Journal of Convex Analysis
Volume 28 (2021), No. 4, 1119–1136

Optimization of Fitzpatrick Functions and a Numerical Minimization Algorithm

Ivan Raykov
University of Arkansas, Pine Bluff, U.S.A.
raykovi@uapb.edu

M. Zuhair Nashed
University of Central Florida, Orlando, U.S.A.

Received: July 27, 2020
Accepted: August 15, 2020

We consider optimization problems for a class of convex functions on $H \times H$ introduced by Simon Fitzpatrick, where H is a real Hilbert space. We show that the minimization problem of Fitzpatrick functions can be transformed from solving of the correspondent differential inclusions (d.i) on $H \times H$, to solving simplified d. i. on H. By using the idea of optimization of Fitzpatrick functions we introduce a numerical algorithm for solving convex smooth optimization problems by reducing the number of the independent variables. We present a comparative study with numerical examples. Finally, we show that Fitzpatrick functions are closely related to Lyapunov functions.

Keywords: Maximal monotone operator, lower semicontinuous convex maps, differential inclusions, optimization problems.

2010 Mathematics Subject Classification: 46N10, 47N10.

1. Introduction

Simon Fitzpatrick developed in [10] the representation of monotone operators on a normed space E in terms of the subdifferentials of convex functions on $E \times E^*$, which has led to definition of the Fitzpatrick functions (see the Definition 3.1 in Section 3). Seminal ideas are often triggered by amazing insights. The Fitzpatrick functions represent an example of such idea. During the last two decades several mathematicians have published deep results considering the Fitzpatrick functions and their properties. The functions have played pivotal roles in advancing our understanding of maximal monotone operators and convex functions in three directions. Firstly, some results and connections have been found between Fitzpatrick functions and other results in the theory of monotone operators and convex analysis. Secondly, several deep results on maximal monotone operators have recently found simpler proofs using the Fitzpatrick functions. Thirdly, as demonstrated in this paper, the Fitzpatrick functions have interesting properties in optimization problems. Penot [13] introduced new representations for maximal monotone operators and related them to previous representations given by Kraus, Fitzpatrick, Martinez-Legas, and

We thank Professor Bruce Lundberg from CSU-Pueblo and Professor Hans D. Mittelmann from ASU for the cooperation and support.

ISSN 0944-6532 / $ 2.50

Théra. He showed the usefulness of such representations for the study of compositions and sums of maximal monotone operators, highlighting the relevance of convex analysis for the study of monotonicity. Simons and Zalinescu [15] showed how the versions of the Fenchel duality theorem due to Rockafellar and Attouch-Brezis can be applied to the Fitzpatrick function determined by a maximal monotone multifunction to obtain a number of results on maximal monotonicity, including sufficient conditions for the sum of maximal monotone multifunctions on a reflexive Banach space to be a maximal monotone, unifying several results of the Attouch-Brezis type that have been obtained in recent years. Bartz et al. [2] studied a sequence of Fitzpatrick functions associated with a monotone operator. The first term of their sequence concides with the original Fitzparick function, and the other terms turn out to be useful for the identification and characterization of cyclic monotonicity properties. Borwein and Dutta [6] studied maximal monotone inclusions from the properties of gap functions. They proposed a natural gap function for an arbitrary maximal monotone inclusion, and showed how naturally this gap function arises from the Fitzpatrick function. For the approach using cones of tangent, see [5].

In the present paper, we consider optimization problems for the Fitzpatrick functions, utilizing specific properties such as maximal monotonicity, convexity and lower-semicontinuity. We generalize and unify preliminary results annouced mainly without proofs at two conferences. The paper also contains new material connecting Lyapunov functions with Fitzpatrick functions. Finally, we present new numerical results, which introduce simplified algorithms for this type of problems. The optimization of Fitzpatrick functions is motivated by considerations which can be best understood after presentation of the main results, and consequently they are deferred to concluding comments at the end of the paper. Concluding remarks present the main topics of the paper with possible future developments and applications.

Our interest in optimizing Fitzpatrick functions is motivated in general by the following considerations:

1. Optimizing Fitzpatrick functions can be simplified from optimization problems on $H \times H$ to optimization problems on H. It would be of interest to explore classes of convex functions on $H \times H$ which have a similar simplification property.
2. We introduce an iterative procedure for solving problems of this type. Because of the convexity the procedure can be simplified. We present a comparative study.
3. We present two examples of Fitzpatrick functions which are also Lyapunov functions. This may lead to further connections with Lyapunov functions.
4. Fitzpatrick functions are convex on $H \times H$ and involve implicit functions.([12]

2. Preliminaries

Let H be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$, and let the norm of $H \times H$ be $\|(x, u)\| := \sqrt{\|x\|^2 + \|u\|^2}$, $(x, u) \in H \times H$. We assume that the convex functions are proper and recall that a mapping T of H into subsets of H is a *monotone operator* provided for each $u \in Tx$ and $v \in Ty$ the following inequality holds:

$\langle u-v, x-y\rangle \geq 0$. If T is monotone and the *graph* of T is not properly contained in the graph of a monotone operator then T is said to be *maximal monotone*. We remind also the next: A vector $g \in H$ is a subgradient of a function $F : H \to R$ if for all $z \in DomF$ $F(z) \geq F(x) + \langle z-x, g\rangle$. The set of all subgradients of F at x is called the subdifferential of F at x and denoted by $\partial F(x)$. The subdifferential $\partial F(x)$ is always a closed convex set, even if F is not convex. This follows from the fact that it is an intersection of infinite set of half spaces:

$$\partial F(x) := \bigcap_{x\in DomF}\{g | F(z) \geq F(x) + \langle z-x, g\rangle\}.$$

It is well known that: if F is a proper convex lower semicntinuous function, then $\partial F : H \Longrightarrow H$ is a maximalmonotone operator.

3. Minimization Problems for a Class Convex Functions on $H\times H$.

In this section we consider minimization problems for a class of convex functions on $H\times H$. We define and study a monotone operator on H using a convex function on $H\times H$, presented in [10].

Definition 3.1. For each convex function f on $H\times H$ we define the operator $T_f x := \{u \in H \mid (u, x) \in \partial f(x, u)\}$ for each $x\in H$.

Fitzpatrick proved that T_f is a monotone operator (see [10] Proposition 2.2).

In this section we will consider next optimization problem

$$\inf_{(x,u)\in H\times H} f(x, u) = f^{inf} \tag{1}$$

supposing its solution exists, i.e. we suppose that there exists at least one point $(x^*, u^*) \in H \times H$ for which the solution of (1) exists.

Theorem 3.2. *Let f and T_f be as in Definition* 3.1 *and let function $f(x, u)$ is defined, proper, convex, and (l.s.c.) and T_f is maximal monotone. Then (d. i.)*

$$\dot{x} \in -T_f(x(t)), \quad x(0) = x_0 \tag{2}$$

and
$$(\dot{x}, \dot{u}) \in -\partial f(x(t), u(t)), (x(0), u(0)) = (x_0, u_0) \tag{3}$$

have unique solutions and $x(\cdot)$ and $(x(\cdot), u(\cdot))$ respectively and they coincide when $(x, u) \in G(T_f)$

Proof. Because the function f is real valued, proper, l.s.c., and convex on $H\times H$, then ∂f will be maximal monotone (see [1] Ch. 6, S. 7, Prop. 6). Since T_f is maximal monotone and ∂f is maximal monotone there exist unique solutions of d. i. (2) and (3), which follows from [1] (Ch. 6, S. 8, Theorem 1). We use the uniqueness of the solutions of (2) and (3) and the definitions of f and T_f. It follows then that the solutions $x(.)$ of d. i. (2) and $(x(.), u(.))$ of d. i. (3) coincide for every $t \in [0, \infty)$ when $(x(t), u(t)) \in G(T_f)$ where T_f satisfies Definition 3.1. □

In the rest of the section we consider a class of d. i. involving the subdifferential of the Fitzpatrick functions in which every solution of the d. i. can be extended up to the solution set.

