\ell}\dot{u}^k\dot{u}^\ell)\overline{x}_i$ (we are using the Einstein summation convention). See, for example [S, page 132] [L, pages 28-30] or [M, page 55]. The vanishing of the formula for $\kappa_g B$ is often taken as the definition of a geodesic. We shall deduce this as a consequence of the Euler-Lagrange equations.

Theorem 8.1. *If C is a curve of shortest length on a surface $\overline{x} = \overline{x}(u^1, u^2)$ then $\kappa_g = 0$ along C.*

Proof : [Of Theorem 8.1] C is given by $\overline{x}(t) = \overline{x}(u^1(t), u^2(t)) = (x^1(\overline{u}(t)), x^2(\overline{u}(t)), x^3(\overline{u}(t)))$. We are looking for a curve on the surface which minimizes the arc length functional

$$J[C] = \int_{t_1}^{t_2} \sqrt{g_{ij}\dot{u}^i\dot{u}^j}\,dt$$

Now

$$F_{u^k} = \frac{1}{2\sqrt{g_{ij}\dot{u}^i\dot{u}^j}}\frac{\partial g_{ij}}{\partial u^k}\dot{u}^i\dot{u}^j, \quad F_{\dot{u}^k} = \frac{1}{\sqrt{g_{ij}\dot{u}^i\dot{u}^j}}g_{kj}\dot{u}^j \quad (k = 1, 2).$$

Hence condition (12) becomes:

$$\frac{d}{dt}(\frac{1}{\sqrt{g_{ij}\dot{u}^i\dot{u}^j}}g_{kj}\dot{u}^j) = \frac{1}{2\sqrt{g_{ij}\dot{u}^i\dot{u}^j}}\frac{\partial g_{ij}}{\partial u^k}\dot{u}^i\dot{u}^j, \quad k = 1, 2.$$

These are not independent:

Use arc length as the curve parameter, i.e. $g_{ij}\dot{u}^i\dot{u}^j = 1$ (where $\dot{}$ is now $\frac{d}{ds}$.) Then the above conditions become

$$g_{kj}\ddot{u}^j + \frac{1}{2}(\frac{\partial g_{kj}}{\partial u^i} + \frac{\partial g_{ki}}{\partial u^j} - \frac{\partial g_{ij}}{\partial u^k})\dot{u}^i\dot{u}^j = 0, \quad k = 1, 2$$

Now "raise indices", i.e. multiply by g^{km} and sum over k. The result is

$$\ddot{u}^m + \Gamma^m_{ij}\dot{u}^i\dot{u}^j = 0, \quad m = 1, 2$$

hence $\kappa_g = 0$ along C. **q.e.d.**

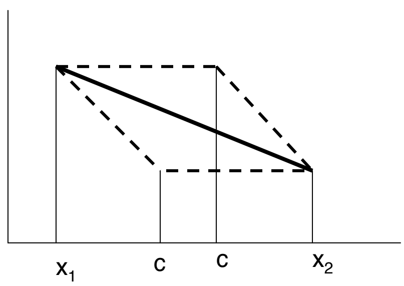
Note that since $x''(s) = \kappa_n\overline{N} + \kappa_g B = \kappa\overline{\wp}(s)$ where $\overline{\wp}$ is the principal normal to C, $\kappa_g = 0$ is equivalent with $\overline{\wp} = \overline{N}$.

An example applying the Weierstrass-Erdmann corner condition (4.2) Consider $J[y] = \int_{x_1}^{x_2}(y' + 1)^2(y')^2dx$ with $y(x_2) < y(x_1)$ to be minimized. Since the integrand is of the form $f(y')$ the extremals are straight lines. Hence the only smooth extremal is the line segment connecting the two points, which gives a positive value for $J[y]$. Next allow one corner, say at $c \in (x_1, x_2)$. Set $y'(c^-) = p_1, y'(c^+) = p_2$ (for the left and right hand slopes at c of a minimizing broken extremal, $y(x)$). Since $f_{y'} = 4(y')^3 + 6(y')^2 + 2y'$, $f - y'f_{y'} = -(3(y')^4 + 4(y')^3 + (y')^2$ the two corner conditions give:

1. $4p_1^3 + 6p_1^2 + 2p_1 = 4p_2^3 + 6p_2^2 + 2p_2$

2. $3p_1^4 + 4p_1^3 + p_1^2 = 3p_2^4 + 4p_2^3 + p_2^2$

   For $p_1 \neq p_2$ these yield $2(\overbrace{p_1^2 + p_1p_2 + p_2^2}^{set=w}) + 3(\overbrace{p_1 + p_2}^{set=u}) + 1 = 0$ and $-3u^3 + 6uw + 4w + u = 0$. The only real solution is $(u, w) = (-1, 1)$, leading to either $p_1 = 0, p_2 = -1$ or $p_1 = -1, p_2 = 0$ (see diagram below).



Each of these broken extremals yields the absolute minimum, 0 for $J[y]$. Clearly there are as many broken extremals as we wish, if we allow more corners, all giving $J[y] = 0$.

9 The first variation of an integral, $\mathfrak{I}(t) = J[y(x,t)] = \int_{x_1(t)}^{x_2(t)} f(x, y(x,t), \frac{\partial y(x,t)}{\partial x})dx$; Application to transversality.

Notation and conventions:

Suppose we are given a 1-parameter family of curves, $y(x,t)$ sufficiently differentiable and simply covering an (x,y)-region, $\mathfrak{F}$. We call $(\mathfrak{F}, \{y(x,t)\})$ a *field.* We also assume $x_1(t)$ and $x_2(t)$ are sufficiently differentiable for $a \le t \le b$. We also assume that for all (x,y) with $x_1(t) \le x \le x_2(t)$, $y = y(x,t), a \le t \le b$ are in $\mathfrak{F}$. y' will denote $\frac{\partial y}{\partial x}$. Finally we assume $f(x,y,y')$ is of class C^3 for all $(x,y) \in \mathfrak{F}$ and for all y'.

Set $Y_i(t) = y(x_i(t),t), i = 1,2$. Then

(14)

$$\frac{dY_i}{dt} = y'(x_i(t),t)\frac{dx_i}{dt} + y_t(x_i(t),t), \quad i = 1,2.$$

The curves $y(x,t)$ join points, $(x_1(t), Y_1(t))$ of a curve C to points $(x_2(t), Y_2(t))$ of a curve D (see the figure below).

We will use the notation $f\Big|^{i}$ or $f\Big|_{i}$ for $f(x_i(t), y(x_i(t),t), y'(x_i(t),t)) \quad i = 1,2$

Definition 9.1. The *first variation* of $\mathfrak{I}(t)$ is

$$d\mathfrak{I} = \frac{d\mathfrak{I}}{dt} \cdot dt$$

Notice this generalizes the definitions in §3 where ϵ is the parameter (instead of t) and the field is given by $y(x,\epsilon) = y(x) + \epsilon\eta(x)$. The curves C and D are both points.

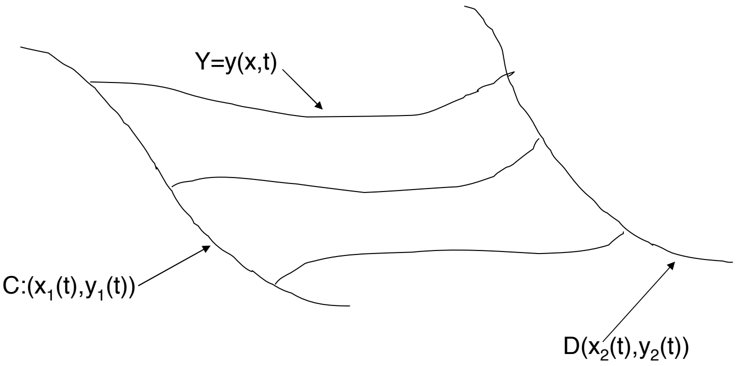


Figure 9: A field, $\mathfrak{F}$.

$$\frac{d\mathfrak{I}}{dt} = \frac{d}{dt}\int_{x_1(t)}^{x_2(t)} f(x, y(x,t), y'(x,t))dx = (f\frac{dx}{dt})\Big|_1^2 + \int_{x_1(t)}^{x_2(t)} (f_y y_t + f_{y'} y_t')dx$$

Using integration by parts and $\frac{\partial}{\partial t}\frac{\partial y}{\partial x} = \frac{\partial}{\partial x}\frac{\partial y}{\partial t}$ we have:

$$\frac{d\mathfrak{I}}{dt} = (f\frac{dx}{dt} + f_{y'}y_t)\Big|_1^2 + \int_{x_1(t)}^{x_2(t)} y_t(f_y - \frac{d}{dx}f_{y'})dx$$

Using (14) this is equal to

(15)

$$[(f - y'f_{y'})\frac{dx}{dt} + f_{y'}\frac{dY}{dt}]\Big|_1^2 + \int_{x_1(t)}^{x_2(t)} y_t(f_y - \frac{d}{dx}f_{y'})dx$$

If $y(x,t)$ is an extremal then $f_y - \frac{d}{dx}f_{y'} = 0$. So for a field of extremals the equations reduce to:

(16)

$$\frac{d\Im}{dt} = [(f - y'f_{y'})\frac{dx}{dt} + f_{y'}\frac{dY}{dt}]\Big|_1^2 = [f\frac{dx}{dt} + f_{y'}(\frac{dY}{dt} - y'\frac{dx}{dt})]\Big|_1^2$$

Notice that in (15) and (16) $(\frac{dx_i}{dt}, \frac{dY_i}{dt})$is the tangent vector to C if $i = 1$ and to D if $i = 2$. $y'(x,t)$ is the "slope function", $p(x,y)$ of the field, $p(x,y) \underset{def}{=} y'(x,t)$.

As a special case of the above we now consider C to be a point. We apply (15) and (16) to derive a necessary condition for an arc from a point C to a curve D that minimizes the integral

$$\int_{x_1(t)}^{x_2(t)} f(x, y(x,t), y'(x,t))dt.$$

The field here is assumed to be a field of extremals, i.e. for each t we have $f_y(x, y(x,t), y'(x,t)) - \frac{d}{dx}f_{y'}(x, y(x,t), y'(x,t)) = 0$. Further the shortest extremal $y(x,t_0)$ joining the point C to D must satisfy $0 = \frac{d\Im}{dt}\Big|_{t=t_0} = [f(x_2, y(x_2(t_0), t_0), y'(x_2(t_0), t_0))\frac{dx_2}{dt}\Big|_{t_0} + f_{y'}(x_2, y(x_2(t_0), t_0), y'(x_2(t_0), t_0))(\frac{dY_2}{dt} - y'(x_2(t_0), t_0)\frac{dx_2}{dt})]\Big|_{t_0}$

Definition 9.2. A curve satisfies the *transversality condition* at t_0 if

$$[f\Big|^2 dx_2 + f_{y'}\Big|^2 (dY_2 - p\Big|^2 dx_2)]\Big|_{t_0} = 0.$$

where p is the slope function.

We have shown that the transversality condition must hold for a shortest extremal from C to the curve D. (9.2) is a condition for the minimizing arc at its intersection with D.

Specifically it is a condition on the direction $(1, p(x(t), y(t)))$ of the minimizing arc, $y(x, t_0)$ and the directions (dx_2, dY_2) of the tangent to the curve D at their point of intersection. In terms of the slopes, p and $Y' = \frac{dY_2}{dx_2}$ it can be written as

$$f\Big|^{2,t_0} - p\Big|^{2,t_0} f_{y'}\Big|^{2,t_0} + f_{y'}\Big|^{2,t_0} Y'\Big|^{t_0} = 0$$

This condition may not be the same as the usual notion of transversality without some condition on f.

Proposition 9.3. *Transversality is the same as orthogonality if and only if $f(x, y, p)$ has the form $f = g(x, y)\sqrt{1 + p^2}$ with $g(x, y) \neq 0$ near the point of intersection.*

Proof : If f has the assumed form then $f_p = \frac{gp}{\sqrt{1+p^2}}$. Then (9.2) becomes

$$g\sqrt{1 + p^2} - \frac{gp^2}{\sqrt{1 + p^2}} + \frac{gpY'}{\sqrt{1 + p^2}} = 0$$

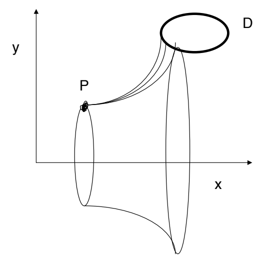
or $pY' = -1$ (assuming $g(x, y) \neq 0$ near the intersection point). So the slopes are orthogonal.

If (9.2) is equivalent to $Y' = -\frac{1}{p}$ then identically in (x, y, p) we have $f - pf_{y'} - \frac{1}{p}f_{y'} = 0$ i.e. $\frac{f_p}{f} = \frac{p}{1+p^2}$ or $\frac{\partial}{\partial p} \ln f = \frac{p}{1+p^2}$, $\quad \ln f = \frac{1}{2}\ln(1 + p^2) + \ln g(x, y)$ or $f = g(x, y)\sqrt{1 + p^2}$.

q.e.d.

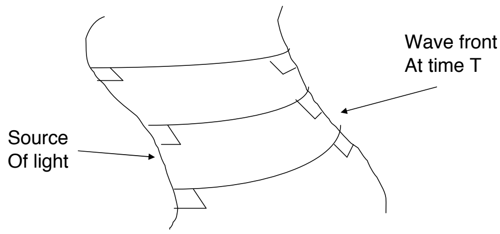
Example: Minimal surface of revolution.

Recall the minimum surface of revolution functional is given by $J[y] = \int_{x_1}^{x_2} y\sqrt{1 + (y')^2}dx$ which satisfies the hypothesis of proposition (9.3). Consider the problem of finding the curve from a point P to a point of a circle D which minimizes the surface of revolution (with horizontal distance $\ll$ the vertical distance). The solution is a catenary from point P which is perpendicular to D.



Remark 9.4. A problem to minimize $\int_{x_1}^{x_2} g(x,y)\sqrt{1+(y')^2}dx$ may always be interpreted as a problem to find the path of a ray of light in a medium of variable velocity $V(x,y)$ (or index of refraction $g(x,y) = \frac{c}{V(x,y)}$) since $dt = \frac{ds}{V(x,y)} = \frac{1}{c}g(x,y)\sqrt{1+(y')^2}dx$. By Fermat's principle the time of travel is minimized by light.

Assume $y(x,t)$ is a field of extremals and C, D are curves (not extremals) lying in the field, $\mathfrak{F}$ with each $y(x,t)$ transversal to both C at $(x_1(t), Y_1(t))$ and D at $(x_2(t), Y_2(t))$. Then since $\frac{d\mathfrak{I}}{dt} = 0$ we have $\mathfrak{I}(t) = \int_{x_1(t)}^{x_2(t)} f(x, y(x,t), y'(x,t))dx = \text{ constant }$. For example if light travels through a medium with variable index of refraction, $g(x,y)$ with light source a curve lighting up at $T_0 = 0$ then at time $T = \int_{x_1}^{x_2} \frac{1}{c}g(x,y)ds$, the "wave front" D is perpendicular to the rays.



10 Fields of Extremals and Hilbert's Invariant Integral.

Assume we have a simply-connected field $\mathfrak{F}$ of extremals for $\int_{x_1}^{x_2} f(x, y, y')dx$[20]. Part of the data are curves, $C : (x_1(t), y_1(t)), D : (x_2(t), y_2(t))$ in $\mathfrak{F}$. Recall the notation: $Y_i(t) = y(x_i(t), t)$.

Consider two extremals in $\mathfrak{F}$

- $y(x, t_P)$ joining P_1 on C to P_2 on D
- $y(x, t_Q)$ joining Q_1 on C to Q_2 on D

with $t_p < t_Q$. Apply (16):

$$\frac{d\mathfrak{I}(t)}{dt} = [f\frac{dx}{dt} + f_{y'}(\frac{dY}{dt} - p\frac{dx}{dt})]\Big|_1^2$$

where $\mathfrak{I}(t) = \int_{x_1(t)}^{x_2(t)} f(x, y(x,t), y'(x,t))dx, \quad p(x, y) = y'(t)$. Hence $\mathfrak{I}(t_Q) - \mathfrak{I}(t_P) = \int_{t_P}^{t_Q} \frac{d\mathfrak{I}}{dt}dt =$

[20]Note that by the assumed nature of $\mathfrak{F}$, each point (x_0, y_0) of $\mathfrak{F}$ lies on exactly one extremal, $y(x, t)$. The simple connectivity condition guarantees that there is a 1-parameter family of extremals from t_Q to t_P.

$$\int_{t_P}^{t_Q}\Big[f(x_2(t),t(x_2(t),t),y'(x_2(t),t))\frac{dx_2}{dt}+$$
$$f_{y'}(x_2(t),t(x_2(t),t),y'(x_2(t),t))(\frac{dY_2}{dt}-p(x_2(t),Y_2(t))\frac{dx_2}{dt})\Big]dt$$
$$-\int_{t_P}^{t_Q}\Big[f(x_1(t),t(x_1(t),t),y'(x_1(t),t))\frac{dx_1}{dt}+$$
$$f_{y'}(x_1(t),t(x_1(t),t),y'(x_1(t),t))(\frac{dY_1}{dt}-p(x_1(t),Y_1(t))\frac{dx_1}{dt})\Big]dt$$

$$=\int_{D_{\widehat{P_2Q_2}}} fdx+f_{y'}(dY-pdx)-\int_{C_{\widehat{P_1Q_1}}} fdx+f_{y'}(dY-pdx).$$

This motivates the following definiton:

Definition 10.1. For any curve, $\mathfrak{B}$, the *Hilbert invariant integral* is defined to be

$$I^*_{\mathfrak{B}}=\int_{\mathfrak{B}} f(x,y,p(x,y))dx+f_{y'}(x,y,p(x,y))(dY-pdx)$$

Then we have shown that

(17)

$$\mathfrak{I}(t_Q)-\mathfrak{I}(t_P)=I^*_{\mathfrak{D}_{\widehat{P_2Q_2}}}-I^*_{\mathfrak{C}_{\widehat{P_1Q_1}}}.$$

From (17) we have the following for a simply connected field, $\mathfrak{F}$:

Theorem 10.2. *$I^*_{\mathfrak{C}_{\widehat{P_1Q_1}}}$ is independent of the path, C, joining the points P_1, Q_1 of $\mathfrak{F}$.*

Proof #1: By (17) $I^*_{\mathfrak{C}_{\widehat{P_1Q_1}}} = I^*_{\mathfrak{D}_{\widehat{P_2Q_2}}} - (\mathfrak{I}(t_Q) - \mathfrak{I}(t_P))$.

The right hand side is independent of C.

Proof #2: For the line integral (10.1)

$$I^*_{\mathfrak{B}} = \int_{\mathfrak{B}} \underbrace{(f - pf_p)}_{P(x,y)} dx + \underbrace{f_p}_{Q(x,y)} dY$$

we must check that $P_y - Q_x = 0$. Now

$$P_y - Q_x = f_y + f_p p_y - p_y f_p - p f_{py} - p f_{pp} p_y - f_{px} - f_{pp} p_x$$

$$= f_y - f_{px} - p f_{py} - f_{pp}(p_x + p p_y)$$

$$= f_y - \frac{d}{dx} f_p$$

This is zero if and only if $p(x, y)$ is the slope function of a field of extremals. **q.e.d.**

I^* in the case of Tangency to a field.

Suppose that $\mathfrak{B} = \mathcal{E}$, a curve that either is an extremal of the field or else shares, at every point (x, y) through which it passes, the triple (x, y, y') with the unique extremal of $\mathfrak{F}$ that passes through (x, y). Thus $\mathfrak{B}$ may be an envelope of the family $\{y(x,t) \big| a \leq t \leq b\}$. Then along $\mathfrak{B} = \mathcal{E}$, $\quad dY = p(x, y)dx$. Hence by (10.1) we have:

(18)

$$I^*_{\mathcal{E}} = \int_{\mathcal{E}} f(x, y, p)dx = \mathfrak{I}[y(x)] \underset{def}{=} \mathfrak{I}(\mathcal{E}).$$

I^* in case of Transversality to the field.

Suppose $\mathfrak{B}$ is a curve satisfying, at every point, transversality, (9.2), then

(19)

$$I^*_{\mathfrak{B}} = \int_{\mathfrak{B}} (f + f_{y'}[\frac{dY}{dx} - p(x,y)])dx = 0$$

11 The Necessary Conditions of Weierstrass and Legendre.

We derive two further necessary conditions, on class $\widetilde{C}^{(1)}$ functions, $y_0(x)$ furnishing a relative minimum for $J[y] = \int_{x_1}^{x_2} f(x, y, y')dx$.

Definition 11.1 (The Weierstrass $\mathcal{E}$ function)**.** Associated to $J[y]$ define the following function in the 4 variables x, y, y', Y':

$$\mathcal{E}(x, y, y', Y') = f(x, y, Y') - f(x, y, y') - (Y' - y')f_{y'}(x, y, y')$$

Taylor's theorem with remainder gives the formula:

$$f(x, y, Y') = f(x, y, y') + f_{y'}(x, y, y')(Y' - y') + \frac{f_{y'y'}(x, y, y' + \theta(Y' - y'))}{2}(Y' - y')^2$$

where $0 < \theta < 1$. Hence we may write the Weierstrass $\mathcal{E}$ function as

$$\mathcal{E}(x, y, y', Y') = \frac{(Y' - y')^2}{2} f_{y'y'}(x, y, y' + \theta(Y' - y')) \tag{20}$$

However the region, $\mathcal{R}$ of admissible triples, (x, y, z) for $J[y]$ may conceivably contain $(x, y, y'$ and (x, y, Y') without containing $(x, y, y' + \theta(Y' - y'))$ for all $\theta \in (0, 1)$. Thus to make sure that Taylor's theorem is applicable, and (20) valid we must assume either:

(21)

- Y' is sufficiently close to y'; since $\mathcal{R}$ is open, it will contain all $(x, y, y' + \theta(Y' - y'))$ if it contains (x, y, y'), or

- Assume that $\mathcal{R}$ satisfies *condition "C"*: If $\mathcal{R}$ contains $(x, y, z_1), (x, y, z_2)$ then it contains the line connecting these two points.

Theorem 11.2 (Weierstrass' Necessary Condition for a Relative Minimum, 1879)**.** *If* $y_0(x)$ *gives a* $\left\{\begin{array}{ll}(i) & \textit{strong}\\ (ii) & \textit{weak}\end{array}\right\}$ *relative minimum for* $\int_{x_1}^{x_2} f(x, y, y')dx$*, then* $\mathcal{E}(x, y_0(x), y_0'(x), Y') \geq 0$ *for all* $x \in [x_1, x_2]$ *and for* $\left\{\begin{array}{ll}(i) & \textit{all } Y'\\ (ii) & \textit{all } Y' \textit{sufficiently close to } y_0'\end{array}\right\}$.

Proof :

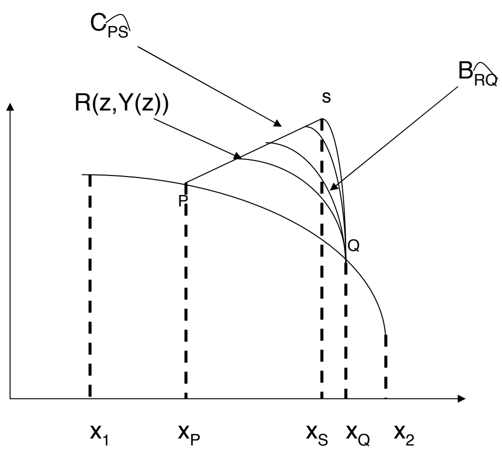


Figure 10: constructions for proof of 11.2

On the $\widetilde{C}^{(1)}$ extremal $E : y = y_0(x)$, let P be any point $(x_p, y_0(x_P))$. Let Q be a nearby point on E with no corners between P and Q such that, say for definiteness $x_P < x_Q$. Let Y_0' be any number $\neq y_0'(x_P)$ (and under alternative (ii) of the theorem sufficiently close to $y_0'(x_P)$). Construct a parametric curve $C_{\widehat{PS}} : x = z, y = Y(z)$ with $x_P < x_S < x_Q$ and $Y'(x_P) = Y_0'$. Also construct a one-parameter family of curves $B_{\widehat{RQ}} : y = y_z(x)$ joining the points $R = R(z, Y(z))$ of $C_{\widehat{PS}}$ to Q. e.g. set $y_z(x) = y_0(x) + \frac{Y(z)-y_0(x)}{z-x_Q}(x - x_Q) \quad z \leq x \leq x_Q$.

Now by hypothesis, E minimizes $J[y]$, hence also, for any point R of $C_{\widehat{PS}}$ we have $J[C_{\widehat{PR}}] + J[B_{\widehat{RQ}}] \geq J[E_{\widehat{PQ}}]$. That is to say

$$\int_{x_P}^{z} f(x, Y(x), Y'(x))dx + J[B_{\widehat{RQ}}] \geq J[E_{\widehat{PQ}}].$$

Set the left-hand side $= L(z)$. Then the right-hand side is $L(x_P)$ hence $L(z) \geq L(x_p)$ for all $z, \quad x_P \leq z \leq x_s$. Since $L'(z)$ exist (as a right-hand derivative at $z = x_P$) this implies $L'(x_P^+) \geq 0$. But $L'(z) = f(z, Y(z), Y'(z)) + \frac{dJ[B_{\widehat{RQ}}]}{dz}$ and

$$L'(x_P^+) = f(x_P, \underbrace{Y(x_P)}_{=y_0(x_P)}, Y'(x_P)) + \frac{dJ[B_{\widehat{RQ}}]}{dz}\Big|_{z=x_P}.$$

Now apply (16)[21] We obtain

$$-\frac{dJ[B_{\widehat{RQ}}]}{dz}\Big|_{z=x_P} = f(x_p, y_0(x_P), y_0'(x_P))\frac{dz}{dz} + f_{y'}(x_p, y_0(x_P), y_0'(x_P))(Y_0' - y_0'(x_P)\frac{dz}{dz}).$$

Hence $L'(x_P^+) \geq 0$ yields

$$f(x_p, y_0(x_p), Y_0') - f(x_p, y_0(x_p), y_0') - f_{y'}(x_p, y_0(x_p), y_0'(x_P))(Y_0' - y_0'(x_P)) \geq 0$$

q.e.d.

[21]Note that the curve D degenerates here to the single point Q.

Corollary 11.3 (Legendre's Necessary Condition, 1786)**.** *For every element* (x, y_0, y_0') *of a minimizing arc* $E : y = y_0(x)$ *we must have* $f_{y'y'}(x, y_0, y_0') \geq 0$.

Proof : $f_{y'y'}(x, y_0, y_0') < 0$ at any $x \in [x_1, x_2]$ then by (20) we would contradict Weierstrass' condition, $\mathcal{E} \geq 0$. **q.e.d.**

Definition 11.4. A minimizing problem for $\int_{x_1}^{x_2} f(x, y, y')dx$ is *regular* if and only if $f_{y'y'} > 0$ throughout the region $\mathcal{R}$ of admissible triples and if $\mathcal{R}$ satisfies condition C of (21).

Note: The Weierstrass condition, $\mathcal{E} \geq 0$ and the Legendre condition, $f_{y'y'} \geq 0$ are necessary for a minimizing arc. For a maximizing arc replace ≥ 0 by ≤ 0.

Examples.

1. $J[y] = \int_{x_1}^{x_2} \frac{dx}{1+(y')^2}$: Clearly $0 < J[y] < x_2 - x_1$. The infemum, 0 is not attainable, but is approachable by oscillating curves, $y(x)$ with large $|y'|$. The supremum, $x_2 - x_1$ is furnished by the maximizing arc $y(x) =$ a constant which applies if and only if $y_1 = y_2$. If $y_1 \neq y_2$ the supremum is attainable only by curves composed of vertical and horizontal segments (and these are not admissible if we insist on curves representable by $y(x) \in \widetilde{C}^1$).

Nevertheless Euler's condition, (3.2) furnishes all straight lines, and Legendre's condition (11.3) is satisfied along any straight line:

- Euler: $f_{y'} - \int_{x_1}^{x_2} f_y d\overline{x} = c = \frac{-2y'}{(1+(y')^2)^2}$ gives $y' =$ constant, i.e. a straight line.

- Legendre: $f_{y'y'} = 2\frac{3(y')^2-1}{(1+(y')^2)^2}$ which is $\begin{cases} > 0, & \text{if } |y'| > \frac{1}{\sqrt{3}} \\ = 0, & \text{if } |y'| = \frac{1}{\sqrt{3}} \\ < 0 & \text{if } |y'| < \frac{1}{\sqrt{3}} \end{cases}$;

Thus lines with slope y' such that $|y'| \geq \frac{1}{\sqrt{3}}$ may furnish minimum (but we know they don't!), lines with slope y' such that $|y'| \leq \frac{1}{\sqrt{3}}$ may furnish a maxima, which we also know they don't unless $y' = 0$. This shows that Euler and Legendre taken together are insufficient to guarantee a maximum or minimum.

The Weierstrass condition is more revealing for this problem. We have

$$\mathcal{E} = \frac{(y' - Y')^2 \cdot ((y')^2 - 1 + 2Y'y')}{(1 + (Y')^2)(1 + (y')^2)}.$$

Because of the factor $(y')^2 - 1 + 2Y'y')$ $\mathcal{E}$ can change sign as Y' varies, at any point of an extremal unless $y' = 0$ at that point. Thus the necessary condition ($\mathcal{E} \geq 0$ or ≤ 0) along the extremal implies that we must have $y' \equiv 0$.

2. $\int_{(0,0)}^{(1,0)} (x(y')^4 - 2y(y')^3)dx$**:** Just in case the previous example may have raised exaggerated hopes this example will show that Weierstrass' condition is insufficient to guarantee a maximum or minimum. The Euler-Lagrange equation gives $f_y - \frac{df_{y'}}{dx} = 6y'y''(y - xy') = 0$ which gives straight lines only. Therefore $y \equiv 0$ is the only candidate joining $(0,0)$ and $(1,0)$. Now

$$\mathcal{E} = x(Y')^4 - 2y(Y')^3 - x(y')^4 + 2y(y')^3 - (Y' - y') \cdot 2(y')^2(xy' - 3y) = x(Y')^4 \text{ along } y = 0.$$

Hence $\mathcal{E} \geq 0$ along $y \equiv 0$ for all Y'. i.e. Weierstrass' necessary condition for a minimum is satisfied along $y = 0$. But $y = 0$ gives $J[y] = 0$, while a broken extremal made of a line from

$(0,0)$ to a point (h,k), say with $h,k>0$ followed by a line from (h,k) to $(1,0)$ gives $J[y]<0$ if h is small. By rounding corners we may get a negative value for $J[y]$ if y is smooth. Thus Weierstrass's condition of $\mathcal{E}\geq 0$ along an extremal is not sufficient for a minimum.

12 Conjugate Points,Focal Points, Envelope Theorems

The notion of an *envelope* of a family of curves may be familiar to you from a course in differential equations. Briefly if a one-parameter family of curves in the xy-plane with parameter t is given by $F(x, y, t) = 0$ then it may be the case that there is a curve C with the following properties.

- Each curve of the family is tangent to C
- C is tangent at each of its points to some unique curve of the family.

If such a curve exist , we call it the envelope of the family. For example the set of all straight lines at unit distance from the origin has the unit circle as its envelope.

If the envelope C exist each point is tangent to some curve of the family which is associated to a value of the parameter. We may therefore regard C as given parametrically in terms of t.

$$x = \phi(t), \quad y = \Psi(t).$$

One easily sees that the envelope, C is a solution of the equations:

$$F(x, y, t) = 0, \quad F_t(x, y, t) = 0.$$

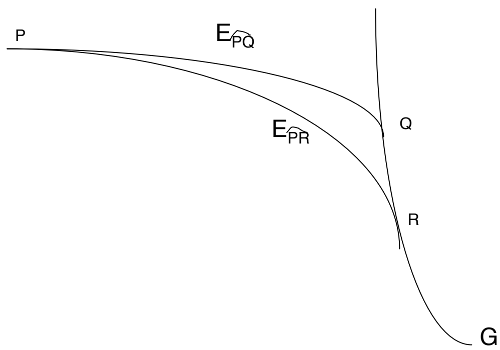
Theorem 12.1. *Let $E_{\widehat{PQ}}$ and $E_{\widehat{PR}}$ be two members of a one-parameter family of extremals through the point P touching an envelope, G, of the family at their endpoints Q and R. Then*

$$\mathfrak{I}(E_{\widehat{PQ}}) + \mathfrak{I}(G_{\widehat{QR}}) = \mathfrak{I}(E_{\widehat{PR}})$$

Proof : By (17) we have

$$\mathfrak{I}(E_{\widehat{PR}}) - \mathfrak{I}(E_{\widehat{PQ}}) = I^*_{G_{\widehat{QR}}}$$

By (18) $I^*_{G_{\widehat{QR}}} = \mathfrak{I}($ **.d.**



(12.1) is an example of an *envelope theorem*. Note that Q precedes R on the envelope G, in the sense that as t increases, E_t induces a path in G transversed from Q to R. The terminology is to say E_t is transversed from P through G to R.

Definition 12.2. The point R is *conjugate* to P along E_{PR}.

Next let $\{y(x,t)\}$ be a 1-parameter family of extremals of $J[y]$ each of which intersects a curve N transversally (cf. (9.2)). Assume that this family has an envelope, G. The point of contact, P_2, of the extremal $E_{\widehat{P_1P_2}}$ with G is called the *focal point* of P_1 on $E_{\widehat{P_1P_2}}$[22]. By (17) we have:

$$\mathfrak{I}(E_{\widehat{Q_1Q_2}}) - \mathfrak{I}(E_{\widehat{P_1P_2}}) = I^*(G_{\widehat{P_2Q_2}}) - I^*(N_{\widehat{P_1Q_1}}).$$

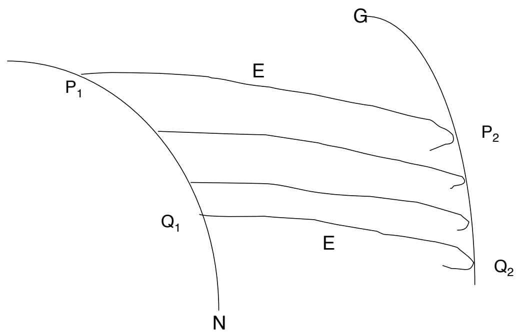


Figure 11: Extremals transversal at one end, tangent at the other.

But by (18): $I^*(G_{\widehat{P_2Q_2}}) = \mathfrak{I}(G_{\widehat{P_2Q_2}})$ and by (19): $I^*(N_{\widehat{P_1Q_1}}) = 0$; hence we have another envelope theorem

Theorem 12.3. *With N, E, and G as above*

$$\mathfrak{I}(E_{\widehat{Q_1Q_2}}) = \mathfrak{I}(E_{\widehat{P_1P_2}}) + \mathfrak{I}(G_{\widehat{P_2Q_2}})$$

Example: Consider the extremals of the length functional $J[y] = \int_{x_1}^{x_2} \sqrt{1 + (y')^2}$, i.e. straight lines, transversal (i.e. $\perp$) to a given curve N. One might think that the shortest distance from a point, 1, to a curve N is given by the straight line from 1 which intersects N at a right angle. For some curves, N this is the case *only if the point,* 1, *is not too far*

[22]The term comes from the analogy with the focus of a lens or a curved mirror.

from N! For some N the lines have an envelope, G called the *evolute* of N (see figure (12)). (12.3) says that the distance along the straight line from 3 to the point 2 on N is the same as going along G from 3 to 5 then following the straight line $E_{\widehat{54}}$ to N. This is the *string property of the evolute.* It implies that a string fastened at 3 and allowed to wrap itself around the evolute G will trace out the curve N.

