Journal of Convex Analysis
Volume 17 (2010), No. 3&4, 1033–1055

Distances between Closed Convex Cones: Old and New Results

Alfredo Iusem
*Instituto de Matemática Pura e Aplicada,
Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, Brazil
iusp@impa.br*

Alberto Seeger
*University of Avignon, Department of Mathematics,
33 rue Louis Pasteur, 84000 Avignon, France
alberto.seeger@univ-avignon.fr*

Dedicated to Hedy Attouch on the occasion of his 60th birthday.

Received: November 27, 2008

There are many ways of measuring the distance between two closed convex cones in a normed space. This work surveys the existing methods and discusses the merit of each one.

Keywords: Convex cone, metric space, Lipschitz equivalence of metrics, geometry of normed spaces, metric projection

1991 Mathematics Subject Classification: 46B40, 52A05

1. Introduction

1.1. Purpose of this work

One of the most fruitful extension of the concept of linear subspace is that of convex cone. The latter expression refers a nonempty set that is stable both under addition and multiplication by positive scalars. Convex cones play nowadays an ubiquitous role in many branches of applied mathematics.

How far is a closed convex cone from another one? This is a question that arises once and over again in the most diverse contexts. Mathematically speaking, the issue at hand is that of introducing a suitable metric on the hyperspace of all closed convex cones in a given normed space. Several metrics have been proposed in the last decades, some of them are quite natural, some others are more fancy. This work[1] attempts to survey the state of the art in the area.

The question that is being addressed is so fundamental that it will be inevitable to discuss also some related issues. General results concerning the geometry of normed spaces will enter into the discussion.

[1]Research carried out within the framework of the Brazil-France Cooperation Agreement in Mathematics.

ISSN 0944-6532 / $ 2.50 © Heldermann Verlag

1.2. Preliminary results on the Pompeiu-Hausdorff metric

We prepare the ground by recalling some facts concerning the classical Pompeiu-Hausdorff metric. A lot of space and time will be saved by proceeding in this way. Let $\mathcal{B}(X)$ denote the collection of all nonempty closed bounded subsets of a given real normed space $(X, \|\cdot\|)$. The commonest way of measuring the distance between two elements C_1, C_2 in $\mathcal{B}(X)$ is by using the expression

$$\mathrm{haus}(C_1, C_2) = \max\left\{\sup_{x\in C_1} \mathrm{dist}(x, C_2),\ \sup_{x\in C_2} \mathrm{dist}(x, C_1)\right\} \tag{1}$$

with $\mathrm{dist}(x, C) = \inf_{u\in C} \|u - x\|$ standing for the distance from the point x to the set C. The bivariate function introduced above is called the *Pompeiu-Hausdorff metric* on $\mathcal{B}(X)$.

Some authors apply the definition (1) also to unbounded sets, but such a strategy is somewhat problematic. We adhere to the traditional rule according to which a metric must be a finite-valued function. The abbreviation "haus" is used for honoring Felix Hausdorff (1868-1942), but the metric itself was introduced for the first time in the literature by the Romanian mathematician Dimitrie Pompeiu (1873-1954). The original definition given by Pompeiu in his 1905 Ph.D. thesis [26] suggests adding excesses instead of taking their maximum.

The next theorem is part of the folklore related to the Pompeiu-Hausdorff metric (cf. [27]). As usual, the symbol B_X stands for the closed unit ball in X.

Theorem 1.1. *Let X be a normed space. Then,*

(a) *The function* $\mathrm{haus} : \mathcal{B}(X) \times \mathcal{B}(X) \to \mathbb{R}$ *satisfies the axioms of a metric.*

(b) *For all C_1, C_2 in $\mathcal{B}(X)$, one can write*

$$\begin{aligned}\mathrm{haus}(C_1, C_2) &= \inf\{r \geq 0 : C_1 \subset C_2 + rB_X,\ C_2 \subset C_1 + rB_X\},\\ &= \sup_{x\in X} |\mathrm{dist}(x, C_1) - \mathrm{dist}(x, C_2)|\,,\end{aligned}$$

where the supremum could equally be taken just over $C_1 \cup C_2$.

There is yet another equivalent formulation of (1) when this metric is restricted to the convex sets in $\mathcal{B}(X)$. The next result, due to Hörmander [13], gives a support function characterization of the Pompeiu-Hausdorff metric. Recall that the support function Ψ^*_C of a set $C \in \mathcal{B}(X)$ is given by

$$\Psi^*_C(y) = \sup_{x\in C} y(x)$$

for all y in the topological dual space X^*.

