On the necessary and sufficient conditions for minima in variational principles

Xinlin Wu
-Nanjing University
Instructor: Joao Rodrigues
-Cambridge University

July 30, 2022

Abstract: In this text we discuss the necessary and sufficent conditions for minima in variational principles. The first part focuses on the necessary conditions and the second on the sufficent conditions. For necessary conditions, we discuss the Euler condition, the Weierstrass condition, the Legendre condition and the Jacobi condition. For the second part, we first introduce the concept of fields. Secondly we give the fundamental sufficiency theorem and sufficent cobinations of conditions. Next, we use convexity to judge whether certain function is the global minimum point. Finally, in the section 2.6, we close the whole text with some examples.

Firstly we define $\mathcal{Y} = \{y$ is piecewise smooth(pws) on $[t_0, t_1] \big| y(t_0)a = a, y(t_1) = b\}$($a$ and b are to be determined), $\mathcal{H} = \{\eta$ is piecewise smooth on $[t_0, t_1]$ but $\eta(t_0) = \eta(t_1) = 0\}$ and ssume $f(t, y, r) \in C^\infty(\mathbb{R}^3)$. In many cases, one problem may be reduced to considering a functional in the from $J(y) = \int_{t_0}^{t_1} f(t, y, \dot{y})dt$ and in order to solve the problem we often need to answer the following questions:

(1) Does there exist $y_0 \in \mathcal{Y}$ s.t. $J(y) \leq J(y), \forall y \in \mathcal{Y}$?

(2) If so, is the minimizing y_0 unique?

(3) How can all such y_0 be determined?

Remark: Note that $J(y_0)$ is a maximum iff $-J(y_0)$ is a minimum. Hence, we do not give a separate discussion of maxima. Secondly, for most cases we will assume y is smooth if no pws spcified.

The following two sections focus on the necessary and sufficient conditions of being an extreme point. When a point involves $[t_0, t_1]$ and the $\dot{y}$, then at any interior point where the derivative of the pws function fails to exist , the stated condition is understood to hold with $\dot{y}$ interpreted as either $\dot{y}^-(t)$ or $\dot{y}^+(t)$. We will present examples after the discussion of the two topics so that we can analysis them more explicitly.

1 Necessary Conditions

1.1 Different kinds of minima of $J(y)$

Def 1.1.1(Global minimum). If $J(y_0) \leq J(y), \forall y \in \mathcal{Y}$, then we call y_0 is the global minumum point of $J(y)$. Before we define another kinds of minima, we should construct two topoloties.

The first one is the topology induced by the norm $||\cdot||_{C[t_0,t_1]}$, where $||f - g||_{C[t_0,t_1]} = \sup_t |f(t) - g(t)|, f, g \in C[t_0, t_1]$ and we use $B_0(x, \delta)$ to denote a neighborhood w.r.t. this topology, i.e. $B_0(x, \delta) = \big\{y \in \mathcal{Y} \big| ||y - x||_{C[t_0,t_1]} < \delta\big\}$ $\forall x \in \mathcal{Y}, \forall \delta > 0$.

Similarly, The second one is the topology induced by the norm $||\cdot||_{C^1[t_0,t_1]}$, where $||f - g||_{C^1[t_0,t_1]} = ||f - g||_{C[t_0,t_1]} + ||f' - g'||_{C[t_0,t_1]}, f, g \in C^1[t_0, t_1]$ and we use $B_1(x, \delta)$ to denote a neighborhood w.r.t. this topology, i.e. $B_1(x, \delta) = \big\{y \in \mathcal{Y} \big| ||y - x||_{C^1[t_0,t_1]} < \delta\big\}$ $\forall x \in \mathcal{Y}, \forall \delta > 0$.

Remark: Note that y can be pws(piecewise smooth), although usually it is smooth, then it will have corners. In this case we just need to change the norm

$||\cdot||_C$ to $||\cdot||_\infty$.

Def 1.1.2(Strong local minimum) If $\exists\delta>0$ s.t. $J(y_0)\le J(y), \forall y\in B_0(y_0,\delta)$, then $J(y_0)$ is called a strong local minimum.

Def 1.1.3(Weak local minimum) If $\exists\delta>0$ s.t. $J(y_0)\le J(y), \forall y\in B_1(y_0,\delta)$, then $J(y_0)$ is called a weak local minimum.

Example 1.1.4 Suppose $J(y)=\int_0^1(\dot{y}^2+\dot{y}^3)dx$. $\mathcal{Y}=\{$y is pws on [0,1]$|$ $y(0)=y(1)=0\}$. Claim: $y_0=0$ is a weak local minimum point. We set $\delta=\frac{1}{2}$. If $||y||_{C^1}<\delta$,

$$J(y)-J(y_0)=\int_0^1(\dot{y}^2+\dot{y}^3)dx\ge\int_0^1(\dot{y}^2-\frac{1}{2}\dot{y}^2)dx=\frac{1}{2}\int_0^1\dot{y}^2dx\ge 0$$

Hence, $y_0=0$ is a weak local minimum point. However, we now show that it is not a strong local minimum point. In fact, $\forall 0<h<1-h^2$, define

$$y(x,h)=\begin{cases}-\frac{x}{h}, & x\in[0,h^2]\\ \frac{h(x-1)}{1-h^2}, & x\in[h^2,1]\end{cases}$$

then,

$$(\dot{y}^2+\dot{y}^3)(x)=\begin{cases}\frac{1}{h^2}-\frac{1}{h^3}\le-\frac{1}{2h^3}, & x\in[0,h^2]\\ \left(\frac{h}{1-h^2}\right)^2+\left(\frac{h}{1-h^2}\right)^3\le 2, & x\in[h^2,1]\end{cases}$$

Note that $||y||_{C[0,1]}\le h$. However,

$$\begin{aligned}J(y)-J(y_0)=J(y)&=\int_0^h\left[\frac{1}{h^2}-\frac{1}{h^3}\right]dx+\int_{h^2}^1\left[\left(\frac{h}{1-h^2}\right)^2+\left(\frac{h}{1-h^2}\right)^3\right]dx.\\ &\le\int_0^{h^2}-\frac{1}{2h^3}dx+\int_{h^2}^1 2dx\le 2-\frac{1}{2h}\to-\infty(h\to 0)\end{aligned}$$

Thus, y_0 is not a strong local minimum point.

Remark: global minimum $\Rightarrow$ strong local minimum $\Rightarrow$ weak local minimum. This is because we need to check all $y\in\mathcal{Y}$ when we want to comfirm certain $y_0\in\mathcal{Y}$ is a global minimum, but when it comes to strong or weak local minimum, we will do less tests. In other words, from left to right, one topology is stronger than another respectively.

1.2 The Euler necessary condition

Lemma 1.2.1(du Bois-Reymond) If $f \in C[t_0, t_1]$ and $\int_{t_0}^{t_1} f(t) \cdot \phi(t)dt = 0$, $\forall \phi \in \mathcal{Y}$, then f is a constant.
Proof: let $C = \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} f(t)$ and $\phi(t) = \int_{t_0}^{t}(f(s) - C)ds$, then $\phi(t_0) = \phi(t_1) = 0$ and hence $\in \mathcal{Y}$. By hypothesis,

$$0 = \int_{t_0}^{t_1} f(t) \cdot \phi(t)dt = \int_{t_0}^{t_1} f(t)(f(t) - C)dt = \int_{t_0}^{t_1} (f(t) - C)^2 dt \qquad (1)$$

The last step follows from $\int C^2 = \int C \cdot f(t)$. Since f is continous and we must have $f - C = 0$, i.e. $f = C$.

$\square$

Remark: We point out that if f is just piecewise smooth, then the conclusion still holds for all t except the corners.

