# 1. CURVES IN $R^n$

**Definition 1.1.** *Let $I = [a, b]$ be an interval on the real line. A curve C in $R^n$ is a $C^1$ mapping $X : I \to R^n$ parametrized by t. For each $t \in I$ , we may write*

$$X(t) = (x_1(t), x_2(t), \dots, x_n(t))$$

*in terms of the Euclidean coordinates $(x_1, \dots, x_n)$ of $R^n$. Thus $X(t)$ is the position vector of a point moving on the curve. We call the parametrization regular if*

$$X'(t) \neq 0 \forall t \ .$$

*We call a regular curve an embedded curve if the mapping is one to one. A regular curve which is not necessarily one to one is sometimes called an immersion.*

**Examples 1.2.**

(1) *Straight line. $x(t) = P + t\vec{v}$ where $P = (P_1, P_2, \dots, P_n))$ is a point of $R^n$ and $\vec{v}$ is a vector in $R^n$ is the equation of the line through P in the direction of $\vec{v}$. Note that $X'(t) = \vec{v}$.*

(2) *Helix. Let $X(t) = (a\cos t, a \sin t, bt)$ where a,b are positive constants. Note that the helix lies on the cylinder $x^2 + y^2 = a^2$ in $R^3$ and rises at a constant rate ($\frac{dz}{dt} = b > 0$) as it turns counterclockwise around the cylinder. Since*

$$|X'(t)| = \sqrt{a^2 + b^2} > 0 \ ,$$

*the parametrization is regular.*

(3) *Cusp in $R^2$. Let $X(t) = (t^2, t^3)$. Then $X'(0) = \vec{0}$ so the parametrization is not regular at the origin. This is a parametrization of the (image) curve $x_2^2 = x_1^3$ which does not have a tangent line at the origin.*

(4) *The folium of Descartes in $R^2$. Let*

$$X(t) = (\frac{3t}{1+t^3}, \frac{3t^2}{1+t^3})$$

**Definition 1.3.** *The length of regular curve is the number*

$$L = \int_a^b |X'(t)| \ dt = \int_a^b \sqrt{x_1'(t)^2 + x_2'(t)^2 + \ldots + x'n(t)^2} \ dt \ .$$

This is the value obtained by the usual limit process if we approximate the curve by inscribed polygons.

If we consider only the part of the curve corresponding to values of t between a and $t_0$, then the arc length of this part is equal to

$$s(t_0) = \int_a^{t_0} |X'(t)| \ dt \ .$$

Thus s is a function of $t_0$ and by the fundamental theorem of calculus,

$$s'(t_0) = |X'(t_0)| \ .$$

If we think of t as time, the quantity $s'(t_0)$ represents the speed at which the point is moving along the curve at time $t = t_0$.

To each point $X(t) = (x_1(t), \dots x_n(t))$ of the curve C, we associate the vector

$$\vec{v}(t) = X'(t) \ .$$

When t represents time, we call $\vec{v}(t)$ the velocity vector. The condition that the curve C is regular is just $\vec{v}(t) \neq 0$. Note that

$$|\vec{v}(t)| = s'(t) \ .$$

The geometric interpretation of regularity is that the curve C has a well defined tangent direction at each point.

**Definition 1.4.** *Let C be a regular curve. The unit tangent vector to C is given by*

$$T = \frac{\vec{v}(t)}{|\vec{v}(t)|} \ .$$

It is often convenient to use the arc length s as a parameter for the curve. Then by the chain rule,

$$\vec{v} = \frac{dX}{dt} = \frac{dX}{ds}\frac{ds}{dt} = s'(t)\frac{dX}{ds} \ .$$

Note that since $|\vec{v}(t)| = s'(t)$ , $|\frac{dX}{ds} = 1$ . Hence,

$$T = \frac{dX}{ds} = X'(s)$$

1.1. **Frenet formulas for plane curves.** Let's now restrict ourselves to $C^2$ (twice continuously differentiable) regular plane curves $X(s) = (x(s), y(s))$ parametrized by arc length. To each point of the curve, we associate two orthogonal unit vectors $e_1(s), e_2(s)$ such that the pair $(e_1, e_2)$ have the same orientation as $R^2$ and such that $e_1$ is the unit tangent vector to the curve. We have just seen that $X'(s)$ is equal to the unit tangent vector $T(s)$ so $e_1(s) = X'(s)$. Note that $e_2(s) = Je_1(s)$ where J represents the operation of rotation of a vector by $\frac{\pi}{2}$ in the positive (counterclockwise) sense. It is easy to see that if $\vec{v} = (v_1, v_2)$, then $Jv = (-v_2, v_1)$. In particular, $e_2(s) = (-y'(s), x'(s))$. The Frenet formulas describe how the frame $e_1, e_2$ changes with s.

**Definition 1.5.** *The curvature vector*

$$\vec{\kappa}(s) = T'(s) = X''(s) = \frac{de_1}{ds}$$

*is the rate of change of T with respect to arc length.*

*Since $T \cdot T = 1$ , $T'(s) \cdot T = 0$, $T'(s)$ is orthogonal to $e_1(s)$ so we may write*

$$\vec{\kappa} = \kappa(s) e_2(s) \ .$$

*The number $\kappa(s)$ is called the signed curvature of the curve. Thus we may write*

$$\frac{de_1}{ds} = \kappa e_2 \ .$$

It remains to compute $\frac{de_2}{ds}$. Since $e_2(s)$ also has length 1, it follows as above that

$$\frac{de_2}{ds} = \alpha(s)e_1(s)$$

for a function $\alpha(s)$ to be determined. Taking the dot product with $e_1(s)$ we find

$$\alpha(s) = e_1(s) \cdot \frac{de_2}{ds} \ .$$

Differentiating the relation $0 = e_1(s) \cdot e_2(s)$ with respect to s gives

$$\alpha(s) = e_1(s) \cdot \frac{de_2}{ds} = -e_2(s) \cdot \frac{de_1}{ds} = -\kappa(s) \ .$$

To summarize, the Frenet formulas are

$$\frac{dX}{ds} = e_1 \ , \ \frac{de_1}{ds} = \kappa e_2 \ , \ \frac{de_2}{ds} = -\kappa e_1$$

where $\kappa(s)$ is the signed curvature of the curve and $e_2(s) = (-y'(s, x'(s))$.

## 2. The Isoperimetric inequality

Of all simple closed plane curves of given perimeter, which one bounds the greatest area? This is sometimes known as Dido's problem because of the story that Queen Dido of Tyre bargained for some land bounded on one side by the (straight) Mediterranean coast and agreed to pay a fixed sum for as much land as could be enclosed by a bull's hide. She cut the hide into thin strips and tied these end to end and formed them into a semicircle , shrewdly using the straight coastline as part of the boundary. It is easy to see that Dido's version of the isoperimetric problem is equivalent to the problem I first stated.

This isoperimetric theorem was known to the ancient Greeks although a rigorous proof did not appear until the nineteenth century. Around 1840 the geometer Jakob Steiner published several ingenious "proofs" based on purely geometric reasoning showing that an extremal curve must be a circle. His colleague Dirichlet tried (unsuccessfully) to persuade him that his proofs were incomplete because the existence of an extremal was not evident. Nevertheless, Steiner's reasoning is valuable and compelling.

The proof will proceed in several steps. The first observation is that the extremal curve $\gamma$ must be convex (the straight segment joining any two points inside $\gamma$ remains inside $\gamma$. Indeed if not (see Fig. 1) this would imply existence of two points on $\gamma$ such that the connecting segment is outside $\gamma$. Then reflecting a region between the segment and the boundary of the shape as on

Fig. 2, it would be possible to increase its area without changing its perimeter. The second step of the proof is the following

**Lemma 2.1.** *Choose points $S$ and $T$ on $\gamma$ such that the points divide the perimeter into two equal parts. Then the segment $ST$ must divide the area of the shape into two equal parts.*

Proof. Assume to the contrary that area S1T is larger than T2S (Fig. 3). Then reflecting S1T in ST will get the shape S1T3 with the same perimeter as S1T2 but a larger area.

This argument shows that it's sufficient to solve Dido's version of the problem: Among all arcs with a given length and endpoints on a line ST find the one that along with the line ST encloses the largest area.

**Lemma 2.2.** *Consider all the arcs with a given length and endpoints on a line. The curve that encloses the maximum area between it and the straight line is a semicircle.*

Proof. It suffices to show that for any point A on an extremal $\gamma$, $\angle SAT$ is a right angle. If not, think of SAT as being hinged at A and "open or close" angle A (for example we can keep RS parallel to the x axis) until the angle becomes right. Let pieces of the arc move along. As an exercise prove that among all triangles with two given sides the one whose sides form a right angle has the largest area. Since the area of the two red regions did not change but the area of the triangle (Fig 5) grew, the whole area between the new arc and the line ST has increased. This shows that unless the curve is a semicircle we can always increase the area in question by moving points S and T. This proves Lemma 2 and with it the isoperimetric theorem.

As mentioned earlier, this elegant proof is flawed because the existence of the extremal is not so clear. In Problem Sheet 1, I have outlined a way to circumvent this difficulty by considering only polygons with n sides and fixed perimeter.

**Theorem 2.3.** *Among all polygons of $n$ sides and fixed perimeter, the regular polygon maximizes the enclosed area.*

Using this result, we can prove the isoperimetric theorem by an approximation argument.

The isoperimetric theorem can be give a more quantitative and useful form called the isoperimetric inequality.

