LECTURES ON OPTIMAL CONTROL THEORY

Terje Sund

March 3, 2014

CONTENTS

1 INTRODUCTION

In the theory of mathematical optimization one try to find maximum or minimum points of functions depending of real variables and of other functions. Optimal control theory is a modern extension of the classical calculus of variations. Euler and Lagrange developed the theory of the calculus of variations in the eighteenth century. Its main ingredient is the Euler equation[1] which was discovered already in 1744. The simplest problems in the calculus of variation are of the type

$$\max \int_{t_0}^{t_1} F(t, x(t), \dot{x}(t))\,\mathrm{d}t$$

where $x(t_0) = x_0$, $x(t_1) = x_1$, t_0, t_1, x_0, x_1 are given numbers and F is a given function of three (real) variables. Thus the problem consists of finding

[1]The Euler equation is a partial differential equation of the form $\frac{\partial F}{\partial x} - \frac{d}{\mathrm{d}t}(\frac{\partial F}{\partial \dot{x}}) = 0$, where F is a function of three real variables, x is an unknown function of one variable t and $\dot{x}(t) = \frac{dx}{\mathrm{d}t}(t)$. Here $\frac{\partial F}{\partial x}$ stands for the partial derivative of F with respect to the second variable x, $\frac{\partial F}{\partial x}(t, x, \dot{x})$, and $\frac{\partial F}{\partial \dot{x}}$ is the partial derivative with respect to the third variable $\dot{x}$.

functions $x(t)$ that make the integral $\int_{t_0}^{t_1} F(t, x(t), \dot{x}(t))\,\mathrm{d}t$ maximal or minimal. Optimal control theory has since the 1960-s been applied in the study of many different fields, such as economical growth, logistics, taxation, exhaustion of natural resources, and rocket technology (in particular, interception of missiles).

2 FUNCTIONS OF SEVERAL VARIABLES.

Before we start on the calculus of variations and control theory, we shall need some basic results from the theory of functions of several variables. Let $A \subseteq \mathbb{R}^n$, and let $F : A \to \mathbb{R}$ be a real valued function defined on the set A. We let $F(\vec{x}) = F(x_1, \ldots, x_n), \vec{x} = (x_1, \ldots, x_n) \in A, ||\vec{x}|| = \sqrt{x_1^2 + \cdots + x_n^2}$. $||\vec{x}||$ is called (the Euklidean) norm of $\vec{x}$.

A vector $\vec{x}_0$ is called an *inner point* of the set A, if there exists a $r > 0$ such that

$$B(\vec{x}_0, r) = \{\vec{x} \in \mathbb{R}^n \,:\, ||\vec{x} - \vec{x}_0|| < r\} \subseteq A.$$

We let

$$A^0 = \{\vec{x}_0 \in A \,:\, \vec{x}_0 \text{ is an inner point of } A\}$$

The set A^0 is called the *interior* of A. Let

$$\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$$

be the vectors of the standard basis of $\mathbb{R}^n$, that is,

$$\vec{e}_i = (0, 0, \ldots, 0, 1, 0, \ldots, 0) \quad (i = 1, 2, \ldots, n),$$

hence $\vec{e}_i$ has a 1 on the i-th entry, 0 on all the other entries.

Definition 1. let $\vec{x}_0 \in A$ be an inner point. The first order *partial derivatives* $F_i'(\vec{x}_0)$ of F at the point $\vec{x}_0 = (x_1^0, \ldots, x_n^0)$ with respect to the i-th variable x_i, is defined by

$$\begin{aligned} F_i'(\vec{x}_0) &= \lim_{h \to 0} \frac{1}{h}[F(\vec{x}_0 + h\vec{e}_i) - F(\vec{x}_0)] \\ &= \lim_{h \to 0} \frac{1}{h}[F(x_1^0, \ldots, x_i^0 + h, x_{i+1}^0, \ldots, x_n^0) - F(x_1^0, \ldots, x_n^0)], \end{aligned}$$

if the limit exists.

We also write

$$F_i'(\vec{x}_0) = \frac{\partial F}{\partial x_i}(\vec{x}_0) = D_i F(\vec{x}_0) \quad (1 \leq i \leq n)$$

for the i-th first order *partial derivative* of of F i $\vec{x}_0$.

Second order partial derivatives are written

$$F_{i,j}'(\vec{x}_0) = \frac{\partial^2 F}{\partial x_j \partial x_i}(\vec{x}_0) = D_{i,j} F(\vec{x}_0) \quad (1 \leq i \leq n)$$

where we let

$$\frac{\partial^2 F}{\partial x_j \partial x_i}(\vec{x}_0) = \frac{\partial}{\partial x_j}\big(\frac{\partial F}{\partial x_i}\big)(\vec{x}_0) = F_j'(F_i'(\vec{x}_0)).$$

The gradient of the function F at $\vec{x}_0$ is the vector

$$\nabla F(\vec{x}_0) = (\frac{\partial F}{\partial x_1}(\vec{x}_0), \ldots, (\frac{\partial F}{\partial x_n}(\vec{x}_0))$$

Example 1. Let $F(x_1, x_2) = 7x_1^2 x_2^3 + x_2^4,\ \vec{x}_0 = (2, 1)$ Then

$$\begin{aligned} \nabla F(x_1, x_2) &= (14x_1 x_2^3, 21x_1^2 x_2^2 + 4x_2^3), \\ \nabla F(x_1, x_2) &= (14x_1 x_2^3, 21x_1^2 x_2^2 + 4x_2^3) \\ &= (28, 28). \end{aligned}$$

Level surfaces: Let $c \in \mathbb{R},\ F : \mathbb{R}^n \to \mathbb{R}$. The set

$$M_c = \{\vec{x} : F(\vec{x}) = c\}$$

is called *the level surface* (or the level hypersurface) of F corresponding to the level c. if $\vec{x}_0 = (x_1^0, x_2^0, \ldots, x_n^0) \in M_c$, then we define *the tangent plane* of the level surface M_c at the point $\vec{x}_0$ as the solution set of the equation

$$\nabla F(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0) = 0.$$

In particular, $\nabla F\vec{x}_0)$ is perpendicular to the tangent plane at $\vec{x}_0$, $\nabla F(\vec{x}_0) \perp$ the tangent plane at $\vec{x}_0$.

Example 2. let $F(x_1, x_2) = x_1^2 x_2$. then

$$\nabla F(x_1, x_2) = (2x_1 x_2, x_1^2) \text{ and } \nabla F(1, 2) = (4, 1).$$

Tangent the equation at the point $(1,2)$ for the level surface (the level curve) $x_1^2 x_2 = 2$ is $\nabla F(1,2) \cdot (\vec{x} - (1,2)) = 0$, hence $(4,1) \cdot (x_1 - 1, x_2 - 1) = 0$, $4x_1 - 4 + x_2 - 2 = 0$, that is $4x_1 + x_2 = 6$.

Example 3. Let $F(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$, and consider the level surface

$$x_1^2 + x_2^2 + x_3^2 = 1,$$

the sphere with radius 1 and centre at the origin. (a) Then

$$\nabla F(x_1, x_2, x_3) = 2(x_1, x_2, x_3)$$
$$\nabla F(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}) = \frac{2}{\sqrt{3}}(1,1,1).$$

The tangent plane at the point $\frac{1}{\sqrt{3}}(1,1,1)$ has the equation:

$$\frac{2}{\sqrt{3}}(1,1,1) \cdot (x_1 - \frac{1}{\sqrt{3}}, x_2 - \frac{1}{\sqrt{3}}, x_3 - \frac{1}{\sqrt{3}}) = 0,$$
$$\frac{2}{\sqrt{3}}(x_1 + x_2 + x_3) = \frac{2}{\sqrt{3}}\frac{3}{\sqrt{3}},$$
$$x_1 + x_2 + x_3 = \sqrt{3}.$$

(b) Here

$$\nabla F(0, \frac{3}{5}, \frac{4}{5}) = 2(0, \frac{3}{5}, \frac{4}{5}) = \frac{2}{5}(0,3,4)$$

The tangent plane at $(0, \frac{3}{5}, \frac{4}{5})$ has the equation

$$\frac{2}{5}(0,3,4) \cdot (x_1 - 0, x_2 - \frac{3}{5}, x_3 - \frac{4}{5}) = 0,$$
$$0 + 3(x_2 - \frac{3}{5}) + 4(x_3 - \frac{4}{5}) = 0,$$
$$3x_2 + 4x_3 = 5.$$

Directional derivatives. Assume that $F : A \subseteq \mathbb{R}^n \to \mathbb{R}$. Let $\vec{a} \in \mathbb{R}^n$ and $\vec{x} \in A^0$ be an inner point of A. The derivative of F *along* $\vec{a}$ at the point $\vec{x}$ is defined as

$$F'_{\vec{a}}(\vec{x}) = \lim_{h \to 0} \frac{1}{h}[F(\vec{x} + h\vec{a}) - F(\vec{x})]$$

if the limit exists.

Remark 1. The average growth of F from $\vec{x}$ of $\vec{x} + h\vec{a}$ is

$$\frac{1}{h}[F(\vec{x} + h\vec{a}) - F(\vec{x})],$$

hence the above definition is natural if we think of the derivative as a rate of change.

Definition 2. We say that F is *continuously differentiable,* or C^1, if the function $F'_{\vec{a}}(\vec{x})$ is continuous at $\vec{x}$ for all $\vec{a} \in \mathbb{R}^n$ and all $\vec{x} \in A^0$, that is if

$$\lim_{\vec{y}\to 0} F'_{\vec{a}}(\vec{x} + \vec{y}) = F'_{\vec{a}}(\vec{x}).$$

if F is C^1, then we can show that the map $\vec{a} \mapsto F'_{\vec{a}}(\vec{x})$ is linear, for all $\vec{x} \in A^0$. We show first the following version of the Mean Value Theorem for functions of several variables,

Proposition 1 (Mean Value Theorem). Assume that F is continuously differentiable on an open set $A \subseteq \mathbb{R}^n$ and that the closed line segment

$$[\vec{x}, \vec{y}] = \{t\vec{x} + (1-t)\vec{y} \,:\, 0 \le t \le 1\}$$

from $\vec{x}$ to $\vec{y}$ is contained in A. Then there exists a point $\vec{w}$ on the open line segment $(\vec{x}, \vec{y}) = \{t\vec{x} + (1-t)\vec{y} \,:\, 0 < t < 1\}$ such that

$$F(\vec{x}) - F(\vec{y}) = F'_{\vec{x}-\vec{y}}(\vec{w})$$

Proof. Set $g(t) = F(t\vec{x} + (1-t)\vec{y}), \quad t \in (0,1)$. Then

$$\begin{aligned}
\frac{1}{h}[g(t+h) - g(t)] &= \frac{1}{h}[F((t+h)\vec{x} + (1-(t+h))\vec{y}) - F(t\vec{x} + (1-t)\vec{y})] \\
&= \frac{1}{h}[F(t)\vec{x} + (1-t)\vec{y} + h(\vec{x} - \vec{y})) - F(t\vec{x} + (1-t)\vec{y})] \\
&\xrightarrow[h\to 0]{} F'_{\vec{x}-\vec{y}}(t\vec{x} + (1-t)\vec{y})
\end{aligned}$$

Since F was continuously differentiable, it is clear that $g'(t)$ is continuous for $0 < t < 1$, and $g'(t) = F'_{\vec{x}-\vec{y}}(t\vec{x} + (1-t)\vec{y})$. Moreover, $g(0) = F(\vec{y})$ and $g(1) = F(\vec{x})$. We apply the "ordinary" Mean Value Theorem to the function g. Hence there exists a $\theta \in (0,1)$ such that

$$\begin{aligned}
&g(1) - g(0) = g'(\theta) \\
&= F'_{\vec{x}-\vec{y}}(\theta\vec{x} + (1-\theta)\vec{y}) = F'_{\vec{x}-\vec{y}}(\vec{w}),
\end{aligned}$$

where $\vec{w} = \theta x + (1-\theta)\vec{y}$ lies on the open line segment $(\vec{x}, \vec{y})$. Hence we have proved the Mean Value Theorem. ■

Proposition 2. Let $\vec{x} \in A^0$ be given. If F is C^1, then

$$\begin{aligned} &(1) \quad F'_{c\vec{a}}(\vec{x}) = cF'_{\vec{a}}(\vec{x}), \quad \text{for all } c \in \mathbb{R} \text{ and } \vec{a} \in \mathbb{R}^n \\ &\qquad \text{and} \\ &(2) \quad F'_{\vec{a}+\vec{b}}(\vec{x}) = F'_{\vec{a}}(\vec{x}) + F'_{\vec{b}}(\vec{x}), \end{aligned}$$

that is, for each fixed $\vec{x}$, the map $\vec{a} \mapsto F'_{\vec{a}}(\vec{x})$ is a linear map of $\mathbb{R}^n$ into $\mathbb{R}$.

Proof. (1) if $c = 0$, it is clear that both sides of (1) are equal to 0. Assume that $c \neq 0$. Then

$$\begin{aligned} F'_{c\vec{a}}(\vec{x}) &= \lim_{h\to 0} \frac{1}{h}[F(\vec{x} + hc\vec{a}) - F(\vec{x})] \\ &= c \cdot \lim_{h\to 0} \frac{1}{hc}[F(\vec{x} + hc\vec{a}) - F(\vec{x})] \\ &\overset{k=hc}{=} c \lim_{k\to 0} \frac{1}{h}[F(\vec{x} + k\vec{a}) - F(\vec{x})] = cF'_{\vec{a}}(\vec{x}). \end{aligned}$$

(2) F is C^1 hence $F'_{\vec{a}}(\vec{x})$ and $F'_{\vec{b}}(\vec{x})$ exist. Since $x \in A^0$, exists $r > 0$ such that

$$||\vec{x} - \vec{y}|| < r \Rightarrow \vec{y} \in A.$$

Choose h such that

$$||h(\vec{a} + \vec{b})|| < r \text{ and } ||h\vec{b}|| < r.$$

Then

$$\vec{x} + h(\vec{a} + \vec{b}) \in A \text{ and } \vec{x} + h\vec{b} \in A,$$

hence

$$\begin{aligned} (*) \quad &F(\vec{x} + h(\vec{a} + \vec{b})) - F(\vec{x}) \\ &F(\vec{x} + h(\vec{a} + \vec{b})) - F(\vec{x} + h\vec{b}) - F(\vec{x})] + [F(\vec{x} + h\vec{b}) - F(\vec{x})] \end{aligned}$$

We apply the Mean Value Theorem to the expression in the first bracket. Thus there exists a $\theta \in (0, 1)$ such that

$$(**) \quad F((\vec{x} + h\vec{b}) + h\vec{a}) - F(\vec{x} + h\vec{b}) = F'_{h\vec{a}}(\vec{x} + h\vec{a} + \theta h\vec{b}).$$

By part (1) we have

$$F'_{h\vec{a}}(\vec{x}+h\vec{a}+\theta h\vec{b})=hF'_{\vec{a}}(\vec{x}+h\vec{a}+\theta h\vec{b})$$

If we divide the equation in $(**)$ by h, then we find

$$\frac{1}{h}[F((\vec{x}+h\vec{b})+\vec{a}))-F(\vec{x}+h\vec{b})]=F'_{\vec{a}}(\vec{x}+h\vec{a}+\theta h\vec{b})\underset{h\to 0}{\longrightarrow}F'_{\vec{a}}(\vec{x}),$$

since F is continuously differentiable. Consequently, dividing by h in $(*)$, it follows that

$$\begin{aligned}&\lim_{h\to 0}\frac{1}{h}[F((\vec{x}+h(\vec{a}+\vec{b}))-F(\vec{x})]\\&=F'_{\vec{a}}(\vec{x})+\lim_{h\to 0}\frac{1}{h}[F(\vec{x}+h\vec{b})-F(\vec{x})]=F'_{\vec{a}}(\vec{x})+F'_{\vec{b}}(\vec{x})\quad\blacksquare\end{aligned}$$

We write $\vec{a}=\sum_{i=1}^{n}a_i\vec{e}_i$. Then

$$F'_{\vec{a}}(\vec{x})=\sum_{i=1}^{n}a_iF'_{\vec{e}_i}(\vec{x})=\sum_{i=1}^{n}a_i\frac{\partial F}{\partial x_i}(\vec{x})=\nabla F(\vec{x})\cdot\vec{a}.$$

As a consequence we have,

Corollary 1. If F is C^1 on $A\subseteq\mathbb{R}^n$, then

$$F'_{\vec{a}}(\vec{x})=\nabla F(\vec{x})\cdot\vec{a},$$

for all $\vec{x}\in A^0$.

