Journal of Convex Analysis
Volume 8 (2001), No. 1, 241–253

Convexity Properties of Some Implicit Functions

U. Würker
Faculty of Mathematics, Chemnitz University of Technology,
09107 Chemnitz, Germany.
e-mail: wuerker@mathematik.tu-chemnitz.de

Received August 26, 1999

We consider implicit functions $y = y(x)$ defined by a system of equations $G_i(x, y) = 0$, $i = 1, \ldots, m$. In the case of convex differentiable functions G_i we establish some sufficient conditions under which the component function $y_k(x)$ is convex or concave. Examples show that without these assumptions $y_k(x)$ can be nonconvex and nonconcave.

For the special case with additive separated convex functions $G_i(x, y) = g_i(x) + h_i(y)$ additional results concerning the gradients ∇g_i and ∇h_i are obtained which can be applied to the differentiable continuation of convex marginal functions in parametric optimization.

Keywords: convex function, implicit function, convex parametric optimization

1991 Mathematics Subject Classification: 26B10, 26B25, 52A20, 90C25

1. Introduction

In convex optimization often one has to consider systems of equations where convex functions are involved. A typical example is the Karush-Kuhn-Tucker system for a convex programming problem which is linear w.r.t. the Lagrange multipliers and nonlinear w.r.t. the optimization variables in the (convex) constraints, gradient and complementarity relations.

We consider the following system of equations

$$G(x, y) = \mathbf{0}\,, \qquad G(x, y) = (\,G_1(x, y), \ldots, G_m(x, y)\,) \tag{1}$$

where $G : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^m$, $n, m \geq 1$, is a continuously differentiable function. As it is well known, the solution set of analogous inequalities $G(x, y) \leq \mathbf{0}$ with convex functions G_i is a convex subset of $\mathbb{R}^{n+m}$. In contrast to this, in general the solution set of the equations (1) does not have any convexity properties even if G is convex. Geometrically, this set is the intersection of the boundaries of convex bodies, whereby quite differently constituted curves and hypersurfaces can occur.

In the literature one can find many papers about the implicit function $y = y(x)$ defined by (1) under suitable additional assumptions, but there are only very few results about convexity properties of $y(x)$. Some investigations for the case $m = 1$ (one convex equation) are part of the papers [1] and [5]. Furthermore, the book [7] contains some related results w.r.t. the boundary of convex sets from the geometrical point of view (also for $m = 1$ only).

For our examination of convexity properties of such solution functions first we cite a general implicit function theorem.

ISSN 0944-6532 / \$ 2.50

Theorem 1.1. *Let $G(x,y)$ be a C^1 function defined in a neighborhood of $\overline{x} \in \mathbb{R}^n$ and $\overline{y} \in \mathbb{R}^m$ taking values in $\mathbb{R}^m$, with $G(\overline{x},\overline{y}) = \mathbf{0}$. Then if $\nabla_y G(\overline{x},\overline{y})$ is invertible there exists a neighborhood U of $\overline{x}$ and a C^1 function $y : U \to \mathbb{R}^m$ such that $\overline{y} = y(\overline{x})$ and $G(x, y(x)) = \mathbf{0}$ for every $x \in U$.*

Furthermore, y is unique in that there exists a neighborhood V of $\overline{y}$ such that there is only one solution $z \in V$ of $G(x,z) = \mathbf{0}$, namely $z = y(x)$. Finally, the differential of y can be computed by implicit differentiation from

$$\nabla_x G_i(x, y(x)) + \sum_{j=1}^{m} \frac{\partial}{\partial y_j} G_i(x, y(x)) \cdot \nabla y_j(x) = \mathbf{0}\,, \quad i = 1, \ldots, m, \quad \forall\, x \in U\,. \tag{2}$$

Remark. For the proof see for instance [3] or [9].

2. Systems of convex equations

Theorem 2.1. *Let the assumptions of Theorem 1.1 be fulfilled, and assume that the functions G_i, $i = 1, \ldots, m$, are convex on the considered neighborhood of $(\overline{x},\overline{y})$. Furthermore, let a $(n+1)$-dimensional affine manifold $H \subset \mathbb{R}^{n+m}$ exist such that*

$$Y := \left\{\ (x,y) \in \mathbb{R}^{n+m} \ : \ y = y(x)\,,\ x \in U\ \right\} \subseteq H\,,\ \dim(H) = n+1\,. \tag{3}$$

Then for each $i = 1, \ldots, m$ the function $y_i = y_i(x)$ is convex or concave on U.

Proof. Because of the structure (3) of the set Y we have the existence of two vectors $\overline{a}, \overline{d} \in \mathbb{R}^m$ such that

$$H = \left\{\ (x,y) \in \mathbb{R}^{n+m} \ : \ x \in \mathbb{R}^n\,,\ y = \overline{a} + t \cdot \overline{d}\,,\ t \in \mathbb{R}^1\ \right\}\,. \tag{4}$$

Then in the case $\overline{d} = \mathbf{0}$ the assertion of the Theorem is obviously true because of $y(x) = \overline{a}$ $\forall\, x \in U$.

In the case $\overline{d} \neq \mathbf{0}$ we can conclude from (4) that a uniquely defined function $t : U \to \mathbb{R}^1$ exists with

$$Y = \left\{\ (x,y) \in \mathbb{R}^{n+m} \ : \ y = y(x) = \overline{a} + t(x) \cdot \overline{d}\,,\ x \in U\ \right\}\,. \tag{5}$$