Proposition 1.2. *Let X be a normed space. If $C_1, C_2 \in \mathcal{B}(X)$ are convex, then*

$$\mathrm{haus}(C_1, C_2) = \sup_{\|y\|_*=1} |\Psi^*_{C_1}(y) - \Psi^*_{C_2}(y)|$$

with $\|\cdot\|_*$ standing for the usual norm in X^*.

A nice property of the Pompeiu-Hausdorff metric is its invariance with respect to isometric transformations. An invertible linear isometry on X is a linear map $U : X \to X$ satisfying:

$$U \circ V = V \circ U = I \ \text{ for some linear map } V : X \to X, \quad \text{and}$$
$$\|Ux\| = \|x\| \ \text{ for all } x \in X.$$

The symbol $\mathrm{Isom}(X)$ indicates the set of all invertible linear isometries on X. Clearly, V is unique. It is denoted by U^{-1} and called the inverse of U. One can check that the inverse of an invertible linear isometry is in turn an invertible linear isometry.

Proposition 1.3. *Let U be an invertible linear isometry on a normed space X. Then,*

$$\mathrm{haus}(U(C_1), U(C_2)) = \mathrm{haus}(C_1, C_2)$$

for all $C_1, C_2 \in \mathcal{B}(X)$.

Proof. The sets $U(C_1)$ and $U(C_2)$ remain in $\mathcal{B}(X)$. In view of Theorem 1.1(b), one has

$$\begin{aligned}
\mathrm{haus}(U(C_1), U(C_2)) &= \sup_{x \in X} |\mathrm{dist}(x, U(C_1)) - \mathrm{dist}(x, U(C_2))| \\
&= \sup_{x \in X} \left|\mathrm{dist}(U^{-1}x, C_1) - \mathrm{dist}(U^{-1}x, C_2)\right| \\
&= \sup_{z \in X} |\mathrm{dist}(z, C_1) - \mathrm{dist}(z, C_2)| \,.
\end{aligned}$$

This completes the proof of the proposition. □

Taking the closed convex hull is a nonexpansive operation with respect to the Pompeiu-Hausdorff metric. This result, which is the object of the next proposition, is well known in the differential inclusion community [19].

Proposition 1.4. *For all $C_1, C_2 \in \mathcal{B}(X)$, one has* $\mathrm{haus}(\overline{\mathrm{co}}(C_1), \overline{\mathrm{co}}(C_2)) \leq \mathrm{haus}(C_1, C_2)$.

The Pompeiu-Hausdorff metric enjoys many additional properties. In this introductory section we have recorded the bare minimum needed to treat the main subjet of this survey.

2. Distance between closed convex cones

Unboundedness has always been a source of mathematical difficulties. Is there a reasonable way of measuring the distance between two closed sets that are not necessarily bounded? We do not need to address this question in such a degree of generality. After all, we are just interested in evaluating the distance between two members in the particular set

$$\Xi(X) \equiv \text{nontrivial closed convex cones in } X.$$

That a convex cone is nontrivial simply means that it is different from the singleton $\{0\}$ and different from the whole space X. Nothing is lost if one removes both trivial convex cones from the discussion.

Several metrics on $\Xi(X)$ are possible, each one of them having its own advantages and inconveniences. Let us review the most important options.

2.1. The truncated Pompeiu-Hausdorff metric ϱ

The class $\Xi(X)$ is formed by unbounded sets and therefore the Pompeiu-Hausdorff metric is not directly applicable. A natural way to overcome this difficulty is to truncate the sets $K_1, K_2 \in \Xi(X)$ by intersecting them with the unit ball B_X. After truncation, one can proceed with the evaluation of the expression

$$\varrho(K_1, K_2) = \mathrm{haus}(K_1 \cap B_X, K_2 \cap B_X). \tag{2}$$

By an obvious reason, one refers to ϱ as the *truncated Pompeiu-Hausdorff metric* on $\Xi(X)$. One could equally well consider the more general term

$$\varrho_t(K_1, K_2) = \mathrm{haus}(K_1 \cap tB_X, K_2 \cap tB_X),$$

but this would add nothing new. A simple homogeneity argument shows that t behaves merely as a multiplicative factor, to wit, $\varrho_t(K_1, K_2) = t\varrho(K_1, K_2)$. This is why the attention is focused only on the expression (2).

It is not clear who was the first person that considered ϱ as tool for measuring distances between closed convex cones. In any case, for closed linear subspaces of a Banach space, the use of ϱ goes back at least to Gurariĭ [11].