Now the evolute is not an extremal! So any line from 1 to 2 (with a focal point, 3 between 1 and 2) has a longer distance to N than going from 1 to 3 along the line, followed by tracing out the line, L from 3 to 5 followed by going along the line from 5 to 4 on the curve.

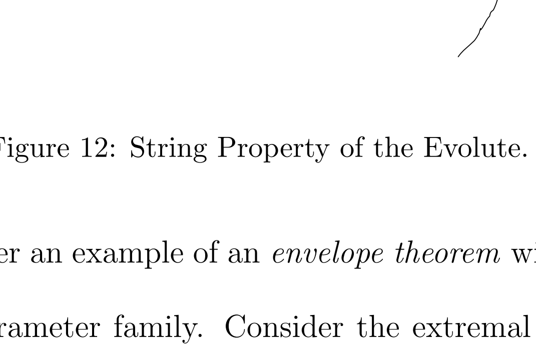


Figure 12: String Property of the Evolute.

Example: We now consider an example of an *envelope theorem* with tangency at both ends of the extremals of a 1-parameter family. Consider the extremal $y = b_0 \cosh \frac{x-a_0}{b_0}$ (a catenary) of the minimum surface of revolution functional $J[y] = \int_{x_1}^{x_2} y\sqrt{1+(y')^2}dx$. Suppose the catenary passes through P_1. The conjugate point P_1^c of P_1 was found by Lindelöf's construction, based on the fact that the tangents to the catenary at P_1 and at P_1^c intersect on the x-axis (we will prove this in §15). We now construct a 1-parameter family of catenaries with the same tangents by projection of the original catenary $\widehat{P_1P_1^c}$, 'contracting' to $T(x_0, 0)$. Specifically we have $y = \frac{b_0}{b}w$ and $x = \frac{b_0}{b}(z - x_0) + x_0$, so (z, w) therefore satisfies

$$w = b\cosh\frac{z - [x_0 - \frac{b}{b_0}(x_0 - a_0)]}{b}$$

which gives another w).

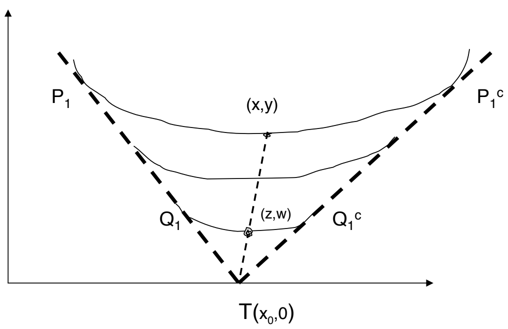


Figure 13: Envelope of a Family of Catenaries.

Now apply (17) and (18):

$$\mathfrak{I}(E_{\widehat{Q_1Q_1^c}}) - \mathfrak{I}(E_{\widehat{P_1P_1^c}}) = \mathfrak{I}(\overline{P_1^cQ_1^c}) - \mathfrak{I}(\overline{P_1Q_1})$$

This is our third example of an *envelope theorem.* Now let $b \to 0$. i.e. $Q_1 \to T, \quad Q_1^c \to 0$. Therefore $\mathfrak{I}(E_{\widehat{Q_1Q_1^c}}) \to 0$. We conclude that

$$\mathfrak{I}(E_{\widehat{P_1P_1^c}}) = \mathfrak{I}(\overline{P_1T}) - \mathfrak{I}(\overline{P_1^cT}) = \mathfrak{I}(\overline{P_1T} + \overline{TP_1^c}).$$

Hence $E_{\widehat{Q_1 Q_1^c}}$ certainly does not furnish an absolute minimum for $J[y] = \int_{x_1}^{x_2} y\sqrt{1+(y')^2}dx$, since the polygonal curve $\overline{P_1 T P_1^c}$ does as well as $E_{\widehat{Q_1 Q_1^c}}$, and Goldschmidt's discontinuous solution, $\overline{P_1 P_1' P_1^{c\prime} P_1^c}$, does better ($P_1'$ is the point on the x-axis below P_1 and $P_1^{c\prime}$ is the point on the x axis below P_1^c).

13 Jacobi's Necessary Condition for a Weak (or Strong) Minimum: Geometric Derivation

As we saw in the last section an extremal emanating from a point P that touches the point conjugate to P may not be a minimum. Jacobi's theorem generalizes the phenomena observed in these examples.

Theorem 13.1 (Jacobi's necessary condition)**.** *Assume that $E_{\widehat{P_1P_2}}$ furnishes a (weak or strong) minimum for $J[y] = \int_{x_1}^{x_2} f(x, y, y')dx$, and that $f_{y'y'}(x, y_0(x)y'(x)) > 0$ along $E_{\widehat{P_1P_2}}$. Also assume that the 1-parameter family of extremals through P_1 has an envelope, G. Let P_1^c be a conjugate point of P_1 on $E_{\widehat{P_1P_2}}$. Assume G has a branch pointing "back" to P_1 (cf. the paragraph after (12.1)). Then P_1^c must come after P_2 on $E_{\widehat{P_1P_2}}$ (symbolically: $P_1 \prec P_2 \prec P_1^c$).*

Proof : Assume the conclusion false. Then let R be a point of G "before" P_1^c (see the diagram). Applying the envelope theorem, (12.1) we have:

$$\mathfrak{J}(E_{\widehat{P_1P_1^c}}) = \mathfrak{J}(E_{\widehat{P_1R}}) + \mathfrak{J}(G_{\widehat{RP_1^c}})$$

Hence

$$\mathfrak{J}(E_{\widehat{P_1P_2}}) = \mathfrak{J}(E_{\widehat{P_1R}}) + \mathfrak{J}(G_{\widehat{RP_1^c}}) + \mathfrak{J}(E_{\widehat{P_1^cP_2}}).$$

Now $f_{y'y'} \neq 0$ at P_1^c; hence $E_{\widehat{P_1P_2}}$ is the unique extremal through P_1^c in the direction of $G_{\widehat{RP_1^c}}$ (see the paragraph after (4)), so that G is not an extremal! Hence for R sufficiently close to P_1^c on G, we can find a curve $F_{\widehat{RP_1^c}}$, within any prescribed weak neighborhood of $G_{\widehat{RP_1^c}}$ such

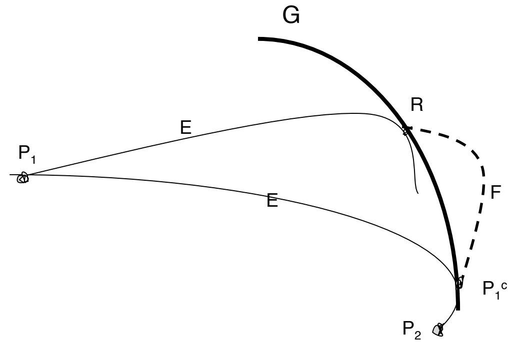


Figure 14: Diagram for Jacobi's necessary condition.

that $\Im(F_{\widehat{RP_1^c}}) < \Im(G_{\widehat{RP_1^c}})$. Then

$$\Im(E_{\widehat{P_1P_2}}) > \Im(E_{\widehat{P_1R}}) + \Im(F_{\widehat{RP_1^c}}) + \Im(E_{\widehat{P_1^cP_2}}),$$

which shows that $E_{\widehat{P_1P_2}}$ does not give even a weak relative minimum for $\int_{x_1}^{x_2} f(x, y, y')dx$. **q.e.d.**

Note: If the envelop G has no branch pointing back to P_1, (e.g. it may have a cusp at P_1^c, or degenerate to a point), the above proof doesn't work. In this case, it may turn out tat P_1^c may coincide with P_2 without spoiling the minimizing property of $E_{\widehat{P_1P_2}}$. For example great semi-circles joining P_1 (= the north pole) to P_2 (= the south pole). $G = P_2$. But in no case may P_1^c lie between P_1 and P_2. For this case there is a different, analytic proof of Jacobi's condition based on the *second variation* which will be discussed §21 below.

Example: An example showing that Jacobi's condition, even when combined with the Euler-Lagrange equation and Lagendre's condition is not sufficient for a strong minimum.

Consider the functional

$$J[y] = \int_{(0,0)}^{(1,0)} ((y')^2 + (y')^3)dx$$

The Euler-Lagrange equation is

$$2y' + 3(y')^2) = \text{ constant}$$

therefore the extremals are straight lines, $y = ax + b$[23]. Hence the extremal joining the end points is $E_0 : y_0 \equiv 0, \quad 0 \le x \le 1$ with $J[y_0] = 0$.

Legendre's necessary condition says $f_{y'y'} = 2(1 + 3y') > 0$ along E_0. This is ok.

Jacobi's necessary condition: The family of extremals through $(0,0)$ is $\{y = mx\}$ therefore there is no envelope i.e. there are no conjugate points to worry about. But E_0 does not furnish a <u>strong</u> relative minimum. To see this consider C_h, the polygonal line from $(0,0)$ to $(1-h, 2h)$ to $(1,0)$ with $h > 0$. We obtain $J[C_h] = 4h(-1 + \frac{h}{1-h} + \frac{2h^2}{(1-h)^2})$ which is < 0 if h is small. Thus E_0 does not give a strong minimum.

However if $0 < \epsilon < 2$ C_h is not in a <u>weak</u> ϵ-neighborhood of E_0 since the second portion of C_h has slope -2. E_0 does indeed furnish a weak relative minimum. In fact for any $y = y(x)$ such that $|y'| < 1$ on $[0,1]$, $((y')^2 + (y')^3) = (y')^2(1 + y') > 0$. Hence $J[y] > 0$. In this connection, note also that $\mathcal{E}(x, 0, 0, Y') = (Y')^2(1 + Y')$.

[23]We already observed that any functional $\int_{x_1}^{x_2} f(x, y, y')dx$ with $f(x, y, y') = g(y')$ has straight line extremals.

14 Review of Necessary Conditions, Preview of Sufficient Conditions.

<u>Necessary Conditions</u> for $J[y_0] = \int_{x_1}^{x_2} f(x,y,y')dx$ **to be minimized by** $E_{\widehat{P_1P_2}} : y = y_0(x)$**:**

$E.(\underline{Euler}, \underline{1744})$	$y_0(x)$ must satisfy $f_{y'} - \int_{x_1}^{x_2} f_y d\overline{x} = c$
	therefore wherever $y_0' \in C[x_1,x_2] : \frac{d}{dx} f_{y'} - f_y = 0.$
$E'.$	therefore wherever $f_{y'y'}(x, y_0(x), y_0'(x)) \neq 0, \quad y'' = \frac{f_y - f_{y'x} - f_{y'y}y'}{f_{y'y'}}.$
$L(\underline{Legender}, \underline{1786})$	$f_{y'y'}(x, y_0(x), y_0'(x)) \geq 0$ along $E_{\widehat{P_1P_2}}$.
$J.(\underline{Jacobi}, \underline{1837})$	No point conjugate to P_1 should precede P_2 on $E_{\widehat{P_1P_2}}$
	if $E_{\widehat{P_1P_2}}$ is an extremal satisfying $f_{y'y'}(x, y_0(x), y_0'(x)) \neq 0$.
$W.(\underline{Weierstrass}, \underline{1879})$,	$\mathcal{E}(x, y_0(x), y_0'(x), Y') \geq 0$ (along $E_{\widehat{P_1P_2}}$), for all Y'
	(or for all Y' of some neighborhood of $y_0'(x)$.

Of these conditions, E and J are necessary for a <u>weak</u> relative minimum. W for <u>all Y'</u> is necessary for a <u>strong</u> minimum; W with <u>all Y' near $y'(x)$</u> is necessary for a <u>weak</u> minimum. L is necessary for a weak minimum.

Stronger versions of conditions L,J,W will figure below in the discussion of sufficient conditions.

(i): Stronger versions of L.

$$\begin{cases} L', & > 0 \text{ instead of } \geq 0 \text{ along } E_{\widehat{P_1P_2}} \\ L_\alpha, & f_{y'y'}(x, y_0(x), Y') \geq 0 \text{ for all } (x, y_0(x)) \text{ of } E_{\widehat{P_1P_2}} \\ & \quad \text{and all } Y' \text{ for which } (x, y_0(x), Y') \text{is an admissible triple.} \\ L_b, & f_{y'y'}(x, y, Y') \geq 0 \text{ for all } (x, y) \text{ near points } (x, y_0(x)) \text{ of } E_{\widehat{P_1P_2}} \\ & \quad \text{and all } Y' \text{ such that } (x, y, Y') \text{ is admissable .} \\ L'_\alpha, L'_b, & \text{replace } \geq 0 \text{ with } > 0 \text{ in } L_\alpha \text{ and } L_b \text{ respectively.} \end{cases}$$

(ii): Stronger versions of J.

$$\begin{cases} J', & \text{On } E_{\widehat{P_1P_2}} \text{ such that } f_{y'y'} \neq 0 : P_1^c \text{ (the conjugate of } P_1\text{) should come } \underline{\text{after}}\ P_2 \\ & \text{if at all. i.e., not only } P_2 \preceq P_1^c \text{ but actually } P_2 \prec P_1^c. \end{cases}$$

(iii): Stronger versions of W.

$$\begin{cases} W', & \mathcal{E}((x, y_0(x), y_0'(x), Y') > 0 \text{ (instead of } \geq 0\text{) whenever } Y' \neq y_0'(x) \text{ along } E_{\widehat{P_1P_2}}. \\ W_b, & \mathcal{E}(x, y, y', Y') \geq 0 \text{ for all } (x, y, y') \text{ near elements } (x, y_0(x), y_0'(x)) \text{ of } E_{\widehat{P_1P_2}}, \\ & \quad \text{and all admissible } Y'. \\ W_b', & W_b \text{ with } > 0 \text{ (in place of } \geq 0\text{) whenever } Y' \neq y'. \end{cases}$$

Two further conditions are isolated for use in sufficiency proofs:

$$\begin{cases} C: \\ \qquad \text{The region} R' \text{ of admissible elements } (x,y,y'), \text{ for the minimizing problem at} \\ \qquad \text{hand has the property that if } (x,y,y_1') \text{ and } (x,y,y_2') \\ \qquad \text{are admissible, then so is every } (x,y,y') \text{ where } y' \text{ is between } y_1 \text{ and } y_2. \\ F: \\ \qquad \text{Possibility of imbedding a smooth extremal } E_{\widehat{P_1P_2}} \text{ in a field of smooth extremals. :} \\ \qquad \text{There is a field, } \mathfrak{F} = (F, y(x,t)), \text{ of extremals, } y(x,t), \text{ which are a one-parameter} \\ \qquad \text{family } \{y(x,t)\} \text{ of them simply and completely covering} \\ \qquad \text{a simply-connected region } F, \text{ such that } E_{\widehat{P_1P_2}} : y = y(x,0) \text{ is interior to } F \\ \qquad \text{and such that } y(x,t) \text{ and } y'(x,t) (= \frac{\partial y(x,t)}{\partial x}) \text{ are of class } C^2. \\ \qquad \text{We also require } y_t(x,t) \neq 0 \text{ on each extremal } y = y(x,t). \end{cases}$$

Significance of C : Referring to (20).

$$\mathcal{E}(x,y,y',Y') = \frac{(Y'-y')^2}{2} f_{y'y'}(x,y,\widetilde{y}'),$$

with $\widetilde{y}'$ between y' and Y' shows that if condition C holds, then $\underline{L' \text{ implies } W'}$ at least for all Y' sufficiently close to $y_0'(x)$.

Significance of F : Preview of sufficiency theorems.

1. Imbedding lemma: $E \& L' \& J' \Rightarrow F$

This will be proven in §16.

2. Fundamental Sufficiency Lemma.

$$[E_{\widehat{P_1P_2}} \text{ satisfies } E'\&F] \Rightarrow [J(C_{\widehat{P_1P_2}}) - J(E_{\widehat{P_1P_2}}) = \int_{x_1}^{x_2} \mathcal{E}(x, Y(x), p(x, Y(x)), Y'(x))dx]$$

where $C_{\widehat{P_1P_2}}$ is any neighboring curve: $y = Y(x)$ of $E_{\widehat{P_1P_2}}$, lying within the field region, F, and $p(x, y)$ is the slope function of the field (i.e. $p(x, y) = y'(x, t)$) of the extremal $y = y(x, t)$ through (x, y).

This will be proven in §17.

From **1 & 2** we shall obtain the first sufficiency theorem:

$E\&L'\&J'\&W_b'$ is sufficient for $E_{\widehat{P_1P_2}}$ to give a strong relative, proper minimum for $\mathcal{I}(E_{\widehat{P_1P_2}}) = \int\limits_{E_{\widehat{P_1P_2}}} f(x, y, y')dx$.

We proved, in (20) that $L' \Rightarrow [W'$ holds for all elements (x, y, y') in some <u>weak</u> neighborhood of $E_{\widehat{P_1P_2}}]$. Hence we have a second sufficiency theorem:

$E\&L'\&J'$ are sufficient for $E_{\widehat{P_1P_2}}$ to give a weak relative, proper minimum.

These sufficiency theorems will be proven in §18.

We conclude with the following:

Corollary 14.1. *If the region of admissible triples, R' satisfies condition C, then $E\&L_b'\&W'$ are sufficient for $E_{\widehat{P_1P_2}}$ to give a strong relative minimum.*

Proof : Exercise.

15 More on Conjugate Points on Smooth Extremals.

Assume we are dealing with a regular problem. i.e. assume that the base function, f of the functional $J[y] = \int_{x_1}^{x_2} f(x, y, y')dx$ satisfies $f_{y'y'}(x, y, y') \neq 0$ for all admissible triples, (x, y, y').

1. By the consequence of Hilbert's theorem, (4), there is a 2-parameter family, $\{y(x, a, b)\}$ of solution curves of

$$y'' = \frac{f_y - f_{y'x} - y' f_{y'y}}{f_{y'y'}}.$$

By the Picard existence and uniqueness theorem for such a second order differential equation, given an admissible element (x_0, y_0, y'_0) there exist a unique a_0, b_0 such that $y_0 = y(x_0, a_0, b_0)$, $y'_0 = y'(x_0, a_0, b_0)$.

More specifically we can construct a two-parameter family of solutions by letting the parameters, a, b be the values assumed by y, y' respectively at a fixed abscissa, x_1, e.g. the abscissa of P_1, the initial point of $E_{\widehat{P_1P_2}}$. If we do this then

$$\left.\begin{aligned} a &= y(x_1, a, b) \\ b &= y'(x_1, a, b) \end{aligned}\right\} \text{identically in } a, b$$

hence

$$\begin{aligned} 1 &= y_a(x_1, a, b), \quad 0 = y_b(x_1, a, b) \\ 0 &= y'_a(x_1, a, b), \quad 1 = y'_b(x_1, a, b) \end{aligned}$$

This shows that we can construct the family $\{y(x, a, b)\}$ in such a way that if we set

$$\text{(22)} \qquad \Delta(x, x_1) \underset{def}{=} \begin{vmatrix} y_a(x, a_0, b_0) & y_b(x, a_0, b_0) \\ y_a(x_1, a_0, b_0) & y_b(x_1, a_0, b_0) \end{vmatrix}$$

where (a_0, b_0) gives the extremal $E_{\widehat{P_1P_2}}$. Then

$$\Delta'(x_1, x_1) = \begin{vmatrix} y'_a(x_1, a_0, b_0) & y'_b(x_1, a_0, b_0) \\ y_a(x_1, a_0, b_0) & y_b(x_1, a_0, b_0) \end{vmatrix} \neq 0$$

2. Now consider the 1-parameter subfamily, $\{y(x,t)\} = \{y(x, a(t), b(t))\}$ of extremals through $P_1(x_1, y_1)$. Thus $a(t), b(t)$ satisfy $y_1 = y(x_1, a(t), b(t))$. e.g. If we construct $y(x, a, b)$ specifically as in **1** above, then $a(t) = y_1$ for the subfamily, and b itself can then serve as the parameter, t, which in this case represents the slope of the particular extremal at P_1.

Let us call a point (x^C, y^C), other than (x_1, y_1) conjugate to $P_1(x_1, y_1)$ if and only if it satisfies

$$\begin{cases} y^C = & y(x^C, t) \\ 0 = & y_t(x^C, t) \end{cases}$$

for some t. The set of conjugate points includes the envelope, G of the 1-parameter subfamily $\{y(x,t)\}$ of extremals through P_1 as defined in §12[24].

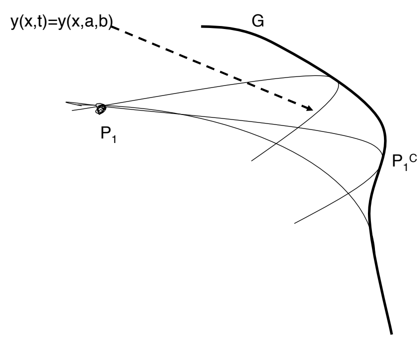


Figure 15:

[24]Note: the present definition of conjugate point is a little more inclusive than the definition in §12.

We now derive a characterization of the conjugate points of P_1, in terms of the function $y(x,a,b)$. Namely at P_1 we have $y_1 = y(x_1, a(t), b(t))$ for all t. Therefore

$$\begin{aligned} 0 =& \quad y_a(x_1,a,b)a' + y_b(x_1,a,b)b' \\ \text{At } P_1^C: \ 0 = y_t(x^C,t) =& \quad y_a(x^C,a,b)a' + y_b(x^C,a,b)b' \end{aligned} \tag{23}$$

View (23) as a homogenous linear system. There is a non-trivial solution, a', b', hence the determinant of the system must be 0. That is to say at any point (x^C, y^C) conjugate to $P_1(x_1, y_1)$

$$\begin{vmatrix} y_a(x^C,a,b) & y_b(x^C,a,b) \\ y_a(x_1,a,b) & y_b(x_1,a,b) \end{vmatrix} = 0 \tag{24}$$

In particular if $E_{\widehat{P_1P_2}}$ corresponds to (a_0, b_0), the conjugate points (x^C, y^C) of P_1 must satisfy $\Delta(x^C, x_1) = 0$ where Δ is defined in (22).

3. Examples.

(i) $J[y] = \int_{x_1}^{x_2} f(y')dx$. The extremals are straight lines. Thus $y(x,a,b) = ax + b, \quad y_a = x, y_b = 1, \quad \Delta(x, x_1) = \begin{vmatrix} x & 1 \\ x_1 & 1 \end{vmatrix} = x - x_1$. Therefore $\Delta(x, x_1) \neq 0$ for all $x \neq x_1$ which confirms, by (24) that no extremal contains any conjugate points of any of its points, P_1. i.e. the 1-parameter subfamily of extremals through P_1 is the pencil of straight lines through P_1 which has no envelope.

(ii) The minimal surface of revolution functional $J[y] = \int_{x_1}^{x_2} y\sqrt{1 + (y')^2}dx$. The extremals are the catenaries, $y(x,a,b) = a\cosh\frac{x-b}{a}$. Hence $y_a = \cosh\frac{x-b}{a}, \quad y_b = -\sinh\frac{x-b}{a}$. Set

$z = \frac{x-b}{a}$, $z_1 = \frac{x_1-b}{a}$ then

$$\Delta(x, x_1) = \begin{vmatrix} \cosh z - z \sinh z & -\sinh z \\ \cosh z_1 - z_1 \sinh z_1 & -\sinh z_1 \end{vmatrix} = \sinh z \cosh z_1 - \cosh z \sinh z_1 + (z - z_1) \sinh z \sinh z_1.$$

Hence for $z_1 \neq 0$ (i.e. $x_1 \neq b$) equation (24): $\Delta(x^C, x_1) = 0$, for the conjugate point, becomes

(25)

$$\coth z^C - z^C = \coth z_1 - z_1.$$

For a given $z_1 \neq 0$, there is exactly one z^C of the opposite sign satisfying (25) (See graph below).

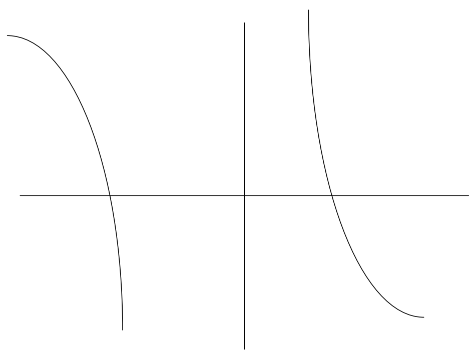


Figure 16: Graph of coth x - x.

Hence if $P_1(x_1, y_1)$ is on the descending part of the catenary, then there is exactly one conjugate point, P_1^C on the ascending part. Furthermore, the tangent line to the catenary at P_1 has equation

$$y - a\cosh\frac{x_1 - b}{a} = (\sinh\frac{x_1 - b}{a})(x - x_1).$$

Hence this tangent line intersects the x-axis at $x_t = x_1 - a\coth\frac{x_1-b}{a}$. Similarly the tangent at P_1^C intersect the x-axis at $x_t^C = x^C - a\coth\frac{x^C-b}{a}$. Subtract and use (25), obtaining $x_t = x_t^C$. Hence the two tangent lines intersect on the x-axis. This justifies Lindelöf's construction of the conjugate point (1860).

16 The Imbedding Lemma.

Lemma 16.1. *For the extremal,* $E_{\widehat{P_1P_2}}$ *of* $J[y] = \int_{x_1}^{x_2} f(x, y, y')dx$*, assume:*

(i) $f_{y'y'}(x, y, y') \neq 0$ *along* $E_{\widehat{P_1P_2}}$ *(i.e. condition* L' *if* $E_{\widehat{P_1P_2}}$ *provides a minimum).*

(ii) $E_{\widehat{P_1P_2}}$ *is free of conjugate points,* P_1^c *of* P_1 *(i.e. condition* J'*).*

Then there exist a simply-connected region F *having* $E_{\widehat{P_1P_2}}$ *in its interior, also a 1-parameter family* $\{y(x,t)\Big| -\epsilon \leq t \leq \epsilon\}$ *of smooth extremals that cover* F *simply and completely, and such that* $E_{\widehat{P_1P_2}}$ *itself is given by* $y = y(x, 0)$*.*

Proof : Since $f_{y'y'}(x, y, y') \neq 0$ on $E_{\widehat{P_1P_2}}$ we have $f_{y'y'}(x, y, y') \neq 0$ throughout some neighborhood of the set of elements, (x, y, y') belonging to $E_{\widehat{P_1P_2}}$. In particular $E_{\widehat{P_1P_2}}$ can be extended beyond its end points, P_1P_2 and belongs to a 2-parameter family of smooth extremals, $y = y(x, a, b)$ as in §15 part **1.** Further, if $E_{\widehat{P_1P_2}}$ is given by $y = y(x, a_0, b_0)$, and if we set

$$\Delta(x, x_1) \underset{def}{=} \begin{vmatrix} y_a(x, a_0, b_0) & y_b(x, a_0, b_0) \\ y_a(x_1, a_0, b_0) & y_b(x_1, a_0, b_0) \end{vmatrix}$$

we may assume, as in §15 **1.** that

$$\Delta'(x_1, x_1) = \begin{vmatrix} y'_a(x_1, a_0, b_0) & y'_b(x_1, a_0, b_0) \\ y_a(x_1, a_0, b_0) & y_b(x_1, a_0, b_0) \end{vmatrix} \neq 0.$$

Also assumption (ii) implies, by §15 part **2.** that

$$\Delta(x, x_1) \underset{def}{=} \begin{vmatrix} y_a(x, a_0, b_0) & y_b(x, a_0, b_0) \\ y_a(x_1, a_0, b_0) & y_b(x_1, a_0, b_0) \end{vmatrix} \neq 0$$

on $E_{\widehat{P_1P_2}}$, even extended some beyond P_2. except at $x = x_1$: $\quad \Delta(x_1, x_1) = 0$.

Hence if we set

$$u(x) \underset{def}{=} \Delta(x, x_1) = ky_a(x, a_0, b_0) + \ell y_b(x, a_0, b_0)$$

where $\begin{cases} k =, & y_b(x_1, a_0, b_0) \\ \ell =, & y_a(x_1, a_0, b_0 \end{cases}$ then

$$\begin{cases} u(x_1) = 0; \quad u(x) \neq 0 \text{ (say } > 0) & \text{for } x_1 \leq x \leq x_2 + \delta_2 \quad (\delta_2 > 0) \\ u' \neq 0, \Rightarrow u(x) \neq 0, \text{ (say } > 0) & \text{for } x_1 - \delta_1 \leq x \leq x_1 + \delta_1 \quad (\delta_1 > 0). \end{cases}$$

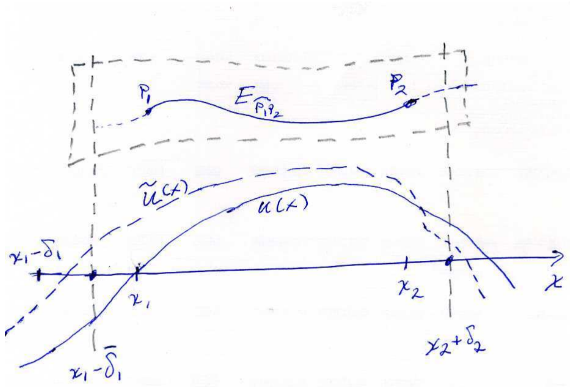
Now choose $\tilde{k}, \tilde{\ell}$ so close to k, ℓ respectively that settin

$$\tilde{u}(x) \overset{def}{=} \tilde{k}y_a(x, a_0, b_0) + \tilde{\ell}y_b(x, a_0, b_0),$$

we still have

$$\begin{cases} \tilde{u}(x) \neq 0 & \text{for } x_1 \leq x \leq x_2 + \delta_2 \quad (\delta_2 > 0) \\ \tilde{u}'(x) \neq 0 & \text{for } x_1 - \delta_1 \leq x \leq x_1 + \delta_1 \quad (\delta_1 > 0). \end{cases}$$

and such that $\tilde{u}(x_1)$ has opposite sign from $\tilde{u}(x_1 - \delta_1)$ and $u(x_1 - \delta_1)$. Then $\tilde{u}(x) = 0$ exactly once in $x_1 - \delta_1 \leq x \leq x_2 + \delta_2$, <u>before</u> x_1. Hence $\tilde{u}(x) \neq 0$ in (say) $x_1 - \overline{\delta}_1 \leq x \leq x_2 + \delta_2 (0 < \overline{\delta}_1 < \delta_1)$ (see the figure below).



With these constants, $\tilde{k}, \tilde{\ell}$, construct the 1-parameter subfamily of extremals $y(x,t) \stackrel{def}{=} y(x, a_0 + \tilde{k}t, b_0 + \tilde{\ell}t), \quad x_1 - \overline{\delta}_1 \leq x \leq x_2 + \delta_2$. Then

$$\begin{aligned} y(x,0) &= \quad y(x, a_0, b_0) \qquad \text{gives } E_{\widehat{P_1P_2}} \text{ (extended beyond } P_1, P_2). \\ y_t(x,0) &= \quad \tilde{k}y_a(x, a_0, b_0) + \tilde{\ell}y_b(x, a_0, b_0) \qquad = \tilde{u}(x). \end{aligned}$$

Hence $y_t(x, 0) \neq 0$, $x_1 - \delta_1 \leq x \leq x_2 + \delta_2$, and therefore by the continuity of y_a, y_b we have $y_t(x,t) \neq 0$, in the simply-connected region $x_1 - \widetilde{\delta}_1 \leq x \leq x_2 + \widetilde{\delta}_2, -\epsilon \leq t \leq \epsilon$ for $\tilde{\delta}_1, \tilde{\delta}_2$ sufficiently small. Now consider the region, F in the (x, y) plane bounded by

$$\left.\begin{array}{ll} x = & x_1 - \widetilde{\delta}_1 \\ x = & x_2 - \widetilde{\delta}_2 \end{array}\right\} \left.\begin{array}{l} y = y(x, -\epsilon) \\ y = y(x, \epsilon) \end{array}\right\};$$

Since $y_t(x,t) > 0$ for any x in $x_1 - \widetilde{\delta}_1 \leq x \leq x_2 + \widetilde{\delta}_2$, and t in $-\epsilon \leq t \leq \epsilon$ (or < 0 in this region), it follows that the lower boundary, $y(x, -\epsilon)$ will sweep through F, covering it completely and simply, as $-\epsilon$ is replaced by t and t is made to go from $-\epsilon$ to ϵ.

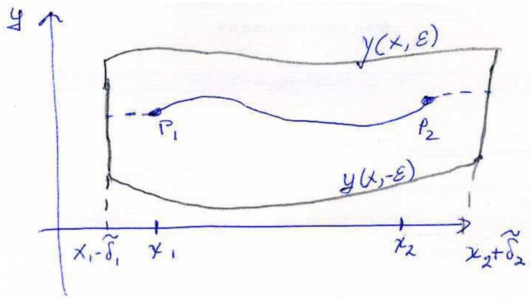


Figure 17: E from P_1 to P_2 imbedded in the Field F

q.e.d.

Note: For the slope function, $p(x,y) = y'(x,t)$ of the field, $\mathcal{F}$ just constructed (where $t = t(x,y)$ is the parameter value corresponding to (x,y) of F - i.e. $y(x,t)$ passes through (x,y)), continuity follows from the continuity and differentiability properties of the 2-parameter family $y(x,a,b)$ as well as from relations $a = a_0 + \widetilde{k}t, b = b_0 + \widetilde{\ell}t$ figuring in the selection of the 1-parameter family, $y(x,t)$.

17 The Fundamental Sufficiency Lemma.

Theorem 17.1. ***(a) Weierstrass' Theorem*** *Assume $E_{\widehat{P_1P_2}}$, a class C^2 extremal of $J[y] = \int_{x_1}^{x_2} f(x, y, y')dx$ can be imbedded in a simply-connected field, $\mathcal{F} = (F; \{y(x,t)\})$ of class C^2 extremals such that condition F (see §14) is satisfied, and let $p(x,y) = y'(x,t)$ be the slope function of the field. Then for every curve $C_{\widehat{P_1P_2}} : y = Y(x)$ of a sufficiently small strong neighborhood of $E_{\widehat{P_1P_2}}$ and still lying in the field region F,*

$$J[C_{\widehat{P_1P_2}}] - J[E_{\widehat{P_1P_2}}] = \int_{\substack{x_1 \\ C_{\widehat{P_1P_2}}}}^{x_2} \mathcal{E}(x, Y(x), p(x, Y(x)), Y'(x))dx \tag{26}$$

(b) Fundamental Sufficiency Lemma *Under the assumptions of (a) we have:*

(i) $\mathcal{E}(x, y, p(x,y), y') \geq 0$ at every (x, y) of F and for every y', implies that $E_{\widehat{P_1P_2}}$ gives a strong relative minimum for $J[y]$.

(ii) If > 0 holds in (i) for all $y' \neq p(x,y)$, the minimum is proper.