If $||\vec{a}||=1$, then $F'_{\vec{a}}(\vec{x})$ is often called the *directional derivative* of F at the point $\vec{x}$ in the direction of $\vec{a}$. From Corollary 1 it follows easily that

Corollary 2. F is continuously differentiable at the point $\vec{x}_0\Leftrightarrow\nabla F$ is continuous at the point $\vec{x}_0\Leftrightarrow$ all the first order partial derivatives of F are continuous at $\vec{x}_0$.

Proof. Let $\vec{a}=(a_1,\ldots,a_n)\in\mathbb{R}^n,\ ||\vec{a}||=1$. Then

$$\begin{aligned}&|F'_{\vec{a}}(\vec{x})-F'_{\vec{a}}(\vec{x}_0)|=|[\nabla F(\vec{x})-\nabla F(\vec{x}_0)]\cdot\vec{a}]|\\=&|\nabla F(\vec{x})-\nabla F(\vec{x}_0)|\cdot||\vec{a}||\cdot|\nabla F(\vec{x})-\nabla F(\vec{x}_0)|.\end{aligned}$$

This proves that $F'_{\vec{a}}(\vec{x})-F'_{\vec{a}}(\vec{x}_0)\underset{\vec{x}\to\vec{x}_0}{\longrightarrow}0\Leftrightarrow\nabla F(\vec{x})-\nabla F(\vec{x}_0)\underset{\vec{x}\to\vec{x}_0}{\longrightarrow}0$. Hence F is continuously differentiable at the point $\vec{x}_0\Leftrightarrow\nabla F$ is continuous at $\vec{x}_0$.

If we let $\vec{a} = \vec{e}_i,\ (1 \leq i \leq n)$, we find that $\frac{\partial F}{\partial x_i}(\vec{x}) = F'_{\vec{e}_i}(\vec{x})$ is continuous at $\vec{x}_0$ for $i = 1, \ldots, n$ if F is continuously differentiable i $\vec{x}_0$. Conversely, if the partial derivatives $\frac{\partial F}{\partial x_i}(\vec{x}_0)$ are continuous for $i = 1, \ldots n$, then

$$F'_{\vec{a}}(\vec{x}) = \sum_{i=1}^{n} a_i F'_{\vec{e}_i}(\vec{x}) = \sum_{i=1}^{n} a_i \frac{\partial F}{\partial x_i}(\vec{x})$$

is continuous at $\vec{x}_0$ for all $\vec{a} \in \mathbb{R}^n$. Hence it follows that F is continuously differentiable at $\vec{x}_0$. ■

Convex sets. A set $S \subseteq \mathbb{R}^n$ is called *convex* if

$$t\vec{x} + (1-t)\vec{y} \in S$$

for all $\vec{x}, \vec{y} \in S$, and all $t \in [0, 1]$.

In other words, S is convex if and only if the closed line segment $[\vec{x}, \vec{y}]$ between $\vec{x}$ and $\vec{y}$ is contained in S for all $\vec{x}, \vec{y} \in S$.

Example 4. Let us show that the unit disc with centre at the origin, $S = \{(x_1, x_2) \,:\, x_1^2 + x_2^2 \leq 1\}$ is a convex subset of $\mathbb{R}^2$.

Let $\vec{x} = (x_1, x_2), \vec{y} = (y_1, y_2) \in S$, $0 \leq t \leq 1$. We must show that

$$\vec{z} = t\vec{x} + (1-t)\vec{y} = t(x_1, x_2) + (1-t)(y_1, y_2) \in S.$$

Using the triangle inequality and that $||t\vec{u}|| = |t| \cdot ||\vec{u}||, \vec{u} \in \mathbb{R}^n$, we find that

$$\begin{aligned} ||\vec{z}|| &= ||t\vec{x} + (1-t)\vec{y}|| \\ &\leq ||t\vec{x}|| + ||(1-t)\vec{y}|| = t||\vec{x}|| + (1-t)||\vec{y}|| \\ &\leq t \cdot 1 + (1-t) \cdot 1 = 1 \end{aligned}$$

Hence $\vec{z} \in S$. This shows that S is convex. ■

Exercise 1. Let $a > 0$ and $b > 0$. Show that the elliptic disc

$$S = \{(x_1, x_2) \,:\, \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} \leq 1\}$$

er a convex subset of $\mathbb{R}^2$.

Concave and convex functions. A C^2 function (that is, a function possessing a continuous second derivative) $f : [a, b] \to \mathbb{R}$ is *concave* if $f''(x) \geq 0$ for all x in the open interval (a, b). The function f is *convex* if $f''(x) \geq 0$ for

all $x \in (a, b)$. If f is convex, we see see geometrically that the chord $[(x, f(x))]$ always lies under or on the graph of f. Equivalently, the inequality

$$(*) \quad tf(x) + (1-t)f(y) \leq f(tx + (1-t)y),$$

holds for all $x, y \in (a, b)$, and all $t \in [0, 1]$. The converse is also right: If $(*)$ holds, then f is concave. For functions of several variables, the second derivative at a vector $\vec{x}$ is no real number, but a bilinear function. As a consequence, we will use the inequality $(*)$ when we define concavity for functions of several variables.

Definition 3. Let S be a convex subset of $\mathbb{R}^n$ and let f be a real valued function defined on S. We say that f is *concave* if f satisfies the inequality

$$(**) \quad f(t\vec{x} + (1-t)\vec{y}) \geq tf(\vec{x}) + (1-t)f(\vec{y}),$$

for all $\vec{x}, \vec{y} \in S$ and all $t \in [0, 1]$. A function f is called *convex* if the opposite inequality holds. The function is *strongly concave* (respectively *strongly convex*) if strict inequality holds in $(**)$.

Remark 2. That f is concave, means geometrically that the tangent plane of the surface $z = f(\vec{x})$ lies over (or on) the surface at every point $(\vec{y}, f(\vec{y}))$: the equation for the tangent plane through the point $(\vec{y}, f(\vec{y}))$is

$$z = \nabla f(\vec{y}) \cdot (\vec{x} - \vec{y}) + f(\vec{y}),$$

hence

$$f(\vec{x}) \leq z = \nabla f(\vec{y}) \cdot (\vec{x} - \vec{y}) + f(\vec{y})$$

for all $\vec{x}$ on the tangent plane.

Proposition 3. Assume that the functions f and g are defined on a convex set $S \subseteq \mathbb{R}^n$. Then the following statements hold

(a) If f and g are concave and $a \geq 0$, $b \geq 0$, then the function $af + bg$ is concave.

(b) If f and g are convex and $a \geq 0$ then the function $af + bg$ is convex.

Proof. (b): Assume that f and g are convex, $a \geq 0, b \geq 0$. Set $h(\vec{x}) = af(\vec{x}) + bg(\vec{x}) \quad (\vec{x} \in S)$, and let $t \in (0, 1)$. For all $\vec{x}, \vec{y} \in S$ then

$$h(t\vec{x} + (1-t)\vec{y}) = af(t\vec{x} + (1-t)\vec{y}) + bg(t\vec{x} + (1-t)\vec{y})$$
$$\leq a[tf(\vec{x}) + (1-t)f(\vec{y})] + b[tg(\vec{x}) + (1-t)g(\vec{y})] = th(\vec{x}) + (1-t)h(\vec{y}).$$

Accordingly h convex.

Part (a) is shown in the same way. ■

The following useful "$2nd$ derivative test" holds:

Proposition. Let $f : S \subseteq \mathbb{R}^2 \to \mathbb{R}$, where S is open and convex. If all the $2nd$ order partial derivatives of f are continuous (that is, f is C^2 on S), then

$$\begin{aligned}
(a) \quad & f \text{ is convex} \Leftrightarrow \frac{\partial^2 f}{\partial x_1^2}(\vec{x}) \geq 0, \frac{\partial^2 f}{\partial x_2^2}(\vec{x}) \geq 0 \text{ and} \\
& \frac{\partial^2 f}{\partial x_1^2}(\vec{x})\frac{\partial^2 f}{\partial x_2^2}(\vec{x}) - \frac{\partial^2 f(\vec{x})}{\partial x_1 \partial x_2}^2 \geq 0 \text{ for all } \vec{x} \in S. \\
(b) \quad & f \text{ is concave} \Leftrightarrow \frac{\partial^2 f}{\partial x_1^2}(\vec{x}) \leq 0, \frac{\partial^2 f}{\partial x_2^2}(\vec{x}) \leq 0 \text{ and} \\
& \frac{\partial^2 f}{\partial x^2}(\vec{x})\frac{\partial^2 f}{\partial x_2^2}(\vec{x}) - \frac{\partial^2 f(\vec{x})}{\partial x_1 \partial x_2}^2 \geq 0 \text{ for all } \vec{x} \in S.
\end{aligned}$$

We vil prove this proposition below (see Proposition 8). First we shall need some results on symmetric 2×2 matrices Let

$$Q = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

be a symmetric 2×2-matrix, and let

$$Q(\vec{x}) = \vec{x}^t Q \vec{x} = (x_1, x_2) Q \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = ax_1^2 + 2bx_1x_2 + cx_2^2, \quad \vec{x} \in \mathbb{R}^2.$$

Definition 4. We say that Q is

(a) positive semidefinite if $\vec{x}^t Q \vec{x} \geq 0$ for all $\vec{x} \in \mathbb{R}^2$.

(b) positive definite if $\vec{x}^t Q \vec{x} > 0$ for all $\vec{x} \in \mathbb{R}^2 \setminus \{0\}$.

(c) negative semidefinite if $\vec{x}^t Q \vec{x} \leq 0$ for all $\vec{x} \in \mathbb{R}^2$.

(d) negative definite if $\vec{x}^t Q \vec{x} < 0$ for all $\vec{x} \in \mathbb{R}^2 \setminus \{0\}$.

We will also write $Q \geq 0$ if Q is positive semidefinite, $Q > 0$ if Q is positive definite, and similarly for (c) and (d).

Proposition 4. Q is positive semidefinite $\Leftrightarrow a \geq 0, c \geq 0$ and $ac - b^2 \geq 0$.

Proof. $\Leftarrow$: assume that $a \geq 0, c \geq 0$ and $ac - b^2 \geq 0$. If $a = 0$ then

$$ac - b^2 = -b^2 \geq 0$$

hence $b = 0$. Therefore

$$\begin{aligned}(x_1, x_2)Q \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= (x_1, x_2) \begin{pmatrix} 0 & 0 \\ 0 & c \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= (x_1, x_2) \begin{pmatrix} 0 \\ cx_2 \end{pmatrix} = cx_2^2 \geq 0,\end{aligned}$$

for all (x_1, x_2). Consequently Q is positive semidefinite.

Assume that $a \neq 0$. Then

$$\begin{aligned}(x_1, x_2) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1, x_2) \begin{pmatrix} ax_1 + bx_2 \\ bx_1 + cx_1 \end{pmatrix} &= ax_1^2 + 2bx_1x_2 + cx_2^2 \\ &= a[x_1^2 + 2\frac{b}{a}x_1x_2 + \frac{c}{a}x_2^2] \\ &= a[x_1^2 + 2\frac{b}{a}x_1x_2 + \frac{b^2}{a^2}x_2^2 + (\frac{c}{a} - \frac{b^2}{a^2})x_2^2] \\ &= a[(x_1 + \frac{b}{a}x_2)^2 + \frac{ac - b^2}{a^2}x_2^2] \geq 0\end{aligned}$$

for all $(x_1, x_2) \in \mathbb{R}^2$. Hence is $Q \geq 0$.

$\Rightarrow$: Assume that $Q \geq 0$, then $\vec{x}^t Q\vec{x} \geq 0$, $\vec{x} \in \mathbb{R}^2$. In particular,

$$(1, 0) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = a \geq 0$$

and

$$(0, 1) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = c \geq 0$$

Assume that $a > 0$: As we have seen above,

$$\begin{aligned}(x_1, x_2) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ = a[(x_1 + \frac{b}{a}x_2)^2 + \frac{ac - b^2}{a^2}x_2^2].\end{aligned}$$

If we let $x_1 = -\frac{b}{a}x_2$, we deduce that $\frac{ac-b^2}{a^2}x_2^2 \geq 0$, for all x_2, hence $ac-b^2 \geq 0$.

Assume that $a = 0$: Then

$$\begin{aligned} 0 \leq \vec{x}^t Q\vec{x} &= ax_1^2 + 2bx_1x_2 + cx_2^2 \\ &= 2bx_1x_2 + cx_2^2 \end{aligned}$$

if $b > 0$, the fact that $x_2 = 1, x_1 = -n$, implies

$$0 \leq -2bn + c \underset{n\to\infty}{\longrightarrow} -\infty,$$

yields a contradiction. Hence $b = 0$, and $ac - b^2 = 0 \geq 0$. ■

Further, it is easy to show

Proposition 5. Q is positive definite $\Leftrightarrow a > 0$ (and $c > 0$) and $ac-b^2 > 0$.

Exercise. Prove Proposition 5.

The following chain rule for real functions of several variables will be useful.

Proposition 6. (Chain Rule) Let f be a real C^1- function defined on an open convex subset S of $\mathbb{R}^n$, and assume that $\vec{r} : [a, b] \to S$ is differentiable. Then the composed function $f \circ \vec{r} = f \circ (r_1, \ldots, r_n) : [a, b] \to \mathbb{R}$ is well defined. Set $g(t) = f(\vec{r}(t))$. Then g is differentiable and

$$g'(t) = \frac{\partial f}{\partial x_1}(\vec{r}(t))r_1'(t) + \cdots + \frac{\partial f}{\partial x_n}(\vec{r}(t))r_n'(t) = \nabla f(\vec{r}(t)) \cdot \vec{r}\,'(t).$$

Proof. consider

$$(*) \quad \frac{g(t+h) - g(t)}{h} = \frac{f(\vec{r}(t+h)) - f(\vec{r}(t))}{h}$$

Since S is open, there exists an open sphere B_r with positive radius and centre at $\vec{r}(t)$ which is contained in S. We choose $h \neq 0$ so small that $\vec{r}(t+h) \in B_r$. Set $\vec{u} = \vec{r}(t+h), \vec{v} = \vec{r}(t)$. Then the numerator on the right hand side of $(*)$ is $f(\vec{u}) - f(\vec{v})$. By applying the Mean Value Theorem to the last difference, we obtain

$$f(\vec{u}) - f(\vec{v}) = \nabla f(\vec{z}) \cdot (\vec{u} - \vec{v}),$$

where $\vec{z} = \vec{u} + \theta(\vec{u} - \vec{v})$ for a $\theta \in (0,1)$. Hence the difference quotient in $(*)$ may be written

$$(**) \quad \frac{1}{h}[g(t+h) - g(t)] = \frac{1}{h}\nabla f(\vec{z}) \cdot [(\vec{r}(t+h)) - (\vec{r}(t))]$$

As $h \to 0$, we find $\vec{u} \to \vec{v}$, so that $\vec{z} \to \vec{u}$. Since$f$ is C^1, is ∇f continuous, $\nabla f(\vec{z}) \to \nabla f(\vec{u}) = \nabla f(\vec{r}(t))$. Therefore the right side of $(**)$ converges to $\nabla f(\vec{r}(t)) \cdot \vec{r}\,'(t)$. This shows that $g'(t)$ exists and is equal to the inner product $\nabla f(\vec{r}(t)) \cdot \vec{r}\,'(t)$. ■

If the function f is defined on an open subset S of $\mathbb{R}^2$ for which the partial derivatives of 2nd order exist on S, then we may ask if the two mixed partial derivative are equal. That is, if $\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1}$. This is not always the case however, the following result which we state without proof, will be sufficient for our needs.

