Journal of Convex Analysis
Volume 24 (2017), No. 2, 525–545

The Monotonicity of Perturbed Gradients of Convex Functions

Cornel Pintea*
Babeş-Bolyai University, Faculty of Mathematics and Computer Science, Str. Kogălniceanu no. 1, 400084 Cluj-Napoca, Romania
cpintea@math.ubbcluj.ro

Tiberiu Trif
Babeş-Bolyai University, Faculty of Mathematics and Computer Science, Str. Kogălniceanu no. 1, 400084 Cluj-Napoca, Romania
ttrif@math.ubbcluj.ro

Received: November 4, 2014
Revised manuscript received: March 8, 2016
Accepted: March 10, 2016

In this paper we perturb the (Minty-Browder monotone) gradients of some convex functions f, defined on convex open subsets of the Euclidean space with suitable linear transformations A and provide sufficient conditions on f and A in order to obtain the η-monotonicity of the perturbation $\nabla f + A$ for some $\eta \in (-1, 0)$. These conditions are given in terms of the magnitude of the Hessian operators associated to f and the spectra of the operators $(A + A^*)/2$ and A^*A and they make the Minty-Browder monotonicity of the perturbation $\nabla f + A$ to fail. While the symmetric perturbations are only considered as instructive examples in the Minty-Browder monotone case, we mostly rely, in the even dimensional context, on linear transformations which can be represented as linear combinations between the identity operator and skew-symmetric isometries. Note that the skew-symmetric isometries of $\mathbb{R}^{2m}$ are complex structures on $\mathbb{R}^{2m}$. In the bidimensional case the rotations do admit such representations and the perturbations of the gradients of convex functions by rotations were treated before in [9]. We also obtain some global injectivity results for suitable linear perturbations of some gradients of convex functions.

Keywords: Minty-Browder monotone operator, η-monotone operator, h-monotone operator

2010 Mathematics Subject Classification: 47H05, 15A18, 15B57, 47A12

1. Introduction

Let $(H, \langle\cdot,\cdot\rangle)$ be a real Hilbert space, and let D be a nonempty subset of H. An operator $T : D \to H$ is called *Minty-Browder monotone* if (see, for instance, [4, 12])

$$\langle Tx - Ty, x - y\rangle \geq 0 \quad \text{for all } x, y \in D.$$

*The first author was supported by a grant of the Romanian National Authority for Scientific Research CNCS - UEFISCDI, project number PN-II-ID-PCE-2011-3-0024.

ISSN 0944-6532 / $ 2.50

The operator T is called *h-monotone* if

$$\frac{Tx - Ty}{\|Tx - Ty\|} \neq \frac{y - x}{\|y - x\|} \quad \text{for all } (x, y) \in (D \times D) \setminus \ker(T),$$

where $\ker(T) := \{(x, y) \in D \times D \mid Tx = Ty\}$. Taking into consideration the Cauchy-Schwarz inequality, as well as the case when equality occurs in it, it is immediately seen that T is h-monotone if and only if

$$\left\langle \frac{Tx - Ty}{\|Tx - Ty\|}, \frac{x - y}{\|x - y\|} \right\rangle > -1 \quad \text{for all } (x, y) \in (D \times D) \setminus \ker(T).$$

Obviously, every Minty-Browder monotone operator is also h-monotone.

Given any number $\eta \in (-1, 1)$, an operator $T : D \to H$ is called *η-monotone* if

$$\left\langle \frac{Tx - Ty}{\|Tx - Ty\|}, \frac{x - y}{\|x - y\|} \right\rangle \geq \eta \quad \text{for all } (x, y) \in (D \times D) \setminus \ker(T).$$

If $\eta \in (-1, 0)$, then the class of η-monotone operators is an intermediate one, lying between the class of Minty-Browder monotone operators and that of h-monotone operators. If $\eta \in [0, 1)$, then the class of η-monotone operators is contained in that of Minty-Browder monotone operators.

The class of h-monotone operators was introduced by G. Kassay and C. Pintea [8]. However, only one example of an h-monotone operator which is not Minty-Browder monotone is provided in [8]. Moreover, that example (see [8, Example 2.5] is given in the framework of the 2-dimensional Euclidean space. The class of η-monotone operators was introduced by D. Marian, I. R. Peter, and C. Pintea [9]. They extended the study of such operators in [10] and provided examples of η-monotone operators which are not Minty-Browder monotone, both in [9] and [10], but again almost all these examples are given in dimension two.

The purpose of the present paper is to fill the gaps by providing examples of η-monotone (respectively h-monotone) operators which are not Minty-Browder monotone in the n-dimensional setting. The η-monotonicity is used to derive several global injectivity results. Our examples are obtained by perturbing the gradient of a differentiable convex function by an appropriate linear operator. Recall that the gradient of a C^1-smooth convex function is well known to be a Minty-Browder monotone operator. In fact, the Minty-Browder monotonicity of ∇f ensures the convexity of f as well [6, Theorem 4.1.4]. For a characterization of the convexity (quasiconvexity) in terms of Minty-Browder monotonicity (quasimonotonicity) of the, possibly multivalued, Clarke subdifferential operator for less regular functions we refere the reader to [5] ([2]).

1.1. Background and notation

Throughout the paper we make use of the notation described below. Let D be a nonempty open convex subset of $\mathbb{R}^n$, and let $f : D \to \mathbb{R}$ be a C^2-smooth convex function. The Hessian matrix of f at an arbitrary point $x \in D$ will be denoted by $H_f(x)$. Recall that $H_f(x)$ is a symmetric matrix and it defines a symmetric bilinear functional

$$\mathcal{H}_f(x) : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R},\ \mathcal{H}_f(x)(u,v) := u \cdot H_f(x) \cdot v^T.$$

Denote by $h_f(x) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ the linear transformation defined by the following equality

$$\mathcal{H}_f(x)(u,v) := \langle h_f(x)u, v\rangle\,, \quad \forall u,v \in \mathbb{R}^n.$$

In fact $h_f(x)u = u \cdot H_f(x)$. Note that the operator $h_f(x)$ is symmetric, i.e. $\langle h_f(x)u, v\rangle = \langle u, h_f(x)v\rangle$, for all $u, v \in \mathbb{R}^n$. We also set

$$\sigma_f := \sup_{z \in D} \|h_f(z)\|.$$

Further, let $A : \mathbb{R}^n \to \mathbb{R}^n$ be a linear operator, and let $T : D \to \mathbb{R}^n$ be the operator defined by $Tx := \nabla f(x) + Ax$. We shall denote by $[A]$ the matrix representation of A with respect to the standard basis of $\mathbb{R}^n$. Let S^{n-1} denote the unit sphere in $\mathbb{R}^n$, and let

$$W(A) := \{\langle Ax, x\rangle \mid x \in S^{n-1}\}$$

be the *numerical range* of A. It is well known that $W(A) = [\lambda_A, \mu_A]$, where λ_A and μ_A denote the smallest and the greatest eigenvalue, respectively, of the symmetric operator $(A + A^*)/2$. From the inequalities [6, p. 154]

$$\lambda_A \le \langle Ax, x\rangle \le \mu_A \quad \text{for all } x \in S^{n-1}$$

one can easily deduce that

$$\lambda_A\|x-y\|^2 \le \langle A(x-y), x-y\rangle \le \mu_A\|x-y\|^2 \quad \text{for all } x, y \in \mathbb{R}^n. \tag{1}$$

Let also λ_A^* and μ_A^* denote the smallest and the greatest eigenvalue, respectively, of the symmetric positive semidefinite operator A^*A. It is well known that

$$\|A\| := \max_{x \in S^{n-1}} \|Ax\| = \sqrt{\mu_A^*} \quad \text{and} \quad \min_{x \in S^{n-1}} \|Ax\| = b_A. \tag{2}$$

Sometimes we set, for brevity, $a_A := \|A\| = \sqrt{\mu_A^*}$ and $b_A := 1/\|A^{-1}\|$ if A is invertible. Since $\|A^{-1}\|$ equals the square root of the greatest eigenvalue of $(A^{-1})^*A^{-1} = (AA^*)^{-1}$, it follows that $b_A = \sqrt{\lambda_A^*}$.

An instructive example covers the class of symmetric matrices.