Proof : (a)

$$J[C_{\widehat{P_1P_2}}] - J[E_{\widehat{P_1P_2}}] \underset{by (18)}{=} J[C_{\widehat{P_1P_2}}] - I^*[E_{\widehat{P_1P_2}}] \underset{by\ Th.(10.2)}{=} J[C_{\widehat{P_1P_2}}] - I^*(C_{\widehat{P_1P_2}}) =$$

$$\int_{\substack{x_1 \\ C_{\widehat{P_1P_2}}}}^{x_2} f(x, Y(x), Y'(x))dx - \int_{\substack{x_1 \\ C_{\widehat{P_1P_2}}}}^{x_2} \Big[f(x, Y(x), p(x, Y(x))dx + f_{y'}(x, Y(x), p(x, Y(x)) \underbrace{dy}_{=Y'(x)dx} - pdx \Big] =$$

$$\int_{\substack{x_1\\ C_{\widehat{P_1P_2}}}}^{x_2} \Big[\underbrace{f(x,Y(x),Y'(x)) - f(x,Y(x),p(x,Y(x)) - (Y'(x)-p(x,Y))f_{y'}(x,Y(x),p(x,Y(x))}_{=}\Big]dx$$

$$\int_{\substack{x_1\\ C_{\widehat{P_1P_2}}}}^{x_2} \overbrace{\mathcal{E}(x,Y(x),p(x,Y(x)),Y'(x))}\,dx.$$

Proof : (b) Hence if $\mathcal{E}(x,Y(x),p(x,Y(x)),Y'(x)) \geq 0$ for all triples (x,y,p) near the elements, $(x,y_0(x),y_0'(x))$ of $E_{\widehat{P_1P_2}}$ and for all Y' (in other words if condition W_b of §14 holds) then $J[C_{\widehat{P_1P_2}}] - J[E_{\widehat{P_1P_2}}] \geq 0$ for all $C_{\widehat{P_1P_2}}$ in a sufficiently small strong neighborhood of $E_{\widehat{P_1P_2}}$; while if $\mathcal{E}(x,Y(x),p(x,Y(x)),Y'(x)) > 0$ for all $Y' \neq p$ then $J[C_{\widehat{P_1P_2}}] - J[E_{\widehat{P_1P_2}}] > 0$ for all $C_{\widehat{P_1P_2}} \neq E_{\widehat{P_1P_2}}$. **q.e.d.**

Corollary 17.2. *Under the assumptions of (a) above we have:*

(i) $\mathcal{E}(x,Y(x),p(x,Y(x)),Y'(x)) \geq 0$ *at every* (x,y) *of* F*, and for every* Y' *in some* ϵ *neighborhood of* $p(x,y)$ *implies that* $E_{\widehat{P_1P_2}}$ *gives a weak minimum.*

(ii) if > 0 *instead of* ≥ 0 *for every* $Y' \neq p(x,y)$ *then the minimum is proper.*

18 Sufficient Conditions.

We need only combine the results of §16 and §17 to obtain sets of sufficient conditions:

Theorem 18.1. *Conditions $E\&L'\&J'$ are sufficient for a weak, proper, relative minimum, i.e $J[E_{\overset{\frown}{P_1P_2}}] < J[C_{\overset{\frown}{P_1,P_2}}]$ for all admissible (i.e. class-$\widetilde{C}^1$) curves, $C_{\overset{\frown}{P_1,P_2}}$ of a sufficiently small weak neighborhood of $E_{\overset{\frown}{P_1P_2}}$.*

Proof : First $E\&L'\&J'$ imply F, by the imbedding lemma of §16. Next, since $f_{y'y'}(x,y,y')$ is continuous and the set of elements $(x, y_0(x), y_0'(x))$ belonging to $E_{\overset{\frown}{P_1P_2}}$ is a compact subset of $\mathbb{R}^3$, L' implies that $f_{y'y'}(x,y,y') > 0$ holds for all triples (x,y,y') within an ϵ neighborhood of the locus of triples $\{(x, y_0(x)y_0'(x))\Big| x_1 \leq x \leq x_2\}$ of $E_{\overset{\frown}{P_1P_2}}$. Hence by (20) it follows that $\mathcal{E}(x,y,y',Y') > 0$ for all (x,y,y',Y') such that $Y' \neq y'$ and sufficiently near to quadruples $(x, y_0(x), y_0'(x), y_0'(x))$ determined by $E_{\overset{\frown}{P_1P_2}}$[25]. Hence the hypothesis to the corollary of (17.2) are satisfied, proving the theorem.

q.e.d.

Example: In this example we show that conditions $E\&L'\&J'$, just shown to be sufficient for a weak minimum, are not sufficient for a strong minimum. We examined the example

$$J[y] = \int_{(0,0)}^{(1,0)} (y')^2 + (y')^3 dx$$

in §13 where it was seen that the extremal $E_{\overset{\frown}{P_1P_2}} : y \equiv 0 (0 \leq x \leq 1)$ satisfies $E\&L'\&J'$, and gives a weak relative minimum (in and $\epsilon = 1$ weak neighborhood), but does not give a strong relative minimum. We will give another example in the next section.

[25]Condition C is not needed here except locally - Y' near y_0'- where it holds because R' is open.

Theorem 18.2. *An extremal,* $E_{\widehat{P_1P_2}} : y = y_0(x)$ *of* $J[y] = \int_{x_1}^{x_2} f(x, y, y')dx$ *furnishes a strong relative minimum under each of the following sets of sufficient conditions:*

(a) $E\&L'\&J'\&W_b$ *(If* W_b' *holds, then the minimum is proper).*

(b) $E\&L_b'\&J'$ *provided the region* R' *of admissible triples,* (x, y, y') *satisfies condition* C *(see §14). The minimum is proper.*

Proof : (a) First $E\&L'\&J' \Rightarrow F$, by the embedding lemma. Now condition W_b (or W_b' respectively) establishes the conclusion because of the fundamental sufficiency lemma, (17).

(b) If R satisfies condition C, then by (20) we have

$$\mathcal{E}(x, y, y', Y') = \frac{(Y' - y')^2}{2} f_{y'y'}(x, y, y' + \theta(Y' - y')).$$

This proves that $L_b' \Rightarrow W_b'$. Also $L_b' \Rightarrow L'$. Hence the hypotheses $E\&L'\&J'\&W_b$ of part (a) holds. Therefore y provides a strong, proper minimum. **q.e.d.**

Regular Problems: A minimum problem for $J[y] = \int_{x_1}^{x_2} f(x, y, y')dx$ is regular if and only if[26]

$$\begin{cases} (i), & \text{The region} R' \text{ of admissible triples satisfies condition } C \text{ of §14.} \\ (ii), & f_{y'y'} > 0 \text{ (or } < 0), \text{ for all } (x, y, y' \in R'. \end{cases}$$

Recall the following consequences of (ii):

1. A minimizing arc $E_{\widehat{P_1P_2}}$ cannot have corners, all $E_{\widehat{P_1P_2}}$ are smooth, (c.f. 4.3).

2. All extremals $E_{\widehat{P_1P_2}}$ are of class C^2, (c.f. 4.4).

[26] The term was used earlier to cover just (ii).

Now part (b) of the last theorem, the fact that $L' \Rightarrow L'_b$ holds in the case of a regular problem, (11.3) and (13.1) proves the following:

Theorem 18.3. *For a regular problem, $E\&L'\&J'$ are a set of conditions that are both necessary and sufficient for a strong proper relative minimum provided that where $E_{\widehat{P_1P_2}}$ meets the envelope G there is a branch of G going "back" to P_1.*

If the proviso does not hold, the necessity of J' must be looked at by other methods. See §21.

The example at the beginning of this section, and an example in §19 show that $E\&L'\&J'$ is not sufficient for a strong relative minimum if the problem is not regular.

19 Some more examples.

(1) An example in which the strong, proper relative minimizing property of extremals follows directly from $L'_b(\Rightarrow W'_b)$ and from the imbeddability condition, F, which in this case can be verified by inspection.

The variational problem we will study is

$$J[y] = \int_{x_1}^{x_2} \frac{\sqrt{1+(y')^2}}{y^2} dx; \quad (x, y) \text{ in the upper half-plane } y > 0.$$

Here $f_y = \frac{-\sqrt{1+(y')^2}}{y^2}$; $f_{y'} = \frac{y'}{y\sqrt{1+(y')^2}}$; $f_{y'y'}(x, y, y') = \frac{1}{y(\sqrt{1+(y')^2})^3} > 0$ (therefore this is a regular problem). The Euler-Lagrange equation becomes

$$0 = 1 + yy'' + (y')^2 = 1 + (\frac{1}{2}y^2)''$$

Therefore extremals are semicircles $(y - c_1) + y^2 = c_2^2; \quad y > 0$.

Now for any P_1, P_2 in the upper half-plane, the extremal (circle) arc $E_{\widehat{P_1P_2}}$ can clearly be imbedded in a simply-connected field of extremals, namely concentric semi-circles. Further, since $f_{y'y'}(x, y, y') > 0$ holds for all triples, (x, y, y') such that $y > 0$, it follows from (20) that $\mathcal{E}(x, y, y', Y') > 0$ if $Y' \neq y'$ for all triples $(x, y, y'$ such that $y > 0$. Hence by the fundamental sufficiency lemma, (17) $E_{\widehat{P_1P_2}}$ gives a strong proper minimum.

Jacobi's condition, J even J' is of course satisfied, being necessary, but we reached our conclusion without using it. The role of J' in proving sufficient conditions was to insure imbeddability, (condition F), which in this example was done directly.

2. We return to a previous example, $J[y] = \int_{x_1}^{x_2} (y'+1)^2(y')^2 dx$ (see the last problem in §8). The Euler-Lagrange equation leads to y' =constant, hence the extremals are straight lines, $y = mx + b$, or polygonal trains. For the latter recall form §8 that at corners, the two slopes must be either $m_1 = y_0'\Big|_{c^-} = 0 \quad m_2 = y_0'\Big|_{c^+} = 1$ or $m_1 = 1 \quad m_2 = 0$. Next: J' is satisfied, since no envelope exist. Next: $f_{y'y'}(x, y, y') = 2(6(y')^2 + 6y' + 1) = 12(y' - a_1)(y' - a_2)$ where $a_1 = \frac{1}{6}(-3 - \sqrt{3}) \approx -0.7887, \quad a_2 = \frac{1}{6}(-3 + \sqrt{3}) \approx -0.2113$. Hence the problem is not regular. The extremals are given by $y = mx + b$. They satisfy condition L' for a minimum if $m < a_1$, or $m > a_2$ and for a maximum if $a_1 < m < a_2$.

Next: $\mathcal{E}(x, y, y', Y') = (Y' - y')^2[(Y')^2 + 2(y'+1)Y' + 3(y')^2 + 4y' + 1]$; The discriminant of the quadratic $[(Y')^2 + 2(y'+1)Y' + 3(y')^2 + 4y' + 1]$ in Y' is $-y'(y'+1)$. Hence (for $Y' \neq y'$) $\mathcal{E} > 0$ if $y' > 0$ or $y' < -1$, while $\mathcal{E}$ can change sign if $-1 < y' < 0$. Hence the extremal $y = mx + b$ with $-1 < m < 0$ cannot give a strong extremum.

We now consider the smooth extremals $y = mx + b$ by cases, and apply the information just developed.

(i) $\underline{m < -1}$ or $\underline{m > 0}$: The extremal $y = mx + b$ can be imbedded in a 1−parameter family (viz. parallel lines), also $\mathcal{E} > 0$ for $Y' \neq y'$, hence $y = mx + b$ furnishes a <u>strong minimum</u> for $J[y]$ if it connects the given end points.

(ii) $\underline{-1 < m < a_1}$ or $\underline{a_2 < m < 0}$: Since $E \& J' \& L'$ (for a minimum) hold for $J[y]$, such an extremal, $y = mx + b$ furnishes a <u>weak</u>, proper relative <u>minimum</u> for $J[y]$; <u>not</u> however a strong relative minimum, since $J[mx + b] > 0$ while $J[C_{\widehat{p_1 P_2}}] = 0$ for a path joining P_1 to P_2

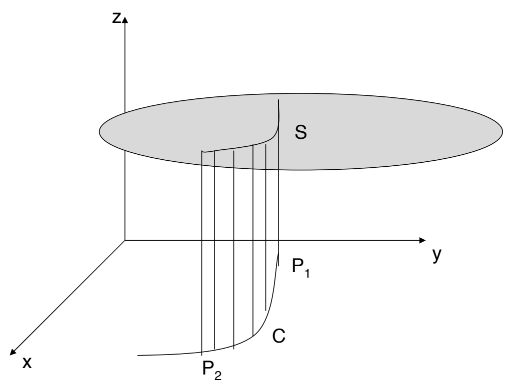
that consists of segments of slopes 0 and -1. Hence, once again, for a problem that is not regular, conditions $E\&J'\&L'$ are not sufficient to guarantee a strong minimum.

(iii) $\underline{a_1 < m < a_2}$: Conditions $L' : f_{y'y'}(x, y, y') < 0$ for a maximum holds, and in fact $E\&J'\&L'$ for a maximum guarantee that $y = mx + b$ furnishes a weak maximum. Not, however a strong maximum. This is clear by looking at $J[y]$ evaluated using a steep zig-zag path from P_1 to P_2. or from the fact that $\mathcal{E}(x, y, y', Y')$ can change sign as Y' varies.

(iv) $\underline{m = -1 \text{ or } a_1 \text{ or } a_2 \text{ or } 0}$: Exercise.

3. All problems of the type $J[y] = \int_{x_1}^{x_2} \eta(x, y)ds$, $\eta(x, y) > 0$, are regular since $f_{y'y'}(x, y, y') = \frac{\eta(x,y)}{(1+(y')^2)^{\frac{3}{2}}} > 0$ for all (x, y, y') and condition C is satisfied. We previously interpreted all such variational problems as being equivalent to minimizing the time it takes for light to travel along a path.

We may also view η as representing a surface S given by $z = \eta(x, y)$. Then the cylindrical surface with directrix $C_{\widehat{p_1P_2}}$ and vertical generators between the (x, y) plane and S has area $J[C_{\widehat{p_1P_2}}]$ (see figure below)



4. An example of a smooth extremal that cannot be imbedded in any family of smooth extremals. For $J[y] = \int_{(x_1,0)}^{(x_2,0)} y^2 dx$, the only smooth extremals are given by $\frac{d}{dx} f_{y'} - f_y = 0$ i.e. y=0. This gives just one smooth extremal, which obviously gives a strong minimum. Hence condition F is not necessary for a strong minimum.

20 The Second Variation. Other Proof of Legendre's Condition.

We shall exploit the necessary conditions $\left.\frac{d^2 J[y_0+\epsilon\eta]}{d\epsilon^2}\right|_{\epsilon=0}$ [27] ≥ 0 for y_0 to provide a minimum for the functional $J[y] = \int_{x_1}^{x_2} f(x,y,y')dx$. As before $\eta(x)$ is in the same class as $y_0(x)$ (i.e. $\widetilde{C}^1$ or C^1) and $\eta(x_1) = \eta(x_2) = 0$.

First

$$\varphi''(\epsilon) \underset{def}{=} \frac{d^2 J[y_0+\epsilon\eta]}{d\epsilon^2} = \frac{d}{dx}\int_{x_1}^{x_2} [f_y(x, y_0+\epsilon\eta, y_0'+\epsilon\eta')\eta + f_{y'}(x, y_0+\epsilon\eta, y_0'+\epsilon\eta')\eta']dx =$$

$$= \int_{x_1}^{x_2} [f_{yy}(x, y_0+\epsilon\eta, y_0'+\epsilon\eta')\eta^2 + 2f_{yy'}(x, y_0+\epsilon\eta, y_0'+\epsilon\eta')\eta\eta' + f_{y'y'}(x, y_0+\epsilon\eta, y_0'+\epsilon\eta')(\eta')^2]dx.$$

Therefore $\varphi''(0) =$

(27)

$$\int_{x_1}^{x_2} [\underbrace{f_{yy}(x, y_0(x), y_0'(x))\eta^2(x) + 2f_{yy'}(x, y_0(x), y_0'(x))\eta(x)\eta'(x) + f_{y'y'}(x, y_0(x), y_0'(x))\eta'(x)^2}_{\text{set this } = 2\Omega(x,\eta,\eta')}]dx.$$

Hence for $y_0(x)$ to minimize $J[y] = \int_{x_1}^{x_2} f(x,y,y')dx$, among all $y(x)$ joining P_1 to P_2 and of class $\widetilde{C}^1$ (or of class C^1), it is necessary that for all $\eta(x)$ of the same class on $[x_1, x_2]$ such that $\eta(x_i) = 0, i = 1, 2$ we have

$$\int_{x_1}^{x_2} \Omega(x, \eta, \eta')dx \geq 0 \tag{28}$$

Set

- $P(x) \underset{def}{=} f_{yy}(x, y_0(x), y_0'(x))$

[27] This is called the second variation of J.

- $Q(x) \underset{def}{=} f_{yy'}(x, y_0(x), y_0'(x))$
- $R(x) \underset{def}{=} f_{y'y'}(x, y_0(x), y_0'(x))$

Since f is of class C^3 at all admissible (x, y, y')[28] and $y_0(x)$ at least of class $\widetilde{C}^1$ on $[x_1, x_2]$, it follows that P, Q, R are at least of class $\widetilde{C}^1$ on $[x_1, x_2]$, hence bounded.

$$\Big|P(x)\Big| \le p, \quad \Big|Q(x)\Big| \le q, \quad \Big|R(x)\Big| \le r \quad \text{for all } x \in [x_1, x_2]$$

We now use (28) to give another proof of Legendre's necessary condition:

$$f_{y'y'}(x, y_0(x), y_0'(x)) \ge 0, \quad \text{for all } x \in [x_1, x_2]$$

Proof : Assume the condition is not satisfied, say it fails at c i.e. $R(x) < 0$. Then, since $R(x) \in \widetilde{C}^1$, there is an interval $[a, b] \subset [x_1, x_2]$ that contains c, <u>possibly at an end point</u>, such that $R(x) \le -k < 0$, for all $x \in [a, b]$. Now set

$$\widetilde{\eta}(x) \underset{def}{=} \begin{cases} \sin^2 \frac{\pi(x-a)}{b-a}, & \text{on } [a, b] \\ 0, & \text{on } [x_1, x_2] \quad \backslash \quad [a, b] \end{cases}$$

Then it is easy to see that $\widetilde{\eta}(x) \in \widetilde{C}^1$ and

$$\widetilde{\eta}^{pr}(x) \underset{def}{=} \begin{cases} \frac{\pi}{b-a} \sin \frac{2\pi(x-a)}{b-a}, & \text{on } [a, b] \\ 0, & \text{on } [x_1, x_2] \quad \backslash \quad [a, b] \end{cases}$$

Hence

$$\varphi''_{\widetilde{\eta}}(0) = \int_a^b (P\widetilde{\eta}^2 + 2Q\widetilde{\eta}\widetilde{\eta}' + R(\widetilde{\eta}')^2) dx$$

[28] Recall being an admissible triple means the triple in in the region where f is C^3.

$$\leq p(b-a) + 2q\frac{\pi}{b-a}(b-a) + (-k)\frac{\pi^2}{(b-a)^2}\overbrace{\int_a^b sin^2\Big(\frac{2\pi(x-a)}{b-a}\Big)dx}^{=\frac{1}{2}(b-a)}$$

$$= p(b-a) + 2q\pi - \frac{k\pi^2}{2(b-a)}$$

and the right hand side is < 0 for $b-a$ sufficiently small, which is a contradiction. **q.e.d.**

21 Jacobi's Differential Equation.

In this section we confine attention to extremals $E_{\overset{\frown}{P_1P_2}}$ of $J[y] = \int_{x_1}^{x_2} f(x,y,y')dx$ along which $f_{y'y'}(x,y,y') \neq 0$ holds. Hence if $E_{\overset{\frown}{P_1P_2}}$ is to minimize, $f_{y'y'}(x,y,y') > 0$ must hold.

1. Consider condition (28), $\int_{(x_1,0)}^{(x_2,0)} \Omega(x,\eta(x),\eta'(x)dx \geq 0$ for all admissible $\eta(x)$ such that $\eta(x_1) = \eta(x_2) = 0$, where

$2\Omega(x,\eta,\eta') = f_{yy}(x,y_0(x),y_0'(x))\eta^2(x)+2f_{yy'}(x,y_0(x),y_0'(x))\eta(x)\eta'(x)+f_{y'y'}(x,y_0(x),y_0'(x))\eta'(x)^2 =$

$P(x)\eta^2+2Q(x)\eta\eta'+R(x)(\eta')^2, \quad R(x) > 0$ on $[x_1,x_2]$. Setting $J^*[\eta] \underset{def}{=} \int_{x_1}^{x_2} \Omega(x,\eta(x),\eta'(x))dx$ we see that condition (28) (which is necessary for y_0 to minimize $J[y]$) implies that the functional $J^*[\eta]$ must be minimized among admissible $\eta(x)$, (such that $\eta(x_1) = \eta(x_2) = 0$) by $\eta_0(x) \equiv 0$ which furnishes the minimum value $J^*[\eta_0] = 0$[29]

This suggest looking at the minimizing problem for $J^*[\eta]$ in the (x,η)-plane, where the end points are $P_1^* = (x_1,0), P_2^* = (x_2,0)$. The base function, $f(x,y,y')$ is replaced by Ω. The associated problem is regular since $\Omega_{\eta'\eta'} = f_{y'y'}(x,y,y') = R(x) > 0$.

Hence: (i) No Minimizing arc, $\eta(x)$ can have corners

and (ii)every extremal for J^* is of class C^2 on $[x_1,x_2]$.

We may set up the Euler-Lagrange equation for J^*.

$$\frac{d}{dx}(R\eta') + (Q' - P)\eta = 0$$

[29]It is amusing to note that if there is an $\widetilde{\eta}$ such that $J^*[\widetilde{\eta}] < 0$ then there is no minimizing η for J^*. This follows from $J^*[k\eta] = k^2 J^*[\eta]$.

i.e.

$$\eta'' + \frac{R'}{R}\eta' + \frac{Q' - P}{R}\eta = 0 \tag{29}$$

This linear, homogeneous 2^{nd} order differential equation for $\eta(x)$ is the *Jacobi differential equation.* It is clearly satisfied by $\eta = 0$.

2.a Since $2\Omega = P(x)\eta^2 + 2Q(x)\eta\eta' + R(x)(\eta')^2$ is homogeneous of degree 2 in $\eta.\eta'$ we have (by Euler's identity for homogeneous functions)

$$2\Omega = \eta\Omega_\eta + \eta'\Omega_{\eta'}.$$

Hence

$$J^*[\eta] = \int_{(x_1,0)}^{(x_2,0)} \Omega dx = \frac{1}{2}\int_{(x_1,0)}^{(x_2,0)} (\eta\Omega_\eta + \eta'\Omega_{\eta'})dx = \frac{1}{2}\int_{(x_1,0)}^{(x_2,0)} \eta(\Omega_\eta - \frac{d}{dx}\Omega_{\eta'})dx$$

(we used integration by parts on $\eta'\Omega_{\eta'}$ which is allowed since we are assuming $\eta \in C^2$. We also used $\eta(x_i) = 0, i = 1, 2$). This shows that for any class-C^2 solution, $\eta(x)$ of the Euler-Lagrange equation, $\frac{d\Omega_{\eta'}}{dx} - \Omega_\eta = 0$ which vanishes at the end points (i.e. any class-C^2 solution of Jacobi's differential equation) we have $J^*[\eta] = 0$.

2.b Let $[a, b]$ be any subinterval of $[x_1, x_2] : x_1 \leq a \leq b \leq x_2$. If $\eta(x)$ satisfies Jacobi's equation on $[a, b]$ and $\eta(a) = \eta(b) = 0$ then it follows that, as in **2.a** that $\int_a^b \Omega dx = \frac{1}{2}\int_a^b \eta(\Omega_\eta - \frac{d}{dx}\Omega_{\eta'})dx = 0$.

Hence if we define $\widetilde{\eta}(x)$ on $[x_1, x_2]$ by

$$\widetilde{\eta}(x) \underset{def}{=} \begin{cases} \eta(x), & \text{on } [a,b] \\ 0, & \text{on } [x_1,x_2]\diagdown[a,b] \end{cases}$$

then $J^*[\widetilde{\eta}] = 0$.

3. We now prove

Theorem 21.1 (Jacobi's Necessary Condition, first version)**.** *If Jacobi's differential equation has a non-trivial solution (i.e. not identically 0) $\mu(x)$ with $\mu(x_1)$ and $\mu(\widetilde{x}_1) = 0$ where $x_1 < \widetilde{x}_1 < x_2$, then there exist $\eta(x)$ such that $\varphi'' < 0$. Hence by (28) $y_0(x)$ cannot furnish even a weak relative minimum for $J[y]$*

Proof : If such a solution $\mu(x)$ exist set $\overline{\eta}(x) = \begin{cases} \mu(x), & \text{on } [x_1,\widetilde{x}_1] \\ 0, & \text{on } [\widehat{x}_1, x_2] \end{cases}$; Then by **2.b** $J^*[\overline{\eta}] = 0$. But $\overline{\eta}$ has a corner at $\widetilde{x}_1$. To see this note that $\overline{\eta}'(\widetilde{x}_1^-) = \mu'(\widetilde{x}_1) \neq 0$. (This is the case since $\mu(x)$ is a <u>non-trivial</u> solution of the Jacobi equation, and $\mu(\widetilde{x}_1) = 0$, hence we cannot have $\mu'(\widetilde{x}_1) = 0$ by the existence and <u>uniqueness</u> theorem for a 2nd order differential equation of the type $y'' = F(x,y,y')$.) But $\overline{\eta}'(\widetilde{x}_1^+) = 0$ so there is a corner at $\widetilde{x}_1$. But by remark (i) in **1.** $\overline{\eta}$ <u>cannot</u> be a minimizing arc for $J^*[\eta]$. Therefore there must exist an admissible function, $\eta(x)$ such that $J^*[\eta] < J^*[\overline{\eta}] = 0$, which means that the necessary condition, $\varphi''(0) \geq 0$ is not satisfied for η. **q.e.d.**

4. We next show that $\widetilde{x}_1(\neq x_1)$ is a zero of a non-trivial solution, $\mu(x)$ of Jacobi's differential equation if and only if the point $\widetilde{P}_1 = (\widetilde{x}_1, y_0(\widetilde{x}_1))$ of the extremal, $E_{\widehat{P_1P_2}}$ of $J[y]$ is conjugate to $P_1 = (x_1, y_1)$ in the sense of §15.

There are three steps:

Step a. Jacobi's differential equation is linear and homogeneous. Hence if $\mu(x)$ and $\nu(x)$ are two non-trivial solutions, both 0 at x_1 then $\mu(x) = \lambda\nu(x)$ with $\lambda \neq 0$. To see this we first note that $\mu'(x_1) \neq 0$ since we know that $\eta_0 \equiv 0$ is the unique solution of the Jacobi equation which vanishes at x_1 and has zero derivative at x_1. Same for $\nu(x)$. Now define λ by $\mu'(x_1) = \lambda\nu'(x_1)$. Set $h(x) = \mu(x) - \lambda\nu(x)$. Then $h(x)$ satisfies the Jacobi equation (since it is a linear differential equation) and $h(x_1) = h'(x_1) = 0$. Therefore $h(x) \equiv 0$.

Step b. Jacobi's differential equation is the Euler-Lagrange equation for a variational problem. Let $\{y(x, a, b)\}$ be the two parameter family of solutions of the Euler-Lagrange equation, (4). Then for all a, b the functions, $y(x, a, b)$ satisfy

$$\frac{d}{dx}[f_{y'}(x, y(x, a, b), y'(x, a, b))] - f_y(x, y(x, a, b), y'(x, a, b)) = 0$$

In particular if $\{y(x, t)\} = \{y(x, a(t), b(t))\}$ is any 1-parameter subfamily of $\{y(x, a, b)\}$ we have, for all t

$$\frac{d}{dx}[f_{y'}(x, y(x, t), y'(x, t))] - f_y(x, y(x, t), y'(x, t)) = 0$$

Assuming $y(x, t) \in C^2$ we may take $\frac{\partial}{\partial t}$. This gives:

(30)

$$\frac{d}{dx}[f_{y'y}(x, y(x, t), y'(x, t)) \cdot y_t(x, t) + f_{y'y'}(x, y(x, t), y'(x, t)) \cdot y'_t(x, t)] - (f_{yy} y_t(x, t) + f_{yy'} \cdot y'_t(x, t))$$

This equation, satisfied by $\omega(x) \underset{def}{=} y_t(x, t_0)$ for any fixed t_0, is called the equation of variation of the differential equation $\frac{d}{dx}(f_{y'}) - f_y = 0$. In fact, $\omega(x) = y_t(x, t_0)$ is the "first variation" for the solution subfamily, $\{y(x,t)\}$ of $\frac{d}{dx}(f_{y'}) - f_y = 0$, in the sense that $y(x, t_0 + dt) = y(x, t_0) + y_t(x, t_0) \cdot \Delta t +$ higher powers (Δt)

Now note that (30) is satisfied by $\omega(x) = y_t(x, t_0)$. Hence also by $y_a(x, a_0, b_0)$ and by $y_b(x, a_0, b_0)$.[30] This is Jacobi's differential equation all over again!

$$\frac{d}{dx}(Q(x)\omega(x) + R(x)\omega') - (P(x)\omega(x) + Q(x)\omega'(x) = 0.$$

Step c. Let $\{y(x,t)\}$ be a family of extremals for $J[y]$ passing through $P_1(x_1, y_1)$. By §15 part **2.**, a point $P_1^C = (x_1^C, y_1^C)$ of $E_{\widehat{P_1P_2}}$ is conjugate to P_1 if and only if it satisfies $y^C = y(x_1^C, t_0)$; $\quad y_t(x_1^C, t_0) = 0$ where $y = y(x, t_0)$ is the equation of $E_{\widehat{P_1P_2}}$. Now by **Step b.** the function $w(x) = y_t(x, t_0)$ satisfies Jacobi's differential equation, it also satisfies $y_t(x_1, t_0) = 0$ since $y(x_1, t) = y_1$ for all t, and it satisfies $y_t(x_1^C, t_0) = 0$. To sum up: If P_1^C is conjugate to P_1 on $E_{\widehat{P_1P_2}} : y = y(x, t_0)$, then we can construct a solution $\omega(x)$ of (29) that satisfies $\omega(x_1) = \omega(x_1^C) = 0$ namely $\omega(x) = y_t(x, t_0)$.

Conversely, assume that there exist $\mu(x)$ such that $\mu(x)$ is a solution of (29) satisfying $\mu(x_1) = \mu(\widetilde{x}_1) = 0$. Then by **4 a (ii)**: $y_t(x, t_0) = \lambda\mu(x)$. Hence $y_t(\widetilde{x}, t_0) = \lambda\mu(\widetilde{x}) = 0$. This implies that $\widetilde{x}_1$ gives a conjugate point $P_1^C = (\widetilde{x}_1, y(\widetilde{x}_1, t_0)$ to P_1 on $E_{\widehat{P_1P_2}}$. Thus we have proven the statement at the beginning of **Step 4.**. Combining this with **Step 3.** we have proven:

[30] The fact that if $y(x,a,b)$ is the general solution of the Euler-Lagrange equation then $y_a(x, a_0, b_0)$ and $y_b(x, a_0, b_0)$ are two solutions of Jacobi's equation is known as Jacobi's theorem. From §15, part **1**, we know that these two solutions may be assumed to be independent, therefore a basis for all solutions.

Theorem 21.2 (Jacobi's Necessary Condition (second version):)**.** *If on the extremal $E_{\widehat{P_1P_2}}$ given by $y = y_0(x)$ along which $f_{y'y'}(x, y, y') > 0$ is assumed, there is a conjugate point, P_1^C of P_1 such that $x_1 < x_1^C < x_2$, then y_0 cannot furnish even a weak relative minimum for $J[y]$.*

5. Some comments. a Compared with the geometric derivation of Jacobi's condition in §13 the above derivation, based on the second variation has the advantage of being valid regardless of whether or not the envelope, G of the family $\{y(x, t)\}$ of extremals through P_1 has a branch point at P_1^C pointing back toward P_1.

b Note that the function $\Delta(x, x_1)$ constructed in §15 (used to locate the conjugate points, P_1^C by solving for x^C in $\Delta(x^C, x_1) = 0$) is a solution of Jacobi's equation since it is a linear combination of the solutions $y_a(x, a_0, b_0)$ and $y_b(x, a_0, b_0)$. Also $\Delta(x_1, x_1) = 0$.

c Note the two aspects of Jacobi's differential equation: It is (i) the Euler-Lagrange equation for the associated functional $J^*(\eta)$ of $J[y]$, and (ii) the equation of variation of the Euler-Lagrange equation for $J[y]$ (see **1.** and **4 b.** above).

6. The stronger Jacobi Condition, $J' : x_1^C \notin [x_1, x_2]$**.** Let:

- J' stand for: No x of $[x_1, x_2]$ gives a conjugate point, $P_1^C(x, y_0(x))$ of P_1 on $E_{\widehat{P_1P_2}}$.
- J_1' stand for: Every non trivial solution, $\mu(x)$ of Jacobi's equation that is $= 0$ at x_1 is free of further solutions on $[x_1, x_2]$.
- J_2' stand for: There exist a solution $\nu(x)$ of the Jacobi equation that is nowhere 0 on $[x_1, x_2]$.

a We proved in **4** **c̲** that $J' \Leftrightarrow J'_1$.

b $J'_1 \Leftrightarrow J'_2$.

We shall need Strum's Theorem

Theorem 21.3. *If $\mu_1(x), \mu_2(x)$ are non trivial solutions of the Jacobi equation such that $\mu_1(x_1) = 0$ but $\mu_2(x_1) \neq 0$, then the zeros of $\mu_1(x), \mu_2(x)$ interlace. i.e. between any two consecutive zeros of μ_1 there is a zero of μ_2 and vice versa.*

Proof : Suppose $x_1, \widetilde{x}_1$ are two consecutive zeros of μ_1 without no zero of μ_2 in $[x_1, \widetilde{x}_1]$. Set $g(x) = \frac{\mu_1(x)}{\mu_2(x)}$. Then $g(x_1) = g(\widetilde{x}_1) = 0$. Hence $g'(x^*) = 0$ at some $x^* \in (x_1, \widetilde{x}_1)$. i.e. $\mu'_1(x^*)\mu_2(x^*) - \mu_1(x^*)\mu'_2(x^*) = 0$. But then $(\mu_1, \mu'_1) = \lambda(\mu_2, \mu'_2)$ at x^*. Thus $\mu_1 - \lambda\mu_2$ is a solution of the Jacobi equation and has 0 derivative at x^*. Therefore $\mu_1 - \lambda\mu_2 \equiv 0$ which is a contradiction. **q.e.d.**

Proof : [of **b.**] $J'_1 \Leftrightarrow J'_2$: Let $\mu(x)$ be a solution of the Jacobi equation such that $\mu(x_1) = 0$ and μ has no further zeros on $[x_1, x_2]$. Then there is x_3 such that $x_2 < x_3$ and
$mu(x) \neq 0$ on $[x_1, x_3]$. Set $\nu(x) =$ a solution of the Jacobi equation such that $\nu(x_3) = 0$. Then by Strum's theorem $\nu(x)$ cannot be 0 anywhere on $[x_1, x_2]$.

$J'_2 \Leftrightarrow J'_1$: Let $\nu(x)$ be a solution of the Jacobi equation such that $\nu(x) \neq 0$ for all $x \in [x_1, x_2]$. Then if $\mu(x)$ is any solution of the Jacobi equation such that $\mu(x_1) = 0$, Strum's theorem does not allow for another zero of $\mu(x)$ on $[x_1, x_2]$. **q.e.d.**

Theorem 21.4 (Jacobi)**.** $J' \Rightarrow \varphi''_\eta(0) > 0$ *for all admissible η not $\equiv 0$.*

Proof : First, if $\omega(x) \in C^1$ on $[x_1, x_2]$, then

$$\int_{x_1}^{x_2} (\omega'\eta^2 + 2\omega\eta\eta')ds = \int_{x_1}^{x_2} \frac{d}{dx}(\omega\eta^2)dx = 0$$

since $\eta(x_i) = 0, i = 1, 2$. Therefore

$$(31)\quad \varphi''_\eta(0) = \int_{x_1}^{x_2} (P\eta^2 + 2Q\eta\eta' + R(\eta')^2)dx = \int_{x_1}^{x_2} [(P+\omega')\eta^2 + 2(Q+\omega)\eta\eta' + R(\eta')^2)]dx$$

for <u>any</u> $\omega \in C^1$ on $[x_1, x_2]$.

