Journal of Mathematical Sciences: Advances and Applications
Volume 19, Number 1, 2013, Pages 13-33

AN INEQUALITY FOR LEGENDRE TRANSFORMATION

ALIREZA BEHTASH

Department of Physics
Middle East Technical University
06531 Ankara
Turkey
e-mail: alireza.behtash@metu.edu.tr

Abstract

Here, we prove a theorem for Legendre transformation of some specific derivative-like sequence chosen as the argument of Legendre transform $f^\star$ of a function f using theory of convex functions, mean value theorem in one dimensional Euclidean space, and finally, a mathematical program established to provide some conditions of local convexity that may be incompatible with the existence of Legendre transformation. We also discuss the useful results of this theorem along with numerous examples. These results aim at providing a new set of Legendre transformations generated by a given convex function, where the variable of the function is regarded as an interval length. This generation is actually based on an appropriate modification of variables, which yields the treatment of Legendre transformations over a specific field of distributions.

1. Introduction

Convex functions are subject to the generic notion of convexity due to having epigraphs, which are always convex sets. Continuity is followed by convexity, whereas differentiability has to be treated carefully as it may

2010 Mathematics Subject Classification: 52A41.

Keywords and phrases: Legendre transformation, local convexity, quasiconvexity, pseudomean value.

Received August 15, 2012; Revised November 13, 2012

happen that a convex function is not differentiable at some point(s). The set of these points is countable. Roughly speaking, given a continuous function *f* over, say, a closed set in $\mathbb{R}$, it can still be locally convex if the set contains a *singular* point for which f' is undefined as long as the epigraph remains convex. If epigraph loses its convexity, then *f* is said to be "quasiconvex". Both convex and quasiconvex functions appear to be extremely useful in analysis and optimization and highly favorable in applied sciences as well. The proof of several famous inequalities such as Jensen's inequality and arithmetic-geometric inequalities makes use of convex functions as a central argument.

Legendre transformation is an another example of the application of convex functions that plays a significant role in theoretical physics and in particular the theory of classical fields. In this paper, we will be concerned with the behaviour of near maximized Legendre transformations, which may reduce to the ordinary Legendre transformations as the maximization condition is retrieved by proving the following theorem:

Theorem 1. *Let* $f : M \to \mathbb{R}$ *for some closed interval M be convex on* $J_i = [x_{i-1}, x_i]$, *where* $\bigcup_{i \in I} J_i \subset M$ *for some index set I. Additionally, suppose* $x_i - x_{i-1} = \delta_i > 0$. *For any* p_i *of the form* $p_i = \dfrac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}}$, *there is a Legendre transformation* $g(p_i) = p_i \chi_i - g^\star(\chi_i)$ *such that* $\chi_i = x_i - \delta_i$.

Since the maximization condition leans on the convexity of functions applied to the Legendre transformation and special criteria exist that must be satisfied by a convex function, a general local convexity program to determine which real-valued functions are mathematically eligible for a local study of Legendre transformations on closed intervals will be given in Section 2. Most of what will appear in this section are extensively discussed in the literature by many authors [5, 9]. Finally, the proof of Theorem 1 will be presented in Section 3.

2. Preliminaries

2.1. Legendre transformation

Definition 1. The Legendre transform $f^\star$ of a continuously differentiable convex function f is defined by $f^\star(p) = \max_x(px - f(x))$, where $px - f(x)$ is maximized with respect to the variable x; hence $p = df / dx$.

A key characteristic of Legendre transformation is that every variable of a continuously differentiable convex function $f : \mathbb{R} \to \mathbb{R}$ has to be able to be written as a function of some other variable belonging to a new function, or the transformed function $f^\star : \mathbb{R} \to \mathbb{R}$. Yet, $f^\star$ is convex if and only if f is convex. Therefore, the property of convexity of f will be preserved under Legendre transformation. In this part, we shall discuss the convexity of functions as a local problem on $\mathbb{R}$ that results in a Legendre transformation including functions f and $f^\star$ both being convex on a given interval with the help of the famous mean value theorem. We recall that here the condition of local convexity refers to the study of convexity on closed intervals, which will be clarified that, as briefly discussed in the previous section, is slightly different from quasiconvexity.

Example 1. Let $f(x) = -\ln(x)$ that is evidently convex on every closed interval $A \subset \mathbb{R}^+$ as is on $\mathbb{R}^+$. According to $p = df/dx = \frac{-1}{x}$, one has $x = \frac{-1}{p}$, giving $f^\star(p) = px - (-\ln(x))$. Thus $f^\star(p) = \ln(\frac{-1}{p}) - 1$, which is also convex on every closed interval $B \subset \mathbb{R}^-$ as is on $\mathbb{R}^-$ itself.

In fact, the local convexity condition says that there is only one of those critical points, i.e., *ps*, at which the Legendre transformation holds within some closed interval, and that it is obviously a maximum, or indeed, the mean value of the function in that interval. Requiring the convexity again, f and $f^\star$ are sometimes defined by the formula

$$f'(x) = [f^{\star'}(x)]^{-1}, \tag{1}$$

where $'$ stands for the derivative of the function with respect to its variable and $[.]^{-1}$ represents the inverse function. Here, it must be noticed that f and $f^{\star}$ satisfy also the following important relation:

$$f = f^{\star\star}, \tag{2}$$

which implies the fact that the Legendre transformation is an involution.