**Theorem 2.4.** *Denote the perimeter and area of a planar region $D$ bounded by a simple curve $\gamma$ by $L$ and $A$, respectively. Then*

$$L^2 \geq 4\pi A$$

*with equality only for a circle.*

2.1. **Hurwitz's proof of the the isoperimetric inequality.** A fairly simple modern analytic proof of the isoperimetric inequality was given by A. Hurwirtz in 1902. It uses the formula for the area A of a domain D in $R^2$ in terms a line integral around its (positively oriented) boundary C:

$$A = \int_C x dy = -\int_C y dx = \frac{1}{2}\int_C x dy - y dx$$

where the line integrals are defined by the usual limit of Riemann sums. For our purposes, if $C : x = x(t)\ ,\ y = y(t)\ a \leq t \leq b$ is a parametrization of a simple curve C, then

$$\int_C x dy = \int_a^b x(t)y'(t)dt\ ,\ \int_C y dx = \int_a^b y(t)x'(t)dt\ .$$

The formulas for A follow from the divergence theorem:

$$\int_D \mathrm{div} V(\vec{x}, y) dx dy = \int_C \vec{V} \cdot \vec{n} ds\ ,$$

where if $V = (V^1, V^2)$ , div $V(\vec{x}, y) = V^1_x + V^2_y$ , $\vec{n}$ is the outer unit normal to C and s is arc length. Choosing $V = (x, y)$ , div $\vec{V} = 2$. Noting that in terms of the parametrization of C,

$$\vec{n} = \frac{(y'(t), -x'(t)}{\sqrt{x'(t)^2 + y'(t)^2}} \text{ and } ds = \sqrt{x'(t)^2 + y'(t)^2}\ dt\ ,$$

the formulas for A follow easily.

Hurwitz's proof also uses a famous inequality known as Wirtinger's inequality.

**Lemma 2.5.** *Let $f(t)$ be a $C^1$ periodic function of period $2\pi$ and average $\int_0^{2\pi} f(t)\ dt = 0$. Then*

$$\int_0^{2\pi} f'(t)^2\ dt \geq \int_0^{2\pi} f(t)^2\ dt$$

*with equality only if $f(t) = a\cos t + b\sin t$.*

Proof. Expand $f(t)$ in a Fourier series

$$f(t) = \sum_{n=1}^{\infty}(a_n \cos nt + b_n \sin nt)$$

(note that the constant term $\frac{1}{2}a_0$ is zero by the average zero condition). Then

$$f'(t) = \sum_{n=1}^{\infty}(nb_n \cos nt - na_n \sin nt)$$

By Parseval's formula,

$$\int_0^{2\pi} f(t)^2 \, dt = \sum_{n=1}^{\infty} (a_n^2 + b_n^2) \ ,$$

$$\int_0^{2\pi} f'(t)^2 \, dt = \sum_{n=1}^{\infty} n^2 (a_n^2 + b_n^2) \ .$$

Hence,

$$\int_0^{2\pi} (f'(t)^2 f(t)^2) \, dt = \sum_{n=1}^{\infty} (n^2 - 1)(a_n^2 + b_n^2) \ ,$$

which is strictly positive unless $a_n = b_n = 0$ for all $n > 1$, proving the lemma.

For simplicity assume that the length L of C is $2\pi$ (convince yourself that we can do this without loss of generality). We may also assume by translating C that

$$\int_0^{2\pi} x(t) \, ds = 0 \ .$$

Then

$$2\pi = \int_0^{2\pi} ((x')^2 + (y')^2) \, ds \ ,$$

and $A = \int_0^{2\pi} x(s) y'(s) \, ds$. Hence,

$$2(\pi - A) = \int_0^{2\pi} ((x')^2 - 2xy' + (y')^2) \, ds = \int_0^{2\pi} ((x')^2 - x^2) \, ds + \int_0^{2\pi} (x - y')^2 \, ds \ .$$

The first integral is nonegative by the lemma and the second integral is also clearly nonnegative. Hence $4\pi A \leq (2\pi)^2 = L^2$ which is the isoperimetric inequality. The equality sign holds only if

$$x = a\cos s + b\sin s \ , \ y' = x$$

which gives

$$y = a\sin s - b\cos s + c \ .$$

Thus $x^2 + (y - c)^2 = a^2 + b^2$ and C is a circle.

## 3. CALCULUS OF VARIATIONS FORMALISM: NECESSARY CONDITIONS

The simplest and most common problems in the Calculus of Variations are of the type where we seek to minimize (or maximize) a functional

$$E(u) = \int_a^b f(x, u(x), u'(x)) \, dx$$

where $f = f(x, z, p)$ is a given (say $C^2$) function of its arguments and the function $u(x)$ belongs to a certain admissible class, for example $C^1$ functions or more generally perhaps piecewise $C^1$ (or even Lipschitz) functions satisfying given boundary conditions $u(a) = A\ ,\ u(b) = B$ where A,B are prescribed.

**Examples 3.1.** *1. The Brachistochrone problem*

formulated by Johann Bernoulli in 1696 as a challenge problem to his colleagues , seeks among all curves joining two given points Pand Q, the path with the property that a frictionless particle acted on by gravity would have the minimum transit time. Earlier, Galileo had guessed that the circle might be the extremum and would likely be better than the straight line path. If we take the higher point as the origin of our coordinate system (in a vertical plane passing through PQ) with the y axis pointing down and the x axis to the right)and the second point as $(x_1, y_1)$ then

$$T = \int_0^L \frac{ds}{v}\ ,$$

where s is arc length ranging from 0 to L and $v = \frac{ds}{dt}$ is the speed along the curve. From Newton's laws or conservation of energy, it is not difficult to show that $v^2 = 2gy$. Assume that the path is a graph $y = y(x)$; then $ds = \sqrt{1 + (y')^2}\ dx$ so

$$T(y) = \frac{1}{\sqrt{2g}} \int_0^{x_1} \sqrt{\frac{1 + (y')^2}{y(x)}}\ dx$$

is the functional to be minimized among all $y(x)$ satisfying $y(0) = 0\ ,\ y(x_1) = y_1$. Note that the integrand is singular at $y = 0$.

2. *Minimum area of a surface of revolution* Consider the surface of revolution obtained by revolving the graph $y = f(x)$ joining $(a, A), (b, B)$ (with $a < b\ ,\ A, B > 0$) about the x-axis. This is given by

$$A(f) = 2\pi \int_a^b f(x)\ \sqrt{1 + (f')^2}\ dx$$

The problem is to determine f so that $A(f)$ is a minimum. Physically, this is a special version of the Plateau problem, where we seek a soap film spanning two wire circles (representing the boundary conditions) which seeks to minimize the surface area. This problem, which we discuss later, is degenerate in the sense (do the experiment) that when the circles are far enough apart, the minimum is not a connected surface but the flat disks spanning the individual circles.

i3. Here is a simple (made-up) example of a functional for which there is no extremum (either max or min). Let

$$E(u) = \int_0^1 \frac{dx}{1 + (u')^2}\ dx$$

where $u(0) = 0$ and $u(1) = 1$ Clearly $0 < I(u) < 1$ since the integrand is positive and less than 1. By taking $u(x) = 0$ on $(0, 1-\epsilon)$ and $u(x) = 1 + \frac{1}{\epsilon}(x-1)$ on $(1-\epsilon, 1)$ we can make $I(u)$ as close to 1 as we please but there is no admissible u with $I(u) = 1$. We can similarly find u with $I(u)$ as close to zero as we please but again there is no admissible u with $I(u) = 0$.

In the following discussion we will let $u(x) = (u^1(x), \ldots, u^N(x))$ be a vector function since the arguments are unchanged and this will allow us greater applicability.

**Definition 3.2.** *Let $I = (a, b)$. A function $u \in C^1(I, R^N)$ is said to be a weak $C^1$ extremal for the functional $E(u) = \int_a^b f(x, u, u')\ dx$ if*

$$E'(u)(\phi) = \int_a^b (f_z(x, u, u')\phi + f_p(x, u, u') \cdot \phi')\ dx = 0\ , \tag{3.1}$$

*for all $\phi \in C_0^\infty(I, R^N)$*

The expression $E'(u)(\phi)$ (sometimes called the first variation) in equation (3.1) is the directional derivative (Gateaux derivative) of the functional E in the direction $\phi$ at $u(x)$. Formally,

$$E'(u)(\phi) = \lim_{t \to 0} \frac{E(u + t\phi) - E(u)}{t} \tag{3.2}$$

These expressions arise naturally from making a variation $u(x, t)$in the direction $\phi(x)$. This is a family of admissible functions $u(x, t)$ for $|t|$ small such that $u(x, 0) = u(x)$ and $\frac{d}{dt}u(x, t)|_{t=0} = \phi(x)$. The simplest type of variation is $u(x, t) = u(x) + t\phi(x)$ with $\phi$ as above. Then

$$E'(u)(\phi) = \frac{d}{dt}E(u(x, t)|_{t=0} = \frac{d}{dt}\int_a^b E(x, u(x, t), u'(x, t)\ dx|_{t=0}$$

Differentiating under the integral sign gives (3.1). The interpretation of weak extremal is that $u(x)$ is a stationary point for the functional $E(u)$.

**Definition 3.3.** *A function $u \in C^1(I, R^N)$ is said to be a weak $C^1$ local minimum for the functional $E(u)$ if $E(u) \leq E(u + t\phi)$ for all $\phi \in C_0^\infty(I, R^N)$ and $|t|$ sufficiently small (depending on $\phi$.) Clearly, a weak minimum is necessarily a weak extremal.*

**Lemma 3.4.** *(Fundamental lemma) Suppose $f \in C^0(I)$ satisfies*

$$\int_a^b f(x)\eta(x)\ dx = 0\ \forall \eta \in C_0^\infty(I)\ . \tag{3.3}$$

*Then $f(x) \equiv 0$ on $I$.*

Proof. Since smooth functions of compact support are dense in the space of piecewise $C^1$ functions which vanish at the endpoints of I, it follows that we can expand the admissible functions $\eta(x)$ to include this broader class. Thus if $f(x_0) > 0$, then $f(x) > 0$ on $(x_0 - 2\delta, x_0 + 2\delta)$ for $\delta$ small. Choose $eta = 1$ on $x_0 - \delta, x_0 + \delta)$ $\eta = 0$ on $I \setminus (x_0 - 2\delta, x_0 + 2\delta)$ and linear in between. This gives $\int_I f\eta \, dx > 0$, a contradiction. Thus $f \equiv 0$ on I.