Because of the differentiability of $y(x)$ on U the function $t(x)$ is differentiable on U, too. Defining the sets

$$\begin{aligned} C_i &= \{\ (x,y) \in \mathbb{R}^n \times \mathbb{R}^m \ : \ G_i(x,y) \leq 0\ \}\,, \\ D_i &= \{\ (x,y) \in C_i \ : \ G_i(x,y) = 0\ \}\,, \\ C &= \bigcap_{i=1}^{m} C_i\,, \quad D = \bigcap_{i=1}^{m} D_i\,, \end{aligned} \tag{6}$$

we find that C_i and $C_i \cap H$ are convex for each $i = 1, \ldots, m$. Furthermore we have $(\overline{x},\overline{y}) \in D_i \subseteq C_i$ and $\nabla_y G_i(\overline{x},\overline{y}) \neq \mathbf{0}$. Hence, Theorem 17.5 from [7] provides that C_i is a $(n+m)$-dimensional closed convex set, D_i is the boundary $\partial C_i = \mathrm{cl}(C_i) \setminus \mathrm{int}(C_i)$ of C_i

and a $(n+m-1)$-dimensional regular manifold of the $\mathbb{R}^{n+m}$. For each point $(x^*, y^*) \in D_i$ there is a unique supporting halfspace $R_i^+(x^*, y^*)$ to C_i at (x^*, y^*) with

$$R_i^+(x^*, y^*) = \left\{ \ (x,y) \in \mathbb{R}^{n+m} \ : \ \left\langle \begin{pmatrix} \nabla_x G_i(x^*, y^*) \\ \nabla_y G_i(x^*, y^*) \end{pmatrix}, \begin{pmatrix} x - x^* \\ y - y^* \end{pmatrix} \right\rangle \leq 0 \ \right\}. \tag{7}$$

Hence we have $D_i \subseteq C_i \subseteq R_i^+(x^*, y^*)$ and $D_i \cap H \subseteq R_i^+(x^*, y^*) \cap H$ with

$$\begin{aligned} R_i^+(x^*, y^*) \cap H = \{ \ & (x,y) \in \mathbb{R}^{n+m} \ : \ x \in \mathbb{R}^n \ , \ y = \overline{a} + t \cdot \overline{d} \ , \ t \in \mathbb{R}^1 \ , \\ & \left\langle \begin{pmatrix} \nabla_x G_i(x^*, y^*) \\ \nabla_y G_i(x^*, y^*) \end{pmatrix}, \begin{pmatrix} x - x^* \\ y - y^* \end{pmatrix} \right\rangle \leq 0 \ \right\}. \end{aligned} \tag{8}$$

Because of $Y \subseteq C_i \cap H \subseteq R_i^+(x^*, y^*) \cap H$ we can choose $(x, y(x)) \in Y$ and $(x^*, y(x^*)) \in Y$ in (8) which leads us to

$$\langle \nabla_x G_i(x^*, y^*), x - x^* \rangle + \left\langle \nabla_y G_i(x^*, y^*), \overline{a} + t(x) \cdot \overline{d} - y^* \right\rangle \leq 0 \tag{9}$$

$\forall \ x \in U$. With respect to $y^* = \overline{a} + t(x^*) \cdot \overline{d}$ from this we get

$$\langle \nabla_x G_i(x^*, y^*), x - x^* \rangle + (t(x) - t(x^*)) \cdot \left\langle \nabla_y G_i(x^*, y^*), \overline{d} \right\rangle \leq 0 \, . \tag{10}$$

Since the vectors $\nabla_y G_i(x^*, y^*)$, $i = 1, \ldots, m$, are linearly independent there is at least one index $i_0 \in \{1, \ldots, m\}$ such that the product $s^* := \left\langle \nabla_y G_{i_0}(x^*, y^*), \overline{d} \right\rangle$ is nonzero. Without loss of generality we can suppose that either $s^* > 0$ or $s^* < 0$ is true for all $x^* \in U$. Hence we can multiply relation (10) with $\dfrac{1}{s^*}$ and we get

$$t(x) - t(x^*) + \frac{1}{s^*} \langle \nabla_x G_{i_0}(x^*, y^*), x - x^* \rangle \begin{cases} \leq 0 \, , & s^* > 0 \, , \\ \geq 0 \, , & s^* < 0 \, , \end{cases} \qquad \forall x, x^* \in U \, . \tag{11}$$

From this it follows directly that the function $t(x)$ is convex or concave on U (depending on the sign of s^*). Considering $y_j(x) = \overline{a}_j + t(x) \cdot \overline{d}_j$, $j = 1, \ldots, m$, from the convexity/concavity of $t(x)$ we get also one of the two properties for the function $y_j(x)$ on U (depending on the sign of the factor $\overline{d}_j$). $\square$

Remark. The assertion of Theorem 2.1 does not mean that all functions $y_i = y_i(x)$, $i = 1, \ldots, m$, are either convex or concave. In general only some functions $y_i(x)$, $i \in I_1$, are convex and the remaining functions $y_i(x)$, $i \in \{1, \ldots, m\} \setminus I_1$, are concave. This is illustrated by the following two examples.

Example 2.2. We consider the system (1) with $n = 1$, $m = 2$ and

$$\begin{aligned} G_1(x,y) &= x^2 + (y_1 - 1)^2 + (y_2 + 1)^2 - 4 \, , \\ G_2(x,y) &= x^2 + (y_1 + 1)^2 + (y_2 - 1)^2 - 4 \, . \end{aligned}$$

For $\overline{x} = 0$ and $\overline{y}_1 = \overline{y}_2 = \pm 1$ all assumptions of Theorem 2.1 are fulfilled, and we get

$$\begin{array}{lll} \text{(a1)} \quad \overline{y} = (1,1) & \Longrightarrow & y_1(x) = \sqrt{1 - \frac{x^2}{2}} \, , \quad y_2(x) = \sqrt{1 - \frac{x^2}{2}} \, ; \\ \text{(a2)} \quad \overline{y} = (-1,-1) & \Longrightarrow & y_1(x) = -\sqrt{1 - \frac{x^2}{2}} \, , \quad y_2(x) = -\sqrt{1 - \frac{x^2}{2}} \, . \end{array}$$

In the case (a1) the functions $y_1(x)$ and $y_2(x)$ are concave, but in the case (a2) both functions are convex (in a neighborhood of $\overline{x}$, respectively).