Theorem 1.2.2 If y_0 is a weak local minimum point, then, $\exists C$-constant s.t.

$$\frac{\partial f}{\partial \dot{y}}(t, y_0(t), \dot{y_0}(t)) = \int_{t_0}^{t} \frac{\partial f}{\partial y}(s, y_0(s), \dot{y_0}(s))ds + C, \forall t \in [t_0, t_1] \qquad (2)$$

Proof: $\forall \eta \in \mathcal{H}$, $\forall |\epsilon| > 0$, $y_0 + \epsilon\eta \in \mathcal{Y}$. Note that $y_0 + \epsilon\eta \to y_0$ in $C^1[t_0, t_1]$. Then, $y_0 + \epsilon\eta \in B_1(y_0, \delta)$ provided $|\epsilon|$ is sufficiently small. Define $F : \mathbb{R} \to \mathbb{R}$

$$F(\epsilon) := \int_{t_0}^{t_1} f(t, y_0 + \epsilon\eta, \dot{y_0} + \epsilon\dot{\eta})dt \qquad (3)$$

By the definition of weak local minimum, F has a local minimum at $\epsilon = 0$. (Here the minimum is in the sense of ordinary functions, not the counterpart of functionals defined above). Then we have $F'(0) = 0$. Also,

$$
\begin{aligned}
F'(\epsilon) &= \int_{t_0}^{t_1} \frac{d}{d\epsilon} f(t, y_0 + \epsilon\eta, \dot{y_0} + \epsilon\dot{\eta}) dt = \int_{t_0}^{t_1} (\frac{\partial f}{\partial y}\eta + \frac{\partial f}{\partial \dot{y}}\dot{\eta}) dt \\
&= \int_{t_0}^{t_1} \frac{\partial f}{\partial \dot{y}}\dot{\eta} dt + \int_{t_0}^{t_1} (\frac{d}{ds}\int_{t_0}^{t} \frac{\partial f}{\partial y}(s) ds)\eta dt \\
&\xlongequal[\text{by part}]{\text{integrate}} \int_{t_0}^{t_1} \frac{\partial f}{\partial \dot{y}}\dot{\eta} dt + [\eta \int_{t_0}^{t} \frac{\partial f}{\partial y}(s)]\Big|_{t_0}^{t_1} - \int_{t_0}^{t_1} \dot{\eta} \int_{t_0}^{t} \frac{\partial f}{\partial y}(s) ds \\
&\xlongequal{\eta(t_0)=\eta(t_1)=0} \int_{t_0}^{t_1} [\frac{\partial f}{\partial \dot{y}} - \int_{t_0}^{t} \frac{\partial f}{\partial y}(s) ds]\dot{\eta} dt \\
&= \int_{t_0}^{t_1} \Big[\frac{\partial f}{\partial \dot{y}}(t, y_0 + \epsilon\eta, \dot{y_0} + \epsilon\dot{\eta}) - \int_{t_0}^{t} \frac{\partial f}{\partial y}(s, y_0, \dot{y_0}) ds\Big]\dot{\eta} dt = 0 \\
&\Rightarrow \frac{\partial f}{\partial \dot{y}}(t, y_0, \dot{y_0}) = \int_{t_0}^{t} \frac{\partial f}{\partial y}(s, y_0(s), \dot{y_0}(s)) ds + C
\end{aligned}
$$

□

Theorem 1.2.3(The Euler-Lagrange equation) If y_0 is a weak local minimum point , then $\frac{\partial f}{\partial y} = \frac{d}{dt}\frac{\partial f}{\partial \dot{y}}$

Proof: Like above,

$$
\begin{aligned}
F'(\epsilon) &= \int_{t_0}^{t_1} (\frac{\partial f}{\partial y}\eta + \frac{\partial f}{\partial \dot{y}}\dot{\eta}) dt \xlongequal[\text{by part}]{\text{integrate}} \int_{t_0}^{t_1} \frac{\partial f}{\partial y}\eta dt + \eta\frac{\partial f}{\partial \dot{y}}\Big|_{t_0}^{t_1} - \int_{t_0}^{t_1} \frac{d}{dt}\frac{\partial f}{\partial \dot{y}}\eta dt \\
&\xlongequal{\eta(t_0)=\eta(t_1)=0} \int_{t_0}^{t_1} \Big[\frac{\partial f}{\partial y}(t, y_0 + \epsilon\eta, \dot{y_0} + \epsilon\dot{\eta}) - \frac{d}{dt}\frac{\partial f}{\partial \dot{y}}(t, y_0 + \epsilon\eta, \dot{y_0} + \epsilon\dot{\eta})\Big]\eta dt = 0
\end{aligned}
$$

By $F'(0) = 0$ and Lemma 1.2.1 we obtain the Euler-Lagrange equation $\frac{\partial f}{\partial y} = \frac{d}{dt}\frac{\partial f}{\partial \dot{y}}$.

□

Remark: In many cases, we will just write down the Euler-Lagrange equations and try to solve them if one can know for sure that there must be a minimum for the problem. In fact, sometimes these equations are solvable and only have a few solutions so that we can just plug in the solutions to verify if one of them exmtremizes $J(y)$.

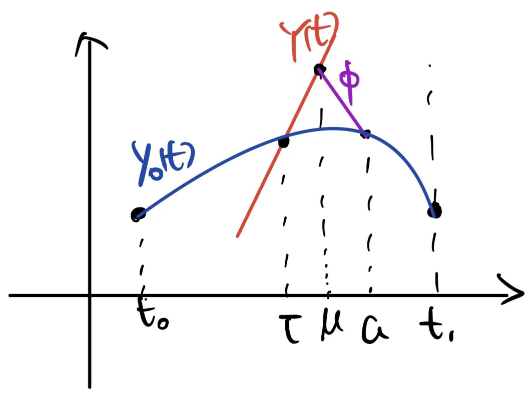
1.3 The Weierstrass necessary condition

$E(t, y, r, q) := f(t, y, q) - f(t, y, r) - (q - r)f_r(t, y, r)$ is called the Weierstrass excess-function or E-function. It will appear again in the part of sufficient conditions.

Theorem 1.3.1 If $y_0 \in \mathcal{Y}$ is a strong local minimum point, then $E(t, y_0, \dot{y_0}, q) \geq 0, \forall t \in [t_0, t_1], \forall q \in \mathbb{R}$.
Proof: Fix $q \in \mathbb{R}, \tau \in [t_0, t_1[$. Choose $a \in]\tau, t_1]$. Define $Y(t) = y_0(\tau) + q(t - \tau)$ and for $u \in [\tau, a[$, define:

$$\phi(t, u) := y_0(t) + \frac{Y(u) - y_0(u)}{a - u}(a - t)$$

$$y(t) := \begin{cases} y_0(t), & t \in [t_0, \tau] \cup [a, t_1] \\ Y(t), & t \in [\tau, u] \\ \phi(t, u), & t \in [u, a] \end{cases}$$



Note that $y(t)$ coincides with $y_0(t)$ except on $[\tau, a]$ with one segment of the curve replaced by two segments of straight lines. Also, y is only a piecewise smooth function, which does not belong to $\mathcal{Y}$. However, we can expand the class $\mathcal{Y}$ to a bigger one with pws functions included and the conclusions for our all discussion will be the same. This is because a pws function can be approximated by smooth functions.

Define

$$
\begin{aligned}
\Phi(u) := J(y) - J(y_0) &= \int_\tau^u [f(t, Y, \dot{Y}) - f(t, y_0, \dot{y_0})]dt \\
&+ \int_u^a [f(t, \phi(t,u), \phi_t(t,u)) - f(t, y_0, \dot{y_0})]dt.
\end{aligned}
$$

$$
\begin{aligned}
\text{Then,} \Phi'(u) &= f(y, Y(u), \dot{Y}(u)) - f(u, y_0(u), \dot{y_0}(u)) + \int_u^a \Big[f_y(t, \phi, \phi_t)\phi_u + f_r(t, \phi, \phi_t)\phi_{tu} \Big] dt \\
&- f(u, \phi(u,u), \phi_t(u,u)) + f(u, y_0(u, \dot{y_0}(u))) \\
&= f(u, Y(u, \dot{Y}(u))) - f(u, \phi(u,u), \phi_t(u,u)) + \int_u^a \Big[f_y \cdot \phi_u + f_r \cdot \phi_{tu} \Big] dt
\end{aligned} \tag{4}
$$

In fact, we can choose a to be sufficiently small so that $||y_0 - y||_{C[t_0,t_1]} < \delta$. Since y_0 is a strong local minimum point and note that $y = y_0$ if $u = \tau$, then $\Phi(\tau)$ is a local minimum. Hence, $\Phi'(\tau) \geq 0$ i.e. $f(\tau, Y(\tau), \dot{Y}(\tau)) - f(\tau, \phi_(\tau, \tau), \phi_t(\tau, \tau)) + \int_\tau^a (f_y \phi_u + f_r \phi_{tu}) dt \geq 0$. Note that :

$$
\begin{aligned}
Y(\tau) &= y_0(\tau) + q(\tau - \tau) = y_0(\tau) \\
\phi(\tau, \tau) &= y_0(\tau) + Y(\tau) - y_0(\tau) = y_0(\tau) \\
\phi_t(\tau, \tau) &= \dot{y_0}(\tau) - \frac{Y(\tau) - y_0(\tau)}{a - \tau} \\
\phi_u(t, u) &= \frac{[\dot{Y}(u) - \dot{y_0}(u)](a - u) + Y(u) - y_0(u)}{(a-u)^2}(a - t) \Rightarrow \phi_u(\tau, \tau) = \dot{Y}(\tau) - \dot{y_0}(\tau)
\end{aligned}
$$

Integrate by parts w.r.t. the second term of the integrand of (4) and use the E-L equation($\frac{d}{dt} f_r = f_y$), set $u = \tau$ in the equation (4) and combine the equations above we can get $E(\tau, y_0(\tau), \dot{y_0}(\tau), \dot{Y}(\tau)) \geq 0$. i.e. $E(\tau, y_0, \dot{y_0}, q) \geq 0$. Note that τ and q are both arbitrary. Hence, by the continuity of E we can know that the inequality holds for all $t \in [t_0, t_1]$ and $q \in \mathbb{R}$.

