The role of perspective functions in convexity, polyconvexity, rank-one convexity and separate convexity

Bernard Dacorogna* and Pierre Maréchal†

October 4, 2007

Abstract

Any finite, separately convex, positively homogeneous function on $\mathbb{R}^2$ is convex. This was first established in [1]. In this paper, we give a new and concise proof of this result, and we show that it fails in higher dimension. The key of the new proof is the notion of *perspective* of a convex function f, namely, the function $(x, y) \to yf(x/y)$, $y > 0$. In recent works [9, 10, 11], the perspective has been substantially generalized by considering functions of the form $(x, y) \to g(y)f(x/g(y))$, with suitable assumptions on g. Here, this *generalized perspective* is shown to be a powerful tool for the analysis of convexity properties of parametrized families of matrix functions.

1 Introduction

In [1], Dacorogna established the following theorem:

Theorem 1 *Let $f \colon \mathbb{R}^2 \to \mathbb{R}$ be separately convex and positively homogeneous of degree one. Then f is convex.*

A rather natural question then arises: does this theorem remain valid in higher dimension ? As we will see, the answer is negative.
In Section 2 of this paper, we provide a new and concise proof of the above theorem, which uses the notion of *perspective* in convex analysis. We then

*Section de Mathématiques, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland. Email: bernard.dacorogna@epfl.ch.

†Institut de Mathématiques, Université Paul Sabatier, F-31062 Toulouse cedex 4, France. Email: marechal@mip.ups-tlse.fr.

establish that the result fails for functions on $\mathbb{R}^n$ as soon as $n \geq 3$. We construct counterexamples in dimension 3 and 4, using ideas from [3]. We also point out that the theorem is false even in dimension 2 if the function is not everywhere finite.

The role of the perspective in the analysis of convexity properties of functions is further explored in the subsequent sections. An overview of a convex analytic operation recently introduced by Maréchal in [9, 10, 11, 12], which generalizes the perspective, is given in Section 3. It is then applied to the study of parametrized families of matrix functions in Section 4.

2 Perspective and separately convex homogeneous functions

Throughout, we denote by $\mathbb{R}_+^*$ (resp. $\mathbb{R}_-^*$) the set of positive (resp. negative) numbers.

2.1 Perspective functions

A standard way to produce a convex and positively homogeneous function on $\mathbb{R}^n \times \mathbb{R}_+^*$ is to form the *perspective* of some convex function f on $\mathbb{R}^n$. This is recalled in the following lemma, whose proof is provided for the sake of completeness.

Lemma 2 *Let $f\colon \mathbb{R}^n \to [-\infty, \infty]$. Then, the function $\breve{f}$ defined by*

$$\breve{f}(x,y) = yf\left(\frac{x}{y}\right), \quad x \in \mathbb{R}^n,\ y \in \mathbb{R}_+^*$$

is convex if and only if f is convex.

PROOF. The *only if* part is obvious (take $y = 1$). Conversely, if f is convex, then

$$\begin{aligned} &((1-\lambda)y_1 + \lambda y_2) f\left(\frac{(1-\lambda)x_1 + \lambda x_2}{(1-\lambda)y_1 + \lambda y_2}\right) \\ = \;& ((1-\lambda)y_1 + \lambda y_2) f\left(\frac{(1-\lambda)y_1}{(1-\lambda)y_1 + \lambda y_2}\frac{x_1}{y_1} + \frac{\lambda y_2}{(1-\lambda)y_1 + \lambda y_2}\frac{x_2}{y_2}\right) \\ \leq \;& (1-\lambda)y_1 f\left(\frac{x_1}{y_1}\right) + \lambda y_2 f\left(\frac{x_2}{y_2}\right) \end{aligned}$$

for all $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^n \times \mathbb{R}_+^*$ and all $\lambda \in (0,1)$. ∎

It is customary to allow y to vanish, in the definition of $\breve{f}$, by letting

$$\breve{f}(x,0) = f0^+(x) := \sup \left\{ f(x+z) - f(z) \;\middle|\; z \in \operatorname{dom} f \right\}$$

Here, $f0^+$ is the recession function of f (see [13], Section 8). Recall that, if f is closed proper convex, then

$$\forall x \in \operatorname{dom} f, \quad (f0^+)(x) = \lim_{y \downarrow 0} y f\left(\frac{x}{y}\right),$$

and that the latter formula holds for all $x \in \mathbb{R}^n$ in the case where the domain of f contains the origin (see [13], Corollary 8.5.2).
In the remainder of this paper, we will always consider $\breve{f}$ to be extended in this way. It is well known that $\breve{f}$ is then closed if and only if f is closed.

2.2 A new proof of Theorem 1

We start with a lemma which allows to obtain convex functions on $\mathbb{R}$ and on $\mathbb{R}^2$ by repasting pieces of a function which is convex on overlapping domains.

Lemma 3 (i) *Let $f\colon \mathbb{R} \to \mathbb{R}$ and let $a, b \in \mathbb{R}$ be such that $a < b$. If f is convex on $(-\infty, b)$ and on (a, ∞), then f is convex on $\mathbb{R}$.*

(ii) *Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be continuous and convex on the open half-planes $\mathbb{R} \times \mathbb{R}_+^*$, $\mathbb{R} \times \mathbb{R}_-^*$, $\mathbb{R}_+^* \times \mathbb{R}$ and $\mathbb{R}_-^* \times \mathbb{R}$. Then f is convex on $\mathbb{R}^2$.*

Proof.

(i) The assumptions imply that f is continuous on $\mathbb{R}$, and that the right (or left) derivative of f exists at every $x \in \mathbb{R}$ and is increasing (see [8], Theorems I-3.1.1 and I-4.1.1 and Remark I-4.1.2). The convexity of f on $\mathbb{R}$ then follows from [8], Theorem I-5.3.1.

