The reasons behind some classical constructions in analysis

V. Milman

Tel-Aviv University

In the memory of a great scientist and a good friend,
Aleksander (Olek) Pełczyński

June 2014, Będlewo, Poland

Instead of an Introduction: some geometric results

What should we call "duality"?

Consider the class $\mathrm{Cvx}(\mathbb{R}^n)$ of all *lower-semi-continuous* convex functions $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. The Legendre transform is the map

$$\mathcal{L}\varphi(x) = \sup_{y \in \mathbb{R}^n} \left[\langle x, y \rangle - \varphi(y) \right].$$

Of course, there are many "Legendre transforms": We may select 0 of the space, a scalar product and a shift for a function.

Theorem (Artstein–Milman)

1. Assume $T : \mathrm{Cvx}(\mathbb{R}^n) \to \mathrm{Cvx}(\mathbb{R}^n)$ satisfies:
   (a) $T \cdot T\varphi = \varphi$ (for any $\varphi \in \mathrm{Cvx}(\mathbb{R}^n)$);
   (b) $\varphi \leq \psi$ implies $T\varphi \geq T\psi$.

   Then T is a Legendre transform. It means that $\mathrm{Cvx}(\mathbb{R}^n)$ has a unique duality structure!

Let us embed $\mathcal{K}(\mathbb{R}^n) = \{K \subseteq \mathbb{R}^n \mid \text{closed convex}\}$ into $\mathrm{Cvx}\,(\mathbb{R}^n)$ by "convex characteristic" functions:

$$K \longrightarrow \mathbf{1}_K^\infty = \begin{cases} 0 & x \in K, \\ +\infty & x \notin K. \end{cases}$$

For $K \in \mathcal{K}_0(\mathbb{R}^n) = \{K \in \mathcal{K}\,(\mathbb{R}^n) \mid 0 \in K\}$, define its gauge function (or Minkowski functional $M\,(\mathbf{1}_K^\infty)$) – 1-homogeneous convex function $\|x\|_K$, "generalized" norm, s.t. $K = \{x \in \mathbb{R}^n \mid \|x\|_K \le 1\}$.

Let $\mathcal{H}_0 = \{\|x\|_K \mid K \in \mathcal{K}_0(\mathbb{R}^n)\}$. Define the **Minkowski map**

$$M(\mathbf{1}_K^\infty) = \|x\|_K \in \mathcal{H}_0$$

and the **support map**

$$S(\mathbf{1}_K^\infty) = \sup\{\langle x, y\rangle \mid y \in K\} \in \mathcal{H}_0.$$

Obviously, $M : \mathcal{K}_0(\mathbb{R}^n) \to \mathcal{H}_0$ is an order preserving map (1-1 and onto) and $S : \mathcal{K}_0(\mathbb{R}^n) \to \mathcal{H}_0$ is an order reversing map.

We continue the theorem:

Theorem (Artstein–Milman)

2. *There is a unique order reversing extension of the support map S to* $\mathrm{Cvx}\left(\mathbb{R}^n\right)$ *which is the Legendre transform.*

And what is the polarity map on $\mathcal{K}_0\left(\mathbb{R}^n\right)$

$$K \to K^\circ = \left\{x \in \mathbb{R}^n \mid \langle x, y\rangle \leq 1 \ \forall y \in K\right\}?$$

$K^\circ$is defined if $0 \in K$. So, let

$$\mathrm{Cvx}_0\left(\mathbb{R}^n\right) = \left\{f \in \mathrm{Cvx}\left(\mathbb{R}^n\right) \mid f(x) \geq 0 \text{ and } f(0) = 0\right\}$$

be the class of geometric convex functions.