□

Remark: Here the assumption of strong local minimum is important. Note q is arbitraty, which implies that the slope of $Y(t)$ can be choosen as very large and then $\Phi(u)$ may fail to have a local minimum at $\Phi(\tau)$. This works as follows: Suppose that y_0' is bounded, say, $|y_0'(t)| \leq C, \forall t \in [t_0, t_1]$. If we choose q s.t. $q > C + 1$, then $Y' = q$ and hence $||y_0 - y||_{C^1} \geq 1$ no matter how small the a is. That is to say, in this case we can not conclude that $\Phi = J(y) - J(y_0) \geq 0$. In other words, the conclusion stated in the theorem may only hold for some

q but not all q if y_0 just furnishes a weak local minimum. However, if we can bounding q to a smaller region then we can conclude the same results.It states as follows:

Theorem 1.3.2 If $y_0 \in \mathcal{Y}$ is a weak local minimum point, then $E(t, y_0, \dot{y_0}, q_t) \geq 0, \forall t \in [t_0, t_1]$, for q_t s.t. $|q_t - \dot{y_0}(t)| < \delta$

This theorem tells us if y_0 is just a weak local minimum point, then we can only state that for each t, there are only some q that make the conclusion holds.

1.4 The Legendre necessary condition

Theorem 1.4.1 If $y_0 \in \mathcal{Y}$ is a local minimum point, then $\frac{\partial^2 f}{\partial \dot{y}^2}(t, y_0, \dot{y_0}) \geq 0$, $\forall t \in [t_0, t_1]$.

Proof: we set $F(q) := f(t, y, q)$ with t, y fixed. Applying Taylor's formula we get,

$$F(q) = F(\dot{y}) + F'(\dot{y})(q - \dot{y}) + F''(\xi)\frac{(q - \dot{y})^2}{2}$$

where $\xi = \dot{y} + \theta(q - \dot{y})$ $0 < \theta < 1$. Then,

$$f(t, y, q) = f(t, y, \dot{y}) + \frac{\partial f}{\partial \dot{y}}(t, y, \dot{y})(q - \dot{y}) + \frac{\partial^2 f}{\partial \dot{y}^2}(t, y, \xi)\frac{(q - \dot{y}^2)}{2} \tag{5}$$

By the definition of E-function,

$$E(t, y, \dot{y}, q) = f(t, y, q) - f(t, y, \dot{y}) - (q - \dot{y})\frac{\partial f}{\partial \dot{y}}(t, y, \dot{y}) \overset{(5)}{=\!=} \frac{\partial^2 f}{\partial \dot{y}^2}(t, y, \xi)\frac{(q - \dot{y})^2}{2}$$

Now we argue by contradiction. If $\frac{\partial^2 f}{\partial \dot{y}^2}(t, y_0, \dot{y_0}) \geq 0$ does not hold, then $\exists t' \in [t_0, t_1]$ s.t. $\frac{\partial^2 f}{\partial \dot{y}^2}(t', y_0(t'), \dot{y_0}(t')) < 0$. By the continuity of $\frac{\partial^2 f}{\partial \dot{y}^2}$, for q sufficiently close to $\dot{y_0}(t')$, we still have $\frac{\partial^2 f}{\partial \dot{y}^2}(t', y_0(t'), \xi) < 0$. Then, $E(t', y(t'), \dot{y}(t'), q) < 0<$ which contradicts to Thm 1.3.2.

□

It is not hard to see that, in fact, the Legendre necessary condition is a weaker version of the Weierstrass necessary condition.

1.5 The Jacobi necessary condition

We continue to consider function $F(\epsilon) = \int_{t_0}^{t_1} f(t, y_0 + \epsilon\eta, \dot{y_0} + \epsilon\dot{\eta})dt$. Here,again, y_0 is a weak local minimum point. Then, $F(0)$ is a local minimum. Thus, $F'(0) = 0, F''(0) \geq 0$. In the preceding section, we only use the first condition and now we turn to the second one. Firstly,

$$\begin{aligned} F''(\epsilon) &= \int_{t_0}^{t_1} (\frac{\partial^2 f}{\partial y^2}\eta^2 + 2\frac{\partial^2 f}{\partial y \partial \dot{y}}\eta\dot{\eta} + \frac{\partial^2 f}{\partial \dot{y}^2}\dot{\eta}^2)dt \\ \Rightarrow F''(0) &= \int_{t_0}^{t_1} [\frac{\partial^2 f}{\partial y^2}(t, y_0, \dot{y_0})\eta^2 + 2\frac{\partial^2 f}{\partial y \partial \dot{y}}\eta\dot{\eta} + \frac{\partial^2 f}{\partial \dot{y}^2}(t, y_0, \dot{y_0})\dot{\eta}^2]dt =: I(\eta) \end{aligned}$$

We write the integrand of $I(\eta)$ by $2\omega(t, \eta, \dot{\eta})$. Thus, $I(\eta) = \int_{t_0}^{t_1} 2\omega(t, \eta, \dot{\eta})dt \geq 0$, $\forall \eta \in \mathcal{H}$. In particular, $I(\eta) = 0$ if $\eta \equiv 0$. Consequently, $I(\eta)$ has the minimum value zero on $\mathcal{H}$. The problem $I(\eta) =$ minimum on $\mathcal{H}$, which is similar to what we originally put forward , is called the accessory minimum problem. The E-L equation of the problem is

$$\omega_\eta = \frac{d}{dt}\omega_{\dot{\eta}} = \omega_{\dot{\eta}t} + \omega_{\dot{\eta}\eta} + \omega_{\dot{\eta}\dot{\eta}}\ddot{\eta} \tag{6}$$

Now introduce two definitions w.r.t. to this E-L equation.

Def 1.5.1 (6) is called the Jacobi equation and any (C^2) solution of this equation is called a Jacobi field.

Def 1.5.2 If u is a jacobi field statisfying $u(t_0) = u(t_2) = 0(t_2 \in]t_0, t_1[)$ but $u(t) \neq 0$ for $t_0 < t < t_2$, then t_2 is said to be a conjugate value to t_0 and $(t_2, y_0(t_2))$ is called a conjugate point to the initial point $(t_0, y_0(t_0))$.

Example 1.5.3 Consider the problem: find the shortest curve that connects $(0, 0)$ and $(1, 1)$. This problem can be reduced to finding the y_0 minimizng the functional $J(y) = \int_0^1 \sqrt{1 + \dot{y}^2}dx$. Set $f(x, y, \dot{y}) = \sqrt{1 + \dot{y}^2}$. We already know that $y_0 = x$ is the extremal curve. Then

$$\omega(x, \eta, \dot{\eta}) = \frac{1}{2}\Big[\frac{\partial^2 f}{\partial y^2}(x, y_0, \dot{y_0})\eta^2 + 2\frac{\partial^2 f}{\partial y \partial \dot{y}}\eta\dot{\eta} + \frac{\partial^2 f}{\partial \dot{y}^2}(x, y_0, \dot{y_0})\dot{\eta}^2\Big] = A\dot{\eta}^2$$

Where A is a constant. Hence, $\omega_{\dot{\eta}} = 2A\dot{\eta}$. The E-L equation w.r.t. $I(\eta)$ is $0 = \frac{d}{dt}\omega_{\dot{\eta}}$ i.e. $2A\dot{\eta} = B$. If u is a solution of this differential equation then

$u = Cx$. Furthermore, if u satisfies $u(0) = u(1) = 0$ then $C = 0$. i.e. $u \equiv 0$. Hence, there is no conjugate point in this problem.

In fact, plus certain positive definiteness that many practical problems hold, we can see that the existence of conjugate points is also a necessary condition.

Theorem 1.5.4 If y_0 is a weak local minimum point and if $\frac{\partial^2 f}{\partial \dot{y}^2}(t, y_0, \dot{y_0}) > 0$, then there is no conjugate point.

Before proving this Thm, two Lemmas are needed.