Proposition 7. Assume that f is a real valued function defined on an open subset S of $\mathbb{R}^2$ and that all the first and second order partial derivatives of f are continuouse on S. Then

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(\vec{x}) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(\vec{x}),$$

for all $\vec{x}$ i S.

Proof. See for instance Tom M. Apostol, Calculus Vol. II, 4.25.

Next we will apply the last four propositions to prove

Proposition 8. Let f be a C^2 function defined on an open convex subset S of $\mathbb{R}^2$. Then the following statements hold,

(a) f is convex $\Leftrightarrow \frac{\partial^2 f}{\partial x_1^2} \geq 0$, $\frac{\partial^2 f}{\partial x_2^2} \geq 0$ and $\Delta_f = \frac{\partial^2 f}{\partial x_1^2}\frac{\partial^2 f}{\partial x_2^2} - (\frac{\partial^2 f}{\partial x_1 \partial x_2})^2 \geq 0$

(b) f is concave $\Leftrightarrow \frac{\partial^2 f}{\partial x_1^2} \leq 0$, $\frac{\partial^2 f}{\partial x_2^2} \leq 0$ and $\Delta_f \geq 0$

(c) $\frac{\partial^2 f}{\partial x_1^2} > 0$ and $\Delta_f > 0 \Rightarrow f$ is strictly convex.

(d) $\frac{\partial^2 f}{\partial x_1^2} < 0$ and $\Delta_f > 0 \Rightarrow f$ is strictly concave.

Alle the above inequalities are supposed to hold at each point of S.

Proof. (a) $\Leftarrow$: Choose two arbitrary vectors $\vec{a}, \vec{b}$ in S, and let $t \in [0,1]$. We consider the function g given by

$$g(t) = f(\vec{b} + t(\vec{a} - \vec{b})) = f(t\vec{a} + (1-t)\vec{b}).$$

Put $\vec{r}(t) = \vec{b} + t(\vec{a} - \vec{b})$, thus $g(t) = f(\vec{r}(t))$. Let $\vec{a} = (a_1, a_2), \vec{b} = (b_1, b_2)$. The Chain Rule yields

$$\begin{aligned} g'(t) &= \nabla f(\vec{r}(t)) \cdot \vec{r}'(t) \\ &\frac{\partial f}{\partial x_1}(\vec{r}(t)) \cdot (a_1 - b_1) + \frac{\partial f}{\partial x_2}(\vec{r}(t)) \cdot (a_2 - b_2) \end{aligned}$$

and

$$\begin{aligned} g''(t) =& (\frac{\partial^2 f(\vec{r}(t))}{\partial x_1^2}, \frac{\partial^2 f(\vec{r}(t))}{\partial x_2 \partial x_1}) \cdot (\vec{a} - \vec{b})(a_1 - b_1) \\ &+ (\frac{\partial^2 f(\vec{r}(t))}{\partial x_1 \partial x_2}, \frac{\partial^2 f(\vec{r}(t))}{\partial x_2^2}) \cdot (\vec{a} - \vec{b})(a_2 - b_2) \\ =& [\frac{\partial^2 f}{\partial x_1^2}(a_1 - b_1)^2 + \frac{\partial^2 f}{\partial x_2 \partial x_1}(a_1 - b_1)(a_2 - b_2) \\ &+ \frac{\partial^2 f}{\partial x_1 \partial x_2}(a_1 - b_1)(a_2 - b_2) + \frac{\partial^2 f}{\partial x_2^2}(a_2 - b_2)^2]_{\vec{r}(t)} \\ =& (\vec{a} - \vec{b})^t \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix}_{\vec{r}(t)} (\vec{a} - \vec{b}) \end{aligned}$$

From the conditions in (a) the "Hessematrix" H of f,

$$H = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix}$$

is positive semidefinite in S. Accordingly $g''(t) \geq 0$, for all $t \in (0, 1)$, so that g is convex. In particular it follows that

$$\begin{aligned} f(t\vec{a} + (1-t)\vec{b}) &= g(t) = g(t \cdot 1 + (1-t) \cdot 0) \\ &\leq tg(1) + (1-t)g(0) = tf(\vec{a}) + (1-t)f(\vec{b}). \end{aligned}$$

Since $\vec{a}$ and $\vec{b}$ were arbitrary in S, we deduce that f is convex.

$\Rightarrow$: In light of Proposition 4 it suffices to prove that the Hessematrix H of f is positive semidefinite in S. Let $\vec{x} \in S$, $h \in \mathbb{R}^2$ be arbitrary. As S is open, there exists $r > 0$ such that $|t| < r \Rightarrow \vec{x} + t\vec{h} \in S$. Define p on the interval $I = (-r, r)$ ved $p(t) = f(\vec{x} + t\vec{h})$.

Claim: p is a convex function.
In order to see this, let $\lambda \in [0,1]$, t and $s \in I$. Then

$$\begin{aligned} p(\lambda t + (1-\lambda)s)) &= f(\vec{x} + \lambda t\vec{h} + (1-\lambda)s\vec{h}) \\ &= f(\lambda(\vec{x} + t\vec{h}) + (1-\lambda)(\vec{x} + s\vec{h})) \\ &\leq \lambda f(\vec{x} + t\vec{h}) + (1-\lambda)f(\vec{x} + s\vec{h}) \\ &= \lambda p(t) + (1-\lambda)p(s), \end{aligned}$$

which proves the claim.

Since p is C^2 and convex, we deduce $p'' \geq 0$. Let $\vec{r}(t) = \vec{x} + t\vec{h},\ t \in I$. By The Chain Rule,

$$p'(t) = \nabla f(\vec{r}(t)) \cdot \vec{h}$$

and as in the first part of the proof,

$$p''(t) = \vec{h}^t H(\vec{r}(t))\vec{h} \quad (\vec{h} \in \mathbb{R}^2 \text{ arbitrary}).$$

As $p''(t) \geq 0$, it follows that $H(\vec{r}(t))$ is positive semidefinite, $t \in I$. In particular,

$$H(\vec{x}) = H(\vec{r}(0)) \geq 0$$

for all $\vec{x} \in S$. Hence part (a) follows.
The proofs of $(b), (c)$ and (d) are similar. ■

The following "gradient inequality" and its consequenses are useful in the calculus of variations.

Proposition 9. (Gradient inequality) Assume that $S \subseteq \mathbb{R}^n$ is convex, and let $f : S \to \mathbb{R}$ be a C^1-function. Then the following equivalence holds

f is concave $\Leftrightarrow$

$$(1) \quad f(\vec{x}) - f(\vec{y}) \leq \nabla f(\vec{y}) \cdot (\vec{x} - \vec{y}) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{y})(x_i - y_i)$$

for all $\vec{x}, \vec{y} \in S$.

Proof. $\Rightarrow$: Assume that f is concave, and let $\vec{x}, \vec{y} \in S$. For all $t \in (0,1)$: we know that

$$tf(\vec{x}) + (1-t)f(\vec{y}) \leq f(t\vec{x} - (1-t)\vec{y}),$$

that is

$$t(f(\vec{x}) - f(\vec{y})) \leq f(t(\vec{x} - \vec{y}) + \vec{y}) - f(\vec{y})$$

hence

$$f(\vec{x}) - f(\vec{y}) \leq \frac{1}{t}[f(\vec{y} + t(\vec{x} - \vec{y})) - f(\vec{y})]$$

If we let $t \to 0$, we find that the expression to the right approaches the derivative of f at $\vec{y}$ along $\vec{x} - \vec{y}$. Hence,

$$\begin{aligned} f(\vec{x}) - f(\vec{y}) &\leq \lim_{t \to 0} \frac{1}{t}[f(\vec{y} + t(\vec{x} - \vec{y})) - f(\vec{y})] = f'_{\vec{x}-\vec{y}}(\vec{y}) \\ &= \nabla f(\vec{y}) \cdot (\vec{x} - \vec{y}) \end{aligned}$$

$\Leftarrow$: Assume that the inequality (1) holds. Let $\vec{x}, \vec{y} \in S$, and let $t \in (0, 1)$. We put

$$\vec{z} = t\vec{x} + (1 - t)\vec{y}.$$

It is clear that $\vec{z} \in S$, since S is convex. By (1),

$$(2) \qquad f(\vec{x}) - f(\vec{z}) \leq \nabla f(\vec{z}) \cdot (\vec{x} - \vec{z})$$

and

$$(3) \qquad f(\vec{y}) - f(\vec{z}) \leq \nabla f(\vec{z}) \cdot (\vec{y} - \vec{z})$$

Multiplying the inequality in (2) by t and the inequality in (3) by $1 - t$, and then adding the two resulting inequalities, we derive that

$$\begin{aligned} (4) \quad & t(f(\vec{x}) - f(\vec{z})) + (1 - t)(f(\vec{y}) - f(\vec{z})) \\ & \leq \nabla f(\vec{z}) \cdot [t(\vec{x} - \vec{z}) + (1 - t)(\vec{y} - \vec{z})] \end{aligned}$$

Here the right side of (4) equals $\vec{0}$, since

$$\begin{aligned} t(\vec{x} - \vec{z}) + (1 - t)(\vec{y} - \vec{z}) &= t\vec{x} + (1 - t)\vec{y} - t\vec{z} - \vec{z} + t\vec{z} \\ &= t\vec{x} + (1 - t)\vec{y} - \vec{z} = \vec{0}. \end{aligned}$$

Rearranging the inequality (4) we hence find,

$$\begin{aligned} tf(\vec{x}) + (1 - t)f(\vec{y}) &\leq f(\vec{z}) + (1 - t)f(\vec{z}) \\ &= f(\vec{z}) = f(t\vec{x} + (1 - t)\vec{y}), \end{aligned}$$

which shows that f is convex. ■

Remark 3. We also have that f is strictly concave $\Leftrightarrow$ proper inequality holds in (1). Corresponding results with the opposite inequality (respectively the opposite proper inequality) in (1), holds for convex (respectively strongly convex) functions.

Definition 5. Let $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$, and let $\vec{x}^* = (x_1^*, \ldots, x_n^*) \in S$. A vector $\vec{x}^*$ is called a *(global) maximum point* for f if $f(\vec{x}^*) \geq f(\vec{x})$ for all $\vec{x} \in S$. Similarly we define global minimum point. $\vec{x}^*$ is called a *local maximum point* for f if there exists a positive r such that $f(\vec{x}^*) \geq f(\vec{x})$ for all $\vec{x} \in S$ that satisfies $||\vec{x} - \vec{x}^*|| < r$.

A *stationary* (or *critical) point* for f is a point $\vec{y} \in S$ such that $\nabla f(\vec{y}) = \vec{0}$, that is, all the first order partial derivatives of f at the point $\vec{y}$ are zero, $\frac{\partial f}{\partial x_i}(\vec{y}) = 0 \quad (1 \leq i \leq n)$.

If $\vec{x}^*$ is a local maximum- or minimum point in the interior S^0 of S, then $\nabla f(\vec{x}^*) = \vec{0}$:

$$\frac{\partial f}{\partial x_i}(\vec{x}^*) = \lim_{h \to 0^+} \frac{1}{h}[f(\vec{x}^*) - f(\vec{x}^* + h\vec{e}_i)] \geq 0$$

and

$$\frac{\partial f}{\partial x_i}(\vec{x}^*) = \lim_{h \to 0^-} \frac{1}{h}[f(\vec{x}^*) - f(\vec{x}^* + h\vec{e}_i)] \leq 0,$$

and hence $\frac{\partial f}{\partial x_i}(\vec{x}^*) = 0, \quad (1 \leq i \leq n)$. For arbitrary C^1-functions f the condition $\nabla f(\vec{x}^*) = \vec{0}$ is not sufficient to ensure that f has a loal maximum or minimum at $\vec{x}^*$, however, if f is convex or concave, the following holds,

Proposition 10. Let f be a real valued C^1 function defined on a convex subset S of $\mathbb{R}^n$, and let $\vec{x}^* \in S$. The following holds

(a) if f is concave, then $\vec{x}^*$is a local maximum point for f
$\Leftrightarrow \nabla f(\vec{x}^*) = \vec{0}$.

(b) if f is convex, is $\vec{x}^*$ a local minimum point for f
$\Leftrightarrow \nabla f(\vec{x}^*) = \vec{0}$.

Proof. (a)
$\Rightarrow$: We have seen above that if $\vec{x}^*$ is a local maximum point for f, then $\nabla f(\vec{x}^*) = \vec{0}$.

$\Leftarrow$: Assume that $\vec{x}^*$ is a stationary point and that f is concave. For all $\vec{x} \in S$ it follows from the gradient inequality in Proposition 9 that

$$f(\vec{x}) - f(\vec{x}^*) \leq \nabla f(\vec{x}^*) \cdot (\vec{x} - \vec{x}^*) = 0,$$

hence $f(\vec{x}) \leq f(\vec{x}^*)$, for all $\vec{x} \in S$.

the proof of (b) is similar.

The next result will also be useful,

Proposition 11. Assume that F is a real valued function defined on a convex subset S of $\mathbb{R}^n$, and that G is a real function defined on an interval in $\mathbb{R}$ that contains the image $f(S) = \{f(\vec{x}) : \vec{x} \in S\}$ of f. Then the following statements hold,

(a) f is concave and G is concave and increasing $\Rightarrow G \circ f$ is concave.

(b) f is convex and G is convex and increasing $\Rightarrow G \circ f$ is convex.

(c) f is concave and G is convex and decreasing $\Rightarrow G \circ f$ is convex.

(d) f is convex and G concave and decreasing $\Rightarrow G \circ f$ is concave.

Proof. (a) Let $\vec{x}, \vec{y} \in S$, and let $t \in (0,1)$. Put $U = G \circ f$. Then

$$\begin{aligned} U(t\vec{x} + (1-t)\vec{y}) &= G(f(t\vec{x} + (1-t)\vec{y}) \\ &\geq G(tf(\vec{x}) + (1-t)f(\vec{y})) \quad (f \text{ concave and } G \text{ increasing}) \\ &\geq tG(f(\vec{x})) + (1-t)G(f(\vec{y})) \qquad (G \text{ concave}) \\ &= tU(\vec{x}) + (1-t)U(\vec{y}), \end{aligned}$$

hence U is concave.
(b) is proved as (a).
(c) and (d): Apply (a) and (b) to the function $-G$. (Notice that G is convex and decreasing $\Leftrightarrow$ $-G$ is concave and increasing.) ■

3 CALCULUS OF VARIATIONS.

We may say that the basic problem of the calculus of variations is to determine the maximum or minimum of an integral of the form

$$(1) \qquad J(x) = \int_{t_0}^{t_1} F(t,\, x(t),\, \dot{x}(t))\, dt,$$

where F is a given C^2 function of three variables and $x = x(t)$ is an unknown C^2 function on the interval $[t_0, t_1]$, such that

$$(2) \qquad x(t_0) = x_0, \text{ and } x(t_1) = x_1,$$

where x_0 and x_1 are given numbers. Additional side conditions for the problem may also be included. Functions x that are C^2 and satisfy the endpoint conditions (2) are called **admissible** functions.