3. There is a **unique** order reversing extension of the polarity map $\left\{\mathbf{1}_K^\infty \to \mathbf{1}_{K^\circ}^\infty \mid K \in \mathcal{K}_0\right\}$ to $\mathrm{Cvx}_0\left(\mathbb{R}^n\right) \backslash 0$ defined by

$$\mathcal{A}f = \sup \frac{\langle x, y\rangle - 1}{f(y)},$$

and $\mathcal{A}(0) := \mathbf{1}_{\{0\}}^\infty$. (By extension we mean that $\mathcal{A}\mathbf{1}_K^\infty = \mathbf{1}_{K^\circ}^\infty$)

There are ONLY two dualities on $\mathrm{Cvx}_0\left(\mathbb{R}^n\right)$ – $\mathcal{L}$ and $\mathcal{A}$:

Theorem (Artstein–Milman)

Let $n \geq 2$. The maps $\mathcal{L}$ and $\mathcal{A}$ are (essentially) the only order reversing involutions on $Cvx_0(\mathbb{R}^n)$. Precisely: if $T : Cvx_0 \to Cvx_0$ is

1. *involution $T \cdot T = Id$.*
2. *order reversing: $\forall f, g \in Cvx_0$ we have $f \leq g \Rightarrow Tf \geq Tg$,*

then $\exists C > 0$ and $B \in GL_n$, symmetric, s.t. either

$$\forall f \in Cvx_0, \quad Tf = \mathcal{L}(f(Bx))$$

or

$$\forall f \in Cvx_0, \quad Tf = C\mathcal{A}(f(Bx)).$$

(when $n = 1$ there are 8 such different dualities)

Consider the order preserving map (involution)

$$\mathcal{J} = \mathcal{L}\mathcal{A} = \mathcal{A}\mathcal{L}$$

which connects two dualities (supporting map – Legendre transform $\mathcal{L}$, and geometric duality$\mathcal{A}$).

$\mathcal{J}$ is a very interesting map $Cvx_0 \to Cvx_0$, order preserving. It is the <u>gauge</u> map:

[Fact]Artstein-Milman. $\mathcal{J}$ is the <u>only</u> order preserving extension of the Minkowski map M onto $Cvx_0(\mathbb{R}^n)$, i.e.

$$\mathcal{J}(\mathbf{1}_K^\infty) \equiv M(\mathbf{1}_K^\infty) = \|x\|_K.$$

So, on the class of convex functions we have the notion of support function, Minkowski functional and polarity!

Classical constructions in analysis which appear (uniquely) from elementary (simplest) properties

Now let us look into the continuation of the previous geometric ideas applied to the problems of **Analysis**.

We start with a characterization of the classical **Fourier transform** $\mathbb{F}$ on $\mathbb{R}^n : \mathbb{F}f = \int e^{-2\pi i \langle x,y \rangle} f(y)dy$. Let S be the Schwartz class of "rapidly" decreasing (infinitely smooth) functions on $\mathbb{R}^n$.

Theorem (Artstein, Faifman, Milman)

Assume we are given a bijective transform $\mathcal{F} : S \to S$, s.t. $\forall f, g \in S$ we have

$$\mathcal{F}(f \cdot g) = \mathcal{F}f * \mathcal{F}g.$$

Then $\exists$ diffeomorphism $\omega : \mathbb{R}^n \to \mathbb{R}^n$ s.t.

either $\forall f \in S$, $\mathcal{F}f = \mathbb{F}(f \circ \omega)$

or $\forall f \in S$, $\mathcal{F}f = \overline{\mathbb{F}(f \circ \omega)}$.

Real linearity and continuity of $\mathcal{F}$ is the automatic consequence. Previous versions contained more conditions and were proved jointly with S. Alesker. Joining these results with the previous theorem we may state that if $\mathcal{F} : S \to S$ s.t. $\forall f, g \in S$

$$\mathcal{F}(f \cdot g) = \mathcal{F}f * \mathcal{F}g,$$

$$\mathcal{F}(f * g) = \mathcal{F}f \cdot \mathcal{F}g,$$

then $\exists$ linear $A \in GL_n$, $|det(A)| = 1$, s.t. either

$$\forall f \in S, \quad \mathcal{F}f = \mathbb{F}(f \circ A)$$

or

$$\forall f \in S, \quad \mathcal{F}f = \overline{\mathbb{F}(f \circ A)}.$$

Derivative; Approach through the chain rule

What algebraic property characterizes derivative on $\mathbb{R}$?

Let

$$C_b^1(\mathbb{R}) = \left\{ f \in C^1(\mathbb{R}) \mid \begin{array}{l} f \text{ bounded from above or from} \\ \text{below (or both)} \end{array} \right\}.$$

write $(f \circ g)(x) := f(g(x))$ (composition).