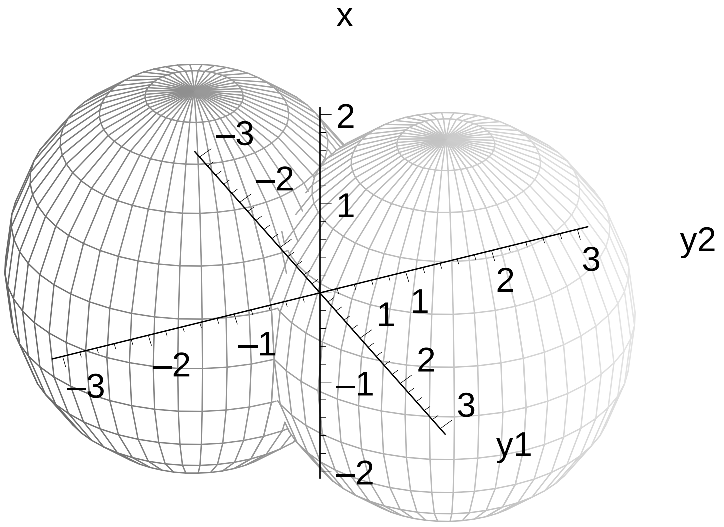

Figure 2.1: Balls $G_i(x, y) = 0$ for Example 2.3

Example 2.3. Let in (1) the relations $n = 1$, $m = 2$ and

$$G_1(x,y) = x^2 + (y_1+1)^2 + (y_2+1)^2 - 4\,,$$
$$G_2(x,y) = x^2 + (y_1-1)^2 + (y_2-1)^2 - 4$$

be valid. Then for $\overline{x} = 0$ and $\overline{y}_1 = -\overline{y}_2 = \pm 1$ all assumptions of Theorem 2.1 are fulfilled with

$$\text{(b1)}\quad \overline{y} = (1,-1) \quad\Longrightarrow\quad y_1(x) = \sqrt{1-\tfrac{x^2}{2}}\,,\quad y_2(x) = -\sqrt{1-\tfrac{x^2}{2}}\,;$$
$$\text{(b2)}\quad \overline{y} = (-1,1) \quad\Longrightarrow\quad y_1(x) = -\sqrt{1-\tfrac{x^2}{2}}\,,\quad y_2(x) = \sqrt{1-\tfrac{x^2}{2}}\,.$$

In both cases (b1) and (b2) one convex and one concave function $y_i(x)$ appears (in a neighborhood of $\overline{x}$, respectively).

Remark. In the case $m > 1$ the additional assumption in Theorem 2.1 (dimension $(n+1)$ of the set Y) seems to be very restrictive because in general the curve $(x, y(x))$ is a subset of $\mathbb{R}^{n+m}$. But the following example shows even the dimension $(n+2)$ of Y can cause the appearance of nonconvex and nonconcave implicit functions $y_i(x)$.

Example 2.4. Let in (1) the relations $n = 1$, $m = 2$ and

$$G_1(x,y) = \sqrt{x^2+y_1^2}\cdot\left(20 + (\arctan\tfrac{x}{y_1})^3\right) - 20 - y_2\,,$$
$$G_2(x,y) = \sqrt{x^2+y_1^2}\cdot\left(20 - 2(\arctan\tfrac{x}{y_1})^3\right) - 20 + 2y_2$$

hold (cf. Figure 2.2). Then the functions $G_i(x, y)$, $i = 1, 2$, are convex and continuously differentiable in a neighborhood of $\overline{x} = 0$ and $\overline{y} = (1, 0)$. Because of

$$\nabla G_1(x,y) = \left(\frac{20x + x(\arctan\frac{x}{y_1})^3 + 3y_1(\arctan\frac{x}{y_1})^2}{\sqrt{x^2+y_1^2}},\ \frac{20y_1 + y_1(\arctan\frac{x}{y_1})^3 - 3x(\arctan\frac{x}{y_1})^2}{\sqrt{x^2+y_1^2}},\ -1\right)$$
$$\nabla G_2(x,y) = \left(\frac{20x - 2x(\arctan\frac{x}{y_1})^3 - 6y_1(\arctan\frac{x}{y_1})^2}{\sqrt{x^2+y_1^2}},\ \frac{20y_1 - 2y_1(\arctan\frac{x}{y_1})^3 + 6x(\arctan\frac{x}{y_1})^2}{\sqrt{x^2+y_1^2}},\ 2\right)$$

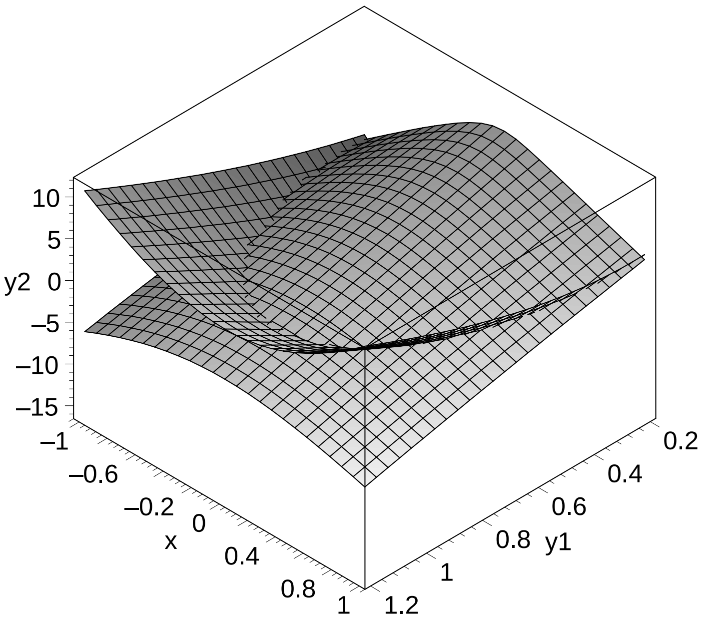

Figure 2.2: Surfaces $G_i(x, y) = 0$ for Example 2.4

we find that $\det(\nabla_y G(\overline{x}, \overline{y})) = \begin{vmatrix} 20 & -1 \\ 20 & 2 \end{vmatrix} = 60 \neq 0$, such that all assumptions of Theorem 1.1 are fulfilled, and we get

$$y_1(x) = \sqrt{1 - x^2}\,, \quad y_2(x) = \left(\arctan \frac{x}{\sqrt{1 - x^2}}\right)^3 , \quad \forall\, x \in U(\overline{x})$$

(cf. Figure 2.3). But the (threedimensional) set Y does not fulfill the additional assumption of Theorem 2.1. Therefore, the function $y_2(x)$ is neither convex nor concave (w.r.t. an arbitrary neighborhood of $\overline{x} = 0$).