(ii) It suffices to see that f is convex on every line $\Delta \subset \mathbb{R}^2$. If Δ is parallel to one of the axes, then either it is contained in one of the four half-spaces under consideration, in which case there is nothing to prove, or it is one of the axes, in which case an obvious continuity argument shows the convexity of f on Δ. If Δ is not parallel to any of the axes, then either it intersects the axes at two distinct points, in which case the convexity of f on Δ is an immediate consequence of Part (i), or it passes through the origin, in which case the convexity of f on Δ results again from the continuity of f. ∎

We are now ready to give our new proof.

Proof of Theorem 1. Since f is finite and separately convex, it is continuous on $\mathbb{R}^2$ (see e.g. [1], Theorem 2.3 page 29). Now, the partial mapping $x \mapsto f(x,1)$ is convex by assumption, and Lemma 2 shows that the mapping

$$(x,y) \mapsto y f\left(\frac{x}{y}, 1\right) = f(x,y)$$

is convex on the open half-plane $\mathbb{R} \times \mathbb{R}_+^*$. Repeating the same reasoning with the partial mappings $x \mapsto f(x,-1)$, $y \mapsto f(1,y)$ and $y \mapsto f(-1,y)$ shows that f is also convex on the open half-planes $\mathbb{R} \times \mathbb{R}_-^*$, $\mathbb{R}_+^* \times \mathbb{R}$ and $\mathbb{R}_-^* \times \mathbb{R}$. The theorem then follows from Lemma 3(ii). ■

2.3 Counterexamples

Notice first that, in Theorem 1, the assumption of finiteness of f is essential. As a matter of fact, it is clear that the *indicator function* of the set

$$E = \left\{ (x,y) \in \mathbb{R}^2 \;\middle|\; xy \geq 0 \right\}$$

is positively homogeneous and separately convex but not convex. Recall that the indicator function of a set E is the function

$$\delta(x|E) = \begin{cases} 0 & \text{if } x \in E, \\ \infty & \text{otherwise.} \end{cases}$$

We now turn to higher dimensional considerations. As announced in the introduction of this paper, Theorem 1 fails for functions on $\mathbb{R}^n$ as soon as $n \geq 3$. Our counterexamples will all be of the form given in the following proposition. We denote by S^{n-1} the unit sphere in $\mathbb{R}^n$ and by $\mathcal{E} = \{e_1, \ldots, e_n\}$ the Euclidean basis of $\mathbb{R}^n$. We also define the sets

$$\mathcal{C} := \left\{ (\xi,\eta) \in S^{n-1} \times S^{n-1} \;\middle|\; \langle \xi, \eta \rangle = 0 \right\}$$

and

$$\mathcal{S} := \left\{ (\xi,\eta) \in S^{n-1} \times S^{n-1} \;\middle|\; \langle \xi, \eta \rangle = 0,\ \exists (t,s) \in \mathbb{R} \times \mathbb{R} \colon t\xi + s\eta \in \mathcal{E} \right\}.$$

Proposition 4 *Let M be an $n \times n$ real symmetric matrix, with eigenvalues $\mu_1 \leq \mu_2 \leq \ldots \leq \mu_n$ and corresponding orthonormal set of eigenvectors $\{\varphi_1, \varphi_2, \ldots, \varphi_n\}$. Let*

$$f(\xi) := \begin{cases} \dfrac{\langle M\xi, \xi \rangle}{\|\xi\|} & \text{if } \xi \neq 0, \\ 0 & \text{if } \xi = 0. \end{cases}$$

Then

$$f \text{ is convex} \iff u \geq 0 \iff 2\mu_1 - \mu_n \geq 0,$$

and

$$f \text{ is separately convex} \iff v \geq 0,$$

where

$$u := \min_{(\xi,\eta)\in\mathcal{C}} \big\{ 2\langle M\eta, \eta\rangle - \langle M\xi, \xi\rangle \big\} \text{ and } v := \min_{(\xi,\eta)\in\mathcal{S}} \big\{ 2\langle M\eta, \eta\rangle - \langle M\xi, \xi\rangle \big\}.$$

PROOF. Since f is continuous on $\mathbb{R}^n$, the convexity properties under consideration may be examined only on every line which does not contain the origin. It follows that f is convex if and only if

$$\inf_{\xi,\lambda\in\mathbb{R}^n\setminus\{0\}} \big\{ \langle \nabla^2 f(\xi)\lambda, \lambda\rangle \big\} \geq 0,$$

and it is separately convex if and only if

$$\inf_{\substack{\xi\in\mathbb{R}^n\setminus\{0\} \\ \lambda\in\mathcal{E}}} \big\{ \langle \nabla^2 f(\xi)\lambda, \lambda\rangle \big\} \geq 0.$$

Straightforward computations show that

$$\begin{aligned} \langle \nabla^2 f(\xi)\lambda, \lambda\rangle &= \frac{1}{\|\xi\|^5}\Big(2\|\xi\|^4 \langle M\lambda, \lambda\rangle - 4\|\xi\|^2 \langle M\xi, \lambda\rangle\langle \xi, \lambda\rangle \\ &\quad - \|\xi\|^2 \|\lambda\|^2 \langle M\xi, \xi\rangle + 3\langle \xi, \lambda\rangle^2 \langle M\xi, \xi\rangle \Big). \end{aligned}$$

Since the above expression is positively homogeneous of degree -1 in ξ, one can add the condition $\|\xi\| = 1$ in the previous *infima*. Furthermore, every λ in $\mathbb{R}^n$ can be written

$$\lambda = t\xi + s\eta \quad \text{with} \quad t, s \in \mathbb{R}, \quad \|\eta\| = 1 \quad \text{and} \quad \langle \xi, \eta\rangle = 0.$$