We say, $T : Dom(T) = C^1(\mathbb{R}) \to Im(T) \subset C(\mathbb{R})$ is **non-degenerate** if

$$\exists x_0 \in \mathbb{R}, h_0 \in C^1(\mathbb{R}), (Th_0)(x_0) = 0, \tag{1}$$

$$\exists x_2 \in \mathbb{R}, h_2 \in C_b^1(\mathbb{R}), (Th_2)(x_2) < 0. \tag{2}$$

(a very weak surjectivity type condition).

Theorem (Artstein, König, Milman)

Let $T : \mathcal{D}om(T) := C^1(\mathbb{R}) \to Im(T) \subset C(\mathbb{R})$ *be an operation satisfying the chain rule*

$$T(f \circ g) = (Tf) \circ g \cdot Tg; \qquad f, g \in \mathcal{D}(T). \tag{3}$$

Assume that T *is non-degenerate in the sense of (1)+(2).*

(Normalizing conditions)

$$T(2Id) = 2$$

(is a constant function), then $Tf = f'$ *is the only solution.*

No linearity or any kind of continuity of T is assumed, but is the consequence (under the above conditions) of the chain rule only.

Without the normalization condition the result is:

$\exists$ strictly monotone continuously differentiable function $G(x)$ and $p > 0$, s.t.

$$Tf = \left|\frac{d(G \circ f)}{dG}\right|^p \cdot sgn(f').$$

Note: If $T(f \circ g) = Tf \cdot Tg \qquad (T : C^k \to C)$ (and not degenerate at any $x \in \mathbb{R}$) then $Tf = 1, \quad \forall f \in C^k$. So, we need $(Tf) \circ g \cdot Tg$ to build a non-degenerate operation.

The same answer may be written differently:

$\exists H \in C(\mathbb{R})$, $H > 0$ and $p > 0$ s.t.

$$Tf = \frac{H(f(x))}{H(x)} \cdot \left|f'\right|^p sgn(f').$$

The chain rule (3) has a natural (and unique) domain on which it acts (with the image inside $C(\mathbb{R})$), and it is $C^1(\mathbb{R})$.

Facts from (A-K-M).

Let $T : L \to C(\mathbb{R})$ satisfy (3) for $f, g, f \circ g \in L$. If:

- $L = C(\mathbb{R})$ and $\exists g_0 \in C$ and $x_0 \in \mathbb{R}$ s.t. $(Tg_0)(x_0) = 0$.Then $T\big|_{C_b(\mathbb{R})} \equiv 0$.
- $C^\infty(\mathbb{R}) \subset L \subset C^1(\mathbb{R}) \implies T$ maybe extended to $C^1(\mathbb{R})$;

So the natural $Dom(T)$ is C^1!

Rigidity of the chain rule

To what extent is the standard chain rule inferred by a much weaker version?

Let $V : C^1(\mathbb{R}) \to C(\mathbb{R})$ be non-degenerate, i.e.

1. For any $x \in \mathbb{R}$ there is $f \in C^1(\mathbb{R})$ such that $Vf(x) \neq 0$, and
2. For any $x \in \mathbb{R}$ there are $y \in \mathbb{R}$ and $f \in C^1(\mathbb{R})$ such that $f(y) = x$ and $Vf(y) \neq 0$.

Theorem (König-Milman)

Assume that $V, T_{1,} T_2 : C^1(\mathbb{R}) \to C(\mathbb{R})$ *are operators such that the equation*

$$V(f \circ g) = (T_1 f) \circ g \cdot (T_2 g)$$

holds for all $f, g \in C^1(\mathbb{R})$. *Assume that* V *is non-degenerate. There is a solution* T *of the chain rule equation* $T(f \circ g) = (Tf) \circ g \cdot Tg$, *s.t.*

$$(Vf)(x) = c_1(f(x))c_2(x) \cdot (Tf)(x)$$

$$(T_1 f)(x) = c_1(f(x)) \cdot (Tf)(x)$$

$$(T_2 f)(x) = c_2(x) \cdot (Tf)(x),$$

i.e. all three a priori different operators are essentially the same (a strong super rigidity).

Stability of the Chain Rule

$T : C^1(\mathbb{R}) \to C(\mathbb{R})$ is locally non-degenerate if $\forall$ open interval $J \subset \mathbb{R}$, $\forall x \in J$, $\exists g \in C^1(\mathbb{R})$, $y \in \mathbb{R}$, s.t. $g(y) = x$, $\mathrm{Im}(g) \subset J$ and $Tg(y) \neq 0$.