Example 1. (Minimal surface of revolution.)
Consider curves $x = x(t)$ in the $tx-$plane, $t_0 \leq t \leq t_1$, all with the same given, end points (t_0, x_0), (t_1, x_1). By revolving such curves around the $t-$axis, we obtain a surface of revolution S.
Problem. Which curve x gives the smallest surface of revolution?
If we assume that $x(t) \geq 0$, the area of S is

$$A(x) = \int_{t_0}^{t_1} 2\pi x\sqrt{1+\dot{x}^2}\, dt$$

Our problem is to determine the curve(s) x^* for which $\min A(x) = A(x^*)$, when $x(t_0) = x_0$, $x(t_1) = x_1$. This problem is of the same type as in (1) above.

We shall next formulate the main result for problems of the type (1).

Theorem 1. ("Main Theorem of the Calculus of Variations")
Assume that F is a C^2 function defined on $\mathbb{R}^3$. Consider the integral

$$(*) \qquad \int_{t_0}^{t_1} F(t, x, \dot{x})\, \mathrm{d}t$$

If a function x^* maximizes or minimizes the integral in $(*)$ among all C^2 functions x on $[t_0,\, t_1]$ that satisfy the end point conditions

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

then x^* satisfies the Euler equation

$$(E) \qquad \frac{\partial F}{\partial x}(t, x(t), \dot{x}(t)) - \frac{d}{dt}(\frac{\partial F}{\partial x}(t, x(t), \dot{x}(t)) = 0 \qquad (t \in [t_0, t_1]).$$

If the function $(x, \dot{x}) \mapsto F(t, x, \dot{x})$ is concave (respectively convex) for each $t \in [t_0, t_1]$, then a function x^* that satisfies the Euler equation (E), solves the maximum (respectively minimum) problem.

Solutions of the Euler equations are called **stationary points** or **extremals**. A proof of Theorem 1 will be given below. First we consider an example.

Example 2. Solve

$$\min_x \int_0^1 (x^2 + \dot{x}^2)\, dt, \quad x(0) = 0, x(1) = e^2 - 1.$$

We let $F(t, x, \dot{x}) = x^2 + \dot{x}^2$. Then $\frac{\partial F}{\partial x} = 2x, \quad \frac{d}{dt}\frac{\partial F}{\partial \dot{x}} = \frac{d}{dt}(2\dot{x}) = 2\ddot{x}$. Hence the Euler equation of the problem is

$$\ddot{x} - x = 0,$$

The general solution is

$$x(t) = Ae^t + Be^{-t}$$

Here $x(0) = 0 = A + B, B = -A$, $x(1) = e^2 - 1 = Ae + Be^{-1} = A(e - e^{-1})$. This implies that $A = e, B = -e$, hence

$$x(t) = e^{t+1} - e^{1-t}$$

is the only possible solution, by the first part of Theorem 1. Here $F(t, x, \dot{x}) = x^2 + \dot{x}^2$ is convex as a function of $(x, \dot{x})$, for each t (this follows from the 2nd derivative test, or from the fact that F is a sum of two convex functions each of them of a single variable: $x \mapsto x^2$, and $\dot{x} \mapsto \dot{x}^2$). As a consequence of the last part of Theorem 1, we conclude that x does in fact solve the problem. The minimum is

$$\int_0^1 (x^2 + \dot{x}^2)\, dt = \int_0^1 [(e^{t+1} - e^{1-t})^2 + (e^{t+1} + e^{1-t})^2]\, dt = \ldots = e^4 - 1.$$

We notice that t does not occur explicitly in the formula for F in this example. Next we shall take a closer look at this particular case.

The case $F = F(x, \dot{x})$ (that is, t does not occur explicitly in the formula of the function F). Such problems are often called *autonomous*.

If $x(t)$ is a solution of the Euler equation, then the Chain Rule yields:

$$\begin{aligned}\frac{d}{dt}[f(x, \dot{x}) - \dot{x}\frac{\partial F}{\partial \dot{x}}(x, \dot{x})] &= \frac{\partial F}{\partial x}\dot{x} + \frac{\partial F}{\partial \dot{x}}\ddot{x} - \ddot{x}\frac{\partial F}{\partial \dot{x}} - \dot{x}\frac{d}{dt}\frac{\partial F}{\partial \dot{x}} \\ &= \dot{x}[\frac{\partial F}{\partial x} - \frac{d}{dt}(\frac{\partial F}{\partial \dot{x}})] = \dot{x} \cdot 0 = 0,\end{aligned}$$

so that

$$F - \dot{x}\frac{\partial F}{\partial \dot{x}} = C \quad \text{(is constant)}.$$

This equation is called a first integral of the Euler equation.

In order to prove Theorem 1 we will need,

The Fundamental Lemma (of the Calculus of Variations). Assume that $f : [t_0, t_1] \to \mathbb{R}$ is a continuous function, and that

$$\int_{t_0}^{t_1} f(t)\mu(t)\,\mathrm{d}t = 0$$

for all C^2 functions μ that satisfies $\mu(t_0) = \mu(t_1) = 0$. Then $f(t) = 0$ for all t in the interval $[t_0, t_1]$.

Proof. If there exists an $s \in (t_0, t_1)$ such that $f(s) \neq 0$, say $f(s) > 0$, then $f(t) > 0$ for all t in some interval $[s-\epsilon, s+\epsilon]$ of s since f is continuous. We choose a function μ such that μ is C^2 and

$$\mu \geq 0,\ \mu(s-\epsilon) = \mu(s+\epsilon) = 0,\ \mu > 0 \text{ on } \ (s-\epsilon, s+\epsilon),$$

and $\mu = 0$ outside the interval $(s-\epsilon, s+\epsilon)$. We may for instance let

$$\begin{cases} (t-(s-\epsilon))^3((s+\epsilon)-t)^3,\ y \in [s-\epsilon, s+\epsilon] \\ 0, \qquad t \notin [s-\epsilon, s+\epsilon]. \end{cases}$$

which is C^2. Then

$$\int_{t_0}^{t_1} f(t)\mu(t)\,\mathrm{d}t = \int_{s-\epsilon}^{s+\epsilon} f(t)\mu(t)\,\mathrm{d}t > 0.$$

since $\mu(t)f(t) > 0$ and $\mu(t)f(t)$ is continuous on the open interval $(s-\epsilon, s+\epsilon)$. Hence we have obtained a contradiction. ■

We shall also need to "differentiate under the integral sign":

Proposition 2. Let g be a C^2 function defined on the rectangle $R = [a, b] \times [c, d]$ i $\mathbb{R}^2$, and let

$$G(u) = \int_a^b g(t, u)\,\mathrm{d}t, \qquad u \in [c, d].$$

Then

$$G'(u) = \int_a^b \frac{\partial g}{\partial u}(t,u)\,\mathrm{d}t.$$

Proof. Let $\epsilon > 0$. Since $\frac{\partial g}{\partial u}$ is continuous, it is also uniformly continuous on R (R is closed and bounded, hence compact). Therefore there exists a $\delta > 0$ such that

$$|\frac{\partial g}{\partial u}(s,u) - \frac{\partial g}{\partial u}(t,v)| < \frac{\epsilon}{b-a},$$

for all (s,u), (t,v) in R such that $||(s,u)-(t,v)|| < \delta$. By the Mean value Theorem there is a $\theta = \theta(t,u,v)$ between u and v such that

$$g(t,v) - g(t,u) = \frac{\partial g}{\partial u}(t,\theta)(v-u).$$

Consequently,

$$\begin{aligned}
&|\int_a^b [\frac{g(t,v)-g(t,u)}{v-u} - \frac{\partial g}{\partial u}(t,u)]\,\mathrm{d}t| \\
&\leq \int_a^b |\frac{\partial g}{\partial u}(t,\theta) - \frac{\partial g}{\partial u}(t,u)|\,\mathrm{d}t \\
&\leq \int_a^b \frac{\epsilon}{b-a}\,\mathrm{d}t = \epsilon, \qquad \text{for } |v-u| < \delta.
\end{aligned}$$

Hence

$$\begin{aligned}
G'(u) &= \lim_{v\to u} \frac{G(v)-G(u)}{v-u} = \lim_{v\to u} \int_a^b \frac{g(t,v)-g(t,u)}{v-u}\,\mathrm{d}t \\
&= \int_a^b \frac{\partial g}{\partial u}(t,u)\,\mathrm{d}t.
\end{aligned}$$

■

Proof of Theorem 1. We will consider the maximum problem (the minimum problem is similar)

$$(1) \quad \begin{cases} \max\limits_x \int_{t_0}^{t_1} F(t,x(t),\dot{x}(t))\,\mathrm{d}t \\ \text{when } \ x(t_0) = x_0 \text{ and } \ x(t_1) = x_1, \ \text{(here } x_0, \text{ and } x_1 \ \text{ are given numbers.} \end{cases}$$

assume that x^* is a C^2 function that solves (1). Let μ be an arbitrary C^2 function on $[t_0, t_1]$ that satisfies $\mu(t_0) = \mu(t_1) = 0$. For each real ϵ the function

$$x = x^* + \epsilon\mu$$

is *an admissible function*, that is, a C^2 function satisfying the endpoint conditions $x(t_0) = x_0$ and $x(t_1) = x_1$. Therefore we must have

$$J(x^*) \geq J(x + \epsilon\mu)$$

for all real ϵ. Let μ be fixed. We will study $I(\epsilon) = J(x^* + \epsilon\mu)$ as a function of ϵ. The function I has a maximum for $\epsilon = 0$. Hence

$$I'(0) = 0.$$

Now

$$I(\epsilon) = \int_{t_0}^{t_1} F(t, x^*(t) + \epsilon\mu(t), \dot{x}^*(t) + \epsilon\dot{\mu}(t))\,\mathrm{d}t$$

Differentiating under the integral sign (see Proposition 2 above) with respect to ϵ we find by the Chain Rule,

$$0 = I'(0) = \int_{t_0}^{t_1} \big[\frac{\partial F^*}{\partial x}\mu(t) + \frac{\partial F^*}{\partial \dot{x}}\dot{\mu}(t)\big]\,\mathrm{d}t$$

where we put $\frac{\partial F^*}{\partial x} = \frac{\partial F}{\partial x}(t, x^*(t), \dot{x}^*(t))$ and $\frac{\partial F^*}{\partial \dot{x}} = \frac{\partial F^*}{\partial \dot{x}}(t, x^*(t), \dot{x}^*(t))$. Hence

$$\begin{aligned} 0 = I'(0) &= \int_{t_0}^{t_1} \frac{\partial F^*}{\partial x}\mu(t)\,\mathrm{d}t + \int_{t_0}^{t_1} \frac{\partial F^*}{\partial \dot{x}}\dot{\mu}(t)\,\mathrm{d}t \\ &= \int_{t_0}^{t_1} \frac{\partial F^*}{\partial x}\mu(t)\,\mathrm{d}t + \Big[\frac{\partial F^*}{\partial x}\mu(t)\Big]_{t_0}^{t_1} - \int_{t_0}^{t_1} \frac{\mathrm{d}}{\mathrm{d}t}\big(\frac{\partial F^*}{\partial \dot{x}}\big)\mu(t)\,\mathrm{d}t \quad \text{(using integration by parts)} \\ &= \int_{t_0}^{t_1} \big[\frac{\partial F^*}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\big(\frac{\partial F^*}{\partial \dot{x}}\big)\big]\mu(t)\,\mathrm{d}t \quad (\text{since } \mu(t_1) = \mu(t_0) = 0) \end{aligned}$$

hence

$$\int_{t_0}^{t_1} \big[\frac{\partial F^*}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\big(\frac{\partial F^*}{\partial \dot{x}}\big)\big]\mu(t)\,\mathrm{d}t = 0$$

for all such C^2 functions μ with $\mu(t_1) = \mu(t_1) = 0$. By the Fundamental Lemma it follows that

$$\frac{\partial F^*}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\big(\frac{\partial F^*}{\partial \dot{x}}\big) = 0, \qquad (2)$$

hence x^* satisfies the Euler equation. Using a quite similar argument we find that an optimal solution x^* for the minimum problem also satisfies (2).

Sufficiency: assume that for every t in $[t_0, t_1]$, $F(t, x, \dot{x})$ is concave with respect to the last two variables $(x, \dot{x})$, and let x^* satisfy the Euler equation with the endpoint conditions of (1). We shall prove that x^* solves the maximum problem. Let x be an arbitrary admissible function for the problem. By concavity the "Gradient inequality" yields,

$$F(t, x, \dot{x}) - F(t, x^*, \dot{x}^*) \leq \frac{\partial F^*}{\partial x}(x - x^*) + \frac{\partial F^*}{\partial \dot{x}}(\dot{x} - \dot{x}^*)$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial F^*}{\partial \dot{x}}\right)(x - x^*) + \frac{\partial F^*}{\partial \dot{x}}(\dot{x} - \dot{x}^*) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial F^*}{\partial \dot{x}}(x - x^*)\right),$$

for all $t \in [t_0, t_1]$. By integration of the inequality we deduce that

$$\int_{t_0}^{t_1} \big(F(t, x, \dot{x}) - F(t, x^*, \dot{x}^*)\big)\,\mathrm{d}t$$
$$\leq \int_{t_0}^{t_1} \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial F^*}{\partial \dot{x}}\Big)(x - x^*)\,\mathrm{d}t = \left[\frac{\partial F^*}{\partial \dot{x}}(x - x^*)\right]_{t_0}^{t_1} = 0,$$

where we used the endpoint conditions for x and x^* at the last step. Hence

$$\int_{t_0}^{t_1} F(t, x, \dot{x})\,\mathrm{d}t \leq \int_{t_0}^{t_1} F(t, x^*, \dot{x}^*)\,\mathrm{d}t$$

as we wanted to prove. ■

Remark. In the above theorem, assume instead that F is concave with respect to $(x, \dot{x})$ (for each $t \in [t_0, t_1]$) only in a certain open and convex subset R of $\mathbb{R}^2$. The above sufficiency proof still applies to the set V of all admissible functions x enjoying the property that $(x(t), \dot{x}(t)) \in R$ for all $t \in [t_0, t_1]$. Hence any admissible solution x^* of the Euler equation that is also an element of V will maximize the integral $\int_{t_0}^{t_1} F(t, x, \dot{x})\,\mathrm{d}t$ among all members of V. Consequently, we obtain a maximum relative to the set V. Note that the proof uses the "Gradient Inequality" (Proposition 9 in Section 1) which requires the region to be open and convex.

Example 3. We will find a solution $x = x(t)$ of the problem

$$\min \int_{-1}^{1} F(t, x, \dot{x})\,\mathrm{d}t$$

where $F(t,x,\dot{x}) = t^2\dot{x}^2 + 12x^2$ and $x(-1) = -1$, $x(1) = 1$.