For additional properties we need the following lemma on convex, finite generated dual cones from linear algebra.

Lemma 2.5. *Let a^i, $i = 1, \ldots, m$, be linearly independent vectors of $\mathbb{R}^N$, and let*

$$\begin{aligned} K &= \operatorname{cone}(a^1, \ldots, a^m)\,, \\ K' &= \{\ z \in \mathbb{R}^N\ :\ \langle z, a^i \rangle = 0\,,\ i = 1 \ldots, m\ \}\,, \\ K^* &= \{\ z \in \mathbb{R}^N\ :\ \langle z, a^i \rangle \leq 0\,,\ i = 1 \ldots, m\ \} \end{aligned} \tag{12}$$

be the convex cone generated by the a^i, the maximal subspace orthogonal to all a^i and the dual cone to K. Furthermore, let the matrices

$$D = (a_{ij})_{i,j=1}^m\,, \quad a_{ij} = \langle a^i, a^j \rangle\,, \qquad D^{-1} = (b_{ij})_{i,j=1}^m \tag{13}$$

be given. Then m linearly independent vectors z^i, $i = 1, \ldots, m$, exist such that

$$K^* = K' + \operatorname{cone}(z^1, \ldots, z^m)\,, \quad z^i = -\sum_{j=1}^m\ b_{ij} \cdot a^i\,,\ i = 1, \ldots, m. \tag{14}$$

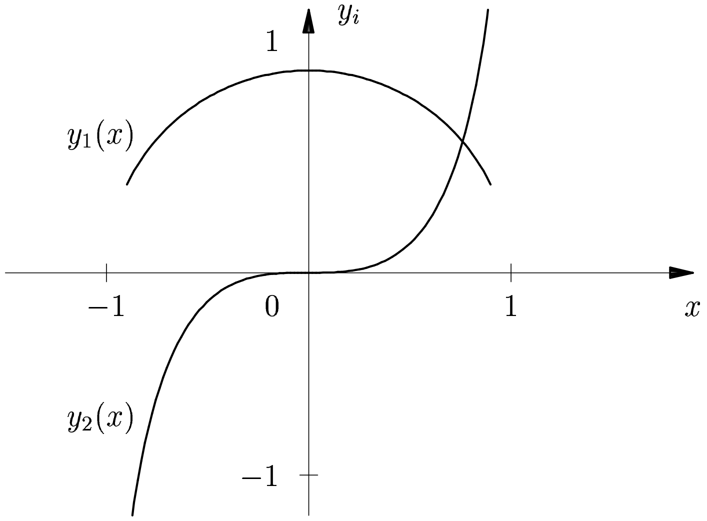

Figure 2.3: Solution functions for Example 2.4

Proof. Let A be the matrix containing the column vectors a^1 to a^m (with rank m). Then from linear algebra we know that the dual cone $K^* = \{ z \in \mathbb{R}^N : z^\top A \leq \mathbf{0} \}$ is representable as algebraic sum of the subspace $K' = \{ z \in \mathbb{R}^N : z^\top A = \mathbf{0} \}$ and a finite generated cone $K^0 = \text{cone}(y^1, \ldots, y^k)$. Hereby the extreme rays $y^s \in K^*$ are uniquely defined (despite of a positive factor) by the rank $(m-1)$ of the set of columns a^j with $\langle y^s, a^j \rangle = 0$ (cf. Theorem 2.16 in [4]).

Because of the linear independence of the m columns a^j there are exactly m different variants for determining the vectors y^s, hence we have $k = m$ and $\text{lin}(a^1, \ldots, a^m) = \text{lin}(y^1, \ldots, y^m)$ (because of $\langle y^s, z \rangle = 0 \ \forall \ z \in K'$). Choosing $y^s = \sum_{i=1}^{m} \alpha_i^{(s)} a^i$ we can calculate the vectors y^s by solving the systems

$$\begin{aligned} \langle y^s, a^j \rangle &= \sum_{i=1}^{m} \alpha_i^{(s)} \langle a^i, a^j \rangle = 0\,, \quad \forall\ j \neq s\,, \\ \langle y^s, a^s \rangle &= \sum_{i=1}^{m} \alpha_i^{(s)} \langle a^i, a^s \rangle = r^{(s)} < 0\,. \end{aligned} \tag{15}$$

With $r^{(s)} = -1$ the system (15) is equivalent with $D \cdot \alpha^{(s)} = -e^s$. Because of the linear independence of the columns a^j the matrix D from (13) is regular, hence $\alpha^{(s)} = -D^{-1} \cdot e^s$ and $y^s = -A \cdot D^{-1} \cdot e^s$. Then with $y^s = z^s$, $s = 1, \ldots, m$, the assertion (14) is proved. ☐

The convexity or concavity of the implicit functions $y_k(x)$ from Theorem 2.1 can be proved also by direct examination of the adequate inequalities. But as Example 2.4 shows therefore some additional assumptions are necessary. By this we ensure that the projection of the $(n+m)$-dimensional set Y into the graph space of a fixed function y_k can be used as supporting halfspace of the epigraph of y_k.