Theorem 2. *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuously differentiable convex function* (*not necessarily strictly convex*). *For the sequence*,

$$p_i = \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}}; \quad i \in I, \tag{3}$$

with I being an index set, *the Legendre transform $f^{\star}$ of f satisfies the inequality*

$$f^{\star}(p_i) > p_i \left\{\frac{f(x_i)x_{i-1} - x_i f(x_{i-1})}{f(x_i) - f(x_{i-1})}\right\}, \tag{4}$$

such that f is convex on the interval $[x_{i-1}, x_i]$ and $f(x_i) \neq f(x_{i-1})$.

To prove this theorem, one has to initially assert the following lemma on the relationship between the convexity of a continues function $f(x)$ on some (closed) interval and the mean value theorem satisfied over that interval, thanks to the continuity of $f(x)$.

Lemma 1. *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuously differentiable function of one variable*, *which is also convex on $[x_0, x_1]$. Suppose that p is the argument of Legendre transform $f^{\star}$ of f, i.e., $f^{\star}(p) = px - f(x)$. Then the following inequality holds for any p*:

$$p > \frac{f(c) - f(x_0)}{c - x_0}, \tag{5}$$

such that $p = f'(c) = \dfrac{f(x_1) - f(x_0)}{x_1 - x_0}$ and $c \in (x_1, x_0)$.

Proof. The proof of this lemma is so easy if one knows that f is said to be convex on $[x_0, x_1]$, if it lies below the straight line segment connecting two points, for any two points in this interval[1]. Mathematically, this can be shown by

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0}(X - x_0) + f(x_0) > f(X), \tag{6}$$

where $X \in (x_0, x_1)$. In principle, the left-hand side of the inequality (6) represents indirectly the equation of a straight line, $F(X)$, that intercepts f at two points x_1 and x_0 so that the convexity of f requires $F(X) > f(X)$ for any $X \in (x_0, x_1)$.

In the original definition of Legendre transformation, i.e., Definition 1, p is taken to be constant. Hence, the process of the maximization on the interval given above results in $p = f'(x)|_c$, where by making use of the mean value theorem c satisfies $f'(c) = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$, so that f must be necessarily differentiable on (x_0, x_1). We see then from the inequality (6) that

$$f'(c)c - f(c) > f'(c)x_0 - f(x_0) \tag{7}$$

(which is evidently correct since $c > x_0$ and f is convex on $[x_0, x_1]$). Putting $p = f'(c)$ in this inequality completes the proof. □

Now, we are able to prove Theorem 2 by using Lemma 1.

Proof of Theorem 2. Following Lemma 1, it can be seen that $p > \frac{pc - f^\star(p) - f(x_0)}{c - x_0}$, with some $c \in (x_0, x_1)$ satisfying the mean value theorem for the function f. So by knowing $c > x_0$, we get

[1] It is kind of tricky to rely on this definition of local convexity of a function. Actually, this has something in common with the case of a function being just quasiconvex; see Program 1.

$$f^{\star}(p) > px_0 - f(x_0), \tag{8}$$

which must not be regarded as a special case of Young's inequality (see Theorem 3). Let f be convex on the intervals $[x_{i-1}, x_i]$ with $i \in I$ so that c_i are now taken from (x_{i-1}, x_i). Modifying $x_0 \to x_{i-1}$ and $x_1 \to x_i$ along with substituting the expression (3) into the inequality (8) yields

$$f^{\star}(p_i) > \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}} x_{i-1} - f(x_{i-1}) \equiv p_i \left\{ \frac{f(x_i)x_{i-1} - x_i f(x_{i-1})}{f(x_i) - f(x_{i-1})} \right\}, \tag{9}$$

in which we also used the condition $f(x_i) \neq f(x_{i-1})$[2]. □

Remark 1. For any function being convex on $[x_{i-1}, x_i]$, it is necessary for $c_i \in (x_{i-1}, x_i)$ to be unique, in the sense that, another $\bar{c}_i \in (x_{i-1}, x_i)$ for which the inequality (7) holds, cannot exist. This tells us how local convexity and mean value theorem are related with each other (see Program 1).

Example 2. The function $f(x) = x^3 - x^2$ is convex on $[\frac{i-1}{i}, \frac{2}{i}]$ for any $i \in I \subseteq \mathbb{N}$; so $p_1 = 2$ and $c_1 = \frac{1}{3} + \frac{\sqrt{7}}{3}$ which from (8), it is obvious that $f^{\star}(p_1 = 2) > 0$ since $f^{\star}(2) = 2 \times (\frac{1}{3} + \frac{\sqrt{7}}{3}) - f(\frac{1}{3} + \frac{\sqrt{7}}{3}) \simeq 0.31 > 0$. However, for instance, on the interval $[-i, 1-i]$, $f(x)$ is not convex and therefore for $p_1 = 2$ and $c_1 = \frac{1}{3} - \frac{\sqrt{7}}{3}$, the inequality (8) will not hold since $f^{\star}(2) = 2 \times (\frac{1}{3} - \frac{\sqrt{7}}{3}) - f(\frac{1}{3} - \frac{\sqrt{7}}{3}) \simeq -0.46 \not> 0$.