(**a**) First we will solve the problem by means of Theorem 1. Here we find

$$\frac{\partial F}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial F}{\partial \dot{x}}\Big) = 24x - \frac{\mathrm{d}}{\mathrm{d}t}\big(t^2 \cdot 2\dot{x}\big) = 24x - 4t\dot{x} - 2t^2\ddot{x},$$

and the Euler equation takes the form

$$t^2\ddot{x} + 2t\dot{x} - 12x = 0.$$

We seek solutions of the type $x(t) = t^k$. This gives

$$(k(k-1) + 2k - 12)t^k = 0, \text{ for all } t,$$

that is, $k^2 + k - 12 = 0$, which has the solutions $k = 3$ and $k = -4$. Consequently, $x(t) = At^3 + Bt^{-4}$ where

$$\begin{aligned} x(-1) &= -A + B = -1 \\ x(1) &= A + B = 1 \end{aligned}$$

Hence $B = 0$ and $A = 1$, so that $x(t) = t^3$ is the only possible solution. Furthermore,

$$\frac{\partial^2 F}{\partial \dot{x}^2} = 2t^2 \geq 0, \; \frac{\partial^2 F}{\partial x^2} = 24 > 0, \; \frac{\partial^2 F}{\partial x \partial \dot{x}} = 0,$$

and the determinant of the Hesse matrix is $24 \cdot 2t^2 = 48t^2 \geq 0$. It follows that F is convex with respect to $(x, \dot{x})$. Hence, in view of Theorem 1, $x(t) = t^3$ gives a minimum. ■

(**b**) Suppose next that x is a solution of the Euler equation to the problem which in addition satisfies the given endpoint conditions. If z is an arbitrary admissible function, then $\mu = z - x$ is a C^2 function that satisfies $\mu(-1) = \mu(1) = 0$. Let us show that

$$J(z) - J(x) = J(x + \mu) - J(x) \geq 0.$$

From this it will follow that x minimizes J. Now

$$\begin{aligned} J(x+\mu) - J(x) &= \int_{-1}^{1} [t^2(\dot{x} + \dot{\mu})^2 + 12(x+\mu)^2 - t^2\dot{x}^2 - 12x^2]\,\mathrm{d}t \\ &= \int_{-1}^{1} [2t^2\dot{x}\dot{\mu} + t^2\dot{\mu}^2 + 24x\mu + 12\mu^2]\,\mathrm{d}t = \int_{-1}^{1} [t^2\dot{\mu}^2 + 12\mu^2]\,\mathrm{d}t + I \end{aligned}$$

where

$$
\begin{aligned}
I &= \int_{-1}^{1} [2t^2\dot{x}\dot{\mu} + 24x\mu]\,\mathrm{d}t \\
&= \Big|_{-1}^{1} 2t^2\dot{x}\mu - \int_{-1}^{1} (2t^2\ddot{x}\mu + 4t\dot{x}\mu)\,\mathrm{d}t + \int_{-1}^{1} 24x\mu\,\mathrm{d}t \quad \text{(using integration by parts)} \\
&= 0 - \int_{-1}^{1} 2\mu[t^2\ddot{x} + 2t\dot{x} - 12x]\,\mathrm{d}t \quad \big(\mu(-1) = \mu(1) = 0\big) \\
&= 0,
\end{aligned}
$$

where the last inequality follows from the fact that x satisfies the Euler equation

$$t^2\ddot{x} + 2t\dot{x} - 12x = 0.$$

Hence

$$J(x+\mu) - J(x) = \int_{-1}^{1} [t^2\dot{\mu}^2 + 12\mu^2]\,\mathrm{d}t > 0$$

if $\mu \neq 0$, hence every solution of the Euler equation that satisfies the endpoint conditions, yields a minimum. We have seen in (a) that $x(t) = t^3$ is the only such solution. ■

Example 4. (No minimum)
Consider the minimum problem

$$\min \int_0^4 (\dot{x}(t)^2 - x(t)^2)\,\mathrm{d}t, \quad x(0) = 0,\ x(4) = 0.$$

The Euler equation

$$(E) \qquad \ddot{x} + x = 0$$

has the characteristic roots $\pm i$, hence the general solution of (E) is

$$x(t) = A\cos t + B\sin t$$

The endpoint conditions yield $A = B = 0$, hence the unique solution is

$$x^*(t) = 0 \quad (\text{for all } t \in [0,4])$$

Consequently, if the minimum exists, then it is equal to zero and is attained at $x^* = 0$. Let us show that the minimum does not exist. To that end, we consider the function

$$x(t) = \sin(\frac{\pi t}{4})$$

which is clearly C^2. Moreover, $x(0) = x(4) = 0$, so that x is admissible. Now

$$\int_0^4 (\dot{x}^2 - x^2)\,\mathrm{d}t = \int_0^4 [\frac{\pi^2}{16}\cos^2(\frac{\pi t}{4}) - \sin^2(\frac{\pi t}{4})]\,\mathrm{d}t$$
$$= \int_0^4 (\frac{\pi^2}{16} - (\frac{\pi^2}{16}+1)\sin^2\frac{\pi t}{4})\,\mathrm{d}t = \frac{\pi^2}{4} - (\frac{\pi^2}{16}+1)\int_0^4 \frac{1}{2}(1-\cos\frac{\pi t}{2})\,\mathrm{d}t$$
$$= \frac{\pi^2}{4} - (\frac{\pi^2}{16}+1)\frac{1}{2}\big[{}_0^4\, t - \frac{2}{\pi}\sin\frac{\pi t}{2}\big] = \frac{\pi^2}{4} - (\frac{\pi^2}{16}+1)\frac{1}{2}[4-0] = -\frac{\pi^2}{8} - 2 < 0$$

Therefore, $x^* = 0$ does not minimize the integral and no minimum exists.

Other endpoint conditions. Next we will consider optimation problems for which the left end point $x(t_0)$ of the admissible functions x is fixed, whereas the right end point $x(t_1)$ is free. In this case we have the following result:

Theorem 2. Assume that x_0 is a given number. A necessary condition for a C^2 function[2] x^* to solve the problem

$$\max\ (\min) \int_{t_0}^{t_1} F(t, x(t), \dot{x}(t))\,\mathrm{d}t, \quad x(t_0) = x_0,\ x(t_1) \text{ is free},$$

is that x^* satisfies the Euler equation and, in addition, the condition

$$(T) \qquad \big(\frac{\partial F}{\partial \dot{x}}\big)_{t=t_1} = 0.$$

If the function $F(t, x, \dot{x})$ is concave (respectively convex) in $(x, \dot{x})$ for each $t \in [t_0, t_1]$, then any admissible function x^* that satisfies the Euler equation and the condition (T), solves the maximum (respectively minimum) problem.

The condition (T) is called the *transversality condition.*

Proof. We will give a proof of the maximum problem. The proof of the minimum problem is similar. Assume that x^* solves the problem. In particular all admissible (C^2) functions x which have the same value $x^*(t_1)$ as x^* at the right endpoint $t = t_1$, satisfy $J(x) \le J(x^*)$, hence x^* is optimal among those functions. However, then x^* satisfies the Euler equation. In the

[2]As can be shown, it suffices to consider C^1-functions. In fact, even piecewise C^1-functions will do.

proof of the Euler equation, see the proof of Theorem 1, we considered the function

$$I(\alpha) = \int_{t_0}^{t_1} F(t, x^* + \alpha\mu, \dot{x}^* + \alpha\dot{\mu})\,\mathrm{d}t$$

After integration by parts and differentiation under the integral, we concluded that

$$0 = I'(0) = \int_{t_0}^{t_1} [\frac{\partial F^*}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}(\frac{\partial F^*}{\partial \dot{x}})]\,\mathrm{d}t + \big[\frac{\partial F^*}{\partial \dot{x}}\mu(t)\big]_{t_0}^{t_1},$$

where

$$(*) \quad \frac{\partial F^*}{\partial x} = \frac{\partial F}{\partial x}(t, x^*(t), \dot{x}^*(t)) \text{ and } \frac{\partial F^*}{\partial \dot{x}} = \frac{\partial F}{\partial \dot{x}}(t, x^*(t), \dot{x}^*(t))$$

Here, at this point, we let μ be a C^2 function such that $\mu(t_0) = 0$ and $\mu(t_1)$ is free. Since x^* satisfies the Euler equation, the integral in $(*)$ must be equal to zero, hence

$$\mu(t_1)\big(\frac{\partial F^*}{\partial \dot{x}}\big)_{t=t_1} = 0.$$

If we choose a μ with $\mu(t_1) \neq 0$, it follows that $(\frac{\partial F^*}{\partial \dot{x}})_{t=t_1} = 0$.

Sufficiency: Assume that $F(t, x, \dot{x})$ is concave in $(x, \dot{x})$, and that x^* is an admissible function which satisfies the Euler equation and the condition (T). The argument led to the inequality

$$\begin{aligned}&\int_{t_0}^{t_1} \big(F(t, x, \dot{x}) - F(t, x^*, \dot{x}^*)\big)\,\mathrm{d}t \\ &\leq \int_{t_0}^{t_1} \frac{\mathrm{d}}{\mathrm{d}t}\big(\frac{\partial F^*}{\partial \dot{x}}\big)(x - x^*)\,\mathrm{d}t = \left[\frac{\partial F^*}{\partial \dot{x}}(x - x^*)\right]_{t_0}^{t_1}\end{aligned}$$

The last part of the proof of Theorem 1, goes through as before. Here we find

$$\left[\frac{\partial F^*}{\partial \dot{x}}(x - x^*)\right]_{t=t_0} = 0$$

since $x(t_0) = x^*(t_0) = x_0$. Finally the condition (T) yields

$$\left[\frac{\partial F^*}{\partial \dot{x}}(x - x^*)\right]_{t=t_1} = 0.$$

Thus it follows that $J(x) \leq J(x^*)$, hence x^* maximizes J. ■

Using techniques similar to those of the previous proof, we can show

Theorem 3.
Assume that t_0, t_1, x_0, and x_1 are given numbers, and let F be a C^2 function defined on $\mathbb{R}^3$. Consider the integral

$$(*) \qquad \int_{t_0}^{t_1} F(t, x(t), \dot{x}(t))\, \mathrm{d}t$$

If a function x^* maximizes or minimizes the integral in $(*)$ among all C^2 functions x on $[t_0,\, t_1]$ that satisfy the end point conditions

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

then x^* satisfies the Euler equation

$$(E) \qquad \frac{\partial F}{\partial x}(t, x(t), \dot{x}(t)) - \frac{d}{dt}(\frac{\partial F}{\partial x}(t, x(t), \dot{x}(t)) = 0 \qquad (t \in [t_0, t_1]).$$

and the transversality condition

$$(T) \qquad (\frac{\partial F}{\partial \dot{x}})_{t=t_1} \leq 0 \quad (= 0 \text{ if } x(t_1) > x_1).$$

If the function $(x, \dot{x}) \mapsto F(t, x, \dot{x})$ is concave (respectively convex) for each $t \in [t_0, t_1]$, then a function x^* that satisfies the Euler equation (E), solves the maximum (respectively minimum) problem.

Exercise 1. Let

$$F(t, x, \dot{x}) = x^2 + 7x\dot{x} + \dot{x}^2 + t\dot{x}$$

(**a**) Find the Euler equation (E) of the function F. Is F concave or convex as a function of the last two variables $(x, \dot{x})$?

(**b**) Find the general solution of (E).

(**c**) Express the integral

$$\int_{t_0}^{t_1} x(t)\dot{x}(t)\, \mathrm{d}t$$

as a function of $x(t_0)$ and $x(t_1)$. Find a solution x of the problem

$$\min \int_0^1 F(t, x(t), \dot{x}(t))\, \mathrm{d}t, \quad x(0) = \frac{1}{2},\, x(1) = \frac{3}{2}$$

Exercise 2. Find a function $F(t, x, \dot{x})$ which has the given differential equation as an Euler equation:

(**a**)

$$\ddot{x} - x = 0$$

(**b**)

$$t^2\ddot{x} + 2t\dot{x} - x = 0$$

(**c**) Decide if the functions F of (**a**) and (**b**) are concave or convex as functions of $(x, \dot{x})$ for fixed t.

(**d**) Find the only C^2-function x^* of the variation problem

$$\min \int_0^1 \big(\frac{1}{2}x^2 + \frac{1}{2}\dot{x}^2\big)\, \mathrm{d}t, \quad x(0) = 0,\, x(1) = e - e^{-1},$$

Justify your answer.

(**e**) Let F be as in (**b**) and solve the problem

$$\min \int_0^{\frac{\pi}{\sqrt{3}}} F(t, x(t), \dot{x}(t))\, \mathrm{d}t, \quad x(0) = 1,\, x(\frac{\pi}{\sqrt{3}}) = 2e^{\frac{-\pi}{2\sqrt{3}}}.$$

Exercise 3. Consider the problem $\min J(x)$, where

$$J(x) = \int_1^2 \dot{x}(1 + t^2\dot{x})\, \mathrm{d}t, \quad x(1) = 3, x(2) = 5.$$

(a) Find the Euler equation of the problem and show that it has exactly one solution x that satisfies the given endpoint conditions.

(b) Apply Theorem 1 to prove that the solution from (a) really solves the minimum problem.

(c) For an arbitrary C^2 function μ that satisfies $\mu(1) = \mu(2) = 0$, show that

$$J(x + \mu) - J(x) = \int_1^2 t^2\dot{\mu}^2\, \mathrm{d}t,$$

where x is the solution from (a). Explain in light of this, that x minimizes J.

Exercise 4. A function F is given by

$$F(x,y) = x^2(1+y^2), \qquad (x,y) \in \mathbb{R}^2$$

(**a**) Show that F is convex in the region

$$R = \{(x,y) \,:\, |y| \leq \frac{1}{\sqrt{3}}\}$$

(**b**) Show that the variation problem

$$\begin{aligned} &\min \int_0^1 x^2(1+\dot{x}^2)\,\mathrm{d}t, \\ &x(0) = x(1) = 1 \end{aligned}$$

has Euler equation

$$(*) \qquad\qquad x\ddot{x} + \dot{x}^2 - 1 = 0$$

(**c**) Find the only possible solution x^* of the variation problem. Let

$$\begin{aligned} V = \{x \,:\, &x \text{ is a } C^2 \text{ function such that } x(0) = x(1) = 1 \\ &\text{and } |\dot{x}(t)| < \frac{1}{\sqrt{3}} \text{ for all } t \in [0,1]\} \end{aligned}$$

Show that $x^* \in V$. Explain that x^* minimizes the integral in (1) among all functions in V.

Exercise 5. (Reduction of order)
Suppose that one solution x_1 of the homogeneous linear differential equation

$$(*) \qquad \ddot{x} + p(t)\dot{x} + q(t)x = 0$$

is known on an interval I where p and q are continuous. Substitute

$$x_2(t) = u(t)x_1(t)$$

into the equation $(*)$ to deduce that x_2 is a solution of $(*)$ if

$$(**) \qquad x_1\ddot{u} + (2\dot{x_1} + px_1)\dot{u} = 0$$

Since $(**)$ is a separable first order equation in $v = \dot{u}$, it is readily solved for the derivative $\dot{u}$ of u and hence for u. This is called the method of **reduction of order**. As u is nonconstant, the general solution of $(*)$ is

$$x = c_1 x_1 + c_2 x_2 \qquad (c_1, c_2 \text{ arbitrary constants}).$$

Exercise 6. Let $\alpha > -1$, and put

$$(*) \qquad F(t, x, \dot{x}) = \frac{1}{2}(1 - t^2)\dot{x}^2 - \frac{1}{2}\alpha(\alpha + 1)x^2 + tx + \frac{1}{2}t^2\dot{x}$$

(**a**) Show that the Euler equation of F is

$$(**) \qquad (1 - t^2)\ddot{x} - 2t\dot{x} + \alpha(\alpha + 1)x = 0$$

This is known as the Legendre equation of order α and is important in many applications.