Let $X_f := \{(x,u) \in H \times H | f(x,u) = f^{inf}\}$ denote the solution set of the optimization problem (1) and $X_{f_\varepsilon} := \{(x,u) \in H \times H | \min_{(y,v)} \|(x,u) - (y,v)\| < \varepsilon$ an ε-neighborhood of X_f.

Lemma 3.3. *Consider d. i.*

$$(\dot{x}, \dot{u}) \in (\Phi(x,u)), (x(0), u(0)) = (x_0, u_0),\ t \geq 0, \tag{4}$$

where $\Phi : H \times H \Longrightarrow R$ is maximal monotone multi-valued map and therefore upper semicontinuous (u.s.c.), and for each $(x,u) \in H \times H$, $\Phi(x,u)$ is bounded, closed, and convex set, $\|\Phi(x,u)\| = \sup_{(w,z)\in\Phi(x,u)} \|(w,z)\| \leq 1 + \|(x,u)\|$.
Let $\theta(x,u) : H \times H \Longrightarrow R_+$ be strictly positive and bounded function

$$0 < \delta \leq \theta(x,u) \leq M < \infty$$

for which $\theta(x(t), u(t))$ is measurable for every absolutely continuous function

$$(x(t), u(t)) | R \Longrightarrow H \times H.$$

Denote $\tau(t) = \int_0^t \theta(x(s), u(s)) ds$ and $t(\tau)$ its inverse function. We consider also the following d. i.:

$$(\dot{z}, \dot{w}) \in \theta(z,w)(\Phi(z,w)), (z(0), w(0)) = (x_0, u_0),\ t \geq 0 \tag{5}$$

If $(x(t), u(t))$ is a solution of d. i. (4) then the function $(z(t), w(t)) = (x(\tau(t)), u(\tau(t))$ almost everywhere satisfies d. i. (5) i.e., the set of trajectories of d. i. (4) and (5) coincide as curves in the phase space $H \times H$.

Proof. Note that, according to Deimling [8], there exists a solution of inclusion (4) if Φ maps bounded sets into bounded sets and satisfies some additional conditions of compactness. $\tau(t)$ is absolutely continuous function for which, under $0 < \delta \leq \theta(x,u)$, there exists an inverse function $t(\tau)$. Let $(x(t), u(t))$ be a solution of (4). The function $\tau(t) = \int_0^t \theta(x(s), u(s)) ds$ is absolutely continuous, and since we have $0 < \delta \leq \theta(x,u) \leq M < \infty$, it satisfies the Lipschitz condition with a constant M. Almost everywhere in t, for its derivative we have $\dot{\tau}(t) = \theta(x(t), u(t)) \geq \delta > 0$. Thus, the function $\tau(t)$ has an inverse function $t(\tau)$ for which

$$0 < \frac{1}{M} \leq \dot{t}(\tau) = \frac{1}{\theta(x(t), u(t))} \leq \frac{1}{\delta} < \infty.$$

Obviously, $t(\tau)$ is an absolutely continuous function. Now, almost everywhere in t we can write

$$\frac{d(x(\tau(t), u(\tau(t)))}{dt} \in \theta(x(\tau(t)), u(\tau(t)))\Phi(x(\tau(t), u(\tau(t))),$$

$$\frac{d(z(t(\tau)), w(t(\tau)))}{d\tau} \in \frac{1}{\theta(z(t(\tau)), w(t(\tau)))}\theta(z(t(\tau)), w(t(\tau)))\Phi(z(t(\tau)), w(t(\tau)))$$

$$= \Phi(z(t(\tau)), w(t(\tau)))$$

Thus, the positive multiplier $\theta(z,w)$ only transforms the time. □

We prove in the next theorem that if the optimization problem (1) has a solution, and therefore the correspondent d. i. (3) has a solution, then the solution of a modification of d. i. (3), satisfying Lemma 3.3, can be extended to the solution set X_f of (3), respectively of (1), in a finite time. Recall that the metric projection of the point x in a Hilbert space H on a subset S is defined with $Pr_S x$, i.e.

$$Pr_S x := \{y \in S| \quad \|x - y\| = \min_{z \in S} \|x - z\|\}.$$

The metric projection on a closed and convex subset of a Hilbert space exists and is unique. We denote here by $m(\mathbf{0}, \partial f(x, u))$ the metric projection of the origin onto $\partial f(x, u)$, that is,

$$m(\mathbf{0}, \partial f(x, u)) = \{(v, y) \in H \times H| \|(0, 0) - (y, v)\| = \min_{(w,z) \in \partial f(x,u)} \|(0, 0) - (w, z)\|\}.$$

We assume that $\mathbf{0}$ is not an element of the subdifferential. We use also the notation

$$D(\mathbf{0}, \partial f(x, u)) = \sup_{(w,z) \in \partial f(x,u)} \|(0, 0) - (w, z)\|.$$

In what follows when x and u depend on a parameter t(time), we will abbreviate the subdifferential in the metric projection and in the definition of D by $\partial f(t)$.

Theorem 3.4. *Let us consider again the d. i.* (3) *where $\partial f(x(.), u(.))$ is a maximal monotone operator. Suppose the map f is proper, convex and l.s.c., and maps (x, u) to a bounded, closed and convex set. For $\varepsilon > 0$ and $(x_0, u_0) \notin X_{f_\varepsilon}$ let $\|m(\mathbf{0}, \partial f(x, u))\| \geq \beta > 0$, $(x, u) \notin X_{f_\varepsilon}$ and let*

$$X_f := \{(x, u) \in H \times H | f(x, u) = f^{inf}\} = \{(z, w) \in H \times H | (0, 0) \in \partial f(z, w)\}.$$

If d. i. (3) *has a solution and the solutions set is bounded for $t \geq 0$, then every solution of*

$$(\dot{x}, \dot{u}) \in -\frac{\partial f(x(t), u(t))}{\|m(\mathbf{0}, \partial f(t))\|^2}, (x(0), u(0)) = (x_0, u_0) \tag{6}$$

can be extended up to the set X_f in a finite time $T \leq (f(x_0, u_0) - \underline{f})$ where $\underline{f}$ is a lower bound of $f(x, u)$.

Proof. Function $\|m(\mathbf{0}, \partial f(x, u))\|^2 \geq \beta^2 > 0$ is strictly positive and l.s.c.. Thus, for every absolutely continuous $(x(t), u(t))$ the function $\|m(\mathbf{0}, \partial f(x(t), u(t)))\|$ is l.s.c.. From Lemma 3.3 the sets of trajectories of d. i. (3) and (6) coincide as curves in $H \times H$. Let $(x(t), u(t))$ be a solution of d. i. (6) and $\underline{f}$ is a lower bound of $f(x, u)$. The function $(f(x, u) - \underline{f})$ is real valued, strictly positive, proper, l.s.c. and convex. Suppose $f(x(t), u(t))$ is differentiable a.e. One writes

$$\begin{aligned}
&(f(x(0), u(0)) - \underline{f}) - (f(x(t), u(t)) - \underline{f}) = f(x(0), u(0)) - f(x(t), u(t)) \\
&= -\int_0^t (f(x(q), u(q)))'_q dq \in -\int_0^t \big\langle \partial f(x(q), u(q)), (\dot{u}(q), \dot{x}(q)) \big\rangle dq \\
&\subset \int_0^t \Big\langle \partial f(x(q), u(q)), \frac{\partial f(x(q), u(q))}{\|m(\mathbf{0}, \partial f(q))\|^2} \Big\rangle dq \subset t \Big[1, \frac{(D(\mathbf{0}, \partial f(t)))^2}{\|m(\mathbf{0}, \partial f(t))\|^2}\Big].
\end{aligned}$$

For every t we have $0 < f(x(t), u(t)) - \underline{f} \leq f(x(0), u(0)) - \underline{f} - t$. Thus in time $T \leq (f(x(0), u(0)) - \underline{f})$, where $\underline{f}$ is a lower bound of $f(x, u)$, the solution $(x(t), u(t))$ of (6), for $(x(0), u(0)) = (x_0, u_0)$, has to reach the set X_ε in time $T \leq (f(x(0), u(0)) - \underline{f}$, because the solutions set of inclusion (3) is bounded. Here $D(\mathbf{0}, \partial f(t)) = \sup_{(w,z) \in \partial f(x(t),u(t))} \|(0, 0) - (w, z)\|$. □

Let us consider the following theorem.

