Fenchel problem of level sets[1,2,3]

T. Rapcsák[4]

[1]The paper is dedicated to the memory of Guido Stampacchia.

[2]The author wishes to thank F. Giannessi for the advice, and K. Balla for the assistance to solve the system of the differential equations in Example 3.1 and for helpful remarks.

[3]Research was supported in part by the Hungarian Scientific Research Fund, Grants OTKA-T043276, OTKA-T043241 and by CNR, Rome, Italy.

[4]Professor, Computer and Automation Research Institute, Hungarian Academy of Sciences, Budapest, Hungary.

Abstract. In the paper, the Fenchel problem of level sets in the smooth case is solved by deducing a new and "nice" geometric necessary and sufficient condition for the existence of a smooth convex function with the level sets of a given smooth pseudoconvex function. The main theorem is based on a general differential geometric tool, the space of paths defined on smooth manifolds. This approach provides a complete geometric characterization of a new subclass of pseudoconvex functions originated from analytical mechanics, and a new view on the convexlike and generalized convexlike mappings in the image analysis.

KeyWords. Fenchel problem of level sets, convex image transformable functions, pseudoconvexity, space of paths.

1. Introduction

In the paper of two parts, a new and "nice" geometric necessary and sufficient condition will be given for the existence of a smooth convex function with the level sets of a given smooth pseudoconvex function, which is a new solution for the second part of the Fenchel problem of level sets in the smooth case. A survey of the Fenchel problem of level sets can be found in Ref. 1. The main theorem is proved by using a general differential geometric tool, the geometry of paths defined on smooth manifolds which is the subject of the first part of the paper (Ref. 2). This approach provides a complete geometric characterization of a new subclass of pseudoconvex functions originated from analytical mechanics, an extension of the local-global property of nonlinear optimization to nonconvex open sets, a powerful tool - the linear connection which does not depend on either the original data or a Riemannian metric - to improve the structure of a function or a problem from optimization point of view.

In Section 2, a new subclass of the pseudoconvex functions originated from analytical mechanics are characterized by the geometry of paths, in Section 3, a new geometric description of the convex image transformable functions can be found in the smooth case, and in Section 4, some concluding remarks end the paper.

2. A new subclass of the pseudoconvex functions

In the theory of convex and classic generalized convex functions, the most important classes are the convex, pseudoconvex and quasiconvex functions. In this part, the geometric characterization given in Theorem 2.4 and Corollary 2.4 in Ref. 2 is specialized to a new subclass of the pseudoconvex functions originated from analytical mechanics. The most commonly used definition of pseudoconvexity was introduced by Mangasarian (Ref. 3) for differentiable functions.

Definition 2.1. A subset A of the n-dimensional real Euclidean space R^n is a convex set if $\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in A$ for every $\mathbf{x}_1, \mathbf{x}_2 \in A$ and $0 \leq \lambda \leq 1$.

A differentiable function f defined on an open convex set $A \subseteq R^n$ is said to be pseudoconvex if for every $\mathbf{x}_1, \mathbf{x}_2 \in A$,

$$f(\mathbf{x}_1) < f(\mathbf{x}_2) \quad \Rightarrow \quad \nabla f(\mathbf{x}_2)(\mathbf{x}_1 - \mathbf{x}_2) < 0. \tag{1}$$

Let $A \subseteq R^n$ be a convex set, $\mathbf{x} \in A$ and $\mathbf{y} = \mathbf{z} - \mathbf{x}$, $\mathbf{z} \in A$. Then,

$$\mathbf{x}(t) = \mathbf{x} + t\mathbf{y} \in A, \qquad t \in [0,1],$$

and f is convex (pseudoconvex) iff every

$$f(\mathbf{x}(t)), \qquad t \in [0,1],$$

is convex (pseudoconvex).

In analytical mechanics, a motion of a system of mass points is a C^2 vector function

$$\mathbf{x}(t), \qquad t \in [0,1],$$

where $\mathbf{x}'(t)$ and $\mathbf{x}''(t)$, $t \in [0,1]$, is the velocity and acceleration vector at t, respectively. In the case of a convex function, every $\mathbf{x}(t)$, $t \in [0,1]$, is a line segment, and

$$\mathbf{x}'(t) = \mathbf{y}, \qquad \mathbf{x}''(t) = 0, \qquad t \in [0,1].$$

From analytical point of view, this motion is of the same velocity at every t.

If the function f is not convex, but convex along every line segment

$$\mathbf{x}(t) = \mathbf{x} + \varphi_{(\mathbf{x},\mathbf{y})}(t)\mathbf{y} \in A, \qquad \varphi'_{(\mathbf{x},\mathbf{y})}(t) > 0, \qquad t \in [0, \hat{t}_{(\mathbf{x},\mathbf{y})}],$$

for some $\hat{t}_{(\mathbf{x},\mathbf{y})} > 0$, then, from analytical point of view, this motion is of the acceleration $\varphi''_{(\mathbf{x},\mathbf{y})}(t)\mathbf{y}$ at every t. A natural question is how to characterize this class of functions.

The following statement characterizes pseudoconvexity, see, e.g., Crouzeix (Ref. 4).