---

[2] However, we can neglect this weak condition by a quick elimination of the ambiguity when considering $f(x_i) = f(x_{i-1})$ for $x_i > x_{i-1}$, as i is some given natural number in the index set I, since the singularity occurring for $f^{\star}(p_i)$ in this case is not instinct and will leave the convexity of $f^{\star}$ intact.

Another property of Legendre transformation is taken from the duality of $f(p)$ and $f^{\star}(p)$ introduced by Young as follows: If $g(p) = f^{\star}(p)$, then by Equation (2), $f(p) = g^{\star}(p)$, which we say f and g are dual. This leads to the following theorem:

Theorem 3 (Young's inequality)**.** *If f and $f^{\star}$ are convex functions and are dual* (*in Young's sense*), *then* $f^{\star}(p) + f(x) \geq px$.

It is important to note that Young's inequality is basically different from our inequality (9) by the remarkable reason that in Young's inequality, $f^{\star}$ and f are considered to be dual, but in ours, f and $f^{\star}$ are not necessarily dual. In Section 3, using both inequalities, we present an interesting set of Legendre transformations based upon the application of convex functions and their properties [1].

2.2. Identifying local convexity and quasiconvexity

Now, we turn to a more or less familiar problem concerning the local convexity of a real-valued function. The problem is that local convexity may not be justified for every function satisfying Equation (5) that is of fundamental importance in the topic of Legendre transformations. It happens when local convexity turns into quasiconvexity, where Legendre transformation might be broken in some sense, which is treated below. As characterized by many other authors [8], the quasiconvex functions are sometimes used to be defined by the following old implication: Suppose that $f : M \to \mathbb{R}$ is differentiable at $x_0 \in M \subset \mathbb{R}$, where M is an open interval (set). With this in mind, the following implication is true [7]:

$$\underbrace{x_0,\, x_1 \in M : f \text{ being locally quasiconvex at } x_1 \text{ (with respect to } M)}_{\Downarrow}$$

$$f(x_0) \leq f(x_1) \Rightarrow f'(x_1)(x_0 - x_1) \leq 0. \tag{10}$$

If f is differentiable on the open convex interval $M \subset \mathbb{R}$, the following general result can also be proved as for the quasiconvexity of f:

$$x_0, x_1 \in I;\, f(x_0) \leq f(x_1) \Rightarrow f'(x_1)(x_0 - x_1) \leq 0 \Leftrightarrow f \text{ being quasiconvex on } M. \tag{11}$$

The second portion of bi-implication (11) holds, if the first one is automatically satisfied by any pair of $\{x_{i-1}, x_i\} \in J_i \subset M$, where i is in the index set I of a closed set J_i and $f(x_{i-1}) \leq f(x_i)$. However, this bi-implication suffers from being applicable to both concave and convex functions in the following simple way: Suppose $J_1 = [x_0, x_1]$ and let f be a monotonically increasing function on J_1. Combining the condition and the consequence of bi-implication (11) together yields $f'(x_1) \geq -f'(c)$, where c is the mean value in J_1. By putting $f'(x_1) = 0$, one can easily find infinite paths on J_1 for which $0 \leq f'(c)$, which might be concave as well. In other words, every convex function is quasiconvex, but the inverse statement may not be true. This is resolved by setting a further condition that comes off as appropriate in treating the issue of quasiconvexity and convexity via a thorough program consisting of 7 conditions and cases, which can assist us to realize whether a real-valued function is convex or just quasiconvex on a closed interval $J_i = [x_{i-1}, x_i]$ by taking advantage of the mean value theorem. This way applying the Legendre transformation to the given function on the considered closed interval can be well determined.

Theorem 4 (**Program 1: identification of local convexity and quasiconvexity**)

(1) *If and only if f is continuous on $J_i = [x_{i-1}, x_i]$* (*the continuity condition*).

(2) *If for all x in J_i, $F \cup f := \{(x, F(x))\} \cup \{(x, f(x))\}$, where F is any line intersecting f, constitutes a compact set of numbers S_i in the plane $J_i \times \mathbb{R}$, namely*, *the boundary of epigraph of f, that is convex*[3].

---

[3] The physical description of this condition is that if is convex, then remembering $F \cap f = \{\overline{P}(x_i, f(x_i)), P(x_{i-1}, f(x_{i-1}))\}$ a point mass moving over f from the point P to $\overline{P}$ will have the sign of its acceleration altered once *only* during its entire motion.

(3) (i) *If f is differentiable on* $J_i = [x_{i-1}, x_i]$ (*the differentiability condition*).

(ii) *If* $f'(x_{i-1}) < f'(x_i)$ *with* $f'(x_i) \neq 0$.

(4) *If f satisfies the following formula*:

$$f'(x_i) > -f'(c_i), \tag{12}$$

where $c_i \in K_i \subset J_i$ *is the unique mean value of f in some open subinterval* K_i *such that* $\sup(K_i) = x_i$, *then f is said to have convexity* (*and*, *therefore*, *quasiconvexity*) *on* J_i.

(5) *If f does not satisfy the condition* 3(i) *and* S_i *is not convex*, *yet the inequality* (12) *holds with* $c_i \notin K_i$ *for any* $i \in I$ *necessarily*, *then f is said to be just quasiconvex on* J_i, *i.e.*, *f is not convex but is quasiconvex on* J_i. *Such* c_i *that exists and may not belong to* K_i *is called a* "*pseudomean value*" *of f*.