Theorem 3.5. *We consider now the optimization problem* (1).
We suppose that f is proper, convex and l.s.c., ∂f is maximal monotone.
Let the solution set X_f be not empty and $(x_0, u_0) \notin X_{f_\varepsilon}$.
Let $\|\partial f(x,u)\| \geq \gamma > 0$, $(x,u) \notin X_{f_\varepsilon}$. Then:

1. *The solution set X_f of* (1) *coincides with the set for which origin belongs to the subdifferential $\partial f(x,u)$.*
2. *For every initial position $(x_0, u_0) \in H \times H$ and every solutions $(x(t), u(t))$ of d. i.*

$$(\dot{x}, \dot{u}) \in -\frac{\partial f(x(t), u(t))}{\|m(\mathbf{0}, \partial f(t))\|^2}, (x(0), u(0)) = (x_0, u_0), \tag{6}$$

where $m(\mathbf{0}, \partial f(t))$ the metric projection of the origin onto the multivaued function $\partial f(x,u)$, that is,

$$m(\mathbf{0}, \partial f(t)) = \{(v,y) \in E^* \times E \| \|(0,0) - (y,v)\| = \min_{(w,z) \in \partial f(x(t),u(t))} \|(0,0) - (w,z)\|\},$$

there exists a finite moment $T \leq f(x_0, u_0) - \underline{f}$, where $\underline{f}$ is a lower bound of $f(x,u)$, for which $(x(T), u(T))$ solves the optimization problem (1).

Proof. By Lemma 3.3 and Theorem 3.4, d. i. (6) and (3) have unique solutions which coincide as curves in $H \times H$ up to the moment T when $(0,0) \in \partial f(x(T), u(T))$. We are going to show that the function $f(x,u)$ is the Lyapunov function for d. i. (6). This method is based on the Lyapunov function theory. Under the Chain rule (see [7]) for the generalized gradientl of $f(x(t), u(t))$ almost everywhere in t we obtain

$$\partial_t f(x(t), u(t)) = \{\langle (\eta, \xi), (\dot{x}(t), \dot{u}(t)) \rangle | (\eta, \xi) \in \partial f(y, v), (y, v) = (x(t), u(t)).$$

D.i. (6) is the Filipov extension of the next differential equation (d.e.) with possibly discontinuous right-hand side, (see [1] and [9]):

$$(\dot{x}, \dot{u}) = -\frac{m(\mathbf{0}, \partial f(t))}{\|m(\mathbf{0}, \partial f(t))\|^2}(x(0), u(0)) = (x_0, u_0).$$

For every $(x(t), u(t))$ which satisfies the above d.e. and by the properties of the metric projection $m(\mathbf{0}, \partial f(t))$ we have

$$\max_{(\eta,\xi) \in \partial f(x(t),u(t))} \langle (\eta, \xi), (\dot{x}(t), \dot{u}(t)) \rangle = \max_{(\eta,\xi) \in \partial f(x(t),u(t))} \left\langle (\eta, \xi), -\frac{m(\mathbf{0}, \partial f(t))}{\|m(\mathbf{0}, \partial f(t))\|^2} \right\rangle = -1.$$

Thus, $f(x,u)$ is the Lyapunov function for d. i. (6). We obtain that $f(x(t), u(t))$ is a monotone decreasing function in t at the solutions of d. i. (6).

Therefore according to [11], Lemma 3.3, and Theorem 3.4 we can find a solution of the optimization problem (1) by solving d. i. (6). □

In the following theorem we denote with $m(\mathbf{0}, \Gamma_{T_f(t)})$ the metric projection of the origin onto the range of the multivalued map $T_f(x(t))$.

Theorem 3.6. *Let f and $T_f x$ satisfy Definition* 3.1 *and Theorem* 3.2. *Consider d. i.* (2) *where the right-hand side is maximal monotone and maps x to a bounded, closed and convex set, $x_0 \notin X_{f_\varepsilon}$ and $\varepsilon > 0$ is arbitrary chosen.*
Let $\|m(\mathbf{0}, T_f(x))\| \geq \eta > 0$, $x \notin X_{f_\varepsilon}$ and

$$X_f := \{(x,u) \in H \times H | f(x,u) = f^{inf}\} = \{u \in T_f(x),\ (u,x) \in \partial f(x,u) \mid 0 \in T_f(x)\}.$$

If d. i. (2) *admits a solution and the set of these solutions is bounded for $t \geq 0$, then every solution of*

$$\dot{x} \in -\frac{T_f(x(t))}{\|m(\mathbf{0}, \Gamma_{T_f(t)})\|^2}, \quad x(0) = x_0,$$

can be extended up to the set X_f in a finite time $T \leq f(x_0, u_0) - \underline{f}$, where $u_0 \in T_f(x_0)$.

We omit the proof, because it is similar to the proof of Theorem 3.5.

4. Minimization problem for a Fitzpatrick function

Let us consider now next Fitzpatrick function.

Definition 4.1. (see [10] 3.1. Definition) Let $T : H \to H$ be a monotone operator.

$$L_T : H \times H \Longrightarrow]-\infty, +\infty] := (x,u) \mapsto \sup_{y \in H} \{\langle u, y \rangle + \langle Ty, x - y \rangle\}.$$

The function $L_T(x,u)$ is called the *Fitzpatrick function* representing a monotone operator T.

It follows from Definition 3.1 that if $D(T) \neq \emptyset$, then the function L_T is proper, lower semicontinuous and convex on $H \times H$ and ∂L_T is maximal monotone (see [14] and [4]).

According to Theorem 3.10 in [10] if T is a maximal monotone operator on H, then L_T is the minimal convex function f on $H \times H$ such that $f(x,u) \geq \langle u, x \rangle$ for all x and u, and $f(y,v) = \langle v, y \rangle$ for all $(y,v) \in G(T)$.

We will consider two cases.

Case 1: The following result will be used in what follows.

Theorem 4.2. (see [10] 3.4. Theorem) *If T is a monotone operator from Definition 4.1 on H and $(x,u) \in G(T)$ then $L_T(x,u) = \langle u, x \rangle$ and $(u,x) \in \partial L_T(x,u)$.*

Proof. We can apply for this case directly Theorem 3.2 from the previous section because when $(x,u) \in G(T)$ then $(u,x) \in \partial L_T(x,u)$. □

Case 2: The following result will be used in what follows.

Lemma 4.3. (see [10] 3.3. Lemma) *If T is the monotone operator from Definition* 4.1 *on H and $(y,v) \in G(T)$ and for some $(x,u) \in H \times H$ we have $L_T(x,u) = \langle v, x - y \rangle + \langle u, y \rangle$ then $(y,v) \in \partial L_T(x,u)$.*

Using Lemma 4.3 we prove the following theorem:

Theorem 4.4. *Let T and L_T be as in Definition 4.1, $D(T) \neq \emptyset$ and T is maximal monotone. And let there exist an open set $S \subset H \times H$ such that for every $(x,u) \in S$ the conditions of Lemma 4.3 will be satisfied and therefore if $(y,v) \in G(T)$, then $(y,v) \in \partial L_T(x,u)$. Suppose $(x,u) \in S$. Then d. i.*

$$\dot{y} \in -T(y(t)), \quad y(0) = y_0 \tag{7}$$

has a unique solution $y(\cdot)$ defined on $[0,\infty)$ and d. i.

$$(\dot{x},\dot{u}) \in -\partial L_T(x(t),u(t)), (x(0),u(0)) = (x_0,u_0) = (y(0),v(0)) = (y_0,v_0) \tag{8}$$

has also a unique solution $(x(\cdot),u(\cdot))$ defined on $[0,\infty)$. The solutions $y(\cdot)$ of (7) and $(x(\cdot),u(\cdot))$ of (8) coincide when $(y,v) \in G(T)$ and $(x,u) \in S$.