Theorem 2.1. Let $A \subseteq R^n$ be an open convex set and $f \in C^2(A, R)$. Then, f is pseudoconvex iff

$$\mathbf{x} \in A,\ \nabla f(\mathbf{x}) = 0 \quad \Rightarrow \quad f \ \text{ has a local minimum at } \mathbf{x}, \tag{2}$$

$$\mathbf{x} \in A,\ \mathbf{v} \in R^n,\ \nabla f(\mathbf{x})\mathbf{v} = 0 \quad \Rightarrow \quad \mathbf{v}^T Hf(\mathbf{x})\mathbf{v} \geq 0. \tag{3}$$

In Theorem 2.1, the openness of the set A cannot be excluded.

Let the augmented Hessian matrix of the function $f \in C^2(A, R)$ be given by

$$Hf(\mathbf{x}; r) = Hf(\mathbf{x}) + r\nabla f(\mathbf{x})^T \nabla f(\mathbf{x}), \qquad \mathbf{x} \in A, \quad r \in R. \tag{4}$$

Definition 2.2. Let H, H_c and H_{lL} be the family of functions $f \in C^2(A, R)$ for which a positive semidefinite augmented Hessian matrix with a function $\psi : A \to R$, a continuous function $\psi : A \to R$ and a locally Lipschitz function $\psi : A \to R$ exists at every $\mathbf{x} \in A$, respectively.

By Fenchel (Ref. 5), a necessary condition for the convexifiability of f over an open convex set A is that $f \in H$, and the second one is the pseudoconvexity of f.

It is obvious that if $f \in C^2(A, R)$ is convex on A, then, $\psi(\mathbf{x}) = 0$, $\mathbf{x} \in A$, can be chosen. The family of the functions H_c was introduced by Avriel and Schaible in 1978 (see, Ref. 6) and characterized by Schaible and Zang (Ref. 7), see, Avriel et al. (Ref. 6). They have shown that the family of the functions H_c in H is described by the property that for every compact convex subset $A' \subseteq A$, there exists a real value $r(A')$ such that $H(\mathbf{x}; r(A'))$ is positive semidefinite on $A' \subseteq A$. The family H

is characterized in Avriel et al. (Ref. 6).

The next statement characterizes the new subclass $H_{lL} \subseteq H_c \subset H$.

Theorem 2.2. Let $A \subseteq R^n$ be an open convex set, $f \in C^2(A, R)$ and $\psi : A \to R$ a locally Lipschitz function. Then, $f \in H_{lL}$ iff for every $\mathbf{x} \in A$ there exists a convex neighborhood $U(\mathbf{x}) \subseteq A$ such that for every pair $(\mathbf{x}, \mathbf{y} = \mathbf{z} - \mathbf{x})$, $\mathbf{z} \in A$, the single variable function

$$f\big(\mathbf{x} + \varphi_{(\mathbf{x},\mathbf{y})}(t)\mathbf{y}\big), \qquad \mathbf{x} + \varphi_{(\mathbf{x},\mathbf{y})}(t)\mathbf{y} \in U(\mathbf{x}), \qquad t \in [0, 1], \tag{5}$$

is convex where $\varphi_{(\mathbf{x},\mathbf{y})} : [0, 1] \to R$, $\varphi_{(\mathbf{x},\mathbf{y})}(0) = 0$, $\varphi'_{(\mathbf{x},\mathbf{y})}(0) = 1$, is a strictly increasing function given by the following differential equation:

$$\left(-\frac{1}{\varphi'_{(\mathbf{x},\mathbf{y})}(t)}\right)' = \psi\big(\mathbf{x} + \varphi_{(\mathbf{x},\mathbf{y})}(t)\mathbf{y}\big)\nabla f\big(\mathbf{x} + \varphi_{(\mathbf{x},\mathbf{y})}(t)\mathbf{y}\big)\mathbf{y}, \qquad t \in [0, 1]. \tag{6}$$

Moreover, if $\psi : A \to R_+$, and

$$\nabla f(\mathbf{x})\mathbf{y} > 0, \tag{7}$$

then, $\varphi_{(\mathbf{x},\mathbf{y})}$ is strictly convex.

In the proof of the theorem, two lemmas are used. The first lemma shows that the subclass of the functions H_{lL} is related to local Γ-convexity.

Lemma 2.1. Let $A \subseteq R^n$ be an open set, $f \in C^2(A, R)$ and $\psi : A \to R$ a locally Lipschitz function. Then, $f \in H_{lL}$ with ψ iff f is locally Γ-convex on A with

$$\Gamma^1 = -\psi(\mathbf{x}) \begin{pmatrix} \dfrac{\partial f(\mathbf{x})}{\partial x_1} & \frac{1}{2}\frac{\partial f(\mathbf{x})}{\partial x_2} & \cdots & \dfrac{1}{2}\dfrac{\partial f(\mathbf{x})}{\partial x_n} \\ \dfrac{1}{2}\dfrac{\partial f(\mathbf{x})}{\partial x_2} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{1}{2}\dfrac{\partial f(\mathbf{x})}{\partial x_n} & 0 & \cdots & 0 \end{pmatrix},$$

$$\vdots \tag{8}$$

$$\Gamma^n = -\psi(\mathbf{x}) \begin{pmatrix} 0 & 0 & \cdots & \dfrac{1}{2}\dfrac{\partial f(\mathbf{x})}{\partial x_1} \\ 0 & 0 & \cdots & \dfrac{1}{2}\dfrac{\partial f(\mathbf{x})}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{1}{2}\dfrac{\partial f(\mathbf{x})}{\partial x_1} & \dfrac{1}{2}\dfrac{\partial f(\mathbf{x})}{\partial x_2} & \cdots & \dfrac{\partial f(\mathbf{x})}{\partial x_n} \end{pmatrix},$$