(6) *If f does not satisfy the condition* 3(i), *and some pseudomean value* c_i *exists in* (12), *which is not contained in either* J_i *or its complement* J_i^c *for any i*, *then f is absolutely convex*.

(7) (i) *Finite sum of functions satisfying the case* 4 *on* J_i *will be convex on* J_i *with* $c_i \in K_i$.

(ii) *Finite sum of functions satisfying the case* 4 *or* 6 *on* J_i *will be convex on* J_i.

Warning. The condition 3(i) is nearly weak as a prerequisite for (local) convexity. This can be directly explored through the last two cases of Program 1 (see P(6) below). However, this weakness is confined to the piecewise differentiable functions including only one singular point in J_i. Otherwise, some particular matching conditions (say, conditions that set $x_i, x_{i-1} \in J_i - \{\text{singular point(s)}\}$) would be required which in turn invalidates the generality of Program 1 at every point of J_i (Figure 1).

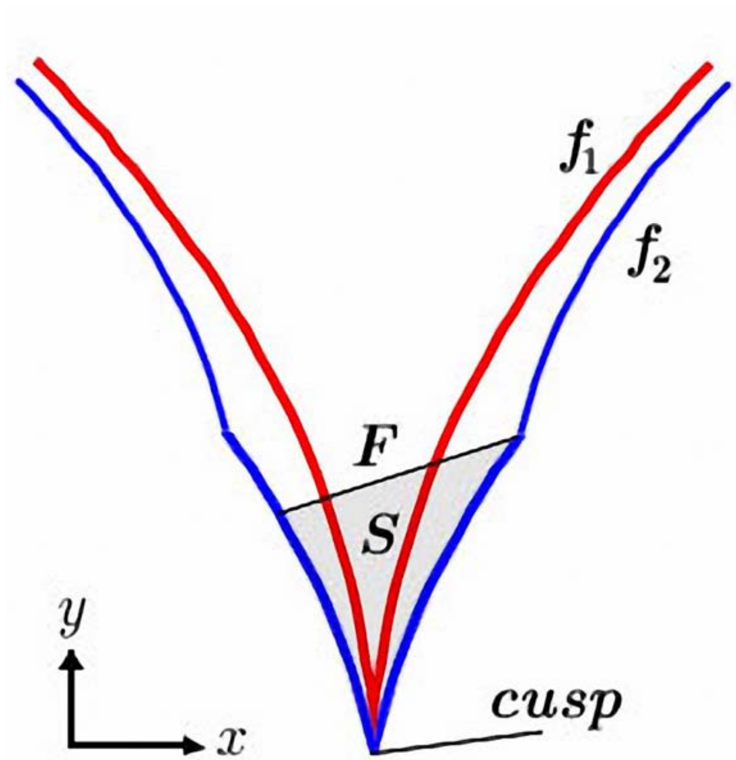

Figure 1. $F \cap f_2$ represents two points at which f_2' is undefined. Equation (12) must be adjusted with a condition at these two points to match with f_2. Also, the point of cusp, whose value makes Equation (12) undefined as well is not a valid input since it does not satisfy the bi-implication (11) of quasiconvexity before being introduced in Equation (12).

Proof. We split the proof of Program 1 into four parts, as proof of each part is shown by a 'P' before the number of cases being ranged from 4 to 7.

P(4). To derive Equation (12), without loss of generality, we take the extra condition $x_0 < x_1$ to hold in the bi-implication (11) so as to keep our previous assumption $x_{i-1} < x_i$. This immediately results in $f'(x_1)(x_0 - x_1) \lesssim 0$ such that $f'(x_1) \gtrsim 0$, which is consistent with the condition of bi-implication (11) (Remark 1), that is, $0 < f(x_1) - f(x_0)$, we find

$$-f'(x_1) < \frac{f(x_1) - f(x_0)}{x_1 - x_0}. \tag{13}$$

Now, we turn back to the inequality (6) and rewrite it as follows:

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} < \frac{f(X) - f(x_1)}{X - x_1}, \tag{14}$$

where $X \in (x_0, x_1)$. The left-hand side of (14) is equal to that of (13) so that $-f'(x_1) < \frac{f(X) - f(x_1)}{X - x_1}$. By making use of i-index transcription of this result and following $K_i = (X, x_i)$ that is an open subinterval of J_i with the characteristic of $\sup(K_i) = x_i$, one finally obtains that $f'(c_i) > 0$ corresponds to the right-hand side of Equation (14) using the conditions 1 and 3 of Program 1, such that $c_i \in K_i$ is unique due to the condition 2 (Remark 1). Since we demanded convexity along with quasiconvexity of *f* on J_i at first, therefore the proof is accomplished. □

P(5). Let $c_i \notin K_i$ be some unique pseudomean value of *f*. This means that there is only one point $x_s \in (x_{i-1}, x_i)$ for which f' does not exist so that 3(i) is weakly violated at x_s. Assume that it is unlikely to have $f'(x_i) = 0$; hence x_i would be taken as a global minimum in J_i that makes x_s a point of discontinuity for *f* in order for f' not to exist, which is in contradiction with the condition 1[4]. Thus f cannot be monotonically increasing on J_i, which proves $f'(x_{i-1}) < 0 < f'(x_i)$ and 3(ii). On the other hand, since $f(x_{i-1}) \leq f(x_i)$, then by P(4), $f'(c_i) > 0$ that supports the condition 4. This together with 3(i) and $f'(x_s) = undefined$ states that f is a curve having a 'cusp' at x_s that calls for the set S_i to be non-convex, leading f to lose its convexity on J_i. Since Equation (12) is still valid, according to P(4) and the implication that *quasiconvexity* $\Rightarrow$ *convexity* may not hold, we are done. □