Lemma 1.5.5 If $J(y_0)$ is a weak local minimum and t_2 is a corner of y_0, then

$$\frac{\partial f}{\partial \dot{y}}(t_2, y_0(t_0), \dot{y_0}^-(t_2)) = \frac{\partial f}{\partial \dot{y}}(t_2, y_0(t_0), \dot{y_0}^+(t_2))$$

Proof: According to Thm 1.2.2. $\forall t \in [t_0, t_1]$

$$\frac{\partial f}{\partial \dot{y}}(t, y_0(t), \dot{y_0}(t)) = \int_{t_0}^{t} \frac{\partial f}{\partial y}(s, y_0(s), \dot{y_0}(s))ds + C \tag{7}$$

where C is a constant. Since f is a smooth function, then f_r is bounded on $[t_0, t_1]$. That is to say, the integral $\int_{t_0}^{t} \frac{\partial f}{\partial y}(s, y_0(s), \dot{y_0}(s))ds$ is continous in t. Now suppoose t_2 is a corner. We let $t \uparrow t_2^-$ and $t \downarrow t_2^+$. On the one hand, since $\frac{\partial f}{\partial \dot{y}}$ is continous, then

$$\lim_{t \uparrow t_2^-} \frac{\partial f}{\partial \dot{y}}(t, y_0(t), \dot{y_0}(t)) = \frac{\partial f}{\partial \dot{y}}(t_2, y_0(t_2), \dot{y_0}^-(t_2))$$
$$\lim_{t \downarrow t_2^-} \frac{\partial f}{\partial \dot{y}}(t, y_0(t), \dot{y_0}(t)) = \frac{\partial f}{\partial \dot{y}}(t_2, y_0(t_2), \dot{y_0}^+(t_2))$$

Since the integral is continous in t, then

$$\lim_{t \uparrow t_2^-} \int_{t_0}^{t} \frac{\partial f}{\partial y}(s, y_0(s), \dot{y_0}(s))ds = \lim_{t \downarrow t_2^+} \int_{t_0}^{t} \frac{\partial f}{\partial y}(s, y_0(s), \dot{y_0}(s))ds$$

Combine equation (7) we can get what we want.

$\square$

Lemma 1.5.6 If $f_{rr}(t, y, r) \neq 0$, $\forall t \in]t_0, t_1], \forall y, r \in \mathbb{R}$, then no function y_0 having a corner can minimize $J(y)$.

Proof: Suppose $J(y_0)$ is a weak local minimum and $(t, y_0(t))$ is a corner. Set $p = \dot{y_0}^-(t)$ and $q = \dot{y_0}^+(t)$. By the preceding lemma, $f_r(t, y_0(t), p) = f_r(t, y_0(t), q)$.

By the mean value Thm, $f_{rr}(t, y_0(t), \xi)(p-q) = 0$. Since the term $f_{rr} \neq 0$, we must have $p = q$. Contradiction.

$\square$

Now we turn to Thm 1.5.4

Proof of Theorem 1.5.4: Argue by contradiction. Suppose there is a value t_2 conjugate to t_0. Then there exists a Jacobi field $u(t)$ vanishing at t_0 and t_2 but nowhere between. Define

$$\eta_0 \in \mathcal{H}, \quad \eta_0(t) := \begin{cases} u(t), & t_0 \leq t \leq t_2 \\ 0, & t_2 < t \leq t_1 \end{cases}$$

Since $u(t)$ solves $\frac{d}{dt}\omega_{\dot{\eta}} = \omega_\eta$, $\eta_0(t)$ again satisfies

$$\omega_\eta(t, \eta_0, \dot{\eta}) = \frac{d}{dt}\omega_{\dot{\eta}}(t, \eta_0, \dot{\eta_0}) \tag{8}$$

on each half interval $[t_0, t_2[\ ,\ ]t_2, t_1[$. A simple fact is $2\omega = \eta\omega_\eta + \dot{\eta}\omega_{\dot{\eta}}$. Plug in (8) we get $2\omega = \eta\frac{d}{dt}\omega_{\dot{\eta}} + \dot{\eta}\omega_{\dot{\eta}} = \frac{d}{dt}(\eta\omega_{\dot{\eta}})$, $t \in [t_0, t_2[\cup]t_2, t_1]$. Therefore,

$$I(\eta_0) = \int_{t_0}^{t_1} 2\omega dt = \int_{t_0}^{t_2} \frac{d}{dt}(\eta\omega_{\dot{\eta}})dt + \int_{t_2}^{t_1} 2\omega(t, \eta_0, \dot{\eta_0})dt = \eta\omega_{\dot{\eta}}\Big|_{t_0}^{t_1} = 0$$

Thus, $I(\eta_0) = 0$, i.e. η_0 minimizing $I(\eta)$ on $\mathcal{H}$. That is to say, η_0 must be one solution of the E-L equation $\omega_\eta = \omega_{\dot{\eta}t} + \omega_{\dot{\eta}\eta}\dot{\eta} + \omega_{\dot{\eta}\dot{\eta}}\ddot{\eta}$. We find by differentiation of ω that $\omega\dot{\eta}\dot{\eta} = \frac{\partial^2 f}{\partial \dot{y}^2}(t, y_0, \dot{y_0})$, which does not equal to 0 by assumption. Thus, the E-L equation is a second order of nonautomous differential equation. If η_0 is a smooth function, then we must have $\dot{\eta}_0(t_2) = \dot{\eta}_0^+(t_2) = 0$. Then we have the initial data $\eta_0(t_1) = 0$ and $\dot{\eta_0}(t_2) = 0$. By the standard existing theorem, the secong order differential equation has a unique solution. However, clearly, $\eta \equiv 0$ soloves this equation coinciding with the initail data. Consequently, $\eta_0 \equiv 0$. This contradicts to the definition of η_0. Therefore, η_0 cannot be smooth. In other words, it is pws so that it has corners. Combine $\frac{\partial^2 f}{\partial \dot{y}^2} > 0$ and Lemma 1.5.6 we find the contradiction.

$\square$

2 Sufficient conditions

2.1 Notations for conditions

Now we turn to discussing the sufficient conditions. In fact, the necessary conditions appeared above will sometimes also be the sufficient conditions. Hence, for simplicity, it is quiet convenient to refer these necessary conditions together with the strengthened forms by some notations.

(i):The Euler condition.

(ii):The Wererstrass condition.

(ii)$'$: $E(t, y_0, \dot{y_0}(t), q) > 0$, $\forall (t, q)$ s.t. $t \in [t_0, t_1]$ and $q \neq \dot{y_0}(t)$.

(ii)$_N$: $E(t, y, p, q) \geq 0$, $\forall (t, y, p, q)$ s.t. $\in [t_0, t_1], q \in \mathbb{R}$ and s.t. $|y - y_0(t)| < \delta_1$, $|p - \dot{y_0}(t)| < \delta_2$ where $\delta_1, \delta_2 > 0$.

(ii)$_R$: $E(t, y, p, q) \geq 0$, $\forall t \in [t_0, t_1], y, p, q \in \mathbb{R}$.

(ii)$'_N$: Condition (ii)$_N$ with the strict inequality whenever $p \neq q$.

(ii)$'_R$: Condition (ii)$_R$ with the strict inequality whenever $p \neq q$.

(iii):The Legendre condition.

(iii)$'$: $\frac{\partial^2 f}{\partial \dot{y}^2} f(t, y_0, \dot{y_0}) > 0$, $\forall t \in [t_0, t_1]$.

(iii)$_N$: $f_{rr}(t, y, p) \geq 0$, $\forall t \in [t_0, t_1]$, $\forall y, p$ s.t. $|y - y_0(t)| < \delta_1$, $|p - \dot{y_0}(t)| < \delta_2$ where $\delta_1, \delta_2 > 0$.

(iii)$'_N$: Condition (iii)$_N$ with the strict inequality.

(iii)$_R$: $f_{rr}(t, y, p) \geq 0$, $\forall t \in [t_0, t_1]$, $\forall y, p \in \mathbb{R}$.

(iii)$'_R$: Condition (iii)$_R$ with the strict inequality.

(iv): There is no conjugate point.

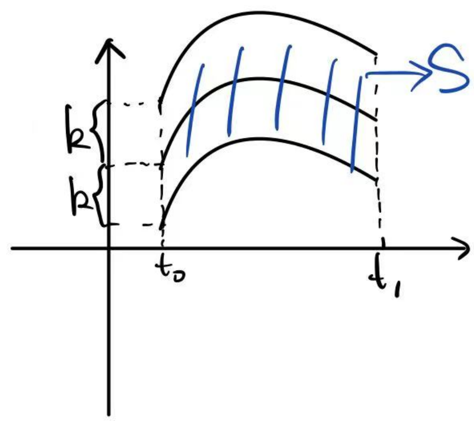
2.2 Fields and the Hilbert integral

Suppose $J(y)$ is the corresponding functional of the fixed endpoints problem. Given a solution y_0 satisfying the E-L equation and a family of a smooth functions $\{\phi(\cdot, \alpha)\}_\alpha$ with the following properties:

(1)$\forall \alpha$, $\phi(t, \alpha)$ solves the E-L equation.