Proof. Since the function L_T is real valued, proper, l.s.c. and convex on $H \times H$, the subdifferential ∂L_T is maximal monotone (see [1], Proposition 6, Section 6.7). Because T is maximal monotone and ∂L_T is maximal monotone the existence of the unique solutions of d. i. (7) and (8) follows from Theorem 1 Section 6.8 of [1]. Next we involve the uniqueness of the solutions of (7) and (8) and the definitions of T and L_T. According to Lemma 4.3 above we know that when $(y,v) \in G(T)$ and $(x,u) \in S$ then $(y,v) \in \partial L_T(x,u)$. Thus the solutions of d. i. (7) and (8) coincide for every $t \in [0,\infty)$ for which $(y(t),v(t)) \in G(T)$ and also because $(y,v) \in \partial L_T(x,u)$ when $(x,u) \in S$. □

According to Theorem 4.4, to solve optimization problem

$$\inf_{(x,u)\in H\times H} L_T(x,u) = L_T^{inf}. \tag{9}$$

we can solve instead of d. i. (8) the easier d. i (7).

In the rest of the section we consider a class of d. i. involving the subdifferential of the Fitzpatrick functions in which every solution of the d. i. can be extended up to the solution set. Let us suppose now that $X := \{(x,u) \in H \times H | L_T(x,u) = L_T^{inf}\}$ be the solution set of the optimization problem (9) and X_ε be an ε-neighborhood of X.

Recall that the metric projection of the point x in a Hilbert space H on a subset S is defined with $Pr_S x$, i.e.

$$Pr_S x := \{y \in S| \quad \|x-y\| = \min_{z\in S} \|x-z\|\}.$$

The metric projection on a closed and convex subset of a Hilbert space exists and is unique. We denote here by $m(\mathbf{0}, \partial L_T(x,u))$ the metric projection of the origin onto $\partial L_T(x,u)$, that is,

$$m(\mathbf{0}, \partial L_T(x,u)) = \{(v,y) \in H \times H | \|(0,0)-(y,v)\| = \min_{(w,z)\in \partial L_T(x,u)} \|(0,0)-(w,z)\|\}.$$

We assume that $\mathbf{0}$ is not an element of the subdifferential. We use also the notation

$$D(\mathbf{0}, \partial L_T(x,u)) = \sup_{(w,z)\in \partial L_T(x,u)} \|(0,0)-(w,z)\|.$$

In what follows when x and u depend on a parameter t(time), we will abbreviate the subdifferential in the metric projection and in the definition of D by $\partial L_T(t)$.

Theorem 4.5. *Let us consider again d. i.* (8) *where* $\partial L_T(x(.), u(.))$ *is a maximal monotone operator. Suppose the map* L_T *is proper, convex and l.s.c., and maps* (x, u) *to a bounded, closed and convex set. For* $\varepsilon > 0$ *and* $(x_0, u_0) \notin X_\varepsilon$ *let* $\|m(\mathbf{0}, \partial L_T(x, u))\| \geq \beta > 0$, $(x, u) \notin X_\varepsilon$ *and let*

$$X := \{(x, u) \in H \times H | L_T(x, u) = L_T^{inf}\} = \{(z, w) \in H \times H | (0, 0) \in \partial L_T(z, w)\}.$$

If the d. i. (8) *has a solution and the solutions set is bounded for* $t \geq 0$, *then every solution of* $(\dot{x}, \dot{u}) \in -\frac{\partial L_T(x(t),u(t))}{\|m(\mathbf{0},\partial L_T(t))\|^2}$, $(x(0), u(0)) = (x_0, u_0)$ *can be extended up to the set* X *in finite time* $T \leq (L_T(x_0, u_0) - \underline{L_T})$ *where* $\underline{L_T}$ *is a lower bound of* $L_T(x, x^*)$.

We omit the proof, because it is similar to the proof of Theorem 3.5.

5. Iterative method

In this section we present an iterative procedure for solving the optimization problems given in Sections 3 and 4.

Definition 4.1, presented in [10], introduces a monotone operator on H by using a convex function on $H \times H$. We denote by $m(\mathbf{0}, \partial f_k)$ the metric projection of the origin onto the multivalued map $\partial f(x_k, u_k)$ and with $m(\mathbf{0}, T_{f_k})$ the metric projection of the origin onto the range of the multivalued map $\partial(\|T_f(x_k)\|^2)$
We replace the following iterative procedure

$$(x_{k+1}, u_{k+1}) = (x_k, u_k) - \frac{f(x_k, u_k)m(\mathbf{0}, \partial f_k)}{\|m(\mathbf{0}, \partial f_k)\|^2}, \; k = 0, 1, \ldots. \tag{10}$$

which is based on the d. i. (6), with the following simpler iterative procedure:

$$x_{k+1} = x_k - \frac{\|T_f(x_k)\|^2 m(\mathbf{0}, T_{f_k})}{\|m(\mathbf{0}, T_{f_k})\|^2}, \; u_k = T_f(x_k), \; k = 0, 1, \ldots. \tag{11}$$

based on d. i.: $\dot{x} \in -\dfrac{T_f(x(t))}{\|m(\mathbf{0}, T_f(t))\|^2}, \quad x(0) = x_0.$

The following theorem asserts that the sequence (11) is monotone.

Theorem 5.1. *Consider the optimization problem* (1) *where* $f(x, u)$ *is a convex function on* $H \times H$ *and the iterative procedure* (11). *Suppose there exist an* ε-*neighborhood ,* X_{f_ε} *of the set* X_f *for which* $\langle(\eta, \xi), (x, u) - (y, v)\rangle \neq 0$, $(\eta, \xi) \in \partial f(x, u)$, $(y, v) \in Pr_{X_f}(x, u)$, $(x, u) \in X_{f_\varepsilon} \setminus X_f$ *and*

$$\liminf_{k\to\infty} \|m(\partial f(x_k, u_k))\| \neq 0,$$

where inf *is taken at all sequences* $(x_k, u_k) \in X_{f_\varepsilon} \setminus X_f$.
We recall that $\langle(v, y), (x, u)\rangle = \langle v, x\rangle + \langle u, y\rangle$ *for all* x *and* y *in* H *and* $u(x)$ *and* $v(y)$ *in* H. *If* $(x_0, u_0) \in X_{f_\varepsilon}$ *and the following inequality is satisfied:*

$$f(x_k, u_k) - \langle m(\mathbf{0}, \partial f_k), (x_k, u_k) - (y_k, v_k)\rangle < \varepsilon^k, \tag{12}$$

where $(y_k, v_k) \in Pr_{X_f}(x_k, u_k)$ *and* $\varepsilon^k < \frac{q}{2} f(x_k, u_k)$, $0 < q < 1$, *then the following inequality*

$$\max \|(x_{k+1}, u_{k+1}) - Pr_{X_f}(x_{k+1}, u_{k+1})\| \leq \min \|(x_k, u_k) - Pr_{X_f}(x_k, u_k)\| \tag{13}$$

is valid for every $k = 0, 1, 2, \ldots$.