We then have:

$$\begin{aligned} |\lambda|^2 &= t^2 + s^2, \\ \langle \xi, \lambda\rangle &= t, \\ \langle M\xi, \lambda\rangle &= t\langle M\xi, \xi\rangle + s\langle M\xi, \eta\rangle, \\ \langle M\lambda, \lambda\rangle &= t^2\langle M\xi, \xi\rangle + 2st\langle M\xi, \eta\rangle + s^2\langle M\eta, \eta\rangle, \end{aligned}$$

so that

$$\begin{aligned}\langle \nabla^2 f(\xi)\lambda, \lambda\rangle &= 2\big(t^2\langle M\xi,\xi\rangle + 2st\langle M\xi,\eta\rangle + s^2\langle M\eta,\eta\rangle\big)\\ &\quad -4t\big(t\langle M\xi,\xi\rangle + s\langle M\xi,\eta\rangle\big)\\ &\quad -(t^2+s^2)\langle M\xi,\xi\rangle + 3t^2\langle M\xi,\xi\rangle\\ &= s^2\big(2\langle M\eta,\eta\rangle - \langle M\xi,\xi\rangle\big).\end{aligned}$$

Therefore, the change of variable $(\xi,\lambda)\to(\xi,\eta)$ shows that f is convex if and only if

$$u = \inf_{(\xi,\eta)\in\mathcal{C}} \big\{2\langle M\eta,\eta\rangle - \langle M\xi,\xi\rangle\big\} \geq 0, \tag{1}$$

and that f is separately convex if and only if

$$v = \inf_{(\xi,\eta)\in\mathcal{S}} \big\{2\langle M\eta,\eta\rangle - \langle M\xi,\xi\rangle\big\} \geq 0.$$

It is clear that both *infima* are attained, and that the *infimum* in (1) is attained for $\eta=\varphi_1$ and $\xi=\varphi_n$, so that f is convex if and only if

$$2\mu_1 - \mu_n \geq 0. \ \blacksquare$$

We now turn to counterexamples to Theorem 1 in higher dimension.

Example 5 **($n=3$)** Let γ be a nonnegative parameter, let $M_\gamma := A+\gamma B$, where

$$A := \begin{bmatrix} 8 & 2 & -1\\ 2 & 8 & -1\\ -1 & -1 & 11\end{bmatrix} \quad\text{and}\quad B := \begin{bmatrix} -1 & 1 & 0\\ 1 & -1 & 0\\ 0 & 0 & 0\end{bmatrix},$$

and let f be as in the above proposition. Finally, let

$$\begin{aligned} u_\gamma &= \min_{(\xi,\eta)\in\mathcal{C}} \big\{2\langle M_\gamma\eta,\eta\rangle - \langle M_\gamma\xi,\xi\rangle\big\},\\ v_\gamma &= \min_{(\xi,\eta)\in\mathcal{S}} \big\{2\langle M_\gamma\eta,\eta\rangle - \langle M_\gamma\xi,\xi\rangle\big\}.\end{aligned}$$

The vectors

$$\varphi_1 = \frac{\sqrt{2}}{2}\begin{pmatrix}1\\-1\\0\end{pmatrix}, \quad \varphi_2 = \frac{\sqrt{3}}{3}\begin{pmatrix}1\\1\\1\end{pmatrix}, \quad \varphi_3 = \frac{\sqrt{6}}{6}\begin{pmatrix}1\\1\\-2\end{pmatrix}$$

form an orthonormal system of eigenvectors for both A and B, with eigenvalues $\{6, 9, 12\}$ and $\{-2, 0, 0\}$, respectively. We clearly have, as in the proposition,

$$\begin{aligned} u_\gamma &= 2(6-2\gamma) - 12 = -4\gamma, \\ v_\gamma &\geq \min_{(\xi,\eta)\in\mathcal{S}} \big\{2\langle A\eta, \eta\rangle - \langle A\xi, \xi\rangle\big\} - \gamma \max_{(\xi,\eta)\in\mathcal{S}} \big\{2\langle B\eta, \eta\rangle - \langle B\xi, \xi\rangle\big\} \\ &\geq v_0 - 2\gamma, \end{aligned}$$

since

$$\max_{(\xi,\eta)\in\mathcal{S}} \big\{2\langle B\eta, \eta\rangle - \langle B\xi, \xi\rangle\big\} \leq \max_{(\xi,\eta)\in\mathcal{C}} \big\{2\langle B\eta, \eta\rangle - \langle B\xi, \xi\rangle\big\} = 2.$$

Moreover, $v_0 > 0$ since $e_1, e_2, e_3 \notin \operatorname{span}\{\varphi_1, \varphi_3\}$. Therefore, choosing $\gamma > 0$ sufficiently small guarantees that

$$v_\gamma > 0 > u_\gamma$$

which, according to the proposition, shows that f_γ is separately convex but not convex. ∎

Example 6 **($n = 4$)** Let

$$M := \begin{bmatrix} 10 & 0 & 0 & 1 \\ 0 & 7 & 2 & 0 \\ 0 & 2 & 7 & 0 \\ 1 & 0 & 0 & 10 \end{bmatrix}$$

and let f be as in the proposition. This function, regarded as a function on the space of real 2×2 matrices, was shown to be rank-one convex but not convex (see [3], Remark 1.9). Since rank one convex functions are trivially separately convex, we have the desired counterexample. ∎

Finally, observe that Theorem 1 can be generalized to an n-dimensional setting as follows:

Theorem 7 *Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be $(n-1)$-partially convex and positively homogeneous of degree one. Then f is convex.*

A function $f\colon \mathbb{R}^n \to [-\infty, \infty]$ is said to be *k-partially convex* if each partial mapping obtained by assigning any prescribed values to $n - k$ variables is convex. As the reader may check, the proof of the latter result is a straightforward adaptation of our proof of Theorem 1.