Theorem 2.6. *Let the assumptions of Theorem* 1.1 *be fulfilled, and assume that the functions* G_i*,* $i = 1, \ldots, m$*, are convex on the considered neighborhood of* $(\overline{x}, \overline{y})$*. Furthermore,* *for arbitrary fixed $x \in U$ and $y = y(x)$ consider the convex polyhedral cones*

$$\begin{aligned} C(x) &= \operatorname{cone}(\nabla G_1(x, y(x)), \ldots, \nabla G_m(x, y(x))) \,, \\ C'(x) &= \{\ z \in \mathbb{R}^{n+m}\ :\ \langle z, \nabla G_i(x, y(x))\rangle = 0\,,\ i = 1 \ldots, m\ \}\,, \\ C^*(x) &= \{\ z \in \mathbb{R}^{n+m}\ :\ \langle z, \nabla G_i(x, y(x))\rangle \leq 0\,,\ i = 1 \ldots, m\ \} \\ &= C'(x) + \operatorname{cone}(z^1(x), \ldots, z^m(x)) \end{aligned} \tag{16}$$

(where the $z^i(x)$, $i = 1, \ldots, m$, are defined according to Lemma 2.5). Finally, let $k \in \{1, \ldots, m\}$ be a fixed index and $d^k = \begin{pmatrix} \nabla y_k(x) \\ -e^k \end{pmatrix} \in \mathbb{R}^{n+m}$ the gradient of the k-th component of the solution function $y(x)$ extended by an appropriate unit vector. Then we have:

(i) *If $\langle d^k, z^i(x)\rangle < 0\ \forall\ i = 1, \ldots, m$, then $y_k(x)$ is convex in a neighborhood of x.*
(ii) *If $\langle d^k, z^i(x)\rangle \leq 0\ \forall\ i = 1, \ldots, m$, then $y_k(x^*) \geq y_k(x) + \langle \nabla y_k(x), x^* - x\rangle$ holds for all x^* from a neighborhood of x.*
(iii) *If $\langle d^k, z^i(x)\rangle > 0\ \forall\ i = 1, \ldots, m$, then $y_k(x)$ is concave in a neighborhood of x.*
(iv) *If $\langle d^k, z^i(x)\rangle \geq 0\ \forall\ i = 1, \ldots, m$, then $y_k(x^*) \leq y_k(x) + \langle \nabla y_k(x), x^* - x\rangle$ holds for all x^* from a neighborhood of x.*

Proof. From (2) we get

$$\begin{pmatrix} \nabla_x G_i(x, y(x)) \\ \nabla_y G_i(x, y(x)) \end{pmatrix} = -\sum_{j=1}^{m} \frac{\partial}{\partial y_j} G_i(x, y(x)) \cdot \begin{pmatrix} \nabla y_j(x) \\ -e^j \end{pmatrix} \quad \forall\ x \in U\,, \tag{17}$$

which gives us together with the linear independence of the m vectors $\nabla_y G_i(x, y(x)) \in \mathbb{R}^m$ the relation $\binom{\nabla y_j(x)}{-e^j} \in \operatorname{lin}(\nabla G_1(x, y(x)), \ldots, \nabla G_m(x, y(x)))$, hence $d^k \in \operatorname{lin}(C(x))$ and $d^k \perp C'(x)$, respectively.

Using the sets Y and $R_i^+(x, y)$ from (3) and (7) in the proof of Theorem 2.1 we find for arbitrary $x^* \in U$ the inclusion

$$(x^*, y(x^*)) \in Y \subseteq \bigcap_{i=1}^{m} R_i^+(x, y(x)) = (x, y(x)) + C^*(x)\,. \tag{18}$$

Because of the structure (16) of the set $C^*(x)$ a vector $r \in C'(x)$ and constants $\alpha_i \in \mathbb{R}^1$, $i = 1, \ldots, m$, exist such that the representation

$$(x^*, y(x^*)) = (x, y(x)) + r + \sum_{i=1}^{m} \alpha_i z^i(x)\,, \quad \alpha_i \geq 0\,,\ i = 1, \ldots, m, \tag{19}$$

holds. From this we get the equation

$$\langle d^k, (x^* - x, y(x^*) - y(x))\rangle = \langle d^k, r + \sum_{i=1}^{m} \alpha_i z^i(x)\rangle = \sum_{i=1}^{m} \alpha_i \langle d^k, z^i(x)\rangle\,. \tag{20}$$

Assume that $\langle d^k, z^i(x)\rangle \le 0 \ \forall\ i = 1, \dots, m$ is true. Then from the nonnegativity of the factors α_i from (20) we get $\langle d^k, (x^* - x, y(x^*) - y(x))\rangle \le 0$, which leads us together with the definition of d^k to

$$\langle \nabla y_k(x), x^* - x\rangle - (y_k(x^*) - y_k(x)) \le 0 \tag{21}$$

and the assertion (ii) is proved. Analogously, in the case of reverse inequality sign we find the assertion (iv), too.

Assuming the strict inequalities $\langle d^k, z^i(x)\rangle < 0 \ \forall\ i = 1, \dots, m$, because of the continuity of all functions we find that $\langle d^k, z^i(x')\rangle \le 0 \ \forall\ i = 1, \dots, m$ holds for all x' from a neighborhood $U(x) \subset U$ of x. Hence we can apply assertion (ii) in all points x' and get

$$y_k(x^*) \ge y_k(x') + \langle \nabla y_k(x'), x^* - x'\rangle \quad \forall x^* \in U \quad \forall x' \in U(x)\,. \tag{22}$$

From this the convexity of y_k on $U(x)$ is proved. Analogously we find (iii) as corollary from (iv). □

For illustration of Theorem 2.6 we apply its assertion to the above example with $(n+2)$-dimensional set Y for which Theorem 2.1 was not applicable.

Example 2.7. Let the functions G_i are given as in Example 2.4 in a neighborhood of $\overline{x} = 0$ and $\overline{y} = (1, 0)$. Then for the sets defined by (16) we have

$$C(\overline{x}) = \text{cone}\left(\begin{pmatrix}0\\20\\-1\end{pmatrix}, \begin{pmatrix}0\\20\\2\end{pmatrix}\right), \quad C'(\overline{x}) = \text{lin}\left(\begin{pmatrix}1\\0\\0\end{pmatrix}\right),$$

$$C^*(\overline{x}) = \text{cone}(z^1(\overline{x}), z^2(\overline{x})) + C'(x)\,, \quad z^1(\overline{x}) = \frac{1}{30}\begin{pmatrix}0\\-1\\10\end{pmatrix}, \; z^2(\overline{x}) = \frac{1}{60}\begin{pmatrix}0\\-1\\-20\end{pmatrix}.$$

Hereby $z^1(\overline{x})$ and $z^2(\overline{x})$ are determined with $D = \begin{pmatrix}401 & 398\\398 & 404\end{pmatrix}$ and $D^{-1} = \frac{1}{3600}\begin{pmatrix}404 & -398\\-398 & 401\end{pmatrix}$ as in formula (14). Hence, from the explicit solution functions $y_1(x) = \sqrt{1-x^2}$ and $y_2(x) = \left(\arctan \frac{x}{\sqrt{1-x^2}}\right)^3$ or from the implicit representation (2) we get $y_i'(\overline{x}) = 0$, which leads to $d^1 = \begin{pmatrix}0\\-1\\0\end{pmatrix}$ and $d^2 = \begin{pmatrix}0\\0\\-1\end{pmatrix}$.

