Algebraically inspired results on convex functions and bodies

Liran Rotem

School of Mathematical Sciences, Tel-Aviv University
Tel-Aviv 6997801, Israel

liranro1@post.tau.ac.il

We show how algebraic identities, inequalities and constructions, which hold for numbers or matrices, often have analogs in the geometric classes of convex bodies or convex functions. By letting the polar body K° or the dual function φ^* play the role the inverses "K^{-1}" and "φ^{-1}", we are able to conjecture many new results, which often turn out to be correct.

As one example, we prove that for every convex function φ one has

$$(\varphi+\delta)^* + (\varphi^*+\delta)^* = \delta,$$

where $\delta(x) = \frac{1}{2}|x|^2$. We also prove several corollaries of this identity, including a Santal type inequality and a contribution to the theory of summands. We proceed to discuss the analogous identity for convex bodies, where an unexpected distinction appears between the classical Minkowski addition and the more modern 2-addition.

In the final section of the paper we consider the harmonic and geometric means of convex bodies and convex functions, and discuss their concavity properties. Once again, we find that in some problems the 2-addition of convex bodies behaves even better than the Minkowski addition.

Keywords: polarity; Legendre transform; summand; geometric mean

1. Introduction

Denote by $\mathrm{Cvx}(\mathbb{R}^n)$ the class of all convex and lower semi-continuous functions $\varphi : \mathbb{R}^n \to (-\infty, \infty]$. If $\varphi_1, \varphi_2 \in \mathrm{Cvx}(\mathbb{R}^n)$ and $\lambda > 0$ we define the functions $\varphi_1+\varphi_2, \lambda\varphi_1 \in \mathrm{Cvx}(\mathbb{R}^n)$ in the obvious, pointwise, way:

$$(\varphi_1+\varphi_2)(x) = \varphi_1(x)+\varphi_2(x)$$
$$(\lambda\varphi)(x) = \lambda\varphi(x).$$

Note that we write $\lambda\varphi$ and not $\lambda\cdot\varphi$, as we reserve the notation $\lambda\cdot\varphi$ for another operation to be defined shortly. The set $\mathrm{Cvx}(\mathbb{R}^n)$ is obviously a cone with respect to these operations.

Given $\varphi_1, \varphi_2 \in \mathrm{Cvx}(\mathbb{R}^n)$ we will write $\varphi_1 \le \varphi_2$ if $\varphi_1(x) \le \varphi_2(x)$ for all $x \in \mathbb{R}^n$. For a function $\varphi \in \mathrm{Cvx}(\mathbb{R}^n)$ the Legendre transform of φ is defined by

$$\varphi^*(x) = \sup_{y\in\mathbb{R}^n} \left(\langle x, y\rangle - \varphi(y)\right),$$

where $\langle\cdot,\cdot\rangle$ denotes the standard Euclidean product on $\mathbb{R}^n$.

As it turns out, once the order $\leq$ is given the Legendre transform does not need to be defined explicitly, but emerges naturally as the "duality transform" on $\mathrm{Cvx}(\mathbb{R}^n)$. More explicitly, in [2] and [4] Artstein-Avidan and Milman proved the following:

Theorem 1.1. *Assume $\mathcal{T} : \mathrm{Cvx}(\mathbb{R}^n) \to \mathrm{Cvx}(\mathbb{R}^n)$ satisfies the following two properties:*

- *$\mathcal{T}$ is an involution, i.e. $\mathcal{T}(\mathcal{T}\varphi) = \varphi$ for all $\varphi \in \mathrm{Cvx}(\mathbb{R}^n)$.*
- *$\mathcal{T}$ is order reversing, i.e. $\varphi_1 \leq \varphi_2$ implies that $\mathcal{T}\varphi_1 \geq \mathcal{T}\varphi_2$.*

Then $\mathcal{T}$ is the Legendre transform up to linear terms. Explicitly, there exists a constant $C \in \mathbb{R}$, a vector $v \in \mathbb{R}^n$, and an invertible symmetric linear transformation $B \in GL(n)$ such that

$$(\mathcal{T}\varphi)(x) = \varphi^*(Bx + v) + \langle x, v\rangle + C.$$

Even though our main theorems will be stated for convex functions, we will also be interested in convex sets. Denote by $\mathcal{K}_0^n$ the class of all closed and convex sets $K \subseteq \mathbb{R}^n$ such that $0 \in K$. One way to produce the structure of an ordered cone on $\mathcal{K}_0^n$ is by embedding $\mathcal{K}_0^n$ into $\mathrm{Cvx}(\mathbb{R}^n)$. The most useful such embedding sends $K \in \mathcal{K}_0^n$ to its support function $h_K \in \mathrm{Cvx}(\mathbb{R}^n)$, defined by

$$h_K(x) = \sup_{y\in K} \langle x, y\rangle .$$

Hence we define $K_1 + K_2$ by the relation $h_{K_1+K_2} = h_{K_1} + h_{K_2}$, λK by the relation $h_{\lambda K} = \lambda h_K$, and $K_1 \leq K_2$ by the inequality $h_{K_1} \leq h_{K_2}$. Of course, one can give more direct definitions: $K_1 + K_2$ is the Minkowski addition

$$K_1 + K_2 = \{x_1 + x_2 :\ x_1 \in K_1,\ x_2 \in K_2\},$$

(or, to be completely rigorous, the closure of this set). λK is the dilation

$$\lambda K = \{\lambda x :\ x \in K\},$$

and $K_1 \leq K_2$ if and only if $K_1 \subseteq K_2$.