$$\mathbf{x} \in R^n.$$

Proof. Based on the locally Lipschitz function ψ and $f \in C^2$, the linear connection (8) determines a space of paths containing all the Γ-geodesics on A. Since,

$$-\nabla f(\mathbf{x})\Gamma(\mathbf{x}) = -\sum_{i=1}^{n} \frac{\partial f(\mathbf{x})}{\partial x_i}\Gamma^i(\mathbf{x}) = \psi(\mathbf{x})\nabla f(\mathbf{x})^T \nabla f(\mathbf{x}), \qquad \mathbf{x} \in A, \tag{9}$$

and by formulas (17) in Ref. 2, we have that

$$D_\Gamma^2 f(\mathbf{x}) = Hf(\mathbf{x}) + \psi(\mathbf{x})\nabla f(\mathbf{x})^T \nabla f(\mathbf{x}), \qquad \mathbf{x} \in A. \tag{10}$$

By Corollary 2.4 in Ref. 2, f is locally (strictly) Γ-convex on A iff $D_\Gamma^2 f$ is positive semidefinite on A, from which the statement follows. □

It is known that every function $f \in H_c$ is pseudoconvex, e.g., Avriel et al. (Ref. 6). Here, a simple geometric proof is given for the class H_{lL}.

Lemma 2.2. Let $A \subseteq R^n$ be an open convex set, $f \in C^2(A, R)$ and $\psi : A \to R$ a locally Lipschitz function. If f is locally Γ-convex on A with the linear connection (8), then f is pseudoconvex on A.

Proof. If f is locally Γ-convex on A with any linear connection (8), then by Corollary 2.1 in Ref 2, (2) holds, and by Lemma 2.1, $f \in H_{lL}$ with a locally Lipschitz function ψ, from which (3) follows. By Theorem 2.1, (2) and (3) hold iff f is pseudoconvex on A. □

Proof of Theorem 2.2.

(I.) If $f \in H_{lL}$, then by Lemma 2.1, f is locally Γ-convex on A with the linear connection (8). The Γ-geodesics $\gamma(t)^T = \mathbf{x}(t)^T = (x_1(t), \ldots, x_n(t))$, $t \in [0,1]$, can be obtained by solving the following system of differential equations:

$$x_i''(t) = -\mathbf{x}'(t)^T \Gamma^i\big(\mathbf{x}(t)\big)\mathbf{x}'(t) = \psi(\mathbf{x}(t))\cdot$$

$$(x_1'(t), \ldots, x_n'(t)) \begin{pmatrix} & \cdots & \frac{1}{2}\frac{\partial f(\mathbf{x}(t))}{\partial x_1} & \cdots & \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2}\frac{\partial f(\mathbf{x}(t))}{\partial x_1} & \cdots & \frac{\partial f(\mathbf{x}(t))}{\partial x_i} & \cdots & \frac{1}{2}\frac{\partial f(\mathbf{x}(t))}{\partial x_n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ & \cdots & \frac{1}{2}\frac{\partial f(\mathbf{x}(t))}{\partial x_n} & \cdots & \end{pmatrix} \begin{pmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{pmatrix}$$

$$= \psi(\mathbf{x}(t)) \sum_{j=1}^{n} \frac{\partial f(\mathbf{x}(t))}{\partial x_j} x_j'(t) x_i'(t) = \psi(\mathbf{x}(t)) x_i'(t) \nabla f(\mathbf{x}(t))\mathbf{x}'(t), \qquad t \in [0,1], \tag{11}$$

on some subinterval of $[0,1]$.

In order to find all the trajectories $\gamma \in C^2$ that pass an arbitrary point $\mathbf{x} \in A$, different cases must be considered. Let $\mathbf{x}(t)$, $t \in [0,1]$, be the solution with the following initial values:

$$\mathbf{x}(0) = \mathbf{x}, \qquad \mathbf{x}'(0) = \mathbf{y}.$$

(i) Assume that $x_i'(0) \neq 0$, $i = 1.\ldots, n$. Then, due to continuity, there exists an interval $[0, t_0] \subseteq [0, 1]$ such that the functions $\mathbf{x}(t)$,$\mathbf{x}'(t)$,$t \in [0, t_0]$, are sign preserving.

By rewriting system (11), we have that

$$\frac{x_i''(t)}{x_i'(t)} = \psi(\mathbf{x}(t))\nabla f(\mathbf{x}(t))\mathbf{x}'(t), \qquad t \in [0, t_0], \quad i = 1, \ldots, n. \tag{12}$$

It follows that

$$\frac{x_i''(t)}{x_i'(t)} = \frac{x_j''(t)}{x_j'(t)}, \qquad t \in [0, t_0], \qquad \forall (i, j) \in \{1, 2, \ldots, n\}^2, \tag{13}$$

from which

$$\frac{x_k''(t)}{x_k'(t)} = \begin{cases} \left(\ln x_k'(t)\right)' & \text{if} \quad x_k'(0) > 0, \\ \\ \left(\ln(-x_k'(t)\right)' & \text{if} \quad x_k'(0) < 0, \end{cases} \qquad t \in [0, t_0], \quad k = 1, 2, \ldots, n.$$