3 Generalized perspective

The notion of perspective has been significantly generalized in [9, 10, 11], where convex functions on $\mathbb{R}^{n+m}$ are obtained from convex functions on $\mathbb{R}^n$ and $\mathbb{R}^m$. We recall here the main features of this construction. Given any function ϕ on $\mathbb{R}^n$, the convex conjugate of ϕ is denoted by $\phi^\star$.

Definition 8 (i) Let $\varphi\colon \mathbb{R}^n \to (-\infty, \infty]$ be proper convex, with $\varphi(0) \leq 0$, and let $\psi\colon \mathbb{R}^m \to \{-\infty\} \cup [0, \infty)$ be proper concave. The pair (φ, ψ) is then said to be of type I, and we denote by $\varphi \triangle \psi$ the function given, on $\mathbb{R}^n \times \mathbb{R}^m$, by

$$(\varphi \triangle \psi)(x, y) := \begin{cases} \psi(y)\varphi\left(\dfrac{x}{\psi(y)}\right) & \text{if } \psi(y) \in (0, \infty), \\ \varphi 0^+(x) & \text{if } \psi(y) = 0, \\ \infty & \text{if } \psi(y) = -\infty. \end{cases}$$

(ii) Let $\varphi\colon \mathbb{R}^n \to (-\infty, \infty]$ be proper convex with $\varphi \geq \varphi 0^+$, and let $\psi\colon \mathbb{R}^m \to [0, \infty]$ be proper convex. The pair (φ, ψ) is then said to be of type II, and we denote by $\varphi \triangle \psi$ the function given, on $\mathbb{R}^n \times \mathbb{R}^m$, by

$$(\varphi \triangle \psi)(x, y) := \begin{cases} \psi(y)\varphi\left(\dfrac{x}{\psi(y)}\right) & \text{if } \psi(y) \in (0, \infty), \\ \varphi 0^+(x) & \text{if } \psi(y) = 0, \\ \infty & \text{if } \psi(y) = \infty \end{cases}$$

in the case where $\varphi \neq \varphi 0^+$, and by

$$(\varphi \triangle \psi)(x, y) := \begin{cases} \varphi(x) & \text{if } y \in \operatorname{cl} \operatorname{dom} \psi, \\ \infty & \text{if } y \notin \operatorname{cl} \operatorname{dom} \psi \end{cases}$$

in the case where $\varphi = \varphi 0^+$.

The condition $\varphi = \varphi 0^+$ is equivalent to positive homogeneity of φ. In Case (ii), the particular definition of $\varphi \triangle \psi$ for positively homogeneous φ coincides with the general one, except when $y \in \operatorname{cl} \operatorname{dom} \psi \setminus \operatorname{dom} \psi$ (the latter set may be nonempty, even if ψ is closed). This definition ensures closedness of $\varphi \triangle \psi$ whenever φ and ψ are closed. The proof of the following theorem can be found in [10].

Theorem 9 (i) *Let* (φ, ψ) *be of type* I*, and suppose that* φ *and* ψ *are closed. Then* $((-\psi)^\star, \varphi^\star)$ *is of type* II*, and the following duality relationships hold:*

$$\begin{aligned} (\varphi \bigtriangleup \psi)^\star(\xi, \eta) &= \big((-\psi)^\star \bigtriangleup \varphi^\star\big)(\eta, \xi) \\ \big((-\psi)^\star \bigtriangleup \varphi^\star\big)^\star(y, x) &= (\varphi \bigtriangleup \psi)(x, y). \end{aligned}$$

Consequently, $\varphi \bigtriangleup \psi$ *is closed proper convex.*

(ii) *Let* (φ, ψ) *be of type* II*, and suppose that* φ *and* ψ *are closed. Then* $(\psi^\star, -\varphi^\star)$ *is of type* I*, and the following duality relationships hold:*

$$\begin{aligned} (\varphi \bigtriangleup \psi)^\star(\xi, \eta) &= \big(\psi^\star \bigtriangleup (-\varphi^\star)\big)(\eta, \xi) \\ \big(\psi^\star \bigtriangleup (-\varphi^\star)\big)^\star(y, x) &= (\varphi \bigtriangleup \psi)(x, y). \end{aligned}$$

Consequently, $\varphi \bigtriangleup \psi$ *is closed proper convex.*

4 Applications

In the forthcoming developments, we intend to demonstrate the relevance of the generalized perspective as a tool for the study of convexity properties of families of matrix functions.
We denote by $M_{m\times n}$ the space of real $m \times n$ matrices, and we write $M_n = M_{n\times n}$. Recall that $\delta(\cdot|C)$ denotes the indicator function of a set C.

Theorem 10 *Let* $f\colon M_n \to (-\infty, \infty]$ *be defined by*

$$f(A) = \begin{cases} \dfrac{\|\mathrm{adj}_s A\|^\gamma}{(\det A)^\alpha} & \text{if } \det A > 0, \\ \delta(\mathrm{adj}_s A \,|\{0\}) & \text{if } \det A = 0, \\ \infty & \text{if } \det A < 0, \end{cases}$$

in which $s \in \{1, \ldots, n-1\}$ *and* $\gamma > \alpha > 0$*. Then the following are equivalent:*

(i) f *is polyconvex;*

(ii) f *is rank-one convex;*

(iii) $\gamma \geq 1 + \alpha$.