Example 1.1. Let $A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a symmetric operator, i.e. $\langle Ax, y\rangle = \langle x, Ay\rangle$, for all $x, y \in \mathbb{R}^n$ or, equivalently, $A = A^*$. Note that A has real eigenvalues in this case, say $\lambda_1, \dots, \lambda_n$, and therefore $\lambda_A = \min\{\lambda_1, \dots, \lambda_n\}$ as well as $\mu_A = \max\{\lambda_1, \dots, \lambda_n\}$. Thus, the eigenvalues of $A^*A = A^2$ are $\lambda_1^2, \dots, \lambda_n^2$

and therefore $\lambda^*_A = \min\{\lambda_1^2, \lambda_2^2, \dots, \lambda_n^2\}$ as well as $\mu^*_A = \max\{\lambda_1^2, \lambda_2^2, \dots, \lambda_n^2\}$. Consequently

$$\|A\| = \sqrt{\mu^*_A} = \max\{|\lambda_1|, |\lambda_2|, \dots, |\lambda_n|\} \quad \text{and} \\ \min_{x\in S^{n-1}} \|Ax\| = \sqrt{\lambda^*_A} = b_A = \min\{|\lambda_1|, |\lambda_2|, \dots, |\lambda_n|\}. \tag{3}$$

If $\lambda_1 \le \lambda_2 \le \dots \le \lambda_n$, then obviously $\lambda_A = \lambda_1$ and $\mu_A = \lambda_n$.

- If $\lambda_1 \le \lambda_2 \le \dots \le \lambda_n \le 0$, then $\lambda_1^2 \ge \lambda_2^2 \ge \dots \ge \lambda_n^2$. Therefore $\lambda^*_A = \lambda_n^2$ as well as $\mu^*_A = \lambda_1^2$. This shows that $\|A\| = -\lambda_1$ and $\min_{x\in S^{n-1}} \|Ax\| = -\lambda_n$ in this case.
- If $0 \le \lambda_1 \le \lambda_2 \le \dots \le \lambda_n$, then $\lambda_1^2 \le \lambda_2^2 \le \dots \le \lambda_n^2$. Consequently $\lambda^*_A = \lambda_1^2$ and $\mu^*_A = \lambda_n^2$. In this case we have $\|A\| = \lambda_n$ and $\min_{x\in S^{n-1}} \|Ax\| = \lambda_1$.

Another example, which is more suitable for our current needs, concerns the linear combinations of the identity operator and the skew-symmetric isometries.

Remark 1.2. Let $B : \mathbb{R}^{2m} \longrightarrow \mathbb{R}^{2m}$ a linear skew-symmetric isometry, i.e.

1. $\langle Bx, y\rangle + \langle x, By\rangle = 0,\ \forall x, y \in \mathbb{R}^{2m}$, or, equivalently $\langle Bx, x\rangle = 0,\ \forall x \in \mathbb{R}^{2m}$;
2. $\langle Bx, By\rangle = \langle x, y\rangle,\ \forall x, y \in \mathbb{R}^{2m}$.

The matrix representation of such an operator, with respect to the standard basis of $\mathbb{R}^{2m}$, is skew-symmetric and orthogonal and these properties characterize completely the class of skew-symmetric operators. Moreover, every linear skew-symmetric isometry is a complex structure on $\mathbb{R}^{2m}$. For instance, if $P \in O(m)$ is an orthogonal matrix, then

$$B := \begin{pmatrix} 0 & -P \\ P^T & 0 \end{pmatrix} \in \mathbb{R}^{2m\times 2m}$$

is orthogonal and skew symmetric.

Example 1.3. If $p, q \in \mathbb{R}$ and $A := pI_{2m} + qB$, where $B : \mathbb{R}^{2m} \longrightarrow \mathbb{R}^{2m}$ is a linear skew-symmetric isometry, then $(A + A^*)/2 = pI_{2m}$ and $A^*A = (p^2 + q^2)I_{2m}$, whence $\lambda_A = \mu_A = p$ and $\lambda^*_A = \mu^*_A = p^2+q^2$. Thus $a_A = b_A = \sqrt{p^2 + q^2}$.

In connection to an arbitrary operator T (not necesarily on a finite dimensional space) we define the following quantities:

$$i_T := \inf \left\{ \left\langle \frac{Tx - Ty}{\|Tx - Ty\|}, \frac{x - y}{\|x - y\|} \right\rangle \,\middle|\, (x, y) \in (D \times D) \setminus \ker(T) \right\},$$

$$s_T := \sup \left\{ \left\langle \frac{Tx - Ty}{\|Tx - Ty\|}, \frac{x - y}{\|x - y\|} \right\rangle \,\middle|\, (x, y) \in (D \times D) \setminus \ker(T) \right\},$$

$$i'_T := \inf \left\{ \frac{\|Tx - Ty\|}{\|x - y\|} \,\middle|\, (x, y) \in (D \times D) \setminus \Delta_D \right\},$$

where $\Delta_D := \{(x,x) \mid x \in D\}$. If $s_T < 1$, then we set

$$\eta_T := \frac{i_T i'_T - 1}{\sqrt{1 - s_T^2}}\,.$$

Example 1.4. If $p, q \in \mathbb{R}$ and $A := pI_{2m} + qB$, where $B : \mathbb{R}^{2m} \longrightarrow \mathbb{R}^{2m}$ is a linear skew-symmetric isometry, then

1. $\langle Ax, Ay\rangle = (p^2+q^2)\langle x,y\rangle$, $\forall x,y \in \mathbb{R}^{2m}$.
2. $\|Ax\| = \sqrt{p^2+q^2}\|x\|$, $\forall x \in \mathbb{R}^{2m}$. Thus $i'_A = \sqrt{p^2+q^2}$.
3. If $p = \cos\theta$ and $q = \sin\theta$ for some arbitrary $\theta \in \mathbb{R}$, then A is an isometry.
4. $\langle Ax, x\rangle = p\|x\|^2$, $\forall x \in \mathbb{R}^{2m}$. Thus $i_A = s_A = \frac{p}{\sqrt{p^2+q^2}}$.
5. $\eta_A = \frac{(p-1)\sqrt{p^2+q^2}}{|q|}$.

Remark 1.5.

1. An operator T is i_T-monotone, whenever $-1 < i_T < 1$.
2. An operator T is Minty-Browder monotone if and only if $i_T \geq 0$.
3. If the operator T is η-monotone, then $i_T \geq \eta$.
4. If $A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is an isometry, then $i_A = \lambda_A$, $s_A = \mu_A$ and $i'_A = 1$.

We close this subsection by recalling one of the main results proved in [9] which involve the above associated values i_T, s_T, i'_T, η_T to an arbitrary operator T and will be used within the section of proofs.

Theorem 1.6 (see [9, Theorem 1.1]). *Let H be a real Hilbert space, let D be a nonempty open convex subset of H, let $T : D \to H$ be a Minty-Browder monotone operator, and let $\hat{T} := T - i_D$, where $i_D : D \hookrightarrow H$ is the inclusion mapping. If $0 < i_T$, $s_T < 1$, and $-1 < \eta_T < 0$, then $\hat{T}$ is η_T-monotone. If, in addition, $i'_T s_T < 1$, then $\hat{T}$ is not Minty-Browder monotone.*

In terms of the associated values i_T, s_T, i'_T, Theorem 1.6 can be stated in the following way:

If $0 < i_T$, $s_T < 1$, and $1 - \sqrt{1 - s_T^2} < i_T i'_T < 1$, then $i_{\hat{T}} \geq \frac{i_T i'_T - 1}{\sqrt{1-s_T^2}}$. If $i'_T s_T < 1$, then $i_{\hat{T}} < 0$.

1.2. Contents of the paper

Clearly, i_T represents the greatest possible value of η for which the operator T remains η-monotone. This property is emphasized by Remark 1.5(3). Having this in mind, it suffices to provide conditions ensuring that $-1 < i_T < 0$, in order to obtain η-monotone operators which are not Minty-Browder monotone.