Like in the case of convex functions, the order $\subseteq$ automatically produces a duality transform on $\mathcal{K}_0^n$, which is the polarity transform

$$K \mapsto K^\circ = \{x \in \mathbb{R}^n :\ \langle x, y\rangle \leq 1 \text{ for all } y \in K\}.$$

More formally:

Theorem 1.2. *Assume $\mathcal{T} : \mathcal{K}_0^n \to \mathcal{K}_0^n$ is an order reversing involution. Then there exists an invertible symmetric linear transformation $B \in GL(n)$ such that $\mathcal{T}K = BK^\circ$ for all $K \in \mathcal{K}_0^n$.*

Different people proved such a theorem for different classes of convex sets. For the class $\mathcal{K}_0^n$ the theorem was proven by Artstein-Avidan and Milman in [3]. Similar theorems on different classes were proven by Gruber ([10]) and by Brczky and

Schneider ([6]) – as the conditions of the theorem are so weak, slightly changing the class changes the proof almost completely.

Let us remark that other embeddings of $\mathcal{K}_0^n$ into $\mathrm{Cvx}(\mathbb{R}^n)$ produce other operations on $\mathcal{K}_0^n$. For example, one may fix $p \geq 1$ and map $K \in \mathcal{K}_0^n$ to $h_K^p \in \mathrm{Cvx}(\mathbb{R}^n)$. This embedding gives us the p-addition $K_1 +_p K_2$ which is defined implicitly by $h^p_{K_1 +_p K_2} = h^p_{K_1} + h^p_{K_2}$, and the p-homothety $\lambda \cdot_p K$ which is defined by $h^p_{\lambda \cdot_p K} = \lambda h^p_K$ (of course, $\lambda \cdot_p K = \lambda^{1/p} K$). The p-addition of convex bodies was introduced by Firey ([8]) and studied extensively by Lutwak ([13],[14]). Notice that the induced "p-order" is still the regular inclusion, so the natural duality transform remains unchanged.

The structure of an ordered cone with a duality transform appears often in mathematics, including in less geometric settings. The simplest example is the positive real numbers $\mathbb{R}_+$ themselves, with the usual addition, multiplication and order, and with the inversion $x \mapsto \frac{1}{x}$ as a duality. Another algebraic example is the class $\mathcal{M}_+^n$ of $n \times n$ positive-definite matrices. Here the addition and multiplication by scalar are the obvious choices, and the order is the matrix order $\preceq$, that is $M_1 \preceq M_2$ if $M_2 - M_1$ is positive definite. The duality is the matrix inversion $M \mapsto M^{-1}$.

The main goal of this paper is to observe some surprising similarities between the algebraic classes of numbers and matrices and the geometric classes of convex functions and convex sets. We will think of the dual function φ^* or the dual body K° as the "inverses" φ^{-1} and K^{-1}, and this intuition will allow us to conjecture new results in convexity. Once conjectured, these results are often not difficult to prove.

To illustrate this point of view let us consider one known example. For two numbers $x, y > 0$ the harmonic mean of x and y is $\left(\frac{x^{-1}+y^{-1}}{2}\right)^{-1}$, and it is well known that the harmonic mean is always smaller than the arithmetic mean, that is

$$\frac{x+y}{2} \geq \left(\frac{x^{-1}+y^{-1}}{2}\right)^{-1}.$$

It is fairly easy to prove a similar result for positive-definite matrices: for every $M, N \in \mathcal{M}_+^n$ we have

$$\frac{M+N}{2} \succeq \left(\frac{M^{-1}+N^{-1}}{2}\right)^{-1}.$$

What are the geometric analogs of this algebraic result? For convex sets, Firey established in [7] that for every $K, T \in \mathcal{K}_0^n$ one has

$$\frac{K+T}{2} \supseteq \left(\frac{K^\circ+T^\circ}{2}\right)^{\circ}.$$

In fact, in his paper Firey writes that "This may be viewed as the analogue, for convex bodies, of the theorem of the arithmetic and harmonic means for positive numbers".

Finally, for convex functions, the analogous result is more or less folklore. For every "reasonable" $\varphi, \psi \in \text{Cvx}(\mathbb{R}^n)$ one has

$$(\varphi^* + \psi^*)^*(x) = (\varphi \Box \psi)(x) = \inf_{y+z=x} (\varphi(y) + \psi(z)).$$

The operation $\varphi \Box \psi$ is known as the infimal convolution of φ and ψ. The word "reasonable" above comes from the fact that $\varphi \Box \psi$ need not be lower semi-continuous, even if φ and ψ are. As this technicality will have no impact on our discussion we ignore it, and refer the reader to section 1.6.2 of [18] for a more careful discussion of this topic.

Similarly, for every $\varphi \in \text{Cvx}(\mathbb{R}^n)$ and every $\lambda > 0$

$$(\lambda\varphi)^*(x) = \lambda\varphi^*\left(\frac{x}{\lambda}\right).$$

As is sometimes customary, we will write $(\lambda\varphi)^* = \lambda \cdot \varphi^*$ and warn the reader not to confuse the two possible multiplications.

Once these facts are established we evidently have

$$\frac{\varphi + \psi}{2} \geq \left(\frac{\varphi^* + \psi^*}{2}\right)^*,$$

since

$$\begin{aligned}\left(\frac{\varphi^* + \psi^*}{2}\right)^*(x) &= \left[\frac{1}{2} \cdot (\varphi \Box \psi)\right](x) = \frac{1}{2}(\varphi \Box \psi)(2x). \\ &= \frac{1}{2} \inf_{y+z=2x} (\varphi(y) + \psi(z)) \leq \frac{\varphi(x) + \psi(x)}{2}\end{aligned}$$

Hence we see that in this example the analogy is perfect, and the same basic inequality holds for positive numbers, positive-definite matrices, convex sets and convex functions.

Another interesting example of treating K° and φ^* as inverses can be found in a recent paper by Molchanov ([16]), where he constructs continued fractions of convex sets and convex functions and uses them to solve "body valued quadratic equations".