**Theorem 3.5.** *(Euler equations) Assume $f(x, z, p)$ is $C^2$ in its arguments in a neighborhood of a weak extremal $u \in C^2(I, R^N)$ Then $u(x)$ satisfies the so called Euler equations*

$$-\frac{d}{dx} f_{p^i}(x, u, u') + f_{z^i}(x, u, u') = 0 \ , \ i = 1, \dots, N \ . \tag{3.4}$$

Proof. In (3.1) choose $\phi = (\phi^1, \dots, \phi^N)$ where $\phi^k = \delta_{ki}\eta \ , \ \eta \in C_0^\infty(I)$. Then integrating by parts we find

$$\int_I (f_{z^i} - \frac{d}{dx} f_{p^i})\eta \, dx = 0 \ .$$

Applying the fundamental lemma gives (3.4).

**Examples 3.6.**

4. Let $f(x, z, p) = \sqrt{1 + p^2} + \omega(x, z)$ , $(N = 1)$ . Then the Euler equation associated to $E(u) = \int_a^b (\sqrt{1 + (u')^2} + \omega(x, u)) \, dx$ is

$$\frac{d}{dx}(\frac{u'}{\sqrt{1 + (u')^2}} = \omega(x, u) \tag{3.5}$$

The quantity

$$\kappa = \frac{d}{dx}(\frac{u'}{\sqrt{1 + (u')^2}}) = \frac{u''}{(1 + (u')^2)^{\frac{3}{2}}}$$

is the curvature of the graph $y = u(x)$ with respect to the upward normal and the Euler equation (3.5) says that the curvature is a prescribed function of position and height.

5. Let $f(x, z, p) = \frac{1}{2}(p^2 + q(x) + z^2)$ , $(N = 1)$ . Then the Euler equation associated to the energy functional

$$E(u) = \int_a^b ((u')^2 + q(x)u^2) \, dx$$

is

$$-u'' + q(x)u = 0 \tag{3.6}$$

6. Let $x = x(t) : I \to R^3$ represent the motion of a particle of mass $m$ in $R^3$ and $V(x) \in C^1(R^3)$ a potential function. Consider the total energy (kinetic+potential) functional

$$E(x) = \int_a^b (\frac{m}{2}|x'|^2 - V(x))\ dt \tag{3.7}$$

Then the Euler equations are just Newton's laws for a point mass in a conservative force field $-\nabla V(x)$:

$$mx''(t) = -\nabla V(x(t) \tag{3.8}$$

We now will present an alternate (weak) form of the Euler equations that applies to weak $C^1$ extremals. To derive it we need an alternate form of the fundamental lemma.

**Lemma 3.7.** *Let $f(x) \in C^0(I)$ satisfy*

$$\int_a^b f(x)\eta'(x)\ dx = 0\ ,$$

*for all test functions $\eta \in C_0^\infty(I)$. Then $f \equiv c$ (constant).*

Proof. Note that $c = \frac{1}{b-a}\int_I f(x)\ dx$ is the precise constant and that

$$\int_a^b (f(x) - c)\eta'(x)\ dx = 0\ .$$

As we have stated earlier we can take a broader class of test functions $\eta$. In particular, define

$$\eta(x) = \int_a^x (f(t) - c)\ dt$$

Then $\eta(a) = \eta(b) = 0$ and $\eta'(x) = f(x) - c$, hence

$$\int_a^b (f(x) - c)^2\ dx = 0$$

and so $f(x) \equiv c$.

**Theorem 3.8.** *(weak Euler equations) Let $u \in C^1(I, R^N)$ be a weak extremal. Then there is a constant vector $\vec{c} \in R^N$ such that*

$$f_p(x, u, u') = \vec{c} + \int_a^x f_z(t, u(t), u'(t))\ dt\ . \tag{3.9}$$

Proof. In the first variation (3.1) we integrate by parts the term

$$\int_a^b f_z(x, u, u')\phi(x)\ dx = -\int_a^b (\int_a^x f_z(t, u(t), u'(t)\ dt)\phi'(x)\ dx\ ,$$

so

$$\int_a^b (f_p - \int_a^x f_z\ dt)\phi'(x) dx = 0\ .$$

The theorem follows just as in the proof of Euler equations using now the alternate form of the fundamental lemma.

Note that if the weak Euler equation (3.9) implies that $f_p(x, u(x), u'(x))$ is $C^1$ since the right hand side is clearly $C^1$. Hence we have

**Corollary 3.9.** *Any $C^1$ weak extremal $u(x)$ satisfies*

$$-\frac{d}{dx} f_{p^i}(x, u, u') + f_{z^i}(x, u, u') = 0\ ,\ i = 1, \ldots, N\ . \tag{3.10}$$

Even if $f$ is $C^2$ in its arguments, there is a catch here in that we cannot differentiate $f(x, u, u')$ unless we know that $u''$ exists. In general this is not true as the following example shows.

**Example 3.10.** *6. (A $C^1$ weak min which is not $C^2$)*

Let

$$E(u) = \int_{-1}^{1} u^2(2x - u')^2\, dx\ ,$$

with boundary conditions $u(-1) = 0$ , $u(1) = 1$. It is easy to see that the unique minimizer is $u(x) = 0$ on $[-1, 0]$ , $u(x) = x^2$ , on $[0, 1]$. This is piecewise $C^2$ but fails to be $C^2$ at the origin.

Formally, we can see what is happening by carrying out the differentiation:

$$\frac{d}{dx} f_p(x, u, u') = f_{px}(x, u, u') + f_{pz}(x, u, u')u' + f_{pp}(x, u, u')u''\ . \tag{3.11}$$

Note that the "coefficient" of the vector $u'(x)$ is the matrix $f_{p^i p^j}(x, u, u')$. If this is invertible, we can solve the weak Euler equations for $u'$ using the implicit function theorem and see that it is $C^1$. More precisely,

**Theorem 3.11.** *Let $u$ be a weak $C^1$ extremal of $E$ on $I$ and suppose $f_p$ is $C^1$ in a neighborhood of $(x, u, u')$ and that $f_{pp}(x, u, u')$ is invertible for all $x \in \overline{I}$. Then $u \in C^2(\overline{I}, R^N)$.*

There is one important simplification to the Euler equations when $f = f(z, p)$ is independent of x. Define

$$\Phi(z, p) = p f_p(z, p) - f(z, p)$$

and observe that if $u \in C^2(I, R^n)$ satisfies the Euler equations, then

$$\frac{d}{dx}\Phi(u, u') = u' \frac{d}{dx} f_p(u, u') + u'' f_p(u, u') - f_z(u, u')u' - f_p(u, u')u'' \tag{3.12}$$

$$= u'\left(\frac{d}{dx} f_p(u, u') - f_z(u, u')\right) = 0\ . \tag{3.13}$$

(3.14)

It follows that $\Phi(u(x), u'(x)) = \vec{c}$ on $I$.

Therefore we have proved

**Theorem 3.12.** *Suppose u is* $C^2$ *and satisfies the Euler equations for* $E(u) = \int_a^b f(u,u')\ dx$. *Then*

$$u' f_p(u,u') - f(u,u') = \vec{c} \text{ on } I \ . \tag{3.15}$$

**Example 3.13.** *Returning to the brachistochrone problem with* $f(z,p) = \sqrt{\frac{1+p^2}{z}}$ *(we drop the coefficient for simplicity), we see that a solution* $y = u(x)$ *of the Euler equation must satisfy*

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

*for some constant c. Simplifying gives*

$$c^2\ u(1+(u')^2) = 1 \ .$$

*It is convenient to express the solution in parametric form* $u = u(t)\ ,\ x = x(t)$ *where we define t by* $u = k(1-\cos t)$ *where we choose* $kc^2 = \frac{1}{2}$. *Then*

$$u' = \frac{du}{dx} = \frac{\frac{du}{dt}}{\frac{dx}{dt}} = \frac{k\sin t}{\dot{x}} \ .$$

*Substiting this above gives*

$$c^2\ k(1-\cos t)(1 + \frac{k^2\sin^2 t}{\dot{x}^2} = 1 \ ,$$

*which after simplification yields*

$$\dot{x}(t) = k(1-\cos t)\ ,\ x(t) = k(t - \sin t) \ .$$

*The parametric equations*

$$x = k(t-\sin t)\ ,\ u = k(1-\cos t) \ ,$$

*define a cycloid.*

## 3.1. **Subsidiary Constraints-Lagrange Multipliers.**

**Theorem 3.14.** *Consider the problem*

$$I(y) = \int_a^b f(x,y,y')\ dx \to \ min\ ,\ y(a) = A, y(b) = B$$

*and* $J(y) = \int_a^b g(x,y,y')\ dx = C$ .
*We suppose y is a weak* $C^1$ *extremal for I but not for J (i.e.* $\delta J(\psi) \neq 0$ *for some* $\eta \in C^1$ *with compact support. Then there exists a multiplier* $\lambda$ *such that y(x) is a weak* $C^1$ *weak extremal for* $I + \lambda J$.

Proof. Let $\tilde{y} = y + \epsilon\eta + \beta(\epsilon)\psi$ . Let

$$J(\epsilon,\beta) \equiv J(\tilde{y}) = \int_a^b g(x, y+\epsilon\eta+\beta\psi, y'+\epsilon\eta'+\beta\psi')\ dx \ ,$$

Then $J(0,0) = C$ and

$$\frac{\partial J}{\partial \beta}|\epsilon = \beta = 0 = \int_a^b (g_y \psi + g_{y'}\psi')\ dx = \delta J(psi) \neq 0\ .$$

Hence by the Implicit function theorem there exists for $\epsilon$ small a unique $C^1$ curve of solutions $\beta = \beta(\epsilon)$ to $J(\epsilon, \beta) \equiv C$. Differentiating with respect to $\epsilon$ gives

$$\beta'(0) = -\frac{\delta J(y)}{\delta J(\psi)} \tag{3.16}$$

Hence $\tilde{y}$ is an admissible variation of y so that

$$\frac{d}{d\epsilon} I(y + \epsilon\eta + \beta(\epsilon)\psi)|_{\epsilon=0} = 0\ .$$

This gives

$$\delta I(y) + \beta'(0)\delta I(\psi) = 0. \tag{3.17}$$

Using (3.16), equation (3.17) may be rewritten

$$\delta(I(y) + \lambda J(y)) = 0$$

(where $\lambda = -\frac{\delta I(\psi)}{\delta J(\psi)}$) which is the Lagrange multiplier rule. The rule is easily extended to the case of many independent integral constraints.