Theorem 1.7. *Let $D \subseteq \mathbb{R}^n$ be a convex open set, let $f : D \longrightarrow \mathbb{R}$ be a C^2-smooth convex function and let $A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a linear transformation. If* *the following inequalities are satisfied*

$$\sigma_f < b_A + \lambda_A \tag{4}$$

$$\inf_{z \in D} \|h_f(z)\| < -\mu_A, \tag{5}$$

then one has

$$-1 < \frac{\lambda_A}{b_A - \sigma_f} \le i_{\nabla f + A} < 0, \tag{6}$$

i.e. $\nabla f + A$ *is* $\frac{\lambda_A}{b_A - \sigma_f}$*-monotone but not Minty-Browder monotone.*

Example 1.8. Let $p, q \in \mathbb{R}$, $p < 0$ and $A := pI_{2m} + qB$, where $B : \mathbb{R}^{2m} \longrightarrow \mathbb{R}^{2m}$ is a linear skew-symmetric isometry. If $0 < \alpha < \min\left\{\sqrt{p^2+q^2}+p,\ -p\right\} = \min\{b_A + \lambda_A, -\mu_A\}$, then $-1 < i_T < 0$, where $T = \alpha I_{2m} + A = \nabla f + A$ and $f : \mathbb{R}^{2m} \to \mathbb{R}$ is the convex function defined by $f(x) := \frac{\alpha}{2}\|x\|^2$. Therefore, T is h-monotone, but not Minty-Browder monotone. The inequalities $-1 < i_T < 0$ follow by means of Theorem 1.7 taking into account that $H_f(x) = \alpha I_{2m}$ and $\sigma_f = \alpha$.

Theorem 1.9. *Let* $D \subseteq \mathbb{R}^n$ *be a convex open set, let* $f : D \longrightarrow \mathbb{R}$ *be a* C^2*-smooth convex function and let* $A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *be a linear transformation. If* $\lambda_A > 0$*, then*

$$i_{\nabla f + A} \ge \frac{\lambda_A}{a_A + \sigma_f} \tag{7}$$

and

$$i'_{\nabla f + A} \le a_A + \sigma_f. \tag{8}$$

If, in addition, $\sigma_f < b_A$*, then*

$$s_{\nabla f + A} \le \frac{\mu_A + \sigma_f}{b_A - \sigma_f} \tag{9}$$

and

$$i'_{\nabla f + A} \ge b_A - \sigma_f. \tag{10}$$

Remark 1.10. The upper estimate $\frac{\mu_A + \sigma_f}{b_A - \sigma_f}$ for $s_{\nabla f + A}$ is completely useless, unless

$$\frac{\mu_A + \sigma_f}{b_A - \sigma_f} < 1, \tag{11}$$

as the inequality $s_T \le 1$ is obvious for every operator T. If $B : \mathbb{R}^{2m} \longrightarrow \mathbb{R}^{2m}$ is a skew-symmetric isometry and $A = pI_{2m} + qB$, $p > 0$, then the inequality (11) is satisfied if

$$\sigma_f < \frac{-p + \sqrt{p^2 + q^2}}{2}. \tag{12}$$

Example 1.11. Let $D \subseteq \mathbb{R}^n$ be a convex open set and let $f : D \longrightarrow \mathbb{R}$ be the convex function

$$f(x_1, \ldots, x_n) = \frac{1}{2}\sum_{i=1}^{n} \alpha_i x_i^2,$$

where $0 < \alpha_1 \leq \cdots \leq \alpha_n$. If $A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is the linear transformation whose matrix representation with respect to the standard basis of $\mathbb{R}^n$ is

$$M^T \cdot \operatorname{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\} \cdot M,$$

$0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ for some orthogonal matrix $M \in O(n)$, then

$$i_{\nabla f + A} \geq \frac{\lambda_1}{\lambda_n + \alpha_n} \tag{13}$$

and

$$i'_{\nabla f + A} \leq \lambda_n + \alpha_n. \tag{14}$$

If, in addition, $\alpha_n < \lambda_1$, then

$$i'_{\nabla f + A} \geq \lambda_1 - \alpha_n. \tag{15}$$

Indeed, by using the obvious fact that $H_f(x) = \operatorname{diag}\{\alpha_1, \ldots, \alpha_n\}$ for every $x \in D$, one can deduce that $\sigma_f = \alpha_n$. Furtheron, we only need to combine Example 1.1 with Theorem 1.9. Note that the estimation

$$s_{\nabla f + A} \leq \frac{\lambda_n + \alpha_n}{\lambda_1 - \alpha_n}, \tag{16}$$

produced by (9) in this particular case, is completely useless since

$$\frac{\lambda_n + \alpha_n}{\lambda_1 - \alpha_n} > 1.$$

Theorem 1.12. *Let $D \subseteq \mathbb{R}^n$ be a convex open set, let $f : D \longrightarrow \mathbb{R}$ be a C^2-smooth convex function and let $A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a linear transformation. If $\lambda_A > 0$, $\mu_A < \min\left\{b_A, \frac{b_A}{a_A}\right\}$, and*

$$a_A - \lambda_A \mu_A < \sqrt{b_A^2 - \mu_A^2}, \tag{17}$$

then for sufficiently small σ_f one has $-1 < \eta_{\nabla f + A} < 0$ and the operator $\hat{T} = \nabla f + A - i_D$ is $\eta_{\nabla f + A}$-monotone, but not Minty-Browder monotone, i.e. $\eta_{\nabla f + A} \leq i_{\hat{T}} < 0$.

1.3. A class of convex functions with $s_{\nabla f} = 1$

The proof of Theorem 1.12 relies on [9, Theorem 1.1] which states that the operator $\hat{T} = T - id$ constructed out of a Minty-Browder monotone operator T on a Hilbert space and the identity map is η_T-monotone, but not Minty-Browder monotone, whenever $0 < i_T, s_T < 1$ and $\eta_T < 0$. As in the previous paper [9] we could have considered the gradients ∇f of convex functions f to play the role of the Minty-Browder monotone operators in Theorem 1.6, for more concrete examples. However the associated upper-bound $s_{\nabla f}$ of such an operator was proved to be 1, at least in the bidimensional case and for strictly convex functions with positive definite Hessian matrices. In order to fulfill all requirements of Theorem 1.6, the gradients of such functions were perturbed in [9] by some suitable rotations. This argument is still a good enough reason to perturb the gradients of the convex functions by certain isometries, in higher dimensional context, as $s_{\nabla f}$ keeps being 1 on a still quite large class of convex functions as we shall see in the following statement.

Theorem 1.13. *Let $D \subseteq \mathbb{R}^n$ be a convex open set and $f : D \longrightarrow \mathbb{R}$ be a convex function with precisely one global minimum point. If the Hessian matrix of f is everywhere positive definite, then $s_{\nabla f} = 1$.*

Proof. Taking into account the positive definiteness of the Hessian matrix $H_f(x)$ it follows that $H_f(x)$ is nondegenerated at every $x \in D$ and in particular it remains non-degenerate at its critical points and, in particular, at its (unique) minimum point p_0. Therefore f is a Morse function and p_0 is a critical point of index zero. Thus, up to some local diffeomorphism, f behaves locally around p_0 like the function $x_1^2 + \cdots + x_n^2$ ([11, p. 44], [13, p. 199]) and the inverse images $f^{-1}(q)$ are, up to diffeomorphisms, $(n-1)$-spheres for $q > f(p_0)$ close enough to $f(p_0)$. Observe that the function $\delta : f^{-1}(q) \longrightarrow \mathbb{R}$, $\delta(x) := \|x - p_0\|^2$ has two critical points $p_1, p_2 \in f^{-1}(q)$ at least. Thus

$$\begin{aligned}(d\delta)_{p_1} = 0 \; &\Longleftrightarrow (d\delta)_{p_1}(v) = 0, \; \forall v \in T_{p_1}(f^{-1}(q)) \\ &\Longleftrightarrow v \cdot (p_1 - p_0) = 0, \; \forall v \in T_{p_1}(f^{-1}(q)),\end{aligned}$$

i.e. $p_1 - p_0 \perp T_{p_1}(f^{-1}(q))$. Consequently the vectors $p_1 - p_0$ and $(\nabla f)(p_1)$ are both outer normal vectors to the hypersurface $f^{-1}(q)$ i.e. $(\nabla f)(p_1) = \lambda(p_1 - p_0)$ for some $\lambda > 0$.