In the next two sections we will explore two new results that follow from the same philosophy. In the next section we will prove a simple yet surprising identity, with an unexpected application to Santal type inequalities. In Section 3 we will return to means of convex sets and convex functions and discuss their concavity properties.

2. A new identity

Notice the following trivial fact: for every $x > 0$ one has

$$\frac{1}{x+1} + \frac{1}{\frac{1}{x}+1} = 1.$$

From here it is easy to prove a similar result for matrices: for every $M \in \mathcal{M}_+^n$ one has

$$(M+I)^{-1} + \left(M^{-1}+I\right)^{-1} = I,$$

where I denotes of course the identity matrix.

In order to understand the counterpart of this algebraic fact for convex functions, we need to understand which function $\delta \in \mathrm{Cvx}\left(\mathbb{R}^n\right)$ plays the role of 1 or I. The number 1 can be characterized as the only positive solution of $x^{-1} = x$. Similarly, I may be characterized as the only matrix $X \in \mathcal{M}_+^n$ such that $X^{-1} = X$. Hence we define δ to be the only solution of the equation $\varphi^* = \varphi$. This solution is $\delta(x) = \frac{1}{2}\left|x\right|^2$, where $|\cdot|$ denotes the Euclidean norm. Our theorem then reads:

Theorem 2.1. *For every $\varphi \in \mathrm{Cvx}\left(\mathbb{R}^n\right)$ one has*

$$(\varphi+\delta)^* + (\varphi^*+\delta)^* = \delta.$$

Proof. Write $\rho = (\varphi+\delta)^* + (\varphi^*+\delta)^*$. On the one hand, directly by the definition of the Legendre transform, we may write

$$\begin{aligned}\rho(x) &= \sup_y \left(\langle x, y\rangle - \varphi(y) - \frac{|y|^2}{2}\right) + \sup_z \left(\langle x, z\rangle - \varphi^*(z) - \frac{|z|^2}{2}\right)\\ &= \sup_{y,z}\left(\langle x, y+z\rangle - \varphi(y) - \varphi^*(z) - \frac{|y|^2}{2} - \frac{|z|^2}{2}\right).\end{aligned}$$

On the other hand, we have

$$\begin{aligned}(\varphi^*+\delta)^* &= (\varphi^*+\delta^*)^* = \varphi\square\delta\\ (\varphi+\delta)^* &= \left((\varphi^*)^*+\delta^*\right)^* = \varphi^*\square\delta.\end{aligned}$$

Hence we can also write

$$\begin{aligned}\rho(x) &= \inf_y\left(\varphi(y) + \frac{|x-y|^2}{2}\right) + \inf_z\left(\varphi^*(z) + \frac{|x-z|^2}{2}\right)\\ &= \inf_{y.z}\left(\varphi(y) + \frac{|x|^2 - 2\langle x,y\rangle + |y|^2}{2} + \varphi^*(z) + \frac{|x|^2 - 2\langle x,z\rangle + |z|^2}{2}\right)\\ &= |x|^2 - \sup_{y,z}\left(\langle x, y+z\rangle - \varphi(y) - \varphi^*(z) - \frac{|y|^2}{2} - \frac{|z|^2}{2}\right) = |x|^2 - \rho(x).\end{aligned}$$

From here we immediately obtain $\rho(x) = \frac{1}{2}\left|x\right|^2 = \delta(x)$ and the proof is complete.□

Let us state one corollary of this theorem:

Theorem 2.2. *Let γ_n be the standard Gaussian measure on $\mathbb{R}^n$, i.e.*

$$\gamma_n(A) = \frac{1}{(2\pi)^{n/2}}\int_A e^{-|x|^2/2}\mathrm{d}x.$$

Then for every $\varphi \in \text{Cvx}(\mathbb{R}^n)$ one has

$$\int e^{-\varphi} d\gamma_n \cdot \int e^{-\varphi^*} d\gamma_n \le \left(\int e^{-\delta} d\gamma_n \right)^2 .$$

Before we prove Theorem 2.2, let us put it in some perspective. In [11], Klartag proves that if $\varphi \in \text{Cvx}(\mathbb{R}^n)$ is even, then for every even log-concave measure μ on $\mathbb{R}^n$ one has

$$\int e^{-\varphi} d\mu \cdot \int e^{-\varphi^*} d\mu \le \left(\int e^{-\delta} d\mu \right)^2 .$$

Our theorem is just the special case $\mu = \gamma_n$. Surprisingly, however, our theorem holds for functions φ that are not necessarily even, and there is no need to translate φ.

Proof. Define $\psi(x) = \frac{1}{2}\varphi\left(\sqrt{2}x\right)$, and notice that $\psi^*(x) = \frac{1}{2}\varphi^*\left(\sqrt{2}x\right)$. Applying Theorem 2.1 for ψ instead of φ we have

$$(\psi + \delta)^* + (\psi^* + \delta)^* = \delta,$$

and by applying duality to both sides we obtain

$$(\psi + \delta) \,\square\, (\psi^* + \delta) = \left[(\psi + \delta)^* + (\psi^* + \delta)^* \right]^* = \delta^* = \delta.$$

Define $f, g \in \text{Cvx}(\mathbb{R}^n)$ by $f = 2 \cdot (\psi + \delta)$ and $g = 2 \cdot (\psi^* + \delta)$. Explicitly this means that

$$f(x) = 2\left(\psi\left(\frac{x}{2}\right) + \delta\left(\frac{x}{2}\right)\right) = \varphi\left(\frac{x}{\sqrt{2}}\right) + \delta\left(\frac{x}{\sqrt{2}}\right)$$
$$g(x) = 2\left(\psi^*\left(\frac{x}{2}\right) + \delta\left(\frac{x}{2}\right)\right) = \varphi^*\left(\frac{x}{\sqrt{2}}\right) + \delta\left(\frac{x}{\sqrt{2}}\right).$$