On the basis of inequalities (12) *and* (13) *we obtain the following inequalities:*

$$\|T_f(x_k)\|^2 - \big\langle m(\mathbf{0}, T_{f_k}), (x_k - y_k)\big\rangle < \varepsilon_1^k, \tag{14}$$

where $(y_k, v_k) \in Pr_{X_f}(x_k, u_k)$ *and* $\varepsilon_1^k < \frac{q}{2}\|T_f(x_k)\|^2$, $0 < q < 1$, *and finally:*

$$\max\, \|x_{k+1} - y_{k+1}\| \leq \min\, \|x_k - y_k\| \quad for\, every\, k = 0, 1, 2, \dots \tag{15}$$

And therefore the iterative procedure (10) *can be replaced with the iterative procedure* (11)*.*

Proof. As long as $f(x,u)$ is convex under (13), for any $(y_k, v_k) \in Pr_{X_f}(x_k, u_k)$, $k = 0, 1, 2, \dots$, we obtain

$$\begin{aligned}
0 \leq A_k = \frac{\|x_{k+1} - y_{k+1}\|^2}{\|x_k - y_k\|^2} &\leq \frac{\|(x_{k+1}, u_{k+1}) - (y_{k+1}, v_{k+1})\|^2}{\|(x_k, u_k) - (y_k, v_k)\|^2} \\
= \frac{\|(x_{k+1}, u_{k+1}) - (y_{k+1}, v_{k+1})\|^2}{\|(x_k, u_k) - (y_k, v_k)\|^2} &\leq \frac{\|(x_{k+1}, u_{k+1}) - (y_k, v_k)\|^2}{\|(x_k, u_k) - (y_k, v_k)\|^2} \\
= \frac{1}{\|(x_k, u_k) - (y_k, v_k)\|^2} &\|(y_k, v_k) - (x_k, u_k) + \frac{f(x_k, u_k) m(\mathbf{0}, \partial f_k)}{\|m(\mathbf{0}, \partial f_k)\|^2}\|^2 \\
= \frac{(f(x_k, u_k))^2}{\|m(\mathbf{0}, \partial f_k)\|^2 \|(x_k, u_k) - (y_k, v_k)\|^2} &+ 1 + 2\frac{f(x_k, u_k)\big\langle m(\mathbf{0}, \partial f_k), (y_k, v_k) - (x_k, u_k)\big\rangle}{\|m(\mathbf{0}, \partial f_k)\|^2 \|(x_k, u_k) - (y_k, v_k)\|^2}.
\end{aligned}$$

Under the inequality $\varepsilon^k \leq \frac{q}{2} F(x_k, u_k)$ and (12), we complete the proof:

$$\begin{aligned}
A_k &\leq \frac{(f(x_k, u_k))^2}{\|m(\mathbf{0}, \partial f_k)\|^2 \|(x_k, u_k) - (y_k, v_k)\|^2} + 1 + 2\frac{(\varepsilon^k - f(x_k, u_k)) f(x_k, u_k)}{\|m(\mathbf{0}, \partial f_k)\|^2 \|(x_k, u_k) - (y_k, v_k)\|^2} \\
&= 1 - \frac{(f(x_k, u_k))^2}{\|m(\mathbf{0}, \partial f_k)\|^2 \|(x_k, u_k) - (y_k, v_k)\|^2} + 2\frac{\varepsilon^k f(x_k, u_k)}{\|m(\mathbf{0}, \partial f_k)\|^2 \|(x_k, u_k) - (y_k, v_k)\|^2} \\
&\leq 1 - (1 - q)\frac{(f(x_k, u_k))^2}{\|m(\mathbf{0}, \partial f_k)\|^2 \|(x_k, u_k) - (y_k, v_k)\|^2} = 1 - (1 - q)(B_k)^2 < 1,
\end{aligned}$$

where
$$B_k = \frac{(f(x_k, u_k))^2}{\|m(\mathbf{0}, \partial f_k)\|^2 \|(x_k, u_k) - (y_k, v_k)\|^2}.$$

Thus, for every $(x_k, u_k) \in X_{f_\varepsilon} \setminus X_f$, $k = 1, 2, \dots$ we obtain

$$0 \leq A_k < 1 - (1 - q)(B_k)^2 < 1,$$

hence there exists $\varepsilon_1^k \leq \varepsilon^k$ such that (15) is satisfied. □

Applications

We apply now the introduced idea with Fitzpatrick functions to use convexity to create simplified algorithm for solving unconstrained smooth minimization problems by reducing the number of independent variables. We present two iterative procedures of one independent variable for solving convex minimization problems of two independent variables. We show with a comparative study that the new iterative procedures solve the convex minimization problems faster with the same precision than the existing procedures.

Comparison algorithms

We compare the results of our algorithm with the results of four other algorithms. Two of them are well-known variants of Newton method from MATLAB Optimization Tool: FMINUNC and FMINSEARCH. The third one is The Proximal-point algorithm presented by Bauschke and Combetts (see Theorem 23.41 from [3]). Similar algorithm is presented also by Pavel an Raykov (see [12]). The fourth one is The Regularized Minimization Algorithm introduced also by Bauschke and Combetts (see Theorem 27.23 from [3])

We present in short the last two algorithms and the statements of the theorems of their covergence here.

We introduce next Theorem 5.2 (Theorem 23.41 from [3]) proving the convergence of the proximal-point algorithm.

Theorem 5.2. (Theorem 23.41 (Proximal-point algorithm) from [3])
Let $A : H \to 2^H$ be a maximally monotone operator such that $zer A \neq \emptyset$, let $(\gamma)_{n\in N}$ be a sequence in R_{++} such that $\Sigma_{n\in N}\gamma_n^2 = +\infty$, and let $x_0 \in H$.
Set $(\forall n \in N)$ $x_{n+1}J_{\gamma_n A}x_n$. Then the following hold:

(i) *$(x_n)_{n\in N}$ converges weakly to a point in $zer A$.*

(ii) *Suppose that A is uniformly monotone on every bounded subset of H. Then $(x_n)_{n\in N}$ coverges strongly to the unique point in $zer A$.*

In the above theorem $J_{\gamma_n A}$ is the *resolvent* of A: $J_{\gamma_n A} = (Id + \gamma_n A)^{-1}$. For the proof see [3].

Method of Regularization of Minimization Problems (see [3])

Let $f \in \Gamma_0(H)$ and suppose that Argmin $f \neq \emptyset$, i.e., the minimization problem $\min_{x\in H} f(x)$ has at least one solution. In order to obtain a specific minimizer, one can introduce, for every $\varepsilon \in]0, 1[$, the regularized problem $\min_{x\in H} f(x) + \varepsilon g(x)$, where $g \in \Gamma_0(H)$. The objective is to choose the regularization function g such that the last minimization problem admits a unique solution x_ε and such that the net $(x_\varepsilon)_{\varepsilon\in]0,1[}$ converges to a specific point of Argmin f. We choose $g = (1/2)\|\cdot\|^2$ (*Tykhonov Regularization* framework.) In this case the regularization is equivalent to find $x_\varepsilon \in H$ such that $0 \in \partial f(x_\varepsilon) + \varepsilon x_\varepsilon$. If we denote by x_0 the minimum norm minimizer of f, Theorem 23.44(i) ([3]) proves that $x_\varepsilon \to x_0$ as $\varepsilon \downarrow 0$.

The next Theorem 5.3 (Theorem 27.23 from [3]) explores the asymptotic behavior of the curve $(x_\varepsilon)_{\varepsilon\in]0,1[}$ for more general choices of the regularization function g.

Theorem 5.3. (Theorem 27.23 from [3]) *Let f and g be in $\Gamma_0(H)$. Suppose that* Argmin $f\cap$ dom $g \neq \emptyset$ *and that g coercive and strictly convex. Then g admits a unique minimizer x_0 over* Argmin$f$ *and, for every $\varepsilon \in]0, 1[$, the regularized problem $\min_{x\in H} f(x)+\varepsilon g(x)$ admits a unique solution x_ε. Moreover, the following statements hold:*

(i) $(x_\varepsilon) \rightharpoonup x_0$ *as* $\varepsilon \downarrow 0$.

(ii) $g(x_\varepsilon) \to g(x_0)$ *as* $\varepsilon \downarrow 0$.

(iii) *Suppose that g is uniformly convex on every closed ball in H.*
Then $(x_\varepsilon) \to x_0$ as $\varepsilon \downarrow 0$.

For the proof see [3].