Now chose a particular ω as follows: Let $\nu(x)$ be a solution of the Jacobi equation that is $\neq 0$ throughout $[x_1, x_2]$. Such a ν exist by the hypothesis, since $J' \Rightarrow J'_2$. Set

$$\omega(x) = -Q - R\frac{\nu'}{\nu}.$$

Then, using

$$\nu'' + \frac{R'}{R}\nu' + \frac{Q' - P}{R}\nu = 0$$

(i.e. the Jacobi equation!) we have

$$\omega'(x) = -Q' - R'\frac{\nu'}{\nu} - R\frac{\nu''\nu - (\nu')^2}{\nu^2} = -P + \frac{R(\nu')^2}{\nu^2}$$

Substituting into (31) we have:

$$(32)\qquad \varphi''_\eta(0) = \int_{x_1}^{x_2} (\frac{R(\nu')^2}{\nu^2}\eta^2 - 2R\frac{\nu'}{\nu}\eta\eta' + R(\eta')^2)dx = \int_{x_1}^{x_2} \frac{R}{\nu^2}(\eta\nu' - \eta'\nu)^2 dx.$$

Since $R(x) > 0$ on $[x_1, x_2]$ and since $\nu' - \eta'\nu)^2 = \nu^4[\frac{d}{dx}(\frac{\eta}{\nu})]^2$ is not $\equiv 0$ on $[x_1, x_2]$[31]. $\varphi''_\eta(0) > 0$ now follows from (32)[32].

[31] otherwise $\eta = c\nu, \quad c \neq 0$ which is impossible since $\eta(x_1) = 0, \nu(x_2) \neq 0$.

[32] The trick of introducing ω is due to Legendre.

7. Two examples

a. $J[y] = \int_{(0,0)}^{(1,0)}((y')^2 + (y')^3)dx$ (already considered in §13 and §18. For the extremal $y_0 \equiv 0$ the first variation, $\varphi'_\eta(0) = 0$ while the second variation, $\varphi''_\eta(0) = 2\int_{x_1}^{x_2}(\eta')^2 dx > 0$ if $\eta \neq 0$. Yes as we know from §18 $y_0(x)$ does not furnish a strong minimum; thus $\varphi''_\eta(0) > 0$ for all admissible η is not sufficient for a strong minimum[33].

b. $J[y] = \int_{x_1}^{x_2} \frac{x}{1+(y')^2}dx$. This is similar to §11 example **1**, from which we borrow some calculations (the additional factor x does not upset the calculations).

- $f_{y'} = \frac{-2xy'}{(1+(y')^2)^2}$
- $f_{y'y'}(x, y, y') = 2x \cdot \frac{3(y')^2-1}{(1+(y')^2)^3}$
- $\mathcal{E} = \frac{x\cdot(y'-Y')^2}{(1+(y')^2)^2(1+(Y')^2)^2}((y')^2 - 1 + 2y'Y')$.

Hence Legendre's condition for a minimum (respectively maximum) is satisfied along an extremal if and only if $(y')^2 \geq \frac{1}{3}$ (respectively $\leq \frac{1}{3}$) along the extremal. Notice that the extremals are the solutions of $\frac{2xy'}{(1+(y')^2)^2}$ =constant, which except for $y' \equiv 0$ does not give straight lines.

Weierstrass' condition for either a strong minimum or maximum is not satisfied unless $y' \equiv 0$ (if $y' \neq 0$ then the factor $((y')^2 - 1 + 2y'Y')$ can change sign as Y' varies.)

To check the Jacobi condition, it is simplest here to set up Jacobi's differential equation. Since $P \equiv 0, Q \equiv 0$ in this case, (29) becomes $\mu'' + \frac{R'}{R}\mu' = 0$ and this does have a solution

[33] Jacobi drew the erroneous conclusion that his theorem implied that condition J' was sufficient for $E_{\widehat{P_1P_2}}$ satisfying also L' to give a strong minimum.

$\mu(x)$ that is nowhere zero. Namely $\mu(x) \equiv 1$. Hence J_2' is satisfied, and so is J' by the lemma in **6. b** above.

Remark 21.5. The point of this example is to demonstrate that Jacobi's theorem provides a convenient way to detect a conjugate point. One does not need to find the envelope (which can be quite difficult). In this example it is much easier to find a solution of the Jacobi differential equation which is never zero. On the other hand it is in general difficult to solve the Jacobi equation.

22 One Fixed, One Variable End Point.

Problem: To minimize $J[y] = \int_{x_1}^{x_2} f(x, y, y')dx$, given that $P_1 = (x_1, y_1)$ is fixed while $p_2(t) = (x_2(t0, y_2(t))$ varies on an assigned curve, $N(a < t < b)$. Some of the preliminary work was done in §9 and §12.

1. Necessary Conditions. Given P_1 fixed, $P_2(t)$ variable on a class C^1 curve N $(x = x_2(t), y = y_2(t)), a < t < b)$. Assume that the curve C from P_1 to $P_2(t_0) \in N$ minimizes $J[y]$ among all curves $y = y(x)$ joining P_1 to N. Then:

(i) C must minimize $J[y]$ also among curves joining the two fixed points P_1 and $P_2(t_0)$. Hence $C = E_{t_0}$, an extremal and must satisfy E, L, J, W.

(ii) Imbedding E_{t_0} in a $1-$ parameter family $\{y(x,t)\}$ of curves joining P_1 to points $P_2(t)$ of N near $P_2(t_0)$, say for all $t \in |t - t_0| < \epsilon$, we form $\mathcal{I}(t) \underset{def}{=} J[y(x,t)]$ and obtain, as in §9 the necessary condition, $\left.\frac{\mathcal{I}(t)}{dt}\right|_{t=t_0} = 0$ (for E to minimize), which as in §9 gives:

$$\text{Condition } \mathbb{T} : f(x_2(t_0), y_2(t_0), \underbrace{y'(x_2(t_0), t_0)}_{=p(x_2(t_0), y_2(t_0))})dx_2 + f_{y'}(x_2, y_2, p)(dY_2 - pdx_2)\Big|_{t=t_0}.$$

This is the transversality condition of §9, at $P_2(t_0)$.

(iii) If $f_{y'y'}(x, y, y') \neq 0$ along E_{t_0}, then a further necessary condition follows from the envelope theorem, (12.1). Let G be the focal curve of N, i.e. the envelope of the $1-$parameter family of extremals transversal to N[34]. Now assume that $E_{\widehat{P_1 P_2}}$ touches G at a point P^f [35] and assume further that G has at P^f a branch "toward" $P_2(t_0)$.

[34] The evolute of a curve is an example of a focal curve.

[35] P^f is called the focal point of P_2 on $E_{\widehat{P_1 P_2}}$.

Condition $\mathbb{K}$:Let P^f be as described above. Then if $P_1 \preceq P^f \prec P_2(t_0)$ on $E_{\widehat{P_1P_2}}$, then not even a weak relative minimum can be furnished by $E_{\widehat{P_1P_2}}$. That is to say a minimizing arc $E_{\widehat{P_1P_2}}$ such that $f_{y'y'}(x,y,y') \neq 0$ cannot have on it a focal point where G has a branch toward $P_2(t_0)$[36].

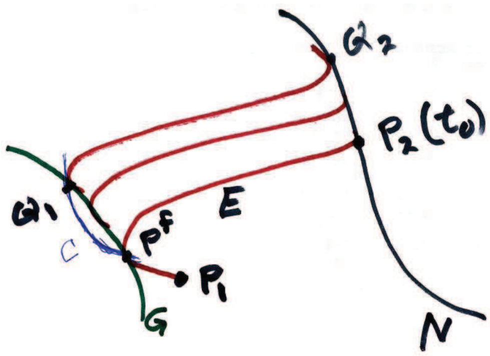
Proof of condition $\mathbb{K}$:

We have from (17) and (12.1) that

$$\mathcal{I}[E_{\widehat{Q_1Q_2}}] - \mathcal{I}[E_{\widehat{P^fP_2(t_0)}}] = I^*(N_{\widehat{P_2Q_2}}) - I^*(G_{\widehat{P^fQ_1}})$$

$$\Rightarrow \mathcal{I}[E_{\widehat{P^fP_2(t_0)}}] = \mathcal{I}(G_{\widehat{P^fQ_1}}) + \mathcal{I}[E_{\widehat{Q_1Q_2}}]$$

(see figure). Now since $f_{y'y'}(x,y,y') \neq 0$ at P^f, the extremal through p^f is unique, hence



G is not an extremal. Therefore there exist a curve $C_{\widehat{P^fQ_1}}$ such that $\mathcal{I}[C_{\widehat{P^fQ_1}}] < \mathcal{I}[G_{\widehat{P^fQ_1}}]$. Such a $C_{\widehat{P^fQ_1}}$ exist in any weak neighborhood of $G_{\widehat{P^fQ_1}}$. Therefore the curve going from P_1

[36] A second proof based on $\frac{d^2\mathcal{I}(t)}{dt^2}$, can be given that shows that if $f_{y'y'}(x,y,y') \neq 0$ on $E_{\widehat{P_1P_2}}$, then $P_1 \prec P^f \prec P_2(t_0)$ is incompatible with $E_{\widehat{P_1P_2}}$ being a minimizing arc, regardless of the branches of G

along E to P^f, followed by $C_{\widehat{P^f Q_1}}$ then by $E_{\widehat{Q_1 Q_2}}$ is a shorter path from P_1 to N. **q.e.d.**

Summarizing Necessary conditions are: E, L, J, W, T, K.

2. One End Point Fixed, One Variable, Sufficient Conditions. Here we assume N to be of class C^2 and regular (i.e. $(x')^2 + (y')^2 \neq 0$). We shall also use a slightly stronger version of condition $\mathbb{K}$ as follows. Assuming $\{y(x,t)\Big| a < t < b\}$ to be the 1−parameter family of extremals meeting N transversally, define a focal point of N to be any (x, y) satisfying, for some $t = t_f$ of (a, b) :

$$\begin{cases} y = y(x, t_f) \\ 0 = y_t(x, t_f) \end{cases}$$

The locus of focal points will then include the envelope, G of **1.** Now conditions $\mathbb{K}$ reads $E_{\widehat{P_1 P_2(t_0)}}$ shall be free of focal points.

Theorem 22.1. *Let* $E_{\widehat{P_1 P_2(t_0)}}$ *intersect* N *at* $P_2(t_0) = (x_2(t_0), y_2(t_0))$. *Then*

(i) If $E_{\widehat{P_1 P_2(t_0)}}$ *satisfies* $E, L', \mathbf{T}, \mathbf{K}$ *and* $f\Big|_{P_2(t_0)} \neq 0$, *then* $E_{\widehat{P_1 P_2(t_0)}}$ *gives a weak relative minimum for* $J[y]$.

(ii) If in addition, W_b *holds on* $E_{\widehat{P_1 P_2(t_0)}}$, *we have a strong relative minimum.*

Proof : In two steps: (i) Will show that $E_{\widehat{P_1 P_2(t_0)}}$ can be imbedded in a simply-connected field $\{y(x,t)\}$ of extremals, all transversal to N and such that $y_t(x,t) \neq 0$ on $E_{\widehat{P_1 P_2(t_0)}}$, and (ii) we will then be able to apply the same methods as in §17, and §18 to finish the sufficiency proof. (i) Sine E, L' hold there exist a two parameter family $\{y(x, a, b)\}$ of smooth extremals of which $E_{\widehat{P_1 P_2(t_0)}}$ is a member for, say, $a = a_0, b = b_0$. We may assume this family so constructed that

$$\Delta'(x_2(t_0), x_2(t_0)) = \begin{vmatrix} y'_a(x_2(t_0), a_0, b_0) & y'_b(x_2(t_0), a_0, b_0) \\ y_a(x_2(t_0), a_0, b_0) & y_b(x_2(t_0), a_0, b_0) \end{vmatrix} \neq 0$$

in accordance with (22).

We want to construct a $1-$parameter subfamily, $\{y(x,t)\} = \{y(x, a(t), b(t))\}$ containing $E_{P_1\widehat{P_2(t_0)}}$ for $t = t_0$ and such that $E_t : y = y(x,t)$ is transversal to N at $P_2(t) = (x_2(t), y_2(t))$.

Denote by $p(t)$ the slope of E_t at $P_2(t)$, i.e. set $p(t) \underset{def}{=} y'(x_2(t), a(t), b(t))$, assuming of course that our subfamily exist. Then the three unknown functions, $a(t), b(t), p(t)$ of which the first two are needed to determine the subfamily, must satisfy the three relations:

$$\text{(33)} \qquad \left\{ \begin{array}{ll} f(x_2(t), y_2(t), p(t)) \cdot x'_t + f_{y'}(x_2(t)mt_2(t), p(t))(Y'_2 - p(t)x'_2) = 0, & \text{by } \mathbb{T} \\ y'(x_2(t), a(t), b(t)) - p(t) = 0, & \\ -Y_2(t) + y(x_2(t), a(t), b(t)) = 0, & \end{array} \right\}.$$

and furthermore $a(t_0) = a_0, b(t_0) = b_0$. Set $p_0 \underset{def}{=} y'(x_2(t_0), a_0, b_0)$.

The above system of equations is sufficient, by the implicit function theorem, to determine $a(t), b(t), p(t)$ as class C^1[37] functions of t, for t near t_0. To see this notice that the system is satisfied at (t_0, a_0, b_0, p_0) and at this point the relevant Jacobian determinant turns out to be non zero as follows. Denoting the members of the left hand side of (33) by L_1, L_2, L_3

[37] The assumption that N is of class C^2 enters here, insuring that $a(t)$ and $b(t)$ are of class C^1.

$$\left.\frac{\partial(L_1,L_2,L_3)}{\partial(p,a,b)}\right|_{t=t_0} = \left|\begin{array}{ccc} f_{y'}\cdot x_2' + f_{y'y'}(x,y,y')\cdot(Y_2' - px_2') - f_{y'}\cdot x_2' & 0 & 0 \\ -1 & y_a' & y_b' \\ 0 & y_a & y_b \end{array}\right|_{t=t_0}$$

$$= [Y_2'(t_0) - p_0 x_2'(t_0))]\cdot f_{y'y'}\Big|_{t_0}\cdot \Delta'(x_2(t_0), x_2(t_0))$$

The first of the three factors is not zero since condition $\mathbb{T}$ at $P_2(t_0)$ and $f\Big|_{P_2(t_0)} \neq 0$, would otherwise imply $x_2' = Y_2'(t_0) = 0$. But N is assumed free of singular points. The second factor is not zero by L'. The third factor is not zero since $\Delta'(x_2(t_0), x_2(t_0)) \neq 0$ by the construction of the family $\{y(x,a,b)\}$.

The subfamily $\{y(x,t)\} = \{y(x,a(t),b(t))\}$ thus constructed has each of its members cutting N transversely at $(x_2(t), Y_2(t))$, by the first relation of (33). Also it contains $E_{\widehat{P_1P_2(t_0)}}$ for $t = t_0$, since $a(t_0) = a_0, b(t_0) = b_0$. Finally, $y_t(x,t_0) \neq 0$ along $E_{\widehat{P_1P_2(t_0)}}$ by condition $\mathbb{K}$. Hence $y_t(x,t) \neq 0$ (say > 0) for all pairs (x,t) near the pairs (x,t_0) of $E_{\widehat{P_1P_2(t_0)}}$. As in the proof of the imbedding lemma in §16, this implies that for some neighborhood, $|t - t_0| < \epsilon$, the $y(x,t)$ cover a simply-connected neighborhood, F, of $E_{\widehat{P_1P_2(t_0)}}$ simply and completely.

(ii)Now let $C_{\widehat{P_1,Q}} : y = Y(x)$ be a competing curve, within the field F, from P_1 to a point Q of N. Then using results of §10 (in particular the invariance of I^*) and the methods of §17 we have

$$J[C_{\widehat{P_1,Q}}] - J[E_{\widehat{P_1P_2(t_0)}}] = J[C_{\widehat{P_1,Q}}] - I^*(E_{\widehat{P_1P_2(t_0)}}) = J[C_{\widehat{P_1,Q}}] - I^*(C_{\widehat{P_1,Q}} \cup \underbrace{N_{\widehat{QP_2(t_0)}}}_{I^*('')=0})$$

$$= \int_{C_{\widehat{P_1,Q}}} \mathcal{E}(x, Y(x), p(x, Y(x)), Y'(x))dx$$

Finally, if L' holds then $\mathcal{E} \geq 0$ in a weak neighborhood of $E_{\widehat{P_1P_2(t_0)}}$; if W_b holds then $\mathcal{E} > 0$ in a strong neighborhood of $E_{\widehat{P_1P_2(t_0)}}$. **q.e.d.**

23 Both End Points Variable

For $J[y] = \int_{P_1(u)}^{P_2(v)} f(x, y, y')dx$ assume $P_1(u)$ varies on a curve $M : x = x_1(u), y = y_1(u); a \leq u \leq b$ and $P_2(v)$ varies on a curve $N : x = x_2(v), y = y_2(v); c \leq v \leq d$. Assume both M and N are class C^2 and regular. Also $x_1(u) \leq x_2(v)$ for all u, v.

<u>Problem:</u> Find a curve $E_{P_1(u_0)\frown P_2(v_0)} : y = y_0(x)$ minimizing $J[y]$ among all curves, $y = y(x)$ joining a point P_1 of M to a point P_2 of N.

1. Necessary Conditions. Denote by T_M, T_N the transversality conditions between the extremal $E_0 \underset{def}{=} E_{P_1(u_0)\frown P_2(v_0)}$ at its intersections, $P_1(u_0)$ with M and $P_2(v_0)$ with N, respectively; also by $\overline{K}_M, \overline{K}_N$, the requirements that E_0 be free of focal points relative to the family of extremals transversal to M, or to N respectively. Furthermore K_M, K_N would be the weakened requirement that is in force only when G_M, G_N have branches toward P_1, P_2 respectively.

Then since a minimizing arc must minimize $J[y]$ among competitors joining P_1 to P_2, also among competitors joining P_1 to N, also among competitors joining P_2 to M, we have:

Theorem 23.1. *$E, L, J, W, T_M, T_N, K_M, K_N$ are necessary conditions for $E_0 = E_{P_1(u_0)\frown P_2(v_0)}$ to minimize $\int_{P_1(u)}^{P_2(v)} f(x, y, y')dx$.*

We derive a further condition

Condition B (due to Bliss) Assume the extremal satisfies the conditions of Theorem (23.1). Assume further that the extremal is smooth and satisfies

(i) $f(x, y_0, y_0')\Big|_{P_1(u_0)} \neq 0$, $f(x, y_0, y_0')\Big|_{P_2(v_0)} \neq 0$;

(ii) $\overline{K}_M$, $\overline{K}_N$;

(iii) E_0 or its extension beyond $P_1(u_0), P_2(v_0)$ contains focal points $P^f_{1(M)}, P^f_{2(N)}$ of P_1, P_2 where the envelopes G_M, G_N do have branches at $P^f_{1(M)}, P^f_{2(N)}$ toward P_1, P_2 respectively.

(iv) L' holds along E_0 and its extensions through $P^f_{1(M)}, P^f_{2(N)}$.

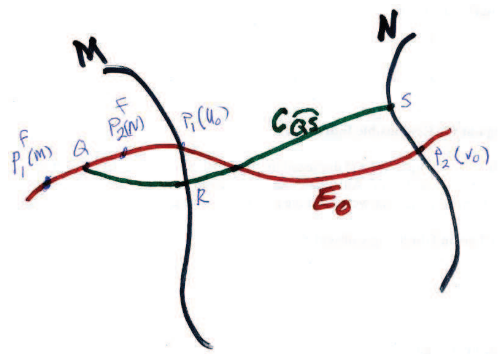
THEN for E_0 to furnish even a weak minimum for $J[y]$, the cyclic order of the four points $P_1, P_2, P^f_{1(M)}, P^f_{2(N)}$ must be:

$$P^f_{1(M)} \prec P_1 \prec P_2 \prec P^f_{2(N)}$$

where any cyclic permutation of this is ok (e.g. $P_1 \prec P_2 \prec P^f_{2(N)} \prec P^f_{1(M)}$).

Proof : Assume **B** is not satisfied, say by having E_0 extended contains the four points in the order

$$P^f_{1(M)} \prec P^f_{2(N)} \prec P_1 \prec P_2 \text{ (see diagram below) .}$$



Choose Q on E_0 (extended) such that $P^f_{1(M)} \prec Q \prec P^f_{2(N)}$. Then:

(1) $E_{Q\widehat{P_1(u_0)}}$ is an extremal joining point Q to curve M and satisfying all the conditions of the sufficiency theorem (22.1). Hence for any curve $C_{\widehat{QR}}$ joining Q to M and sufficiently near $E_{Q\widehat{P_1(u_0)}}$ we have

$$J[C_{\widehat{QR}}] \geq J[E_{Q\widehat{P_1(u_0)}}].$$

So E_0 does not give even a weak minimum. **q.e.d.**

2. Sufficient Conditions. We only state a theorem giving sufficient conditions. For its proof, which depends on the treatment of conditions $\overline{K}$, and B in terms of the second variations see [B46, Pages 180-184] or [Bo].

Notice that condition B includes conditions $\overline{K}_M$ and $\overline{K}_N$.

Theorem 23.2. *If $E_{\widehat{P_1P_2}}$ connecting P_1 on M to P_2 on N is a smooth extremal satisfying; $E, L', f \neq 0$ at P_1 and at $p_2, T_M, T_N,$ and B then $E_{\widehat{P_1P_2}}$ gives a weak relative minimum for $J[y]$. If W_b' is added, then $E_{\widehat{P_1P_2}}$ provides a strong minimum.*

24 Some Examples of Variational Problems with Variable End Points

First some general comments. Let $y = y(x, a, b)$ represent the two-parameter family of extremals. To determine the two variables, a_0, b_0 for suitable candidates, $E_0 : y = y(x, a_0, b_0)$ that are supposed to minimize we have two conditions to begin with:

1. In the case of P_1 fixed, P_2 variable on N, E_0 must pass through P_1 and be transversal to N at P_2.

2. In the case of P_1, P_2 both variable, on M, N respectively, E_0 must be transversal to M at P_1 and to N at P_2.

We study the problem of the shortest distance from a given point to a given parabola. Let The fixed point be $P_1 = (x_1, y_1)$ and the end point vary on the parabola, $N : x^2 = 2py$. $J[y] = \int_{x_1}^{x_2} f(x, y, y')dx$ and the extremals are straight lines. Transversality = perpendicular to N. The focus curve in this case is the evolute of the parabola (see the example in §12). We shall use the following fact:

- The evolute, G of a curve, N is the locus of centers of curvature of N.

To determine G, the centers of curvature, (ξ, η) corresponding to a point (x, y) of N, satisfies:

1. $\frac{\eta-y}{\xi-x} = -\frac{1}{y'(x)} = -\frac{p}{x}$, and

2. $(\xi - x)^2 + (\eta - y)^2 = \rho^2 = \frac{(x^2+p^2)^3}{p^4}$

where the center of curvature $= \rho = \frac{(1+(y')^2)^{\frac{3}{2}}}{y''} = \frac{1}{p^2}(x^2 + p^2)^{\frac{3}{2}}$.

From (1) and (2) we obtain parametric equations for G, with x as parameter:

$$G : \xi = -\frac{x^3}{p^2}; \quad \eta = p + \frac{3}{2p}x^2$$

or

$$27p\xi^2 - 8(\eta - p)^3 = 0.$$

Thus G is a semi-cubical parabola, with cusp at $(0, p)$ (see the diagram below).

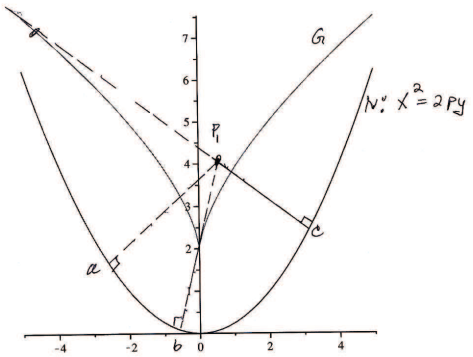


Figure 18: Shortest distance from P_1 to the Parabola, N.

Now consider a point, P_1, and the minimizing extremal(s) (i.e. straight lines) from P_1 to N. Any such extremal must be transversal (i.e. $\perp$) to N, hence tangent to G. On the basis of this there are three cases:

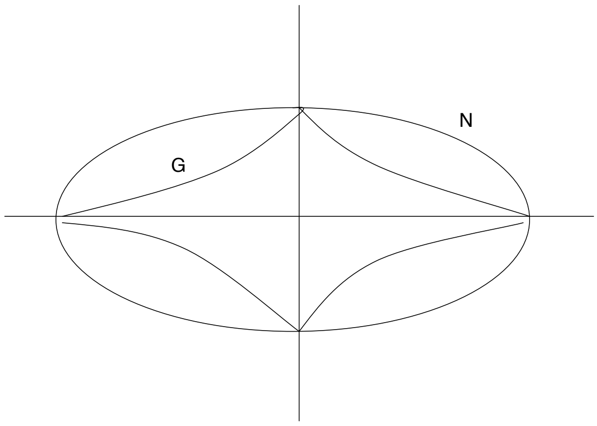
I P_1 lies inside G (this is the case drawn in the diagram). Then there are 3 lines from P_1 and $\perp$ to N. The one that is tangent to G <u>between</u> P_1 and N (hitting N at b) does not even give a weak relative minimum. The other two both give strong, relative minima. These are equal if P_1 lies on the axis of N, otherwise the one <u>not</u> crossing the axis gives the absolute minimum (in the diagram it is the line hitting N at c).

II P_1 lies outside G. Then there is just one perpendicular segment from P_1 to N, which gives a strong relative minimum and is in fact the absolute minimum.

III P_1 lies on G. Then if P_1 is not at the cusp, there are two perpendiculars from P_1 to N. The one not tangent to G at P_1 gives the minimum distance from P_1 to N. The other one does not proved even a weak relative minimum. If P_1 is at the cusp, there is only 1 perpendicular, which does give the minimum. Notice that this does not violate the Jacobi condition since the evolute does not have a branch going back, so the geometric proof does not apply, while the analytic proof required the focal point to be strictly after P_1.

Some exercises:

1) Analyze the problem of the shortest distance from a point to an ellipse.



2) Discuss the problem of finding the shortest path from the circle $M : x^2 + (y - q)^2 = r^2$ where $q - r > 0$, to the parabola $N : x^2 = 2py$ (You must start by investigating common normals.)

3) A particle starts at rest, from $(0, 0)$, and slides down alon a curve C, acted on by gravity, to meet a given vertical line, $x = a$. Find C so as to minimize the time of descent.

4) Given two circles in a vertical plane, determine the curve from the first to the second down which a particle wile slide, under gravity only, in the shortest time.

25 Multiple Integrals

We discuss double integrals. The treatment for $n > 2$ dimensions is similar.

The Problem: Given $f(x, y, u, p, q)$ of class C^3 for $(x, y, u) \in G$ (a 3-dimensional region) and for all p, q. We are also given a closed, class C^1 space curve $\mathfrak{K}$ in G with projection $\partial\mathcal{D}$ in the (x, y) plane. $\partial\mathcal{D}$ is a simple closed class-C^1 curve bounding a region $\mathcal{D}$. Then among all class C^1 functions, $u = u(x, y)$ on $\mathcal{D}$ such that $\mathcal{K} = u(\mathcal{D})$ (i.e. functions that "pass through $\mathcal{K}$"), to find $u_0(x, y)$ giving a minimum for $J[u] = \iint\limits_{\mathcal{D}} f(x, y, u(x, y), u_x(x, y), u_y(x, y))dxdy$.

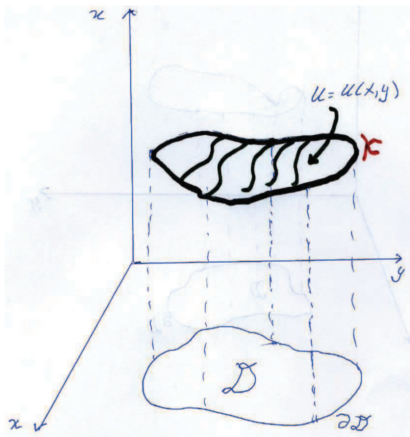


Figure 19: To find surface, $u = u(x, y)$ minimizing $J[u]$.

The first step toward proving necessary conditions on u_0 similar to the one dimensional case is to prove an analogue of the fundamental lemma, (2.1).

Lemma 25.1. *Let $M(x,y)$ be continuous on the bounded region $\mathcal{D}$. Assume that $\iint\limits_{\mathcal{D}} M(x,y)\eta(x,y)dxdy = 0$ for all $\eta \in C^q$ on $\mathcal{D}$ such that $\eta = 0$ on $\partial\mathcal{D}$. Then $M(x,y) \equiv 0$ on $\mathcal{D}$.*

Proof : Write $(x,y) = P, (x_0,y_0) = P_0, \sqrt{(x-x_0)^2+(y-y_0)^2} = \rho(P,P_0)$. Assume we have $M(P_0) > 0$ at $P_0 \in \mathcal{D}$. Hence $M(P) > 0$ on $\{P|\rho(P,P_0) < \delta\}$ for some $\delta > 0$. Set

$$\eta_0 = \begin{cases} [\delta^2 - \rho^2(P,P_0)]^{q+1} & \text{if } \rho(P,P_0) \leq \delta \\ 0 & \text{if } \rho(P,P_0) > \delta \end{cases}$$

Then $\eta_0 \in C^q$ everywhere, $\eta = 0$ on $\partial\mathcal{D}$ and $\iint\limits_{\mathcal{D}} M(x,y)\eta(x,y)dxdy > 0$, a contradiction. **q.e.d.**

Clearly the Lemma and proof generalizes to any number of dimensions. Now assume that $\dot{u}(x,y)$ is of class C^2 on $\mathcal{D}$, passes through $\mathcal{K}$ over $\partial\mathcal{D}$ and furnishes a minimum for $J[u] = \iint\limits_{\mathcal{D}} f(x,y,u(x,y),u_x(x,y),u_y(x,y))dxdy$. among all $\widetilde{C}^1$ $u(x,y)$ that pass through $\mathcal{K}$ above $\partial\mathcal{D}$. Let $\eta(x,y)$ be any fixed C^2 function on $\mathcal{D}$ such that $\eta(\partial\mathcal{D}) = 0$. We imbed $\dot{u}$ in the one-parameter family

$$\{u = \dot{u} + \epsilon\eta | -\epsilon_0 < \epsilon < \epsilon_0, \epsilon_0 > 0\}$$

Then $\varphi_\eta(\epsilon) \underset{def}{=} J[\dot{u} + \epsilon\eta]$ must have a minimum at $\epsilon = 0$. Hence $\varphi'_\eta(0) = 0$ is necessary for $\dot{u}$ to furnish a minimum for $J[u]$. Denoting $f(x,y,\dot{u}(x,y),\dot{u}_x(x,y),\dot{u}_y(x,y))$by $\dot{f}$ etc, we have:

$$\varphi'_\eta(0) = \iint\limits_{\mathcal{D}} (\dot{f}_u\eta + \dot{f}_p\eta_x + \dot{f}_q\eta_y)dxdy = \iint\limits_{\mathcal{D}} \dot{f}_u\eta + [\frac{\partial}{\partial x}(\dot{f}_p\eta) + \frac{\partial}{\partial y}(\dot{f}_q\eta) - \eta(\frac{\partial \dot{f}_p}{\partial x} + \frac{\partial \dot{f}_q}{\partial y})]dxdy.$$

Now $\iint\limits_{\mathcal{D}} \frac{\partial}{\partial x}(\dot{f}_p\eta) + \frac{\partial}{\partial y}(\dot{f}_q\eta)dxdy = \iint\limits_{\partial\mathcal{D}} \eta(\dot{f}_p dy - \dot{f}_q dx) = 0$. So we have $\iint\limits_{\mathcal{D}} \eta(\dot{f}_u - \frac{\partial \dot{f}_p}{\partial x} - \frac{\partial \dot{f}_q}{\partial y})dxdy = 0$ for every $\eta \in C^2$ satisfying $\eta = 0$ on $\partial\mathcal{D}$. Hence by (25.1) we have for $\dot{u}(x,y) \in C^2$ to minimize $J[u]$, it is necessary that $\dot{u}$ satisfy the two dimensional Euler-Lagrange differential equation:

$$\dot{f}_u - \frac{\partial \dot{f}_p}{\partial x} - \frac{\partial \dot{f}_q}{\partial y} = 0. \tag{34}$$

Expanding the partial derivatives in (34) and using the customary abbreviations $p \underset{def}{=} u_x, q \underset{def}{=} u_y, r \underset{def}{=} u_{xx}, s \underset{def}{=} u_{xy}, t \underset{def}{=} u_{yy}$ we see that (34) is a 2-nd order partial differential equation:

$$\dot{f}_{pp}\dot{r} + 2\dot{f}_{pq}\dot{s} + \dot{f}_{qq}\dot{t} + \dot{f}_{up}\dot{p} + \dot{f}_{uq}\dot{q} + (\dot{f}_{xp} + \dot{f}_{yq} - \dot{f}_u) = 0 \tag{35}$$

This differential equation is said to be"quasi-linear". i.e. it is linear in the variables u_{xx}, u_{yy}, u_{xy}.

The partial differential equation, (34) or (35) together with the boundary condition (u has assigned value on $\partial\mathcal{D}$) constitutes a type of boundary value problem known as a Dirichlet problem. An important special case is to find a C^2 function on the disk, D of radius R satisfying $u_{xx} + u_{yy} = 0$ in D such that u assumes assigned, continuous boundary values on the bounding circle. Such functions are called harmonic. This particular Dirichlet problem is solved by the Poisson integral

$$u(r,\theta) = \frac{1}{2\pi}\int_0^{2\pi} u(R,\varphi)\frac{(R^2-r^2)d\varphi}{R^2-2Rrcos(\theta-\varphi)+r^2}$$

where $[r,\theta]$ are the polar coordinates of (x,y) with respect to the center of D, and $u(R,\varphi)$ gives the assigned boundary values.

Example 1: To find the surface bounded by $\mathcal{K}$ of minimum surface area. Here $J[u] = \iint\limits_{\mathcal{D}} \sqrt{1+u_x^2+u_y^2}dxdy$ Therefore $f_u \equiv 0, f_p = \frac{p}{\sqrt{1+p^2+q^2}}, f_q = \frac{q}{\sqrt{1+p^2+q^2}}$ and (34) becomes

$$\frac{\partial}{\partial x}\frac{p}{\sqrt{1+p^2+q^2}} + \frac{\partial}{\partial y}\frac{q}{\sqrt{1+p^2+q^2}} = 0$$

$$\Rightarrow r(1+q^2) - 2pqs + t(1+p^2) = 0$$

To give a geometric interpretation of the Euler-Lagrange equation (due to Meusnier (1776)) we have to review some differential geometry.

Let κ_1, κ_2 be the principal curvatures at a point, i.e. the largest and smallest curvatures of the normal sections through the point). If the surface is given parametrically by $\overline{x}(u,v) = (x_1(u,v), x_2(u,v), x_3(u,v))$ then mean curvature, $H = \frac{1}{2}(\kappa_1+\kappa_2)$ is given by

$$\frac{Ee-2Ff+Gg}{2(EG-F^2)}$$

where $E = \overline{x}_u\cdot\overline{x}_u, F = \overline{x}_u\cdot\overline{x}_v, G = \overline{x}_v\cdot\overline{x}_v$ are the coefficients of the first fundamental form (i.e. E, F, G are g_{11}, g_{12}, g_{22}), and $e = \overline{x}_{uu}\cdot\overline{N}, f = \overline{x}_{uv}\cdot\overline{N}, g = \overline{x}_{vv}\cdot\overline{N}$ ($\overline{N} = \frac{\overline{x}_u\times\overline{x}_v}{|\overline{x}_u\times\overline{x}_v|}$) are the coefficients of the second fundamental form. (See for example [O, Page 212]).