Taking into account the obvious fact $(\nabla f)(p_0) = 0$, we obtain

$$\begin{aligned}\left\langle \frac{(\nabla f)(p_1) - (\nabla f)(p_0)}{\|(\nabla f)(p_1) - (\nabla f)(p_0)\|}, \frac{p_1 - p_0}{\|p_1 - p_0\|} \right\rangle &= \left\langle \frac{(\nabla f)(p_1)}{\|(\nabla f)(p_1)\|}, \frac{p_1 - p_0}{\|p_1 - p_0\|} \right\rangle \\ &= \left\langle \frac{\lambda(p_1 - p_0)}{\|\lambda(p_1 - p_0)\|}, \frac{p_1 - p_0}{\|p_1 - p_0\|} \right\rangle = 1. \qquad \square\end{aligned}$$

Remark 1.14. In the proof of Theorem 1.13 we do not realy use the spherical shape of the inverse image $f^{-1}(q)$, but we only used its quality to be a compact differential hypersurface in $\mathbb{R}^n$. The differential structure on the inverse image $f^{-1}(q)$ is emphasized by the inverse image theorem [14, p. 409], whenever q is a regular value of f. The closedness of the inverse image $f^{-1}(q)$ is due to the continuity of f and its boundedness can be easily proved for values $q > f(p_0)$, sufficiently close to $f(p_0)$. Note that the set of regular values of f is dense in its range, as the set of its critical values has measure zero due to the Sard theorem [1, p. 34]. Thus, f has regular values $q > f(p)$ arbitrarily close to $f(p_0)$. Let us finaly observe that the inverse image $f^{-1}(q)$ is the boundary of the sublevel set $f^{-1}(-\infty, q)$, whenever $q \in \mathrm{Im}(f)$ and the sublevel set $f^{-1}(-\infty, q)$ is open, convex [3, p. 75] and bounded for $q > f(p_0)$ sufficiently close to $f(p_0)$.

2. The proofs

Proof of Theorem 1.7. Note first that (5) implies that $\mu_A < 0$, whence $\lambda_A < 0$. On the other hand, by (4) we have

$$b_A - \sigma_f > -\lambda_A > 0.$$

This guarantees the validity of the first inequality in (6). In order to prove the second inequality in (6), let $x, y \in D$ be such that $Tx \neq Ty$. Then we have

$$\begin{aligned}\left\langle \frac{Tx - Ty}{\|Tx - Ty\|}, \frac{x - y}{\|x - y\|} \right\rangle &= \left\langle \frac{\nabla f(x) - \nabla f(y) + A(x - y)}{\|Tx - Ty\|}, \frac{x - y}{\|x - y\|} \right\rangle \\ &= \frac{\langle \nabla f(x) - \nabla f(y),\, x - y\rangle}{\|Tx - Ty\|\,\|x - y\|} + \frac{\langle A(x - y),\, x - y\rangle}{\|Tx - Ty\|\,\|x - y\|}.\end{aligned}$$

Taking into account that $\nabla f : D \to \mathbb{R}^n$ is Minty-Browder monotone, i.e., that

$$\langle \nabla f(x) - \nabla f(y),\, x - y\rangle \geq 0,$$

by virtue of (1) we deduce that

$$\left\langle \frac{Tx - Ty}{\|Tx - Ty\|}, \frac{x - y}{\|x - y\|} \right\rangle \geq \lambda_A \frac{\|x - y\|}{\|Tx - Ty\|}. \tag{18}$$

On the other hand, we have

$$\begin{aligned}\|Tx - Ty\| &= \|\nabla f(x) - \nabla f(y) + A(x - y)\| \\ &\geq \|A(x - y)\| - \|\nabla f(x) - \nabla f(y)\|,\end{aligned}$$

whence

$$\frac{\|Tx - Ty\|}{\|x - y\|} \geq \left\| A\left(\frac{x - y}{\|x - y\|}\right)\right\| - \frac{\|\nabla f(x) - \nabla f(y)\|}{\|x - y\|}.$$

Since

$$\left\| A\left(\frac{x - y}{\|x - y\|}\right)\right\| \geq b_A \quad \text{and} \quad \frac{\|\nabla f(x) - \nabla f(y)\|}{\|x - y\|} \leq \sigma_f,$$

we get

$$\frac{\|Tx - Ty\|}{\|x-y\|} \geq b_A - \sigma_f,$$

i.e.,

$$\frac{\|x-y\|}{\|Tx - Ty\|} \leq \frac{1}{b_A - \sigma_f}. \tag{19}$$

By (18) and (19) we conclude that

$$\left\langle \frac{Tx - Ty}{\|Tx-Ty\|}, \frac{x-y}{\|x-y\|} \right\rangle \geq \frac{\lambda_A}{b_A - \sigma_f} \quad \text{for all } (x,y) \in (D \times D) \setminus \ker(T).$$

This inequality shows that the second inequality in (6) holds.

It remains to prove that $i_{\nabla f + A} < 0$. To this end, we notice that by (5) it follows that there is a point $z_0 \in D$ as well as a ball $D_0 \subseteq D$ centered at z_0 such that $\sup_{z \in D_0} \|h_f(z)\| < -\mu_A$. Then, due to (1), for all $x, y \in D$ with $x \neq y$ we have

$$\begin{aligned}
\langle Tx - Ty, x-y \rangle &= \langle \nabla f(x) - \nabla f(y), x - y\rangle + \langle A(x-y), x-y\rangle \\
&\leq \|\nabla f(x) - \nabla f(y)\| \, \|x-y\| + \mu_A \|x-y\|^2 \\
&\leq \|x-y\|^2 \left(\sup_{z \in D_0} \|h_f(z)\| + \mu_A \right) \\
&< 0.
\end{aligned}$$

Therefore $i_{\nabla f + A} < 0$. □

Example 2.1. Let $p, q \in \mathbb{R}$ and $A := pI_{2m} + qB$, where $B : \mathbb{R}^{2m} \longrightarrow \mathbb{R}^{2m}$ is a linear skew-symmetric isometry. Observe that the function $f : I^{2m} \longrightarrow \mathbb{R}$ given by

$$f(x_1, \ldots, x_{2m}) = \frac{1}{2k} \sum_{i=1}^{2m} \alpha_i x_i^{2k},$$

is convex whenever $\alpha_1, \ldots, \alpha_{2m} > 0$, where $I = (-1, 1)$ and $k \geq 1$. Moreover, we have

$$\nabla f(x) = \left(\alpha_1 x_1^{2k-1}, \ldots, \alpha_{2m} x_{2m}^{2k-1} \right)$$

and $H_f(x_1, \ldots, x_{2m}) = (2k-1)\operatorname{diag}\left\{ \alpha_1 x_1^{2k-2}, \ldots, \alpha_{2m} x_{2m}^{2k-2} \right\}$ for all $(x_1, \ldots, x_{2m}) \in I^{2m}$ and therefore

$$\begin{aligned}
\|h_f(x_1, \ldots, x_{2m})\| &= (2k-1) \cdot \max\left\{ \alpha_1 x_1^{2k-2}, \ldots, \alpha_{2m} x_{2m}^{2k-2} \right\} \\
&\leq (2k-1) \cdot \max\left\{ \alpha_1, \ldots, \alpha_{2m} \right\}.
\end{aligned}$$

According to Theorem 1.7, if

$$\max\{\alpha_1, \ldots, \alpha_{2m}\} < \frac{\min\left\{ -p, \sqrt{p^2 + q^2} + p \right\}}{2k-1},$$

then the operator $T : \mathbb{R}^{2m} \to \mathbb{R}^{2m}$ given by

$$\begin{aligned} T(x_1, \dots, x_{2m}) &:= (\nabla f)(x_1, \dots, x_{2m}) + A(x_1, \dots, x_{2m}) \\ &= \left(\alpha_1 x_1^{2k-1}, \dots, \alpha_n x_n^{2k-1}\right) + A(x_1, \dots, x_{2m}), \end{aligned}$$

satisfies $-1 < i_{_T} < 0$. Therefore, T is h-monotone, but not Minty-Browder monotone.

Proof of Theorem 1.9. Let $x, y \in D$ be such that $Tx \neq Ty$. By (18) we have

$$\begin{aligned} \left\langle \frac{Tx - Ty}{\|Tx - Ty\|}, \frac{x - y}{\|x - y\|} \right\rangle &\geq \lambda_A \frac{\|x - y\|}{\|Tx - Ty\|} \\ &\geq \lambda_A \frac{\|x - y\|}{\|\nabla f(x) - \nabla f(y)\| + \|A(x - y)\|}. \end{aligned}$$

Since $\|\nabla f(x) - \nabla f(y)\| \leq \sigma_{_f} \|x - y\|$ and $\|A(x - y)\| \leq a_{_A} \|x - y\|$, we deduce that

$$\left\langle \frac{Tx - Ty}{\|Tx - Ty\|}, \frac{x - y}{\|x - y\|} \right\rangle \geq \frac{\lambda_A}{a_A + \sigma_{_f}}$$

for all $(x, y) \in (D \times D) \setminus \ker(T)$. Therefore, (7) holds.