We present now two examples:

Example 1. Let us solve the unconstrained convex optimization problem:

$$\min f_p(x,y), \text{ where } f_p(x,y) = 4(x+p)^2 + 4y^2 - 4(x+p)y - 12y,$$

p is a parameter of disturbances. The partial derivatives should be equal to zero at the min $f_p(x,y)$:

$$f_{p_x}(x,y) = 8(x+p) - 4y = 0 \quad \text{and} \quad f_{p_y}(x,y) = 8y - 4(x+p) - 12 = 0.$$

We have to find the conditions for the parameter of distubances such that to have a monotone convergence algorithm. Because both partial derivatives should be equal to zero at the solution point, then we have

$$y_p = 2(x+p) \quad \text{and} \quad y_p = \frac{x+p+3}{2}. \tag{16}$$

They are monotone, because $f_p(x,y)$ is convex. These two lines intersect at different solution points for a given value of the parameter p and we use a sequence $\{p_k\}$ to create an iterative procedure depending only on x_{p_k} and the parameter of disturbances p_{k+1} on each step. As the solution is a unique one when $p=0$ we can find sufficient conditions for the parameter of disturbances p to ensure convergence. We create an iterative procedute for $x_{p_{k+1}}$ for finding the minimum of $f_p(x,y)$:

$$x_{p_{k+1}} = x_{p_k} - \frac{\frac{9}{4}(x_{p_k} + p_{k+1} - 1)^2}{\frac{9}{2}(x_{p_k} + p_{k+1} - 1)} = x_{p_k} - \frac{x_{p_k} + p_{k+1} - 1}{2}$$

To ensure convergence we have to have $\max \|(x_{p_k}, y_{p_k}) - (x_{p_{k+1}}, y_{p_{k+1}})\| < \varepsilon_k$, where the sequence $\{\varepsilon_k\}$ is convergent to zero, i.e. the $\liminf_{k\to\infty} \varepsilon_k = 0^+$, $\varepsilon_k > 0$ for every positive integer k. Because

$$\max |x_{p_k} - x_{p_{k+1}}| \le \max \|(x_{p_k}, y_{p_k}) - (x_{p_{k+1}}, y_{p_{k+1}})\| < \varepsilon_k,$$

means that $|x_{p_k} - x_{p_{k+1}}| = |x_{p_k} - 1 + p_{k+1}| < 2\varepsilon_k$. We receive

$$|x_{p_k} - 1 + p_{k+1}| \le |x_{p_k} - 1| + |p_{k+1}| < 2\varepsilon_k^1$$

where the sequence $\{\varepsilon_k^1\}$ is also convergent to zero, i.e. the $\liminf_{k\to\infty} \varepsilon_k^1 = 0^+$, $\varepsilon_k^1 \ge \varepsilon_k > 0$ for every positive integer k, such that when $|x_{p_k} - 1| + |p_{k+1}| < 2\varepsilon_k^1$ then $|x_{p_k} - 1 + p_{k+1}| < 2\varepsilon_k$, or

$$0 \le |p_{k+1}| < 2\varepsilon_k^1 - |x_{p_k} - 1|, \quad \text{or} \quad |x_{p_k} - 1| - 2\varepsilon_k^1 \le p_{k+1} \le 2\varepsilon_k^1 - |x_{p_k} - 1|.$$

From equations (16) it follows that when p_k approaches 0 and x_{p_k} approaches $x^* = 1$ then y_{p_k} approaches $y^* = 2$. We can simplify the iterative procedure by presenting the disturbances on each step k for any $x_k \in \{x_{p_k}\}$ with the addition of one parameter p_{k+1}: $x_{p_{k+1}} = x_k - \frac{x_k - 1}{2} + p_{k+1}$. To obtain for any $x_{k+1} \in \{x_{p_{k+1}}\}$ $\lim_{k\to\infty} |x_k - x_{k+1}| \to 0$, p_{k+1} should satisfy the inequalities

$$-\left|\frac{x_k - 1}{2}\right| < p_{k+1} < \left|\frac{x_k - 1}{2}\right|$$

for every $x_k \in \{x_{p_k}\}$.

Example 2. We consider the following unconstrained convex optimization problem in which the surface is flatter close to the solution point: $\min f_p(x,y)$, where

$$f_p(x,y) = (x+p)^6 + e^{(x+p)} + (x+p)^2 + 2(x+p)y + y^2 + y^4,$$

p is again a parameter of disturbances.

The partial derivatives should be equal to zero at the $\min f_p(x,y)$

$$f_{p_x}(x,y) = 6(x+p)^5 + e^{(x+p)} + 2(x+p) + 2y = 0 \text{ and } f_{p_y}(x,y) = 4y^3 + 2y + 2(x+p) = 0.$$

Because both partial derivatives should be equal to zero at the solution point, then we have

$$y_p = -\frac{6(x+p)^5 + e^{(x+p)} + 2(x+p)}{2} \quad \text{and} \quad y_p = \Big(\frac{6(x+p)^5 + e^{(x+p)}}{4}\Big)^{1/3} \tag{17}$$

These two lines intersect at different solution points for a given valueof the parameter p and we use a sequence $\{p_k\}$, $(|p_k| \downarrow 0)$ to create an iterative procedure depending only on x_{p_k} and the parameter of disturbances p_{k+1} on each step. As the solution is a unique one when $p=0$ we can find sufficient conditions for the parameter of disturbances p to ensure convergence. We create an iterative procedute for $x_{p_{k+1}}$ for finding the minimum of $f_p(x,y)$:

$$\begin{aligned}
F_{p_{k+1}} &= \Big\{\Big(\frac{6(x_{p_k}+p_{k+1})^5 + e^{(x_{p_k}+p_{k+1})}}{4}\Big)^{1/3} \\
&\qquad + \frac{6(x_{p_k}+p_{k+1})^5 + 2(x_{p_k}+p_{k+1}) + e^{(x_{p_k}+p_{k+1})}}{2}\Big\}^2, \\
G_{p_{k+1}} &= 2\Big\{\Big(\frac{6(x_{p_k}+p_{k+1})^5 + e^{(x_{p_k}+p_{k+1})}}{4}\Big)^{1/3} \\
&\qquad + \frac{6(x_{p_k}+p_{k+1})^5 + 2(x_{p_k}+p_{k+1}) + e^{(x_{p_k}+p_{k+1})}}{2}\Big\}, \\
H_{p_{k+1}} &= \Big\{\Big(\frac{1}{12}\Big)\Big[\frac{4}{6(x_{p_k}+p_{k+1})^5 + e^{(x_{p_k}+p_{k+1})}}\Big]^{2/3} (30(x_{p_k}+p_{k+1})^4 + e^{(x_{p_k}+p_{k+1})}) \\
&\qquad + 0.5(30(x_{p_k}+p_{k+1})^4 + e^{(x_{p_k}+p_{k+1})} + 2)\Big\}.
\end{aligned}$$

Then $x_{p_{k+1}} = x_{p_k} - \frac{F_{p_{k+1}}}{G_{p_{k+1}} H_{p_{k+1}}}$. To ensure convergence we should have

$$\max \|(x_{p_k}, y_{p_k}) - (x_{p_{k+1}}, y_{p_{k+1}})\| < \varepsilon_k,$$

where the sequence $\{\varepsilon_k\}$ is convergent to zero, i.e. the $\liminf_{k\to\infty} \varepsilon_k = 0^+$, $\varepsilon_k > 0$ for every positive integer k. Because

$$\max |x_{p_k} - x_{p_{k+1}}| \le \max \|(x_{p_k}, y_{p_k}) - (x_{p_{k+1}}, y_{p_{k+1}})\| < \varepsilon_k,$$

means that $|x_{p_k} - x_{p_{k+1}}| = \frac{F_{p_{k+1}}}{|G_{p_{k+1}}| H_{p_{k+1}}} < \varepsilon_k$, since $F_{p_{k+1}} \ge 0$ and $H_{p_{k+1}} \ge 0$ for every p_{k+1} satisfying the convergence conditions. Therefore we have to restrict p_{k+1} such that, when $|p_{k+1}| < \varepsilon^1_{k+1}$, then $\frac{F_{p_k}}{|G_{p_k}| H_{p_k}} < \varepsilon_k$. Here $\{\varepsilon^1_k\}$ is a decreasing sequence with positive elemets that ensures these requirements. From equations

(17) it follows that when p_k approaches 0 and x_{p_k} approaches $x^* = -0.5493$ then y_{p_k} approaches $y^* = 0.4107$.