Since $\left(\frac{1}{2} \cdot f\right) \square \left(\frac{1}{2} \cdot g\right) = \delta$ we may apply the Prkopa-Leindler (see, e.g., Theorem 7.1.2 of [18]) inequality and obtain

$$\int e^{-f(x)} \mathrm{d}x \cdot \int e^{-g(x)} \mathrm{d}x \le \left(\int e^{-\delta(x)} \mathrm{d}x \right)^2 .$$

Applying the change of variables $x = \sqrt{2}y$ to both sides and dividing both sides by $(2\pi)^n$ one obtains the result. □

Before moving on to discuss convex sets, let us show one possible extension of Theorem 2.1. For functions $\varphi_1, \varphi_2 \in \text{Cvx}(\mathbb{R}^n)$ we say that φ_1 is a summand if φ_2 if there exists $\varphi_3 \in \text{Cvx}(\mathbb{R}^n)$ such that $\varphi_1 + \varphi_3 = \varphi_2$. In other words, φ_1 is a summand of φ_2 if $\varphi_2 - \varphi_1 \in \text{Cvx}(\mathbb{R}^n)$. Theorem 2.1 immediately implies that if δ is a summand of φ, then φ^* is a summand of δ. As the next theorem shows, the converse of this statement is also true:

Theorem 2.3. *Fix $\varphi \in \text{Cvx}(\mathbb{R}^n)$. Then δ is a summand of φ if and only if φ^* is a summand of δ.*

Proof. As mentioned above, the "only if" follows immediately from Theorem 2.1. Indeed, if δ is a summand of φ then $\varphi = \psi + \delta$ for some $\psi \in \mathrm{Cvx}(\mathbb{R}^n)$. It follows that

$$\varphi^* + (\psi^* + \delta)^* = (\psi + \delta)^* + (\psi^* + \delta)^* = \delta$$

and φ^* is a summand of δ.

For the other implication, assume that φ^* is a summand of δ, so $\delta - \varphi^*$ is convex. Since φ is lower semi-continuous and δ is continuous the function $\varphi - \delta$ is also lower semi-continuous, and we only need to prove that $\varphi - \delta$ is convex. Convexity of $\delta - \varphi^*$ implies that for every $z, w \in \mathbb{R}^n$ and every $0 < \lambda < 1$ one has

$$\delta(\lambda z + (1-\lambda)w) - \varphi^*(\lambda z + (1-\lambda)w) \le \lambda(\delta(z) - \varphi^*(z)) + (1-\lambda)(\delta(w) - \varphi^*(w)).$$

Simplifying, this implies that

$$\varphi^*(\lambda z + (1-\lambda)w) + \frac{\lambda(1-\lambda)}{2}|z-w|^2 \ge \lambda\varphi^*(z) + (1-\lambda)\varphi^*(w).$$

Now, fix $x, y \in \mathbb{R}^n$ and $0 < \lambda < 1$. Writing $\psi = \varphi - \delta = \varphi^{**} - \delta$, and denoting $\lambda\psi(x) + (1-\lambda)\psi(y)$ by ψ_λ, we have

$$\begin{aligned}
\psi_\lambda &= \lambda \sup_z [\langle x, z\rangle - \varphi^*(z)] + (1-\lambda)\sup_w [\langle y, w\rangle - \varphi^*(w)] - \lambda\delta(x) - (1-\lambda)\delta(y) \\
&= \sup_{z,w} [\lambda\langle x, z\rangle + (1-\lambda)\langle y, w\rangle - (\lambda\varphi^*(z) + (1-\lambda)\varphi^*(w))] - \lambda\delta(x) - (1-\lambda)\delta(y) \\
&\ge \sup_{z,w}\left[\lambda\langle x, z\rangle + (1-\lambda)\langle y, w\rangle - \varphi^*(\lambda z + (1-\lambda)w) - \frac{\lambda(1-\lambda)}{2}|z-w|^2\right] \\
&\quad -\lambda\delta(x) - (1-\lambda)\delta(y)
\end{aligned}$$

To simplify this expression we make a change of variables $u = \lambda z + (1-\lambda)w$ and $v = z - w$. We have $z = u + (1-\lambda)v$ and $w = u - \lambda v$, so

$$\begin{aligned}
\psi_\lambda \ge \sup_{u,v} \Bigg[& \lambda\langle x, u\rangle + \lambda(1-\lambda)\langle x, v\rangle + (1-\lambda)\langle y, u\rangle - \lambda(1-\lambda)\langle y, v\rangle \\
& - \frac{\lambda(1-\lambda)}{2}|v|^2 - \lambda\frac{|x|^2}{2} - (1-\lambda)\frac{|y|^2}{2} - \varphi^*(u)\Bigg] \\
= \sup_{u,v} \Bigg[& \langle \lambda x + (1-\lambda)y, u\rangle - \frac{\lambda(1-\lambda)}{2}(x-y-v)^2 - \frac{|\lambda x + (1-\lambda)y|^2}{2} - \varphi^*(u)\Bigg].
\end{aligned}$$

Now the supremum is obviously attained when $v = x - y$ and we are left with

$$\begin{aligned}
\psi_\lambda &\ge \sup_u [\langle \lambda x + (1-\lambda)y, u\rangle - \varphi^*(u)] - \frac{|\lambda x + (1-\lambda)y|^2}{2} \\
&= \varphi^{**}(\lambda x + (1-\lambda)y) - \delta(\lambda x + (1-\lambda)y) \\
&= \psi(\lambda x + (1-\lambda)y).
\end{aligned}$$

Hence ψ is convex, and the proof is complete. □

Remark 2.1. If one ignores questions of smoothness, it is possible to give a more transparent proof of Theorem 2.3. Indeed, assume that φ and φ^* are both smooth. Convexity of $\delta - \varphi^*$ is then equivalent to the requirement that $\nabla^2\varphi^*(y) \preceq I$ for all $y \in \mathbb{R}^n$. Similarly, convexity of $\varphi - \delta$ is equivalent to $\nabla^2\varphi(x) \succeq I$ for all $x \in \mathbb{R}^n$. Since

$$\nabla^2\varphi^*(y) = \left[\nabla^2\varphi\left(\nabla\varphi^*(y)\right)\right]^{-1},$$

and since the map $A \mapsto A^{-1}$ is order reversing on $\mathcal{M}_+^n$, the theorem follows immediately.