(2)$\exists \alpha_0$ s.t. $\phi(t, \alpha_0) = y_0(t)$.

(3)$\partial_\alpha \phi \neq 0$. The relation $y - \phi(t, \alpha) = 0$ defines implicitly a function $\alpha : S \to \mathbb{R}$ where $S = \{(t, y) : t_0 \leq t \leq t_1, y_0(t) - k < y < y_0(t) + k\}$ where $k \in]0, +\infty]$.



The property (3) can be understood as follows:

Consider $F(t, y, \alpha) := y - \phi(t, \alpha)$. Since $\partial_\alpha \phi \neq 0$ and $F(t, y_0(t), \alpha_0) = 0$, by the implicit function theorem we know $\exists \delta_1 > 0, \delta_2 > 0$ s.t. for any (t', y) with $t' \in B(t, \delta_1)$ and $y \in B(y_0(t), \delta_2)$, there exists a unique α s.t. $F(t', y, \alpha) = 0$ i.e. $y = \phi(t', \alpha)$. But here we require a stronger condition that $B(t, \delta_1) = [t_0, t_1]$.

Define a function $p : S \to \mathbb{R}$, which is called the slope function, by

$$p(t, y) := \phi_t(t, \alpha(t, y)) \tag{9}$$

Since $y = \phi(t, \alpha(t, y))$, $\forall (t, y) \in S$, then

$$p(t, \phi(t, \alpha)) = \phi_t(t, \alpha) \tag{10}$$

Taking partial derivatives w.r.t. t on both sides we obtain:

$$p_t(t, \phi(t, \alpha)) + p_y(t, \phi(t, \alpha))\phi_t(t, \alpha) = \phi_{tt}(t, \alpha) \tag{11}$$

Again, by $y = \phi(t, \alpha)$ and (10) we get

$$p_t(t, y) + p(t, y)p_y(t, \alpha) = \phi_{tt}(t, \alpha(t, y)) \tag{12}$$

Def 2.2.1(Field) The pair (S, p) defined above is called a field $\mathcal{F}$ about y_0 and y_0 is said to be embeded in the field $\mathcal{F}$.

Example 2.2.2: $J(y) = \int_0^1 \sqrt{1+\dot{y}^2}dt$ with endpoints $(0,0)$ and $(1,1)$. Then $y_0(t) = t$ sloves the E-L equation. Define $\phi(t,\alpha) = t+\alpha$. Then $\phi(t,0) = y_0(t)$. The E-L equation of $J(y)$ is $\frac{d}{dt}\frac{\dot{y}}{\sqrt{1+\dot{y}^2}} = 0$. Then, $\phi(t,\alpha)$ solves this equation for any α. Set $y = \phi(t,\alpha)$ then $\alpha = y - t$, which can be defined in the entire strip $S = \{(t,y) : 0 \le t \le 1\}$. Hence, $\{\phi(t,\alpha)\}_\alpha$ defines a field (S,p). On the other hand, the family $\{\alpha t\}_\alpha$ fails to define a field since if we write $y = \alpha t$, then $\alpha = \frac{y}{t}$, which is not defined at the point $(0,0)$.

A natural questions is: under what conditions do there exist a field? We have the following theorem.

Theorem 2.2.3: If y_0 satisfies conditions (i),(iii)$'$ and (iv), then there exists a field $\mathcal{F}$ about y_0.

Now turn to the Hilbert integral

Def 2.2.4(The Hilbert integral) Given a field $\mathcal{F} = (S,p)$. The Hilbert integral w.r.t. to the field is difined by

$$\begin{aligned} J^*(y) :&= \int f(t,y,p(t,y)) + (\dot{y} - p(t,y))f_r(t,y,p(t,y))dt \\ &= \int \big[f(t,y,p(t,y)) - p(t,y)f_r(t,y,p(t,y))\big]dt + f_r(t,y,p(t,y))dy \\ &:= \int Pdt + Qdy \end{aligned}$$

Note $J^*(y)$ is a curvilinear integral and hence independent of the choice of a pws function y through fixed endpoints iff $P_y \equiv Q_t$. This condition applied here can be expressed in the form

$$f_y = f_{rt} + f_{ry}p + f_{rr}p_t + pp_y \tag{13}$$

We show that if $(t, y(t)) \in S$ then this condition will be valid.

Theorem 2.2.5 Given any pws functions y_1 and y_2 in the set S associated with a field $\mathcal{F}$ s.t. $y_1(t)$ and $y_2(t)$ have comman end values, then $J^*(y_1) = J^*(y_2)$.
Proof: Suppose $\{\phi(t,\alpha)\}$ generates the field $\mathcal{F}$. We only need to verify that (8) is valid. For any y s.t. $(t,y(t)) \in S$, we know $y(t) = \phi(t, \alpha(t,y))$. Note that $\phi(t,\alpha)$

solves the E-L equation, i.e. $\frac{d}{dt} f(t, \phi(t, \alpha), \phi_t(t, \alpha)) = f_y(t, \phi(t, \alpha), \phi_t(t, \alpha))$. i.e.

$$f_{rt} + f_{ry}\phi_t + f_{rr}\phi_{tt} = f_y \tag{14}$$

Pluging the equations (10)(12) into (13) we can get (14).

□

2.3 The fundamental sufficiency theorem

Recall that the Weierstrass excess-function is defined by

$$E(t, y, r, q) = f(r, y, q) - f(t, y, r) - (q - r) f_r(t, y, r)$$

It is so named because its average value represented by the integral in the followig theorem is a measure of the excess of $J(y)$ over $J(y_0)$.

Theorem 2.3.1(Wererstrass-Hilbert) If $y_0 \in \mathcal{Y}$ is embedded in the field $\mathcal{F} = (S, p)$ and $y \in \mathcal{Y}$ with $(t, y(t)) \in S$, then

$$J(y) - J(y_0) = \int_{t_0}^{t_1} E(t, y, p(t, y), \dot{y}) dt \tag{15}$$

Proof: Suppose $\phi(t, \alpha(t, y_0)) = \phi(t, \alpha_0) = y_0(t)$. Then

$$p(t, y_0) = \partial_t[\phi(t, \alpha(t, y_0)] = \partial_t \phi(t, \alpha_0) = \dot{y_0}(t)$$

By the definition of the Hilbert integral,

$$J^*(y_0) = \int_{t_0}^{t_1} f(t, y_0, p(t, y_0)) + (\dot{y_0} - p(t, y_0)) f_r dt = \int_{t_0}^{t_1} f(t, y_0, \dot{y_0}) dt = J(y_0)$$

By theorem 2.2.5, $J^*(y) = J^*(y_0)$. It follows that

$$\begin{aligned} J(y) - J(y_0) &= J(y) - J^*(y_0) = J(y) - J^*(y) \\ &= \int_{t_0}^{t_1} \left[f(t, y, \dot{y}) - f(t, y, q) - (\dot{y} - p) f_r \right] dt = \int_{t_0}^{t_1} E(t, y, p, \dot{y}) dt \end{aligned}$$

□

Theorem 2.3.2(Fundamental Sufficiency Theorem) If $y_0 \in \mathcal{Y}$ is embedded in the field $\mathcal{F} = (S, p)$, $y \in \mathcal{Y}$ with $(t, y(t)) \in S$ and $E(t, y(t), p(t, y(t)), \dot{y}(t)) \geq 0 (> 0$ provided $\dot{y(t)} \neq p(t, y(t)))$, then $J(y_0) \leq J(y) (< J(y))$.

Proof: It follows immediately from the preceding theorem.

$\square$.

Remark: It follows from this theorem that $J(y_0)$ is a strong local minimum. This is because, under the hypotheses of the thorem, for any $y \in \mathcal{Y}$ with $||y - y_0||_{C[t_0,t_1]} < k$, i.e. $(t, y(t)) \in S$, we have $J(y_0) \leq J(y)$. This conclusion by itself does not exclude the possibility that $J(y_0)$ is actually a global minimum. If $k = +\infty$ then, S is an infinite strip and hence the inequality on the E-function holds for all $y \in \mathcal{Y}$.

The next theorem is defiend for the weak local minimum.