Proof. It is well known that polyconvexity implies rank-one convexity (see [1]). Let us prove that (ii) implies (iii). Assuming that f is rank-one convex, let $A \in M_n$ and let $u, v \in \mathbb{R}^n$ be such that $\det(A + tu \otimes v) > 0$ for all $t > 0$. By assumption, the function

$$\phi(t) := f(A + tu \otimes v) = \frac{\|\mathrm{adj}_s(A + tu \otimes v)\|^\gamma}{(\det(A + tu \otimes v))^\alpha}, \quad t > 0$$

is convex. By Proposition 16 (see the appendix),

$$\|\mathrm{adj}_s(A + tu \otimes v)\|^2 = at^2 + bt + c,$$

and $\det(A + tu \otimes v) = dt + e$ with $d, e \in \mathbb{R}$. Consequently,

$$\phi(t) = (at^2 + bt + c)^{\gamma/2}(dt + e)^{-\alpha}.$$

Now, a direct computation shows that

$$\begin{aligned}\phi''(t) = {} & (at^2 + bt + c)^{\gamma/2-2} \times \\ & (dt + e)^{-\alpha-2}\left[P(t) + a^2d^2(\gamma^2 - \gamma - 2\alpha\gamma + \alpha(\alpha+1))t^4\right],\end{aligned}$$

in which P is a polynomial of degree less than or equal to 3. For ϕ'' to be nonnegative (on $\mathbb{R}_+^*$), it is necessary that

$$\gamma^2 - \gamma - 2\alpha\gamma + \alpha(\alpha+1) \geq 0,$$

that is, that $(\gamma - \alpha)^2 \geq \gamma - \alpha$. But this implies in turn that $\gamma \geq 1 + \alpha$.
It remains to show that (iii) implies (i). On the one hand, it is clear that the function φ defined on $M_{C_n^s}$ by $\varphi(\xi) = \|\xi\|^\gamma$ is convex and satisfies $\varphi(0) \leq 0$. On the other hand, (iii) implies that $\beta := \alpha/(\gamma - 1) \in (0, 1]$, and the function ψ defined on $\mathbb{R}$ by

$$\psi(y) = \begin{cases} y^\beta & \text{if } y \geq 0, \\ -\infty & \text{otherwise} \end{cases}$$

is closed proper concave and nonnegative on its domain. Theorem 9(i) then shows that

$$(\varphi \bigtriangleup \psi)(\xi, d) = \begin{cases} \dfrac{\|\xi\|^\gamma}{d^\alpha} & \text{if } d > 0, \\ \delta(\xi|\{0\}) & \text{if } d = 0, \\ \infty & \text{if } d < 0 \end{cases}$$

is closed proper convex, and the conclusion follows from the fact that

$$f(A) = (\varphi \bigtriangleup \psi)(\mathrm{adj}_s A, \det A).$$

Notice that, since $\varphi \bigtriangleup \psi$ is lower semi-continuous, so is f. ∎

Another application of the generalized perspective is the following.

Theorem 11 *Let $f_\alpha(A) := (|A|^2 + 2|\det A|^{2\alpha})^{1/2}$, $A \in M_2$, where α is a nonnegative parameter. Then*

(1) *f_α is convex if and only if $\alpha \in \{0, 1/2\}$;*

(2) *f_α is polyconvex if and only if f_α is rank-one convex if and only if $\alpha \in \{0, 1/2\} \cup [1, \infty)$.*

Proof. Step 1. We first prove by contradiction that f_α rank-one convex implies $\alpha \in \{0, 1/2\} \cup [1, \infty)$. So assume that $\alpha \in (0, 1/2) \cup (1/2, 1)$, and consider

$$A := \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad u = v := \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{so that} \quad A + tu \otimes v = \begin{bmatrix} 1 & 0 \\ 0 & t \end{bmatrix}.$$

Then $|A + tu \otimes v|^2 = 1 + t^2$ and $\det(A + tu \otimes v) = t$, so that

$$\phi(t) := f_\alpha(A + tu \otimes v) = (1 + t^2 + 2(t^2)^\alpha)^{1/2}.$$

We may restrict attention to positive t, for which $\phi(t) := f_\alpha(A + tu \otimes v) = (1 + t^2 + 2t^{2\alpha})^{1/2}$, and show that ϕ'' takes negative values. A straightforward computation shows that

$$t^2\phi^3(t)\phi''(t) = 2\alpha(2\alpha - 1)t^{2\alpha} + 4(\alpha^2 - \alpha)t^{4\alpha} + 2(2\alpha^2 - 3\alpha + 1)t^{2\alpha+2} + t^2.$$

Suppose that $\alpha \in (0, 1/2)$. Then, for small values of t, the dominant term in the above expression is $2\alpha(2\alpha - 1)t^{2\alpha}$. Since $2\alpha - 1 < 0$, we see that $t^2\phi^3(t)\phi''(t)$ is negative for small enough $t > 0$. Suppose now that $\alpha \in (1/2, 1)$. Then, for large values of t, the dominant term is

$$2(2\alpha^2 - 3\alpha + 1)t^{2\alpha+2}.$$

Since $2\alpha^2 - 3\alpha + 1 < 0$, we see that $t^2\phi^3(t)\phi''(t)$ is negative for large enough t.

Step 2. Next, we prove that if $\alpha \in \{0, 1/2\}$, then f_α is convex. Let $\lambda_1(A) \leq \lambda_2(A)$ be the singular values of A. Then $f_\alpha(A) = (\lambda_1^2(A) + \lambda_2^2(A) + 2(\lambda_1(A)\lambda_2(A))^{2\alpha})^{1/2}$. Theorem 7.8 in [5] then shows that the convexity of f_α is equivalent to that of

$$g_\alpha(x, y) := (x^2 + y^2 + 2(xy)^{2\alpha})^{1/2}$$

on $\mathbb{R}_+^2$. As a matter of fact, g_α is clearly symmetric and componentwise increasing. Therefore, we need only check the convexity of g_0 and $g_{1/2}$. But $g_0(x, y) = (2 + x^2 + y^2)^{1/2}$ and $g_{1/2}(x, y) = x + y$ on $\mathbb{R}_+^2$. The convexity of both functions being clear, the desired result is established.