P(6). Following P(5), if c_i does not exist anywhere, again a point $x_s \in (x_{i-1}, x_i)$ must be found in J_i for which f' is undefined. From the differentiability condition and 3(ii), we apparently have $f'(x_{i-1}) = B < 0$ and $f'(x_i) = A > 0$, where $A, B = constant$. Let $J_i = P_i \cup Q_i$ such that $P_i \cap Q_i = \emptyset$. Suppose $x_i \in Q_i$ and $x_{i-1} \in P_i$: by the continuity condition, one has $f(x_i) \cup f(x_{i-1}) = f(x)$ such that $x \in J_i$. Now $F \cup f$ shows a set of points constructing a triangle that is obviously a convex polygon (S_i is convex), that can be understood from the countable number of right and left derivatives in the interior of J_i as well. Consequently, f satisfies the translation $A = f'(x_i) > -\dfrac{f(X) - f(x_i)}{X - x_i}$ of Equation (12), where $X \in (x_{i-1}, x_i)$. When $A = 0$, $f(X) < f(x_i) =$ *constant* if and only if $X < x_i$. But since $B < 0$, $P_i \cap Q_i \neq \emptyset$. If $Q_i \subset P_i$, every $X \in P_i - Q_i - \{x_i, x_{i-1}\}$ is obviously greater than $x_i \in Q_i$, which is a contradiction (Figure 2). So $A \gneq 0$ and the proof is complete. □

---

[4] In fact, it is not proved yet whether f is just quasiconvex on J_i to confidently say that the minimum of f is its global minimum in J_i (Lemma 1 of [6]). However, that 'global' mentioned in the proof is to assure that c_i is unique over M.

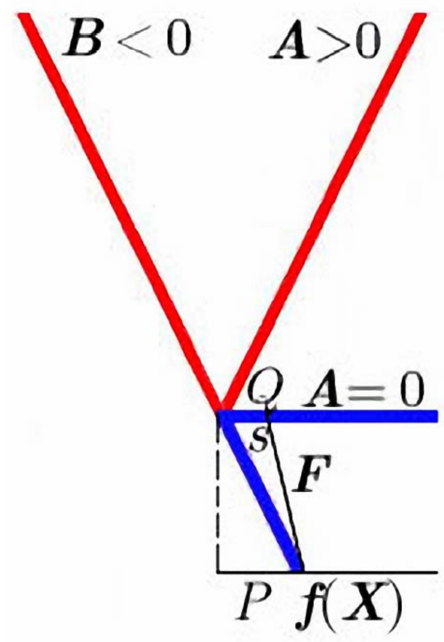

Figure 2. Illustration of P(6).

P(7). (i) Let $G(x) = \sum_{n \geq 1} f_n(x)$ such that $f_n'(x_i) > -f_n'(c_{n_i})$. Thus $G'(x_i) = \sum_{n \geq 1} f_n'(x_i) > -\sum_{n \geq 1} f_n'(c_{n_i}) = -\frac{\sum_{n \geq 1} f_n(X) - \sum_{n \geq 1} f_n(x_i)}{X - x_i} = -\frac{G(X) - G(x_i)}{X - x_i} = -G'(c_i)$ with $c_i \in K_i$, which results in the fact that G is also convex on J_i.

(ii) To not encounter the problem of matching condition, suppose that $f_1, f_3, f_5, \dots, f_{j-1}$ are all convex functions including one joint singularity at x_s in J_i and let $f_2, f_4, f_6, \dots, f_j$ be all convex smooth functions with one joint global minimum at x_s. Hence, $G'(x_s) \equiv \sum_{n \geq 1}^{j} f_n'(x_s) =$ *undefined*. This confirms $f'(x_s) \neq f'(x_i) \neq 0$; so by following P(4) and the former case, there is a $c_i \in K_i$ that satisfies $G'(x_i) > -f'(c_i)$ for any $X \in (x_{i-1}, x_i)$ and this completes the proof. □

Corollary 1. *Let* $f'(x_i) \neq 0$ *be modified by* $f'(x_{i-1}) \neq 0$ *in condition* 3(ii) *of Program* 1. *The following formula can be replaced with Equation* (12):

$$f'(x_{i-1}) < -f'(c_{i-1}), \tag{15}$$

where $f(x_i) \leq f(x_{i-1})$ *and* $c_{i-1} \in K_{i-1} \subset J_i$ *is a unique mean value of* f *in some open subinterval* K_{i-1} *such that* $\inf(K_{i-1}) = x_{i-1}$.