Remark 2.2. It follows from the above discussion that the map

$$\varphi \mapsto (\varphi^* + \delta)^*$$

is an order preserving bijection between $\mathrm{Cvx}\left(\mathbb{R}^n\right)$ and the class of convex summands of δ. The standard duality relation $\psi \leftrightarrow \psi^*$ on $\mathrm{Cvx}\left(\mathbb{R}^n\right)$ corresponds under this map the the obvious duality $\varphi \leftrightarrow \delta - \varphi$ on the class of summands. Furthermore, using this map one can deduce from Theorem 1.1 a theorem characterizing *all* order-reversing involutions on the class of summands.

We conclude this section by discussing convex sets, for which the situation is a bit more complicated. Notice that in Theorem 1.2 there was no mention of an addition operation on $\mathcal{K}_0^n$ - the polarity map $K \mapsto K^\circ$ is uniquely and naturally defined given no other structure except the inclusion. So, while there is only one reasonable candidate for the inversion operation, there are many possible candidates for the addition operation. Usually in convexity the natural choice is the Minkowski addition, but it turns out that the 2-addition is better behaved in our case:

Theorem 2.4. *For every $K \in \mathcal{K}_0^n$ one has*

$$(K +_2 D)^\circ +_2 (K^\circ +_2 D)^\circ = D.$$

Here D denotes the unit Euclidean ball and $+_2$ is the 2-sum as defined in the introduction.

Proof. Simply apply Theorem 2.1 to $\varphi = \frac{1}{2}h_K^2$. Using the fact that $\left(\frac{1}{2}h_A^2\right)^* = \frac{1}{2}h_{A^\circ}^2$, the definition of the 2-addition and the fact that $\delta = \frac{1}{2}h_D^2$, one gets

$$\frac{1}{2}h^2_{(K+_2D)^\circ +_2}\frac{1}{2}h^2_{(K^\circ+_2D)^\circ} = \frac{1}{2}h_D^2,$$

and the result follows. □

Since the Minkowski addition appears much more often in convexity than the 2-addition, it may seem reasonable to conjecture that Theorem 2.4 remains true if the 2-sum is replaced by the more conventional 1-sum:

$$(K + D)^\circ + (K^\circ + D)^\circ = D. \tag{2.1}$$

Numerical evidence seems to suggest that this is true when K is a "symmetric enough" planar convex set. For example, when taking the unit square $Q = [-1,1]^2$ we indeed have

$$(Q + D)^\circ + (Q^\circ + D)^\circ = D,$$

as can be seen in Figure 1.

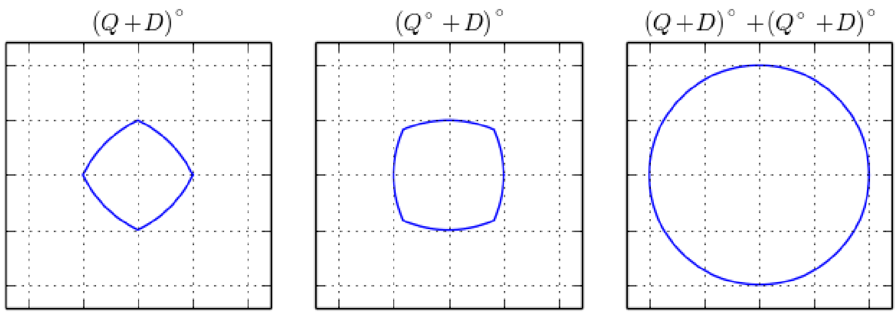

Fig. 1. Checking the identity for $Q = [-1,1]^2$

However, it turns out that in general the identity (2.1) is false. As a counterexample, take $K = \{0\} \times \mathbb{R}^{n-1} \in \mathcal{K}_0^n$. In this case a direct computation shows that

$$T = (K + D)^\circ + (K^\circ + D)^\circ = [-1,1] \times D_{n-1},$$

where D_{n-1} is the unit Euclidean ball in dimension $n-1$.

Notice that we still have $T \subseteq \sqrt{2}D$, and this is not a coincidence: for every $A, B \in \mathcal{K}_0^n$ one has

$$A +_2 B \subseteq A + B \subseteq \sqrt{2} \cdot (A +_2 B),$$

so from Theorem 2.4 we immediately obtain that for every $K \in \mathcal{K}_0^n$ one has

$$\frac{1}{\sqrt{2}} D \subseteq (K + D)^\circ + (K^\circ + D)^\circ \subseteq \sqrt{2}D.$$

I do not have a satisfactory explanation for the fact that (2.1) seems to hold for the square and other planar bodies with many symmetries.

Finally, we note that Theorem 2.3 also yields a corollary for convex sets, in exactly the same way that Theorem 2.1 yields Theorem 2.4. We say that $A \in \mathcal{K}_0^n$ is a 2-summand of $B \in \mathcal{K}_0^n$ if there exists $C \in \mathcal{K}_0^n$ such that $A +_2 C = B$. The result then reads:

Corollary 2.1. *Fix $K \in \mathcal{K}_0^n$. Then D is a 2-summand of K if and only if K° is a 2-summand of D.*

When 2-summands are replaced with 1-summands (simply known as summands) the situation is different. In one direction, Hug proved in his Habilitationsschrift (habilitation thesis) that if K° is a summand of D then D is a summand of K. A proof of this result now appears as Proposition A.3 in [9]. The opposite direction, however, is false. To see this, notice that

$$D + \left(\{0\} \times \mathbb{R}^{n-1}\right) = [-1,1] \times \mathbb{R}^{n-1},$$

so D is a summand of $K = [-1,1] \times \mathbb{R}^{n-1}$, even though $K^\circ = [-1,1] \times \{0\}^{n-1}$ is not a summand of D.