By integrating equations (13), we have that

$$\ln \frac{x_i'(t)}{x_i'(0)} = \ln \frac{x_j'(t)}{x_j'(0)}, \qquad t \in [0, t_0], \qquad \forall (i,j) \in \{1, 2, \ldots, n\}^2,$$

i.e.,

$$x_i'(t) = \frac{x_i'(0)}{x_j'(0)} x_j'(t), \qquad t \in [0, t_0], \qquad \forall (i,j) \in \{1, 2, \ldots, n\}^2. \tag{14}$$

By integrating (14) on $[0, t] \subseteq [0, t_0]$, we can express every component of the trajectory as a linear function of any other component as follows:

$$x_i(t) = \frac{x_j(t) - x_j(0)}{x_j'(0)} x_i'(0) + x_i(0), \qquad t \in [0, t_0], \qquad \forall (i,j) \in \{1, 2, \ldots, n\}^2,$$

from which

$$\mathbf{x}(t) = \mathbf{x}_0 + \varphi_{(\mathbf{x},\mathbf{y})}(t)\mathbf{y}, \qquad t \in [0, t_0], \tag{15}$$

where

$$\varphi_{(\mathbf{x},\mathbf{y})}(t) = \frac{x_1(t) - x_1(0)}{x_1'(0)}, \qquad t \in [0, t_0].$$

Consequently, $\varphi_{(\mathbf{x},\mathbf{y})}(0) = 0,\ \varphi'_{(\mathbf{x},\mathbf{y})}(0) = 1$.

(ii) Assume that $x_i'(0) \neq 0, i \in I_1 \neq \emptyset$, $x_i'(0) = 0, i \in I_2$, and $I_1 \cup I_2 = \{1, 2, \ldots, n\}$. Then, due to the continuity of the derivatives, there exists an interval $[0, t_0] \subseteq [0, 1]$ such that $\mathbf{x}(t)$,$t \in [0, t_0]$; $x_i'(t)$,$t \in [0, t_0]$, $i \in I_1$, are sign preserving, and due to the uniqueness of the solution, $x_i'(t) = 0$, $t \in [0, t_0]$, $i \in I_2$. Thus, $x_i(t)$, $t \in [0, t_0]$, $i \in I_1$, can be given by (15), while

$$x_i(t) = x_i(0) + \varphi_{(\mathbf{x},\mathbf{y})}(t) x_i'(0) = x_i(0), \qquad t \in [0, t_0], \qquad i \in I_2.$$

(iii) Let $x_i'(0) = 0$, $i = 1, \ldots, n$. Then, there exists an interval $[0, t_0]$ such that

$$x_i'(t) = 0, \qquad t \in [0, t_0], \qquad i = 1, \ldots, n,$$

and the trajectory consists of one point:

$$\mathbf{x}(t) = \mathbf{x}_0, \qquad t \in [0, t_0].$$

Let $\varphi_{(\mathbf{x},0)}(t) = t$, $t \in [0, t_0]$.

By introducing the parameter transformation $\tau = \dfrac{t}{t_0}$, $\tau \in [0, 1]$, in (15), the differential equations do not change but the interval of the parameter becomes $[0, 1]$.

A consequence of the above analysis is that the system of differential equations (11) can be reduced to the following single differential equation:

$$\frac{\varphi''_{(\mathbf{x},\mathbf{y})}(t)}{\left(\varphi'_{(\mathbf{x},\mathbf{y})}(t)\right)^2} = \left(-\frac{1}{\varphi'_{(\mathbf{x},\mathbf{y})}(t)}\right)' = \psi\big(\mathbf{x}+\varphi_{(\mathbf{x},\mathbf{y})}(t)\mathbf{y}\big)\nabla f\big(\mathbf{x}+\varphi_{(\mathbf{x},\mathbf{y})}(t)\mathbf{y}\big)\mathbf{y}, \qquad t\in[0,1]. \tag{16}$$

By (16) and $\varphi'_{(\mathbf{x},\mathbf{y})}(0)=1$, the function $\varphi_{(\mathbf{x},\mathbf{y})}$ is strictly increasing.

By Lemma 2.2, f is pseudoconvex on A, thus, by the first-order characterization of the pseudoconvex functions,

$$\mathbf{x},\mathbf{z}\in A \text{ and } \nabla f(\mathbf{x})(\mathbf{z}-\mathbf{x})>0 \qquad \Rightarrow \qquad \nabla f(\mathbf{z})(\mathbf{z}-\mathbf{x})>0.$$

Let us choose $\mathbf{z}=\mathbf{x}+\varphi_{(\mathbf{x},\mathbf{y})}(t)\mathbf{y}\in A$, then,

$$\nabla f(\mathbf{x})\mathbf{y}>0 \qquad \Rightarrow \qquad \nabla f\big(\mathbf{x}+\varphi_{(\mathbf{x},\mathbf{y})}(t)\mathbf{y}\big)\mathbf{y}>0, \qquad \forall t\in[0,1]. \tag{17}$$

Since $\psi : A \rightarrow R_+$, (16) and (17) result in a strictly convex function $\varphi_{(\mathbf{x},\mathbf{y})} : [0,1]\rightarrow R$.