For all $x, y \in D$ with $x \neq y$ we have

$$\frac{\|Tx - Ty\|}{\|x - y\|} \leq \frac{\|\nabla f(x) - \nabla f(y)\|}{\|x - y\|} + \frac{\|A(x - y)\|}{\|x - y\|} \leq \sigma_{_f} + a_A.$$

This proves the validity of (8).

Note that the assumption $\lambda_A > 0$ together with (1) ensures that A is injective, hence invertible. Since

$$\|x\| = \|A^{-1}(Ax)\| \leq \|A^{-1}\| \, \|Ax\|,$$

it follows that

$$\|Ax\| \geq b_{_A} \|x\| \quad \text{for all } x \in \mathbb{R}^n. \tag{20}$$

In order to establish (9) and (10), let $x, y \in D$ such that $x \neq y$. Taking into account (20) we get

$$\begin{aligned} \|Tx - Ty\| &\geq \|A(x - y)\| - \|\nabla f(x) - \nabla f(y)\| \\ &\geq b_{_A} \|x - y\| - \sigma_{_f} \|x - y\| = (b_A - \sigma_{_f}) \|x - y\|. \end{aligned}$$

From this inequality we deduce that (10) holds and that whenever $Tx \neq Ty$ one has

$$\left\langle \frac{Tx - Ty}{\|Tx - Ty\|}, \frac{x - y}{\|x - y\|} \right\rangle \leq \frac{\langle \nabla f(x) - \nabla f(y), x - y \rangle + \langle A(x - y), x - y \rangle}{(b_A - \sigma_{_f}) \|x - y\|^2}.$$

Using (1) and

$$\begin{aligned}\langle \nabla f(x) - \nabla f(y),\, x - y\rangle &\leq \|\nabla f(x) - \nabla f(y)\|\, \|x - y\| \\ &\leq \sigma_f \|x - y\|^2\end{aligned}$$

we obtain

$$\left\langle \frac{Tx - Ty}{\|Tx - Ty\|}, \frac{x - y}{\|x - y\|} \right\rangle \leq \frac{\mu_A + \sigma_f}{b_A - \sigma_f}$$

for all $(x, y) \in (D \times D) \setminus \ker(T)$. Therefore, (9) holds. $\square$

Example 2.2. Let $f : I^n \longrightarrow \mathbb{R}$ be the convex function given by

$$f(x_1, \ldots, x_n) = \frac{1}{2k} \sum_{i=1}^{n} \alpha_i x_i^{2k},$$

where $I = (-1, 1)$, $0 < \alpha_1 \leq \cdots \leq \alpha_n$ and $k \geq 1$. If $0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ and $A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is the linear transformation whose matrix representation with respect to the standard basis of $\mathbb{R}^n$ is $M^T \cdot \mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\} \cdot M$, for some orthogonal matrix $M \in O(n)$, then

$$i_{\nabla f + A} \geq \frac{\lambda_1}{\lambda_n + (2k-1)\alpha_n} \tag{21}$$

and

$$i'_{\nabla f + A} \leq \lambda_n + (2k-1)\alpha_n. \tag{22}$$

If, in addition, $(2k-1)\alpha_n < \lambda_1$, then

$$i'_{\nabla f + A} \geq \lambda_1 - (2k-1)\alpha_n. \tag{23}$$

Indeed, by using the obvious fact that

$$H_f(x_1, \ldots, x_n) = (2k-1)\,\mathrm{diag}\left\{\alpha_1 x_1^{2k-2}, \ldots, \alpha_n x_n^{2k-2}\right\}$$

one can show that $\sigma_f \leq (2k-1)\alpha_n$. Further on, we only need to combine Example 1.1 with Theorem 1.9. For $k > 1$ the boundedness of the domain I^n of f is essential in order to get $\sigma_f < \infty$. For $k = 1$ the boundedness of this domain plays no role, as the Hessian matrix of f is the constant matrix $\mathrm{diag}\{\alpha_1, \ldots, \alpha_n\}$ in this case. This situation was emphasized by Example 1.11. Note that the estimation

$$s_{\nabla f + A} \leq \frac{\lambda_n + (2k-1)\alpha_n}{\lambda_1 - (2k-1)\alpha_n}, \tag{24}$$

produced by (9) in this particular case, is completely useless since

$$\frac{\lambda_n + (2k-1)\alpha_n}{\lambda_1 - (2k-1)\alpha_n} > 1.$$

Proof of Theorem 1.12. Note first that the assumption $\lambda_A > 0$ and the estimation (7) ensures that $i_{\nabla f+A} > 0$. On the other hand, due to (9), one has $s_{\nabla f+A} < 1$ as soon as

$$\sigma_f < \frac{b_A - \mu_A}{2}. \tag{25}$$

Since $i_{\nabla f+A} \leq s_{\nabla f+A}$, it follows that the validity of $i'_{\nabla f+A} s_{\nabla f+A} < 1$ implies that $\eta_{\nabla f+A} < 0$. Let us find conditions ensuring the validity of $i'_{\nabla f+A} s_{\nabla f+A} < 1$. By virtue of (8) and (9), this inequality holds as soon as

$$(a_A + \sigma_f)\,\frac{\mu_A + \sigma_f}{b_A - \sigma_f} < 1. \tag{26}$$

If σ_f satisfies (25), then $\sigma_f < b_A$, so (26) is equivalent to

$$\sigma_f^2 + (a_A + \mu_A + 1)\sigma_f + a_A\mu_A - b_A < 0. \tag{27}$$

The second degree equation $s^2 + (a_A + \mu_A + 1)s + a_A\mu_A - b_A = 0$ has the discriminant $\Delta = (a_A - \mu_A)^2 + 2(a_A + \mu_A) + 4b_A + 1 > 0$, so it possesses two real roots s_1 and s_2. Since $s_1 + s_2 = -(a_A + \mu_A + 1) < 0$ and $s_1 s_2 = a_A\mu_A - b_A < 0$ (because $\mu_A < b_A/a_A$), it follows that $s_1 < 0 < s_2$. Therefore, (27) holds if

$$\sigma_f < \frac{\sqrt{(a_A - \mu_A)^2 + 2(a_A + \mu_A) + 4b_A + 1} - (a_A + \mu_A + 1)}{2}. \tag{28}$$

In other words, the conditions $i'_{\nabla f+A} s_{\nabla f+A} < 1$ and $\eta_{\nabla f+A} < 0$ are both fulfilled as soon as σ_f satisfies (28).

Finally, we have to find conditions under which

$$-1 < \eta_{\nabla f+A} \quad \Leftrightarrow \quad \frac{1 - i_{\nabla f+A} i'_{\nabla f+A}}{\sqrt{1 - s^2_{\nabla f+A}}} < 1. \tag{29}$$

Taking into account the estimations from Theorem 1.9, we have

$$\begin{aligned}
\frac{1 - i_{\nabla f+A} i'_{\nabla f+A}}{\sqrt{1 - s^2_{\nabla f+A}}} &\leq \frac{1 - \dfrac{\lambda_A}{a_A + \sigma_f}\,(b_A - \sigma_f)}{\sqrt{1 - \left(\dfrac{\mu_A + \sigma_f}{b_A - \sigma_f}\right)^2}} \\
&= \frac{b_A - \sigma_f}{\sqrt{(b_A + \mu_A)(b_A - \mu_A - 2\sigma_f)}} \left(1 - \frac{\lambda_A(b_A - \sigma_f)}{a_A + \sigma_f}\right).
\end{aligned}$$

If σ_f satisfies (25), i.e., $b_A - \sigma_f > \mu_A + \sigma_f$, then we get

$$\begin{aligned}
\frac{1 - i_{\nabla f + A} i'_{\nabla f + A}}{\sqrt{1 - s^2_{\nabla f + A}}} &< \frac{b_A - \sigma_f}{\sqrt{(b_A + \mu_A)(b_A - \mu_A - 2\sigma_f)}} \left(1 - \frac{\lambda_A(\mu_A + \sigma_f)}{a_A + \sigma_f}\right) \\
&< \frac{b_A - \sigma_f}{\sqrt{(b_A + \mu_A)(b_A - \mu_A - 2\sigma_f)}} \left(1 - \frac{\lambda_A \mu_A}{a_A + \sigma_f}\right) \\
&= \frac{b_A - \sigma_f}{a_A + \sigma_f} \cdot \frac{a_A - \lambda_A \mu_A + \sigma_f}{\sqrt{(b_A + \mu_A)(b_A - \mu_A - 2\sigma_f)}} \\
&< \frac{a_A - \lambda_A \mu_A + \sigma_f}{\sqrt{(b_A + \mu_A)(b_A - \mu_A - 2\sigma_f)}} .
\end{aligned}$$