Step 3. We now prove that, if $\alpha \geq 1$, then f_α is polyconvex. Let

$$\varphi(x) := (x^2 + 2)^{1/2},\ x \in \mathbb{R}, \quad \text{and} \quad \psi(\delta) := |\delta|^\alpha.$$

Both functions are closed proper convex and nonnegative. Furthermore, the recession function of φ is given by $\varphi 0^+(x) = |x|$. Thus $\varphi \geq \varphi 0^+$, and the function $h := \varphi \bigtriangleup \psi$ satisfies:

$$h(x, \delta) = |\delta|^\alpha(\left(\frac{x}{|\delta|^\alpha}\right)^2 + 2)^{1/2} = (x^2 + 2|\delta|^{2\alpha})^{1/2}.$$

By Theorem 9, h is convex. Now, there is no doubt that $x \mapsto h(x, \delta)$ is an increasing function. Consequently,

$$(A, \delta) \mapsto h(\|A\|, \delta)$$

is convex on $M_2 \times \mathbb{R}$, and the polyconvexity of f_α follows.

Step 4. Finally, we prove that f_α is not convex for $\alpha \geq 1$. In order to achieve this goal, we consider again the function g_α defined in Step 2, and show that its Hessian matrix H fails to be positive semi-definite. We have:

$$H = \begin{bmatrix} g_{\alpha xx} & g_{\alpha xy} \\ g_{\alpha xy} & g_{\alpha yy} \end{bmatrix},$$

in which $g_{\alpha xx} := \partial^2 g_\alpha / \partial^2 x$, $g_{\alpha xy} := \partial^2 g_\alpha / \partial x \partial y$ and $g_{\alpha yy} := \partial^2 g_\alpha / \partial^2 x$ satisfy

$$\begin{aligned}
x^2 g_\alpha^3(x, x) g_{\alpha xx}(x, x) &= 2(4\alpha^2 - 4\alpha + 1)x^{4\alpha+2} + 4\alpha(\alpha - 1)x^{8\alpha} + x^4, \\
x^2 g_\alpha^3(x, x) g_{\alpha xy}(x, x) &= 4\alpha(2\alpha - 1)x^{4\alpha+2} + 4\alpha^2 x^{8\alpha} - x^4, \\
x^2 g_\alpha^3(x, x) g_{\alpha yy}(x, x) &= 2(4\alpha^2 - 4\alpha + 1)x^{4\alpha+2} + 4\alpha(\alpha - 1)x^{8\alpha} + x^4.
\end{aligned}$$

We see that, if $w := (-1,1)$, then

$$x^2 g_\alpha^3(x,x)\langle w, H(x,x)w\rangle = 4\big((1-2\alpha)x^{4\alpha+2} - 2\alpha x^{8\alpha} + x^4\big).$$

For small values of x, the dominant term is $-4\alpha x^{8\alpha}$. This shows that $\langle w, H(x,x)w\rangle$ takes negative values, and the proof is complete. ∎

5 Appendix: Adjugate matrix, polyconvex and rank-one convex functions

We recall here a few basic facts about adjugate matrices, polyconvex and rank-one convex matrix functions. For a more complete exposition, the reader is referred to [1]. Some of the missing proofs may also be found in [7].

5.1 Adjugate matrices

Let $m \in \mathbb{N}^*$. For all $s \in \{1, \dots, m\}$, we endow the set

$$I_{m,s} := \left\{ (i_1, \dots, i_s) \in \mathbb{N}^s \;\middle|\; 1 \le i_1 < \dots < i_s \le m \right\}$$

with the inverse lexicographical order, which we denote by $\prec$. It is clear that,

$$\operatorname{card} I_{m,s} = C_m^s := \frac{m!}{s!(m-s)!}.$$

Let $\alpha = \alpha_{m,s}$ be the unique bijection from $\{1, \dots, C_m^s\}$ to $I_{m,s}$ such that

$$i > j \implies \alpha_{m,s}(i) \succ \alpha_{m,s}(j).$$

Let $A \in M_{m\times n}$. The *adjugate of order s of A* is the $C_m^s \times C_n^s$-matrix $\mathrm{adj}_s A$ given by

$$(\mathrm{adj}_s A)_{ij} := (-1)^{i+j} \det\left(A_{\alpha_{m,s}(i)\alpha_{n,s}(j)}\right),$$

in which $A_{\alpha_{m,s}(i)\alpha_{n,s}(j)}$ denotes the submatrix corresponding to $\alpha_{m,s}(i) = (i_1, \dots, i_s)$ and $\alpha_{n,s}(j) = (j_1, \dots, j_s)$, that is,

$$A_{\alpha_{m,s}(i)\alpha_{n,s}(j)} := \begin{pmatrix} A_{i_1 j_1} & \dots & A_{i_1 j_s} \\ \vdots & & \vdots \\ A_{i_s j_1} & \dots & A_{i_s j_s} \end{pmatrix} \in M_{s\times s}.$$

Now, let $\mathcal{A}_{m\times n} := M_{m\times n} \times M_{C_m^2\times C_n^2} \times \ldots \times M_{C_m^{m\wedge n}\times C_n^{m\wedge n}}$, and let

$$\begin{array}{rccl} \mathrm{adj}\colon & M_{m\times n} & \longrightarrow & \mathcal{A}_{m\times n} \\ & A & \longmapsto & \mathrm{adj}\,A := (A, \mathrm{adj}_2 A, \ldots, \mathrm{adj}_{m\wedge n} A). \end{array}$$