We can simplify the iterative procedure by presenting the disturbances on each step k for any $x_k \in \{x_{p_k}\}$ with addition a one parameter p_{k+1}. We use the following simplified analogical notations:

$$F_k = \Big\{\Big(\frac{6x_k^5 + e^{x_k}}{4}\Big)^{1/3} + \frac{6x_k^5 + 2x_k + e^{x_k}}{2}\Big\}^2,$$

$$G_k = 2\Big\{\Big(\frac{6x_k^5 + e^{x_k}}{4}\Big)^{1/3} + \frac{6x_k^5 + 2x_k + e^{x_k}}{2}\Big\},$$

and
$$H_k = \Big\{\Big(\frac{1}{12}\Big)\Big[\frac{4}{6x_k^5 + e^{x_k}}\Big]^{2/3}(30x_k^4 + e^{x_k}) + 0.5(30x_k^4 + e^{x_k} + 2)\Big\}.$$

and receive the iterative procedure: $x_{p_{k+1}} = x_k - \frac{F_k}{G_k H_k} + p_{k+1}$. To get for any $x_k \in \{x_{p_k}\}$ and $x_{k+1} \in \{x_{p_{k+1}}\}$ the limit $\lim\limits_{k\to\infty} |x_k - x_{k+1}| \to 0$ where $|x_k - x_{k+1}| = \Big|\frac{F_k}{|G_k|H_k} - p_{k+1}\Big|$, p_{k+1} should satisfy the next inequalities:

$$-\frac{F_k}{|G_k|H_k} < p_{k+1} < \frac{F_k}{|G_k|H_k}$$

6. Numerical Results

We present in this section a comparative study of results with our algorithm with the results from four other algorithms for Examples considered at the end of Section 5 using MATLAB. Our results are for both Examples with initial point $(x_0, y_0) = (5, 5)$ and with precision $\varepsilon = 0.000001$.

Numerical results for Example 1.

(a) The results with our algorithm are: $x = 1.0000$ and $y = 2.0000$ for time $t = 1.8896 \times 10^{-4}$ sec.

(b) The results with FMINUNC (MATLAB) are: $x = 1.0000$ and $y = 2.0000$ for time $t = 0.0089$ sec.

(c) The results with FMINSEARCH (MATLAB) are: $x = 1.0000$ and $y = 2.0000$ for time $t = 0.0025$ sec.

(d) The results with Proximal-point algorithm are: $x = 1.0000$ and $y = 2.0000$ for time $t = 2.0426 \times 10^{-3}$ sec.

(e) The results with Regularization algorithm are: $x = 1.0000$ and $y = 2.0000$ for time $t = 0.0143$ sec.

These results are summarized in Table 1:

Algorithm	x - results	y - results	time
Our Algorithm	1.0000	2.0000	1.8896 $\times 10^{-4}$ sec.
FMINUNC	1.0000	2.0000	0.0089 sec.
FMINSEARCH	1.0000	2.0000	0.0026 sec.
Proximal-point	1.0000	2.0000	2.0426$\times 10^{-3}$ sec.
Regularization	1.0000	2.0000	0.0143 sec.

The advantages of our algorithm and the proximal-point algorithm are obvious. See next Figure 1.

$$f(x, y) = 4x^2 + 4y^2 - 4xy - 12y$$

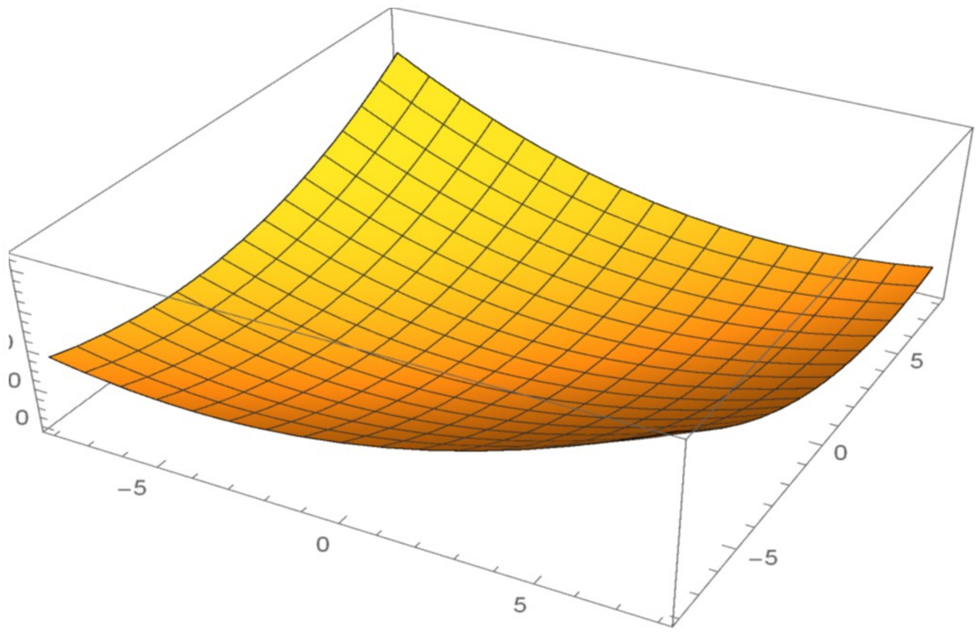

Figure 1

Numerical Results for Example 2.

We compare our results with the results with the four other algorithms next:

(a) The results with our algorithm are: $x = -0.5493$ and $y = 0.4107$ for time $t = 3.6133 \times 10^{-4}$ sec.

(b) The results with FMINUNC (MATLAB) are: $x = -0.5487$ and $y = 0.4117$ for time $t = 0.0110$ sec.

(c) The results with FMINSEARCH (MATLAB) are: $x = -0.5493$ and $y = 0.4107$ for time $t = 0.0034$ sec.

(d) The results with the Proximal-point algorithm are: $x = -0.5493$ and $y = 0.4107$ for time $t = 3.2726 \times 10^{-3}$ sec.

(e) The results with Regularization algorithm are: $x = -0.5469$ and $y = 0.4120$ for time $t = 0.0175$ sec.

These results are summarized in the following Table 2.

Algorithm	x - results	y - results	time
Our Algorithm	-0.5493	0.4107	3.6133 $\times 10^{-4}$ sec.
FMINUNC	-0.5487	0.4117	0.0110 sec.
FMINSEARCH	-0.5493	0.4107	0.0034 sec.
Proximal-point	-0.5493	0.4107	3.2726$\times 10^{-3}$ sec.
Regularization	-0.5469	0.4120	0.0175 sec.