By the Whitehead theorem, formulas (16) and (17) in Ref. 2, a convex neighbourhood exists around every point $\mathbf{x}\in A$ such that relations (5) and (6) hold.

(II.) If the single variable functions (5) are convex and the differential equations (6) hold, then by twice differentiating (5), we obtain that

$$\mathbf{y}^T Hf\big(\mathbf{x}+\varphi_{(\mathbf{x},\mathbf{y})}(t)\mathbf{y}\big)\mathbf{y}+\frac{\varphi''_{(\mathbf{x},\mathbf{y})}(t)}{\big(\varphi'_{(\mathbf{x},\mathbf{y})}(t)\big)^2}\nabla f\big(\mathbf{x}+\varphi_{(\mathbf{x},\mathbf{y})}(t)\mathbf{y}\big)\mathbf{y}\geq 0, \qquad \mathbf{y}\in R^n. \tag{18}$$

By substituting the right-hand side of (6) for $\dfrac{\varphi''_{(\mathbf{x},\mathbf{y})}}{(\varphi'_{(\mathbf{x},\mathbf{y})})^2}$ in formula (18), we have that $f \in H_{lL}$. □

Corollary 2.1. Let $A \subseteq R^n$ be an open convex set and $f \in C^2(A, R)$. Then, f is convex iff the single variable functions defined by (5), (6) and (7) are locally convex on A for every locally Lipschitz function $\psi : A \to R_{\geq}$.

In the next example, the space of paths is determined based on a linear connection (8). The paths are line segments parametrized by a function φ which is the solution of differential equation (6).

Example 2.1. If

$$f(x,y) = xy, \qquad \psi(x,y) = -\frac{1}{xy}, \qquad (x,y) \in R^2_+,$$

where R^2_+ denotes the positive orthant in R^2, then,

$$\nabla f(x, y) = (y, x), \qquad (x, y) \in R^2_+.$$

The linear connection defined by (8) is as follows:

$$\Gamma^1(x, y) = \frac{1}{xy}\begin{pmatrix} y & \frac{1}{2}x \\ \frac{1}{2}x & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{x} & \frac{1}{2y} \\ \frac{1}{2y} & 0 \end{pmatrix}, \qquad (x, y) \in R^2_+,$$

$$\Gamma^2(x, y) = \frac{1}{xy}\begin{pmatrix} 0 & \frac{1}{2}y \\ \frac{1}{2}y & 0 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{2x} \\ \frac{1}{2x} & \frac{1}{y} \end{pmatrix}, \qquad (x, y) \in R^2_+.$$

The Γ-geodesics $\gamma(t)^T = (x(t), y(t))$, $t \in [0, 1]$, i.e., the space of paths can be obtained by solving the following system of differential equations:

$$x''(t) = (x'(t), y'(t))^T \begin{pmatrix} \frac{1}{x(t)} & \frac{1}{2y(t)} \\ \frac{1}{2y(t)} & 0 \end{pmatrix} \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix}$$

$$= \frac{\left(x'(t)\right)^2}{x(t)} + \frac{x'(t)y'(t)}{y(t)}, \qquad t \in [0, 1], \tag{19a}$$

$$y''(t) = (x'(t), y'(t))^T \begin{pmatrix} 0 & \dfrac{1}{2x(t)} \\ \dfrac{1}{2x(t)} & \dfrac{1}{y(t)} \end{pmatrix} \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix}$$

$$= \frac{\big(y'(t)\big)^2}{y(t)} + \frac{x'(t)y'(t)}{x(t)}, \qquad t \in [0,1]. \tag{19b}$$

In order to find all the trajectories $\gamma \in C^2$ that pass an arbitrary point $(x_0, y_0) \in R^2_+$ different cases must be considered. Let

$$\gamma(t)^T = (x(t), y(t)), \qquad t \in [0,1],$$

be the solution with the following initial values:

$$\gamma(0)^T = (x(0), y(0)), \qquad \gamma'(0)^T = (x'(0), y'(0)).$$

(i) Assume that $x'(0) \neq 0$, $y'(0) \neq 0$ and $C = x'(0)y(0) - y'(0)x(0) \neq 0$, i.e., $\dfrac{x'(0)}{x(0)} \neq \dfrac{y'(0)}{y(0)}$. Then,

$$\frac{\varphi''(t)}{(\varphi'(t))^2} = \frac{\big(y(0) + \varphi(t)y'(0)\big)x'(0) + \big(x(0) + \varphi(t)x'(0)\big)y'(0)}{\big(x(0) + \varphi(t)x'(0)\big)\big(y(0) + \varphi(t)y'(0)\big)}$$

$$= \frac{x'(0)}{x(0) + \varphi(t)x'(0)} + \frac{y'(0)}{y(0) + \varphi(t)y'(0)}, \qquad t \in [0,1]. \tag{20}$$

Note that we seek for the solution with $\varphi(0) = 0$ and $\varphi'(0) = 1$. By integrating the equation

$$\Big(\ln \varphi'(t)\Big)' = \Big(\ln\big(x(0) + \varphi(t)x'(0)\big)\Big)' + \Big(\ln\big(y(0) + \varphi(t)y'(0)\big)\Big)', \qquad t \in [0,1], \tag{21}$$

we obtain that

$$\varphi'(t) = \frac{x(0) + \varphi(t)x'(0)}{x(0)} \cdot \frac{y(0) + \varphi(t)y'(0)}{y(0)}, \qquad t \in [0,1], \tag{22}$$