(**b**) Let $\alpha = 1$ in $(**)$. Show that in this case the function $x_1(t) = t$ is a solution of the equation $(**)$. Then use the method of reduction of order (Exercise 5) to derive the second solution

$$x_2(t) = 1 - \frac{1}{2}t \ln |\frac{1 + t}{1 - t}|, \quad \text{for } t \neq \pm 1.$$

(**c**) Find all possible C^2-solutions of the variation problem

$$\max \int_2^3 F(t, x, \dot{x})\, \mathrm{d}t, \quad x(2) = 0,\ x(3) = \frac{3}{2}\ln\frac{3}{2} - \frac{1}{2}.$$

(**d**) Find the optimal x of the problem in (**c**)

Exercise 7. Find the unique C^2- solution x of the minimum problem

$$\min \int_1^e (x^2 + tx\dot{x} + t^2\dot{x}^2)\, \mathrm{d}t, \qquad x(1) = 0,\ x(e) = e^{r_1} - e^{r_2},$$

where $r_1 = \frac{1}{2}(-1 + \sqrt{3})$ and $r_2 = \frac{1}{2}(-1 - \sqrt{3})$

Exercise 8. Let

$$F(t, x, \dot{x}) = x^2 + \frac{1}{2}t(t - 1)\dot{x}^2,$$

and consider the problem

$$(1) \quad \min \int_2^3 F(t, x(t), \dot{x}(t))\, \mathrm{d}t, \quad x(2) = 0,\ x(3) = \ln \frac{25}{27}.$$

(**a**) Find the Euler equation (E) of the problem in (1).

(**b**) Show that (E) has a polynomial solution x_1.

(**c**) Use reduction of order (see Exercise 5) to find another solution x_2 of (E) such that $x_2(t) = v(t)x_1(t),\ t \in [2, 3]$.

(**d**) What is the general solution of (E) ? Find the only solution x^* of (E) that satisfies the endpoint conditions in (1).

(**e**) Decide if x^* minimizes the integral in (1).

Remark. The Euler equation (E) of Exercise 8 (a) is a special case of Gauss' hypergeometric differential equation

$$(G) \quad t(1-t)\ddot{x} + [c - (a+b+1)t]\dot{x} - abx = 0$$

where a, b, and c are constants. Hypergeometric differential equations occur frequently in both mathematics and physics. In particular, they are obtained from certain interesting partial differential equations by separation of variables.

Exercise 9. In the followig we shall study maxima and minima of

$$(*) \qquad J(x) = \int_1^2 F(t, x(t), \dot{x}(t))\, \mathrm{d}t, \qquad x(1) = -\frac{1}{3},\ x(2) \text{ is free .}$$

Here

$$F(t, x, \dot{x}) = \frac{1}{2}\dot{x}^2 + \frac{1}{6}g(t)\dot{x}^3,$$

and g is a positive $(g > 0)$ C^2-function on $[1, 2]$.

(**a**) Find the Euler equation (E) of the problem. Express $\dot{x}$ in terms of g.

From now on we let

$$g(t) = \frac{1}{t^2}, \quad \text{for all } t \in [1, 2].$$

(**b**) Find the general solution of (E) in this special case.

(**c**) Show that (E) has exactly two solutions x_1 and x_2 which satisfy the endpoint conditions of $(*)$.

(**d**) Consider the two following convex sets of functions,

$$R_1 = \{x \,:\, x \in C^2[1,2], x(1) = -\frac{1}{3}, x(2) \text{ is free, } \dot{x}(t) > -1 \quad (\text{for all } t \in [1,2])\}$$

and

$$R_2 = \{x \,:\, x \in C^2[1,2], x(1) = -\frac{1}{3}, x(2) \text{ is free, } \dot{x}(t) < -1 \quad (\text{for all } t \in [1,2])\}$$

Decide if the function $(x, \dot{x}) \mapsto F(t, x, \dot{x})$ $(t \in [1,2]$ arbitrary, fixed) is concave or convex on R_1 or on R_2. Does $J(x)$ have a maximum or a minimum on any of the two sets R_1, R_2 ? Find the maximum and the minimum of $J(x)$ on R_1 and on R_2 if they exist.

Remark. The Euler equation of Exercis 9 (a) (with $y = \dot{x}$) is a simple special case of an Abel differential equation of the 2nd kind:

$$(1) \qquad (y+s)\dot{y} + (p + qy + ry^2) = 0,$$

where p, q, r, and s are C^1-functions of a real variable on an open interval I. Such equations can always be reduced to the simpler form

$$(2) \qquad z\dot{z} + (P + Qz) = 0$$

via a substitution $y = \alpha + \beta z$, where α and β are C^1-functions on I. In fact, we may let

$$\alpha = -s, \beta(t) = e^{-\int r(t)\,\mathrm{d}t}, P = (p - qs + rs^2)e^{2\int r(t)\,\mathrm{d}t}, Q = (q - 2rs - \dot{s})e^{\int r(t)\,\mathrm{d}t}.$$

The simple special case (E) of Exercise 9 (a) may be solved without such a reduction. Abel studied equation (1) in the 1820-s and he made the mentioned reduction to (2). (Œuvres complétes de Niels Henrik Abel, Tome Second, p. 26-35.)

Weierstrass' sufficiency condition.[3]

Example 5. Consider the problem of maximizing or minimizing the integral

$$J(x) = \int_0^1 (\dot{x}^2 - x^2)\,\mathrm{d}t, \quad x(0) = x(1) = 0.$$

[3]This section is optional.

The Euler equation of the problem is

$$\ddot{x} + x = 0$$

The extremals (that is, the solutions of the Euler equation) are

$$x(t) = A \sin t + B \cos t, \quad A \text{ and } B \text{ arbitrary constants.}$$

Among all extremals only the zero function x^*, $x^*(t) = 0$ for all $t \in [0, 1]$, satisfies the endpoint conditions. It is easy to see, using the second-derivative test, that $F(t, x, \dot{x}) = \dot{x}^2 - x^2$ is neither concave nor convex i $(x, \dot{x})$. Hence we cannot use Theorem 1 to decide if x^* gives a maximum or a minimum for the problem.

Exercise 10. Show that the function $F(t, x, \dot{x}) = \dot{x}^2 - x^2$ of the last Example is neither convex nor concave.

The sufficiency condition stated in Theorem 3 below, is often useful for solving problems of the type

$$(1) \quad \max \text{ or } \min J(x)$$

where

$$J(x) = \int_{t_0}^{t_1} F(t,\, x(t),\, \dot{x}(t))\, dt.$$

As above F is assumed to be a given C^2 function of three variables and $x = x(t)$ is an unknown C^2 function on the interval $[t_0, t_1]$, such that

$$(2) \quad x(t_0) = x_0,\, x(t_1) = x_1,$$

where x_0 and x_1 are given numbers.

The solutions of the Euler equation of this problem (without any endpoint conditions), are called *extremals*.

Definition. we say that an extremal x^* for

$$J(x) = \int_{t_0}^{t_1} F(t,\, x(t),\, \dot{x}(t))\, dt, \quad x(t_0) = x_0,\, x(t_1) = x_1,$$

is a *relative maximum* if there exists an $r > 0$ such that $J(x^*) \geq J(x)$ for all C^2 functions x with $x(t_0) = x_0$, $x(t_1) = x_1$ that satisfies $|x^*(t) - x(t)| < r$ for all $t \in [0, 1]$. A relative minimum x^* is defined similarly.

Theorem 4. (Weierstrass' sufficiency condition). Assume that an extremal x^* for the problem (1) satisfies the endpoint conditions (2) and that there exists a parameter family $x(\cdot,\alpha)$ [4], $-\epsilon < \alpha < \epsilon$, of extremals such that

(1) $x(t,0) = x^*(t), \quad t \in [t_0, t_1]$,

(2) $x(\cdot,\alpha)$ is differentiable with respect to α and $\dfrac{\partial x}{\partial \alpha}(t,\alpha)\big|_{\alpha=0} \neq 0$, $t_0 \leq t \leq t_1$,

(3) if $\alpha_1 \neq \alpha_2$, then the two curves given by $x(\cdot,\alpha_1)$ and $x(\cdot,\alpha_2)$have no common points,

(4) $\dfrac{\partial^2 F}{\partial \dot{x}^2}(t,x,\dot{x}) < 0$ for all $t \in [t_0,t_1]$ and all $x = x(\cdot,\alpha),\ \alpha \in (-\epsilon,\epsilon)$

Then x^* is a relative maximum for $J(x)$ with the given endpoint conditions. If the condition

(4′) $\dfrac{\partial^2 F}{\partial \dot{x}^2}(t,x,\dot{x}) > 0$ for all $t \in [t_0,t_1]$ and all $x = x(\cdot,\alpha),\ \alpha \in (-\epsilon,\epsilon)$,

holds instead of (4), then x^* is a relative minimum.

Example 6. Consider the extremals

$$x(t) = A\cos t + B\sin t, \quad t \in [0,1].$$

of Example 5. The parameter family of extremals $x(\cdot,\alpha)$ given by

$$(*) \qquad x(t,\alpha) = \alpha\cos t,\ \alpha \in (-1,1),$$

satisfies $x(\cdot,0) = x^* = 0$, and is differentiable with respect to α with $\frac{\partial x(t,\alpha)}{\partial \alpha}\big|_{\alpha=0} = \cos t \neq 0$ for all $t \in [0,1]$. Moreover, it is clear that the family of curves given by $(*)$ have no common points. Finally we have

$$\frac{\partial^2 F}{\partial \dot{x}^2}(t,x,\dot{x}) = \frac{\partial}{\partial \dot{x}}(2\dot{x}) = 2 > 0.$$

Hence, by the Weierstrass sufficiency condition, $x^* = 0$ is a relative minimum point for $J(x)$.

[4] For fixed α, $x(\cdot,\alpha)$ denotes the function $t \mapsto x(t,\alpha)$

Exercise 11. Let

$$J(x) = \int_{t_0}^{t_1} \sqrt{1+\dot{x}^2}\,\mathrm{d}t, \quad x(t_0) = x_0, x(t_1) = x_1, \text{ where } t_0 \neq t_1.$$

$J(x)$ gives the arc length of the curve $x(t)$ between the two given points (t_0, x_0) and (t_1, x_1). We wish to determine the curve x^* that yields the minimal arc length.

(**a**) Show that the Euler equation of the problem may be written

$$(E) \qquad \frac{\dot{x}}{\sqrt{1+\dot{x}}} = c, \quad \text{where } c \text{ is a constant different from } \pm 1.$$

(**b**) Show that (E) has a unique solution x^* that satisfies the given endpoint conditions.

(**c**) Show that among the solutions of (E) there exists a parameter family given by $x(t, \alpha) = x^*(t) + \alpha,\ -1 < \alpha < 1$, where x^* is as in (b), and that the conditions of Theorem 3 are satified such that x^* gives a (relative) minimum.

(**d**) Show that $F(t, x, \dot{x})$ is convex i $(x, \dot{x})$. Conclude by Theorem 1 that x^* solves the minimum problem.

Exercise 12.

(**a**) Verify that $x_1(t) = t^{-1}\cos t$ is a solution of

$$(*) \qquad t\ddot{x} + 2\dot{x} + tx = 0, \qquad (t \neq 0).$$

(**b**) Find another solution x_2 of the equation in $(*)$ of the form

$$x_2 = ux_1,$$

using reduction of order (see Exercise 5).

Next let

$$F(t, x, \dot{x}) = \frac{1}{2}t^2\dot{x}^2 + \frac{1}{3}t^3x\dot{x},$$

and consider the problem

$$(**) \qquad \min \int_{\frac{\pi}{6}}^{\frac{\pi}{2}} F(t, x, \dot{x})\,\mathrm{d}t,\ x(\frac{\pi}{6}) = 0,\ x(\frac{\pi}{2}) = 1.$$

(**c**) Write down the Euler equation of $(**)$ and find its general solution.

(**d**) Find the only possible solution x^* of the variation problem in $(**)$ and decide if x^* is optimal.

4 OPTIMAL CONTROL THEORY.

In the Calculus of Variations we studied problems of the type

$$\max/\min \int_{t_0}^{t_1} f(t, x(t), \dot{x}(t))\,\mathrm{d}t, \quad x(t_0) = x_0,\ x(t_1) = x_1 \text{ (or } x(t_1) \text{ free)}.$$

We assumed that all the relevant functions were of class C^2 (twice continuously differentiable). If we let

$$u = \dot{x},$$

then the above problem may be reformulated as

$$(0) \qquad \max/\min \int_{t_0}^{t_1} f(t, x(t), u(t))\,\mathrm{d}t, \quad \dot{x} = u,\ x(t_0) = x_0, x(t_1) = x_1 (\text{or } x(t_1) \text{ free}).$$

We shall next study a more general class of problems, problems of the type:

$$(1) \qquad \begin{cases} \max/\min \int_{t_0}^{t_1} f(t, x(t), u(t))\,\mathrm{d}t, \quad \dot{x}(t) = g(t, x(t), u(t)), \\ x(t_0) = x_0,\ x(t_1) \text{ free}, \\ u(t) \in \mathbb{R}, t \in [t_0, t_1],\ u \text{ piecewise continuous}. \end{cases}$$

Later we shall also consider such problems with other endpoint conditions and where the functions u can take values in more general subsets of $\mathbb{R}$.

Such problems are called **control problems**, u is called a **control variable** (or just a control). The control region is the common range of the possible controls u.[5] Pairs of functions (x, u) that satisfies the given endpoint conditions and, in addition, the **equation of state** $\dot{x} = g(t, x, u)$, are called **admissible pairs**.

We shall always assume that the controls u are piecewise continuous, that is, u may possess a finite number of jump discontinuities. (In more advanced texts measurable controls are considered.) In the calculus of variations we always assumed that $u = \dot{x}$ was of class C^1.

In the optimal control theory it proves very useful to apply an auxiliary function H of four variables defined by

$$H(t, x, u, p) = f(t, x, u) + pg(t, x.u),$$

called the **Hamilton function** or **Hamiltonian** of the given problem.

[5] Variable control regions may also be considered

Pontryagin's Maximum principle gives conditions that are necessary for an admissible pair (x^*, u^*) to solve a given control problem:

Theorem (The Maximum Principle I)

Assume that (x^*, u^*) is an optimal pair for the problem in (1). Then there exist a continuous function $p = p(t)$, such that for all $t \in [t_0, t_1]$, the following conditions are satisfied:

(a) $u^*(t)$ maximizes $H(t, x^*(t), u, p(t)),\ u \in \mathbb{R}$, that is,

$$H(t, x^*(t), u, p(t)) \leq H(t, x^*(t), u^*(t), p(t)), \quad \text{for all } u \in \mathbb{R}.$$

(b) The function p (called the **adjoint function**) satisfies the differential equation

$$\dot{p}(t) = -\frac{\partial H}{\partial x}(t, x^*(t), u^*(t), p(t)) \qquad (\text{written } -\frac{\partial H^*}{\partial x}),$$

except at the discontinuities of u^*.