Proof. We use other definition of implication (10) and bi-implication (11), that in terms of i-index transcription is: A continuous function $f : M \to \mathbb{R}$ being differentiable at $x_{i-1} \in M \subset \mathbb{R}$ (again, M is an open interval), is said to be *locally pseudoconvex* at x_{i-1} (with respect to M) if and only if [8]

$$x_i \in M;\ f(x_i) < f(x_{i-1}) \Rightarrow f'(x_{i-1})(x_i - x_{i-1}) < 0. \tag{16}$$

By Theorem 3 of [3], if f is quasiconvex at x_{i-1} and if $f'(x_{i-1}) \neq 0$, then f is pseudoconvex at x_{i-1}. Suppose $x_i, x_{i-1} \in J_i \subset M$ wherein J_i is a closed interval and let f be quasiconvex on J_i. Hence, implication (16) will be true at each $x_{i-1} \in J_i$. Continuing with the same line of work that we followed in P(4), introducing implication (16) and assumption $x_i > x_{i-1}$ into the combination of inequalities (13) and (14) with a little change leads to an inequality of the form

$$f'(x_{i-1}) < -\frac{f(X) - f(x_{i-1})}{X - x_{i-1}},$$

where $f'(x_{i-1}) \lessgtr 0$ and $X \in (x_{i-1}, x_i)$. Taking $K_{i-1} = (x_{i-1}, X)$ and putting $-f'(c_{i-1})$ equal to the right-hand side of the above inequality with $c_{i-1} \in K_{i-1}$ complete the proof [5]. □

[5] As an example of this corollary, see Figure 2, in the case of $A = 0$ and $B \lessgtr 0$.

Fact 1. Except for the cases 4 and 7(i) of Program 1, the Legendre transform $f^{\star}$ of f, for the cases 5, 6 and above all else the case 7(ii) cannot be constructed on J_i. □

Example 3. $f(x) = \dfrac{\sqrt{|x|}}{x^2 + 1}$ is quasiconvex on $[\frac{-1}{9i}, \frac{1}{4i}]$ for any $i \in \mathbb{N}$, but is not convex on the same interval because $c_1 \simeq 0.36 \notin (\frac{-1}{9}, \frac{1}{4})$ and putting the exact values of $p_1 = \frac{3546}{9061}$, $x_0 = \frac{-1}{9}$, and $f(x_0) = \frac{27}{82}$ into the inequality (5) leads to the correct solution $X \in (\frac{-1}{9}, \frac{1}{4}) \cup (0.48, \infty)$, where the allowed range is considered to be the first one in the union. Thus, the Legendre transform $f^{\star}$ cannot be defined on $[\frac{-1}{9i}, \frac{1}{4i}]$.

It is remarkable to note that in general for periodic functions being locally quasiconvex in some $J_i \in \mathbb{R}$, e.g., a cycloid, the pseudomean value c_i will belong to J_i^c. As seen in Example 3, a more generalized argument can be given by stating that the existence of at least one extermum point of f in J_i^c leads to $c_i \in J_i^c$, which guarantees the quasiconvexity of the function in J_i following case 5 of Program 1. This characteristic of some quasiconvex functions is central in the study of quasiconvex optimization problems.

3. The Main Result

In this section, we will try to obtain a new class of Legendre transformations using the study of the convexity over closed intervals with the help of the results of Theorem 2 that were established in the previous sections. This will be done via proving Theorem 1.

Proof of Theorem 1. From inequality (8), we have

$$f^{\star}(p_i) > p_i\chi_i - f(\chi_i); \quad p_i = \frac{f(\chi_i + \delta_i) - f(\chi_i)}{\delta_i}. \tag{17}$$

By preserving the new variables p_i and χ_i, we modify (9) to get to a Legendre transformation of f. To do so, one has to just add a term $C > 0$ to the right-hand side of (9), where C is a concave function of some arbitrary variables x_i and $\ell_i (0 < \ell_i \leq \delta_i)$ with $i \in \mathbb{N}$:

$$f^{\star}(p_i) = p_i\chi_i - f(\chi_i) + C(x_i, \ell_i). \tag{18}$$

To prove that Equation (18) does not bother the intrinsic qualification of the Legendre transformation, we start with maximizing Equation (18) with respect to p_i and χ_i

$$f^{\star'}(p_i) = \frac{d}{dp_i}(p_i\chi_i - f(\chi_i) + C(x_i, \ell_i)) = \chi_i, \tag{19}$$

$$f'(\chi_i) = \frac{d}{d\chi_i}(p_i\chi_i - f^{\star}(p_i) + C(x_i, \ell_i)) = p_i, \tag{20}$$

that are obviously true.

Let us take $C(x_i, \ell_i)$ out of Equation (20) and try to re-cast it into a Legendre transformation. Applying Definition 1 to Equation (20) yields

$$f(\chi_i) = p_i[f'(p_i)]^{-1} - f^{\star}(p_i). \tag{21}$$

By differentiating this equation with respect to χ_i (on the left-hand side) and correspondingly with respect to p_i (on the right-hand side), we get

$$\begin{aligned} f'(\chi_i) &= [f'(p_i)]^{-1} + p_i[f''(p_i)]^{-1} - f^{\star'}(p_i) \\ &= p_i[f^{\star''}(p_i)], \end{aligned} \tag{22}$$

where again we made use of Equation (2) twice. Suppose f is convex on $J_i = [\chi_i, x_i]$ and x_i is held to be constant. To make χ_i variable, let $\chi_i = x_i - \ell_i$. In a way similar to the differentiation of Equation (21), integrating Equation (21) from $\chi_{i_0} = \chi_i - \ell_0$ to $\chi_{i_1} = \chi_i - \ell_1$ results in