By using $C \neq 0$ and the relation

$$\frac{\varphi'(t)}{(x(0) + \varphi(t)x'(0))(y(0) + \varphi(t)y'(0))} = \frac{1}{C}\left(\frac{\varphi'(t)x'(0)}{x(0) + \varphi(t)x'(0)} - \frac{\varphi'(t)y'(0)}{y(0) + \varphi(t)y'(0)}\right)$$

$$= \frac{1}{x(0)y(0)}, \qquad t \in [0,1], \tag{23}$$

and by integrating (23), we get that there exists a $t_0 > 0$ such that

$$\varphi(t) = \frac{e^{\left(\frac{x'(0)}{x(0)} - \frac{y'(0)}{y(0)}\right)t} - 1}{\frac{x'(0)}{x(0)} - \frac{y'(0)}{y(0)} e^{\left(\frac{x'(0)}{x(0)} - \frac{y'(0)}{y(0)}\right)t}}, \qquad t \in [0, t_0]. \tag{24}$$

It follows that $\varphi'(t) > 0$, $t \in [0, t_0]$, consequently, φ is strictly increasing.

(ii) Assume that $x'(0) \neq 0$, $y'(0) \neq 0$ and $C = x'(0)y(0) - y'(0)x(0) = 0$. Then, from (22),

$$\frac{\varphi'(t)}{\left(1 + \frac{x'(0)}{x(0)}\varphi(t)\right)^2} = -\frac{x(0)}{x'(0)}\left(\frac{1}{1 + \frac{x'(0)}{x(0)}\varphi(t)}\right)' = 1, \qquad t \in [0,1]. \tag{25}$$

By integrating (25), we have that there exists a $t_0 > 0$ such that

$$\varphi(t) = \frac{t}{1 - \frac{x'(0)}{x(0)}t}, \qquad t \in [0, t_0]. \tag{26}$$

(iii) Assume that $x'(0) \neq 0$ and $y'(0) = 0$. Then, from (22),

$$\frac{x'(0)\varphi'(t)}{x(0) + \varphi(t)x'(0)} = \Big(\ln(x(0) + \varphi(t)x'(0))\Big)' = \frac{x'(0)}{x(0)}, \qquad t \in [0,1]. \tag{27}$$

By integrating (27), we have that there exists a $t_0 > 0$ such that

$$\varphi(t) = \frac{x(0)}{x'(0)}\left(e^{\frac{x'(0)}{x(0)}t} - 1\right), \qquad t \in [0, t_0]. \tag{28}$$

(iv) Assume that $x'(0) = 0$ and $y'(0) \neq 0$. Exactly as before, we get that

$$\varphi(t) = \frac{y(0)}{y'(0)}\left(e^{\frac{y'(0)}{y(0)}t} - 1\right), \qquad t \in [0, t_0]. \tag{29}$$

(v) Let $x'(0)=0$ and $y'(0)=0$. Then, from (20),

$$\varphi(t)=t, \qquad t\in[0,1]. \tag{30}$$

We remark that the continuation of solution (24) can be determined. If $\dfrac{x'(0)}{y'(0)}<0$, then φ exists on $(-\infty,\infty)$; if $\dfrac{x'(0)}{y'(0)}>1$, $C>0$, then φ exists on $(-\infty,\hat{t})$, $\hat{t}>0$; if $0<\dfrac{x'(0)}{y'(0)}<1$, $C>0$, then φ exists on $(\tilde{t},\infty)$, $\tilde{t}<0$; if $\dfrac{x'(0)}{y'(0)}>1$, $C<0$, then φ exists on $(\bar{t},\infty)$, $\bar{t}<0$; if $0<\dfrac{x'(0)}{y'(0)}<1$, $C<0$, then φ exists on $(-\infty,t^*)$, $t^*>0$; and the values $\hat{t},\tilde{t},\bar{t},t^*$ can be calculated.

3. Convex image transformable functions

An important subclass of the smooth pseudoconvex functions consists of the convex twice continuously differentiable (C^2) functions obtained by one-to-one increasing image transformations belonging to the convex image transformable functions, or in other terminology, the G-convex functions. A famous result, the first complete set of necessary and sufficient conditions for the convexifiability of C^2 functions was derived by Fenchel (Refs. 5,8) developed further in different directions, the detailed analysis of which can be found in the literature, see, e.g, Avriel et al. (Ref. 6). In this part, the main result is a new geometric characterization of the smooth convex image transformable functions, based on the space of paths. Let $\mathrm{Im}_f(A)$ denote the image of the function f on the set $A \subseteq R^n$.

Definition 3.1. A nonconvex function is convex image transformable if it can be transformed into a convex function by a one-to-one increasing or decreasing transformation of its image (their ranges).