Then we have: Because of $\langle d^1, z^1(\overline{x})\rangle = \frac{1}{30} > 0$ and $\langle d^1, z^2(\overline{x})\rangle = \frac{1}{60} > 0$ the function $y_1(x)$ is concave in a neighborhood of $\overline{x} = 0$ (case (iii) in Theorem 2.6).

Since $\langle d^2, z^1(\overline{x})\rangle = -\frac{1}{3} < 0$ and $\langle d^2, z^2(\overline{x})\rangle = \frac{1}{3} > 0$, as expected for the (nonconvex and nonconcave) function $y_2(x)$ none of the assertions of Theorem 2.6 is applicable.

Remark. The properties (i) to (iv) in Theorem 2.6 are sufficient but in general not necessary for the convexity or concavity of $y_k(x)$, respectively. For instance, to obtain an affinlinear (convex and concave) function $y_k(x)$ the assumptions (ii) and (iv) has to be valid simultaneously. From this we have $\langle d^k, z^i(x)\rangle = 0\ \forall\ i = 1,\dots,m$ which gives us together with $d^k \perp C'(\overline{x})$ and $z^i(x) \perp C'(\overline{x})\ \forall\ i$ the contradiction $d^k = \mathbf{0}$. Hence, the assumptions of Theorem 2.6 are not suitable to prove the linearity of the implicit function $y_k(x)$.

3. Systems with additive convex functions

In section 2 we have seen that the convexity or concavity of implicit functions can only be obtained with the help of additional assumptions. In the following we consider the special case when two groups of variables are additively separated in the equations. The aim of the following section is to get statements about the gradients of the involved functions which can be useful in the investigation of subdifferentials of marginal functions (for decomposition approaches).

Theorem 3.1. *Let the functions* $g_i(x) : \mathbb{R}^1 \to \mathbb{R}^1$ *and* $h_i(y) : \mathbb{R}^m \to \mathbb{R}^1$, $i = 1, 2$, *be convex and continuously differentiable in a neighborhood of the points* $\overline{x} \in \mathbb{R}^1$ *and* $\overline{y} \in \mathbb{R}^m$, *respectively. Furthermore, suppose that for the system*

$$g_i(x) + h_i(y) = 0, \quad i = 1, 2, \tag{23}$$

a continuous function $y(x)$ *exists with*

$$g_i(x) + h_i(y(x)) = 0, \quad i = 1, 2, \quad \forall\ x \in U(\overline{x}) \tag{24}$$

which is differentiable on $U(\overline{x}) \setminus \{\overline{x}\}$. *Finally let at least one of the following conditions* (i) *to* (iii) *be fulfilled:*

(i) $\exists\ M > 0,\ \delta > 0\ :\ |x - \overline{x}| < \delta \implies ||\nabla y(x)|| < M \quad \forall\ x \neq \overline{x}$;
(ii) $\nabla h_1(\overline{y}) = \mathbf{0}$;
(iii) $m = 1$.

Then the implication

$$G := \nabla h_1(\overline{y}) = \nabla h_2(\overline{y}) \qquad \implies \qquad g_1'(\overline{x}) = g_2'(\overline{x}) \tag{25}$$

holds.

Proof. Differentiating the equations (24) w.r.t. x we get

$$\nabla h_i(y(x)) \cdot \nabla y(x) = -g_i'(x) \qquad \forall\ x \in U(\overline{x}) \setminus \{\overline{x}\}\,. \tag{26}$$

Since $-g_i'(x)$ is continuous in a neighborhood of $\overline{x}$, the expression $\lim\limits_{x\to\overline{x}} \nabla h_i(y(x)) \cdot \nabla y(x)$ exists independently from the existence of the limit $= \lim\limits_{x\to\overline{x}} \nabla y(x)$. If assumption (i) is fulfilled, then there is at least one finite accumulation point $V \in \mathbb{R}^m$ of the gradients $\nabla y(x)$, such that the limit in relation (26) is realized:

$$\lim_{x\to\overline{x}} \nabla h_i(y(x)) \cdot \nabla y(x) = G \cdot V = -g_i'(\overline{x})\,, \quad i = 1, 2\,. \tag{27}$$

Hereby the vector V depends only from the series $\nabla y(x)$ and not from the index i. Hence from (27) we get directly the assertion (25).

(ii) In the case $G = \mathbf{0}$ we have $h_i(y) \geq h_i(\overline{y})\ \forall\ y \in U(\overline{y})$, hence

$$g_i(x) = -h_i(y(x)) \leq -h_i(\overline{y}) = g_i(\overline{x})\,, \quad i = 1,2\,, \quad \forall\ x \in U(\overline{x})\,. \tag{28}$$

Because of the convexity of $g_i(x)$ from this we get $g_i(x) = g_i(\overline{x})\ \forall\ x \in U(\overline{x})$, and this gives us with $g_i'(\overline{x}) = 0$, $i = 1, 2$, the validness of the assertion (25).