Now if the surface is given by $z = z(x,y)$ we revert to a parametrization; $\overline{x}(u,v) = (u,v,z(u,v))$, hence setting $z_1 = p, z_2 = q, z_{11} = r, z_{12} = s, z_{22} = t$, we have:

$$\overline{x}_u = (1,0,p), \overline{x}_v = (0,1,q), \overline{x}_{uu} = (0,0,r), \overline{x}_{uv} = (0,0,s), \overline{x}_{vv} = (0,0,t), \overline{N} = \frac{(-p,-q,1)}{\sqrt{1+p^2+q^2}}.$$

Plugging this into the formula for H we obtain:

$$H = \frac{1}{2(1+p^2+q^2)^{\frac{3}{2}}}[r(1+q^2) - 2pqs + t(1+p^2)]$$

Definition 25.2. A surface is *minimal* if $H \equiv 0$.

The Euler-Lagrange equation for the minimum surface area problem implies:

Theorem 25.3 (Meusnier)**.** *An area-minimizing surface must be minimal.*

The theorem implies that no point of the surface can be elliptic (i.e. $\kappa_1\kappa_2 > 0$). Rather they all look flat, or saddle shaped. Intuitively one can see that at an elliptic point, the surface can be flattened to give a smaller area.

Example 2: To minimize $J[u] = \iint\limits_{\mathcal{D}} (u_x^2 + u_y^2)dxdy = \iint\limits_{\mathcal{D}} |\nabla u|^2 dxdy$. Here $u(x,y)$ passes through an assigned space curve, $\mathcal{K}$ with projection $\partial\mathcal{D}$. (34) becomes

$$u_{xx} + u_{yy} = 0$$

This is Laplace's equation. So $\dot{u}(x,y)$ must be harmonic.

A physical interpretation of this example comes from the the interpretation of $\iint\limits_{\mathcal{D}} |\nabla u|^2 dxdy$ as the energy functional. A solution to Laplace's equation makes the energy functional stationary, i.e. the energy does not change for an infinitesimal perturbation of u.

3: Further remarks.(a) If only $\dot{u} \in \widetilde{C}^1 \in \mathcal{D}$ is assumed for the minimizing function, a first necessary condition in integral form, in place of (34) was obtained by A. Haar (1927).

Theorem 25.4. *For any $\Delta \subset \mathcal{D}$ with boundary curve $\partial\Delta \in \widetilde{C}^1$*

$$\iint\limits_{\Delta} \dot{f}_u dx dy = \int\limits_{\partial\Delta} \dot{f}_p dy - \dot{f}_q dx$$

must hold.

This integral condition is to (34) as, theorem (3.2) is to corollary (3.3).

Furthermore, if $\frac{\partial}{\partial x}\dot{f}_p, \frac{\partial}{\partial y}\dot{f}_q$ are continuous, (34) follows from Haar's condition. See [AK, Pages 208-210] for proofs.

(b) The analogue of Legendre's condition:

Theorem 25.5. *If $\dot{u}(x, y)$ is to give a weak extremum for $J[u]$, we must have*

$$\dot{f}_{pp}\dot{f}_{qq} - \dot{f}_{pq}^2 \geq 0$$

at every interior point of $\mathcal{D}$. For a minimum (respectively maximum), we must further have $\dot{f}_{pp} \geq 0$ (or ≤ 0 respectively).

See [AK, Pages 211-213].

Note that the first part of the theorem is equivalent to u being a Morse function.

(c) For a sufficiency theorem for certain regular double integral problems, by Haar, see [AK, Pages 215-217].

26 Functionals Involving Higher Derivatives

Problem: Given $J[y] = \int_{x_1}^{x_2} f(x, y(x), y'(x), \cdots, y^{(m)}(x))dx$, obtain necessary conditions on $y_0(x)$ minimizing $J[y]$ among $y(x)$ satisfying the boundary conditions:

(36) $$y^{(k)}(x_i) = y_{ik}(\text{ assigned constants }) \quad i = 1, 2; k = 0, 1, \cdots, m-1$$

A carefree derivation of the relevant Euler-Lagrange equation is given in **1 a.** where we don't care about economizing on the differentiability assumptions. A less carefree derivation follows in **1 b.**, where we use Zermelo's lemma to obtain an integral equation version of the Euler-Lagrange equations. In **2. & 3.**, auxiliary material, of interest in its own right, is developed, ending with a proof of Zermelo's lemma for general m.

1 a. Assume first that $f \in C^m$ for (x, y) in a suitable region and for all $(y', \cdots, , y^{(m)}$; also that the minimizing function, $y_0(x)$ is a C^{2m} function. Imbed $y_0(x)$ as usual, in 1-parameter family $\{y_\epsilon\} = \{y - 0 + \epsilon\eta\}$, where $n \in C^{2m}$ and $\eta^{(k)})(x_i) = 0$ for $i = 1, 2; k = 0, \cdots m-1$. Set

$$\varphi_\eta(\epsilon) = J[y_0 + \epsilon\eta].$$

Then the necessary condition, $\varphi'_\eta(0) = 0$ becomes:

$$\int_{x_1}^{x_2} (\dot{f}_y\eta + \dot{f}_{y'}\eta' + \cdots + \dot{f}_{y^{(m)}}\eta^{(m)})dx.$$

Repeated integration by parts (always integrating the $\eta^{(k)}$ factor) and use of the boundary condition yields

$y_0(x)$ must satisfy

$$\dot{f}_y - \frac{d}{dx}\dot{f}_{y'} + \cdots + (-1)^m \frac{d^m}{dx^m}\dot{f}_{y^{(m)}}. \tag{37}$$

. For $y_0 \in C^{2m}$ and $f \in C^m$, the indicated differentiation can be carried out to convert th Euler-Lagrange equation, (37) into a partial differential equation for y_0 of order $2m$. For the $2m$ constrains in it general solution, the conditions (36) impose $2m$ condition.

1 b. Assuming less generous differentiability conditions- say merely $f \in C^1$, and for the minimizing function: $y_0 \in \widetilde{C}^m$ - we can derive an integral equation as a necessary condition that reduces to (36) in the presence of sufficiently favorable conditions. Proceed as in **1 a.** up to $\int_{x_1}^{x_2}(\dot{f}_y\eta + \dot{f}_{y'}\eta' + \cdots + \dot{f}_{y^{(m)}}\eta^{(m)})dx$. Now use integration by part, always differentiating the $\eta^{(k)}$ factor and use the boundary conditions repeatedly until only $\eta^{(m)}$ survives to obtain:

$$0 = \int_{x_1}^{x_2} [\dot{f}_{y^{(m)}} - \int_{x_1}^{x} \dot{f}_{y^{(m-1)}} d\overline{x} + \int_{x_1}^{x}\int_{x_1}^{\overline{x}} \dot{f}_{y^{(m-2)}} d\overline{\overline{x}} d\overline{x} - \cdots + (-1)^m \int_{x_1}^{x}\int_{x_1}^{\overline{x}} \cdots \int_{x_1}^{\overset{(m-1)}{x}} \dot{f}_y \overset{(m)}{x} \cdots \overline{x}]\eta^{(m)} dx$$

where $\overset{(k)}{x}$ denotes a variable with k bars.

Apply Zermelo's lemma (proven in part **3 b.**) to obtain:

$$\dot{f}_{y^{(m)}} - \int_{x_1}^{x} \dot{f}_{y^{(m-1)}} d\overline{x} + \int_{x_1}^{x}\int_{x_1}^{\overline{x}} \dot{f}_{y^{(m-2)}} d\overline{\overline{x}} d\overline{x} - \cdots + (-1)^m \int_{x_1}^{x}\int_{x_1}^{\overline{x}} \cdots \int_{x_1}^{\overset{(m-1)}{x}} \dot{f}_y \overset{(m)}{x} \cdots \overline{x} \tag{38}$$

$$= c_0 + \cdots + c_{m-1}(x - x_1)^{m-1}$$

If we allow for the left-hand side of this integral equation to be in $\widetilde{C}^m$ e.g. by assuming $f \in C^m$ in a suitable (x, y) region and for all $(y', \cdots, y^{(m)})$, then (38) yields (37) at any x where the left hand side is continuous, i.e. at any x where $y_0^{(m)}(x)$ is continuous.

2. We now develop auxiliary notions needed for the proofs of the lemmas in **3.**

Let $\varphi_1(x), \cdots, \varphi_m(x)$ be $\widetilde{C}^0$ functions on $[x_1, x_2]$. They are linearly dependent on $[x_1, x_2]$ if and only if there exist $(\lambda_1, \cdots, \lambda_m) \neq (0, \cdots, 0)$ such that $\Sigma_{i=1}^m \lambda_1 \varphi_i(x) = 0$ at all x of the set S of continuity of all $\varphi_i(x)$. Otherwise we say the functions are linearly independent on $[x_1, x_2]$. Define a scaler product by

$$\langle \varphi_i(x), \varphi_j(x) \rangle = \int_{x_1}^{x_2} \varphi_i(x)\varphi_j(x)dx$$

The Gram determinant of the φ_i is

$$G_{[x_1,x_2]}(\varphi_1(x), \cdots, \varphi_m(x)) = \begin{vmatrix} \langle \varphi_1, \varphi_1 \rangle & \cdots & \langle \varphi_1, \varphi_m \rangle \\ \vdots & & \vdots \\ \langle \varphi_m, \varphi_1 \rangle & \cdots & \langle \varphi_m, \varphi_m \rangle \end{vmatrix}$$

Theorem 26.1. *$\varphi_1(x), \cdots, \varphi_m(x)$ are linearly dependent on $[x_1, x_2]$ if and only if $G_{[x_1,x_2]}(\varphi_1(x), \cdots, \varphi_m(x)) = 0$.*

Proof : (i) Assume linear dependence. Then there exist $\lambda_1, \cdots \lambda_m$ such that $\Sigma_{i=1}^m \lambda_1 \varphi_i(x) = 0$ on S. Apply $\langle -, \varphi_j \rangle$ to this relation. This shows that the resulting linear, homogeneous $m \times m$ system has a not trivial solution, $\lambda_1, \cdots \lambda_m$, hence its determinant, $G_{[x_1,x_2]}(\varphi_1(x), \cdots, \varphi_m(x))$ must be zero.

(ii) Assume $G = 0$. Then the system of equations, $\Sigma_{i=1}^m \lambda_1 \langle \varphi_i(x), \varphi_j(x) \rangle = 0, j = 1, \cdots, m$ has a non-trivial solution, say $\lambda_1, \cdots \lambda_m$. Multiply the $j-$th equation by λ_j and sum over j, obtaining

$$0 = \sum_{i,j=1}^{m} \lambda_i \lambda_j \langle \varphi_i, \varphi_j \rangle = \int_{x_1}^{x_2} \Big(\Sigma_{k=1}^m \lambda_k \varphi_k(x) \Big)^2 dx.$$

Hence $\Sigma_{k=1}^m \lambda_k \varphi_k(x) = 0$ on S. **q.e.d.**

The above notions, theorem and proof generalize to vector valued functions $\{\overline{\varphi}_i(x) = (\varphi_{i1}(x), \cdots, \overline{\varphi}_{in}(x))\}$ of class $\widetilde{C}^0$ on $[x_1, x_2]$. Where we define

$$\langle \overline{\varphi}_i, \overline{\varphi}_j \rangle = \int_{x_1}^{x_2} \overline{\varphi}_i \cdot \overline{\varphi}_j dx.$$

3 a. The generalized lemma of the calculus of variations:

Theorem 26.2. *Let $\varphi_1(x), \cdots, \varphi_m(x)$ be linear independent $\widetilde{C}^0$ functions on $[x_1, x_2]$. Let $M(x) \in \widetilde{C}^0$ on $[x_1, x_2]$. Assume $M \perp \eta(x)$ for every $\eta \in \widetilde{C}^0$ that satisfies $\eta(x) \perp \varphi_i, i = 1, \cdots, m$. Then on the common set of continuity of $\{M(x), \varphi_i(x), i = 1, \cdots m\}$ $M(x) = \Sigma_{i=1}^m c_i \varphi_i(x)$*

Proof : Consider the linear $m \times m$ system in the unknowns $C_1, \cdots C_m$

$$C_1 \langle \varphi_1, \varphi_i \rangle + \cdots + C_m \langle \varphi_m, \varphi_i \rangle = \langle M, \varphi_i \rangle, i = 1, \cdots m. \tag{39}$$

Its determinant, $G_{[x_1,x_2]}(\varphi_1(x), \cdots, \varphi_m(x)) \neq 0$. Hence there is a unique solution $c_1, \cdots, c_m$. Using this construct $\eta_0(x) = M(x) - \Sigma_{j=1}^m c_j \varphi_j(x)$. Then by (39) $\eta_0 \perp \varphi_i, i = 1, \cdots, m$. Hence by hypothesis $M \perp \eta_0$. Therefore

$$\int_{x_1}^{x_2} [M(x) - \Sigma_{j=1}^m c_j \varphi_j(x)]^2 dx = \int_{x_1}^{x_2} \eta_0(x) [M(x) - \Sigma_{j=1}^m c_j \varphi_j(x)] dx = 0,$$

whence $M(x) - \Sigma_{j=1}^m c_j \varphi_j(x) = 0$ **q.e.d.**

Note: This may be generalized to vector valued functions, $\overline{\varphi}_i, i = 1, \cdots, m; \overline{M}$.

3 b. Zermelo's lemma:

Theorem 26.3. *Let* $M(x) \in \widetilde{C}^0$ *on* $[x_1, x_2]$. *Assume* $\int_{x_1}^{x_2} M(x)\eta^{(m)}(x)dx = 0$ *for all* $\eta \in \widetilde{C}^m$ *satisfying* $\eta^{(k)}(x_i) = 0, i = 1, 2, k = 0, \cdots, m-1$ *Then* $M(x) = \Sigma_{i=0}^{m-1} c_i(x - x_1)^i$.

Proof : Consider the m functions, $\varphi_1(x) = 1, \varphi_2(x) = x - x_1, \cdots, \varphi_m(x) = (x - x_1)^{m-1}$. They are linearly independent on $[x_1, x_2]$. Now every $\widetilde{C}^0$ function, $\xi(x)\perp$ to all the $\varphi_i(x)$ satisfies $\int_{x_1}^{x_2} \xi(x)(x - x_1)^{i-1}dx = 0, i = 1, \cdots, m.$. Hence if we set

$$\eta(x) = \int_{x_1}^{x} \xi(t)(x - t)^{m-1}dt$$

then $\eta(x)$ satisfies the conditions of the hypothesis since $\eta^{(k)}(x) = (m-1)(m-2)\cdots(m-k)\int_{x_1}^{x} \xi(t)(x-t)^{m-k-1}dt$ for $k = 0, \cdots, m-1$ and we have $\eta^{(m)}(x) = (m-1)!\xi(x)$. By the hypothesis of Zermelo's lemma, $\int_{x_1}^{x_2} M\eta^{(m)}dx = 0$. Thus $\int_{x_1}^{x_2} M\xi dx = 0$ holds for any ξ that is $\perp$ to $\varphi_1, \cdots \varphi_m$. The theorem now follows by the generalized lemma of the calculus of variations. **q.e.d.**

4. Multiple integrals involving higher derivatives: We give a simple example. Consider:

$$J[u] = \iint\limits_{\mathcal{D}} f(x, y, u(x, y), u_x, u_y, u_{xx}, u_{xy}, u_{yy})dxdy$$

where $u(x, y)$ ranges over functions such that the surface $u(x, y)$ over the (x, y) region, $\mathcal{D}$, is bounded by an assigned curve $\mathcal{K}$ whose projection to the (x, y) plane is $\partial\mathcal{D}$.

Assume $f \in C^2$, a minimizing $\dot{u}(x, y)$ must satisfy the fourth order partial differential equation

$$\dot{f}_u - \frac{\partial}{\partial x}\dot{f}_{u_x} - \frac{\partial}{\partial y}\dot{f}_{u_y} + \frac{\partial^2}{\partial x^2}\dot{f}_{u_{xx}} + \frac{\partial^2}{\partial x \partial y}\dot{f}_{u_{xy}} + \frac{\partial^2}{\partial y^2}\dot{f}_{u_{yy}} = 0$$

The proof is left as an exercise.

27 Variational Problems with Constraints.

Review of the Lagrange Multiplier Rule. Given $f(u,v), g(u,v)$, C^1 functions on an open region, $\mathcal{D}$. Set $S_{g=0} = \{(u,v)|g(u,v)=0\}$. We obtain a necessary condition on (u_0,v_0) to give an extremum for $f(u,v)$ subject to the constraint: $g(u,v)=0$.

Theorem 27.1 (Lagrange Multiplier Rule)**.** *Assume that the restriction of $f(u,v)$ to $\mathcal{D}\bigcap S_{g=0}$ has a relative extremum at $P_0 = (u_0,v_0) \in \mathcal{D}\bigcap S_{g=0}$ [i.e. Say $f(u_0,v_0) \le f(u,v)$ for all $(u,v)\in\mathcal{D}$ satisfying $g(u,v)=0$ sufficiently near (u_0,v_0).] Then there exist λ_0,λ_1 not both zero such that*

$$\lambda\nabla f\Big|_{P_0} + \lambda_1\nabla g\Big|_{P_0} = (0,0).$$

Proof : Assume the conclusion false, ie. assume $\nabla f\Big|_{P_0}, \nabla g\Big|_{P_0}$ are linearly independent. Then the determinant,

$$\Delta = \begin{vmatrix} f_u & f_v \\ g_u & g_v \end{vmatrix}$$

$\neq 0$ at P_0. Hence, by continuity it is not zero in some neighborhood of P_0. Now consider the mapping given by

$$(40)\qquad \begin{cases} L_1(u,v) = & f(u,v) - f((u_0,v_) & = h \\ L_2(u,v) = & g(u,v) & = k \end{cases}$$

from the (u,v) plane to the (h,k) plane. It maps $(u_0,v_)$ to $(0,0)$, and its Jacobian determinant is δ which is not zero at and near (u_0,v_0). Hence by the open mapping theorem any sufficiently small open neighborhood of (u_0,v_0) is mapped $1-1$ onto some open neighborhood of $(0,0)$. Hence, in particular any neighborhood of (u_0, v_0) contains points (u_1, v_1) mapped by (40) onto some $(h_1, 0)$ with $h_1 > 0$, and also points (u_2, v_2) mapped by (40) onto some $(h_2, 0)$ with $h_2 < 0$. Thus $f(u_1, v_1) > f(u_0, v_0)$, $f(u_2, v_2) < f(u_0, v_))$ and $g(u_i, v_i) = 0, i = 1, 2$. Which is a contradiction. **q.e.d.**

Corollary 27.2. *If we add the further assumption that* $\nabla g\Big|_{P_0} \neq (0,0)$ *then the conclusion is that there is a* λ *such that* $\nabla(f + \lambda g)\Big|_{P_0} = (0,0)$.

Proof : λ_0 cannot be zero.

Recall the way the Lagrange multiplier method works, one has the system of equations:

$$\begin{cases} f_u + \lambda g_u &= 0 \\ f_v + \lambda g_v &= 0 \end{cases}$$

Which gives two equations in three unknowns, u_0, v_0, λ plus there is the third condition given by the constraint, $g(u_0, v_0) = 0$.

Generalization to q variables and $p(< q)$ constraints:

Theorem 27.3 (Lagrange Multiplier Rule)**.** *Set* $(u_1, \cdots, u_q) = \overline{u}$. *Let* $f(\overline{u}), g^{(1)}(\overline{u}), \cdots, g^{(p)}(\overline{u})(p < q)$ *all be class* C^1 *in an open* $\overline{u}$ *region,* $\mathcal{D}$. *Assume that* $\dot{\overline{u}} = (\dot{u}_1, \cdots, \dot{u}_q)$ *gives a relative extremum for* $f(\overline{u})$*, subject to the* p *constraints,* $g^{(i)}(\overline{u}) = 0, i = 1, \cdots, p$ *in* $\mathcal{D}$. *[i.e. assume that the restriction of* f *to* $\mathcal{D} \bigcap S_{g^{(1)} = \cdots = g^{(p)} = 0}$ *has a relative extremum at* $\dot{\overline{u}} \in \mathcal{D} \bigcap S_{g^{(1)} = \cdots = g^{(p)} = 0}$*]. Then there exist* $\lambda_0, \cdots, \lambda_p$*, not all zero such that* $\nabla(\lambda_0 f + \Sigma_{j=1}^p \lambda_j g^{(j)})\Big|_{\dot{\overline{u}}} = \overline{0} = (\overbrace{0, \cdots, 0}^{q})$.

Proof : Assume false. i.e. assume the $(p+1)$ vectors $\nabla f\Big|_{\dot{\overline{u}}}, \nabla g^{(1)}\Big|_{\dot{\overline{u}}}, \cdots, \nabla g^{(p)}\Big|_{\dot{\overline{u}}}$ are linearly independent. Then the matrix

$$\begin{pmatrix} f_{u_1} & \cdots & f_{u_q} \\ g^{(1)}_{u_1} & \cdots & g^{(1)}_{u_q} \\ \vdots & & \vdots \\ g^{(p)}_{u_1} & \cdots & g^{(p)}_{u_q} \end{pmatrix}$$

has rank $p+1(\leq q)$. Hence the matrix has a non-zero $(p+1)\times(p+1)$ subdeterminant at, and near $\dot{\overline{u}}$. Say the subdeterminant is given by the first $p+1$ columns. Consider the mapping

$$\text{(41)} \qquad \begin{cases} L_1 = & f(u_1,\cdots,u_{p+1},\dot{u}_{p+2},\cdots,\dot{u}_q) - f(\dot{\overline{u}} & = h \\ L_2 = & g^{(1)}(u_1,\cdots,u_{p+1},\dot{u}_{p+2},\cdots,\dot{u}_q) & = k_1 \\ \vdots & & \vdots \\ L_{p+1} = & g^{(p)}(u_1,\cdots,u_{p+1},\dot{u}_{p+2},\cdots,\dot{u}_q) & = k_p \end{cases}$$

from $(u_1,\cdots,u_{p+1})$ space to $(h,k_1,\cdots,k_p)$ space. This maps $\dot{\overline{u}}=(\dot{u}_1,\cdots,\dot{u}_q)$ to $(0,\cdots,0)$ and the Jacobian of the system is not zero at and near $\dot{\overline{u}}=(\dot{u}_1,\cdots,\dot{u}_q)$. Therefore by the open mapping theorem any sufficiently small neighborhood of $\dot{\overline{u}}=(\dot{u}_1,\cdots,\dot{u}_q)$ is mapped $1-1$ onto an open neighborhood of $(0,\cdots,0)$. As in the proof of (27.1) this contradicts the assumption that $\dot{\overline{u}}=(\dot{u}_1,\cdots,\dot{u}_q)$ provides an extremum. **q.e.d.**

Similar to the corollary to theorem (27.1) we have:

Corollary 27.4. *In addition assume* $\nabla g^{(1)}|_{\dot{\overline{u}}},\cdots,\nabla g^{(p)}|_{\dot{\overline{u}}}$ *are linearly independent, then we may take* $\lambda_0 = 1$.

There are three types of Lagrange problems: Isoperimetric problems (where the constraint is given in terms of an integral); Holonomic problems (where the constraint is given in terms os a function which does not involve derivatives) and Non-holonomic problems (where the constraint are given by differential equations).

Isoperimetric Problems This class of problems takes its name from the prototype mentioned in §1 example VI.

The general problem is to find among all admissible $Y(x) = (y_1(x), \cdots, y_n(x))$ with $Y(x) \in \widetilde{C}^1$ satisfying end point conditions, $Y(x_i) = (y_1^{(i)}, \cdots, y_n^{(i)}), i = 1, 2$ and "isoperimetric constraints", $G^{(j)}[Y] = \int_{x_1}^{x_2} g^{(j)}(x, Y(x), Y'(x))dx = L_j, j = 1, \cdots, p$, the $\dot{Y}(x)$ minimizing $J[Y] = \int_{x_1}^{x_2} f(x, Y(x), Y'(x)dx$.

Remark 27.5. The end point conditions are not considered to be among the constraints.

We derive necessary conditions for a minimizing $\dot{Y}$, first for $n = 1, p = 1$.

Given $f(x, y, y'), g(x, y, y')$ of class C^2 in an (x, y) region, $\mathcal{R}$, for all y' to find, among all admissible $y(x)$ satisfying the constraint $G[y] = \int_{x_1}^{x_2} g(x, y, y')dx = L$, a function, $y_0(x)$ minimizing $j[y] = \int_{x_1}^{x_2} f(x, y, y')dx$.

Toward obtaining necessary conditions, imbed a candidate, $y_0(x)$ in a $2-$parameter family $\{y_0(x) + \epsilon_1\eta_1(x) + \epsilon_2\eta_2(x)\}$, where $\eta_i \in \widetilde{C}^1$ on $[x_1, x_2]$ and $\eta_i(x_j) = 0$ for $i = 1, 2; j = 1, 2$ and where each member of the family satisfies the constraint $G[y] = L$. i.e. ϵ_i are are assumed to satisfy the relation:

$$\gamma(\epsilon_1, \epsilon_2) = G[y_0 + \epsilon_1\eta_1 + \epsilon_2\eta_2] = L.$$

Then for $y_0(x)$ to minimize $J[y]$ subject to the constraint $G[y] = L$, the function

$$\varphi(\epsilon_1, \epsilon_2) = J[y_0 + \epsilon_1\eta_1 + \epsilon_2\eta_2]$$

must have a relative minimum at $(\epsilon_1, \epsilon_2) = (0, 0)$, subject to the constraint $\gamma(\epsilon_1, \epsilon_2) - L = 0$. By the Lagrange multiplier rule this implies that there exist λ_0, λ_2 both not zero such that

$$\text{(42)} \qquad \begin{cases} \lambda_0 \frac{\partial\varphi}{\partial\epsilon_1} + \lambda_1 \frac{\partial\gamma}{\partial\epsilon_1}\Big|_{(\epsilon_1,\epsilon_2)=(0,0)} & = 0 \\ \lambda_0 \frac{\partial\varphi}{\partial\epsilon_2} + \lambda_1 \frac{\partial\gamma}{\partial\epsilon_2}\Big|_{(\epsilon_1,\epsilon_2)=(0,0)} & = 0 \end{cases}$$

Since φ, γ depend on the functions, η_1, η_2 chosen to construct the family, we also must alow for λ_0, λ_1 to be functionals: $\lambda_i = \lambda_i[\eta_1, \eta_2], i = 0, 1$. Now, as usual we have, for $i = 1, 2$ $\frac{\partial\varphi}{\partial\epsilon_i}\Big|_{(0,0)} = \int_{x_1}^{x_2} (\dot{f}_y - \frac{d}{dx}\dot{f}_{y'})\eta_i dx$. Hence equation (42) yields

$$\begin{cases} \lambda_0[\eta_1, \eta_2] \int_{x_1}^{x_2} (\dot{f}_y - \frac{d}{dx}\dot{f}_{y'})\eta_1 dx + \lambda_1[\eta_1, \eta_2] \int_{x_1}^{x_2} (\dot{g}_y - \frac{d}{dx}\dot{g}_{y'})\eta_1 dx & = 0 \\ \lambda_0[\eta_1, \eta_2] \int_{x_1}^{x_2} (\dot{f}_y - \frac{d}{dx}\dot{f}_{y'})\eta_2 dx + \lambda_1[\eta_1, \eta_2] \int_{x_1}^{x_2} (\dot{g}_y - \frac{d}{dx}\dot{g}_{y'})\eta_2 dx & = 0 \end{cases}$$

The first of these relations shows that the ratio $\lambda_0 : \lambda_1$ is independent of η_2; the second shows that it is independent of η_1. Thus λ_0, λ_1 may be taken as constants, and we may rewrite (say) the first relations as:

$$\int_{x_1}^{x_2} [\lambda_0(\dot{f}_y - \frac{d}{dx}\dot{f}_{y'}) + \lambda_1(\dot{g}_y - \frac{d}{dx}\dot{g}_{y'})]\eta_1 dx = 0$$

for all admissible $\eta_1(x)$ satisfying $\eta_1(x_1) = \eta_1(x_2) = 0$. Therefore by the fundamental lemma of the calculus of variations we have

$$\lambda_0(\dot{f}_y - \frac{d}{dx}\dot{f}_{y'}) + \lambda_1(\dot{g}_y - \frac{d}{dx}\dot{g}_{y'})$$

wherever on $[x_1, x_2]$ $y_0(x)$ is continuous. If $\dot{g}_y - \frac{d}{dx}\dot{g}_{y'}$ is not identically zero on $[x_1, x_2]$, then we cannot have $\lambda_0 = 0$, and in this case we may divide by λ_0. Thus we have proven the following

Theorem 27.6 (Euler Multiplier Rule). *Let $y_0(x) \in \widetilde{C}^1$ furnish a weak relative minimum for $J[y] = \int_{x_1}^{x_2} f(x, y, y')dx$ subject to the isoperimetric constraint $G[y] = \int_{x_1}^{x_2} g(x, y, y')dx = L$. Then either $y_0(x)$ is an extremal of $G[y]$, or there is a constant λ such that $y_0(x)$ satisfies*

$$(\dot{f}_y + \lambda\dot{g}_y) - \frac{d}{dx}(\dot{f}_{y'} + \lambda\dot{g}_{y'}) = 0. \tag{43}$$

Note on the procedure: The general solution of (43) will contain two arbitrary constants, a, b plus the parameter λ : $y = y(x; a, b, \lambda)$. To determine (a, b, λ) we have the two end point conditions, $y(x_i; a, b, \lambda) = y_i, i = 1, 2$ and the constraint $G[y(x_i; a, b, \lambda)] = L$.

An Example: Given $P_1 = (x_1, y_1), P_2 = (x_2, y_2)$ and $L > |\overline{P_1P_2}|$, find the curve $\mathcal{C}$: $y = y_0(x)$ of length L from P_1 to P_2 that together with the segment $\overline{P_1P_2}$ encloses the maximum area. Hence we wish to find $y_0(x)$ maximizing $J[y] = \int\limits_{(x_1,y_1)}^{(x_2,y_2)} ydx$ subject to $G[y] = \int\limits_{(x_1,y_1)}^{(x_2,y_2)} \sqrt{1 + (y')^2}dx = L$. The necessary condition (43) on $y_0(x)$ becomes

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

Integrate[38] to give $x - \frac{\lambda y'}{\sqrt{1+(y')^2}} = c_1$; solve for y': $y' = \frac{x-c_1}{\sqrt{\lambda^2-(x-c_1)^2}}$; integrate again, $(x-c_1)^2 = (y-c_2)^2 = \lambda^2$. The center (c_1, c_2) and radius λ are determined by $y(x_i) = y_i, i = 1, 2$ and the length of $\mathcal{C} = L$.

2. Generalization: Given$f(x, Y, Y'), g^{(1)}(x, Y, Y'), \cdots, g^{(p)}(x, Y, Y')$ all of class C^2 in a region, $\mathcal{R}_{(x,Y)}$, and for all Y', where $Y = (y_1, \cdots, y_n), Y' = (y'_1, \cdots, y'_n)$. To find among all admissible (say class $\widetilde{C}^1$) $Y(x)$ passing through $P_1 = (x, y_1^{(1)}, \cdots, y_n^{(1)})$ and $P_2 = (x_2, y_1^{(2)}, \cdots, y_n^{(2)})$ and satisfying p isoperimetric constraints, $G_j[y] = \int_{x_2}^{x_2} g^{(j)}(x, Y(x), Y'(x))dx = L_j, j = 1, \cdots, p$, a curve $\dot{Y}(x)$ minimizing $J[y] = \int_{x_2}^{x_2} f(x, Y(x), Y'(x))dx$.

Theorem 27.7 (Generalized Euler Multiplier Rule)**.** *If $\dot{Y}(x) \in \widetilde{C}^1$ furnishes a weak, relative extremum for $J[y] = \int_{x_2}^{x_2} f(x, Y(x), Y'(x))dx$ subject to p isoperimetric constraints, $G_j[Y] = \int_{x_2}^{x_2} g^{(j)}(x, Y(x), Y'(x))dx = L_j, j = 1, \cdots, p$, then either $\dot{Y}(x)$ is an extremal of some $GY \underset{def}{=} \Sigma_{j=1}^{p} \lambda_j G_j[Y]$, or there exist $\lambda_1, \cdots, \lambda_p$ not all zero such that*

$$(\dot{f}_{y_i} + \Sigma_{j=1}^{p} \lambda_j \dot{g}_{y_j}^{(j)}) - \frac{d}{dx}(\dot{f}_{y'_i} + \Sigma_{j=1}^{p} \lambda_j \dot{g}_{y'_j}^{(j)}) = 0, i = 1, \cdots, n. \tag{44}$$

Proof : To prove the i-t of the n necessary conditions, imbed $\dot{Y}(x)$ in a $(p+1)$-parameter family, $\{\overset{\epsilon}{Y}(x)\} = \{\dot{Y} + \epsilon_1 H_1(x) + \cdots \epsilon_{p+1} H_{p+1}(x)\}$, where $H_k(x) = (0, \cdots, \overbrace{\eta_k(x)}^{i\text{–th place}}, \cdots, 0)$ with $\eta_k(x0 \in \widetilde{C}^1$ and $\eta_k(x_1) = \eta_k(x_2) = 0$ and with the parameters $\epsilon_i, i = 1, \cdots, p+1$ subject to p constraints $\gamma^{(j)}(\epsilon_1, \cdots, \epsilon_{p+1}) \underset{def}{=} G_j[\overset{\epsilon}{Y}(x)] = L_j$. If $\varphi(\epsilon_1, \cdots, \epsilon_{p+1}) = J[\overset{\epsilon}{Y}(x)]$ is to have an extremum subject to the p constraints $\gamma^{(j)}(\bar{\epsilon}) = L_j$, at $(\bar{\epsilon}) = (\bar{0})$. Then by the Lagrange

[38] Alternatively, assume $y \in C^2$; then $1 - \lambda \frac{d}{dx}(\frac{y'}{\sqrt{1+(y')^2}}) = 0$ gives $\frac{y''}{(1+(y')^2)^{\frac{3}{2}}} = \frac{1}{\lambda}$; The left side is the curvature. Therefore $\mathcal{C}$ must have constant curvature $\frac{1}{\lambda}$, i.e. a circle of radius λ.

multiplier rule there exist $\lambda_0, \cdots, \lambda_p$ not all 0 such that

$$\lambda_0 \frac{\partial \varphi}{\partial \epsilon_k}\Big|_{\bar{\epsilon}=\bar{0}} + \lambda_1 \frac{\partial \gamma^{(1)}}{\partial \epsilon_k}\Big|_{\bar{\epsilon}=\bar{0}} + \cdots + \lambda_p \frac{\partial \gamma^{(p)}}{\partial \epsilon_k}\Big|_{\bar{\epsilon}=\bar{0}} = 0, \quad k = 1, \cdots, p+1,$$

where $\lambda_0, \cdots, \lambda_p$ must to begin with be considered as depending on $\eta_1(X), \cdots, \eta_{p+1}(x)$. This becomes, after the usual integrations by parts,

$$\lambda_0[\eta_1, \cdots, \eta_{p+1}] \int_{x_1}^{x_2} (\dot{f}_{y_i} - \frac{d}{dx}\dot{f}_{y_i'})\eta_k dx + \Sigma_{j=1}^{p} \lambda_j[\eta_1, \cdots, \eta_{p+1}] \int_{x_1}^{x_2} (\dot{g}_{y_i}^{(j)} - \frac{d}{dx}\dot{g}_{y_i'}^{(j)})\eta_k dx = 0$$

Ignoring the last of these $p+1$ relations, we se that the ratios $\lambda_0 : \lambda_1 : \cdots : \lambda_p$ are independent of η_{p+1}; similarly they are independent of the other η_k. Hence they are constants. Apply the fundamental lemma to

$$\int_{x_1}^{x_2} [\lambda_0(\dot{f}_{y_i} - \frac{d}{dx}\dot{f}_{y_i'}) + \Sigma_{j=1}^{p} \lambda_j (\dot{g}_{y_i}^{(j)} - \frac{d}{dx}\dot{g}_{y_i'}^{(j)})]\eta_k dx = 0$$

now yields the $i-$th equation of (44), except for the possible occurrence of $\lambda_0 = 0$. But $\lambda_0 = 0$ is ruled out if $\dot{Y}(x)$ is not an extremal of $\Sigma_{j=1}^{p} \lambda_j G_j[Y]$. **q.e.d.**

There are two other types of constraints subject to which $J[y] = \int_{x_1}^{x_2} f(x, Y, Y')$ may have to be minimized. These are both called Lagrange problems. **(1)-Finite, or holonomic constraints:** e.g. for $n = 2, p(\text{number of constraints}) = 1$, minimize $J[y(x), z(x)] = \int_{x_1}^{x_2} f(x, y(x), z(x), y'(x)z'(x))dx$ among admissible (say C^2) $Y(x)$ satisfying

$$\begin{cases} y(x_1) = y_1, y(x_2) = y_2, & \text{E} \\ g(x, y(x), z(x)) = 0, & \text{C} \end{cases}$$

where f, g are assumed to be C^2 in some (x, y, z) region, $\mathcal{R}$ and for all y', z'.