$$
\begin{aligned}
f(\chi_{i_1}) - f(\chi_{i_0}) &= \int_{p_{i_0}}^{p_{i_1}} p_i[f^{\star''}(p_i)]dp_i \\
&= p_{i_1} f^{\star'}(p_{i_1}) - p_{i_0} f^{\star'}(p_{i_0}) - \int_{p_{i_0}}^{p_{i_1}} f^{\star'}(p_i)dp_i \\
&= p_{i_1} f^{\star'}(p_{i_1}) - p_{i_1} f^{\star'}(p_{i_0}) + f^{\star}(p_{i_0}) - f^{\star}(p_{i_1}) \\
&= p_{i_1}\chi_{i_1} - p_{i_0}\chi_{i_0} + f^{\star}(p_{i_0}) - f^{\star}(p_{i_1}),
\end{aligned} \tag{23}
$$

wherein the second integration was taken by parts and the last expression was derived by using the maximization of Legendre transform $f^{\star}$ at points p_{i_0} and p_{i_1}. Re-arrange Equation (23) as

$$
f(\chi_{i_0}) + f^{\star}(p_{i_0}) - p_{i_0}\chi_{i_0} = f(\chi_{i_1}) + f^{\star}(p_{i_1}) - p_{i_1}\chi_{i_1} = L; \tag{24}
$$

in which L is some constant and is needed to be identified. Say, for $i \to \infty$, $\ell_i \to \delta_i$, then by repeating Equation (23) for any pair of χ_{i_j}, with $j \in \mathbb{N}$ as before, one will find

$$
f(\chi_{i_j}) + f^{\star}(p_{i_j}) - p_{i_j}\chi_{i_j} = L. \tag{25}
$$

Thus L does not depend on the index j and we can remove it from the result derived above. Now, we have to show what sign L has. Intuitively, if $f^{\star}(p_i)$ and $f(\chi_i)$ are dual in Young's sense, by Theorem 3, $f^{\star}(p_i) + f(\chi_i) - p_i\chi_i = L \geq 0$ and if they are not necessarily dual, then by Theorem 2, $L > 0$. From an intuitive point of view, one can assert this straightly using the local convexity condition imposed on $f(\chi_i)$, i.e., f being convex on J_i.

To do so, let conditions 1-4 of Program 1 be satisfied by f and take $x_i = \chi_i + \delta_i$. Therefore, $f'(\chi_i) = f'(\chi_i + \delta_i) - \kappa$, where $\kappa > 0$. Inserting this expression into Equation (12) with the help of Corollary 1 yields

$$
f'(c_i) - f'(c_{i-1}) < \kappa.
$$

Note that now 3(ii) is accompanied with an extra condition $f'(\chi_i + \delta_i) \neq 0$ and that $f(\chi_i) \leq f(\chi_i + \delta_i)$ might not be satisfied since $\chi_i + \delta_i$ is no longer taken to be constant, which brings it to the form $\chi_i + \ell_i$, where again for $i \to \infty$, $\ell_i \to \delta_i$. Having this in mind, we take the limit of this inequality when $X \to \chi_i$

$$f'(\chi_i) - \frac{f(\chi_i + \ell_i) - f(\chi_i)}{\ell_i} < \lim_{X \to \chi_i} \kappa = \kappa. \tag{26}$$

For simplicity, let ℓ_i be small (say, when i is in some $S \subseteq \mathbb{N}$) and consequently, we subtract some small parameter $\epsilon > 0$ from the right-hand side of Equation (26) to make its sides equal; then we integrate the result with respect to χ_i

$$-\mathcal{O}(\ell_i) = (\kappa - \epsilon)\chi_i. \tag{27}$$

This equation treats five kinds of a convex function on J_i for different signs of $\mathcal{O}(\ell_i)$, $(\kappa - \epsilon)$, and χ_i. Suppose the case of $\mathcal{O}(\ell_i)$, $\chi_i > 0$ and $\kappa < \epsilon$ belonging to an increasing function. Without loss of generality, we set the first derivative to be linear by $(\epsilon - \kappa) \sim \ell_i$, then one obtains

$$f'(\chi_i) = p_i \approx 2\chi_i - \mathcal{O}((\ell_i)^2). \tag{28}$$

Introducing this expression into Equation (25) and transforming $L \to C$ (x_i, ℓ_i) yield

$$f(\chi_i) = 2\chi_i^2 - f^\star(p_i) - \chi_i\mathcal{O}((\ell_i)^2) + C; \tag{29}$$

where it is obvious that $C(x_i, \ell_i) = (x_i - \ell_i)\mathcal{O}((\ell_i)^2) > 0$. Hence by decomposing C as $C = C_1 + C_2$ and letting $g(\chi_i) = f(\chi_i) - C_1$ and $g^\star(p_i) = f^\star(p_i) - C_2$ in Equation (18) (since $-C$ is convex, $-C_1$ and $-C_2$ must be convex as well, by the case 7(i) of Program 1), we are done.

$\square$

In principle, this result emphasizes that there is a Legendre transformation in which the condition of maximization of $p_i x_i - f$, when we are in transition to $p_i \chi_i - g$ must be neglected. This means that the significance of what now we are faced with, is that the endpoints of the closed interval wherein $g(\chi_i)$ is calculated, play the main role in determining the ingredients of Legendre transformations. However, $g \to f$ needs $\chi_i \to x_i$ to bring p_i back to its maximized value that no longer makes a Legendre transformation be endpoints-dependent, since it is a point only where $p_i = f'(x_i)$. A more intuitive review on the outcomes of this theorem is summarized in the following example.