Theorem 2.3.3 Suppose $y_0 \in \mathcal{Y}$ is embedded in the field $\mathcal{F} = (S, p)$, If $\exists \delta > 0$, s.t $\forall t \in [t_0, t_1]$, $\forall y \in \mathcal{Y}$ with $(t, y(t)) \in S$ and $||\dot{y} - \dot{y_0}||_{C[t_0,t_1]} < \delta$ we have $E(t, y(t), p(t, y(t)), \dot{y}(t)) \geq 0 (> 0$ provided $\dot{y(t)} \neq p(t, y(t)))$, then $J(y_0) \leq J(y) (< J(y))$.

2.4 Sufficient combinations of conditions

Theorem 2.4.1 Suppose $y_0 \in \mathcal{Y}$ satisfies (i),(ii)$_N$ ((ii)$'_N$), (iii)$'$ and (iv), then $J(t_0)$ is a strong(proper strong) local minimum.
Proof: By theorem 2.2.3, a field about y_0 exists. Then, for any $y \in \mathcal{Y}$ s.t. $||y - y_0||_{C[t_0,t_1]}$ is below some positive constant, we can know $(t, y(t)) \in S$. Thus, by theorem 2.3.1

$$J(y) - J(y_0) = \int_{t_0}^{t_1} E(t, y(t), p(t, y(t)), \dot{y}(t)) dt$$

By the condition (ii)$_N$ and the continuity of p, $E(t, y, p, \dot{y}) \geq 0$ whenever $||y - y_0||_{C[t_0,t_1]}$ is sufficiently small. Then the conclusion follows from theorem 2.3.2. A similar arguement applies to the case where the alternative hyphthesis (ii)$'_N$ holds.

$\square$

Theorem 2.4.2 If $y_0 \in \mathcal{Y}$ satisfies conditions (i), (iii)$'$, (iii)$_N$ and (iv) ((i), (iii)$'_N$ and (iv)), then $J(y_0)$ is a strong(proper strong) local minimum.
Proof: Note that condition (iii)$'_N \Rightarrow$ (iii)$'$. Hence, under either of the alternative hypotheses, we have (i), (iv) and (iv) and Consequently having a field by theorem 2.2.3.

If we check the proof of theorem 1.4.1, we can find the following relation:

$$E(t,y,\dot{y},p)=\frac{(p-\dot{y})^2}{2}\frac{\partial^2 f}{\partial \dot{y}^2}(t,y,\dot{y}+\theta(p-\dot{y})) \quad \text{where} \quad \theta\in]0,1[\tag{16}$$

Then, conditions $(\text{iii})_N$ and $(\text{iii})'_N$ will respectively imply $(\text{ii})_N$ and $(\text{ii})'_N$. Thus, this theorem follows from theorem 2.3.2.

□

Theorem 2.4.3 If $y_0 \in \mathcal{Y}$ satisfies conditions (i), (iii)$'$ and (iv), then $J(y_0)$ is a proper weak local minimum.

Proof: Again, by theorem 2.2.3, there exists a field about y_0. By (iii)$'$ we have $\frac{\partial^2 f}{\partial \dot{y}^2}(f,t,y_0,\dot{y_0})>, \forall t\in[t_0,t_1]$. Note that $p(t,y_0(t))=\dot{y_0}(t)$. Hence, by the continuity of p, if $||\dot{y}-\dot{y_0}||_{C[t_0,t_1]}$ is sufficiently small, then

$$||\dot{y}+\theta(p-\dot{y})-\dot{y_0}||_{C[t_0,t_1]}\leq||\dot{y}-\dot{y_0}||+||\theta(q-\dot{y})||_{C[t_0,t_1]}$$

will again be sufficiently small.

By the continuity of $\frac{\partial^2 f}{\partial \dot{y}}(t)$ and t lying in a compact set, then $\exists C>0$ s.t. $\frac{\partial^2 f}{\partial \dot{y}}(t,y_0(t),\dot{y_0}(t))\geq 0$ for all $t\in[t_0,t_1]$. Combine the arguments above, if $||y-y_0||_{C^1[t_0,t_1]}$ is sufficiently small, so are $||y-y_0||$ and $||\dot{y}+\theta(p-\dot{y})-\dot{y_0}||$. Thus, by the continuity of $\frac{\partial^2 f}{\partial \dot{y}}$ and the relation (16), $E(t,y,\dot{y},p)>0$.

It follows from theorem 2.3.3 that $J(y_0)$ is a proper weak local minimum.

□

Theorem 2.4.4 Suppose $y_0\in\mathcal{Y}$ is embedded in the field (S,p) where S is the infinite strip $\{(t,y): t_0\leq t\leq t_1\}$, then $(\text{ii})_R$ or $(\text{iii})_R$ is sufficient for $J(y_0)$ be a global minimum, and either $(\text{ii})'_R$ or $(\text{iii})'_R$ is sufficient for a proper globla minimum. Moreover, under either of the last two conditions, y_0 is the unique function furnishing the proper global minimum.

Proof: These various conclusions follow immediately from the fundamental sufficiency theorem together with the comments after it and the relation (11).

□

2.5 Convexity

Most sufficient conditions in the preceding sections only can ensure that certain $J(y_0)$ is a local minimum. Now we consider under what additional conditions that it will become a global minimum.

Def 2.5.1(Convex set) Suppose $K \in \mathbb{R}^n$. K is said convex if $\forall x, y \in K, \forall t \in [0,1]$, we have $tx + (1-t)y \in K$.

Def 2.5.2(Convex function) Suppose $f : K \to \mathbb{R}^n$. f is said convex on K If

$$f(x + h(y-x)) \leq (1-h)f(x) + hf(y) \quad \forall x, y \in K, \forall h \in [0,1] \tag{17}$$

Theorem 2.5.3 Suppose f is $C^2(K)$. Then f is convex on $K \Leftrightarrow \forall x \in K$ $\forall y \in K$, $\alpha^T Hess(f)(x)\alpha \geq 0$, where $\alpha = y - x$.

Proof:$\Rightarrow$: $\forall x, y \in K$, since f is convex on K, then (17) holds, i.e.

$$\frac{f(x + h(y-x)) - f(x)}{h} \leq f(y) - f(x)$$

Applying differential mean value theorem, the inequality above can be rewritten as $Df(\xi_1)(y-x) \leq Df(\xi_2)(y-x)$ where $\xi_1 = x + \theta_1 h(y-x)$ and $\xi_2 = x + \theta_2(y-x)$. Hence, $\sum_i [f_i(\xi_2) - f_i(\xi_1)](y_i - x_i) \geq 0$. Apply differential mean value theorem again we obtain $\sum_{ij} f_{ij}(\eta_{ij})(y_i - x_i)(y_j - x_j)$. We let $h \to 0$, then $\eta_{ij} \to x$ and hence $\sum_{ij} f_{ij}(x)(y_i - x_i)(y_j - x_j) \geq 0$ i.e. $\alpha^T Hess(f)(x)\alpha \geq 0$.

$\Leftarrow$: Argue by contradiction and run a similar process we can complete the proof.

$\square$

Remark: If x is a interior point then the sufficent condition in this theorem is equivalent to $Hess(f)(x) \geq 0$.

Sometimes the condition in theorem 2.5.3 is very difficult to satisfy. Thus, we now give a weaker version of convexity.

Def 2.5.4(Convex at a point) Suppose $f : K \to \mathbb{R}^n$. f is said convex at $x \in K$ if

$$\forall y \in K, \exists \epsilon > 0, \forall 0 \leq h < \epsilon, f(x + h(y-x)) \leq (1-h)f(x) + hf(y)$$

The corresponding weaker version of theorem 2.5.3 turns out to be:

Theorem 2.5.5 Suppose f is $C^2(K)$. Then f is convex at x_0 if $\exists \delta > 0$ s.t. $\forall x \in B(x_0, \delta)$, $Hess(f)(x) \geq 0$.

Proof: $\forall y \in K$, first we assume y in $B(x_0, \delta)$, then $\forall 0 \leq h \leq 1, x_0 + h(y - x_0) \in B(x_0, \delta)$. Run a similar process as theorem 2.5.3 we can obtain the desired inequality. For $y \notin B(x_0, \delta)$, we can choose ϵ small enough s.t. $\forall 0 \leq h < \epsilon$, we have $h|y - x_0| < \delta$ and hence $x_0 + h(y - x_0) \in B(x_0, \delta)$.

$\square$

Def 2.5.6(Deformation) $\forall y_0, y \in \mathcal{Y}$, define $z : [0,1] \times [t_0, t_1] \to \mathbb{R}^n$ s.t. $z(h,t) = (1-h)y(t) + hy_0(t)$, $h \in [0,1], t \in [t_0, t_1]$. Note that, for fixed h, $z(,\cdot) \in \mathcal{Y}$. Then , the mapping $\zeta : [0,1] \to \mathcal{Y}$, where $\zeta(h) = z(h, \cdot)$ is called a deformation. It is so named because the function y is deformed continuously onto y_0 as h changes from 0 to 1.