3.2. **Endpoint variations and Transversality.** We now consider more general admissible variations in which the endpoints are allowed to vary in a general way. So consider the functional $I(y) = \int_a^b F(x, y, y')\ dx$ for a piecewise $C^1$ curve $y(x)$ with endpoints $A = (a, \lambda)$ and $B = (b, \mu)$ We want to compute $\delta I(y)$ for a variation in which one or both of the endpoints can move.

We consider a family of curves

$$y(x, \epsilon) := y(x, a(\epsilon), \lambda(\epsilon), b(\epsilon), \mu(\epsilon) \tag{3.18}$$

for which $y(x, 0) = y(x)$ . Then we set

$$J(\epsilon) = \int_{a(\epsilon}^{b(\epsilon)} F(x, y(x, \epsilon), y'(x, \epsilon))\ dx$$

and compute the total derivative

$$\frac{dJ}{d\epsilon}|_{\epsilon=0} = \delta I(\delta y(x)) \quad + \quad \frac{\partial I}{\partial a}\delta a \tag{3.19}$$

$$+\frac{\partial I}{\partial \lambda}\delta\lambda \quad + \quad \frac{\partial I}{\partial b}\delta b \frac{\partial I}{\partial \mu}\delta\mu\ . \tag{3.20}$$

We obtain

$$J'(0) = F|_{x=b}\delta b - F|_{x=a}\delta a + \int_a^b (F_y \eta(x) + F_{y'}\eta'(x))\ . \tag{3.21}$$

Here we have used the abreviated notation $\eta(x) = \frac{d}{d\epsilon}y(x,\epsilon)|_{\epsilon=0}$ and have interchanged the order of differentiation

$$\frac{d}{d\epsilon}y'(x,\epsilon) = \frac{d}{dx}\frac{d}{d\epsilon}y(x,\epsilon) = \eta'(x)$$

at $\epsilon = 0$. Therefore we can in (3.21) integrate the last term of the integral by parts to obtain

$$J'(0) = F|_{x=b}\delta b - F|_{x=a}\delta a + \int_a^b (F_y - \frac{d}{dx}F_{y'})\eta \; dx + F_{y'}\eta|_a^b \; . \tag{3.22}$$

It remains only to observe that

$$\eta|_{x=a} \;=\; \delta\lambda - y'(a)\delta a \tag{3.23}$$

$$\eta|_{x=b} \;=\; \delta\mu - y'(b)\delta b \tag{3.24}$$

Inserting (3.24) and (3.24) into (3.22) gives

$$J'(0) = (F - y'F_{y'})|_{x=b}\delta b - (F - y'F_{y'})|_{x=a}\delta a + F_{y'}|_{x=b}\delta\lambda \tag{3.25}$$

$$-F_{y'}|_{x=a}\delta\mu + \int_a^b (F_y - \frac{d}{dx}F_{y'})\eta \; dx \tag{3.26}$$

Now assume that y itself is a $C^2$ extremal for I; then $J'(0) = 0$ and by making a variation fixing A and B , y must satisy the E-L equations and the last integral term in (3.26) vanishes. Suppose also that endpoint A is fixed (so $\delta a = \delta\mu = 0$) and endpoint B varies along a curve $\tau(x, y) = 0$. Then

$$\tau_x\delta b + \tau_y\delta\lambda = 0 \; . \tag{3.27}$$

Then multiplying (3.26) by $\tau_y$ and using (3.27) we obtain the **transversality condition**

$$\tau_x(b, y(b))F_{y'}(b, y(b), y'(b)) = \tau_y(b, y(b))(F - y'F_{y'})(b, y(b)y'(b)) \; . \tag{3.28}$$

**Example 3.15.** *(Geometric optics)*
*We can interpret our conditions for a light ray emanating from a single point and traveling in a plane medium with index of refraction* $n(x, y)$ *and terminating in a curve described by the graph of a function* $y = \psi(x)$*. According to Fermat's principle, the light propagates along the trajectory which minimizes the time of travel.*

*Let* $I(y) = \int_a^b n(x, y)\sqrt{1 + y'^2}\; dx$ *and let* $\tau(x, y) = y - \psi(x)$ *Then*

$$F_{y'} = n(x, y)\frac{y'}{\sqrt{1 + y'^2}} = \frac{y'}{1 + y'^2}F$$

*and*

$$F - y'F_{y'} = \frac{F}{1 + y'^2} \; .$$

*Substitution in (3.28) gives the orthogonality condition* $y'(b) = -\frac{1}{\psi'(b)}$.

We note for later use when we study Hamilton-Jacobi theory that formula (3.26) remains valid when $y = (y_1, \ldots, y_n)$ is a curve in $R^n$ (so $\lambda$ and $\mu$ are vectors in $R^n$) . In terms of new variables $p_i = F_(y'_i)$ , $i = 1, \ldots, n$ we suppose that the jacobian determinant $\det F_{y'_i y'_j} \neq 0$ so we can (locally) solve for $y'_1, \ldots, y'_n$ in terms of the other variables $x, y_1, \ldots, y_n, p_1, \ldots, y_n$. Introduce the Hamiltonian $H(x, y, p)$ through the Legendre transform $H = -F + p \cdot y'$ where the $y'$ are regarded as functions of $x, y, p$. Now we assume that the point A is fixed and the final point $B = (x, y)$ is variable and that the "action" integral

$$S(x, y) = \int_{x_0}^{x} F(x, y, y')\ dx$$

is evaluated on a unique extremal passing through A and B so that S becomes a single valued function (called the geodesic distance) between A and B. Then formula (3.26) can be rewrittten as

$$dS(x, y) = pdy - Hdx \tag{3.29}$$

or in other words,

$$\frac{\partial S}{\partial x} = -H(x, y, p)\ ,\ \frac{\partial S}{\partial y} = p\ . \tag{3.30}$$

From (3.30) we conclude that S satisfies the Hamilton-Jacobi equation.

$$S_x + H(x, y, S_y) = 0 \tag{3.31}$$

3.3. **Parameter variations.** Let $u \in C^1(I, R^N)$ , $I = (a, b)$ and let $x = \psi(t, \epsilon)$ be a $C^1$ diffeomorphism of I onto itself (for $|\epsilon|$ small) such that $\psi(a, \epsilon) = a$ , $\psi(b, \epsilon) = b$ , $\psi(t, 0) = t$. Assume also $\frac{d\psi}{d\epsilon}(\cdot, \epsilon)$ is $C^1$. Then $\psi(\cdot, \epsilon)$ is called an admissible parameter variation. Now set $v(t, \epsilon) = u(\psi(t, \epsilon))$ and note that $v(t, 0) \equiv u(t)$ , $v(a, \epsilon) = u(a)$ , $v(b, \epsilon) = u(b)$ and that $v$ is $C^1$ close to $u$ (see below for details).

Set $\lambda(t) = \frac{d\psi}{d\epsilon}(t, \epsilon)$ and

$$\Psi(\epsilon) = \int_a^b f(t, v(t, \epsilon), \dot{v}(t, \epsilon))\ dt = I(v(t, \epsilon))\ . \tag{3.32}$$

**Theorem 3.16.** *Let u be a weak $C^1$ local minimum of I with fixed endpoints. Then*

$$0 = \int_a^b \{[u' f_p(u) - f(u)]\lambda'(x) - f_x(u)\lambda(x)\}\ dx\ . \tag{3.33}$$

Proof. We have $x = \psi(t, \epsilon)$ with inverse function $t = \tau(x, \epsilon) = \tau(\psi(t, \epsilon), \epsilon)$. We write $' = \frac{d}{dx}$ , $\cdot = \frac{d}{dt}$ and so

$$1 = \tau'(\psi(t, \epsilon), \epsilon)\dot{\psi}(t, \epsilon)\ ,$$

$$0 = \tau_\epsilon() + \tau'()\psi_\epsilon \; .$$

Furthermore since $t = \psi(t, 0) = x$ we have $\tau'(x, 0) = 1$ and so $\tau_\epsilon(x, 0) = -\lambda(x)$. Thus

$$\tau(x, \epsilon) = x - \epsilon\lambda(x) + \dots \; .$$

Recalling the definition of $\Psi(\epsilon)$ in (3.32) we apply the change of variable $t = \tau(x, \epsilon)$. Then,

$$\Psi(\epsilon) = \int_a^b f(\tau(x,\epsilon), u(x), \frac{u'(x)}{\tau'(x,\epsilon)})\tau'(x,\epsilon)\; dx$$

Now we use that

$$\frac{d}{d\epsilon}\tau'(x,\epsilon) = \frac{d}{dx}\frac{d}{d\epsilon}\tau(x,\epsilon)$$

and

$$\frac{d}{d\epsilon}\tau'(x,\epsilon)|_{\epsilon=0} = -\lambda'(x)$$

to compute

$$0 = \Psi'(0) = \int_a^b \{f(u)(-\lambda'(x)) + f_x(u)(-\lambda(x)) + u'(x)f_p(u)\lambda'(x)\}\; dx \; .$$

After rearranging terms, this is (3.33).

**Corollary 3.17.** *Let $u$ be a $C^1$ weak local minimum of $I$ with fixed endpoints. Then there a vector $\vec{c}$ so that*

$$f(x,u,u') - u'f_p(x,u,u') - \int_a^x f_x(t, u(t), u'(t)\; dt = \vec{c}\; . \tag{3.34}$$

*In particular, if $f = f(z, p)$ is independent of $x$,*

$$f(x,u,u') - u'f_p(x,u,u' = \vec{c}\; . \tag{3.35}$$

3.4. **Weierstrass' necessary condition.** The Weierstrass excess function or E-function is a measure of the convexity of $f(x, z, p)$ in the $p$ variable and is defined by

$$E(x,z,p,q) := f(x,z,q) - f(x,z,p) - f_p(x,z,p)\cdot(q-p) \tag{3.36}$$

Note that for fixed $(x, z)$ , $f(x, z, p)$ is convex in $p$ if $E(x, z, p, q) \geq 0$ , $\forall q$.