Theorem 3.1. Let $f \in H_{lL}$ be a real-valued function defined on an open convex set $A \subseteq R^n$. Then, f is convex image transformable by a one-to-one increasing function $\phi \in C^2(\mathrm{Im}_f(A), R)$ iff for every $\mathbf{x} \in A$, there exists a convex neighborhood $U(\mathbf{x}) \subseteq A$ such that for every pair $(\mathbf{x}, \mathbf{y} = \mathbf{z} - \mathbf{x})$, $\mathbf{z} \in A$, the single variable function

$$f\big(\mathbf{x} + \varphi_{(\mathbf{x},\mathbf{y})}(t)\mathbf{y}\big), \qquad \mathbf{x} + \varphi_{(\mathbf{x},\mathbf{y})}(t)\mathbf{y} \in U(\mathbf{x}), \qquad t \in [0,1], \tag{31}$$

is convex where $\varphi_{(\mathbf{x},\mathbf{y})} : [0,1] \to R$, $\varphi_{(\mathbf{x},\mathbf{y})}(0) = 0$, $\varphi'_{(\mathbf{x},\mathbf{y})}(0) = 1$, is a strictly increasing function given by the following differential equation:

$$\varphi'_{(\mathbf{x},\mathbf{y})}(t) = \frac{1}{\phi'\big(f(\mathbf{x})\big)}\phi'\big(f(\mathbf{x}+\varphi_{(\mathbf{x},\mathbf{y})}(t)\mathbf{y})\big), \qquad t \in [0,1]. \tag{32}$$

Moreover, if $\phi : A \to R_+$, and

$$\nabla f(\mathbf{x})\mathbf{y} > 0, \tag{33}$$

then, $\varphi_{(\mathbf{x},\mathbf{y})}$ is strictly convex.

The proof of the theorem is based on two lemmas. In the following well-known characterization of the convex image transformable functions, twice differentiability is used.

Lemma 3.1. Let $A \subseteq R^n$ be an open convex set and $f \in C^2(A,R)$. Then, f is convex image transformable by a one-to-one increasing or decreasing function $\phi \in C^2(\mathrm{Im}_f(A), R)$ iff the matrices

$$H\phi(f(\mathbf{x})) = \phi'(f(\mathbf{x}))Hf(\mathbf{x}) + \phi''(f(\mathbf{x}))\nabla f(\mathbf{x})^T \nabla f(\mathbf{x}), \qquad \mathbf{x} \in A, \tag{34}$$

are positive semidefinite on A, where ∇f and Hf denote the gradient (row) vector and the Hessian matrix, respectively.

Lemma 3.2. Let $A \subseteq R^n$ be an open convex set and $f : A \to R, f \in H_{lL}$ or $-f \in H_{lL}$. Then, f is (strictly) convex image transformable by a one-to-one increasing or decreasing function $\phi \in C^2(\mathrm{Im}_f(A), R)$ iff there exists a locally Lipschitz function $\theta \in C(\mathrm{Im}_f(A), R)$ such that $f \in H_{lL}$ or $-f \in H_{lL}$ with the function $\psi(\mathbf{x}) = \theta(f(\mathbf{x}))$, $\mathbf{x} \in A$.

Proof.

(I.) Let the function f be convex image transformable by a one-to-one increasing or decreasing function $\phi \in C^2(\mathrm{Im}_f(A), R)$. Then, by Lemma 3.1, $f \in H_{lL}$ iff ϕ is increasing, and $-f \in H_{lL}$ iff ϕ is decreasing, moreover,

$$\psi(\mathbf{x}) = \theta(f(\mathbf{x})) = \frac{\phi''(f(\mathbf{x}))}{\phi'(f(\mathbf{x}))}, \qquad \mathbf{x} \in A, \tag{35}$$

in the former case, and

$$\psi(f(\mathbf{x})) = \theta(f(\mathbf{x})) = -\frac{\phi''(f(\mathbf{x}))}{\phi'(f(\mathbf{x}))}, \qquad \mathbf{x} \in A, \tag{36}$$

in the latter case.

(II.) Let $f \in H_{lL}$ or $-f \in H_{lL}$ with a $\theta(f(\mathbf{x}))$, $\mathbf{x} \in A$. It will be shown in the increasing case that the function $f \in C^2(A, R)$ is (strictly) convex image transformable by

$$\phi(t) = C_1 \int e^{\int \theta(\tau) d\tau} dt + C_2, \qquad t \in \mathrm{Im}_f(A), \qquad C_1 > 0, \quad C_2 \in R. \tag{37}$$

By Lemma 3.1, it is sufficient to prove that the differential equation

$$\frac{\phi''(t)}{\phi'(t)} = \theta(t), \qquad t \in \mathrm{Im}_f(A), \tag{38}$$

can be explicitly solved. Since,

$$\frac{\phi''(t)}{\phi'(t)} = \begin{cases} \left(\ln \phi'(t)\right)' & \text{if} \quad \phi'(t) > 0, \\ \\ \left(\ln(-\phi'(t)\right)' & \text{if} \quad \phi'(t) < 0, \end{cases} \qquad t \in \mathrm{Im}_f(A),$$

$$\phi'(t) = C_1 e^{\int \theta(t) dt}, \qquad t \in \mathrm{Im}_f(A), \qquad C_1 > 0,$$

$$\phi(t) = C_1 \int e^{\int \theta(\tau) d\tau} dt + C_2, \qquad t \in \mathrm{Im}_f(A), \qquad C_1 > 0, \quad C_2 \in R,$$

the statement is proved. The proof of the decreasing case ($-f \in H_{lL}$) is similar.