Example 4. The relativistic kinetic energy of a particle as a function of momentum at the point q_i in the presence of some potential $V := V(q_i)$, that is, $\mathcal{H}(p_i, q_i) = \sqrt{p_i^2 + m^2} + V$, if $c = 1$, is generated by

$$\mathcal{H}(p_i, q_i) = p_i \beta_i - \mathcal{L}(\beta_i, q_i), \tag{30}$$

where $\mathcal{L}(\beta_i, q_i) = -m\sqrt{1 - \beta_i^2} - V$ is the corresponding Lagrangian and β_i is defined by [4]

$$\beta_i = \frac{\partial \mathcal{H}}{\partial p_i} = \frac{\sqrt{p_i^2 + m^2}}{p_i}. \tag{31}$$

Here, β_i is the relativistic velocity of the particle at q_i. Let $\beta_i \to \widetilde{\beta}_i = \beta_i + \delta\beta_i$. Thus, $p_i \to \widetilde{p}_i = p_i + \delta p_i = \dfrac{\delta \mathcal{H}}{\delta \beta_i}$ and the Hamiltonian with transformed momentum satisfies[6]

[6] In the case of a non-relativistic Hamiltonian, the quantities $\mathcal{H}(p_i)$ and $\mathcal{L}(\dot{q}_i)$ are dual in Young's sense, where p_i and $\dot{q}_i$ are proportional by a positive constant (the mass m). When these quantities are candidates for a Legendre transformation, since the potential V plays no role, or is just a spectator [10], by Theorem 3, *ms* can be removed from the sides of Young's inequality, and consequently, the inequality is just satisfied whenever one puts $p_i \dot{q}_i = \mathcal{H}(p_i) + \mathcal{L}(\dot{q}_i)$.

$$\mathcal{H}(\tilde{p}_i) < \tilde{p}_i \tilde{\beta}_i - \mathcal{L}(\tilde{\beta}_i). \tag{32}$$

In general, since we took $\tilde{\beta}_i$ to be increasing over the interval $[\tilde{\beta}_i - \delta\beta_i, \tilde{\beta}]$ the sign of inequality flipped. Since we measure $\mathcal{L}(\tilde{\beta}_i)$ on the given interval, Equation (30) was not maximized with respect to $\tilde{\beta}_i$. So adding a convex function $C(\beta_i, \delta\beta_i) > 0$ to the left-hand side of Equation (32) together with following Theorem 1 yield

$$H(\tilde{p}_i) = \tilde{p}_i \tilde{\beta}_i - L(\tilde{\beta}_i).$$

Now, the canonical variables $(\tilde{p}_i, \tilde{\beta}_i)$ are defined by

$$\tilde{p}_i = \frac{\delta \mathcal{L}}{\delta \beta_i}, \quad \tilde{\beta}_i = \frac{\delta \mathcal{H}}{\delta p_i}.$$

Because we did not have q_i and t involved in the calculation procedure leading to these relations, the other canonical equations of Hamilton are written as usual.

Similarly, one can prove the following result for some function g being convex on $J_i = [\chi_i, \chi_i + \delta_i]$, which also consists of a removable singularity.

Fact 2. Let $\chi_i = \dfrac{f(x_i)x_{i-1} - x_i f(x_{i-1})}{f(x_i) - f(x_{i-1})}$ be the argument of the function g being convex on J_i and let p_i be again of the form (3). If f is also convex on $[x_{i-1}, x_i]$, then the following statements are true:

(1) $g(\chi_i) = p_i \chi_i + C$ such that $C > 0$ is some convex function and $dC / d\chi_i = 0$.

(2) $g^\star(p_i) = p_i \chi_i - g(\chi_i)$ so that $g^\star(p_i) = -C$. □

References

[1] V. I. Arnold, Mathematical Methods of Classical Mechanics, Springer-Verlag, New York, (1978), 14-16.

[2] J. P. Crouzeix and J. A. Ferland, Criteria for quasiconvexity and pseudoconvexity: Relationships and comparisons, Math. Programming 23 (1982), 193-205.

[3] G. Giorgi, A note on quasiconvex functions that are pseudoconvex, Trabajos de Investigación Operativa 2 (1987), 111-116.

[4] H. Goldstein, C. Poole and J. Safko, Classical Mechanics, 3rd Edition, Addison Wesley, (2000), 313.

[5] J.-B. Hiriart-Urruty and C. Lemaréchal, Convex Analysis and Minimization Algorithms I, Springer, 1993.

[6] D. G. Luenberger, Quasi-Convex Programming, SIAM J. Appl. Math. 16 (1968), 1090-1095.

[7] O. L. Mangasarian, Nonlinear Programming, McGraw-Hill, New York, 1969.

[8] B. Martos, Nonlinear Programming, Theory and Methods, North-Holland Pub. Co., 1975.

[9] C. P. Niculescu and L.-E. Persson, Convex Functions and their Applications: A Contemporary Approach, Springer, 2004.

[10] R. K. P. Zia, E. F. Redish and S. R. McKay, Making Sense of the Legendre Transform (2009). arXiv:0806.1147v2

■