3. Concavity of means

For positive real numbers $x, y > 0$, one may consider their arithmetic mean $\frac{1}{2}(x+y)$, their harmonic mean $\left[\frac{1}{2}\left(x^{-1}+y^{-1}\right)\right]^{-1}$, and their geometric mean $\sqrt{xy}$. Viewed as functions $(\mathbb{R}_+)^2 \to \mathbb{R}_+$, it is easy to check that all three means are concave functions.

For positive-definite matrices, the situation is a bit more complicated. The arithmetic and harmonic mean are easy to define as $A(M,N) = \frac{1}{2}(M+N)$ and $H(M,N) = \left[\frac{1}{2}\left(M^{-1}+N^{-1}\right)\right]^{-1}$. The geometric mean however is less obvious – $\sqrt{MN}$ is not the correct definition, as the product MN is in general not positive-definite and so its square root is not well defined. It turns out that the correct definition is

$$G(M,N) = M^{1/2}\left(M^{-1/2}NM^{-1/2}\right)^{1/2}M^{1/2}.$$

This definition was first given by Pusz and Woronowicz ([17]), who also proved that G is concave. See [1] for a more readable proof of the concavity of H and G, and see [12] for a survey explaining why G is the "correct" definition of the geometric mean of positive-definite matrices.

For convex sets $K, T \in \mathcal{K}_0^n$, their arithmetic mean is of course $A(K,T) = \frac{1}{2}(K+T)$ and their harmonic mean is $H(K,T) = \left(\frac{1}{2}\left(K^\circ + T^\circ\right)\right)^\circ$, as explained in the introduction. The geometric mean of convex sets is a more delicate construction that was defined in [15]. We say that $K \in \mathcal{K}_0^n$ is a convex body (and not just a convex set) if K is compact and contains 0 at its interior. For given bodies $K, T \in \mathcal{K}_0^n$, one looks at the sequences $\{A_n\}_{n=0}^\infty$ and $\{H_n\}_{n=0}^\infty$ defined by

$$\begin{aligned} A_0 &= K & H_0 &= T \\ A_{n+1} &= \frac{A_n + H_n}{2} & H_{n+1} &= \left(\frac{A_n^\circ + H_n^\circ}{2}\right)^\circ. \end{aligned}$$

The sequences $\{A_n\}_{n=1}^\infty$ and $\{H_n\}_{n=1}^\infty$ converge to a common limit, which we call the geometric mean of K and T and denote by $G(K,T)$.

In [15] it is established that the geometric mean of convex bodies shares many of the basic properties of the geometric mean of numbers and matrices. As it turns out, concavity is *not* one of those properties:

Theorem 3.1. *For convex sets, neither the harmonic mean H nor the geometric mean G are concave functions of their arguments*

Proof. We will give counterexamples in dimension $n = 2$. Let us denote the unit ball of the ℓ_p norm on $\mathbb{R}^2$ by B_p^2 (so $B_\infty^2 = [-1,1]^2$ and $B_2^2 = D$).

For the harmonic mean, take

$$\begin{array}{ll} A_1 = \mathbb{R} \times \{0\} & A_2 = B_\infty^2 \\ B_1 = B_\infty^2 & B_2 = \{0\} \times \mathbb{R}. \end{array}$$

Then

$$H\left(\frac{A_1+B_1}{2}, \frac{A_2+B_2}{2}\right) = H\left(\mathbb{R} \times \left[-\frac{1}{2}, \frac{1}{2}\right], \left[-\frac{1}{2}, \frac{1}{2}\right] \times \mathbb{R}\right) = B_1^2,$$

while

$$\frac{H(A_1,A_2)+H(B_1,B_2)}{2} = \frac{([-2,2]\times\{0\}) + (\{0\}\times[-2,2])}{2} = B_\infty^2.$$

Since $B_1^2 \not\supseteq B_\infty^2$ we found our counterexample. If one wants a counterexample with proper bodies, just approximate A_1 and B_2 with such bodies in an arbitrary way.

For the geometric mean the computations are a bit more complicated. We fix $\epsilon > 0$ and define

$$\begin{array}{ll} A_1 = \left[-\frac{1}{\epsilon}, \frac{1}{\epsilon}\right] \times \left[\epsilon^3, \epsilon^3\right] & A_2 = [-\epsilon, \epsilon]^2 \\ B_1 = [-\epsilon, \epsilon]^2 & B_2 = \left[\epsilon^3, \epsilon^3\right] \times \left[-\frac{1}{\epsilon}, \frac{1}{\epsilon}\right]. \end{array}$$

In [15] it is proved that

$$d\left(G\left([-R,R] \times \left[-\frac{1}{R}, \frac{1}{R}\right], \left[-\frac{1}{R}, \frac{1}{R}\right] \times [-R,R]\right), B_2^2\right) \le \sqrt{1+\frac{1}{R^2}}, \tag{3.1}$$

where d denotes the geometric distance between convex bodies. Denoting $R = \frac{1}{\epsilon}$ and applying the linear map $T_\epsilon(x,y) = \left(x, \epsilon^2 y\right)$ to all of the bodies we obtain

$$d\left(G\left(A_1, A_2\right), T_\epsilon\left(B_2^2\right)\right) \le \sqrt{1+\epsilon^2},$$

which implies that $G(A_1, A_2) \to [-1,1] \times \{0\}$ as $\epsilon \to 0$. Similarly we have $G(B_1, B_2) \to \{0\} \times [-1,1]$ as $\epsilon \to 0$.