Theorem (König-Milman)

Fix $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ *and* $B : \mathbb{R}^3 \to \mathbb{R}$ *such that* $\forall f, g \in C^1$ *and* $\forall x \in \mathbb{R}$

$$T(f \circ g)(x) = Tf \circ g(x) \cdot Tg(x) + B(x, f \circ g(x), g(x)).$$

Assume that T *is locally non-degenerate and* Tf *depends non-trivially on* f'.

Then $B = 0$ *(and* T *satisfies the chain rule).*

Even more rigidity

Consider the "chain rule inequality"

$$T\,(f\circ g)\le (Tf)\circ g\cdot Tg \tag{$*$}$$

for $T : C^1(\mathbb{R}) \to C(\mathbb{R})$, $\mathrm{Dom}(T) = C^1(\mathbb{R})$.

Assume that T satisfies the following:

- non-degeneration: $\forall$ open interval $I \subset \mathbb{R}$, $\forall x \in I$, $\exists g \in C^1(\mathbb{R})$ s.t. $g(x) = x$, $\mathrm{Im}(g) \subset I$ and $Tg(x) > 1$.
- T is pointwise continuous: $\forall f, f_n \in C^1(\mathbb{R})$ s.t $f_n \to f$, $f_n' \to f'$ uniformly on compact subsets we have $(Tf_n)(x) \to (Tf)(x)$ pointwise for all $x \in \mathbb{R}$.

Theorem (König-Milman)

For T as above assume also $\exists x \in \mathbb{R}$ s.t $T(-Id)(x) < 0$. Then $\exists H \in C(\mathbb{R})$, $H > 0$, $\exists p > 0$ and $A \geq 1$ s.t

$$Tf = \begin{cases} \frac{H \circ f}{H} \left| f' \right|^p & \text{for } f' \geq 0 \\ -A \frac{H \circ f}{H} \left| f' \right|^p & \text{for } f' < 0. \end{cases}$$

Note:

- For $A = 1$, T satisfies the chain rule equation: we have equality in $(*)$.
- For both f and g non-decreasing we automatically have equality in $(*)$.
- Actually, the same is true if for some $C > 0$

$$T(f \circ g) \leq C \cdot (Tf) \circ g \cdot Tg$$

and even much more generally (the answer is slightly modified).

The following classically sound functional statement is used:

We say that $K : \mathbb{R} \to \mathbb{R}$ is submultiplicative if

$$K(\alpha\beta) \le K(\alpha)K(\beta), \qquad \forall \alpha, \beta \in \mathbb{R}.$$

Theorem (König-Milman)

Let K be submultiplicative, measurable and continuous at 0 and at 1. Assume $K(-1) < 0 < K(1)$. Then $\exists p > 0$ s.t.

$$K(\alpha) = \begin{cases} \alpha^p & \text{for } \alpha \ge 0 \\ -A|\alpha|^p & \text{for } \alpha < 0 \end{cases}$$

(and $K(-1) = -A \le -1$).

[every assumption in the theorem is needed]

As a corollary, K must be multiplicative on $\mathbb{R}^+$.

Approach through Leibniz rule

Theorem (König-Milman)

Suppose $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ *satisfies the Leibniz product formula*

$$T(f \cdot g) = Tf \cdot g + f \cdot Tg; \qquad f, g \in C^1(\mathbb{R}).$$

Then there are continuous functions $c(x), d(x)$ *such that*

$$Tf(x) = c(x)f'(x) + d(x)f(x)\ln|f(x)|, \qquad f \in C^1(\mathbb{R}), x \in \mathbb{R}.$$

(If T *also maps* $C^2(\mathbb{R})$ *into* $C^1(\mathbb{R})$*, then* $Tf = cf'$*.)*

We see that there exist (only!) two domains on which Leibniz's rule acts:

1. $C(\mathbb{R})$ with the only solution being the entropy function (Goldmann-Šemrl);
2. $C^1(\mathbb{R})$ with 2 solutions, f' and $f \cdot \ln|f|$, and their linear combination.