Therefore, (29) holds whenever

$$a_A - \lambda_A \mu_A + \sigma_f < \sqrt{b_A^2 - \mu_A^2 - 2\sigma_f (b_A + \mu_A)}.$$

Since $a_A \geq b_A$ and $\lambda_A \leq \mu_A < 1$, we have

$$a_A - \lambda_A \mu_A > b_A - \mu_A > 0,$$

hence the above inequality is equivalent to

$$\sigma_f^2 + 2(a_A - \lambda_A \mu_A + b_A + \mu_A)\sigma_f + (a_A - \lambda_A \mu_A)^2 - (b_A^2 - \mu_A^2) < 0. \tag{30}$$

The second degree equation

$$s^2 + 2(a_A - \lambda_A \mu_A + b_A + \mu_A)s + (a_A - \lambda_A \mu_A)^2 - (b_A^2 - \mu_A^2) = 0$$

has the discriminant

$$\Delta = 8(b_A + \mu_A)(a_A + b_A - \lambda_A \mu_A) > 0,$$

so it possesses two real roots s_1 and s_2. Taking into consideration that $s_1 + s_2 < 0$ and $s_1 s_2 = (a_A - \lambda_A \mu_A)^2 - (b_A^2 - \mu_A^2) < 0$ (because of our assumption (17)), it follows that $s_1 < 0 < s_2$. Hence (30) is satisfied if

$$\sigma_f < \sqrt{2(b_A + \mu_A)(a_A + b_A - \lambda_A \mu_A)} - (a_A + b_A + \mu_A - \lambda_A \mu_A). \tag{31}$$

In conclusion, if σ_f satisfies the conditions (25), (28), and (31), then all hypotheses of Theorem 1.6 are fulfilled. By applying Theorem 1.6, we obtain the conclusion of our theorem. $\square$

Remark 2.3. The hypotheses towards the linear perturbation of the gradient in Theorem 1.12 are fulfilled by any linear transformation $A : \mathbb{R}^{2m} \longrightarrow \mathbb{R}^{2m}$ of type $A = pI_{2m} + qB$ with $p, q \in \mathbb{R}$ satisfying

$$0 < p < 1, \qquad 0 < |q| < 1, \qquad p^2 + |q| \geq 1, \tag{32}$$

where $B : \mathbb{R}^{2m} \to \mathbb{R}^{2m}$ is a skew-symmetric isomtery.

Indeed, in this case we have $\lambda_A = \mu_A = p$ and $\lambda_A^* = \mu_A^* = p^2 + q^2$, whence $a_A = b_A = \sqrt{p^2 + q^2}$. Condition (17) is equivalent to

$$\sqrt{p^2 + q^2} - p^2 < |q|$$

and this inequality holds because

$$\sqrt{p^2 + q^2} < \sqrt{p^2 + |q|} \leq p^2 + |q|.$$

In such a case the upper-bounds (25), (28), and (31) of σ_f can be concentrated in

$$\sigma_f < \min\left\{ \frac{\sqrt{p^2 + q^2} - p}{2},\ E(p, q), F(p, q) \right\}, \tag{33}$$

where, setting $\bar{p} := \sqrt{p^2 + q^2}$,

$$E(p, q) := \frac{1}{2}\left(\sqrt{(p - \bar{p})^2 + 2(p + \bar{p}) + 4\bar{p} + 1} - (1 + p + \bar{p}) \right) \tag{34}$$

and

$$F(p, q) := \sqrt{2(p + \bar{p})(2\bar{p} - p^2)} - (2\bar{p} + p - p^2). \tag{35}$$

If $p = \cos\theta$ and $q = \sin\theta$ for some $\theta \in \mathbb{R}$, then the upper-bound (33) becomes

$$\sigma_f < \min\left\{ \sin^2\frac{\theta}{2}, \frac{\sqrt{8 + \cos^2\theta} - 2 - \cos\theta}{2}, 2\cos\frac{\theta}{2}\sqrt{1 + \sin^2\theta} - 2\cos^2\frac{\theta}{2} - \sin^2\theta \right\}, \tag{36}$$

as

$$E(\cos\theta, \sin\theta) = \frac{\sqrt{8 + \cos^2\theta} - 2 - \cos\theta}{2}$$

and

$$F(\cos\theta, \sin\theta) = 2\cos\frac{\theta}{2}\sqrt{1 + \sin^2\theta} - 2\cos^2\frac{\theta}{2} - \sin^2\theta.$$

Consequently, by applying Theorem 1.12 we obtain the following:

Proposition 2.4. *Let $D \subseteq \mathbb{R}^n$ be a convex open set, let $f : D \longrightarrow \mathbb{R}$ be a C^2-smooth convex function. If $B : \mathbb{R}^{2m} \to \mathbb{R}^{2m}$ is a skew-symmetric isometry, $\theta \in (-\pi/2, 0) \cup (0, \pi/2)$, and the condition* (36) *is satisfied, then the operator*

$$\hat{T} := \nabla f + (\cos\theta - 1)I_{2m} + (\sin\theta)B = \nabla f - 2\sin^2\frac{\theta}{2} I_{2m} + (\sin\theta)B$$

is η-monotone for some $\eta \in (-1, 0)$, but not Minty-Browder monotone, i.e. $-1 < i_{\hat{T}} < 0$.

3. Global injectivity via η-monotonicity

Proposition 3.1. *Let $D \subseteq \mathbb{R}^n$ be a convex open set, let $f : D \longrightarrow \mathbb{R}$ be a C^2-smooth convex function and let $A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a linear transformation. If the following inequality is satisfied*

$$\sigma_f < \min \{b_A + \lambda_A, -\mu_A\}, \tag{37}$$

then the operator $T := \nabla f + A$ is injective.

Proof. We show first that the Jacobi matrix of T

$$J_T(x) = H_f(x) + A$$

is invertible for all $x \in D$. Suppose the contrary, i.e., that there exists some point $x \in D$ such that $J_T(x)$ is not invertible. Then there exists $h \in S^{n-1}$ such that $J_T(x)h = 0$, i.e.,

$$H_f(x)h = -Ah.$$

From this equality it follows that

$$\langle H_f(x)h, h\rangle = -\langle Ah, h\rangle \geq -\mu_A,$$

whence

$$-\mu_A \leq \langle H_f(x)h, h\rangle \leq \|h_f(x)h\| \leq \|h_f(x)\|.$$

But this inequality contradicts our assumption (37). The obtained contradiction shows that $J_T(x)$ is invertible for all $x \in D$, i.e., T is a local diffeomorphism. On the other hand, according to Theorem 1.7 the operator T is h-monotone. By applying [8, Corollary 4.12], we conclude that T is injective. $\square$

Theorem 3.2. *Let $D \subseteq \mathbb{R}^{2m}$ be a convex open set, let $f : D \longrightarrow \mathbb{R}$ be a C^2-smooth convex function and let $B : \mathbb{R}^{2m} \longrightarrow \mathbb{R}^{2m}$ be a linear skew-symmetric isometry. If $A = pI_{2m} + qB$ with $p, q \in \mathbb{R}$ satisfying*

$$0 < p < 1, \qquad 0 < |q| < 1, \qquad p^2 + |q| \geq 1, \tag{38}$$

and $1 - p$ is not an eigenvalue for any operator $h_f(x) + qB$, as x covers D, then the operator $\hat{T} = \nabla f + A - i_D$ is injective, whenever

$$\sigma_f < \min \left\{ \frac{\sqrt{p^2 + q^2} - p}{2}, \ E(p, q), F(p, q) \right\}, \tag{39}$$

where $E(p, q)$ and $F(p, q)$ are given by (34) *and* (35).