**Theorem 3.18.** *(Weierstrass necessary condition) Suppose $u \in C^1(I, R^N)$ is a $C^0$ (strong)relative minimum for I. Then $E(x, u, u', q) \geq 0\; \forall q \in R^N$ .*

Proof. Fix $x_0 \in (a, b)$ , $q \in R^N$ and define for $\epsilon$ , $h > 0$ fixed small ,

$$\eta(x) = \begin{cases} x - x_0 & \text{if } x_0 \leq x \leq x_0 + \epsilon\; h \\ \frac{\epsilon}{1-\epsilon}(x_0 + h - x) & \text{if } x_0 + \epsilon\; h \leq x \leq x_0 + h \\ 0 & \text{otherwise} \end{cases}$$

Now define the variation $\tilde{u} = u + \eta\ (q-p)$, where $p = u'(x_0)$ . Note that $\tilde{u}$ is only $C^0$ close to $u$ so is only valid because we assume $u$ is a strong local minimum. Hence

$$0 \leq I(\tilde{u}) - I(u) = \int_{x_0}^{x_0+h} (f(x, u + \eta\ (q-p), u' + \eta'\ (q-p)) - f(x, u, u'))\ dx$$

Define a change of variable $x = x_0 + hs$; then

$$0 \leq \int_0^1 (f(x_0 + hs, u(x_0 + hs) + h\gamma(s)(q-p), u'(x_0 + hs) + \gamma'(s))(q-p))\ ds$$

$$- \int_0^1 f(x_0 + hs, u(x_0 + hs), u'(x_0 + hs))\ ds\ .$$

where

$$\gamma(s) = \begin{cases} s & \text{if } 0 \leq s \leq \epsilon \\ \frac{\epsilon}{1-\epsilon}(1-s) & \text{if } \epsilon \leq s \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Now split the integral (for the first term only) over $(0, \epsilon)$ and $(\epsilon, 1)$:

$$0 \leq \int_0^\epsilon f(x_0 + hs, u(x_0 + hs) + hsq, u'(x_0 + hs) + (q-p))\ ds$$

$$+ \int_\epsilon^1 f(x_0 + hs, u(x_0 + hs) + h\frac{\epsilon}{1-\epsilon}(1-s)(q-p), u'(x_0 + hs) - \frac{\epsilon}{1-\epsilon}(q-p))\ ds$$

i

$$- \int_0^1 f(x_0 + hs, u(x_0 + hs), u'(x_0 + hs))\ ds\ .$$

Taking the limit as $h \longrightarrow 0$ gives

$$0 \leq \epsilon f(x_0, u(x_0), q) + (1-\epsilon) f(x_0, u(x_0), u'(x_0) - \frac{\epsilon}{1-\epsilon}(q-p)) - f(x_0, u(x_0), u'(x_0))$$

Finally, simplifying and dividing by $\epsilon$ and taking limits as $\epsilon \longrightarrow 0$ gives

$$0 \leq f(x_0, u(x_0), q) - f(x_0, u(x_0), u'(x_0)) - f_p(x_0, u(x_0), u'(x_0))(q - u'(x_0))\ .$$

which proves the result.

3.5. **The second variation and Legendre's necessary condition.** In this section we suppose that $u \in C^1(I, R^N)$ is a weak local minimizer in the sense that $I(u) \leq I(u + \epsilon\phi)$ , $\forall \phi \in C_0^1(I, R^N)$ and $\epsilon$ sufficiently small. We assume also that $f(x, z, p)$ is $C^2$ in its arguments.

**Theorem 3.19.** *(Legendre necessary condition) Under the above assumptions,*

$$f_{p^i p^j}(x, u, u')\zeta_i \zeta_j \geq 0\ ,\ \forall \zeta \in R^N\ .$$

Proof. $h(\epsilon) = I(u+\epsilon\phi)$ satisfies $h'(0) = 0$ , $h''(0) \geq 0$ . Hence differentiating twice under the integral sign gives

$$0 \leq I''(u,\phi) := \int_a^b (f_{p^ip^j}(u)\phi^{i'}\phi^{j'} + 2f_{z^ip^j}(u)\phi^i\phi^{j'} + f_{z^iz^j}(u)\phi^i\phi^j)\ dx \tag{3.37}$$

Now fix $x_0 \in (a,b)$ , $\zeta \in R^N$ and define $\phi = \eta(x)\zeta$ where

$$\eta(x) = \begin{cases} \frac{\lambda}{\delta}(x - x_0 + \delta) & \text{if } x_0 - \delta \leq x \leq x_0 \\ \frac{\lambda}{\delta}(x_0 + \delta - x) & \text{if } x_0 \leq x \leq x_0 + \delta \\ 0 & \text{otherwise} \end{cases}$$

By the usual density arguments, this piecewise $C^1$ test function is allowable. Hence,

$$0 \leq \frac{\lambda^2}{\delta^2}\int_{x_0-\delta}^{x_0+\delta} (f_{p^ip^j}(u)\zeta^i\zeta^j + 2f_{z^ip^j}(u)\zeta^i\zeta^j \cdot O(\delta) + f_{z^iz^j}(u)\zeta^i\zeta^j \cdot O(\delta^2))\ dx\ .$$

After cancelling the $\frac{\lambda^2}{\delta^2}$ factor, we divide by $\delta$ and take the limit as $\delta \to 0$. This gives

$$f_{p^ip^j}(x_0, u(x_0), u'(x_0))\zeta_i\zeta_j \geq 0\ .$$

Since $x_0$ is arbitrary, this completes the proof.

3.6. **The Erdmann corner conditions.** Thus far, we have not spoken about the possibility of piecewise $C^1$ extremals. For later applications, we now allow this and derive the so called corner conditions, the necessary conditions that must hold at the "corner points" (the points where the left and right hand derivatives exist but may be different).

**Theorem 3.20.** *(Erdmann corner conditions)Let $u$ be a piecewise $C^1$ extremal. Then if $x_0$ is a corner point,*

$$f_p(x_0, u(x_0), u'(x_0^-)) = f_p(x_0, u(x_0), u'(x_0^+)) \tag{3.38}$$

$$u'(x_0^-)f_p(x_0, u(x_0), u'(x_0^-)) - f(x_0, u(x_0), u'(x_0^-)) = \tag{3.39}$$

$$= u'(x_0^+)f_p(x_0, u(x_0), u'(x_0^+)) - f(x_0, u(x_0), u'(x_0^+) \tag{3.40}$$

Proof. First note that if $u$ is only a piecewise $C^1$ extremal then from the proof of Theorem 3.8 of the weak Euler-Lagrange equations, we find that

$$f_p(x, u, u') = \vec{c} + \int_a^x f_z(t, u(t), u'(t))\ dt$$

holds everywhere except possibly at the corners. However the right hand side is continuous at a corner point $x_0$ so the left and right hand limits of $f_p$ must agree and part i. is proved.

Similarly, from the derivation of Corollary 3.15 by parameter variations, we find that

$$f(x,u,u') - u' f_p(x,u,u') - \int_a^x f_x(t,u(t),u'(t)\ dt = \vec{c}$$

holds except possibly at the corner points. Again the right hand side is continuous at $x_0$ so the left and right hand limits of the left hand side must agree, proving part ii.

3.7. **Jacobi's necessary condition.** The derivation of Legendre's necessary condition used the nonnegativity of the second variation

$$0 \le I''(u,\phi) := \int_a^b (f_{p^ip^j}(u)\phi^{i'}\phi^{j'} + 2f_{z^ip^j}(u)\phi^i\phi^{j'} + f_{z^iz^j}(u)\phi^i\phi^j)\ dx\ .$$

Set

$$F(x,\phi,\phi') = f_{p^ip^j}(u)\phi^{i'}\phi^{j'} + 2f_{z^ip^j}(u)\phi^i\phi^{j'} + f_{z^iz^j}(u)\phi^i\phi^j \tag{3.41}$$

**Definition 3.21.** *The functional* $J(\phi) = \int_a^b F(x,\phi,\phi')\ dx$ *is called the Jacobi accessory functional. For* $u = u_0$ *a fixed extremal, the Euler-Lagrange equation of the accessory functional* $J(\phi)$

$$-\frac{d}{dx}F_p + F_z = -\frac{d}{dx}(f^o_{pp}\phi' + f^o_{zp'}\phi) + f^o_{zz}\phi + f^o_{zp'}\phi' = 0 \tag{3.42}$$

*is called the Jacobi equation. It is a homogeneous second order system of nonlinear ode's and so has the trivial solution* $\phi \equiv 0$.

For $u_0(x)$ a strong or weak relative minimums of $I(u)$, the accessory variational functional $J(\phi)$ must have minimum value 0. Of course, $\phi \equiv 0$ is a solution and the question is whether there are other nontrivial solutions.

**Proposition 3.22.** *Let* $c \in (a,b]$ *and suppose* $\eta$ *is a solution of the Jacobi equation satisfying* $\eta(a) = \eta(c) = 0$ *Then*

$$\int_a^c F(x,\eta,\eta')\ dx = 0\ .$$

Proof. Since $F$ is homogeneous of degree 2 in $(\eta,\eta')$, we have

$$2F(x,\eta,\eta') = \eta F_\eta(x,\eta,\eta') + \eta' F_{\eta'}(x,\eta,\eta')\ .$$

Thus,

$$\begin{aligned} 2\int_a^c F(x,\eta,\eta')\ dx &= \int_a^c (\eta\ F_\eta + \eta'\ F_{\eta'})\ dx \\ &= \int_a^c \eta(F_\eta - \frac{d}{dx}F_{\eta'})\ dx + \eta F_\eta|_a^c = 0 \end{aligned}$$

since $\eta$ vanishes at the endpoints and satisfies the Jacobi equation.