Notice that the constraint (C) requires the curve $x = x; y = y(x); z = z(x)$ to lie on the surface given by $g(x, y, z) = 0$. The end point conditions, (E) are complete in the sense that $z(x_1), z(x_2)$ are determined by the combination of (E),(C).

(2)-Non-holonomic constraints. i.e. differential-equation type constraints: Again we consider $n = 2, p = 1$. To minimize $J[y, z]$ as above subject to

$$\begin{cases} y(x_1) = y_1, y(x_2) = y_2, z(x_2) = z_2, & \mathrm{E}' \\ q(x, y(x), z(x), y'(x), z'(x)) = 0, & \mathrm{C}' \end{cases}$$

with $q(x, y, z, y', z') \in C^2$ for $(x, y, z) \in \mathcal{R}$ and for all y', z'.

Notice that the end point conditions, E' are appropriate, in that C' allows solving, say for $z' = Q(x, y(x), y'(x), z(x))$. Thus for given $y(x)$, the corresponding $z(x)$ is obtained as the solution of a first order differential equation, $z' = \widetilde{Q}(x, z)$, for which the single boundary condition $z(x_2) = z_2$ is appropriate to determine a unique solution, $z(x)$.

For both the Lagrange problems the first necessary conditions on a minimizing curve, viz. the appropriate Euler-Lagrange multiplier rule is similar to the corresponding result for the case of isoperimetric constraints, with the difference that the Lagrange multiplier, λ must be now set up as a function of $x : \lambda = \lambda(x)$.

Theorem 27.8. *(i) Let $\overline{Y} = (\overline{y}(x), \overline{z}(x))$ be a minimizing curve for the holonomic problem, (1) above. Assume also that $\overline{g}_z = g_z(x, \overline{y}(x), \overline{z}(x)) \neq 0$[39] on $[x_1, x_2]$. Then there exist $\lambda(x) \in C^0$ such that $\overline{Y}(x)$ satisfies the Euler-Lagrange -equations for a free minimizing*

[39] The proof is similar if we assume $\overline{g}_y \neq 0$.

problem

$$\int_{x_1}^{x_2} (f(x,y,z,y',z') + \lambda(x)g(x,y,z))dx.$$

(ii) Let $\overline{Y}(x) = (\overline{y}(x), \overline{z}(x))$ be a minimizing curve for the non-holonomic problem, (2) above. Assume that $\overline{q}_{z'} \neq 0$ on $[x_1, x_2]$[40]*. Then there exist $\lambda(x) \in C^0$ such that $\overline{Y}$ satisfies:*

$$(\overline{f}_{z'} + \lambda(x)\overline{q}_{z'})\Big|_{x=x_1} = 0$$

and the necessary Euler-Lagrange -equation for a free minimum of

$$\int_{x_1}^{x_2} (f(x,y,z,y',z') + \lambda(x)q(x,y,z,y',z'))dx.$$

Proof : of (i) Say $g(x,y,z) \in C^2$. Since $\overline{g}_z \neq 0$ on $[x_1, x_2]$, where $(\overline{y}(x), \overline{z}(x))$ is the assumed minimizing curve, we still have $g_z \neq 0$ for all (x,y,z) sufficiently near points $(x, \overline{y}(x), \overline{z}(x))$. Hence we can solve $g(x,y,z) = 0$ for $z : z = \varphi(x,y)$. Hence for any curve $(y(x), z(x))$ on $g(x,y,z) = 0$ and sufficiently near $(\overline{y}(x), \overline{z}(x))$ $z'(x) = \varphi_x + \varphi_y y'(x); \varphi(x) = -\frac{g_x}{g_z}; \varphi_y = -\frac{g_y}{g_z}$. Then for any sufficiently close neighboring curve $(y(x), z(x))$ on $g(x,y,z) = 0$

$$J[y(x), z(x)] = \int_{x_1}^{x_2} f(x, y(x), z(x), y'(x), z'(x))dx =$$

$$\int_{x_1}^{x_2} \underbrace{f(x, y(x), \varphi(x, y(x)), y'(x), \varphi_x(x, y(x)) + \varphi_y(x, y(x))y'(x))}_{=F(x,y(x),y'(x))} dx$$

Thus $\overline{y}(x)$ must minimize $\int_{x_1}^{x_2} F(x, y(x), y'(x))dx$ (We have reduced the number of component functions by 1) and therefore must satisfy the Euler-Lagrange equation for F. i.e.

$$\overline{f}_y + \overline{f}_z\overline{\varphi}_y + \overline{f}_{z'}(\overline{\varphi}_{xy} + \overline{\varphi}_{yy}\overline{y}') - \frac{d}{dx}(\overline{f}_{y'} + \overline{f}_{z'}\overline{\varphi}_y) = 0$$

[40]The proof is similar if we assume $\overline{q}_{z'} \neq 0$.

or

$$\overline{f}_y + \overline{f}_z \overline{\varphi}_y - \frac{d}{dx}\overline{f}_{y'} - \overline{\varphi}_y \frac{d}{dx}\overline{f}_{z'} = 0.$$

Now note that $\overline{\varphi}_y = -\frac{\overline{g}_y}{\overline{g}_z}$:

$$(\overline{f}_y - \frac{d}{dx}\overline{f}_{y'}) - \frac{\overline{g}_y}{\overline{g}_z}(\overline{f}_z - \frac{d}{dx}\overline{f}_{z'}) = 0.$$

Setting $-\lambda(x) = \frac{\overline{f}_z - \frac{d}{dx}\overline{f}_{z'}}{\overline{g}_z}$ (note that $\lambda(x) \in C^0$), the last condition becomes $\overline{f}_y - \frac{d\overline{f}_{y'}}{dx} + \lambda(x)\overline{g}_y = 0$. Hence (since $\overline{g}_{y'} = \overline{g}_{z'} = 0$) we have

$$\begin{cases} \overline{f}_y + \lambda(x)\overline{g}_y - \frac{d}{dx}(\overline{f}_{y'} + \lambda(x)\overline{g}_{y'}) & = 0 \\ \overline{f}_z + \lambda(x)\overline{g}_z - \frac{d}{dx}(\overline{f}_{z'} + \lambda(x)\overline{g}_{z'}) & = 0 \end{cases}$$

which are the Euler-Lagrange equations for $\int_{x_1}^{x_2}(f(x,y,z,y',z') + \lambda(x)g(x,y,z))dx$.

Proof of (ii): Let $(\overline{y}(x), \overline{z})$ be the assumed minimizing curve for $J[y,z]$, satisfying $q(x,y,z,y',z') = 0$. As usual we imbed $\overline{y}(x)$ in a $1-$parameter family $\{y\epsilon(x)\} = \{\overline{y}(x) + \epsilon\eta(x)\}$, where $\epsilon \in C^2$ and satisfies the usual boundary condition, $\eta(x-i) = 0, i = 1,2$. We proceed to obtain a matching z_ϵ for each y_ϵ by means of the differential equation

$$\overset{(\epsilon)}{q} \underset{def}{=} q(x, y_\epsilon(x), z_\epsilon(x), y'_\epsilon(x), z'_\epsilon(x)) = 0 \tag{45}$$

To obtain z_ϵ, note that: (i) for $\epsilon = 0$: $\quad y_0(x) = \overline{y}(x)$ and (45) has the solution $z_0(x) = \overline{z}(x)$, satisfying $\overline{z}(x_2) = z_2$: (ii) $\overline{q}_{z'} \neq 0$ by hypothesis, therefore (q being of class C^2), $q_{z'}(x, y_\epsilon(x), z_\epsilon(x), y'_\epsilon(x), z'_\epsilon(x)) \neq 0$ for ϵ sufficiently small, and (z, z') sufficiently near

$(\overline{z}(x), \overline{z}'(x))$. For such ϵ, z, z' we can solve (45) for z' obtaining an explicit differential equation

$$z' = Q(x, y_\epsilon(x), z, y_\epsilon(x)) = \widetilde{Q}(x, z; \epsilon)$$

To this differential equation whose right-hand side depends on a parameter, ϵ, apply the relevant Picard Theorem, [AK, page 29], according to which there is a unique solution, $z_\epsilon(x)$ such that $z_\epsilon(x_2) = z_2$ and $\frac{\partial z_\epsilon}{\partial \epsilon}$ exists and is continuous. Note that $z_\epsilon(x_2) = z_2$ for all ϵ implies $\frac{\partial z_\epsilon}{\partial \epsilon}\Big|_{x=x_2} = 0$.

We have constructed a 1-parameter family of curves, $\{Y_\epsilon(x)\} = \{y_\epsilon(x), z_\epsilon(x)\} = \{\overline{y}(x) + \epsilon\eta(x), z_\epsilon(x)\}$ that consists of admissible curves all satisfying E' and C'. Since, among them, $(\overline{y}(x), \overline{z}(x))$ (for $\epsilon = 0$) minimizes $J[y, z]$, we must have

$$0 = \frac{dJ[y_\epsilon(x), z_\epsilon(x)]}{d\epsilon}\Big|_{\epsilon=0} = \int_{x_1}^{x_2} (\overline{f}_y \eta + \overline{f}_{y'}\eta' + \overline{f}_z \frac{\partial z_\epsilon}{\partial \epsilon}\Big|_{\epsilon=0} + \overline{f}_{z'} \frac{\partial z'_\epsilon}{\partial \epsilon}\Big|_{\epsilon=0}) dx$$

Combining this with $0 = \frac{d\overset{(\epsilon)}{q}}{d\epsilon}\Big|_{\epsilon=0} = \overline{q}_y \eta + \overline{q}_{y'}\eta' + \overline{q}_z \frac{\partial z_\epsilon}{\partial \epsilon}\Big|_{\epsilon=0} + \overline{q}_{z'} \frac{\partial z'_\epsilon}{\partial \epsilon}\Big|_{\epsilon=0}$ the necessary condition yields

$$\int_{x_1}^{x_2} [(\overline{f}_y + \lambda \overline{q}_y)\eta + (\overline{f}_{y'} + \lambda \overline{q}_{y'})\eta' + (\overline{f}_z + \lambda \overline{q}_z)\frac{\partial z_\epsilon}{\partial \epsilon}\Big|_{\epsilon=0}) + (\overline{f}_{z'} + \lambda \overline{q}_{z'})\frac{\partial z'_\epsilon}{\partial \epsilon}\Big|_{\epsilon=0}] dx = 0$$

where at this point $\lambda(x)$ my be any integrable function of x. Assuming $\lambda(x) \in C^1$, we integrate by parts, as usual using the boundary conditions, $\eta(x_i) = 0, i = 1, 2$; $\frac{\partial z_\epsilon(x_2)}{\partial \epsilon} = 0$ we obtain:

$$\int_{x_1}^{x_2}\Big([(\overline{f}_y+\lambda\overline{q}_y)-\frac{d}{dx}(\overline{f}_{y'}+\lambda\overline{q}_{y'})]\eta+[(\overline{f}_z+\lambda\overline{q}_z)-\frac{d}{dx}(\overline{f}_{z'}+\lambda\overline{q}_{z'})]\frac{\partial z_\epsilon}{\partial\epsilon}\Big|_{\epsilon=0}\Big)dx$$

$$-(\overline{f}_{z'}+\lambda\overline{q}_{z'})\frac{\partial z_\epsilon}{\partial\epsilon}\Big|_{\epsilon=0,x=x_1}=0 \tag{46}$$

On the arbitrary C^1 function, $\lambda(x)$ we now impose the conditions:

1. $(\overline{f}_{z'}+\lambda(x)\overline{q}_{z'})\Big|_{x=x_1}=0$[41]

2. $(\overline{f}_z+\lambda(x)\overline{q}_z)-\frac{d}{dx}((\overline{f}_{z'}+\lambda(x)\overline{q}_{z'})=0,\ \text{all } x\in[x_1,x_2]$

Then (2) is a 1-st order differential equation for $\lambda(x)$ and (1) is just the right kind of associated boundary condition on $\lambda(x)$ (at x_1). Thus (1) and (2) determine $\lambda(x)$ uniquely in $[x_1,x_2]$, by the Picard theorem. With this λ (46) becomes

$$\int_{x_1}^{x_2}(\overline{f}_y+\lambda(x)\overline{q}_y)-\frac{d}{dx}(\overline{f}_{y'}+\lambda(x)\overline{q}_{y'})\eta_(x)dx=0$$

for all $\eta(x)\in C^2$ such that $\eta(x_i)=0, i=1,2$. Hence by the fundamental lemma $(\overline{f}_y+\lambda\overline{q}_y)-\frac{d}{dx}(\overline{f}_{y'}+\lambda(x)\overline{q}_{y'})=0$, which together with (a) and (b) above prove the conclusion of part (ii) of the theorem. **q.e.d.**

Note on the procedure: The determination of $\overline{y}(x),\overline{z}(x),\lambda(x)$ satisfying the necessary conditions is, by the theorem, based on the following relationns:

[41]Recall that $\overline{q}_{z'}\neq 0$ by hypothesis.

(i) Holonomic case:

$$\begin{cases} \overline{f}_y + \lambda \overline{g}_y - \frac{d}{dx}\overline{f}_{y'} & = 0 \\ \overline{f}_z + \lambda \overline{g}_z - \frac{d}{dx}\overline{f}_{z'} & = 0 \\ g(x, \overline{y}, \overline{z}) & = 0 \end{cases} \quad \text{plus} \quad \Big\{ \; y(x_1) = y_1, \quad y(x_2) = y_2 \quad ;$$

(ii) Non-holonomic case:

$$\begin{cases} \overline{f}_y + \lambda \overline{q}_y - \frac{d}{dx}(\overline{f}_{y'} + \lambda \overline{q}_{y'}) & = 0 \\ \overline{f}_z + \lambda \overline{q}_z - \frac{d}{dx}(\overline{f}_{z'} + \lambda \overline{q}_{z'}) & = 0 \\ q(x, \overline{y}, \overline{z}, \overline{y}', \overline{z}') & = 0 \end{cases} \quad \text{plus} \quad \begin{cases} y(x_1) = y_1, y(x_2) = y_2, z(x_2) = z_2 \\ (\overline{f}_{z'} + \lambda \overline{q}_{z'})\Big|_{x=x_1} = 0 \end{cases}$$

Examples: (a) Geodesics on a surface. Recall that for a space curve, $\mathcal{C} : (x(t), y(t), z(t)), t_1 \leq t \leq t_2$, the unit tangent vector $\overrightarrow{T}$, is given by

$$\overrightarrow{T} = \frac{1}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}(\dot{x}(t), \dot{y}(t), \dot{z}(t)) = (\frac{dx}{ds}, \frac{dy}{ds}, \frac{dz}{ds})$$

($s =$ arc length).

The vector $\overrightarrow{P} \underset{def}{=} \frac{d\overrightarrow{T}}{ds} = (\frac{d^2x}{ds^2}, \frac{d^2y}{ds^2}, \frac{d^2z}{dx^2}) = \frac{d}{dt}(\frac{\dot{x}(t)}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}}, \frac{\dot{y}(t)}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}}, \frac{\dot{z}(t)}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}})\frac{dt}{ds}$ is the principal normal. The plane through $Q_0 = (x(t_0), y(t_0), z(t_0))$ spanned by $\overrightarrow{T}(t_0), \overrightarrow{P}(t_0)$ is the osculating plane of the curve $\mathcal{C}$ at Q_0.

If the curve lies on a surface $S : g(x, y, z) = 0$, the direction of the principal normal, $\overrightarrow{P}_0$ at the point Q_0 does not in general coincide with the direction of the surface normal, $\overrightarrow{N}_0 = (g_x, g_y, g_z)\Big|_{Q_0}$ at Q_0. The two directions do coincide, however, for any curve, $\mathcal{C}$ on S that is the shortest connection on S between its end points. To prove this using (i) we assume $\mathcal{C}$ minimizes $J[x(t), y(t), z(t)] = \int_{t_1}^{t_2} \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} dt$, subject to $g(x, y, z) = 0$. By the theorem, necessary conditions on $\mathcal{C}$ are (since $f_x = g_{\dot{x}} = 0$, $f_{\dot{x}} = \frac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}}$, etc.):

$$(47) \qquad \begin{cases} \lambda(t) g_x &= \frac{d}{dt}\left(\frac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}}\right) \\ \lambda(t) g_y &= \frac{d}{dt}\left(\frac{\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}}\right) \\ \lambda(t) g_x &= \frac{d}{dt}\left(\frac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}}\right) \end{cases}$$

That is to say $\lambda(t)\nabla_g = \frac{ds}{dt}\overrightarrow{P}$ or $\lambda(t)\overrightarrow{N} = \frac{ds}{dt}\overrightarrow{P}$. Hence, without having to solve them, the Euler-Lagrange equations, (47) tell us that for a geodesic $\mathcal{C}$ on a surface, S, the osculating plane of $\mathcal{C}$, is always $\perp$ to the tangent plane of S.

(b) Geodesics on a sphere. For the particular case $g(x, y, z) = x^2, y^2, z^2 - R^2$ we shall actually solve the Euler-Lagrange equations. Setting $\mu = \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}$, system (47) yields

$$\frac{\frac{d}{dt}\left(\frac{\dot{x}}{\mu}\right)}{2x} = \frac{\frac{d}{dt}\left(\frac{\dot{y}}{\mu}\right)}{2y} = \frac{\frac{d}{dt}\left(\frac{\dot{z}}{\mu}\right)}{2z}$$

whence

$$\frac{\ddot{x}y - x\ddot{y}}{\dot{x}y - x\dot{y}} = \frac{\ddot{x}z - y\ddot{z}}{\dot{y}z - y\dot{z}},$$

where the numerators are the derivatives of the denominators. So we have

$$\ln(\dot{x}y - x\dot{y}) = \ln(\dot{y}z - y\dot{z}) + \ln c_1,$$

and $\frac{\dot{y}}{y} = \frac{(x+c_1)^\bullet}{x+c_1 z}$. Hence $x - c_2 y + c_1 z = 0$: a plane through the origin. Thus geodesics are great circles.

(c) An example of Non-holonomic constraints. Consider the following special non-holonomic problem: Given $f(x,y,y'), g(x,y,y')$, minimize $J[y,z]=\int_{x_1}^{x_2} f(x,y,y')dx$ subject to

$$\left.\begin{array}{ll} y(x_1)=y_1, & y(x_2)=y_2 \\ z(x_1)=0, & z(x_2)=L \end{array}\right\}(E'); \quad q(x,y,z,y',z')=z'-g(x,y,y')=0(C')$$

The end point conditions, (E') contain more than they should in accordance with the definition of a non-holonomic problem (viz., the extra condition $z(x_1)=0$), which is alright if it does not introduce inconsistency. In fact by (C') we have:

$$\int_{x_1}^{x} g(\overline{x},y,y')d\overline{x} = \int_{x_1}^{x_2} z'd\overline{x} = z(x)-z(x_1)$$

Hence

$$\int_{x_1}^{x_2} g(x,y,y')dx = z(x_2)-z(x_1)=L;$$

Therefore this special non-holonomic problem is equivalent with the isoperimetric problem.

Note: In turn, a holonomic problem may in a sense be considered as equivalent to a problem with an infinity of isoperimetric constraints, see [GF][remark 2 on page 48]

(d) Another non-holonomic problem. Minimize

$$J[y,z]=\int_{x_1}^{x_2} f(x,y,z,z')dx \quad \text{subject to} \left\{\begin{array}{ll} \left.\begin{array}{ll} y(x_1)=y_1, & y(x_2)=y_2 \\ z(x_1)=z_1, & z(x_2)=z_2 \end{array}\right\} & (E') \\ q(x,y,z,y',z')=z-y'=0 & (C') \end{array}\right.$$

Substituting from (C') into $J[y,z]$ and (E'), this becomes:

Minimize $\int_{x_1}^{x_2} f(x,y,y',y'')dx$ subject to $y(x_i)=y_i; \quad y'(x_i)=z_i, \quad i=1,2.$

Hence this particular non-holonomic Lagrange problem is equivalent to the higher derivative problem treated earlier.

Some exercises:

(a) Formulate the main theorem of this section for the case of n unknown functions, $y_1(x), \cdots, y_n(x)$ and $p(< n)$ holonomic or p non-holonomic constraints.

(b) Formulate the isoperimetric problem (n functions, p isoperimetric constraints) as a non-holonomic Lagrange problem.

(c) Formulate the higher derivative problem for general m as a non-holonomic Lagrange problem.

28 Functionals of Vector Functions: Fields, Hilbert Integral, Transversality in Higher Dimensions.

0- Preliminaries: Another look at fields, Hilbert Integral and Transversality for $n = 1$.

(a) Recall that a simply connected field of curves in the plane was defined as: A simply connected region, $\mathcal{R}$ plus a 1-parameter family $\{y(x,y)\}$ of sufficiently differentiable curves- the trajectories of the field- covering $\mathcal{R}$ simply and completely.

Given such a field, $\mathfrak{F} = (\mathcal{R}, \{y(x,t)\})$ we then defined a slope function, $p(x,y)$ as having, at any (x,y) of $\mathcal{R}$, the value $p(x,y) = y'(x,t_0)$, the slope at (x,y) of the field trajectory, $y(x,t_0)$ through (x,y).

With an eye on $n > 1$, we shall be interested in the converse procedure: Given (sufficiently differentiable) $p(x,y)$ in $\mathcal{R}$, we can construct a field of curves for which p is the slope function, viz. by solving the differentiable equation $\frac{dy}{dx} = p(x,y)$ in $\mathcal{R}$; its 1-parameter family $\{y(x,t)\}$ of solution curves is, by the Picard existence and uniqueness theorem, precisely the set of trajectories for a field over $\mathcal{R}$ with slope function p.

(b) Next, given any slope function, $p(x,y)$ in $\mathcal{R}$, and the functional $J[y] = \int_{x_1}^{x_2} f(x,y,y')dx$ where $f \in C^3$ for $(x,y) \in \mathcal{R}$ and for all y', construct the Hilbert integral

$$I_B^* = \int_B \langle [f(x,y,p(x,y)) - p(x,y)f_{y'}(x,y,p(x,y))]dx + f_{y'}(x,y,p(x,y))dy$$

for any curve B in $\mathcal{R}$. Recall from (10.2), second proof, that I_B^* is independent of the path

B in $\mathcal{R}$ if and only if $p(x, y)$ is the slope function of a field of extremals for $J[y]$. i.e. if and only if $p(x, y)$ satisfies

$$f_y((x, y, p(x, y)) - \frac{d}{dx} f_{y'}(x, y, p(x, y)) = 0 \quad \text{for all } (x, y) \in \mathcal{R}.$$

(c) In view of (a) and (b) a 1-parameter field of extremals of $J[y] = \int_{x_1}^{x_2} f(x, y, y')dx$, in brief $\underline{\text{a field for } J[y]}$, can be characterized by: $\underline{\text{A simply connected region, } \mathcal{R}\text{, plus a function,}}$ $\underline{p(x, y) \in C^1 \text{ such that the Hilbert integral is independent of the path } B \in \mathcal{R}.}$ This characterization serves, for $n = 1$ as an alternative definition of a field for the functional $J[y]$.

For $n > 1$, the analogue of this characterization will be taken as the actual definition of a field for $J[y]$. But for $n > 1$ this is not equivalent to postulating an $(x, y_1, \cdots, y_n)$ region $\mathcal{R}$ plus an n-parameter family of extremals for $J[y]$ covering $\mathcal{R}$ simply and completely. It turns out to be more demanding than this.

(d) (Returning to $n = 1$.) Recall, §9, that a curve $N : x = x_N(t), y = y_N(t)$ is $\underline{\text{transversal}}$ to a field for $J[y]$ with slope function $p(x, y)$ if and only if the line elements (x, y, dx, dy) of N satisfy

$$\overbrace{[f(x, y, p(x, y)) - p(x, y)f_{y'}(x, y, p(x, y))]}^{=A(x,y)} dx + \overbrace{f_{y'}(x, y, p(x, y))}^{=B(x,y)} dy = 0$$

Hence, given $p(x, y)$, we can determine the transversal curves, N to the field of trajectories as the solution of the differential equation $Adx + Bdy = 0$.

This differential equation is written as a $\underline{\text{Pfaffian equation}}$. A Pfaffian equation has an integrating factor, $\mu(x, y)$ if $\mu(Adx + Bdy)$ is exact, i.e. if $Adx + Bdy = d\varphi(x, y)$. A two dimensional Pfaffian equation always has such an integrating factor. To see this consider $\frac{dy}{dx} = -\frac{A}{B}$ (or $\frac{dx}{dy} = -\frac{B}{A}$). This equation has a solution, $F(x, y) = c$. Then $\frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy = 0$. This is a Pfaffian equation which is exact (by construction) and has the same solution as the original differential equation. As a consequence the two equations must differ by a factor (the integrating factor). In summary, for $n = 1$, one can always find a curve, N which is transversal to a field.

By contrast for $n > 1$ a Pfaffian equation, $A(x, y, z)dx + B(x, y, z)dy + C(x, y, z)dz = 0$ has, in general no integrating factors (i.e. is not integral). A necessary and sufficient condition for integrability is the case $n = 2$ is that $\overrightarrow{V} \cdot curl\overrightarrow{V} = 0$, where $\overrightarrow{V} = (A, B, C)$.[42] A sufficient condition is exactness of the differential form on the left, which for $n = 2$ means $curl\overrightarrow{V} = 0$.

(e) Let N_1, N_2, be two transversals, to a field for $J[y]$. Along each extremal $E_t = E_{\widehat{P_1(t)P_2(t)}}$ joining $P_1(t)$ of N_1 to $P_2(t)$ of N_2, we have $I^*_{E_{\widehat{P_1P_2}}} = J[E_{\widehat{P_1P_2}}] = \mathfrak{I}(t)$ (18), and by (16): $\frac{\mathfrak{I}(t)}{dt} = 0$ Hence I^* and J have constant values along the extremals from N_1 to N_2.

Conversely given on transversal, N_1, construct a curve N as the locus of points P, on the extremals transversal to N_1, such that $I^*_{E_{\widehat{P_1P_2}}} = J[E_{E_{\widehat{P_1P_2}}}] = c$ where c is a constant. The equation of the curve, N is then $\int_{(x_1(t),y_1(t))}^{(x,y)} f(x, y(\overline{x}, t), y'(\overline{x}, t))d\overline{x} = c$. Then N is also a transversal curve to the given field, $(R, \{y(x, t)\})$ of $J[y]$. This follows since $\frac{d\mathfrak{I}}{dt} = 0$ by the construction of N and (15).

[42]See, for example, [A].

1-Fields, Hilbert Integral, and Transversality for $n \geq 2$.

(a) Transversality In $n+1$- space, $\mathbb{R}^{n+1}$ with coordinates $(x, Y) = (x, y_1, \cdots, y_n)$, consider two hypersurfaces $S_1 : W_1(x, Y) = c_1, \quad S_2 : W_2(x, Y) = c_2$; also a 1-parameter family of sufficiently differentiable curves, $Y(x, t)$ with initial points along a smooth curve $C_1 : (x_1(t), Y_1(t))$ on S_1, and end points on $C_2 : (x_2(t), Y_2(t))$, a smooth curve on S_2.

Set $\mathfrak{I}(t) = J[Y(x,t)] = \int_{x_1(t)}^{x_2(t)} f(x, Y(x,t), Y'(x,t))dx$, where $f \in C^3$ in an $(x, Y)-$region that includes S_1, S_2 and the curves $Y(x, t)$, for all Y'. Then

$$\frac{d\mathfrak{I}}{dt} = \Big[(f - \sum_{i=1}^{n} y_i' f_{y_i'})\frac{dx}{dt} + \sum_{i=1}^{n} f_{y_i'}\frac{dy_i}{dt}\Big]_1^2 + \int_{x_1(t)}^{x_2(t)} \sum_{i=1}^{n} \frac{\partial y_i}{\partial t}\Big[f_{y_i} - \frac{d}{dx} f_{y_i'}\Big]dx. \tag{48}$$

(The same proof as for (15)). Hence if $Y(x, t_0)$ is an extremal, $f_{y_i} - \frac{d}{dx} f_{y_i'} = 0$ and we have

$$\frac{d\mathfrak{I}}{dt}\Big|_{t=t_0} = \Big[(f - \sum_{i=1}^{n} y_i' f_{y_i'})\frac{dx}{dt} + \sum_{i=1}^{n} f_{y_i'}\frac{dy_i}{dt}\Big]_1^2. \tag{49}$$

Definition 28.1. An extremal curve, $\overline{Y}(x)$ of $J[Y]$ and a hypersurface $S : W(x, Y) = c$ are transversal at a point (x, Y) where they meet if and only if

$$\Big[(f - \sum_{i=1}^{n} y_i' f_{y_i'})\delta x + \sum_{i=1}^{n} f_{y_i'}\delta y_i\Big] = 0$$

holds for all directions $(\delta x, \delta y_1, \cdots, \delta y_n)$ for which $(\delta x, \delta y_1, \cdots, \delta y_n) \cdot \nabla W\Big|_{(x,Y)} = 0$

Application of (48) and (49). Let $Y(x, t_1, \cdots, t_n)$ be an n parameter family of extremals for $J[Y]$. Assume $S_0 : W(, Y) = c$ is a hypersurface transversal to all $Y(x, t_1, \cdots, t_n)$. Then there exist a 1-parameter family $\{S_k\}$ of hypersurfaces S_k transversal to the same extremals defined as follows: Assign k, proceed on each extremal, E from its point P_0 of intersection with S_0 to a point P_k such that $J[E_{\widehat{P_1P_2}}] = \int_{E\,P_0}^{P_k} f(x, Y, Y')dx = k$. This equation defines a hypersurface, S_k, the locus of all points P_k such that $\frac{d\mathfrak{I}(t)}{dt} = 0$ holds for any curves, C_1, C_2 on S_0, S_k respectively. Hence by (48) and (49), the transversality consition must hold for S_k if it holds for S_0.

Exercise: Prove that if $J[Y] = \int_{x_1}^{x_2} n(x, Y)\sqrt{1 + (y_1')^2 + \cdots + (y_n')^2}dx$ then transversality is equivalent to orthogonality. Is the converse true?

(b) Fields, Hilbert Integral for $n \geq 2$. Preliminary remark: For a plane curve $C : y = y(x)$, oriented in the direction of increasing x, the tangent direction can be specified by the pair $(1, y') = (1, p)$ of direction numbers, where p is the slope of C. Similarly, for a curve, $C : x = x, y_1 = y_1(x), \cdots, y_n = y_n(x)$ of $\mathbb{R}^{n+1}$, oriented in the direction of increasing x, the tangent direction can be specified by $(1, y_1', \cdots, y_n') = (1, p_1, \cdots, p_n)$. The n-tuple $(p_1, \cdots, p_n)$ is called the slope of C.

Definition 28.2. By a field for the functional $J[y] = \int_{x_1}^{x_2} f(x, Y, Y')dx$ is meant:

(i) a simply connected region, R of $(x, y_1, \cdots, y_n)$-space such that $f \in C^3$ in R.

plus

(ii) a slope function $P(x, Y) = (p_1(x, Y), \cdots, p_n(x, Y))$ of class C^1 in R for which the Hilbert Integral

$$I_B^* = \int_B \langle [f(x, y, p(x, y)) - p(x, y) f_{y'}(x, y, p(x, y))]dx + f_{y'}(x, y, p(x, y))dy$$

is independent of the path B in R.

This definition is motivated by the introductory remarks at the beginning of this section. It is desirable to have another characterization of this concept, in terms of a certain type of family of curves covering R simply and completely. We give two such characterizations.

To begin with the curves, or trajectories, of he field will be solutions of the 1-st order system $\frac{dY}{dx} = P(x, Y)$, i.e. of

$$\frac{dy_i}{dx} = p_i(x, y_1, \cdots, y_n), \quad i = 1, \cdots n.$$

By the Picard theorem, there is one and only one solution curve through each point of R. The totality of solution curves is an n-parameter family $\{Y(x, t_1, \cdots, t_n)\}$.

Before characterizing these trajectories further, note that along each of them, $I_T^* = J[T]$ holds. This is clear if we substitute $dy_i = p_i(x, Y)dx, (i = c, \cdots, n)$ into the Hilbert integral. Now for the first characterization:

Theorem 28.3. *A. (i) The trajectories of a field for $J[Y]$ are extremals of $J[Y]$.*

(ii) This n-parameter subfamily of (the $2n$-parameter family of all) extremals does have transversal hypersurfaces.

B. *Conversely: If a simply-connected region, R of $\mathbb{R}^{n+1}$ is covered simply and completely by an n-parameter family of extremals for $J[y]$ that does admit transversal hypersurfaces, then R together with the slope function $P(x,Y)$ of the family consists of a field for $J[Y]$.*

Comment:

(a) For $n = 1$, a 1-parameter family of extremals, given by $\frac{dy}{dx} = p(x,y)$, always has transversal curves, since $A(x,y)dx + B(x,y)dy$ is always integrable (see part (d) of the preliminaries). But for $n > 1$ the transversality condition is of the form

$$A(x,Y)dx + B_1(x,Y)dy_1 + \cdots + B_n(x,Y)dy_n = 0$$

which has solutions surface, S (transversal to the extremals) only if certain integrability conditions are satisfied. (see part (d) of the preliminaries).

(b) The existence of a transversal surface is related to the Frobenius theorem. Namely one may show (for $n = 2$) that the Frobenius theorem implies that if there is a family of surfaces transversal to a given field then the curl condition in §28 (0-(d)) is satisfied.

Proof : Recall for a line integral, $\int_B C_1(u_1,\cdots,u_m)du_1 + \cdots + C_m(u_1,\cdots,u_m)du_m$ to be independent of path in a simply connected region, R it is necessary and sufficient that the $\frac{1}{2}m(m-1)$ exactness conditions $\frac{\partial C_i}{\partial u_j} = \frac{\partial C_j}{\partial u_i}$ hold in R.