Proof. If σ_f is subject to the relation (39), then the requirements (38) ensure us that $i_{\hat{T}} > -1$ and the h-monotonicity of $\hat{T}$ therefore. On the other hand the Fréchet differential $(d\hat{T})_x$ of $\hat{T}$ at $x \in D$ is $h_f(x) + (p-1)I_{2m} + qB$. Since

$1-p$ is not an eigenvalue for any of the operators $h_f(x)+qB$, it follows that $[h_f(x)+qB](y)\neq(1-p)y$, for every $x\in D$ and every $y\in\mathbb{R}^{2m}\setminus\{0\}$, i.e.

$$\ker(d\hat{T})_x=\ker[h_f(x)+(p-1)I_{2m}+qB]=\{0\}.$$

Consequently, the Fréchet differentials $(d\hat{T})_x$, $x\in D$ are all isomorphisms and the operator $\hat{T}$ itself is a local diffeomorphism. The global injectivity of $\hat{T}$ follows now via [8, Corollary 4.12]. □

Proposition 3.3. *Let $D_1,\ D_2\subseteq\mathbb{R}^m$ be open subsets of $\mathbb{R}^m$ and $f_1:D_1\longrightarrow\mathbb{R}$, $f_2:D_2\longrightarrow\mathbb{R}$ be C^2-smooth functions. Further let $f:D_1\times D_2\longrightarrow\mathbb{R}$, $f(x,y)=f_1(x)+f_2(y)$ and $\hat{T}:=\nabla f+(p-1)I_{2m}+qB$, where $p,q\in\mathbb{R}$ and $B:\mathbb{R}^{2m}\longrightarrow\mathbb{R}^{2m}$ is the skew-symmetric isometry whose matrix representation with respect to the standard basis of $\mathbb{R}^{2m}$ is*

$$[B]=\begin{pmatrix} O_m & -P\\ P^T & O_m\end{pmatrix},$$

with $P\in O(m)$ is an orthogonal matrix. Then the following assertions are true:

1. $H_f(x,y)=\begin{pmatrix} H_{f_1}(x) & O_m\\ O_m & H_{f_2}(y)\end{pmatrix}.$
2. $(d\hat{T})_{(x,y)}=h_f(x,y)+(p-1)I_{2m}+qB$ *and therefore*
$$[(d\hat{T})_{(x,y)}]=\begin{pmatrix} H_{f_1}(x)+(p-1)I_m & -qP\\ qP^T & H_{f_2}(y)+(p-1)I_m\end{pmatrix}.$$
3. $\det(d\hat{T})_{(x,y)}=\det\left([H_{f_1}(x)+(p-1)I_m]\cdot[H_{f_2}(y)+(p-1)I_m]+q^2I_m\right).$
4. $\sigma_f\leq\max\{\sigma_{f_1},\sigma_{f_2}\}.$

Proof. While the first two items are obvious, for the third item one can use the formula for the determiant of a matrix of blocks [7, p. 164], i.e.

$$\det\begin{pmatrix} A & B\\ C & D\end{pmatrix}=\det(AD-BC),\quad A,B,C,D\in\mathbb{R}^{m\times m}.$$

For the last item we only need to observe that

$$\begin{aligned}
&\|h_f(x,y)(z_1,\dots,z_{2m})\|^2\\
&=\|h_{f_1}(x)(z_1,\dots,z_m)\|^2+\|h_{f_2}(y)(z_{m+1},\dots,z_{2m})\|^2\\
&\leq\|h_{f_1}(x)\|^2\cdot\|(z_1,\dots,z_m)\|^2+\|h_{f_2}(y)\|^2\cdot\|(z_{m+1},\dots,z_{2m})\|^2\\
&\leq\ \max\left\{\sigma_{f_1}^2,\sigma_{f_2}^2\right\}\cdot\|(z_1,\dots,z_{2m})\|^2,
\end{aligned}$$

i.e. $\|h_f(x,y)\|^2\leq\left(\max\{\sigma_{f_1},\sigma_{f_2}\}\right)^2$, for all $(x,y)\in D_1\times D_2$. □

The formula of the item *3.* becomes effective when that determinant can be computed. This is the case when, for example, the product $[H_{f_1}(x) + (p-1)I_m] \cdot [H_{f_2}(y) + (p-1)I_m]$ is a diagonal matrix. We provide below two examples when this product is a diagonal matrix.

Example 3.4. Let $D \subseteq \mathbb{R}^m$ be an open convex subset of $\mathbb{R}^m$, let $f : D \longrightarrow \mathbb{R}$ be a C^2-smooth convex function and $g : D \times \mathbb{R}^m \longrightarrow \mathbb{R}$, $g(x,y) = f(x) + \frac{\alpha}{2}\|y\|^2$, $\alpha > 0$. If $B : \mathbb{R}^{2m} \longrightarrow \mathbb{R}^{2m}$ is the skew-symmetric isometry such that

$$[B] = \begin{pmatrix} O_m & -P \\ P^T & O_m \end{pmatrix},$$

with $P \in O(m)$ an orthogonal matrix, then the operator

$$\hat{T} := \nabla g + (p-1)I_{2m} + qB = (\nabla f, \alpha I_m) + (p-1)I_{2m} + qB, \quad p, q \in \mathbb{R} \quad (40)$$

is injective whenever

$$0 < p < 1, \qquad 0 < |q| < 1, \qquad p^2 + |q| \geq 1,$$

and

$$\alpha < 1 - p \quad \text{and} \quad \sigma_f < \min\left\{ \frac{\sqrt{p^2+q^2} - p}{2}, 1-p, E(p,q), F(p,q) \right\},$$

where $E(p,q)$ and $F(p,q)$ are given by (34) and (35).

In order to justify the statement, we first diagonalize the Hessian matrix

$$H_f(x) = (M(x))^T \cdot \operatorname{diag}\{\lambda_1(x), \ldots, \lambda_m(x)\} \cdot M(x)$$

by means of some orthogonal matrix $M(x)$ and nonnegative eigenvalues $\lambda_1(x), \ldots, \lambda_m(x)$ of $H_f(x)$. Observe that

$$\begin{aligned} \left[(d\hat{T})_{(x,y)}\right] &= H_g(x,y) + (p-1)I_{2m} + q[B] \\ &= \begin{pmatrix} H_f(x) & O_m \\ O_m & \alpha I_m \end{pmatrix} + (p-1)\begin{pmatrix} I_m & O_m \\ O_m & I_m \end{pmatrix} + q\begin{pmatrix} O_m & -P \\ P^T & O_m \end{pmatrix} \\ &= \begin{pmatrix} H_f(x) + (p-1)I_m & -qP \\ qP^T & (\alpha+p-1)I_m \end{pmatrix}. \end{aligned}$$

Therefore

$$\begin{aligned} &\det(d\hat{T})_{(x,y)} \\ &= \det\begin{pmatrix} H_f(x) + (p-1)I_m & -qP \\ qP^T & (\alpha+p-1)I_m \end{pmatrix} \\ &= \det\left((\alpha+p-1)\left(H_f(x) + (p-1)I_m\right) + q^2 I_m\right) \\ &= \det\left[\operatorname{diag}\{(\alpha+p-1)(\lambda_1(x)+p-1)\right. \\ &\qquad \left. + q^2, \ldots, (\alpha+p-1)(\lambda_n(x)+p-1) + q^2\}\right] \\ &= \left[(\alpha+p-1)(\lambda_1(x)+p-1)+q^2\right] \cdot \ldots \cdot \left[(\alpha+p-1)(\lambda_n(x)+p-1)+q^2\right]. \end{aligned}$$

Taking into account that $\|h_f(x)\| = \max\{\lambda_1(x), \dots, \lambda_m(x)\}$ and the hypothesis $\sigma_f < 1-p$ it follws that $\lambda_i(x) + 1 - p < 0$ for every $i \in \{1, \dots, m\}$ and every $x \in D$. Consequently the products $(\alpha+p-1)(\lambda_i(x)+p-1)$ are strictly greater than zero for all $i \in \{1, \dots, m\}$ and all $x \in D$, whence $\det(d\hat{T})_{(x,y)} > 0$ for all $(x, y) \in D \times \mathbb{R}^m$. On the other hand the hypothesis

$$\sigma_f < \min\left\{\frac{\sqrt{p^2+q^2}-p}{2}, E(p,q), F(p,q)\right\}$$

ensures the h-monotonicity of $\hat{T}$, via Theorem 1.12 and Remark 2.3. By applying [8, Corollary 4.12], we conclude that T is injective.