Remark: Suppose $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ satisfies the Leibniz rule and the chain rule operator equations

$$T(f \cdot g) = Tf \cdot g + f \cdot Tg,$$

$$T(f \circ g) = (Tf) \circ g \cdot Tg,$$

for all $f, g \in C^1(\mathbb{R})$. Then T is either identically 0 or the derivative, $Tf = f'$.

However:

Let $C(\mathbb{R}_{>1})_+ := \{f \in C(\mathbb{R}_{>1}) | \quad f > 1\}$ and $H(x) := x \ln |x|$. Then the operation T defined by $Tf(x) := H(f(x))/H(x)$ yields a map $T : C(\mathbb{R}_{>1})_+ \to C(\mathbb{R})$ satisfying the Leibniz rule and the chain rule functional equations which is different from the derivative on the subset of positive $C^1(\mathbb{R}_{>1})$ functions.

We know a description of **all** solutions T and A_i of

$$T(f \cdot g) = Tf \cdot A_1 g + Tg \cdot A_2 f$$

and there are just three more families of such solutions! And only one of them is non-degenerate. Moreover, we again observe "super-rigidity" of a "Leibniz rule" structure.

Theorem (König-Milman)

Let $V, T_1, T_2 : C^1(\mathbb{R}) \to C(\mathbb{R})$ *be operators such that*

$$V(f \cdot g) = (T_1 f) \cdot g + f \cdot (T_2 g) \tag{4}$$

is satisfied for all $f, g \in C^1(\mathbb{R})$*. Then there are continuous functions* $a, b, c_1, c_2 \in C(\mathbb{R})$ *such that with*

$$(Tf)(x) := b(x)f'(x) + a(x)f(x)\ln|f(x)|$$

we have

$$(Vf)(x) = (Tf)(x) + (c_1(x) + c_2(x))f(x)$$

$$(T_1 f)(x) = (Tf)(x) + c_1(x)f(x)$$

$$(T_2 f)(x) = (Tf)(x) + c_2(x)f(x).$$

The formula for $(Tf)(x)$ *represents the general solution of (4) in the case when* $V = T_1 = T_2 = T$*.*

It is surprising how rigid are simple relations which define (almost uniquely) basic operations/constructions in geometry and analysis. One more example:

Theorem (König-Milman)

Let $k \in \mathbb{N}$, $T : C^k(\mathbb{R}) \to C(\mathbb{R})$ be an operator and $B : \mathbb{R} \times \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}$ be a function such that

$$T(f \cdot g)(x) = Tf(x) \cdot g(x) + f(x) \cdot Tg(x) + B(x, f(x), \dots, f^{(k-1}(x), g(x), \dots g^{(k-1)}(x))$$

holds for all $f, g \in C^k(\mathbb{R})$ and $x \in \mathbb{R}$. Let T annihilate all polynomials of order $\leq k-1$. Then $Tf = d \cdot f^{(k)}$ for $d \in C(\mathbb{R})$, and B has the form

$$B(x, f(x), \dots, g^{(k-1)}(x)) = d(x) \sum_{j=1}^{k-1} \binom{k}{j} f^{(j)}(x) g^{(k-j)}(x).$$

And the Laplacian case

Theorem (König-Milman)

Let $n \in \mathbb{N}$. Let $T : C^2(\mathbb{R}^n, \mathbb{R}) \to C(\mathbb{R}^n, \mathbb{R})$ be an operator and B be a function $B : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ s.t.

$$T(f \cdot g)(x) = Tf(x) \cdot g(x) + f(x) \cdot Tg(x) + B(x, f(x), f'(x), g(x), g'(x)$$

holds for all $f, g \in C^2(\mathbb{R}^n, \mathbb{R})$ and all $x \in \mathbb{R}^n$. Let T annihilate all affine functions and be orthogonally invariant, i.e. $T(f \circ \varphi) = (Tf) \circ \varphi$ for all $\varphi \in O(n)$. Then T is a multiple of the Laplacian: there is $d \in C(\mathbb{R}_{\geq 0}, \mathbb{R})$ such that

$$Tf(x) = d(||x||)\Delta f(x),$$

$$B(x, f(x), f'(x), g(x), g'(x)) = d(||x||)\langle f'(x), g'(x)\rangle.$$

holds for all $f, g \in C^2(\mathbb{R}^n, \mathbb{R})$,$x \in \mathbb{R}^n$, and $||x||$ is Euclidean norm.