(c) The function p obeys the condition

$$(T) \qquad p(t_1) = 0 \qquad (\text{the transversality condition})$$

Remark.
For the sake of simplicity we will frequently use the notations

$$\frac{\partial H^*}{\partial x} = \frac{\partial H}{\partial x}(t, x^*(t), u^*(t), p(t))$$

and similarly,

$$f^* = f(t, x^*(t), u^*(t)),\ g^* = g(t, x^*(t), u^*(t)),\ H^* = H(t, x^*(t), u^*(t), p(t))$$

Remark.
For the varition problem in (0), where $\dot{x} = u$, the Hamiltonian H will be particularly simple,

$$H(t, x, u, p) = f(t, x, u) + pu,$$

hence, by the Maximum Principle, it is necessary that

$$(i) \qquad \dot{p}(t) = -\frac{\partial H^*}{\partial x} = -\frac{\partial f^*}{\partial x}$$

Since $u \in \mathbb{R}$, we must have (there are no endpoints to consider here)

$$0 = \frac{\partial H^*}{\partial u} = \frac{\partial f^*}{\partial u} + p(t).$$

or

$$(ii) \qquad p(t) = -\frac{\partial f^*}{\partial u} = -\frac{\partial f^*}{\partial \dot{x}},$$

in order that $u = u^*(t)$ shall maximize H. Equations (ii) and (i) now yield

$$\dot{p}(t) = -\frac{d}{dt}(\frac{\partial f^*}{\partial \dot{x}}) = -\frac{\partial f^*}{\partial x},$$

Hence we have deduced the Euler equation

$$(E) \qquad \frac{\partial f^*}{\partial x} - \frac{d}{dt}(\frac{\partial f^*}{\partial \dot{x}}) = 0$$

As $p(t) = -\frac{\partial f^*}{\partial \dot{x}}$, the condition (T) above yields the "old" transversality condition $(\frac{\partial f^*}{\partial \dot{x}})_{t=t_1} = 0$ from the Calculus of Variations. This shows that The Maximum Principle implies our previous results from The Calculus of Variations. ■

The following theorem of Mangasarian supplements the Maximum Principle with sufficient conditions for optimality.

Mangasarian's Theorem (First version)
Let the notation be as in the statement of the Maximum Principle. If the map $(x, u) \mapsto H(t, x, u, p(t))$ is concave for each $t \in [t_0, t_1]$, then each admissible pair $(x*, u^*)$ that satisfies conditions $(a), (b)$, and (c) of the Maximum Principle, will give a maximum.

Example 1
We will solve the problem

$$\max \int_0^T [1 - tx(t) - u(t)^2]\, \mathrm{d}t, \qquad \dot{x} = u(t), x(0) = x_0, x(T) \text{ free},$$

where x_0 and T are positive constants.
Solution.
The Hamiltonian is

$$H(t, x, u, p) = 1 - tx - u^2 + pu$$

In order that $u = u^*(t)$ shall maximize $H(t, x^*(t), u, p(t))$, it is necessary that

$$\frac{\partial H}{\partial u}(t, x^*(t), u, p(t)) = 0,$$

that is $-2u + p(t) = 0$, or $u = \frac{1}{2}p(t)$. This yields a maximum since $\frac{\partial^2 H^*}{\partial u^2} = -2 < 0$. Hence

$$u^*(t) = p(t)/2, \qquad t \in [0, T].$$

Furthermore,

$$\frac{\partial H}{\partial x} = -t = -\dot{p}(t),$$

so that $\dot{p} = t$, and $p(t) = \frac{1}{2}t^2 + A$. In addition, the condition (T) gives $0 = p(T) = \frac{1}{2}T^2 + A$, hence $A = -\frac{1}{2}T^2$. Thus

$$p(t) = \frac{1}{2}(t^2 - T^2)$$

and

$$\dot{x}^*(t) = u^*(t) = \frac{1}{2}p(t) = \frac{1}{4}(t^2 - T^2),$$

$$x^*(t) = \frac{1}{12}t^3 - \frac{1}{4}T^2 t + x_0.$$

Here $H(t, x, u, p(t)) = 1 - tx - u^2 + pu$ is concave in (x, u) for each t, being the sum of the two concave functions $(x, u) \mapsto pu - tx$ (which is linear) and $(x, u) \mapsto 1 - u^2$, (which is constant in x and concave as a function of u). As an alternative, we could have used the second derivative test. Accordingly Mangasarian's Theorem shows that (x^*, u^*) is an optimal pair for the problem.

Example 2

$$\max \int_0^T (x - u^2)\,\mathrm{d}t, \quad \dot{x} = x + u, x(0) = 0, x(T) \text{ free}, u \in \mathbb{R}.$$

Here

$$H(t, x, u, p) = x - u^2 + p(x + u) = -u^2 + pu + (1 + p)x,$$

which is concave in (x, u). Let us maximize H with respect to u: $\frac{\partial H}{\partial u} = -2u + p = 0$ if and only if $u = \frac{1}{2}p$. This gives a maximum as $\frac{\partial^2 H}{\partial u^2} = -2 < 0$. As a consequence we find $\dot{x} = x + \frac{1}{2}p$, $\dot{p} = -\frac{\partial H^*}{\partial x} = -(1 + p)$ Hence

$$\int \frac{\mathrm{d}p}{1 + p} = \int (-1)\,\mathrm{d}t, \ \log|1 + p| = -t + C, \ |1 + p| = Ae^{-t}$$

From (T) we find $p(T) = 0$, thus $1 = Ae^{-T,}$ $A = e^T$. Accordingly, $1 + p = \pm e^{T-t}$, $p(t) = e^{T-t} - 1$ ($p(T) = 0$ hence the plus sign must be used.) It follows that $\dot{x} - x = \frac{1}{2}p = \frac{1}{2}[e^{T-t} - 1]$, $\frac{\mathrm{d}}{\mathrm{d}t}(e^{-t}x) = \frac{1}{2}[e^{T-2t} - e^{-t}]$,

$$x(t) = -\frac{1}{4}e^{T-t} + \frac{1}{2} + De^t$$

Now

$$x(0) = 0 = -\frac{1}{4}e^T + \frac{1}{2} + D,\ D = \frac{1}{2}(-1 + \frac{1}{2}e^T) = \frac{1}{4}e^T) = \frac{1}{4}e^T - \frac{1}{2},$$

hence

$$x^*(t) = -\frac{1}{4}e^{T-t} + \frac{1}{4}e^{T+t} - \frac{1}{2}e^t + \frac{1}{2} = \frac{1}{4}[e^{T+t} - e^{T-t}] + \frac{1}{2}(1 - e^t)$$

■

Next we willl study control problems of the form

$$(2) \qquad \begin{cases} \max/\min \int_{t_0}^{t_1} f(t, x(t), u(t))\,\mathrm{d}t, \quad \dot{x}(t) = g(t, x(t), u(t)), \\ x(t_0) = x_0, x(t_1) \text{ free}, \\ u(t) \in U, t \in [t_0, t_1], u \text{ piecewise continuous}, \\ U \text{ a given interval in } \mathbb{R} \end{cases}$$

We shall assume that f and g are C^1-functions and that $x(t_1)$ satisfies exactly one of the following three terminal conditions:

$$(3) \qquad \begin{cases} (i) & x(t_1) = x_1, \text{ where } x_1 \text{ is given} \\ (ii) & x(t_1) \geq x_1, \text{ where } x_1 \text{ is given} \\ (iii) & x(t_1) \text{ is free} \end{cases}$$

In the present situation it turns out that we must consider two different possibilities for the Hamilton function in order to obtain the correct Maximum Principle.

$$(A) \qquad H(t, x, u, p) = f(t, x, u) + pg(t, x, u) \qquad \text{(normal problems)}$$

or

$$(B) \qquad H(t, x, u, p) = pg(t, x, u) \quad \text{(degenerate problems)}$$

Case (A) is by far the most interesting one. Consequently, we shall almost always discuss normal problems in our applications. Note that in case (B) the Hamiltonian does not depend on the integrand f.

Remark. If $x(t_1)$ is free (as in (iii)), then the problem will always be normal. More general, if $p(t) = 0$ for some $t \in [t_0, t_1]$, then it can be shown that the problem is normal.

Pontryagin's Maximum Principle takes the following form.

Theorem (The Maximum Principle II) Assume that (x^*, u^*) is an optimal pair for the control problem in (2) and that one of the terminal conditions $(i), (ii)$, or (iii) is satisfied. Then there exists a continuous and piecewise differentiable function p such that $p(t_1)$ satisfies exactly one of the following transversality conditions:

$$\begin{cases} (i') & \text{no condition on } p(t_1), \text{ if } x(t_1) = x_1 \\ (ii') & p(t_1) \geq 0 \text{ if } x(t_1) \geq x_1, (p(t_1) = 0 \text{ if } x(t_1) > x_1) \\ (iii') & p(t_1) = 0 \text{ if } x(t_1) \text{ is free.} \end{cases}$$

In addition, for each $t \in [t_0, t_1]$, the following conditions must hold
(**a**) $u = u^*(t)$ maximizes $H(t, x^*(t), u, p(t)),\ u \in U$, that is,

$$H(t, x^*(t), u, p(t)) \leq H(t, x^*(t), u^*(t), p(t)), \quad \text{for all } u \in U.$$

(**b**) The function p (called the **adjoint function**) satisfies the differential equation

$$\dot{p}(t) = -\frac{\partial H}{\partial x}(t, x^*(t), u^*(t), p(t)) \qquad (\text{written } -\frac{\partial H^*}{\partial x}),$$

except at the discontinuity points of u^*.
(**c**) If $p(t) = 0$ for some $t \in [t_0, t_1]$, then the problem is normal.

Once again concavity of the Hamiltonian H in (x, u) yields sufficiency for the existence of an optimal pair:

Mangasarian's Theorem II. Assume that (x^*, u^*) is an admissible pair for the maximum problem given in (2) and (3) above. Assume further that the problem is normal.

(**A**) If the map $(x, u) \mapsto H(t, x, u, p(t))$ is concave and nonlinear for each $t \in [t_0, t_1]$, then each admissible pair $(x*, u^*)$ that satisfies conditions

$(a), (b)$, and (c) of the Maximum Principle II, will maximize the integral $\int_{t_0}^{t_1} f(t, x, \dot{x})\,\mathrm{d}t$.

(**B**) If the map $(x, u) \mapsto H(t, x, u, p(t))$ is linear for each $t \in [t_0, t_1]$, then each admissible pair $(x*, u^*)$ that satisfies conditions $(a), (b)$, and (c) of the Maximum Principle II, will either maximize or minimize the integral $\int_{t_0}^{t_1} f(t, x, \dot{x})\,\mathrm{d}t$.

In contrast to the Maximum Principle, the proof of Mangasarian's Theorem is neither particularly long nor difficult. For the proof we shall need the following

Lemma. If ϕ is a concave real-valued C^1-function on an interval I, then

ϕ has a maximum at $x_0 \in I \iff \phi'(x_0)(x_0 - x) \geq 0$, for all $x \in I$.

Proof. We let $a < b$ be the endpoints of I.
$\Longrightarrow$: Assume that $x : 0$ is a max point for ϕ. Since the derivative ϕ' exists there are exactly three possibilities:

(1) $x_0 = a$ and $\phi'(x_0) \leq 0$,
(2) $a < x_0 < b$ and $\phi'(x_0) = 0$, and
(3) $x_0 = b$ and $\phi'(x_0) \geq 0$,

In all three cases we have $\phi'(x_0)(x_0 - x) \geq 0$.
$\Longleftarrow$: Suppose that $\phi'(x_0)(x_0 - x) \geq 0$. Again there are three possibilities:

(1) $x_0 = a$: Then $x_0 - x \leq 0$, hence $\phi'(x_0) \leq 0$. Since ϕ is concave, the tangent lies above or on the graph of ϕ at each point. Hence ϕ' decreases (this may also be shown analytically) so that

$$\phi'(x) \leq \phi'(a) = \phi'(x_0) \leq 0, \text{ for all } x \in I.$$

Hence ϕ decreases and $x_0 = a$ is a maximum point.

(2) $a < x_0 < b$: For any $x \in I$ which satisfies $x_0 < x$ we have $x_0 - x < 0$, hence $\phi'(x_0) \leq 0$. On the other hand, if $x_0 > x$, then $x - x_0 > 0$, and hence $\phi'(x_0) \geq 0$ too. It follows that $\phi(x_0) = 0$. As ϕ' is decreasing, x_0 must be a max point for ϕ.

(3) $x_0 = b$: Then $x_0 - x \geq 0$, hence $\phi'(x_0) \geq 0$. Since ϕ' is decreasing, $\phi'(x) \geq 0$ for all $x \in I$. Therefore ϕ is increasing and $x_0 = b$ must be a maximum point. ■

Proof of Mangasarian's Theorem. (a) Assume that (x^*, u^*) is an admissible pair and satisfies conditions (a), (b), and (c) in the hypothesis of

the Maximum Principle. Assume further that the map

$$(x, u) \longmapsto H(t, x, u, p(t))$$

is concave for each $t \in [t_0, t_1]$. Let (x, u) be any admissible pair for the control problem. We must show that

$$\Delta = \int_{t_0}^{t_1} f(t, x^*(t), u^*(t))\,\mathrm{d}t - \int_{t_0}^{t_1} f(t, x(t), u(t))\,\mathrm{d}t \geq 0$$

Let us introduce the simplified notations

$$f^* = f(t, x^*(t), u^*(t)),\ f = f(t, x(t), u(t)),$$

and similarly for g^*, g, H^*, and H. Thus

$$H^* = H(t, x^*(t), u^*(t), p(t)),\ H = H(t, x(t), u(t), p(t)).$$

Then

$$H^* = f^* + pg^*,$$

hence

$$f^* = H^* - pg^* = H^* - p\dot{x}^* \quad \text{since } \dot{x}^* = g^* \text{ by condition } (b))$$

and

$$f = H - p\dot{x}.$$

It follows that

$$\begin{aligned}\Delta &= \int_{t_0}^{t_1} (f^* - f)\,\mathrm{d}t = \int_{t_0}^{t_1} (H^* - p\dot{x}^* - H + p\dot{x})\,\mathrm{d}t \\ &= \int_{t_0}^{t_1} p(\dot{x}^* - \dot{x})\,\mathrm{d}t - \int_{t_0}^{t_1} (H - H^*)\,\mathrm{d}t\end{aligned}$$

Since H is concave with respect to (x, u) the "gradient inequality" (see Proposition 9 of Section 1) holds

$$\begin{aligned}H - H^* &\leq \frac{\partial H^*}{\partial x}(x - x^*) + \frac{\partial H^*}{\partial u}(u - u^*) \\ &= -\dot{p}(x - \dot{x}) + \frac{\partial H^*}{\partial u}(u - u^*) \\ &\leq -\dot{p}(x - \dot{x}) \qquad \text{(by the last Lemma)},\end{aligned}$$

except at the discontinuities of u^*. For the last inequality above we used that $\frac{\partial H^*}{\partial u}(u^* - u) \geq 0$ which holds since H is concave in (x, u), hence in u, Lemma. Consequently,

$$\begin{aligned}
\Delta &\geq \int_{t_0}^{t_1} p(\dot{x} - \dot{x}^*)\,\mathrm{d}t + \int_{t_0}^{t_1} \dot{p}(x - x^*)\,\mathrm{d}t \\
&\int_{t_0}^{t_1} \frac{d}{dt}[p(x - x^*)]\,\mathrm{d}t \qquad (\text{ here we used that the problem is normal }). \\
&= |_{t_0}^{t_1} p(x - x^*) \\
&= p(t_1)[x(t_1) - x^*(t_1)] - p(t_0)[x(t_0) - x^*(t_0)] \\
&= p(t_1)[x(t_1) - x^*(t_1)] \qquad (\text{since } x(t_0) = x_0 = x^*(t_0))
\end{aligned}$$

By the terminal conditions,

$$\begin{aligned}
&(i) \quad x(t_1) = x_1 = x^*(t_1) : \ \Delta \geq 0 \text{ is clear.} \\
&(ii) \quad x(t_1) \geq x_1 : \ \text{ If } x^*(t_1) = x_1, \text{ then} \\
&\qquad x(t_1) \geq x_1 = x^*(t_1), \text{ and since } p(t_1) \geq 0 \text{ we find that } \Delta \geq 0 \\
&(iii) \quad x(t_1) \text{ is free: } p(t_1) = 0, \text{ hence } \Delta \geq 0.
\end{aligned}$$

Hence we have shown that the pair (x^*, u^*) is optimal.