On the other hand, it follows again from (3.1) that

$$G\left(\frac{A_1+A_2}{2}, \frac{B_1+B_2}{2}\right) = \frac{1+\epsilon^2}{2} G\left(\left[-\frac{1}{\epsilon}, \frac{1}{\epsilon}\right] \times [-\epsilon,\epsilon], [-\epsilon,\epsilon] \times \left[-\frac{1}{\epsilon}, \frac{1}{\epsilon}\right]\right) \to \frac{1}{2} B_2^2$$

as $\epsilon \to 0$.

We see that the concavity inequality

$$G\left(\frac{A_1+A_2}{2}, \frac{B_1+B_2}{2}\right) \supseteq \frac{G(A_1,A_2)+G(B_1,B_2)}{2}$$

becomes, as $\epsilon \to 0$, the incorrect inequality $\frac{1}{2}B_2^2 \supseteq \frac{1}{2}B_\infty^2$. It follows that for some small enough $\epsilon > 0$ we obtain the desired counterexample. □

Things become better when one consider the geometric mean of convex functions instead of convex bodies. For given $\varphi, \psi \in \mathrm{Cvx}(\mathbb{R}^n)$ we look at the sequences $\{\alpha_n\}_{n=0}^\infty$, $\{\eta_n\}_{n=0}^\infty$ defined by

$$\begin{aligned} \alpha_0 &= \varphi & \eta_0 &= \psi \\ \alpha_{n+1} &= \frac{\alpha_n + \eta_n}{2} & \eta_{n+1} &= \left(\frac{\alpha_n^* + \eta_n^*}{2}\right)^* . \end{aligned}$$

If the functions φ, ψ are everywhere finite then these sequences will converge to a common limit, with essentially the same proof as the proof of convergence for bodies: the sequence $\{\alpha_n\}_{n=1}^\infty$ is decreasing and bounded from below by η_1, the sequence $\{\eta_n\}_{n=1}^\infty$ is increasing and bounded from above by α_1, so both sequences converge (pointwise). The relation $\alpha_{n+1} = \frac{1}{2}(\alpha_n + \eta_n)$ and the fact that our functions our finite implies that the limits coincide. We denote this limit by $G(\varphi, \psi)$ and call it the geometric mean of φ and ψ.

Perhaps the most interesting case to consider is the case $\varphi = h_K^p$ and $\psi = h_T^p$ for some convex bodies K and T and some $p \geq 1$. In fact, when $p = 2$ this was studied by Asplund ([5]) much before [15].

As it turns out, means of convex functions are better behaved than means of convex bodies:

Theorem 3.2.

(1) The harmonic mean of convex functions is concave. More explicitly, fix $\varphi_0, \varphi_1, \psi_0, \psi_1 \in \mathrm{Cvx}(\mathbb{R}^n)$ and $0 < \lambda < 1$. Define $\varphi_\lambda = (1-\lambda)\varphi_0 + \lambda\varphi_1$ and $\psi_\lambda = (1-\lambda)\psi_0 + \lambda\psi_1$. Then

$$H(\varphi_\lambda, \psi_\lambda) \supseteq (1-\lambda) \cdot H(\varphi_0, \psi_0) + \lambda H(\varphi_1, \psi_1).$$

(2) The geometric mean of convex functions is concave. In the same notations as above, we have

$$G(\varphi_\lambda, \psi_\lambda) \supseteq (1-\lambda) \cdot G(\varphi_0, \psi_0) + \lambda G(\varphi_1, \psi_1)$$

whenever all the geometric means in this expression are well defined.

Proof.

(1) As explained in the introduction,

$$[H(\varphi, \psi)](x) = \left[\frac{1}{2} \cdot (\varphi \Box \psi)\right](x) = \frac{1}{2} \inf_{y+z=2x} (\varphi(y) + \psi(z)).$$

Hence we have

$$\begin{aligned} H\left(\varphi_\lambda,\psi_\lambda\right)(x) &= \frac{1}{2}\inf_{y+z=2x}\left[(1-\lambda)\varphi_0(y)+\lambda\varphi_1(y)+(1-\lambda)\psi_0(z)+\lambda\psi_1(z)\right] \\ &\geq \frac{1}{2}\inf_{\substack{y+z=2x\\ u+v=2x}}\left[(1-\lambda)\varphi_0(y)+\lambda\varphi_1(u)+(1-\lambda)\psi_0(z)+\lambda\psi_1(v)\right] \\ &= \frac{1}{2}\left[(1-\lambda)\inf_{y+z=2x}\left(\varphi_0(y)+\psi_0(z)\right)+\lambda\inf_{u+v=2x}\left(\varphi_1(u)+\psi_1(v)\right)\right] \\ &= \left[(1-\lambda)\,H(\varphi_0,\psi_0)+\lambda H\left(\varphi_1,\psi_1\right)\right](x), \end{aligned}$$

so H is concave.

(2) Consider the function $F:\mathrm{Cvx}\left(\mathbb{R}^n\right)^2\to\mathrm{Cvx}\left(\mathbb{R}^n\right)^2$ defined by

$$F\left(\varphi,\psi\right)=\left(\frac{\varphi+\psi}{2},\left(\frac{\varphi^*+\psi^*}{2}\right)^*\right).$$

By (1) we know that F is concave, meaning each of its components is concave. It is also trivial to see that each of the components of F is monotone increasing in its arguments.