(iii) For $m = 1$ only the case $G \neq \mathbf{0}$ is to be proved. Then for each of the two equations $g_i(x) + h_i(y) = 0$, $i = 1, 2$, an implicit function $y = y(x)$ exists in $U(\overline{x})$, and

$$y'(\overline{x}) = -\frac{g_i'(\overline{x})}{h_i'(\overline{y})} = -\frac{g_i'(\overline{x})}{G}\,, \quad i = 1, 2. \tag{29}$$

Because of the uniqueness of the solution function $y(x)$ from (29) we get $g_i'(\overline{x}) = -G \cdot y'(\overline{x}) = \text{const.}$, and the assertion of the Theorem is proved. □

Remark. The main result of Theorem 3.1 is the validness of the property (25) even in the case when $y(x)$ is not differentiable at $\overline{x}$. Hereby the boundedness of the gradients $\nabla y(x)$ in a neighborhood of $\overline{x}$ is essential for the proof (at least in the case $G \neq \mathbf{0}$ and $m > 1$). If none of the conditions (i) to (iii) is fulfilled then the assertion (25) can be true or false:

Example 3.2. We consider (23) with $m = 2$ and

$$\begin{aligned} g_1(x) &= -x - 1\,, & h_1(y) &= y_1^2 + y_2^2\,, \\ g_2(x) &= x^2 - x - 1\,, & h_2(y) &= y_1^2 + 2y_2 - 1\,. \end{aligned}$$

Then for $\overline{x} = 0$ and $\overline{y} = (0, 1)$ we have the gradients $G = \nabla h_1(\overline{y}) = \nabla h_2(\overline{y}) = (0, 2)$ and the solution function

$$y_1(x) = \begin{cases} \sqrt{-x - x^2}, & x \leq 0, \\ -\sqrt{3x - x^2}, & 0 < x, \end{cases} \qquad y_2(x) = 1 - |x|$$

(cf. Figure 3.1). Here both components of $y(x)$ are continuously differentiable for each parameter $x \in (-1, 1) \setminus \{0\}$. For $x = 0$ we have

$$\lim_{x \uparrow 0} y_1'(x) = \lim_{x \downarrow 0} y_1'(x) = -\infty\,, \quad \lim_{x \uparrow 0} y_2'(x) = 1 \neq \lim_{x \downarrow 0} y_2'(x) = -1\,.$$

Despite of this nondifferentiability and unboundedness of $\nabla y(x)$ at $x = 0$ the assertion (25) from Theorem 3.1 is true (since $g_1'(0) = g_2'(0) = -1$).

Example 3.3. Let in (23) the relations $m = 2$ and

$$\begin{aligned} g_1(x) &= -x - 20\,, & h_1(y) &= \sqrt{y_1^2 + y_2^2} \cdot (20 + (\arctan \tfrac{y_1}{y_2})^3)\,, \\ g_2(y) &= 2x - 20\,, & h_2(x) &= \sqrt{y_1^2 + y_2^2} \cdot (20 - 2(\arctan \tfrac{y_1}{y_2})^3) \end{aligned}$$

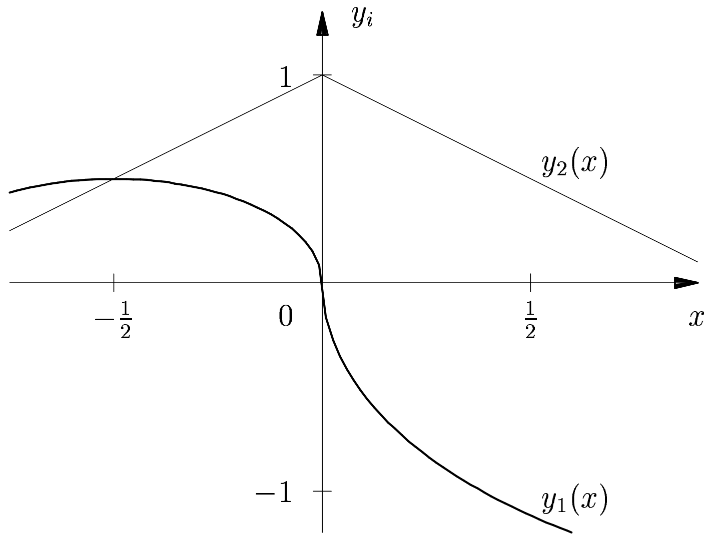

Figure 3.1: Solution functions for Example 3.2

hold. Then the functions $g_i(x)$ and $h_i(y)$, $i = 1, 2$, are convex and continuously differentiable in a neighborhood of $\overline{x} = 0$ and $\overline{y} = (0, 1)$. Because of

$$\nabla h_1(y) = \left(\frac{20y_1 + y_1(\arctan \frac{y_1}{y_2})^3 + 3y_2(\arctan \frac{y_1}{y_2})^2}{\sqrt{y_1^2 + y_2^2}}, \frac{20y_2 + y_2(\arctan \frac{y_1}{y_2})^3 - 3y_1(\arctan \frac{y_1}{y_2})^2}{\sqrt{y_1^2 + y_2^2}} \right),$$

$$\nabla h_2(y) = \left(\frac{20y_1 - 2y_1(\arctan \frac{y_1}{y_2})^3 - 6y_2(\arctan \frac{y_1}{y_2})^2}{\sqrt{y_1^2 + y_2^2}}, \frac{20y_2 - 2y_2(\arctan \frac{y_1}{y_2})^3 + 6y_1(\arctan \frac{y_1}{y_2})^2}{\sqrt{y_1^2 + y_2^2}} \right),$$

we have $G = \nabla h_1(\overline{y}) = \nabla h_2(\overline{y}) = (0, 20)$, and the system (23) has the solution

$$y_1(x) = \sin \sqrt[3]{x}\,, \quad y_2(x) = \cos \sqrt[3]{x}\,, \quad \forall\; x \in U(\overline{x})$$

(cf. Figure 3.2). Hence, the unboundedness of the gradients $\nabla y(x)$

$$\lim_{x \to \overline{x}} ||\nabla y(x)|| = \lim_{x \to \overline{x}} \frac{1}{3\sqrt[3]{x^2}} = +\infty$$

together with $G \neq \mathbf{0}$ and $m \neq 1$ contradicts the assumptions of Theorem 3.1. In this case the assertion (25) is false (since $g_1'(\overline{x}) = -1 \neq g_2'(\overline{x}) = 2$).

Now we can generalize the result of Theorem 3.1 to the multidimensional case ($x \in \mathbb{R}^n$, $n \geq 1$).