(b) We shall just sketch the proof. Assume that the Hamiltonian H is linear and hence both convex and concave in (x, u). Notice that the corresponding minimum problem is

$$\min \int_{t_0}^{t_1} f(t, x(t), u(t))\,\mathrm{d}t = -\max \int_{t_0}^{t_1} -f(t, x(t), u(t))\,\mathrm{d}t$$

It is straight forward to verify that, under the given hypothesis, the max and min problems yield the same candidate(s) (x^*, u^*) for optimal pair(s). If there are more than one such pair (x^*, u^*), then the max is obtained at the pair that gives the largest integral. ■

Example 3 (SSS 12.4 Eksempel 1)
We shall solve the control problem

$$\max \int_0^1 x(t)\,\mathrm{d}t, \quad \dot{x} = x + u,\ x(0) = 0, x(1) \geq 1, u(t) \in [-1, 1] = U, \text{ for all } t \in [0, 1].$$

We will assume (as usual) that the problem is normal. Hence the Hamiltonian is

$$H(t,x,u,p) = x + p(x+u) = (1+p)x + pu$$

which is linear in (x,u)m hence it is concave (for all $t \in [0,1]$). Thus Mangasarian's Theorem applies and any admissible pair (x^*,u^*) that satisfies the conditions of the Maximum Principle will solve the problem. Here the transversality condition

$$(iii') \quad p(1) \geq 0 (p(1) = 0 \text{ if } x(1) > 1)$$

applies. Since H is linear in u, the maximum is attained at one of the endpoints $u = 1$ or $u = -1$, hence

$$u^*(t) = \begin{cases} 1, & \text{if } p(t) > 0 \\ -1, & \text{if } p(t) < 0 \\ \text{undetermined if } p(t) = 0. \text{ We let } u^*(t) = 1 \text{ in this case.} \end{cases}$$

Furthermore, for $x = x^*, u = u^*$,

$$\frac{\partial H}{\partial x} = 1 + p = -\dot{p} \quad \text{(from condition } (b) \text{ in the Maximum Principle)}$$

Hence $\dot{p} + p = -1$ which yields $p(t) = Ae^{-t} - 1$, where $p(1) = Ae^{-1} - 1 \geq 0$, so that $A \geq e$. Therefore

$$p(t) = Ae^{-t} - 1 \geq e^{1-t} - 1 = h(t),$$

where

$$h'(t) = -e^{1-t} < 0, \text{ and } h(1) = 0,\ h(0) = e - 1 > 0.$$

Hence $h(t) \in [0, e-1]$. In particular,

$$p(t) \geq h(t) > 0 \quad \text{for } t \in [0,1),\ p(1) \geq 0.$$

Accordingly

$$u = u^*(t) = 1, \quad t \in [0,1].$$

It follows that

$$\dot{x}^* - x^* = 1, \quad x^*(t) = Be^t - 1.$$

Further,

$$x^*(0) = 0 \Rightarrow B = 1 \Rightarrow x^*(t) = e^t - 1.$$

Thus $x^*(1) = e - 1 > 0$, so that $p(1) = 0$, and $A = e$. We have found

$$x^*(t) = e^t - 1, \ u^*(t) = 1, \text{ for all } t \in [0, 1]$$

and this is an optimal pair by Mangasarian's Theorem. ■

An inspection of the proof of Mangasarian's Theorem reveals that it will work locally if the function $(x, u) \mapsto H(t, x, u, p(t))$ is concave only in some open and convex subset V of $\mathbb{R}^2$. In fact, under these hypothesis the Gradient Inequality (Proposition 1.9) is valid. This yields

Corollary. Assume that (x^*, u^*) is an admissible pair for the maximum problem given in (2) and (3) above. Assume further that the problem is normal. Let V be an open and convex subset of $\mathbb{R}^2$ and let

$$W = \{(x, u) : (x, u) \text{ is admissible and } (x(t), u(t)) \in V \text{ for all } t \in [t_0, t_1]\}.$$

If for each $t \in [t_0, t_1]$ the map $(x, u) \mapsto H(t, x, u, p(t))$ is concave on V, then each admissible pair (x^*, u^*) in W that satisfies conditions (a), (b), and (c) of the Maximum Principle II, will maximize the integral $\int_{t_0}^{t_1} f(t, x, \dot{x})\, \mathrm{d}t$ among all admissible pairs (x, u) in W.

Exercise 1. Solve the control problem

$$\max \int_0^2 (x - 2u)\, \mathrm{d}t, \quad \dot{x} = x + u, u \in [0, 1],$$

and where $x(0) = 0, x(2)$ is free.

Exercise 2. Consider the control problem

$$\max \int_0^2 (x - 2u)\, \mathrm{d}t, \quad \dot{x} = x + u, u \in [0, 1],$$

and where $x(0) = 0, x(2) = e^2 - 3$.

Using the Maximum Principle, try to find a solution candidate (x^*, u^*), an associated adjoint function p, and a point t_* in $(0, 2)$ such that

$$(1) \qquad p(t) \begin{cases} < 2, & \text{if } t \in [0, t_*) \\ > 2, & \text{if } t \in (t_*, 2] \end{cases}$$

or

$$(2) \qquad p(t) \begin{cases} > 2, & \text{if } t \in [0, t_*) \\ < 2, & \text{if } t \in (t_*, 2] \end{cases}$$

(a) Decide if (1) or (2) yield a possible solution (x^*, u^*).
(b) Solve the control problem.

Exercise 3. Consider the system of differential equations

$$(1) \qquad \begin{cases} \dot{x} = 2\varepsilon y - 2\sqrt{x} \\ \dot{y} = 2x + \frac{y}{\sqrt{x}} - 1\,, \quad x > 0, \end{cases}$$

where ϵ is a fixed but arbitrary real number.

(**a**) Solve the nonlinear system in (1) by the method of elimination. If $\epsilon = 1$ the general solution of (1) may be written as

$$x(t) = Ae^{2t} + Be^{-2t}$$

$$y(t) = Ae^{2t} - Be^{-2t} + \sqrt{Ae^{2t} + Be^{-2t}}$$

Consider next the control problem

$$(2) \qquad \max \int_0^1 (x - x^2 - u^2)\,\mathrm{d}t, \quad \dot{x} = -2\sqrt{x} - 2u,\ x(0) = 1,\ x(1) = 0$$

(**b**) Show that the Maximum Principle leads to the system

$$(3) \qquad \begin{cases} \dot{x} = 2p - 2\sqrt{x} \\ \dot{p} = 2x + \frac{p}{\sqrt{x}} - 1 \end{cases}$$

where p denotes the adjoint function. (You can assume that this is a normal problem.)

(**c**) Find the only possible solution (x^*, u^*) of the control problem in (2).
(**d**) Show that, for each fixed $t \in [0, 1]$, the function $(x, u) \mapsto H(t, x, u, p(t))$ is concave on the set

$$R(t) = \{(x, u) \in \mathbb{R}^2 \,:\, 4x^{3/2} \geq p(t)\}$$

(**e**) Next let

$$W = \{(x, u) \,: (x, u) \text{ is an admissible pair and } 4x(t)^{3/2} > p(t), \text{ for all } t \in [0, 1]\}$$

Show that the pair (x^*, u^*) of (**c**) is an element of W. Explain that (x^*, u^*) maximizes the integral $\int_0^1 (x - x^2 - u^2)\,\mathrm{d}t$ among all pairs (x, u) in W.

Arrow's condition.
Mangasarian's Theorem requires that the Hamilton function is concave with respect to (x, u). In many important applications this condition is not satisfied. Consequently, we will consider a weaker but related condition that in some cases is sufficient for optimality. Define

$$\hat{H}(t, x, p) = \max_{u \in U} H(t, x, u, p)$$

where we assume that the maximum exists. Then we have

Arrow's Sufficiency Condition. In the control problem (2) above assume that the Hamilton function is given by

$$H(t, x, u, p) = f(t, x, u) + pg(t, x, u).$$

(that is, the problem is normal). If the conditions of the Maximum Principle are satisfied for an admissible pair (x^*, u^*) and the function $x \mapsto \hat{H}(t, x, p(t))$ is concave for each $t \in [t_0, t_1]$, then (x^*, u^*) is optimal for the problem.

Example. We shall apply the Maximum Principle to find a possible optimal pair for the problem. Then we will show that this pair is indeed optimal.

$$\max \int_{-1}^{1} (tx - u^2)\,\mathrm{d}t, \quad \text{where } \dot{x} = x + u^2,\ u(t) \in [0, 1] \text{ for all } t \in [-1, 1],$$

and with the endpoint conditions $x(-1) = -2e^{-1} - 1$, $x(1)$ free.
Solution:
Assume that (x^*, u^*) is an optimal pair. The Hamilton function is

$$H(t, x, u, p) = tx - u^2 + p(x + u^2) = (t + p)x + (p - 1)u^2.$$

According to the Maximum Principle there exists a continuous function $p(t)$ such that for each t the control $u = u^*(t)$ maximizes $H(t, x^*(t), u, p(t))$. Here

$$\frac{\partial H}{\partial u} = 2(p - 1)u = 0 \text{ if and only if } \ u = 0 \text{ or } \ p = 1.$$

The second derivative of H with respect to u is equal to $2(p - 1)$ and it is negative $\Leftrightarrow p < 1$. Consequently, whenever

$p = p(t) > 1$:
$u = u^*(t) = 0$ yields a minimum for H. Further, $u = u^*(t) = 1$ gives a maximum, and for all such t the function x^* must satisfy the equation

$\dot{x}^* - x^* = 1$, hence $x^*(t) = Ae^t - 1$. Next, if

$p = p(t) < 1$:
$u = u^*(t) = 0$ maximizes H, so that $\dot{x}^* - x^* = 0$, $x^*(t) = Be^t$.

Furthermore,

$$-\dot{p}(t) = \frac{\partial H}{\partial x} = t + p(t),$$

hence $\dot{p}(t) + p(t) = -t$, $p(t) = 1 - t + Ce^{-t}$. Here $p(1) = C = 0$, since $x(1)$ is free. Therefore,

$$p(t) = 1 - t, \quad t \in [-1, 1].$$

Thus $p(t) > 1 \Leftrightarrow t < 0$, so that $p(t) > 1$ on $[-1, 0)$ and $p(t) < 1$ on $(0, 1]$.

$t \in [-1, 0)$:
$p(t) > 1, x^*(t) = Ae^t - 1$, where $x(-1) = Ae^{-1} - 1 = -2e^{-1} - 1, A = -2$. Hence $x^*(t) = -2e^t - 1, \quad u^*(t) = 1$.

$t \in (0, 1]$:
$p(t) < 1, x^*(t) = Be^t$. Continuity of x^* at $t = 0$ gives $x^*(0) = B = -2 - 1 = -3$. Therefore,

$$x^*(t) = -3e^t, u^*(t) = 0$$

$t = 1$:
$u^*(1)$ can be assigned any value in $[0, 1]$, since H is constant in u. Let us choose the value $u^*(1) = 1$.

Then the pair (x^*, u^*) is the only possible candidate for a solution.

It remains to prove that (x^*, u^*) is in fact optimal. We see that $H(t, x, u, p(t))$ is concave in $(x, u) \Leftrightarrow p(t) \leq 1 \Leftrightarrow t \in [0, 1]$. Hence Mangasarian's Theorem does not apply. However, we observe that

$$\hat{H}(t, x, p(t)) = \max_{u \in [0,1]} H(t, x, u, p(t)) = (t + p(t)x) + p(t) - 1 \text{ if } p(t) > 1$$

and

$$\hat{H}(t, x, p(t)) = (t + p(t))x, \text{ if } p(t) \leq 1,$$

whence $\hat{H}$ is linear and is therefore concave in x. Accordingly, Arrow's Theorem implies that the pair (x^*, u^*) solves the problem.

Exercise 4. Solve the control problem in the last example if the end point conditions instead are

$$x(-1) = 0, \ x(1) = e^2 - e^{1+\frac{1}{e}}.$$

Hint: Try with $p(t) < 1$ on an interval $(t_0, 1]$, where p denotes the adjoint function.

REFERENCES

[EP] Edwards & Penney, Elementary Differential Equations, 2009. Upper Saddle River, N.J. : Pearson Prentice Hall. ISBN: 978-0-13-235881-1.

[GF] Gelfand, I. M., Fomin, S. V., Calculus of variations, Prentice-Hall, 1963.

[O] Osgood, W. F., Sufficient Conditions in the Calculus of Variations, The Annals of Mathematics, 2nd Series, Vol. 2, No. 1/4, 1900-1901 (105-129) (Online at: www.jstor.org/stable/2007189)

[LM] Lee, E. B., Markus, L., Foundations of Optimal Control Theory, John Wiley and Sons: New York, 1967.

[PBGM] Pontryagin, Boltyanskii, Gamkrelidze, Mischchenko, The Mathematical Theory of Optimal Processes, Interscience Publishers (Wiley), 1962.

[SSS] Sydsæter K., Seierstad A., Strøm, A., Matematisk Analyse, Bind 2, Gyldendal Akademisk 2004.

[SS] Seierstad, A., Sydsæter, K., Optimal Control Theory with Economic Applications, North-Holland 1987. ISBN: 0-444-87923-4.

[T] Troutman, J. L., Variational Calculus and Optimal Control, Second Edition, Undergraduate Texts in Mathematics, Springer-Verlag New York, 1996.

Answers to selected exercises.

Chapter 2.

1.

(**a**) $\ddot{x} - x = -\frac{1}{2}$. Neither convex nor concave.
(**b**) $Ae^t + Be^{-t} + \frac{1}{2}$
(**c**) $\frac{1}{2}[x(t_1)^2 - x(t_0)^2]$
(**d**) $\frac{e^t - e^{-t}}{e - e^{-1}} + \frac{1}{2}$

2.

(**a**) $\frac{1}{2}(x^2 + \dot{x}^2)$ is one such function.
(**b**) $-\frac{1}{2}(x^2 + t^2\dot{x}^2)$ is one such function.
(**c**) The first function is convex, the second is concave (for all t).

(**d**) $e^t - e^{-t}$.
(**e**) $\cos \frac{1}{2}\sqrt{3}t + 2 \sin \frac{1}{2}\sqrt{3}t$

3.
(**a**) $\frac{d}{dt}(2t^2\dot{x} + 1) = 0.$ $\quad 7 - 4/t$

4.
(**c**) $\sqrt{t^2 - t + 1}$

6.
(**d**) $x(t) = \frac{1}{2}(\ln 3 - 1)t + 1 - \frac{1}{2}t \ln \frac{t+1}{t-1}$

7. $x(t) = t^{r_1} - t^{r_2}$.

8.

9.

12.
(**b**) $t^{-1} \sin t$
(**c**) $t^{-1}(A \cos t + B \sin t))$
(**d**) $\frac{\pi}{2}t^{-1}(\sin t - \frac{1}{\sqrt{3}} \cos t)$ yields a minimum by the Weierstrass sufficiency condition applied to $x(t, \alpha) = \frac{\pi}{2}t^{-1}[-\frac{1}{\sqrt{3}}(1 - \alpha) \cos t + (1 + \alpha) \sin t]$.