The function z has the following simple properties:

- $z(0,t) = y_0(t), z(1,t) = y(t)$
- $|z(h,t) - y(t)| = h|y_0(t) - y(t)| \leq \big[||y||_{C[t_0,t_1]} + ||y_0||_{C[t_0,t_1]}\big]h \to 0$
- $|z(h,t) - y(t)| = h|\dot{y_0}(t) - \dot{y}(t)| \leq \big[||\dot{y}||_{C[t_0,t_1]} + ||\dot{y_0}||_{C[t_0,t_1]}\big]h \to 0$

Def 2.5.7 We say J is (strictly)convex at $y_0 \in \mathcal{Y}$ if $\forall y \in \mathcal{Y}, \exists \epsilon > 0$ s.t.

$$\Phi(h) := J(y_0) + h(J(y) - J(y_0)) - J(z(h,\cdot)) \geq 0 (> 0) \quad \text{provided } 0 < h < \epsilon$$

Theorem 2.5.8 If $J(y_0)$ is a weak local minimum and J is (strictly)convex at y_0, then $J(y_0)$ is a (proper)global minimum.

Proof: By the def 2.5.2, $\forall y \in \mathcal{Y}, \exists \epsilon > 0, \forall 0 < h < \epsilon$ we have $\Phi(h) \geq 0$ i.e.

$$h(J(y) - J(y_0)) \geq J(z(h,\cdot)) - J(y_0)$$

By the properties of z, $\exists h_0 > 0$ s.t. $\forall 0 < h < h_0$, $z(h,\cdot) \in B_1(y_0, \delta)$. Since $J(y_0)$ is a local minimum, then $J(z(h,\cdot)) - J(y_0) \geq 0$. Hence, $\forall y \in \mathcal{Y}, J(y) - J(y_0) \geq 0 (> 0)$.

$\square$

Theorem 2.5.9 $J(y_0)$ is a proper global minimum $\Leftrightarrow$

(a)$J(y_0)$ is a weak local minimum.

(b)J is strictly convex at y_0

Proof: $\Leftarrow$ It is clear by the preceding theorem.

$\Rightarrow$: (a) is clear.

$$\Phi'(h) = J(y) - J(y_0) - \int_{t_0}^{t_1} \left[f_y(t, z(h,t), \dot{z}(h,t))z_h + f_r(t, z(h,t), \dot{z}(h,t))z_{ht}\right]dt$$

$$\Rightarrow \Phi'(0) = J(y) - J(y_0) - \int_{t_0}^{t_1} f_y(t, y_0, \dot{y_0})(y_0 - y) + f_r(t, y_0, \dot{y_0}(\dot{y_0} - \dot{y})dt$$

$$\xlongequal{\text{E-L equation}} J(y) - J(y_0) - \int_{t_0}^{t_1} \frac{d}{dt}(f_r)(y_0 - y) + f_r(\dot{y_0} - \dot{y})dt$$

$$\xlongequal[\text{by part}]{\text{integrate}} J(y) - J(y_0) - f_r(y_0 - y)\Big|_{t_0}^{t_1} + \int_{t_0}^{t_1} \left[f_r(\dot{y_0} - \dot{y}) - f_r(\dot{y_0} - \dot{y})\right]dt$$

$$= J(y) - J(y_0) > 0 \quad \text{since } J(y_0) \text{ is a proper global minima}$$

Now we have $\Phi(0) = 0$, $\Phi'(0) > 0$. Hence, $\exists \epsilon > 0, \forall 0 < h < \epsilon$, $\Phi(h) > 0$. This completes the proof of (b).

$\square$

2.6 Examples

Example 2.6.1 $f(t, y, \dot{y}) = \dot{y}^{\frac{2}{3}}$ with end points (0,0) and (1,0).
Discussion: It is clear that $J(y) \geq 0$ and $y_0(t) = 0 \in \mathcal{Y}$. Hence $J(y_0)$ is the global minimum. The E-L equation is: $\frac{d}{dt}\dot{y}^{\frac{1}{2}} = 0$ and y_0 satisfies this equation. However, $\frac{\partial^2 f}{\partial \dot{y}^2} = \frac{3}{4}\frac{1}{\sqrt{\dot{y}^2}}$. which goes to infinity if $y = y_0$. Hence, the Legendre condition does not apply to the present example. This is because our basic assumption is that f si smooth or at least all derivatives of f appearing in the theorem exist and are continous.

Example 2.6.2 $f(t, y, \dot{y}) = \dot{y}^4 - \frac{2\dot{y}^2 t}{y^2+1}$ with end point (1,0) and (2,0).
Discussion: The E-L equation turns out to be

$$\frac{d}{dt}(4\dot{y}^3 - \frac{4\dot{y}t}{y^2+1}) = \frac{4\dot{y}^2 ty}{(y^2+1)^2}$$

It's difficult to obtain a general solution but we pick one trivial solution $y_0(t) = 0$. Then, $\frac{\partial^2 f}{\partial \dot{y}^2} = 12\dot{y}^2 - \frac{4t}{y^2+1} \Rightarrow \frac{\partial^2 f}{\partial \dot{y}^2}(t, y_0, \dot{y_0}) = -4t$. Thus, condition (iii)$'$ for a maximum is satisfied. Since we do not know whether $J(y_0)$ is a local minimum, we cannot directly use theorem 1.5.4 to see whether there is no conjugate point about y_0. Note that

$$2\omega = \frac{\partial^2 f}{\partial y^2}(t, y_0, \dot{y_0})\eta^2 + 2\frac{\partial^2 f}{\partial y \partial \dot{y}}(t, y_0, \dot{y_0})\eta\dot{\eta} + \frac{\partial^2 f}{\partial \dot{y}^2}(t, y_0, \dot{y_0})\eta^2 = -4t\dot{\eta}^2$$

$\Rightarrow \omega = 2t\dot{\eta}^2$.
The Jacobi equation turns out to be $\frac{d}{dt}(t\dot{\eta}) = 0$. i.e. $\eta = A\log t + B$.
If $\eta(t_0) = \eta(t_1) = 0$ i.e. $\eta(1) = \eta(2) = 0$, then $A = B = 0$. Hence there is no conjucate point. Thus, y_0 satisfies (i), (iii)$'$, (iv) and hence a field about y_0 exists according to theorem 2.2.3.Finally we examine the E-function.

$$\begin{aligned}
&E(t, y, r, q) = f(t, y, q) - f(t, y, r) - (q - r)f_r(t, y, r) \\
&E(t, y, p, \dot{y}) = f(t, y, \dot{y}) - f(t, y, p) - (y - p)f_r(t, y, p) \\
&\text{By (11), } E = (t, y, p, \dot{y}) = \frac{(p - \dot{y})^2}{2}\frac{\partial^2 f}{\partial y^2}(t, y, \xi) = \frac{(p - \dot{y})^2}{2}(-4t) < 0
\end{aligned}$$

Hence, the companion theorem for maximum to theorem 2.3.2 applies and $J(y_0)$ is a strong local minimum.

Example 2.6.3 $f(t, y, \dot{y}) = y^2 = t^2\dot{y} + k^2(\dot{y} - 1)^2(k \neq 0)$ with end point (0,0) and (1,1).
Discussion: The E-L equation is $2t + 2k^2(\dot{y} - 1) = 2y$ i.e. $k^2\ddot{y} - y = -t$. The general solution is $y = Ae^{\frac{t}{k}} + Be^{-\frac{t}{k}} + t$. We set $y_0 = \frac{2}{\alpha}(e^{\frac{t}{k}} + e^{-\frac{t}{k}}) + t$ and the one-parameter family $\phi(t, \alpha) = \alpha \cosh \frac{t}{k} + t$. It is not hard to see that $\{\phi(t, \alpha\})_\alpha$ generates a field about y_0 on the infinite strip $\{(t, y) : 0 \leq t \leq 1\}$. Furthermore, since $\frac{\partial^2 f}{\partial \dot{y}^2} = 2k^2 > 0$, then y_0 satisfies (iii)$'_R$. By theorem 2.4.4, $J(y_0)$ is a proper global minimum.