**Definition 3.23.** *Let $u_0$ be a $C^2$ relative minimum of $I(u)$ and suppose there exists $\eta \in C^2$ a solution of the Jacobi equation (for $u_0$) on $(a,b)$ satisfying $\eta(a) = \eta(c) = 0$ and $\eta$ not identically 0 on any subinterval of $(a,c)$. Then we call $c$ a conjugate point of $a$.*

**Theorem 3.24.** *(Jacobi's necessary condition) Suppose $u_0(x)$ is a $C^2$ relative minimum for $I(u)$ on $[a,b]$ satisfying the regularity condition*

$$\det f_{pp}(c) = \det f_{pp}(c, u_0(c), u_0'(c)) \neq 0 \ .$$

*Then $c$ cannot be conjugate to $a$.*

Proof. If $c$ is conjugate to $a$ then there exists $\eta(x)$ a solution of the Jacobi equation for the extremal $u_0(x)$ on $[a,c]$ vanishing at $x = a$ and $x = c$. Extend $\eta$ to be identically 0 on $[c,b]$ to obtain a piecewise $C^1$ function $\tilde{\eta}$ on $[a,b]$ with possible corner at $x = c$. According to Proposition 3.19 , $J(\tilde{\eta}) = 0$ and so $\tilde{\eta}$ is a minimimizer of J since $u + 0$ is a (strong or weak) relative minimimum. If $c$ is a conjugate point, then $\eta'(c) \neq 0$ for otherwise $\eta \equiv 0$. Hence $\tilde{\eta}$ has a true corner at $c$ and so must satisfy the Erdmann corner condition

$$f_{pp}(c)\tilde{\eta}'(c) = F_{\eta'}(c, \tilde{\eta}(c), \tilde{\eta}'(c^-)) = F_{\eta'}(c, \tilde{\eta}(c), \tilde{\eta}'(c^+)) = 0 \ .$$

By our regularity assumption $\tilde{\eta}'(c) = 0$, a contradiction.

**Example 3.25.** *(geodesics on $S^2$)*

Introduce a parametrization $(x,y)$ on the sphere $S^2$ where $x \in (0, 2\pi)$ is the longitude and $y \in (-\frac{\pi}{2}, \frac{\pi}{2})$ is the latitude and the corresponding point has coordinates

$$X = (\cos x \cos y, \sin x \cos y, \sin y) \ .$$

We consider curves lying near the equator $y_0 \equiv 0$ which can be represented as a graph $y = y(x)$. Then the length functional becomes

$$L(y) = \int_0^b \sqrt{\cos^2 y(x) + (y')^2} \ dx \ .$$

For the geodesic $y_0(x) \equiv 0$ , $F(x, \eta, \eta') = (\eta')^2 - \eta^2$ and the Jacobi equation is $\eta'' + \eta = 0$. The solutions are all of the form

$$\eta = A \cos x + B \sin x \ .$$

The nontrivial solutions with $\eta(0) = 0$ are thus of the form $\eta = B \sin x$. In particular there are no conjugate points $c$ if $b < \pi$ and the first conjugate point is $c = \pi$ when $b \geq \pi$. Thus we see that the great circles cannot be minimizing past the antipodal point and geodesics between any two antipodal points are minimizing but not unique.

## 4. SUFFICIENT CONDITIONS AND FIELDS OF EXTREMALS

In his lectures on the Calculus of Variations, Weierstrass introduced the idea of a field of extremals and proposed a method to prove that a given extremal is a relative minimum using his E function which we have already discussed. His idea is roughly as follows. Let $u_0$ be a given extremal for $I(u)$ on $[a,b]$ and let $u(x)$ be a potential competitor close to $u_0$ with the same endpoints $(a,u_0(a))$ , $(b,u_0(b))$. Suppose further that for each $x \in (a,b]$ there is a unique extremal $\psi(t;x)$ , $a \le t \le x$ joining $(a,u_0(a))$ and $(x,u(x))$ and that $\psi$ is $C^2$ in $t,x$. In particular,

$$\psi(a,x) = u_0(a) \text{ and } \psi(t;b) = u_0(t) \ . \tag{4.1}$$

Define

$$\sigma(x) = -\{\int_a^x f(\psi(t;x)\ dt + \int_x^b f(u(t))\ dt\ .\}$$

Then,

$$\sigma(a) = I(u)\ ,\ \sigma(b) = -I(u_0) \text{ and } I(u) - I(u_0) = \sigma(b) - \sigma(a)\ .$$

Thus we need to investigate

$$\sigma'(x) = f(u(x)) - f(\psi(x;x) - \int_a^x \frac{\partial}{\partial x} f(\psi(t;x))\ dt\ . \tag{4.2}$$

Using that $\frac{\partial}{\partial t} f_p(\psi(t;x)) = f_z(\psi(t;x))$ we compute

$$\frac{\partial}{\partial x} f(t,\psi(t;x),\psi'(t;x)) = f_z\ \psi_x + f_p\ (\psi')_x = \frac{\partial}{\partial t}\{f_p(\psi(t;x))\psi_x\}\ . \tag{4.3}$$

From (4.1) we see that

$$\psi_x(a;x) = 0\ ,\ \psi(x;x) = u(x)\ , \text{ and so } \psi_x(x;x) = u'(x) - \psi'(x;x)\ . \tag{4.4}$$

Combining (4.2),(4.3),(4.4) gives

$$\begin{aligned}\sigma'(x) &= f(x,u,u') - f(x,u,\psi'(x;x)) - f_p(x,u,\psi'(x;x)(u'(x) - \psi'(x;x)) \\ &= E(x,u(x),\psi'(x;x);u'(x))\ .\end{aligned}$$

Thus the condition $E \ge 0$ which is implied by the convexity of $f(x,u,p)$ in $p$ would imply $I(u) \ge I(u_0)$, provided that such a "field of extremals" defined by $\psi(t;x)$ exists. In general it is difficult to insure that $\psi(t;x)$ exists but in simple cases it may be possible.

**Example 4.1.**

Consider $I(u) = \frac{1}{2}\int_0^b ((u')^2 - u^2)\ dx$ with boundary conditions $u(0) = u(b) = 0$. The Euler Lagrange equation is $u'' + u = 0$ and we have already seen that if $b < \pi$ the unique solution is $u_0 \equiv 0$. Now consider

$$\psi(t;x) = \frac{u(x)}{\sin x}\ \sin t\ , \tag{4.5}$$

the unique extremal joining $(0,0)$ to $(x,u(x))$. Then

$$\begin{aligned}
&(4.6) \qquad \sigma(x) &&= -\{\int_0^x ((\psi')^2-\psi^2)\ dt+\int_x^b ((u')^2-u^2)\ dt\} \\
&(4.7) &&= -\{\frac{u^2(x)}{\sin^2 x}\int_0^x ((\psi')^2-\psi^2)\ dt+\int_x^b ((u')^2-u^2)\ dt\} \\
&(4.8) &&= -u^2(x)\cot x-\int_x^b ((u')^2-u^2)\ dt
\end{aligned}$$

Hence,

$$\sigma'(x)=-2uu'\cot x+u^2(x)\csc^2 x+(u')^2(x)-u^2(x)=(u(x)\cot x-u'(x))^2\geq 0\ ,$$

with equality if and only if $u'(x)=u(x)\cot x$. But then $u'(0)=u(0)=0$ and so $u\equiv 0$.

## 5. HAMILTON JACOBI THEORY

5.1. **Lagrangian Mechanics and Stationary Action.** Lagrange observed that Newton's second law of motion $m\ddot{x}=f(t,x)$ describing the motion of a particle of mass m moving under the influence of non-uniform, possibly time-dependent external forces $f(t,x)=-\nabla_x U(t,x)$ (U is called a potential function for f) can be obtained as the Euler- Lagrange equations of the "action integral"

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

where the so-called Lagrangian

$$L(t,x,\dot{x}=T-U=\frac{1}{2}m\dot{x}^2-U(t,x)\ .$$

Here T is the kinetic energy and U is the potential energy of the system. When the system is conservative, i.e L is independent of t, the the total energy $E=T+U$ is conserved.

**Example 5.1.** *(Central forces) Consider a system of N bodies bound by gravitational forces*

$$f_{ij}=m_i m_j\frac{r_i-r_j}{|r_i-r_j|^3}$$

*between any pair of masses $m_i$ and $m_j$, where $r_i=(x_i,y_i,z_i)$ is the position vector in $R^3$ of the mass $m_i$. The corresponding potential for this N mass system is*

$$U=-\frac{1}{2}\sum_{i,j=1}^{N} G\frac{m_i m_j}{|r_i-r_j|}\ ,$$

*where G is Newton's universal gravitational constant. The kinetic energy is given by* $T = \frac{1}{2}\sum_{i=1}^{N} m_i \dot{r_i}^2$ *and* $L = T - U$*. The E-L equations gives a system of N vector equations:*

$$m_i \ddot{r_i} - \sum_{j \neq i} G m_i m_j \frac{r_i - r_j}{|r_i - r_j|^3} = 0$$

**Example 5.2.** *(Spring-mass system-no friction or damping) Consider a system of N masses* $m_i$ *with coordinates* $x_i$ *lying on an axis and joined by springs. Each spring generates a force* $f_i$ *proportional to* $x_i - x_{i+1}$ *where* $x_i^0 - x_{i+1}^0 = l_i$ *corresponds to the resting position of the spring. Then*

$$T = \frac{1}{2}\sum_{i=1}^{N} m_i \dot{x_i}^2 \ ,$$

$$U = \frac{1}{2}k_1(x_2 - x_1)^2 + \ldots + \frac{1}{2}k_{N-1}(x_N - x_{N-1})^2 \ ,$$

*and*

$$L = T - U = \frac{1}{2}\sum_{i=1}^{N}(m_i \dot{x_i}^2 - k_i(x_{i+1} - x_i)^2) \ .$$

*This gives as E-L equations:*

$$m_1 \ddot{x_1} + k_1(x_1 - x_2) = 0$$

$$m_2 \ddot{x_2} + k_2(x_2 - x_3) - k_1(x_1 - x_2) = 0$$

$$\cdot$$

$$\cdot$$

$$m_N \ddot{x_N} - k_{N-1}(x_{N-1} - x_N) = 0 \ .$$

*When all the* $m_i = m$ *and all* $k_i = k$ *the equations simplify considerably.*

However most dynamical systems must satisfy additional constraints (for example consider the motion of a simple pendulum) which makes the application of Lagrange's method impractical. Instead, suppose the state of a constrained system is described by n independent "generalized coordinates" $q_1, \ldots q_n$ corresponding to n "degrees of freedom". The position vector x is determined by a transformation x=x(q) with associated velocity $\dot{x} = \sum \frac{\partial x}{\partial q_i}\dot{q_i}$. The kinetic energy assumes the form

$$T = \frac{1}{2}\sum m_i \dot{x_i}^2 = \sum_{j,k} a_{jk}(q)\dot{q_j}\dot{q_k} \ ,$$

where

$$a_{jk} = \frac{1}{2}\sum_i m_i \frac{\partial x_i}{\partial q_j}\frac{\partial x_i}{\partial q_k} \ ,$$

and the (transformed) potential energy is now $W(q) = U(x(q))$.