The space $\mathcal{A}_{m\times n}$ is isomorphic to $\mathbb{R}^\tau$, where $m \wedge n := \min\{m, n\}$ and

$$\tau = \tau(m,n) = mn + C_m^2 C_n^2 + \cdots + C_m^{m\wedge n} C_n^{m\wedge n} = \sum_{k=1}^{m\wedge n} C_m^k C_n^k.$$

We identify $\mathcal{A}_{m\times n}$ with the set of bloc diagonal matrices

$$\mathrm{bloc}(m \times n; C_m^2 \times C_n^2; \ldots; C_m^{m\wedge n} \times C_n^{m\wedge n})$$

and $\mathrm{adj}\,A$ with the bloc matrix

$$\begin{bmatrix} A & 0 & \ldots & 0 \\ 0 & \mathrm{adj}_2 A & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \mathrm{adj}_{m\wedge n} A \end{bmatrix} \in M_{m_0\times n_0},$$

where $m_0 := \sum_{k=1}^{m\wedge n} C_m^k$ and $n_0 := \sum_{k=1}^{m\wedge n} C_n^k$. In the case where $m = n$, $\tau = \sum_{k=1}^{n} \big(C_n^k\big)^2$ and $m_0 = n_0 = \sum_{k=1}^{n} C_n^k$. In this case, we put $\mathcal{A}_n := \mathcal{A}_{n\times n}$ and $\tau(n) := \tau(m,n)$. Let us review a few basic facts about adjugate matrices.

Theorem 12 *Let $A \in M_{l\times m}$ and $B \in M_{m\times n}$. Then,*

$$\forall s \in \big\{1, \ldots, \min\{l, m, n\}\big\}, \quad \mathrm{adj}_s AB = \mathrm{adj}_s A\, \mathrm{adj}_s B\,.$$

Theorem 13 *Let $A \in M_{m\times n}(\mathbb{R})$ and $s \in \{1, \ldots, m \wedge n\}$. Then*

$$\mathrm{adj}_s A^t = (\mathrm{adj}_s A)^t.$$

Theorem 14 *Let $A \in M_n(\mathbb{R})$ and $s \in \{1, \ldots, n\}$. If A is diagonal, then so is $\mathrm{adj}_s A$. More precisely,*

$$\mathrm{adj}_s \mathrm{diag}\, a = \mathrm{diag}\left(\prod_{j\in\alpha(1)} a_j, \ldots, \prod_{j\in\alpha(C_n^s)} a_j \right),$$

where $\alpha = \alpha_{n,s}$ is defined as above. In particular, $\mathrm{adj}_s I_n = I_{C_n^s}$.

Theorem 15 *Let* $A \in M_n(\mathbb{R})$.

(i) *If* $A \in \mathrm{GL}(n)$, *then* $\mathrm{adj}_s A \in \mathrm{GL}(C_n^s)$ *and* $(\mathrm{adj}_s A)^{-1} = \mathrm{adj}_s A^{-1}$ *for all* $s \in \{2, \ldots, n\}$, *so that* $\mathrm{adj}A \in \mathrm{GL}(\sum_{s=1}^n C_n^s)$ *and* $(\mathrm{adj}A)^{-1} = \mathrm{adj}A^{-1}$.

(ii) *If* $A \in \mathrm{O}(n)$, *then* $\mathrm{adj}_s A \in \mathrm{O}(C_n^s)$ *for all* $s \in \{2, \ldots, n\}$, *so that* $\mathrm{adj}A \in \mathrm{O}(\sum_{s=1}^n C_n^s)$.

(iii) *If* $A \in \mathrm{SO}(n)$, *then* $\mathrm{adj}_s A \in \mathrm{SO}(C_n^s)$ *for all* $s \in \{2, \ldots, n\}$, *so that* $\mathrm{adj}A \in \mathrm{SO}(\sum_{s=1}^n C_n^s)$.

Proposition 16 *Let* $A \in M_n$, $u, v \in \mathbb{R}^n$ *and* $s \in \{1, \ldots, n\}$. *Then, for all* $t \in \mathbb{R}$,

$$\mathrm{adj}_s(A + tu \otimes v) = (1-t)\,\mathrm{adj}_s A + t\,\mathrm{adj}_s(A + u \otimes v)\,.$$

In particular,

$$\det(A + tu \otimes v) = (1-t)\det A + t\det(A + u \otimes v).$$

Proof. Let us write $u \otimes v = PEP^{-1}$, where $P \in \mathrm{GL}(n)$ and $E = (E_{ij})$ is such that $E_{11} = 1$ and all other entries are zero. We then have

$$A + tu \otimes v = P(A' + tE)P^{-1},$$

and Theorems 12 and 15(i) show that

$$\mathrm{adj}_s(A + tu \otimes v) = \mathrm{adj}_s P\,\mathrm{adj}_s(A' + tE)(\mathrm{adj}_s P)^{-1}.$$

It is clear that $\mathrm{adj}_s(A' + tE)$ depends affinely on t:

$$\mathrm{adj}_s(A' + tE) = A_0 t + B_0, \quad \text{with} \quad A_0, B_0 \in M_{C_n^s}.$$

Therefore, letting $\xi := \mathrm{adj}_s P\,A_0(\mathrm{adj}_s P)^{-1}$ and $\eta := \mathrm{adj}_s P\,B_0(\mathrm{adj}_s P)^{-1}$, we see that

$$\mathrm{adj}_s(A + tu \otimes v) = \xi t + \eta$$

and the choices $t = 0$ and $t = 1$ yield the desired formula. ∎

5.2 Polyconvex and rank-one convex functions

A function $f\colon M_{N\times n} \to [-\infty, \infty]$ is said to be polyconvex if there exist a convex function

$$F\colon \mathcal{A}_{N\times n} \to [-\infty, \infty]$$

such that $f = F \circ \mathrm{adj}$. As in convex analysis, we will say that a function $f\colon M_{N\times n} \to [-\infty, \infty]$ is *proper* if it is nowhere equal to $-\infty$ and not identically equal to ∞.