These same conditions are sufficient, but not necessary for the Pfaffian equation

$$C_1(u_1,\cdots,u_m)du_1 + \cdots + C_m(u_1,\cdots,u_m)du_m = 0$$

to be integrable, since if they hold, $C_1(u_1,\cdots,u_m)du_1+\cdots+C_m(u_1,\cdots,u_m)du_m = 0$ is an exact differential. Integrability means there is an integrating factor, $\mu(u_1,\cdots,u_m)$ and a function, $\varphi(u_1,\cdots,u_m)$ such that $\mu(C_1(u_1,\cdots,u_m)du_1+\cdots+C_m(u_1,\cdots,u_m)du_m) = d\varphi$ in R.

A (i) The invariance of the Hilbert integral

$$I_B^* = \int_B \overbrace{[f(x,Y,P(x,Y)) - \sum_{i=1}^n p_i(x,Y)f_{y_i'}(x,Y,P(x,Y))]}^{=A}\,dx + \sum_{i=1}^n \overbrace{f_{y_i'}(x,Y,P(x,Y))}^{=B_i}\,dy_i$$

is equivalent to the $\frac{1}{2}n(n+1)$ conditions (a) $\frac{\partial A}{\partial y_j} = \frac{\partial B_j}{\partial x}$ and (b) $\frac{\partial B_j}{\partial y_k} = \frac{\partial B_k}{\partial y_j}$. Now (b) gives: $\frac{\partial}{\partial y_k}f_{y_j'} = \frac{\partial}{\partial y_j}f_{y_k'}$.

From (a):

$$f_{y_j} + \sum_{i=1}^n f_{y_i'}\frac{\partial p_i}{\partial y_j} - \sum_{i=1}^n \frac{\partial p_i}{\partial y_j}f_{y_i'} - \sum_{i=1}^n p_i\frac{\partial}{\partial y_j}f_{y_i'} = f_{y_j'x} + \sum_{i=1}^n f_{y_j'y_i'}\frac{\partial p_i}{\partial x}$$

$$\text{or } f_{y_j} - \sum_{i=1}^n p_i\frac{\partial}{\partial y_j}f_{y_i'} = f_{y_j'x} + \sum_{i=1}^n f_{y_j'y_i'}\frac{\partial p_i}{\partial x}$$

hence (using (b)):

$$f_{y_j} = f_{y_j'x} + \sum_{i=1}^n f_{y_j'y_i'}\frac{\partial p_i}{\partial x} + \sum_{i=1}^n p_i\frac{\partial}{\partial y_i}f_{y_j'}$$

where $\frac{\partial}{\partial y_i}f_{y_j'} = f_{y_j'y_i} + \sum_{i=1}^n f_{y_j'y_i'}\frac{\partial p_k}{\partial y_i}$ Hence

$$f_{y_j} = f_{y_j'x} + \sum_{i=1}^n p_i f_{y_j'y_i} + \sum_{i=1}^n f_{y_j'y_i'}(\frac{\partial p_i}{\partial x} + \sum_{\ell=1}^n \frac{\partial p_i}{\partial y_\ell}p_\ell)$$

But $p_i = \frac{dy_i}{dx}$ along a trajectory and $(\frac{\partial p_i}{\partial x} + \sum_{\ell=1}^n \frac{\partial p_i}{\partial y_\ell}p_\ell) = \frac{d^2y_i}{dx^2}$ along a trajectory. So any trajectory satisfies

$$f_{y_j} = f_{y_j'x} + \sum_{i=1}^n f_{y_j'y_i'}\frac{\partial y_i}{\partial x} + \sum_{\ell=1}^n \frac{d^2y_i}{dx^2}$$

i.e. $f_{y_j} - \frac{d}{dx} f_{y'_j} = 0$ for all j, so that the trajectories are extremals.

A (ii) Follows from the remark at the beginning of the proof that the exactness conditions are sufficient for integrability; note that the left-hand side in the transversality relation is the integrand of the Hilbert integral.

B Now we are given:

(a) An n-parameter family $Y = Y(x, t_1, \cdots, t_n)$ of extremals with slope function $P(x, Y) = Y'(x, T)$

(b) The family has transversal surfaces, i.e. the Pfaffian equation

$$[f(x, Y, P(x, Y)) - \sum_{i=1}^{n} p_i(x, Y) f_{y'_i}] dx + \sum_{i=1}^{n} f_{y'_i} dy_i = 0$$

i.e. there exist an integrating factor, $\mu(x, Y) \neq 0$ such that

(b1) $\frac{\partial}{\partial y_j}(\mu(f(x, Y, P(x, Y)) - \sum_{i=1}^{n} p_i(x, Y) f_{y'_i}) = \frac{\partial}{\partial x}(\mu f_{y'_i})$ and

(b2) $\frac{\partial(\mu f_{y'_i})}{\partial y_j} = \frac{\partial(\mu f_{y'_j})}{\partial y_i}$

Since R is assumed simply connected, the invariance of I^*_B in R will be established if we can prove the exactness relations (γ):

$(\gamma 1)$ $\frac{\partial}{\partial y_j}(f - \sum_i p_i f_{y'_i}) = \frac{\partial}{\partial x} f_{y'_j}$ and $\qquad (\gamma 2)$ $\frac{\partial f_{y'_i}}{\partial y_j} = \frac{\partial f_{y'_j}}{\partial y_i}$.

Thus we must show that (a) and (b) imply (γ).

By (a): $f_{y_j} - \frac{d}{dx} f_{y'_j} = 0, \quad j = 1, \cdots, n$ for $Y' = P$ since we have extremals; i.e.,

(a') $f_{y_j} - f_{y'_j \cdot x} - \sum_{k=1}^{n} f_{y'_j y_k} \cdot p_k - \sum_{k=1}^{n} f_{y'_j y'_k} \cdot (\frac{\partial p_k}{\partial x} + \sum_{\ell=1}^{n} \frac{\partial p_k}{\partial y_\ell} p_\ell) = 0.$

Now expand (b1) and rearrange, obtaining

(b1') $\mu[f_{y_j} - \sum_{k=1}^{n} p_i \frac{\partial f_{y_i'}}{\partial y_j} - f_{y' \cdot x} - \sum_{k=1}^{n} f_{y_j' y_k'} \cdot p_k] = \mu_x f_{y_j'} - \mu_{y_j}(f - \sum_{i=1}^{n} p_i f_{y_i'})$.

Next we expand (b2) to obtain

(b2') $\mu(\frac{\partial f_{y_j'}}{\partial y_i} - \frac{\partial f_{y_i'}}{\partial y_j}) = \mu_{y_j} f_{y_i'} - \mu_{y_i} f_{y_j'}$.

Multiply (b2') by p_i and sum from 1 to n. This gives

(b2")$\mu(\sum_{i=1}^{n} p_i \frac{\partial f_{y_j'}}{\partial y_i} - \sum_{i=1}^{n} p_i \frac{\partial f_{y_i'}}{\partial y_j}) = \mu_{y_j} \sum_{i=1}^{n} p_i f_{y_i'} - f_{y_j'} \sum_{i=1}^{n} \mu_{y_i} p_i$.

Now subtract (b2") from (b1'):

$$\mu[f_{y_j} - f_{y_j' \cdot x} - \sum_{k=1}^{n} f_{y_j' y_k'} \cdot \frac{\partial p_k}{\partial x} - \sum_{i=1}^{n} p_i \frac{\partial f_{y_j'}}{\partial y_i}] = f_{y_j'}[\mu_x + \sum_{i=1}^{n} \mu_{y_i} p_i] - \mu_{y_j} \cdot f.$$

The factor of μ on the left side of the equality is zero by (a'). The factor of $f_{y_j'}$ on the right side of the equality $= \frac{d\mu}{dx}$ along an extremal (where $P = Y'$). Hence we have, for $j = 1, \cdots, n$

$\mu_{y_j} \cdot f = f_{y_j'} \cdot \frac{d\mu}{dx}\Big|_{\text{in direction of extremal}}$.

Substitute this into $f \cdot$(b2') we obtain:

$\mu f \cdot (\frac{\partial f_{y_j'}}{\partial y_i} - \frac{\partial f_{y_i'}}{\partial y_j}) = 0$. Hence, since $\mu \neq 0$, $f \neq 0$, and $f \in C^3$ ($\gamma 2$) follows. Further, using ($\gamma 2$), it is seen that ($\gamma 1$) reduces to (a). **q.e.d.**

Two examples.

(a) An example of an $n-$paramter family of extremals covering a region, R simply and completely (for $n = 2$) that does not have transversal surfaces, hence is not a field for its functional.

Let $J[y, z] = \int_{x_1}^{x_2} \sqrt{1 + (y')^2 + (z')^2} dx$ (the arc length functional in $\mathbb{R}^3$). The extremals are the straight lines in $\mathbb{R}^3$. Let L_1 be the positive $z-$axis, L_2 the line $y = 1, z = 0$. We shall show that the line segments joining L_1 to L_2 constitute a simple and complete covering of the region R

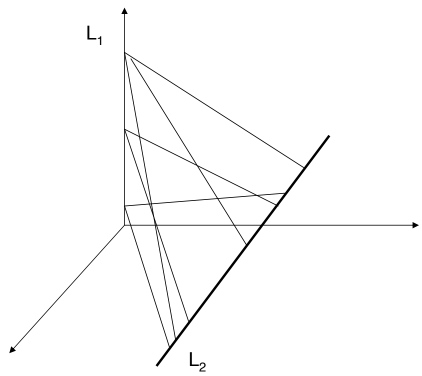


Figure 20: An example of an n-parameter family that is not a field.

To see this we parameterize the lines from L_1 to L_2 by the points, $(0, 0, a) \in L_1, \quad (b, 1, 0) \in L_2$. The line is given by $x = bs, y = x, z = a - as, \quad 0 < s < 1$. It is clear that every point in R lies on a unique line. We compute the slope functions, p_y, p_z for this field by projecting the line through (x, y, z) into the (x, y) plane and the (x, z) plane.

$p_y(x,y,z) = \frac{1}{b} = \frac{y}{x}$. while $p_z(x,y,z) = -\frac{a}{b} = -\frac{zy}{x(1-y)}$. Substituting the slope functions into the Hilbert integral:

$$I_B^* = \int_{x_1}^{x_2} [f(x,y,z,p_y,p_z) - (p_y f_{y'} + p_z f_{z'})]dx + f_{y'}dy + f_{z'}dz$$

yields (a rather complicated) line integral which, by direct computation is not exact. Hence this $2-$ parameter family of extremals does not admit any transversal surface (in this case, transversal = orthogonal).

(b) An example in the opposite direction: Assume that $\{Y(x,t_1,\cdots,t_n)\}$ is an $n-$parameter family of extremals for $J[Y]$ all passing through the same point $P_0 = (x_0, Y_0)$ and covering a simply-connected region, R (from which we exclude the vertex P_0) simply and completely. Then as in the application at the end of **1. (a)**, we can construct transversal surfaces to the family (the surface S_0 of **1. (a)** here degenerates into a single point, P_0, but this does not interfere with the construction of S_k). Hence by part B of the theorem, the $n-$parameter family is a field for $J[Y]$. This is called a central field.

Definition 28.4. An $n-$parameter family of extremals for $J[Y]$ as described in part B of the theorem is called a Mayer family.

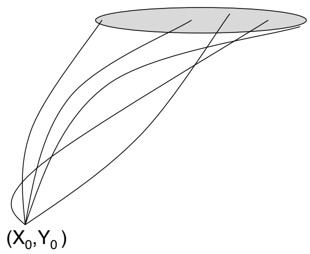


Figure 21: Central field.

We conclude this section with a third characterization of a Mayer family:

c. Suppose we are given an $n-$parameter family $\{Y(x,t_1,\cdots,t_n)\}$ of extremals for $J[Y]$, covering a simply-connected $(x,y_1,\cdots,y_n)$-region, R simply and completely and such that the map given by

$$(x,t_1,\cdots,t_n)\mapsto(x,y_1=y_1(x,t_1,\cdots,t_n),\cdots,y_n=y_n(x,t_1,\cdots,t_n))$$

is one to one from a simply connected $(x,t_1,\cdots,t_n)$-region $\widetilde{R}$ onto R. Using $dy_i=y_i'(x,T)dx+\sum\limits_{k=1}^{n}\frac{\partial y_i(x,T)}{\partial t_k}dt_k$, we change the Hilbert integral I_B^*, formed with the slope function $P(x,Y)=Y'(x,T)$ of the given family, into the new variables, $x,t_1,\cdots,t_n$

$$\begin{aligned}I_b^*&=\int_B[f(x,Y,P(x,Y))-\sum_{i=1}^{n}p_i(x,Y)f_{y_i'}(x,Y,P(x,Y))]dx+\sum_{i=1}^{n}f_{y_i'}(x,Y,P(x,Y))dy_i\\&=\int_{\widetilde{B}}[f(x,T(x,T),Y'(x,T)]dx+\sum_{i,k=1}^{n}f_{y_i'}\frac{\partial y_i}{\partial t_k}dt_k\end{aligned}$$

where $\widetilde{B}$ is the pre-image in $\widetilde{R}$ of B.

Now by the definition and theorem in part **b**,of this section, the given family is a Mayer family if and only if I_B^* is independent of the path in R, hence if and only if the equivalent integral, $\int_{\widetilde{B}}[f(x,T(x,T),Y'(x,T)]dx+\sum\limits_{i,k=1}^{n} f_{y'_i}\frac{\partial y_i}{\partial t_k}dt_k$ is independent of the path $\widetilde{B}$ in $\widetilde{R}$. Hence if and only if the integrand in $\int_{\widetilde{B}}[f(x,T(x,T),Y'(x,T)]dx+\sum\limits_{i,k=1}^{n} f_{y'_i}\frac{\partial y_i}{\partial t_k}dt_k$ satisfies the exactness

$$\text{condition: } \begin{cases} \text{(E1):} \quad \frac{\partial}{\partial t_k}f(x,Y(x,T),Y'(x,T)) = \frac{\partial}{\partial x}\Big(\sum\limits_{i=1}^{n} f_{y'_i}(x,Y(x,T),Y'(x,T))\frac{\partial y_i}{\partial t_k}\Big) \\ \qquad\qquad k=1,\cdots,n \\ \text{(E2):} \quad \frac{\partial}{\partial t_r}\Big(\sum\limits_{i=1}^{n} f_{y'_i}(x,Y(x,T),Y'(x,T))\frac{\partial y_i}{\partial t_s}\Big) = \frac{\partial}{\partial t_s}\Big(\sum\limits_{i=1}^{n} f_{y'_i}(x,Y(x,T),Y'(x,T))\frac{\partial y_i}{\partial t_k}\Big) \\ \qquad\qquad r,s=1,\cdots,n; \quad r<s \end{cases}$$

(E1) is equivalent to the Euler-Lagrange equations (check this), hence yield no new conditions since $Y(x,T)$ are assumed to be extremals of $J[Y]$. However (E2) gives new conditions. Setting $v_i \underset{def}{=} f_{y'_i}(x,Y(x,T),Y'(x,T))$ we have the Hilbert integral is invariant if and only if

$$[t_x,t_r] \underset{def}{=} \sum\limits_{i=1}^{n}\Big(\frac{\partial y_i}{\partial t_s}\frac{\partial v_i}{\partial t_r} - \frac{\partial y_i}{\partial t_r}\frac{\partial v_i}{\partial t_s}\Big) = 0 \quad (r,s=0,\cdots,n).$$

Hence the family of extremals $Y(x,T)$ is a Mayer family if and only if all Lagrange brackets, $[t_r,t_k]$ (which are functions of (x,T)), vanish.

29 The Weierstrass and Legendre Conditions for $n \geq 2$ Sufficient Conditions.

These topics are covered in some detail for $n = 1$ in sections 11, 17and 18. Here we outline some of the modifications for $n > 1$.

1. a. The Weierstrass function for the functional $J[Y] = \int_{x_1}^{x_2} f(x, Y)dx$ is the following function of $3n + 1$ variables:

Definition 29.1.

$$\mathcal{E}(x, Y, Y', Z') = f(x, Y, Z') - f(x, Y, Y') - \sum_{i=1}^{n} (z'_i - y'_i) f_{y'_i}(x, Y, Y'). \tag{50}$$

By Taylor's theorem with remainder there exist θ such that $0 < \theta < 1$ and:

$$\mathcal{E}(x, Y, Y', Z') = \frac{1}{2}\Big[\sum_{i,k=1}^{n} (z'_i - y'_i)(z'_k - y'_k) f_{y'_i y'_k}(x, Y, Y' + \theta(Z' - Y'))\Big]. \tag{51}$$

However, to justify the use of Taylor's theorem and the validity of (51) we must assume either that Z' is sufficiently close to Y' (recall that the region R of admissible (x, Y, Y') is open in $\mathbb{R}^{2n+1}$), or that the region, R is convex in Y'.

b.

Theorem 29.2 (Weierstrass' Necessary Condition)**.** *If $Y_0(x)$ gives a strong (respectively weak) relative minimum for $\int_{x_1}^{x_2} f(x, Y, Y')dx$, then $\mathcal{E}(x, Y, Y', Z') \geq 0$ for all $x \in [x_1, x_2]$ and all Z' (respectively all Z' sufficiently close to $Y_0'(x)$).*

Proof : See the proof for $n = 1$ in §11 or [AK] §17.

Corollary 29.3 (Legendre's Necessary Condition). *If $Y_0(x)$ gives a weak relative minimum for $\int_{x_1}^{x_2} f(x, Y, Y')dx$, then $F_x(\overrightarrow{\rho}) = \sum_{i,k=1}^{n} f_{y_i' y_k'}(x, Y_0(x), Y_0'(x))\rho_i\rho_k \geq 0$, for all $x \in [x_1, x_2]$ and for all $\overrightarrow{\rho} = (\rho_1, \cdots, \rho_n)$.*

Proof : Assume we had $F_x(\overrightarrow{\rho}) < 0$ for some $\overline{x} \in [x_1, x_2]$ and for some $\overrightarrow{\rho}$. Then $F_x(\lambda\overrightarrow{\rho}) < 0$ for all $\lambda > 0$. set $Z_\lambda' = Y_0'(\overline{x}) + \lambda\overrightarrow{\rho}$; then by (51)

$$\mathcal{E}(\overline{x}, Y_0(\overline{x}), Y_0'(\overline{x}), Z_\lambda') = \frac{1}{2}\lambda^2[\sum_{i,k=1}^{n}(\rho_i\rho_k f_{y_i' y_k'}(\overline{x}, Y_0(\overline{x}), Y_0'(\overline{x} + \theta\lambda\overrightarrow{\rho})].$$

Since $\lambda^2 \sum_{i,k=1}^{n} \rho_i\rho_k f_{y_i' y_k'}(\overline{x}, Y_0(\overline{x}), Y_0'(\overline{x})) = F_{\overline{x}}(\lambda\overrightarrow{\rho}) < 0$ for all $\lambda > 0$ and since the $f_{y_i' y_k'}$ are continuous, it follows that for all sufficiently small $\lambda > 0$, we have $\mathcal{E}(\overline{x}, Y_0(\overline{x}), Y_0'(\overline{x}), Z_\lambda') < 0$. Contradicting Weierstrass' necessary condition for a weak minimum. **q.e.d.**

Remark 29.4. Legendre's condition says that the matrix $\left[f_{y_i' y_k'}(x, Y_0(x), Y_0'(x))\right]$ is the matrix of a positive-semidefinite quadratic form. This means that the n, real characteristic roots of the real, symmetric matrix $\left[f_{y_i' y_k'}(x, Y_0(x), Y_0'(x))\right]$, $\lambda_1, \cdots, \lambda_n$ are all ≥ 0. Another necessary and sufficient condition for the matrix to be positive-semidefinite is that all n determinants of the upper right $m \times m$ submatrices are ≥ 0, $m = 1, \cdots, n$.

Example 29.5. (a) Legendre's condition is satisfied for the arc length functional, $J[y_1, y_2] = \int_{x_1}^{x_2} \sqrt{1 + (y_1')^2 + (y_2')^2}dx$ and $\mathcal{E}(x, Y, Y', Z') > 0$ for all $Z' \neq Y'$.

(b) Weierstrass' necessary condition is satisfied for all functionals of the form

$$\int_{x_1}^{x_2} n(x, Y)\sqrt{1 + (y_1')^2 + \cdots + (y_n')^2}dx$$

The proofs are the same as for the case $n = 1$.

2. Sufficient Conditions.

(a) For an extremal $Y_0(x)$ of $J[Y] = \int_{x_1}^{x_2} f(x, Y, Y')dx$, say that condition F is satisfied if and only if $Y_0(x)$ can be imbedded in a field for $J[Y]$ (cf. §28#1). such that $Y = Y_0(x)$ is one of the field trajectories.

(b)

Theorem 29.6 (Weierstrass). *Assume that the extremal* $Y_0(x)$ *of* $J[y] = \int_{(x_1,Y_1)}^{(x_2,Y_2)} f(x, Y, Y')dx$ *satisfies the imbeddability condition, F. Then for any* $\widetilde{Y}(x)$ *of a sufficiently small strong neighborhood of* $Y_0(x)$*, satisfying the same endpoint condition:*

$$J[\widetilde{Y}(x)] - J[Y_0(x)] = \int_{(x_1,Y_1)}^{(x_2,Y_2)} \mathcal{E}(x, \widetilde{Y}(x), P(x, \widetilde{Y}(x)), \widetilde{Y}'(x))dx$$

where $P(x, Y)$ *is the slope function of the field.*

Proof : Almost the same as for $n = 1$.

(c) **Fundamental Sufficiency Lemma.** Assume that (i) the extremal, $E : Y = Y_0(x)$ of $J[Y]$ satisfies the imbeddability condition, F, and that (ii) $\mathcal{E}(x, Y, P(x, Y), Z') \geq 0$ holds for all (x, Y) near the pairs $(x, Y_0(x))$ of E and for all Z' (or all Z' sufficiently near $Y_0'(x)$. Then: $Y_0(x)$ gives a strong (or weak respectively) relative minimum for $J[Y]$. **Proof :** Follows from (b) as for the case $n = 1$.

(d) Another sufficiency theorem.

Theorem 29.7. *If the extremal, $Y_0(x)$ for $J[Y]$ satisfies the imbeddability condition, F, and condition*

$$L'_n : \sum_{i,k=1}^{n} \rho_i \rho_k f_{y'_i y'_k}(x, Y_0(x), Y'_0(x)) > 0$$

for all $\overrightarrow{\rho} \neq \overrightarrow{0}$ and all $x \in [x_1, x_2]$, then Y_0 gives a weak relative minimum for $J[Y]$.

Proof : *Use a continuity argument and (51) as for the case $n = 1$. See also [AK][pages 60-62].*

Further comments:

a. A functional $J[Y] = \int_{x_1}^{x_2} f(x, Y, Y')dx$ is called *regular* (*quasi-regular* respectively) if and only if the region R is convex in Y' and $\sum_{i,k=1}^{n} \rho_i \rho_k f_{y'_i y'_k}(x, Y_0(x), Y'_0(x)) > 0$ (or ≥ 0) holds for all $\overrightarrow{\rho} \neq \overrightarrow{0}$ and for all (x, Y, Y') of R.

b. The Jacobi condition, $\mathcal{J}$ is quite similar for $n > 1$ to the case $n = 1$. Conditions sufficient to yield the imbeddability condition are implied by $\mathcal{J}$ and therefore further sufficiency conditions can be obtained much as in §16 and §18.

30 The Euler-Lagrange Equations in Canonical Form.

Suppose we are given $f(x,y,z)$ with $f_{zz} \neq 0$ in some (x,y,z)-region, R. Map R onto an (x,y,v)-region, $\overline{R}$ by

$$\begin{cases} x = x \\ y = y \\ v = f_z(x,y,z) \end{cases}$$

Note that the Jacobian, $\frac{\partial(x,y,v)}{\partial(x,y,z)} = f_{zz} \neq 0$. So the mapping is $1-1$ in a region and we may solve for z:

$$\begin{cases} x = x \\ y = y \\ z = p(x,y,v) \end{cases}$$

Note that f_{zz} only gives a local inverse. Introduce a function H in $\overline{R}$

$$H(u,y,v) \underset{def}{\equiv} p(x,y,v) \cdot v - f(x,y,p(x,y,v))$$

equivalently

$$H = zf_z - f$$

H is called the *Hamiltonian* corresponding to the functional $J[y]$. Then

$$H_x = p_x v - f_x - f_z p_x, \quad H_y = p_y v - f_y - f_z p_y, \quad H_v = p_v v + p - f_z p_v$$

i.e.

$$(H_x, H_y, H_v) = (-f_x, -f_y, p) = (-f_x, -f_y, z)$$

The case $n > 1$ is similar. Given $f(x, Y, Z) \in C^3$ in some $(x, Y, Z)-$region R, assume $\left|f_{y_i' y_k'}(x, Y, Z)\right| \neq 0$ in R. Map R onto an $(x, Y, V)-$region $\overline{R}$ by means of :

$$\begin{cases} x = x \\ y = y \\ v_k = f_{z_k}(x, Y, Z) \end{cases} \leftrightarrow \begin{cases} x = x \\ y = y \\ z_k = p_k(x, Y, V) \end{cases}$$

Where we assume the map from R to $\overline{R}$ is one to one. (Note that the Jacobian, $\left|f_{y_i' y_k'}(x, Y, Z)\right| \neq 0$ implies map from R to $\overline{R}$ is one to one locally only.)

We now define a function, H by

$$\text{(52)} \quad H(x, Y, V) = \sum_{k=1}^{n} p_k(x, Y, V) \cdot v_k - f(x, Y, P(x, Y, V)) = \sum_{k=1}^{n} z_k f_{z_k}(x, Y, Z) - f(x, Y, Z)$$

The calculation of the first partial derivatives is similar to the case of $n = 1$:

$$\text{(53)} \qquad (H_x; H_{y_1}, \cdots, H_{y_n}; H_{v_1}, \cdots, H_{v_n}) = (-f_x; -f_{y_1}, \cdots, -f_{y_n}; p_1, \cdots, p_n)$$

$$= (-f_x; -f_{y_1}, \cdots, -f_{y_n}; z_1, \cdots, z_n).$$

Example 30.1. If $f(x, Y, Z) = \sqrt{1 + z_1^2 + \cdots z_n^2}$, then $H(x, Y, V) = \pm\sqrt{1 - v_1^2 + \cdots - v_n^2}$.

The transformation from (x, Y, Z) and f to (x, Y, V) and H is <u>involutory</u> i.e. a second application leads from (x, Y, V) and H to (x, Y, Z) and f.

We now transform the Euler-Lagrange equation into canonical variables (x, Y, V) coordinates, also replacing the "Lagrangian", f in terms of the Hamiltonian, H. $\frac{df_{z_i}}{dx} - f_{y_i} = 0$ with $z_i = y_i'(x)$ becomes:

$$\frac{dy_i}{dx} = H_{v_i}(x, Y, V), \quad \frac{dv_i}{dx} = -H_{y_i}(x, Y, V), \quad i = 1, \cdots, n. \tag{54}$$

These are the Euler-Lagrange equations for $J[Y]$ in canonical form. For another aspect of (54) see [GF][§18.2]

Recall (§54, #1c) that an n-parameter family of extremals $Y(x, t_1, \cdots, t_n)$ for $J[Y] = \int_{x_1}^{x_2} f(x, Y, Y')dx$ is a Mayer family if and only if it satisfies

$$0 = [t_s, t_r] = \sum_{i=1}^{n} \Big(\frac{\partial y_i}{\partial t_s} \frac{\partial v_i}{\partial t_r} - \frac{\partial y_i}{\partial t_r} \frac{\partial v_i}{\partial t_s} \Big) = 0 \quad (r, s = 0, \cdots, n)$$

where $v_i = f_{y_i'}(x, Y(x, T), Y'(x, T)), \quad i = 1, \cdots, n.$

Now in the canonical variables, (x, Y, V), the extremals are characterized by (54). Hence the functions $y_i(x, T), v_i(x, T)$ occurring in $[t_s, t_r]$ satisfy the Hamilton equations:

$$\frac{dy_i(x, T)}{dx} = H_{v_i}(x, Y(x, T), V(x, T)), \frac{dv_i(x, T)}{dx} = -H_{y_i}(x, Y(x, T), V(x, T)).$$

Using this we can now prove:

Theorem 30.2 (Lagrange)**.** *For a family $Y(x, T)$ of extremals, the Lagrange brackets $[t_s, t_r]$ (which are functions of (x, T)) are independent of x along each of these extremals. In particular they are constant alone each of the extremals.*

Proof : By straightforward differentiation, using the Hamilton equations and the identity

$$\frac{\partial G(x, Y(x, T), V(x, T))}{\partial t_p} = \sum_{i=1}^{n} (G_{y_i} \frac{\partial y_i}{\partial t_p} + G_{v_i} \frac{\partial v_i}{\partial t_p})$$

applied to $p = r$ and $G = \frac{\partial H}{\partial t_s}$ as well as to $p = s, G = \frac{\partial H}{\partial t_r}$ implies:

$$\frac{d}{dx}[t_s, t_r] = \sum_{i=1}^{n}[\frac{\partial}{\partial t_s}(\overbrace{\frac{dy_i}{dx}}^{=H_{v_i}}) \cdot \frac{\partial v_i}{\partial t_r} + \frac{\partial y_i}{\partial t_s}\frac{\partial}{\partial t_r}(\overbrace{\frac{dv_i}{dx}}^{=-H_{y_i}}) - \frac{\partial}{\partial t_r}(\overbrace{\frac{dy_i}{dx}}^{=H_{v_i}}) \cdot \frac{\partial v_i}{\partial t_s} - \frac{\partial y_i}{\partial t_r}\frac{\partial}{\partial t_s}(\overbrace{\frac{dv_i}{dx}}^{=-H_{y_i}})]$$

$$= \sum_{i=1}^{n}[(\frac{\partial H}{\partial t_s})_{y_i}\frac{\partial y_i}{\partial t_r} + (\frac{\partial H}{\partial t_s})_{v_i}\frac{\partial v_i}{\partial t_r}] - \sum_{i=1}^{n}[(\frac{\partial H}{\partial t_r})_{y_i}\frac{\partial y_i}{\partial t_s} + (\frac{\partial H}{\partial t_r})_{v_i}\frac{\partial v_i}{\partial t_s}]$$

where, e.g. $(\frac{\partial H}{\partial t_s})_{v_i}$ denotes the partial derivative with respect to v_i. But this is $\frac{\partial^2 H}{\partial t_r \partial t_s} - \frac{\partial^2 H}{\partial t_s \partial t_r}$ which is zero. **q.e.d.**

As an application to check the condition for a family of extremals to be a Mayer family, i.e. $[t_s, t_r] = 0$ for all r, s it is sufficient to establish $[t_s, t_r]$ at a single point, $(x_0, Y(x_0, T))$ of each extremal.

For instance, if all $Y(x, T)$ pass through one and the same point, (x_0, Y_0) of $\mathbb{R}^{n+1}$, then $Y(x, T) = Y_0$ for all T. Hence $\frac{\partial y_i}{\partial t_s}\Big|_{(x_0,T)} = 0$ for all i, s. This implies $[t_s, t_r] = 0$ and we have a second proof that the family of all extremals through (x_0, Y_0) is a (central) field.

31 Hamilton-Jacobi Theory

31.1 Field Integrals and the Hamilton-Jacobi Equation.

There is a close connection between the minimum problem for a functional $J[Y] = \int_{x_1}^{x_2} f(x, Y, Y')dx$ on the one hand, and the problem of solving a certain 1−st order partial differential equation on the other. This connection is the subject of Hamilton-Jacobi theory.

(a) Given a simply-connected field for the functional $J[Y] = \int_{x_1}^{x_2} f(x, Y, Y')dx$ let $Q(x, Y)$ be its slope function[43]. Then in the simply connected field region D, the Hilbert integral:

$$I_B^* = \int_B \left\langle [f(x, Y, Q) - \sum_{i=1}^{n} q_i(x, Y) f_{y_i'}(x, Y, Q)]dx + \sum_{i=1}^{n} f_{y_i'}(x, Y, Q)dy_i \right\rangle$$

is independent of the path B. Hence, fixing a point $P^{(1)} = (x_1, Y_1)$ of D we may define a function $W(x, Y)$ in D by

Definition 31.1. The *field integral*, $W(x, Y)$ of a field is defined by

$$W(x, Y) = \int_{(x_1, Y_1)}^{(x,Y)} \left\langle [f(x, Y, Q) - \sum_{i=1}^{n} q_i(x, Y) f_{y_i'}(x, Y, Q)]dx + \sum_{i=1}^{n} f_{y_i'}(x, Y, Q)dy_i \right\rangle$$

where the integral may be evaluated along any path in D from (x_1, Y_1) to (x, Y). It is determined by the field to within an additive constant (the choice of $P^{(1)}$!)

[43]The slope function had been denoted by $P(x, Y)$. We change to $Q(x, Y)$ to avoid confusion with P that appears when on converts to canonical variables.

Since the integral in (31.1) is independent of path in D, its integrand must be an exact differential so that

$$dW(x,Y) = [f(x,Y,Q) - \sum_{i=1}^{n} q_i(x,Y) f_{y_i'}(x,Y,Q)]dx + \sum_{i=1}^{n} f_{y_i'}(x,Y,Q) dy_i$$

hence:

(55)

(a) $W_x(x,Y) = f(x,Y,Q) - \sum_{i=1}^{n} q_i(x,Y) f_{y_i'}(x,Y,Q)$

(b) $W_{y_i} = f_{y_i'}(x,Y,Q)$

We shall abbreviate the $n-$tuple $(W_{y_1}, \cdots, W_{y_n})$ by $\nabla_Y W$.

Now assume we are working in a neighborhood of the point (x_0, Y_0) where $\left| f_{y_i' y_k'}(x_0, Y_0, Q(x_0, Y_0)) \right| \neq 0$. e.g. small enough so that the system of equations $f_{y_i'}(x,Y,Z) = v_i, \quad i = 1, \cdots, n$ can be solved for $Z : z_k = p_k(x,Y,V)$. Then in this neighborhood (55)[(b)] yields

$$q_k(x,Y) = p_k(x,Y,\nabla_Y W), \quad k = 1, \cdots, n$$

Substitute this into (55)[(a)]:

$$W_x(x,Y) = f(x,Y,P(x,Y,\nabla_Y W) - \sum_{i=1}^{n} p_i(x,Y,\nabla_Y W) \overbrace{f_{y_i'}(x,Y,\nabla_Y W)}^{=W_{y_i}}$$

Recalling the definition of the Hamiltonian, H, this becomes

Theorem 31.2. *Every field integral, $W(x, Y)$ satisfies the first order, Hamilton-Jacobi partial differential equation:*

$$W_x(x, Y) + H(x, Y, \nabla_Y W) = 0$$

(b) If $W(x, Y)$ is a specific field integral of a field for $J[Y]$ then in addition to satisfying the Hamilton-Jacobi differential equation, W determines a $1-$parameter family of transversal hypersurfaces. Specifically we have:

Theorem 31.3. *If $W(x, Y)$ is a specific field integral of a field for $J[Y]$, then each member of the $1-$ parameter family of hypersurfaces, $W(x, Y) = c$ (c a constant) is transversal to the field.*

Proof : $W(x, Y) = c$ implies

$0 = dW(x, Y) = [f(x, Y, Q) - \sum_{i=1}^{n} q_i(x, Y) f_{y_i'}(x, Y, Q)]dx + \sum_{i=1}^{n} f_{y_i'}(x, Y, Q) dy_i$ for any $(dx, dy_1, \cdots, dy_n)$ tangent to the hypersurface $W = c$. But this is just the transversality condition, (28.1).