Corollary 3.4. *Let the functions $g_i(x) : \mathbb{R}^n \to \mathbb{R}^1$ and $h_i(y) : \mathbb{R}^m \to \mathbb{R}^1$, $i = 1, 2$, be convex and continuously differentiable in a neighborhood $U(\overline{x}) \subseteq \mathbb{R}^n$ and $U(\overline{y}) \subseteq \mathbb{R}^m$, respectively. Let $T \subseteq \mathbb{R}^n$ be an open subset of $U(\overline{x})$ such that $\overline{x}$ is an accumulation point of T, and let a differentiable function $y(x)$ exist for which*

$$g_i(x) + h_i(y(x)) = 0, \quad i = 1, 2, \quad \forall\; x \in T\,. \tag{30}$$

Finally let at least one of the following conditions (i') *to* (iii') *be fulfilled:*

(i') $\exists\; M > 0\,, \delta > 0\; :\; ||x - \overline{x}|| < \delta \implies ||\nabla y(x)|| < M\; \forall\; x \in T$;

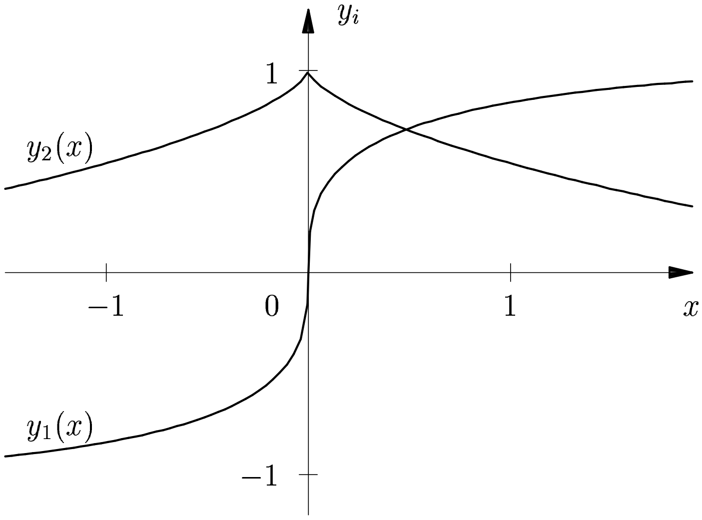

Figure 3.2: Solution functions for Example 3.3

(ii') $\nabla h_1(\overline{y}) = \mathbf{0}$ *and* $y(x)$ *is continuous for each* $x \in U(\overline{x})$;
(iii') $m = 1$.

Then the implication

$$G := \nabla h_1(\overline{y}) = \nabla h_2(\overline{y}) \qquad \Longrightarrow \qquad \nabla g_1(\overline{x}) = \nabla g_2(\overline{x}) \tag{31}$$

holds.

Proof. In the case (i') we can use the relation (26) as in the proof of Theorem 3.1 for each $x \in T$. Then at least one finite accumulation point $V \in \mathbb{R}^{m \cdot n}$ of the matrix $\nabla y(x)$ exists such that from formula (27) the assertion (31) follows.

In the case (iii') the vector equation (29) leads us directly to the assertion.

For the remaining case (ii') we need the existence of a (nondifferentiable but continuous) function $y(x)$ in the whole neighborhood $U(\overline{x})$. Only under this assumption we can use the estimation (28) from the proof of Theorem 3.1 obtaining $g_i(x) = g_i(\overline{x})\ \forall\ x \in U(\overline{x})$ and hence $\nabla g_i(\overline{x}) = \mathbf{0}$, $i = 1, 2$. □

Remark. In Corollary 3.4 the solution function $y(x)$ is not supposed to be differentiable at the accumulation point $\overline{x}$. Nevertheless the property (31) of the gradients can be ensured even at such points if $x(y)$ is differentiable at least in some open subset of the neighborhood of $\overline{x}$. This result can be applied to the differentiable continuation of some marginal functions at the boundary of stability sets (cf. [11]).

The statements of Section 3 for the system (23) cannot simply transfered to the general convex equations (1) from Section 2 since the additive separability of the involved functions is essential for the proofs of Theorem 3.1 and Corollary 3.4 in the cases $G = \mathbf{0}$ and $m > 1$.

References

[1] Z. Balogh, G. Kassay: On convexity of implicit functions, Prepr., "Babes-Bolyai" Univ., Fac. Math. Phys., Res. Semin. 1990, No.7 (1990) 77–82.

[2] B. Bank, J. Guddat, D. Klatte, B. Kummer, K. Tammer: Nonlinear Parametric Optimization, Akademie-Verlag, Berlin 1982.

[3] G. M. Fichtenholz: Differential- und Integralrechnung, 3 Bände, Deutscher Verlag der Wissenschaften, Berlin 1977, 1978.

[4] D. Gale: The Theory of Linear Economic Models, McGraw-Hill, New York 1960.

[5] L. Grygarova: The implicit function theorem for a convex function, Seminarber., Humboldt-Univ. Berlin, Sekt. Math. 94 (1988) 45–58.

[6] F. Nozicka, J. Guddat, H. Hollatz, B. Bank: Theorie der linearen parametrischen Optimierung, Akademie-Verlag, Berlin, 1974.

[7] F. Nozicka, L. Grygarova, K. Lommatzsch: Geometrie konvexer Mengen und konvexe Analysis, Akademie-Verlag, Berlin, 1988.

[8] R. T. Rockafellar: Convex Analysis, Princeton University Press, Princeton, NJ, 1970.

[9] R. S. Strichartz: The Way of Analysis, Jones and Bartlett, Boston, London 1995.

[10] U. Würker: Properties of Some Convex Marginal Functions Without Constant Rank Regularity, In: P. Kall, H.-J. Lüthi (eds.): Operations Research Proceedings 1998, Selected Papers of the International Conference on Operations Research, Zürich, August 31 – September 3, 1998, 73–82.

[11] U. Würker: Konstruktion und Anwendung konvexer differenzierbarer Zwischenfunktionen in der Dekomposition, Habilitationsschrift, TU Chemnitz, Faculty of Mathematics, 2000.