Let $f\colon M_{N\times n} \to [-\infty, \infty]$. Following [1], we define the *polyconvex conjugate* of f as the function $f^P\colon \mathcal{A}_{N\times n} \to [-\infty, \infty]$ given for all $X \in \mathcal{A}_{N\times n}$ by

$$f^P(X) := \sup\left\{ \langle X, \mathrm{adj} A\rangle - f(A) \;\middle|\; A \in M_{N\times n} \right\}.$$

As the supremum of a family of affine functions, it is a closed convex function. We will see below that, if f is proper and minorized by a polyaffine function, then f^P is also proper.

Proposition 17 *Let $f\colon M_{N\times n} \to (-\infty, \infty]$ be proper. The following conditions are equivalent.*

(i) *There exists a convex function $c\colon \mathcal{A}_{N\times n} \to (-\infty, \infty]$ such that, for all $A \in M_{N\times n}$, $f(A) \geq c(\mathrm{adj} A)$ (f has a polyconvex minorant);*

(ii) *there exists $X_0 \in \mathcal{A}_{N\times n}$ and $K \in \mathbb{R}$ such that, for all $A \in M_{N\times n}$, $f(A) \geq \langle X_0, \mathrm{adj} A\rangle - K$ (f has a polyaffine minorant).*

Under these equivalent conditions, the fonction f^P is closed proper convex.

The *polyconvex biconjugate* of f is defined to be the function $f^{PP}\colon M_{N\times n} \to [-\infty, \infty]$ given by

$$f^{PP}(A) := (f^P)^\star(\mathrm{adj} A) = \sup\left\{ \langle X, \mathrm{adj} A\rangle - f^P(X) \;\middle|\; X \in \mathcal{A}_{N\times n} \right\}.$$

If f is proper and minorized by some polyaffine function, then f^P and $(f^P)^\star$ are closed proper convex, and f^{PP} is closed proper polyconvex.

Proposition 18 *Let $f\colon M_{N\times n} \to (-\infty, \infty]$.*

(i) $f^{PP} \leq f$;

(ii) *if f is proper and has a polyaffine minorant, then $f^{PPP} := (f^{PP})^P = f^P$;*

(iii) *if there exists* $F\colon \mathcal{A}_{N\times n} \to (-\infty,\infty]$ *closed proper convex such that* $f = F \circ \operatorname{adj}$, *then* $f^{PP} = f$.

Finally, a function $f\colon M_{N\times n} \to \mathbb{R}$ is said to be *rank-one convex* if it is convex in every direction of rank one, that is to say, if

$$f\big(\alpha\xi + (1-\alpha)\eta\big) \leq \alpha f(\xi) + (1-\alpha) f(\eta)$$

for every $\alpha \in (0,1)$, $\xi, \eta \in M_{N\times n}$ with $\operatorname{rk}[\xi - \eta] \leq 1$.
Recall that *convexity* implies *polyconvexity*, which in turn implies *rank-one convexity* [1].

Acknowledgements: This research started during the visit of the second author at the EPFL. The authors wish to thank J.-B. Hiriart-Urruty for interesting discussions and comments.

References

[1] B. Dacorogna, *Direct Methods in the Calculus of Variations*, Springer-Verlag, 1989.

[2] B. Dacorogna, *On rank-one convex functions which are homogeneous of degree one*, in *Calculus of variations, applications and computations (Pont-à-Mousson, 1994)*, Pitman Research Notes in Mathematics Series, **326**, Longman Science and Technology, Harlow, pp. 84-99, 1995.

[3] B. Dacorogna and J.-P. Haeberly, *On convexity properties of homogeneous functions of degree one*, Proceedings of the Royal Society of Edinburgh, **126A**, pp. 947-956, 1996.

[4] B. Dacorogna & H. Koshigoe, *On the different notions of convexity for rotationally invariant functions*, Annales de la Faculté des Sciences de Toulouse, **II(2)**, pp. 163-184, 1993.

[5] B. Dacorogna & P. Marcellini, *Implicit Partial Differential Equations*, Birkhäuser, 1999.

[6] B. Dacorogna & P. Maréchal, *Convex* $\mathrm{SO}(N) \times \mathrm{SO}(n)$*-invariant functions and refinements of Von Neumann's inequality*, to appear in Annales de la Faculté des Sciences de Toulouse.

[7] B. Dacorogna & P. Maréchal, *A Note on Spectrally Defined Polyconvex Functions*, to appear in Collana della Facolta di Economia dell'Universita degli Studi del Sannio, Edizioni Scientifiche Italiane, 2006.

[8] J.-B. Hiriart-Urruty & C. Lemaréchal, *Convex Analysis and Minimization Algorithms, I and II*, Springer-Verlag, 1993.

[9] P. Maréchal, *On the Convexity of the Multiplicative Potential and Penalty Functions and Related Topics*, Mathematical Programming, **89A**, pp. 505–516, 2001.

[10] P. Maréchal, *On a Functional Operation Generating Convex Functions. Part I: Duality*, Journal of Optimization Theory and Applications, **126(1)**, pp. 175–189, 2005.

[11] P. Maréchal, *On a Functional Operation Generating Convex Functions. Part II: Algebraic Properties*, Journal of Optimization Theory and Applications, **126(2)**, pp. 357–366, 2005.

[12] P. Maréchal, *On a Class of Convex Sets and Functions*, Set Valued Analysis, **13**, pp. 197–212, 2005.

[13] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.