(c) Geodesic Distance. Given a field for $J[Y]$, and two transversal hypersurfaces S_i : $W(x, Y) = c_i, \quad i = 1, 2$. For an extremal of the field joining point P_1 of S_1 to point P_2 of S_2 we have $J[E_{\widehat{P_1P_2}}] = \int_{x_1}^{x_2} f(x, Y, Y')dx = I^*_{E_{\widehat{P_1P_2}}} = \int_{x_1}^{x_2} dW = W(x_2, Y_2) - W(x_1, Y_1) = c_2 - c_1$. The extremals of the field are also called the "geodesics" from S_1 to S_2, and $c_2 - c_1 = J[E_{\widehat{P_1P_2}}]$ the geodesic distance from S_1 to S_2.

In particular by the geodesic distance from a point P_1 to a point P_2 with respect to $J[Y]$, we mean the number $J[E_{\widehat{P_1P_2}}]$. Note that $E_{\widehat{P_1P_2}}$ can be thought of as imbedded in the central field of extremals through P_1 at least for P_2 in a suitable neighborhood of P_1.

(d) We saw that every field integral, $W(x, Y)$ for $J[Y]$ satisfies the Hamilton-Jacobi equation: $W_x + H(x, Y, \nabla_Y W) = 0$ where H is the Hamiltonian function for $J[Y]$. We next show that, conversely, every C^2 solution$U(x, Y)$ of the Hamilton-Jacobi equation is the field integral for some field of $J[Y]$.

Theorem 31.4. *Let $U(x, Y)$ be C^2 in an (x, Y)-region D and assume that U satisfies $U(x, Y) + H(x, Y, \nabla_Y U) = 0$. Assume also that the (x, Y, V)-region,*
$\overline{R} = \{(x, Y, \nabla_Y U)\Big|(x, Y) \in D\}$ is the one to one image under the canonical mapping (see §30), of an (x, Y, Z)-region R consisting of non-singular elements for $J[Y]$ (i.e. such that $\Big|f_{y'_i y'_k}(x, Y, Z)\Big| \neq 0$ in R). Then: there exist a field for $J[Y]$ with field region D and field integral $U(x, Y)$.

Proof : We must exhibit a slope function, $Q(x, Y)$ in D such that the Hilbert integral, I^*_B formed with this Q is independent of the path in D and is in fact equal to $\int_B dU$.

Toward constructing such a $Q(x, Y)$: Recall from **(a)** above that if $W(x, Y)$ is a field integral, then the slope function Q can be expressed in terms of W and in terms of the canonical mapping functions $p_k(x, Y, V)$ by means of: $q_k(x, Y) = p_k(x, Y, \nabla_Y W)$ (see

(55) and the paragraph following). Accordingly, we try: $q_k(x,Y) \underset{def}{=} p_k(x,Y,\nabla_y U)$, $\quad k = 1,\cdots,n$ as our slope function. Solving this for $\nabla_Y U$ (cf. §30)) gives

$$U_{y_i} = f_{y_i'}(x,Y,Q(x,Y)), \quad i = 1,\cdots,n$$

Further, by hypothesis, and recalling the definition of the Hamiltonian, H,

$$U_x = -H(x,Y,\nabla_Y U) = f(x,Y,\overbrace{P(x,Y,\nabla_Y U)}^{=Q(x,Y)}) - \sum_{k=1}^{n} \underbrace{p_k(x,Y,\nabla_Y U)}_{=q_k(x,Y)} \cdot \underbrace{U_{y_k}}_{=f_{y_k'}(x,Y,Q(x,Y))}$$

$$= f(x,Y,Q(x,Y)) - \sum_{k=1}^{n} q_k(x,Y) f_{y_k'}(x,Y,Q(x,Y)).$$

Hence

$$dU(x,Y) = U_x dx + \sum_{i=1}^{n} U_{y_i} dy_i =$$

$$[f(x,Y,Q(x,Y)) - \sum_{k=1}^{n} q_k(x,Y) f_{y_k'}(x,Y,Q(x,Y))]dx + \sum_{i=1}^{n} f_{y_i'}(x,Y,Q(x,Y))dy_i$$

and $\int_B dU = I_B^*$ where I_B^* is formed with the slope function, Q defined above. **q.e.d.**

(e) Examples (for $n = 1$): Let $J[Y] = \int_{x_1}^{x_2} \sqrt{1+(y')^2}$: $H(x,y,v) = \pm\sqrt{x - v^2}$. The extremals are straight lines. Any field integral, $W(x,y)$ satisfies the Hamilton-Jacobi equation: $W_x \pm \sqrt{1 - W_y^2} = 0$ i.e.

$$W_x^2 + W_y^2 = 1$$

Note that there is a 1− parameter family of solutions:$W(x,y;\alpha) = -x\sin\alpha + y\cos\alpha$. From (31.4) we know that for each α there is a corresponding field with W as field integral. From (31.3) we know that $W =$ constant =straight lines with slope $\tan\alpha$ are transversal to the field. Hence the field corresponding to this field integral must be a family of parallel lines with slope $-\cot\alpha$. In particular the Hilbert integral (associated to this field) from any point to the line $-x\sin\alpha + y\cos\alpha = c$ is constant.

In the other direction, suppose we are given the field of lines passing though a point (a, b). We may base the Hilbert integral at the point $P^{(1)} = (a, b)$. Since the Hilbert integral is independent of the path we may define the field integral by evaluating the Hilbert integral along the straight line to (x, y). But along an extremal the Hilbert integral is $J[y]$ which is the distance function. In other words $W(x, y) = \sqrt{(x-a)^2 + (y-b)^2}$. Notice that this W satisfies the Hamilton-Jacobi equation and the level curves are transversal to the field. Furthermore if we choose $P^{(1)}$ to be a point other than (a, b) (so the paths defining the field integral cut across the field) the field integral is constant on circles centered at (a, b).

31.2 Characteristic Curves and First Integrals

In order to exploit the Hamilton-Jacobi differential equation we have to digress to explain some techniques from partial differential equations. The reference here is [CH][Vol. 2, Chapter II].

1.(a) For a first order partial differential equation in the unknown function $u(x_1, \cdots, x_2) = u(X)$:

$$G(x_1, \cdots, x_m, u, p_1, \cdots, p_m) \quad p_i \underset{def}{=} \frac{\partial u}{\partial x_i}, \tag{56}$$

the following system of simultaneous 1−st order ordinary differential equations (in terms of an arbitrary parameter, t, which may be x_1) plays an important role and is called the system of characteristic differential equations for (56).

$$\text{(57)} \qquad \frac{dx_i}{dt} = G_{p_i}; \quad \frac{du}{dt} = \sum_{i=1}^{m} p_i G_{p_i}; \quad \frac{dp_i}{dt} = -(G_{x_i} + p_i G_u) \quad (i = 1, \cdots, m)$$

Any solution, $x_i = x_i(t), u = u(t), p_i = p_i(t)$ of (57) is called a characteristic strip of (56). The corresponding curves, $x_i = x)i(t)$ of $(x_1, \cdots, x_m)$-space, $\mathbb{R}^m$ are the characteristic ground curves of (56).

Roughly speaking the m-dimensional solution surfaces $u = u(x_1, \cdots, x_m)$ of (56) in $\mathbb{R}^{m+1}$ is foliated by characteristic curves of (56).

(b) In the special case when G does not contain u (i.e. $G_u = 0$) the characteristic system minus the redundant equation for $\frac{du}{dt}$ becomes:

$$\frac{dx_i}{dt} = G_{p_i}; \quad \frac{dp_i}{dt} = -G_{x_i}.$$

(c) In particular, let us construct the characteristic system for the Hamilton-Jacobi equation: $W_x + H(x, Y, \nabla_Y W) = 0$. Note the dictionary (where for the parameter, t, we choose x):

$$\begin{array}{ccccccccccccccccc}
x_1, & x_2, & \cdots, & x_m; & u; & p_1, & p_2, & \cdots, & p_m; & G_{x_1}, & G_{x_2}, & \cdots, & G_{x_m}; & G_{p_1}, & G_{p_2}, & \cdots, & G_{p_m} \\
\updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\
x, & y_1, & \cdots, & y_n; & W; & W_x, & W_{y_1}, & \cdots, & W_{y_n}; & H_x, & H_{y_1}, & \cdots, & H_{y_n}; & 1, & H_{v_1}, & \cdots, & H_{v_n}
\end{array}$$

The remaining characteristic equations of (57) include

$$\frac{dy_i}{dx} = H_{v_i}(x, Y, V); \quad \frac{dv_i}{dx} = -H_{y_i}(x, Y, V)$$

where $V = \nabla_Y W$. So we have proven:

Theorem 31.5. *The extremals of $J[Y]$ in canonical coordinates, are the characteristic strips for the Hamilton-Jacobi equation.*

2 (a). Consider a system of m simultaneous $1-$order ordinary differential equations.

$$\frac{dx_i}{dt} = g_i(x_1, \cdots, x_m) \quad i = 1, \cdots, m \tag{58}$$

where the $g_i(x_1, \cdots, x_m)$ are sufficiently continuous or differentiable in some $X = (x_1, \cdots, x_m)-$ region D. The solutions $X = X(t) = (x_1(t), \cdots, x_m(t))$ of (58) may be interpreted as curves, parameterized by t in $(x_1, \cdots, x_m)$-space $\mathbb{R}^m$.

Definition 31.6. A <u>first integral</u> of the system (58) is any class C^1 function $u(x_1, \cdots, x_m)$ that is constant along every solution curve, $X = X(t)$, i.e. any $u(X)$ such that

$$0 = \frac{du(X(t))}{dt} = \sum_{i=1}^{m} u_{x_i} \frac{dx_i}{dt} = \sum_{i=1}^{m} u_{x_i} g_i(X) = \nabla_u \cdot \overrightarrow{G} \tag{59}$$

where $\overrightarrow{G} = (g_1, \cdots, g_m)$.

Note that the relation of $\nabla_u \cdot \overrightarrow{G} = 0$ to the system (58) is that of a 1-st order partial differential equation (in this case linear) to the first half of its system of characteristic equations discussed in #**1.** above.

(b) Since the $g_i(X)$ are free of t, we may eliminate t altogether from (58) and, assuming (say) $g_1 \neq 0$ in D, make x_1 the new independent variable with equations:

$$\frac{dx_2}{dx_1} = h_2(X), \cdots, \frac{dx_m}{dx_1} = h_m(X), \text{ where } h_i = \frac{g_i}{g_1} \tag{60}$$

For this system, equivalent to (58), we have by the Picard existence and uniqueness theorem:

Proposition 31.7. *Through every $(x_1^{(0)}, \cdots, x_1^{(0)})$ of D there passes one and only one solution curve of (58).*

Their totality is an $(m-1)-$parameter family of curves in D. For instance, keeping $x_1^{(0)}$fixed, we can write the solution of (60) that at $x_1 = x_1^{(0)}$ assumes assigned values $x_2 = x_2^{(0)}, \cdots, x_m = x_m^{(0)}$ as:

$$x_2 = \widetilde{x}_2(x_1; x_2^{(0)}, \cdots, x_m^{(0)}), \cdots, x_m = \widetilde{x}_m(x_1; x_2^{(0)}, \cdots, x_m^{(0)})$$

with $x_2^{(0)}, \cdots, x_m^{(0)}$ the $m-1$ parameters of the family.

An $m-1$-parameter family of solutions curves of the system (60) of $m-1$ differential equations is called a general solution of that system if the $m-1$ parameters are independent.

(c)

Theorem 31.8. *Let $u^{(1)}(X), \cdots, u^{(m-1)}(X)$ be $m-1$ solutions of (59), i.e. first integrals of (58), or (60) that are functionally independent in D, i.e. the $(m-1)\times m$ matrix*

$$\begin{pmatrix} u^{(1)}_{x_1} & \cdots & u^{(1)}_{x_m} \\ \vdots & & \vdots \\ u^{(m-1)}_{x_1} & \cdots & u^{(m-1)}_{x_m} \end{pmatrix}$$

is of maximum rank, $m-1$ throughout D. Then:

(i) Any other solution, $u(X)$ of (59) (i.e., any other first integral of (58)) is of the form $u(X) = \varphi(u^{(1)}(X), \cdots, u^{(m-1)}(X))$ in D;

(ii) The equations $u^{(1)}(X) = c_1, \cdots, u^{(m-1)}(X) = c_{m-1}$ represent an $m-1$-parameter family of solution curves, i.e. general solutions of (58), or (60).

Remark: Part (ii) says that the solutions of (60) may be represented as the intersection of $m-1$ hypersurfaces in $\mathbb{R}^m$.

Proof : (i) By hypothesis, the $m-1$ vectors $\nabla u^{(1)}, \cdots, \nabla u^{(m-1)}$ of $\mathbb{R}^m$ are linearly independent and satisfy $\nabla u^{(i)} \cdot \overrightarrow{G} = 0$; thus they span the $m-1$-dimensional subspace of $\mathbb{R}^m$ that is perpendicular to $\overrightarrow{G}$. Now if $u(X)$ satisfies $\nabla u \cdot \overrightarrow{G} = 0$, then ∇u lies in that same subspace. Hence ∇u is a linear combination of $\nabla u^{(1)}, \cdots, \nabla u^{(m-1)}$. Hence

$$\begin{vmatrix} u_{x_1} & \cdots & u_{x_m} \\ u^{(1)}_{x_1} & \cdots & u^{(1)}_{x_m} \\ \vdots & & \vdots \\ u^{(m-1)}_{x_1} & \cdots & u^{(m-1)}_{x_m} \end{vmatrix} = 0$$

in D. This together with the rank $m-1$ part of the hypothesis, implies a functional dependence of the form $u(X) = \varphi(u^{(1)}(X), \cdots, u^{(m-1)}(X))$ in D.

(ii) If $\mathcal{K} : X = X(t) = (x_1(t), \cdots, x_m(t))$ is the curve of intersection of the hypersurfaces $u^{(1)}(X) = c_1, \cdots, u^{(m-1)}(X) = c_{m-1}$, then $X(t)$ satisfies $\frac{du^{(i)}(X)}{dt} = 0 (i = 1, \cdots, m-1)$, i.e. $\nabla u^{(i)} \cdot \frac{dX}{dt} = 0$. Hence both $\overrightarrow{G}$ and $\frac{dX}{dt}$ are vectors in $\mathbb{R}^m$ perpendicular to the $(m-1)-$ dimensional subspace spanned by $\nabla u^{(1)}, \cdots, \nabla u^{(m-1)}$. Therefore $\frac{dX}{dt} = \lambda(t)\overrightarrow{G}(X)$, which agrees with (58) to within a re-parametrization of $\mathcal{K}$; hence $\mathcal{K}$ is a solution curve of (58).

q.e.d.

Corollary 31.9. *If for each $C = (c_1, \cdots, c_{m-1})$ of an $(m-1)$-dimensional parameter region, the $(m-1)$ functions $u^{(1)}(X, C), \cdots, u^{(m-1)}(X, C)$ satisfy the hypothesis of the theorem, the* *conclusion (ii) may be replaced by:*

(ii′) The equations $u^{(1)}(X,C) = 0, \cdots, u^{(m-1)}(X,C) = 0$ *represent an* $(m-1)$*-parameter family of solution curves, i.e. a general solution of (58), or (60).*

31.3 A theorem of Jacobi.

Now for the main theorem of Hamilton-Jacobi theory.

Theorem 31.10 (Jacobi)**.** *(a) Let* $W(x, Y, a_1, \cdots, a_r)$ *be an* r*-parameter family* $(r \le n)$ *of solutions of the Hamilton-Jacobi differential equation:* $W_x + H(x, Y, \nabla_Y W) = 0$*. Then: each of the* r*-th first partials,* $W_{a_j} \quad j = 1, \cdots, r$ *is a first integral of the canonical system:*$\frac{dy_i}{dx} = H_{v_i}(x, Y, V)$, $\frac{dv_i}{dx} = -H_{y_i}(x, Y, V)$ *of the Euler-Lagrange equations.*

(b) Let $A = (a_1, \cdots, a_n)$ *and* $W(x, Y, A)$ *be an* n*-parameter family of solutions of the Hamilton-Jacobi differential equation for* $(x, Y) \in D$*, and for* A *ranging over some* n*-dimensional region,* $\mathcal{A}$*. Assume also that in* $D \times \mathcal{A}$*,* $W(x, Y, A)$ *is* C^2 *and satisfies* $|W_{y_i a_k}| \neq 0$.

Then:

$$y_i = y_i(x, A, B) \text{ as given implicitly by } W_{a_i}(x, Y, A) = b_i$$

and

$$v_i = W_{y_i}(x, Y, A)$$

represents a general solution (i.e. $2n$*-parameter family of solution curves) of the canonical system of the Euler-Lagrange equation. The* $2n$ *parameters are* $\{a_i, b_j\} \quad i, j = 1, \cdots, n$.

Proof : (a) We must show that $\frac{d}{dx}W_{a_j} = 0$ along every extremal, i.e. along every solution curve of the canonical system. Now along an extremal we have:

$$\frac{d}{dx}W_{a_j}(x,Y,A) = W_{a_j x} + \sum_{k=1}^{n} W_{a_j y_k}\frac{dy_k}{dx} = W_{a_j x} + \sum_{k=1}^{n} W_{a_j y_k}\cdot H_{v_k}.$$

But by hypothesis, W satisfies $W_x + H(x,Y,\nabla_Y W) = 0$ which, by taking partials with respect to a_j, implies $W_{a_j x} + \sum_{k=1}^{n} W_{a_j y_k}\cdot H_{v_k} = 0$ proving part (a).

(b) In accordance with theorem (31.8) and corollary (31.9), it suffices to show (since the $2n$ canonical equations in §30 play the role of system (60))that

(1) for any fixed choice of the $2n$ parameters $a_1,\cdots,a_n,b_1,\cdots,b_n$ the $2n$ functions

$$W_{a_i}(x,Y,A) - b_i, \quad W_{y_i}(x,Y,A) - v_i \quad (i = 1,\cdots,n)$$

are first integrals of the canonical system

and

(2) that the relevant functional matrix has maximum rank, $2n$ (the matrix is a $2n\times(2n+1)$ matrix).

Proof of (1). The first n functions, $W_{a_i}(x,Y,A) - b_i$ are 1-st integrals of the canonical system by (a) above. As to the last n:

$$\frac{d}{dx}[W_{y_i}(x,Y,A) - v_i]\Big|_{\text{along extremal}} = W_{y_i x} + \sum_{k=1}^{n} W_{y_i y_k}\frac{dy_k}{dx} - \frac{dv_i}{dx}$$

$$= W_{y_i x} + \sum_{k=1}^{n} W_{y_i y_k}H_{v_k} - H_{y_i} = \frac{\partial}{\partial y_i}[W_x + H(x,Y,\nabla_Y W)] = 0.$$

To prove (2): The relevant functional matrix, of the $2n$ functions $W_{a_i},\cdots,W_{a_n}$, $W_{y_1} - v_1,\cdots,W_{y_n} - v_n$ with respect to the $2n+1$ variables, $x,y_1,\cdots,y_n,v_1,\cdots,v_n$ is the

$2n \times (2n+1)$ matrix

$$\begin{pmatrix} W_{a_1x} & W_{a_1y_1} & \cdots & W_{a_1y_n} & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \\ W_{a_nx} & W_{a_ny_1} & \cdots & W_{a_ny_n} & 0 & \cdots & 0 \\ W_{y_1x} & W_{y_1y_1} & \cdots & W_{y_1y_n} & -1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & 0 & \ddots & 0 \\ W_{y_nx} & W_{y_ny_1} & \cdots & W_{y_ny_n} & 0 & 0 & -1 \end{pmatrix}$$

The $2n \times 2n$ submatrix obtained by ignoring the first column is non singular since its determinant equals $\pm$ the determinant of the $n \times n$ submatrix $[W_{a_iy_k}]$ which is not zero by hypothesis. **q.e.d.**

31.4 The Poisson Bracket.

(a) A function, $\varphi(x, Y, V)$ is a first integral of the canonical system: $\frac{dy_i}{dx} = H_{v_i}(x, Y, V)$, $\frac{dv_i}{dx} = -H_{y_i}(x, Y, V)$ if and only if

$$0 = \frac{d\varphi}{dx}\Big|_{\text{along extremals}} = \varphi_x + \sum_{i=1}^{n}(\varphi_{y_i}\frac{dy_i}{dx} + \varphi_{v_i}\frac{dv_i}{dx})\Big|_{\text{along extremals}} = \varphi_x + \sum_{i=1}^{n}(\varphi_{y_i}H_{v_i} - \varphi_{v_i}H_{y_i})$$

In particular if $\varphi_x = 0$, then $\varphi(Y, V)$ is a first integral of the canonical system if and only if $[\varphi, H] = 0$ where $[\varphi, H] = 0$ is the <u>Poisson bracket</u> of $\varphi, H : [\varphi, H] \underset{def}{=} \sum_{i=1}^{n}(\varphi_{y_i}H_{v_i} - \varphi_{v_i}H_{y_i})$.

(b) Now consider a functional of the form $J[Y] = \int_{x_1}^{x_2} f(Y, Y')dx$. Note that $f_x \equiv 0$. Then the Hamiltonian $H(x, Y, V)$ for $J[Y]$ likewise satisfies $H_x \equiv 0$ (see (53)). Now it is clear that

$[H, H] = 0$. Hence (a) implies that, in this case, $H(Y, V)$ is a first integral of the canonical Euler-Lagrange equations for the extremal $J[Y] = \int_{x_1}^{x_2} f(Y, Y')dx$.

31.5 Examples of the use of Theorem (31.10)

Part (b) of the Jacobi's theorem spells out an important application of the Hamilton-Jacobi equation. Namely an alternative method for obtaining the extremals of $J[Y]$: If we can find a suitable[44] n-parameter family of solutions, $W(x, Y, A)$ of the Hamilton-Jacobi equation. Then the $2n$-parameter family of extremals is represented by $W_{a_i}(x, Y, A) = b_i, W_{y_i}(x, Y, A) = v_i$.

(a) $J[Y] = \int_{x_1}^{x_2} \sqrt{1 + (y')^2}dx$. We have remarked in §31.1 **(e)** that $H(x, y, v) = \pm\sqrt{1 - v^2}$ in this case and the Hamilton-Jacobi equation is $W_x^2 + W_y^2 = 1$. In §31.2 we showed that the extremals of $J[Y]$ (i.e. straight lines) are the characteristic strips of the Hamilton-Jacobi equation. In the notation of §31.2 $G_x = G_y = 0$, and the last two of the characteristic equations, (57) become $\frac{dp_1}{dt} = 0, \frac{dp_2}{dt} = 0$. Therefore $p_1 = W_x = \alpha, p_2 = W_y = \sqrt{1 - \alpha^2}$. Thus $W(x, y, \alpha) = \alpha x + \sqrt{1 - \alpha^2}y = x \cos a + y \sin a$ is a 1-parameter solution family of the Hamilton-Jacobi equation. Hence by theorem (31.10) the extremals of $J[Y]$ are given by $W_x = b; \quad W_y = v$ i.e. by $-x \sin\alpha + y \cos\alpha = b; \quad \sin\alpha = v = f_{y'} = \frac{y'}{\sqrt{1+(y')^2}}$. The first equation gives all straight lines, i.e. <u>all</u>extremals of $J[Y]$. The second is equivalent to

[44]Note that the family $W(x, Y) + c$ for c a constant does not qualify as a suitable parameter. It would contribute a row of zeros in $[W_{y_i a_k}]$. Also recall that a field determines its field integral to within an additive constant only. So if $W(x, Y)$ satisfies the Hamilton-Jacobi equation, so does $W + c$.

$\tan\alpha = y'$, merely gives a geometric interpretation of the parameter α.

(b) $J[Y] = \int_{x_1}^{x_2} y\sqrt{1+(y')^2}dx$ (the minimal surface of revolution problem). Here

$$f(x,y,z) = y\sqrt{1+z^2}; \quad v = f_z = \frac{yz}{\sqrt{1+z^2}}.$$

Therefore $z = \frac{v}{\sqrt{y^2-v^2}}$ and $H(x,y,v) = zf_z - f = \frac{v^2}{\sqrt{y^2-v^2}} - \frac{y^2}{\sqrt{y^2-v^2}} = -\sqrt{y^2-v^2}$. Hence the Hamilton-Jacobi equation becomes $W_x - \sqrt{y^2 - W_y^2} = 0$ or $W_x^2 + W_y^2 = y^2$. In searching for a 1$-$parameter family of solutions it is reasonable to try $W(x,y) = \varphi(x) + \psi(y)$ which gives $(\varphi')^2(x) + (\psi')^2(x) = y^2$ or $(\varphi')^2(x) = y^2 - (\psi')^2(x) = \alpha^2$. Hence $\varphi(x) = \alpha x$, $\psi(y) = \int\sqrt{y^2-\alpha^2}dy$ and $W(x,y,\alpha) = \alpha x + \int\sqrt{y^2-\alpha^2}dy$. Therefore $W_\alpha(x,y,\alpha) = x - \int\frac{\alpha}{\sqrt{y^2-\alpha^2}}dy$. Hence by Jacobi's theorem the extremals are given by

$$W_\alpha = b = x = \alpha\cosh^{-1}\frac{y}{\alpha}; \quad W_y = \sqrt{y^2-\alpha^2} = v = \frac{yy'}{\sqrt{1+(y')^2}}.$$

The first equation is the equation of a catenary, the second interprets the parameter α.

We leave as an exercise to generalize this example to $J[y] = \int_{x_1}^{x_2} f(y)\sqrt{1+(y')^2}dx$.

(c) <u>Geodesics on a Louville surface.</u> These are surfaces $X(p,q) = (x(p,q), y(p,q), z(p,q))$ such that the first fundamental form $dx^2 = Edp^2 + 2Fdpdq + Gdq^2$ is equal to $dx^2 = [\varphi(p) + \psi(q)][dp^2 + dq^2]$. [For instance any surface of revolution $X(r,\theta) = (r\cos\theta, r\sin\theta, z(r))$. This follows since $ds^2 = d^2 + r^2d\theta^2 + (z')^2(r) = r^2[d\theta^2 + \frac{1+(z')^2(r)}{r^2}dr^2]$ which setting $\frac{1+(z')^2(r)}{r^2}dr^2 = d\rho^2$ gives $\rho = \int_{r_0}^{r}\frac{1+(z')^2(\bar r)}{\bar r^2}d\bar r^2$. We write r as a function of ρ, $r = h(\rho)$ and we have $dx^2 = h^2(\rho)[d\theta^2 + d\rho^2]$.]

The geodesics are the extremals of $J[q] = \int_{p_1}^{p_2} \overbrace{\sqrt{\varphi(p)+\psi(q)}\sqrt{1+(\frac{dq}{dp})^2}dp}^{=ds}$. We compute:

$$v = f_{q'} = \sqrt{\varphi+\psi}\frac{q'}{\sqrt{1+(q')^2}}; \quad H(p,q,v) = q'v - f = \frac{v^2 - (\varphi(p)+\psi(q))}{\sqrt{\varphi+\psi-v^2}}.$$

The Hamilton-Jacobi equation becomes: $W_p^2 + W_q^2 = \varphi(p) + \psi(q)$. As in example (b), we try $W(p,q) = \mu(p) + \omega(q)$, obtaining

$$\mu(p) = \int_{p_0}^{p} \sqrt{\varphi(\overline{p})+\alpha}d\overline{p}, \quad \omega(q) = \int_{q_0}^{q} \sqrt{\psi(\overline{q})-\alpha}d\overline{q}$$

and the extremals (i.e. geodesics) are given by:

$$W_\alpha(p,q,\alpha) = \frac{1}{2}[\int_{p_0}^{p} \frac{d\overline{p}}{\sqrt{\varphi(\overline{p})+\alpha}} - \int_{q_0}^{q} \frac{d\overline{q}}{\sqrt{\psi(\overline{q})+\alpha}}] = b.$$

32 Variational Principles of Mechanics.

(a) Given a system of n particles (mass points) in $\mathbb{R}^3$, consisting at time t of masses m_i at $(x_i(t), y_i(t), z_i(t))$, $\quad i = 1, \cdots, n$. Assume that he masses move on their respective trajectories, $(x_i(t), y_i(t), z_i(t))$, under the influence of a conservative force field; this means that there is a function, $U(t, x_1, y_1, z_1, \cdots, x_n, y_n, z_n)$ of $3n+1$ variables -called the potential energy function of the system such that the force, $\overrightarrow{F}_i$ acting at time t on particle m_i at $(x_i(t), y_i(t), z_i(t))$ is given by $\overrightarrow{F}_i = (-U_{x_i}, -U_{y_i}, -U_{z_i}) = -\nabla_x U$.

Recall that the kinetic energy, T, of the moving system, at time t, is given by

$$T = \frac{1}{2}\sum_{i=1}^{n} m_i|\overrightarrow{v}_i|^2 = \frac{1}{2}\sum_{i=1}^{n} m_i(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2)$$

where as usual $(\dot{\ }) = \frac{d}{dt}$.

By the Lagrangian L of the system is meant $L = T - U$. By the action, A of the system, from time t_1 to time t_2, is meant

$$A[C] = \int_{t_1}^{t_2} L dt = \int_C (T - U)dt.$$

Here C is the trajectory of $(m_1, \cdots, m_n)$ in $\mathbb{R}^{3n}$.

Theorem 32.1 (Hamilton's Principle of Least Action:)**.** *The trajectories* $C : (x_i(t), y_i(t), z_i(t)), i = 1, \cdots, n$ *of masses* m_i *moving in a conservative force filed with potential energy* $U(t, x_1, y_1, z_1, \cdots, x_n, y_n, z_n)$ *are the extremals of the action functional* $A[C]$*.*

Proof : By Newton's second law, the trajectories are given by

$$m_i \cdot (\ddot{x}_i, \ddot{y}_i, \ddot{z}_i) = -(U_{x_i}, U_{y_i}, U_{z_i}), \quad i = 1, \cdots, n$$

On the other hand, the Euler-Lagrange -equation for extremals of $A[C]$ are

$$0 = \frac{d}{dt} L_{\dot{x}_i} - L_{x_i} = \frac{d}{dt}(T_{\dot{x}_i} - \overbrace{U_{\dot{x}_i}}^{=0}) - (\overbrace{T_{x_i}}^{=0} - U_{x_i}) = m_i \ddot{x}_i + U_{x_i}$$

(with similar equations for y and z). The Lagrange equations of motion agree with the trajectories given by Newton's law. **q.e.d.**

Note: (a) that Legendre's necessary condition for a minimum is satisfied. i.e. the matrix $[f_{y'_i y'_j}]$ is positive definite.

(b) The principle of least action is only a necessary condition (Jacobi's necessary condition may fail). Hence the trajectories given by Newton's law may not minimize the action.

The <u>canonical variables</u>: $v_1^x, v_1^y, v_1^z, \cdots, v_n^x, v_n^y, v_n^z$ are, for the case of $A[C]$:

$$v_i^x = L_{\dot{x}_i} = m_i \dot{x}_i; \quad v_i^y = L_{\dot{y}_i} = m_i \dot{y}_i; \quad v_i^z = L_{\dot{z}_i} = m_i \dot{z}_i, \quad i = 1, \cdots, n.$$

Thus v_i^x, v_i^y, v_i^z are the components of the <u>momentum</u> $m_i \overrightarrow{v}_i = m_i(\dot{x}_i, \dot{y}_i, \dot{z}_i)$ of the i-th particle. The canonical variables, v_i are therefore often called <u>generalized momenta</u>, even for other functionals.

The Hamiltonian H for $A[C]$ is

$$H(t, x_1, y_1, z_1, \cdots, x_n, Y_n, z_n, v_1^x, v_1^y, v_1^z, \cdots, v_n^x, v_n^y, v_n^z) = \sum_{i=1}^{n} (\dot{x}_i v_i^x + \dot{y}_i v_i^y + \dot{z}_i v_i^z) - L$$

$$= \overbrace{\sum_{i=1}^{n} m_i(\dot{x}_i, \dot{y}_i, \dot{z}_i)}^{=2T} - (T - U) = T + U$$

hence H =(kinetic energy + potential energy)=<u>total energy</u> of the system.

Note that if $U_t = 0$ then $H_t = 0$ also (since $T_t = 0$). In this case the function H is a first integral of the canonical Euler-Lagrange -equations for $A[C]$. i.e. H remains constant along the extremals of $A[C]$, which are the trajectories of the moving particles. This proves the

<u>Energy Conservation law:</u> If the particles $m_1, \cdots, m_n$ move in a conservative force field whose potential energy U does not depend on time, then the total energy, H remains constant during the motion.

Note: Under suitable other conditions on U, further conservation laws for momentum or angular momentum are obtainable. See for example [GF][pages 86-88] where they are derived from Noether's theorem (ibid. §20, pages 79-83): Assume that $J[Y] = \int_{x_1}^{x_2} f(x, y, y')dx$ is invariant under a family of coordinate transformations:

$$x^* = \Phi(x, Y, Y'; \epsilon), y_i^* = \Psi_i(x, Y, Y'; \epsilon)$$

where $\epsilon = 0$ gives the identity transformation. Set $\varphi(x, Y, Y') = \frac{\partial \Phi}{\partial \epsilon}\Big|_{\epsilon=0}, \psi_i(x, Y, Y') = \frac{\partial \Psi_i}{\partial \epsilon}\Big|_{\epsilon=0}$. Then

$$(f - \sum_{i=1}^{n} y_i' f_{y_i'})\varphi + \sum_{i=1}^{n} f_{y_i'} \psi_i$$

is a first integral of the canonical Euler-Lagrange -equations for $J[Y]$.

33 Further Topics:

The following topics are natural to be treated here but for lack of time. All unspecified references are to [AK]

1. Invariance of the Euler-Lagrange -equation under coordinate transformation: §5 pages 14-20.
2. Further treatment of problems in parametric form: §13, §14, §16 pages 54-58, 63-64.
3. Inverse problem of the calculus of variations: §A-5, pages 164-167.
4. Direct method of the calculus of variations: Chapter IV, pages 127-160. (Also [GF][Chapter 8].
5. Strum-Liouville Problems as variational problems: §§A33-A36, pages 219-233. Also [GF][§41], and [CH][Vol I, chapter 6].
6. Morse theory. [M]

References

[A] T. Amaranath *An Elementary Course in Partial Differential Equations*, Alpah Science, (2003).

[AK] N.I. Akhiezer *The Calculus of Variations*, Blaisdell Publishing Co., (1962).

[B46] G.A. Bliss *Lectures on the Calculus of Variations*, University of Chicago Press, Chicago, 1946).

[B62] G.A. Bliss *Calculus of Variations* Carus Mathematical Monographs, AMS. (1962).

[Bo] O. Bolza *Lectures on the Calculus of Variations*, University of Chicago Press, (1946).

[CA] C. Carathéodory *Calculus of Variations and Partial Differential Equations of the first order* Halden-Day series in Mathematical Physics, Halden-Day (1967).

[CH] Courant, Hilbert *Methods of Mathematical Physics*, Interscience, (1953).

[GF] I.M. Gelfand, S.V. Fomin *Calculus of Variations* Prentice Hall, (1963).

[GH] M. Giaquinta, S Hildebrant *Calculus of Variations: Vol I,II* Springer (1996).

[L] Langwitz *Differential and Riemannian Geometery* Academic Press (1965).

[M] J. Milnor *Morse Theory* Princeton University Press, (1969).

[Mo] M. Morse *Variational Analysis: Critical Extremals and Sturmian Extensions*, Dover.

[O] B. O'Neill *Elementary Differential Geometry* Academic Press, (1966).

[S] D. Struik *Differential Geometry* Addison Wesley (1961).