Hamilton's principle of **Stationary Action** is now expressed through the action functional on paths from an initial point $A = q(t_0)$ to a final point $B = q(t_1)$,

$$S(q) = \int_{t_0}^{t_1} L(t, q, \dot{q}\ dt$$

with Lagrangian $L(t, q, \dot{q}) = T - U$. It says that the motion should make S stationary with respect to admissible variations so that the $q_j$ should satisfy the E-L equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} = \frac{\partial L}{\partial q_j}\ ,\ j = 1, \ldots n\ .$$

**Example 5.3.** *(simple pendulum-one-degree of freedom)*
*Consider a simple pendulum consisting of a bob of mass* $m$ *suspended from a string of length* $l$ *of negligible mass. A natural choice for the generalized coordinate is the angle* $\theta$ *that the pendulum makes with the (downward) vertical. Then the kinetic energy is given by* $T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) = \frac{1}{2}m\dot{s}^2 = \frac{1}{2}ml^2\dot{\theta}^2$ *and the potential energy* $U = mgl(1 - \cos\theta)$. *The action integral is then*

$$S(\theta) = \int_{t_0}^{t_1} (\frac{1}{2}ml^2 - mgl(1 - \cos\theta)\ dt\ .$$

*Hamilton's principle implies that the equation of motion for the pendulum is given by*

$$-\frac{d}{dt}L_{\dot{\theta}} + L_{\theta} = -ml^2\ddot{\theta} - mgl\sin\theta = 0\ ,$$

*or*

$$\ddot{\theta} + \frac{g}{l}\sin\theta = 0\ .$$

**Example 5.4.** *(planar motion in a central force field).*

*Consider the motion in a plane of a particle of mass* $m$ *affected by a force* $F = -\frac{k}{r^3}\vec{r}$ *where* $\vec{r}$ *is the positon vector of the particle and* $r = |\vec{r}|$ *is the distance to the origin. As generalized coordinates we choose polar coordinates* $(r, \theta)$ *which are conected to Cartesian coordinates by* $x = r\cos\theta\ ,\ y = r\sin\theta$. *Then*

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2\ ,$$

*and the potential energy is* $U = -\frac{k}{r}$. *The Lagrangian is thus*

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2) + \frac{k}{r}\ ,$$

*and the E-L equations of motion are given by*

$$\begin{aligned} -\frac{d}{dt}(L_{\dot{r}}) + L_r &= -\frac{d}{dt}(m\dot{r} + mr\dot{\theta}^2 - \frac{k}{r^2} = 0 \\ -\frac{d}{dt}(L_{\dot{\theta}}) + L_{\theta} &= -\frac{d}{dt}(mr^2\dot{\theta} = 0\ , \end{aligned}$$

*that is,*

$$m\ddot{r} - mr\dot{\theta}^2 + \frac{k}{r^2} = 0 \ ,$$

$$mr^2\dot{\theta} = C_0 \ .$$

*Here $C_0$ is the angular momentum of the system. It follows that the differential of area swept out by the position vector through an angle $d\theta$ is given by $dA = \frac{1}{2}r^2 d\theta$ and so $2m\dot{A} = 2m\frac{1}{2}r^2\dot{\theta} = C_0$. This shows that the rate at which area is swept out is constant, which is Kepler's second law of planetary motion. To derive Kepler's first law (that the planets move in elliptical orbit with the sun at one focus) is harder (we give a brief sketch) and begins by eliminating $\dot{\theta}$ from the second equation which can then be integrated (multiply both sides by $\dot{r}$ to obtain*

$$\dot{r} = \sqrt{C_1 + \frac{2k}{mr} - \frac{C_0^2}{mr^2}} = \sqrt{(C_1 + \frac{k}{mC_0^2}) - \frac{C_0^2}{m}(\frac{1}{r} - \frac{k}{C_0^2})} \ .$$

*The constant $C_1$ is the total energy which is conserved. Now make the substitution*

$$\frac{1}{r} - \frac{k}{C_0^2} = \frac{\sqrt{m}}{C_0}\sqrt{C_1 + \frac{k}{C_0^2}}\cos\theta \ ,$$

*which gives*

$$\dot{r} = \sqrt{C_1 + \frac{k^2}{mC_0^2}}\sin\theta \ .$$

*Then since*

$$\frac{dr}{d\theta} = \frac{\dot{r}}{\dot{\theta}} = \frac{m}{C_0}r^2\dot{r} \ ,$$

*we can integrate this to obtain*

$$\frac{1}{r} = \frac{m}{C_0}\sqrt{C_1 + \frac{k^2}{mC_0^2}}\cos\theta + C_2 \ ,$$

*which can be written in the form*

$$r(\theta) = \frac{r_0}{1 + e\cos\theta}$$

*which is an elliptical orbit (when the parameters are such that the orbit is periodic).*

5.2. **The Legendre transform.** Given a $C^2$ function $f(x_1, \ldots, x_n)$ of n variables consider the change of variables $y = \nabla x$ which is locally 1-1 if $\det f_{x_i x_j} \neq 0$. We associate to f, a new function $\eta(y)$ called the Legendre transform of f (sometimes written $\eta = f^*$) by the relation

$$f(x) + \eta(y) = x \cdot y \ .$$

You can think of $(y, \eta(y))$ as parametrizing the tangent planes to the graph of the function $f(x)$. Diferentiating this relation with respect to $y_j$ using the chain rule gives

$$\sum_i f_{x_i} \frac{\partial x_i}{\partial y_j} + \eta_{y_j} + \sum_i \frac{\partial x_i}{\partial y_j} y_i \ .$$

Since $y_i = f_{x_i}$ this gives $x_j = \eta_{y_j}$ or $x = \nabla_y \eta(y)$. Thus we have the symmetry that $f = \eta^*$. Note also that we have the relation

$$\delta_{ik} = \sum_j f_{x_i x_j} \frac{\partial x_j}{\partial y_k} = \sum_j f_{x_i x_j} \frac{\partial^2 \eta}{\partial y_j \partial y_k} \ .$$

This says that the Hessian of $\eta$ (in the y variables) is the inverse of the Hessian of f (in the x variables). In particular, if f is a $C^2$ convex function, then so is $\eta$.

5.3. **Derivation of Hamilton's equation.** We will use the Legendre tranform of the Lagrangian $L(t, q, \dot{q})$ with respect to the $\dot{q}$ variable to convert the E-L equations (n second order ode's) into Hamiltons equations (2n first order equations). Introduce the generalized momenta

$$p := L_{\dot{q}}(t, q, \dot{q}) \ . \tag{5.1}$$

We assume that for all $(t, q, \dot{q})$ that the equation (5.1) can be uniquely solved for $\dot{q}$ as a smooth (vector) function

$$\dot{q} = G(t, q, p) \ . \tag{5.2}$$

**Definition 5.5.** *The Hamiltonial $H$ associated with the Lagrangian $L$ is given by*

$$H(t, q, p) = p \cdot G(t, q, p) - L(t, q, G(t, q, p)) \tag{5.3}$$

**Theorem 5.6.** *The functions $q(t)$ and $p(t)$ satisfy Hamilton's equations*

$$\dot{q}(t) = H_p(t, q, \dot{q}) \tag{5.4}$$

$$\dot{p}(t) = -H_q(t, q, p) \tag{5.5}$$

Proof. Note that from (5.1), (5.2) and the definition of H,

$$\frac{\partial H}{\partial q_i}(t, q, p) = \sum_k p_k \frac{\partial G^k}{\partial q_i} - \frac{\partial L}{\partial q_i}$$

$$= -\sum_k \frac{\partial L}{\partial \dot{q}_k} \frac{\partial G^k}{\partial q_i} = \frac{\partial L}{\partial q_i}$$

by (5.1). Likewise,

$$\frac{\partial H}{\partial p_i}(t, q, p) = G^i(t, q, p) + \sum_k p_k \frac{\partial G^k}{\partial p_i} - \sum_k \frac{\partial L}{\partial \dot{q}_k}$$

$$= -\sum_k \frac{\partial L}{\partial \dot{q_k}} \frac{\partial G^k}{\partial q_i} = G^i(t,q,p)$$

by (5.1). Thus,

$$\frac{\partial H}{\partial p_i}(t,q(t),p(t)) = G^i(t,q(t),p(t)) = \dot{q(t)}$$

and similarly

$$\frac{\partial H}{\partial q_i}(t,q(t),p(t)) = \frac{\partial L}{\partial q_i}(t,q(t),G(t,q(t),p(t))$$

$$= \frac{\partial L}{\partial q_i}(t,q(t),\dot{q(t)}) = -\frac{d}{dt}(L_{\dot{q_i}}(t,q(t),\dot{q(t)}) = -\dot{p_i(t)} \ ,$$

where in the last steps we have used the E-L equations and (5.1). Finally, observe that if H is independent of t, then

$$\frac{d}{dt}H(q(t),p(t)) = H_{q_i}\dot{q}_i + H_{p_i}\dot{p}_i = H_q \cdot H_p - H_p \cdot H_q = 0 \ ,$$

so H is a constant of the motion.