The advantages of our algorithm and of the Proximal-point algorithm are clear again. See next Figure 2.

$$f(x, y) = x^6 + e^x + x^2 + 2xy + y^2 + y^4$$

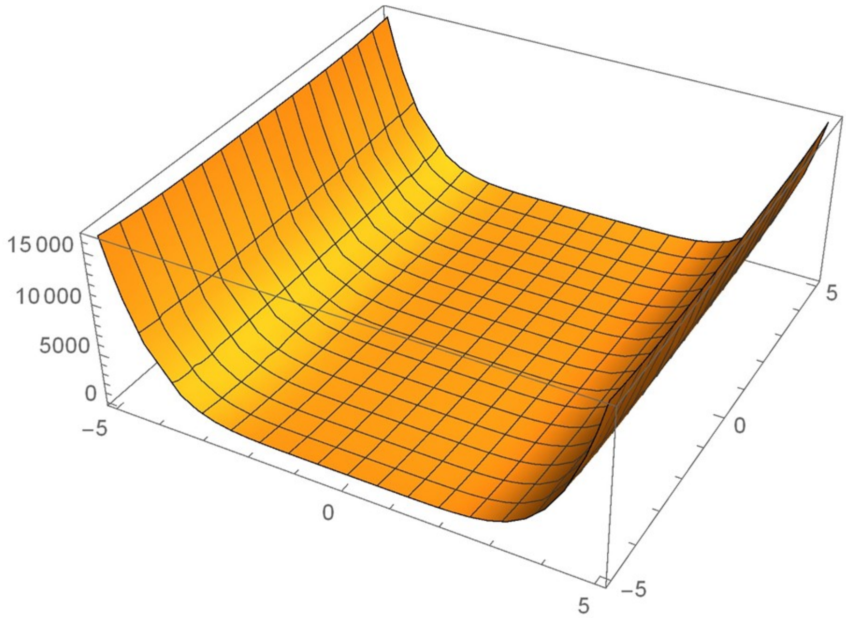

Figure 2

7. A new example of a Fitzpatrick function

We proved in Theorem 3.5 that the solution of optimization problem (1) can be found by solving d.i. (3) and also by solving d.i. (6) instead of d.i. (3). We showed that the solutions of d. i. (3) and (6) coincide. We proved there also that $f(x, u)$ is a Lyapunov function for d. i. (6). Therefore this is an example that the Fitzparick function appears to be a Lyapunov fuction. We showed in Theorem 5.1 that to solve the problem numerically we can substitute its corresponding iterative procedure (10) which is based on the d. i. (6), with the iterative procedure (11) A Lyapunov function for d. i. (9) is $(\|T_f(x)\|)^2$. Ivanov and Raykov in [11] developed a parametric Lyapunov function method for solving nonlinear systems in Hilbert space. We show close connections between the Lyapunov function created to solve the given problem and Fitzpatrick functions.

Statement of the Lyapunov Function Method

Consider $f_i(x, u), i = 1, 2, \ldots, m$ be nonsmooth scalar functions which are defined on $H \times H$ where $\mathbf{H}$ is a Hilbert space and let us consider the following system of equations

$$f_i(x, u) = 0, \quad i = 1, 2, \ldots, m. \tag{18}$$

A family of functions depending on parameter $p(\varepsilon)$ is constructed:

$$V_{p(\varepsilon)}(x, u) = \sum_{i=1}^{m} \alpha_i |f_i(x, x^*)|^{1+p_i(\varepsilon)} \geq 0, \tag{19}$$

where $\alpha_i > 0$, and $p(\varepsilon) = (p_1(\varepsilon), p_2(\varepsilon), \ldots, p_m(\varepsilon))$ is a vector parameter such that $p_i(\varepsilon) \in (p_i - \varepsilon, p_i + \varepsilon)$ and $p_i - \varepsilon > -1, \varepsilon > 0, \quad i = 1, 2, \ldots, m$.

According to [7] the generalized gradient is

$$\partial f(x,u) = \{(\eta,\zeta) \in \mathbf{H}\times\mathbf{H} \,|\, f^0((x,u);(t,s)) \geq \big\langle(\eta,\zeta),(t,s)\big\rangle\, \forall\,(t,s)\in \mathbf{H}\times\mathbf{H}\}$$

We consider functions $f_i(x,u)$, $i=1,2,\dots,m$ that are regular (see [7]) and assume that $V_{p(\varepsilon)}(x,u)$ is a convex function. Note that the convex function is regular. We denote by $m(\mathbf{0},\partial V_{p(\varepsilon)}(x,u)$ the metric projection of the origin to the set $\partial V_{p(\varepsilon)}(x,u)$. Let $0 \notin \partial V_{p(\varepsilon)}(x_0,u_0)$. Consider the following d. i.:

$$(\dot{x},\dot{u}) \in -\frac{\partial V_{p(\varepsilon)}(x,u)}{\|m(\mathbf{0},\partial V_{p(\varepsilon)}(x,u))\|^2}, \quad (x(0),u(0)) = (x_0,u_0). \tag{20}$$

As long as $f_i(x)$ are regular functions, according to [8] we have

$$\partial V_{p(\varepsilon)}(x,u) = \sum_{i=1}^{m} \alpha_i \|f_i(x,u)\|^{p_i(\varepsilon)}(\operatorname{sign} f_i(x,u))\partial f_i(x,u)(1+p_i(\varepsilon)). \tag{21}$$

The d. i. (20) is the Filippov extension of the following d.e. with possibly discontinuous right-hand side, see [1] and [9]:

$$(\dot{x},\dot{u}) = -\frac{m(\partial V_{p(\varepsilon)}(x,u))}{\|m(\mathbf{0},\partial V_{p(\varepsilon)}(x,u))\|^2}, \quad (x(0),u(0)) = (x_0,u_0) \tag{22}$$

As long as $V_{p(\varepsilon)}(x)$ is a convex function, the subdifferential $\partial V_{p(\varepsilon)}(x)$ is a maximal monotone operator and the d. i.

$$(\dot{x},\dot{u}) \in -\partial V_{p(\varepsilon)}(x,u), \quad (x(0),u(0)) = (x_0,u_0) \tag{23}$$

has an unique solution, see [1]. We see that so created Lyapunov function $V_{p(\varepsilon)}(x,u)$ is also an example of a Fitzpatrick function. Using the Lyapunov Function Method, the number of the equations is independent from the number of independent variables.

We could consider at the end a simplification of optimization problem (1) using $T_f x$ from Definition 3.1. By solving d.i. (2) we will solve problem (1), because $(\|T_f x\|)^2$ is a Lyapunov function.

References

[1] J. P. Aubin, I. Ekeland: *Applied Nonlinear Analysis*, Wiley-Interscience, New York (1984).

[2] S. Bartz, H. H. Bauschke, J. M. Borwein, S. Reich, X. Wang: *Fitzpatrick functions, cyclic monotonicity and Rockafellar's antiderivative*, Nonlinear Analysis 66 (2007) 1198–1223.

[3] h. H. Bauschke, P. L. Combettes: *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, Berlin (2017).

[4] H. H. Bauschke, D. A. McLaren, H. S. Sendov: *Fitzpatrick functions: Inequalities, examples, and remarks on a problem by S. Fitzpatrick*, J. Convex Analysis 13/3 (2006) 499–523.

[5] M. S. Bazara, J. J. Goode, M. Z. Nashed: *On the cones of tangents with applications to mathematical programming*, 13/4 (1974) 389–426.

[6] J. M. Borwein, J. Dutta: *Maximal monotone inclusions and Fitzparick functions*, J. Optimization Theory Appl. 171/3 (2016) 757–784.

[7] F. H. Clarke: *Optimization and Nonsmooth Analysis*, Wiley & Sons, New York (1983).

[8] K. Deimling: *Multivalued Differential Equations*, de Gruyer, Berlin (1992).

[9] A. F. Filippov: *Differential Equations with Discontinues Right Hand Side (Russian)*, Nauka, Moscow (1985).

[10] S. Fitzpatrick: *Representing monotone operators by convex functions*, in: *Workshop/Miniconference on Functional Analysis and Optimization*, Canberra 1988, Proc. Centre for Mathematical Analysis 20, Australian National University, Canberra (1988) 59–65.

[11] R. P. Ivanov, I. Raykov: *Parametric Lyapunov function method for solving nonlinear systems in Hilbert spaces*, Numer. Funct. Analysis Optimization 17 (1996) 893–901.

[12] N. Pavel, I. Raykov: *Parametric proximal-point algorithms*, J. Optimization Theory Appl. 139/1 (2008) 85–107.

[13] J.-P. Penot: *The relevance of convex analysis for the study of monotonicity*, Nonlinear Analysis Theory Meth. Appl. 58 (2004) 855–871.

[14] R. T. Rockafellar: *On the maximal monotonicity of subdifferential mappings*, Pacific J. Math. 33 (1970) 209–216.

[15] S. Simons, C. Zălinescu: *Fenchel duality, Fitzpatrick functions and maximal monotonicity*, J. Nonlinear Convex Analysis 6/1 (2005) 1–22.