Example 2.6.4 $f(t, y, \dot{y}) = \dot{y}^2 + y\dot{y} + y^2$ with endpoints (0,0) and (1,1).
Discussion: The E-L equation is $\frac{d}{dt}(2\dot{y} + y) = \dot{y} + 2y \Rightarrow \ddot{y} - y = 0 \Rightarrow y_0(t) = A\cosh t + B\sinh t$. Plug in the endpoints we obtain $y_0(t) = C\sinh t$ where $C = \frac{2}{e - e^{-1}}$. $\frac{\partial^2 f}{\partial \dot{y}^2}(t, y_0, \dot{y_0}) = 2 > 0$.

$$2\omega = \frac{\partial^2 f}{\partial y^2}(t, y_0, \dot{y_0})\eta^2 + 2\frac{\partial^2 f}{\partial y \partial \dot{y}}(t, y_0, \dot{y_0})\eta\dot{\eta} + \frac{\partial^2 f}{\partial \dot{y}^2}(t, y_0, \dot{y_0})\eta^2 = 2\eta^2 + 2\eta\dot{\eta} + 2\dot{\eta}^2$$

$\Rightarrow \omega = \eta^2 + \eta\dot{\eta} + \dot{\eta}^2$. The Jacobi equation turns out to be the same as the E-L equation above. Thus, the general solution is $\eta = A\cosh t + B\sinh t$. If $\eta(0) = \eta(1) = 0$, then $A = B = 0 \Rightarrow \eta \equiv 0$. Hence there is no conjucate point.

To summarize, y_0 satisfies conditions (i), (iii)$'$ and (iv) and it follows from theorem 2.4.3 that $J(y_0)$ is a proper local minimum.
Claim: $J(y_0)$ is a proper global minimum. Note that f is in fact a function of two variables. A simple calculation shows that

$$Hess(f) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

Hence, it follows from theorem 2.5.4 that f is convex on the whole $\mathbb{R}^2$. Next,

$$\begin{aligned} &\Phi(h) = J(y_0) + h(J(y) - J(y_0)) - J(z(h,\cdot)) \\ &= \int_0^1 \Big[f(t, y_0(t), \dot{y_0}(t)) + h(f(t, y(t), \dot{t}(t)) - f(t, y_0(t), \dot{y_0}(t))) - f(t, z(\cdot, t), z_t(\cdot, t) \Big] dt \end{aligned}$$

For fixed $t \in [0,1]$, we set $p = (y(t), \dot{y}(t)), q = (y_0(t), \dot{y_0}(t))$. Since f is convex on $\mathbb{R}^2$, then, for $\forall h \in [0,1]$,

$$\begin{aligned} & f(q + h(p-q)) \leq (1-h)f(q) + hf(p) \\ \Rightarrow \quad & f(p) + h(f(q) - f(p)) - f(p + h(p-q)) \geq 0 \end{aligned}$$

That is to say, $\Phi(h) \geq 0$. By theorem 2.5.8, $J(y_0)$ is a global minimum.

Now we turn to several practical problems. In many physical problems, one may know for sure that there is must a minimum for the functional $J(y)$. Under this hypothesis, the issue will be much easier: if there exists only one local minimum(often obtained from the E-L equation), then we can conclude that it is exactly the global minimum.Consequently,we only need to find the local minimum. In general, when it comes to practical problems, some auxiliary methods are conducive to finding the minumum.

Example 2.6.5(The minimal surface of revolution) Let a surface be generated by the revolution about the $y-$axis of some curve with end points (x_1, y_1) and (x_2, y_2) where $x_1, x_2 > 0$ in the xy-plane. Our goal is to find the smooth curve for which the surface area is a minimum.

The area of the surface can be written as $J = 2\pi \int_{x_1}^{x_2} x\sqrt{1+\dot{y}^2}$. From the E-L equation we obtain $\dot{y}(x) = \frac{C}{\sqrt{x^2 - C^2}}$. The solution is $y_0(x) = Carccosh\frac{x}{C} + d$,

C, d is determined by the initial data:

$$\begin{cases} y_1 = Carccosh\frac{x_1}{C} + d \\ y_2 = Carccosh\frac{x_2}{C} + d \end{cases} \tag{18}$$

Here we assume the equations have solutions.

Now we calculate ω:

$$2\omega = \frac{\partial^2 f}{\partial y^2}(t, y_0, \dot{y_0})\eta^2 + 2\frac{\partial^2 f}{\partial y \partial \dot{y}}(t, y_0, \dot{y_0})\eta\dot{\eta} + \frac{\partial^2 f}{\partial \dot{y}^2}(t, y_0, \dot{y_0})\eta^2 = x(1 + \dot{y_0}^2)^{-\frac{2}{3}}\dot{\eta}^2$$

Then the Jocobi equation turns out to be: $\frac{d}{dx}\left[x(1 + \dot{y_0}^2)^{-\frac{2}{3}}\dot{\eta}\right] = 0$ i.e. $\dot{\eta}(x) = \frac{Ax^2}{(x^2-C^2)^{\frac{2}{3}}}$ where A is a constant. Integrate on both sides we obtain

$$\eta(x) = A\Big[\frac{-x}{\sqrt{x^2 - C^2}} + arccosh\frac{x}{C}\Big] + B$$

Suppose $\eta(x)$ satisfies $\eta(x_1) = \eta(x_2) = 0$. If $A \neq 0$ then $\dot{\eta}(x) > 0$ for all $x > 0$. In other words, $\eta(x)$ is strictly monotone increasing on $]0, +\infty[$. Hence it is impossible that $\eta(x_1) = \eta(x_2) = 0$ since $x_1 \neq x_2$. That is to say, A must be zero. Thus, $\eta \equiv B$. Plug in $\eta(x_1) = 0$ we get $\eta = 0$. Hence, there is no conjucate point, i.e. condition (iv) holds. Next, note that

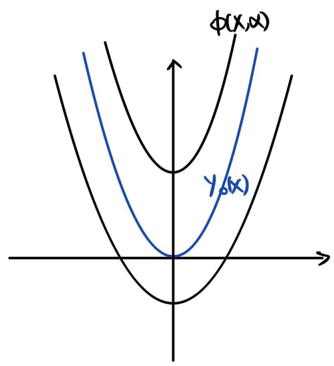
$$\frac{\partial^2 f}{\partial \dot{y}^2} = \frac{x}{(1 + \dot{y}^2)^{\frac{2}{3}}} > 0$$

Thus, condition $(\text{iii})'_N$ holds. According to theorem 2.4.2, $J(y_0)$ is a proper strong local minimum. In this problem it is very hard to say certain y_0 furnishes the global minimun since it is possible that there are two sets of solutions of (18). That is to say, there may exist two y_0 satisfying the E-L equation with the same form. However, the argument above does not distinguish them. Consequently, we can only conclude that $J(y_0)$ is a local minimum.

Example 2.6.6(The principle of Least Potential Energy) Our goal is to determine the shape of an inextensinble, flexible chain of length l and mass λ per unit length ρ that hangs in the uniform gravitational field of the earth with the end points (-a,b) and (a,b). Equivalently, we need to find the minimum of the functional $J(y) = \int_{-a}^{a} (\rho g y - \lambda)\sqrt{1+\dot{y}^2}dx$. From the E-L equation we get

$$\rho g y - \lambda = A\sqrt{1+\dot{y}^2} \Rightarrow y(x) = \frac{\lambda}{\rho g} + \frac{A}{\rho g}\cosh\frac{\rho g x}{A}$$

Now we choose b s.t. $y_0(0) = 0$ we get $A = -\lambda$ and hence $y_0(x) = \frac{\lambda}{\rho g}(1 - \cosh\frac{\rho g x}{\lambda})$. Here $\lambda < 0$. Now consider $\phi(x,\alpha) = \frac{\lambda}{\rho g} + \frac{\alpha}{\rho g}\cosh\frac{\rho g x}{\alpha}$, $\alpha > 0, x \in \mathbb{R}$.



Note that $\phi(x,-\lambda) = y_0(x)$ and $\phi(x,\alpha)$ solves the E-L equation and the family $\{\phi(x,\alpha)\}$ generates a field about y_0 with $S = \{(x,y) : x \in [-a,a], y \in]\frac{\lambda}{\rho g}, +\infty[\}$. Furthermore,

$$\frac{\partial^2 f}{\partial \dot{y}^2} = \frac{\rho g y - \lambda}{(1+\dot{y})^{\frac{2}{3}}} = \frac{-\lambda\sqrt{1+\dot{y}^2}}{(1+\dot{y})^{\frac{2}{3}}} = \frac{-\lambda}{1+\dot{y}^2} > 0$$

Hence, condition $(\text{iii})'_R$ is satisfied. It follows from (16) and theorem 2.3.2 that $J(y_0)$ is a proper strong local minimum.

References

[1] George M.Ewing. [M].Calculus of Variations with Applications.

[2] Gongqing Zhang. [M].Notes on Variational Principles.

[3] V.l. Arnold. [M].Mathematical Methods of Classical Mechanics.