**Example 5.7.** *(Geometric optics)*
*Consider again the Lagragian $L(x,y,y') = n(x,y)\sqrt{1+y'^2}$ for the problem of geometric optics we have previously considered. Then the dual variable $p$ is given by*

$$p = L_{y'} = \frac{n(x,y)y'}{\sqrt{1+y'^2}}$$

*and we can solve for $y'$:*

$$y' = \frac{p}{\sqrt{n^2-p^2}} \ ,$$

*(which is the first canonical equation) and find the Hamiltonian*

$$H(x,y,p) = py' - L = -\sqrt{n^2-p^2} \ .$$

*Then the second canonical equation is*

$$p' = -H_y = -\frac{n(x,y}{\sqrt{n^2-p^2}}\frac{\partial n}{\partial y} \ .$$

*Since H is independent of x, it is constant along an optimal trajectory. If we denote by $\alpha(x,y)$ the angle of inclination of the optimal path, i.e. $y' = \tan\alpha$, then*

$$p = n(x,y)\sin\alpha \ \textit{and} \ H = -n(x,y)\cos\alpha \ .$$

5.4. **Canonical transformations and Generating functions.** We look for transformations $(q, p) \to (q^*, p^*)$ under which the canonical equations preserve their form, that is we require

$$\dot{q^*} = H^*_{p^*} \ , \ \dot{p^*} = -H^*_{q^*} \tag{5.6}$$

for a new Hamiltonian $H^*(t, q^*, p^*)$. To find such canonical transformations, we use the fact that Hamilton's equation are the E-L equations of the actional functional

$$A(q, p) = \int_{t_0}^{t_1} (p \cdot \dot{q} - H)\ dt \tag{5.7}$$

in which the $q_i$ , $p_i$ are regarded as 2n independent functions. We want the new variables $q^*$ , $p^*$ to satisfy (5.6) which suggests that we write

$$A^*(q^*, p^*) = \int_{t_0}^{t_1} (p^* \cdot \dot{q^*} - H^*)\ dt \tag{5.8}$$

where $(q^*, p^*)$ are functions of $(t, q, p)$ . Thus (5.7) and (5.8) represent two different variational problems involving the same variables and we require that these variational problems be equivalent. It is not difficult to see that this is the case if and only if the integrands of the two functionals differ by a total differential.

**Definition 5.8.** *Let $S(t, q, q^*)$ be an arbitrary smooth function and put $p = \frac{\partial S}{\partial q}$ , $p^* = \frac{\partial S}{\partial q^*}$. We assume that this defines a unique transformation $p = p(t, q^*, p^*)$ , $q = q(t, q^*, p^*)$ The function S is called a generating function.*

So in order that (5.7) and (5.8) be equivalent, we require that

$$p \cdot \dot{q} - H(t, q, p) = p^* \cdot \dot{q^*} - H^*(t, q^*, p^*) + \frac{dS}{dt} \tag{5.9}$$

Since

$$\frac{dS}{dt} = S_t + S_q \dot{q} + S_{q^*} \dot{q^*}$$

equation5.9) becomes

$$(p - S_q) \cdot \dot{q} - (p^* + S_{q^*}) \cdot \dot{q^*} - H + H^* - S_t = 0 \tag{5.10}$$

Therefore we have the

**Theorem 5.9.** *Let $S(t, q, q^*)$ be an arbitrary generating function with $p = \frac{\partial S}{\partial q}$ , $p^* = \frac{\partial S}{\partial q^*}$. Then this defines a canonical transformation with $H^* = H + S_t$.*

We note for later reference that in order to make $H^* \equiv 0$ so that the new canonical equations become trivial, we must choose S in a special way.

**Corollary 5.10.** *$H^* \equiv 0$ if and only if S satisfies the Hamilton-Jacobi equation $S_t + H(t, q, S_q) = 0$ for all values of the parameters $q^*$.*

We recall that we have already seen that the action integral is closely related to the Hamilton-Jacobi equation. Suppose that $A = (t_0, \lambda)$ and $B = (t, q)$ can be connected by a unique extremal (we regard $\lambda$ as parameters and think of $(t, q)$ as the variables in some neighborhood of an endpoint $(t_1, q_1)$.) Then

$$S(t,q) = \int_{t_0}^{t} (p \cdot \dot{q} - H)\ dt$$

evaluated on a unique extremal passing through A and B becomes a single valued function (called the geodesic distance) between A and B. Then formula (3.26) can be rewrittten as

$$dS(t,q) = pdq - Hdt \tag{5.11}$$

or in other words,

$$\frac{\partial S}{\partial t} = -H(t,q,p)\ ,\ \frac{\partial S}{\partial q} = p\ . \tag{5.12}$$

From (5.12) we conclude that S satisfies the Hamilton-Jacobi equation,

$$S_t + H(t,q,S_q) = 0\ . \tag{5.13}$$

### 5.5. **Complete Integrals of the Hamilton-Jacobi equation.**

**Definition 5.11.** *Let $I(t,q,\lambda) - b$ be a $C^2$ solution of the Hamilton Jacobi equation for all values of the n parameters $\lambda$ (perhaps in some region) where b is an additive constant. We call I a complete integral if*

$$\det I_{q\lambda} \neq 0. \tag{5.14}$$

Right away we want to note that

**Theorem 5.12.** *Each derivative $\frac{\partial I}{\partial \lambda_i}$ is a first integral of Hamilton's equations, i.e. $\frac{\partial I}{\partial \lambda_i}$=constant along each extremal.*

Proof. We must show

$$\frac{d}{dt}(\frac{\partial I}{\partial \lambda_i} = 0\ ,\ i = 1, \dots, n. \tag{5.15}$$

Calculating the left side of (5.15) we find

$$\frac{d}{dt}(\frac{\partial I}{\partial \lambda_i} = I_{\lambda_i t} + \sum_{j=1}^{n} I_{\lambda_i q_j} \dot{q}_j\ . \tag{5.16}$$

Now differentiate the equation $I_t = -H(t,q,I_q)$ with respect to $\lambda_i$:

$$I_{t\lambda_i} = -\sum_{j=1}^{n} H_{p_j} I_{q_j \lambda_i} \tag{5.17}$$

Substitution of (5.17) into (5.16) gives

$$\frac{d}{dt}(\frac{\partial I}{\partial \lambda_i} = -\sum_{j=1}^{n} H_{p_j} I_{q_j\lambda_i} + \sum_{j=1}^{n} I_{\lambda_i q_j}\dot{q}_j \tag{5.18}$$

$$= \sum_{j=1}^{n} I_{q_j\lambda_i}(\dot{q}_j - H_{p_j}) = 0 \tag{5.19}$$

since $\dot{q}_j - H_{p_j}) = 0$ along each extremal. This proves the theorem.

We now prove Jacobi's theorem:

**Theorem 5.13.** *Let $I(t, q, \lambda)$ be a complete integral of the Hamilton Jacobi equation and let $\mu = (\mu - 1, \ldots, \mu_n)$ be n arbitrary parameters. Then the functions*

$$q_i = q_i(t, \lambda, \mu) \tag{5.20}$$

*defined (implicitly) as solutions of the envelope equations*

$$I_{\lambda_i} = \mu_i \ , \ i = 1, \ldots, n \tag{5.21}$$

*together with the functions $p_i = I_{q_i}(t, q, \lambda)$ (where the $q_i$ are given by (5.20)) constitute a general solution of Hamilton's equations.*

Proof. We make a canonical transformation by choosing $S = I(t, q, \lambda)$ as the generating function with $\lambda$ as the new momenta and $\mu$ as the new generalized coordinates. Then according to Theorem 5.9 and Corollary 5.10 ,

$$p_i = I_{q_i} \ , \ \mu_i = I_{\lambda_i} \ , \ H^* = H(t, q, p) + I_t = 0 \ .$$

Therefore in the new variables the canonical equations become $\dot{q^*} = \dot{p^*} = 0$ with solution $\mu$ = constant and $\lambda$ = constant along each extremal. Therefore we again obtain the same n first integrals $I_{\lambda_i} = \mu_i$ which we can use to determine the functions (5.20) of the 2n parameters $\lambda, \mu$. As before we set $p = I_q(t, q, \lambda)$ where the $q_i$ are given by (5.20) obtaining a general solution of Hamilton's equations.

**Example 5.14.** *(The two body problem)*
*Consider the problem of determining the motion of two bodies of mass $m_1$ and $m_2$ acted upon by the Newtonian gravitational force. We fix the origin at the center of mass and let $r_i$ , $i = 1, 2$ denote the distance of mass $m_i$ from the origin. Then since $m_1 r_1 = m_2 r_2$,*

$$F = \frac{G m_1 m_2}{(r_1 + r_2)^2} = \frac{G m_1 m_2^3}{(m_1 + m_2)^2} \frac{1}{r_1^2}. \tag{5.22}$$

*Thus the problem is reduced to the motion of a fixed mass at the origin attracting a mass m as in Example 5.4. Thus in polar coordinates $(r, \theta)$,*

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2) + \frac{k}{r} \ , \ H = \frac{m}{2}(p_r^2 + \frac{1}{r^2}p_\theta^2) - \frac{k}{r} \tag{5.23}$$

*The Hamilton-Jacobi equation is then*

$$I_t + \frac{m}{2}(I_r^2 + \frac{1}{r^2}I_\theta^2) - \frac{k}{r} = 0 \tag{5.24}$$

*and we can find a complete integral in the form* $I = \alpha t + \beta\theta + f(r)$. *We find*

$$I = \alpha t + \beta\theta + \int_{r_0}^{r} \sqrt{\frac{2k}{m\rho} - \frac{\beta^2}{\rho^2} - \frac{2\alpha}{m}}\ d\rho \tag{5.25}$$

*To solve for the extremals, we differentiate with respect to the parameters:*

$$t \quad - \quad \int_{r_0}^{r} \frac{d\rho}{\sqrt{\frac{2k}{m\rho} - \frac{\beta^2}{\rho^2} - \frac{2\alpha}{m}}} = t_0 \tag{5.26}$$

$$\theta \quad - \quad \beta\int_{r_0}^{r} \frac{d\rho}{\rho^2\sqrt{\frac{2k}{m\rho} - \frac{\beta^2}{\rho^2} - \frac{2\alpha}{m}}} = \theta_0 \tag{5.27}$$

*where* $t_0$ *and* $\theta_0$ *are arbitrary constants. The second equation (5.27) gives the path of the particle while the first (5.27) gives* $r$ *as a function of time. Using the substitution* $\rho = \frac{1}{\sigma}$ *we can explicitly integrate (5.27) obtaining*

$$r = \frac{\delta}{1 - e\sin(\theta - \theta_0)}$$

*where* $\delta = \frac{m\beta^2}{k}$ *and* $e = \sqrt{1 - \frac{2\alpha}{k^2}m\beta^2}$.