$\square$

Proof of Theorem 3.1. Let $A \subseteq R^n$ be an open convex set and $f : A \to R, f \in H_{lL}$. Then, by Lemma 3.2, f is (strictly) convex image transformable by a one-to-one increasing function $\phi \in C^2(\mathrm{Im}_f(A), R)$ iff $f \in H_{lL}$ with a function $\psi(\mathbf{x}) = \theta(f(\mathbf{x}))$, $\mathbf{x} \in A$, where $\theta \in C(\mathrm{Im}_f(A), R)$ is a locally Lipschitz function.

By Theorem 2.2, formulas (5), (6) and (7) are valid, therefore, by (35) and (38), the following differential equation holds:

$$\left(-\frac{1}{\varphi'_{(\mathbf{x},\mathbf{y})}(t)}\right)' = \frac{\phi''\big(f(\mathbf{x}+\varphi_{(\mathbf{x},\mathbf{y})}(t)\mathbf{y})\big)}{\phi'\big(f(\mathbf{x}+\varphi_{(\mathbf{x},\mathbf{y})}(t)\mathbf{y})\big)} \nabla f\big(\mathbf{x}+\varphi_{(\mathbf{x},\mathbf{y})}(t)\mathbf{y}\big)\mathbf{y}, \qquad t \in [0,1], \tag{39}$$

which is equivalent to

$$\left(\ln \varphi'_{(\mathbf{x},\mathbf{y})}(t)\right)' = \left(\ln \phi'\big(f(\mathbf{x}+\varphi_{(\mathbf{x},\mathbf{y})}(t)\mathbf{y})\big)\right)', \qquad t \in [0,1], \tag{40}$$

from which

$$\varphi'_{(\mathbf{x},\mathbf{y})}(t) = \frac{1}{\phi'\big(f(\mathbf{x})\big)} \phi'\big(f(\mathbf{x}+\varphi_{(\mathbf{x},\mathbf{y})}(t)\mathbf{y})\big), \qquad t \in [0,1], \tag{41}$$

which is the statement. □

Example 3.1. If

$$f(x,y) = xy, \qquad (x,y) \in R^2_+, \qquad \phi(t) = -\ln t, \qquad t \in (0,+\infty),$$

where ϕ is a one-to-one decreasing image transformation function and R^2_+ denotes the positive orthant in R^2, then,

$$\phi'(t) = -\frac{1}{t}, \qquad t \in (0,+\infty).$$

By using the notations of Example 2.1, from (32) we obtain that

$$\varphi'(t) = \frac{x(0)y(0)}{\big(x(0)+\varphi(t)x'(0)\big)\big(y(0)+\varphi(t)y'(0)\big)}, \qquad t \in [0,1]. \tag{42}$$

We remark that equation (40) is the inverse of (22) because the image transformation is decreasing.

4. Concluding remarks

In the paper, it is shown that the smooth convex and the smooth convex image transformable functions, as well as a new subclass of the pseudoconvex functions are of the same character and can be originated from the geometry of paths based on smooth manifolds. The common root is that these types of convexity are of geodesic convexity, but while convexity is related to line segments, the geodesic segments of the Euclidean metric, convex image transformability and a subclass of pseudoconvexity to a system of paths generated by a linear connection determined by the gradient of the given function and the image transformation or a locally Lipschitz function, respectively. An important remark is that these convexity notions are based on the differentiable structure of the manifolds but not on the metric structure. The main result related to convex image transformability is a new geometric solution of the Fenchel problem of level sets in the smooth case. An open question is how to weaken the assumptions of this result.

Some further open questions are as follows:

(Q1) How large is the difference between pseudoconvex and convex functions?

(Q2) How can the geometric characterization of H_{lL} be extended to H_c? $H_{lL} \subset H_c$?

(Q3) How to characterize the subclass of the pseudoconvex functions where the single variable functions determined by (5), (6) and (7) are not only locally, but globally convex in the given convex set?

(Q4) How to solve the Fenchel problem of level sets in the case of neither closed nor open convex sets?

References

(1) Rapcsák, T., Survey of the Fenchel Problem of Level Sets, Variational Analysis and Applications, Edited by F. Giannessi and A. Maugeri, Kluwer Academic Publishers, Dordrecht, Holland. (in print)

(2) Rapcsák, T., Geometry of paths and the Fenchel problem of level sets I., Journal of Optimization Theory and Application. (accepted for publication)

(3) Mangasarian, O. L., Pseudo-Convex Functions, SIAM Journal on Control, Vol. 3, pp. 281-290, 1965.

(4) Crouzeix, J.P., Characterizations of Generalized Convexity and Generalized Monotonicity, A Survey, Generalized Convexity, Generalized Monotonicity, Edited by J.P. Crouzeix, J.E. Martinez-Legaz and M. Volle, Kluwer Academic Publishers, Dordrecht, Holland, pp. 237-256, 1998.

(5) Fenchel, W., Convex Cones, Sets and Functions, Mimeographed Lecture Notes, Princeton University Press, Princeton, New Jersey, 1951.

(6) Avriel, M., Diewert, W.E., Schaible, S. and Zang, I., Generalized Concavity, Plenum Press, New York, 1988.

(7) Schaible, S. and Zang, I., On the Convexifiability of Pseudoconvex C^2-Functions, Mathematical Programming, Vol. 19, pp. 289-299, 1980.

(8) Fenchel, W., Über Konvexe Funktionen mit vorgeschriebenen Niveaumannigfaltigkeiten, Mathematische Zeitschrift, Vol. 63, pp. 496-506, 1956.