Example 3.5. Let $I_1, \dots, I_{2m} \subseteq \mathbb{R}$ be open intervals, let $f_j : I_j \longrightarrow \mathbb{R}$ be C^2-smooth convex functions and let $f : I_1 \times \cdots \times I_{2m} \longrightarrow \mathbb{R}$ be the C^2-smooth convex function given by $f(x_1, \dots, x_{2m}) = f_1(x_1) + \cdots + f_{2m}(x_{2m})$. If

$$\begin{aligned} &\sup(f_j'') \\ <\ &\min\left\{\sin^2\frac{\theta}{2}, \frac{\sqrt{8+\cos^2\theta}-2-\cos\theta}{2}, 2\cos\frac{\theta}{2}\sqrt{1+\sin^2\theta}-2\cos^2\frac{\theta}{2}-\sin^2\theta\right\}, \end{aligned} \tag{41}$$

then the operator $\hat{T} := \nabla f - 2\sin^2\frac{\theta}{2} I_{2m} + (\sin\theta)B$ is injective, where $B : \mathbb{R}^{2m} \longrightarrow \mathbb{R}^{2m}$ is the skew-symmetric isometry of $\mathbb{R}^{2m}$, whose matrix representation whith respect to the standard basis of $\mathbb{R}^{2m}$is

$$\begin{pmatrix} O_m & -P \\ P^T & O_m \end{pmatrix},$$

where $P \in O(m)$ is an orthogonal matrix.

Indeed, the h-monotonicity of such an operator was justified above taking into account that

$$\begin{aligned} &H_f(x_1, \dots, x_{2m}) \\ =\ &\operatorname{diag}\{f_1''(x_1), \dots, f_{2m}''(x_{2m})\} \\ =\ &\begin{pmatrix} \operatorname{diag}\{f_1''(x_1), \dots, f_m''(x_m)\} & O_m \\ O_m & \operatorname{diag}\{f_{m+1}''(x_{m+1}), \dots, f_{2m}''(x_{2m})\} \end{pmatrix}. \end{aligned}$$

and the inequality (36) follows from the requirement (41) as

$$\|h_f(x)\| = \max\{f_1''(x_1), \dots, f_{2m}''(x_{2m})\}, \quad \forall x = (x_1, \dots, x_{2m}) \in I^{2m}.$$

On the other hand $1 - \cos\theta = 2\sin^2\frac{\theta}{2}$ is not an eigenvalue for any of the operators

$$h_f(x) + (\sin\theta)B,$$

as the characteristic polynomial $\mathfrak{c}(\lambda) = \det\left[h_f(x) + (\sin\theta)B - \lambda I_{2m}\right]$ of $h_f(x)+(\sin\theta)B$ evaluated at $2\sin^2\frac{\theta}{2}$ is

$$\begin{aligned}
&\mathfrak{c}\left(2\sin^2\frac{\theta}{2}\right)\\
&= \det\left(h_f(x) + (\sin\theta)B - 2\sin^2\frac{\theta}{2}I_{2m}\right)\\
&= \begin{vmatrix} \operatorname{diag}\{f_1''(x_1),\dots,f_m''(x_m)\}-2\sin^2\frac{\theta}{2}I_m & (-\sin\theta)P \\ (\sin\theta)P^T & \operatorname{diag}\left\{f_{m+1}''(x_{m+1}),\dots,f_{2m}''(x_{2m})\right\}-2\sin^2\frac{\theta}{2}I_m \end{vmatrix}\\
&= \det\Bigg[\left(\operatorname{diag}\{f_1''(x_1),\dots,f_m''(x_m)\}-2\sin^2\frac{\theta}{2}I_m\right)\\
&\qquad\left(\operatorname{diag}\left\{f_{m+1}''(x_{m+1}),\dots,f_{2m}''(x_{2m})\right\}-2\sin^2\frac{\theta}{2}I_m\right)+(\sin^2\theta)I_m\Bigg]\\
&= \left[\left(f_1''(x_1)-2\sin^2\frac{\theta}{2}\right)\left(f_{m+1}''(x_{m+1})-2\sin^2\frac{\theta}{2}\right)+\sin^2\theta\right]\\
&\qquad\cdots\left[\left(f_m''(x_m)-2\sin^2\frac{\theta}{2}\right)\left(f_{2m}''(x_{2m})-2\sin^2\frac{\theta}{2}\right)+\sin^2\theta\right]>0.
\end{aligned}$$

The latest inequality follows since every difference $f_j'' - 2\sin^2\frac{\theta}{2}$ is strictly less than zero and the products $\left(f_j''(x_j) - 2\sin^2\frac{\theta}{2}\right)\left(f_{m+j}''(x_{m+j}) - 2\sin^2\frac{\theta}{2}\right)$ are strictly greater than zero therefore.

In the particular case of the convex functions $f_j : I \longrightarrow \mathbb{R}$, $f_j(x) = \frac{1}{2k}\alpha_j x^{2k}$, where $I = (-1, 1)$, the operator $\hat{T} : I^{2m} \longrightarrow \mathbb{R}^{2m}$ given by

$$\begin{aligned}
&\hat{T}(x_1,\dots,x_{2m})\\
&= \left(\nabla f - 2\sin^2\frac{\theta}{2}I_{2m} + (\sin\theta)B\right)(x_1,\dots,x_{2m})\left(\alpha_1 x_1^{2k-1},\dots,\alpha_{2m}x_{2m}^{2k-1}\right)\\
&= 2\sin^2\frac{\theta}{2}(x_1,\dots,x_{2m})+\sin\theta(-x_{m+1},\dots,-x_{2m},x_1,\dots,x_{2m}),
\end{aligned}$$

is not Minty-Browder monotone but $-1 < i_{\hat{T}} < 0$, and h-monotone therefore, whenever $0 < \alpha_1, \dots, \alpha_n$ and

$$\begin{aligned}
&\max\{\alpha_1,\dots,\alpha_n\} \qquad (42)\\
&\le \frac{\min\left\{\sin^2\frac{\theta}{2}, \frac{\sqrt{8+\cos^2\theta}-2-\cos\theta}{2}, 2\cos\frac{\theta}{2}\sqrt{1+\sin^2\theta}-2\cos^2\frac{\theta}{2}-\sin^2\theta\right\}}{2(2k-1)}.
\end{aligned}$$

Besides the h-monotonicity of $\hat{T}$, mentioned above, the operator $\hat{T}$ is also injective under the requirement (42).

References

[1] T. Aubin: A Course in Differential Geometry, Graduate Studies in Mathematics 27, American Mathematical Society, Providence (2001).

[2] D. Aussel, J.-N. Corvellec, M. Lassonde: Subdifferential characterization of quasiconvexity and convexity, J. Convex Analysis 1 (1994) 195–201.

[3] S. Boyd, L. Vandenberghe: Convex Optimization, Cambridge University Press, Cambridge (2004).

[4] F. E. Browder: Nonlinear elliptic boundary value problems, Bull. Amer. Math. Soc. 69 (1963) 862–874.

[5] R. Correa, A. Jofre, L. Thibault: Characterization of lower semicontinuous convex functions, Proc. Amer. Math. Soc. 116 (1992) 67–72.

[6] J.-B. Hiriart-Urruty, C. Lemaréchal: Fundamentals of Convex Analysis, Grundlehren. Text Editions, Springer, Berlin (2001).

[7] K. Hoffman, R. Kunze: Linear Algebra, 2nd Ed., Prentice-Hall, Englewood Cliffs (1971).

[8] G. Kassay, C. Pintea: On preimages of a class of generalized monotone operators, Nonlinear Anal. 73 (2010) 3537–3545.

[9] D. Marian, I. R. Peter, C. Pintea: A class of generalized monotone operators, J. Math. Anal. Appl. 421 (2015) 1827–1843.

[10] D. Marian, I. R. Peter, C. Pintea: Operations with monotone operators and the monotonicity of the resulting operators, Monatsh. Math. 181 (2016) 143–168.

[11] Y. Matsumoto: An Introduction to Morse Theory, Iwanami Series in Modern Mathematics 208, American Mathematical Society, Providence (2002).

[12] G. J. Minty: On a "monotonicity" method for the solution of non-linear equations in Banach spaces, Proc. Natl. Acad. Sci. USA 50 (1963) 1038–1041.

[13] R. S. Palais, C.-L. Terng: Critical Point Theory and Submanifold Geometry, Lecture Notes in Mathematics 1353, Springer, Berlin (1988).

[14] M. M. Postnikov: Geometry VI. Riemanniann Geometry, Encyclopaedia of Mathematical Sciences 91, Springer, Berlin (2001).