From these two facts, it follows that $F^{(2)}=F\circ F$ is concave and increasing: we have

$$F(\varphi_\lambda,\psi_\lambda)\geq(1-\lambda)\cdot F\left(\varphi_0,\psi_0\right)+\lambda F(\varphi_1,\psi_1)$$

(again, $\geq$ means component-wise comparison), so

$$\begin{aligned} F^{(2)}\left(\varphi_\lambda,\psi_\lambda\right) &\geq F\left((1-\lambda)\cdot F\left(\varphi_0,\psi_0\right)+\lambda F(\varphi_1,\psi_1)\right) \\ &\geq (1-\lambda)\,F^{(2)}\left(\varphi_0,\psi_0\right)+\lambda F^{(2)}\left(\varphi_1,\psi_1\right) \end{aligned}$$

and $F^{(2)}$ is concave. The fact that $F^{(2)}$ is increasing is even easier.

Proceeding by induction, we see that $F^{(m)}=\underbrace{F\circ F\circ\cdots\circ F}_{m\text{ times}}$ is concave and increasing for every $m\geq 1$. But then

$$\widetilde{F}(\varphi,\psi)=\lim_{m\to\infty}F^{(m)}(\varphi,\psi)=(G(\varphi,\psi),G(\varphi,\psi))$$

is also concave and increasing. In particular, G is concave. □

Like in the previous section, the result for functions gives a result for 2-sum of convex bodies. For convex bodies K and T we define their 2-geometric mean $G_2(K,T)$ as the joint limit of

$$\begin{aligned} A_0 &= K & H_0 &= T \\ A_{n+1} &= \frac{A_n+_2 H_n}{\sqrt{2}} & H_{n+1} &= \left(\frac{A_n^\circ+_2 H_n^\circ}{\sqrt{2}}\right)^\circ. \end{aligned}$$

Prop 3.1. G_2 is 2-concave: For every convex bodies K_0,K_1,T_0 and T_1

$$G_2\left(K_\lambda,T_\lambda\right)\supseteq\sqrt{1-\lambda}G_2(K_0,T_0)+_2\sqrt{\lambda}G_2(K_1,T_1),$$

where $K_\lambda=\sqrt{1-\lambda}K_0+_2\sqrt{\lambda}K_1$ and $T_\lambda=\sqrt{1-\lambda}T_0+_2\sqrt{\lambda}T_1$.

Proof. Apply Theorem 3.2 to $\varphi_i=\frac{1}{2}h_{K_i}^2$, $\psi_i=\frac{1}{2}h_{T_i}^2$, $i=0,1$. □

Acknowledgement

I would like to thank Vitali Milman for many hours of illuminating discussions on these topics. I would also like to thank Shiri Artstein-Avidan and Bo'az Klartag for their useful suggestions regarding the results of Section 2. Finally, I would like to thank Rolf Schneider for referring me to the work of Daniel Hug and the paper [9].

The author is supported by ISF grant 826/13, BSF grant 2012111, and the Adams Fellowship Program of the Israel Academy of Sciences and Humanities.

References

[1] Tsuyoshi Ando. Concavity of certain maps on positive definite matrices and applications to Hadamard products. *Linear Algebra and its Applications*, 26:203–241, aug 1979.

[2] Shiri Artstein-Avidan and Vitali Milman. A characterization of the concept of duality. *Electronic Research Announcements in Mathematical Sciences*, 14:42–59, 2007.

[3] Shiri Artstein-Avidan and Vitali Milman. The concept of duality for measure projections of convex bodies. *Journal of Functional Analysis*, 254(10):2648–2666, may 2008.

[4] Shiri Artstein-Avidan and Vitali Milman. The concept of duality in convex analysis, and the characterization of the Legendre transform. *Annals of Mathematics*, 169(2):661–674, mar 2009.

[5] Edgar Asplund. Averaged norms. *Israel Journal of Mathematics*, 5(4):227–233, oct 1967.

[6] Károly J. Böröczky and Rolf Schneider. A characterization of the duality mapping for convex bodies. *Geometric and Functional Analysis*, 18(3):657–667, aug 2008.

[7] William J. Firey. Polar means of convex bodies and a dual to the Brunn-Minkowski theorem. *Canadian Journal of Mathematics*, 13:444–453, 1961.

[8] William J. Firey. p-means of convex bodies. *Mathematica Scandinavica*, 10:17–24, 1962.

[9] Ferenc Fodor, Daniel Hug, and Ines Ziebarth. The volume of random polytopes circumscribed around a convex body. *Mathematika*, 2015.

[10] Peter M. Gruber. The endomorphisms of the lattice of norms in finite dimensions. *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, 62(1):179–189, 1992.

[11] Bo'az Klartag. Marginals of Geometric Inequalities. In Vitali Milman and Gideon Schechtman, editors, *Geometric Aspects of Functional Analysis, Israel Seminar 2004-2005*, volume 1910 of *Lecture Notes in Mathematics*, pages 133–166. Springer, Berlin, Heidelberg, 2007.

[12] Jimmie Lawson and Yongdo Lim. The Geometric Mean , Matrices , Metrics , and More. *The American Mathematical Monthly*, 108(9):797–812, 2001.

[13] Erwin Lutwak. The Brunn-Minkowski-Firey theory I: Mixed volumes and the Minkowski problem. *Journal of Differential Geometry*, 38(1):131–150, 1993.

[14] Erwin Lutwak. The Brunn-Minkowski-Firey theory II: Affine and geominimal surface areas. *Advances in Mathematics*, 118(2):244–294, 1996.

[15] Vitali Milman and Liran Rotem. Non-standard constructions in convex geometry; geometric means of convex bodies. *Submitted.*

[16] Ilya Molchanov. Continued fractions built from convex sets and convex functions. *Communications in Contemporary Mathematics*, oct 2014.

[17] Wieslaw Pusz and Stanislaw L. Woronowicz. Functional calculus for sesquilinear

forms and the purification map. *Reports on Mathematical Physics*, 8(2):159–170, oct 1975.

[18] Rolf Schneider. *Convex Bodies: The Brunn-Minkowski Theory.* Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, second edition, 2014.