Proposition 2.1. *Let X be a normed space. Then,*

(a) *The function $\varrho : \Xi(X) \times \Xi(X) \to \mathbb{R}$ satisfies the axioms of a metric.*
(b) *$\varrho(K_1, K_2) \leq 1$ for all $K_1, K_2 \in \Xi(X)$.*
(c) *For all $U \in \mathrm{Isom}(X)$ and $K_1, K_2 \in \Xi(X)$, one has $\varrho(U(K_1), U(K_2)) = \varrho(K_1, K_2)$.*

Proof. That ϱ is a metric follows directly from Theorem 1.1. That ϱ is bounded by 1 is due to the fact that $\mathrm{dist}(x, K \cap B_X) \leq 1$ whenever $x \in B_X$ and $K \in \Xi(X)$. In order to prove (c), consider $K_1, K_2 \in \Xi(X)$ and write

$$\begin{aligned} \varrho(U(K_1), U(K_2)) &= \sup_{x \in X} |\mathrm{dist}(x, U(K_1) \cap B_X) - \mathrm{dist}(x, U(K_2) \cap B_X)| \\ &= \sup_{x \in X} |\mathrm{dist}(x, U(K_1 \cap B_X)) - \mathrm{dist}(x, U(K_2 \cap B_X))| \\ &= \mathrm{haus}(U(K_1 \cap B_X), U(K_2 \cap B_X)). \end{aligned}$$

Proposition 1.3 completes the proof. □

2.2. The uniform metric μ

Another natural way of measuring the distance between a pair K_1, K_2 of elements in $\Xi(X)$ is by means of the expression

$$\mu(K_1, K_2) = \sup_{\|x\|=1} |\mathrm{dist}(x, K_1) - \mathrm{dist}(x, K_2)|. \tag{3}$$

The right-hand side of (3) corresponds to the uniform distance between the functions $\mathrm{dist}[\,\cdot\,, K_1]$ and $\mathrm{dist}[\,\cdot\,, K_2]$ when restricted to the unit sphere. In view of this observation, one refers to μ as the *uniform metric* on $\Xi(X)$.

Remark 2.2. If we were interested in measuring the distance between two arbitrary nonempty closed sets, say C_1 and C_2, then we could consider

$$\mu_{\text{sum}}(C_1, C_2) = \sum_{t=1}^{\infty} \frac{1}{2^t}\, \mu_t(C_1, C_2)$$

with $\mu_t(K_1, K_2) = \sup_{\|x\| \le t} |\text{dist}(x, K_1) - \text{dist}(x, K_2)|$. Another option for consideration would be using the continuous version

$$\mu_{\text{RW}}(C_1, C_2) = \int_0^{\infty} e^{-t} \mu_t(C_1, C_2)\, dt$$

suggested by Rockafellar and Wets [27].

Let us come back to the conic framework. There is no need of truncating K_1 and K_2 in the definition of μ. In fact, there is an implicit truncation procedure in (3) because the supremum is taken over a bounded set and not over the whole space X.

Proposition 2.3. *Let X be a normed space. Then,*

(a) *The function $\mu : \Xi(X) \times \Xi(X) \to \mathbb{R}$ satisfies the axioms of a metric.*
(b) *$\mu(K_1, K_2) \le 1$ for all $K_1, K_2 \in \Xi(X)$.*
(c) *For all $U \in \text{Isom}(X)$ and $K_1, K_2 \in \Xi(X)$, one has $\mu(U(K_1), U(K_2)) = \mu(K_1, K_2)$.*

Proof. Parts (a) and (b) are trivial. For all $U \in \text{Isom}(X)$ and $K_1, K_2 \in \Xi(X)$, one has

$$\begin{aligned} \mu(U(K_1), U(K_2)) &= \sup_{\|x\|=1} |\text{dist}(x, U(K_1)) - \text{dist}(x, U(K_2))| \\ &= \sup_{\|x\|=1} \left|\text{dist}(U^{-1}x, K_1) - \text{dist}(U^{-1}x, K_2)\right| \\ &= \sup_{\|z\|=1} |\text{dist}(z, K_1) - \text{dist}(z, K_2)| \end{aligned}$$

This completes the proof of (c). □

The use of μ as tool for measuring distances between closed convex cones is documented in [27, 30] and in many other references. Are ϱ and μ different metrics after all? The following result is specific to Hilbert spaces and does not extend to general normed spaces.

Proposition 2.4. *If X is a Hilbert space, then $\mu(K_1, K_2) = \varrho(K_1, K_2)$ for all $K_1, K_2 \in \Xi(X)$.*

Proof. In a Hilbert space setting the metric projection onto a closed convex cone is a norm-reducing map. As a consequence, for any $K \in \Xi(X)$, one has

$$\text{dist}(x, K \cap B_X) = \text{dist}(x, K) \quad \text{for all } x \in B_X. \tag{4}$$

In such a case, Theorem 1.1 yields

$$\begin{aligned} \varrho(K_1, K_2) &= \sup_{\|x\| \le 1} |\text{dist}(x, K_1 \cap B_X) - \text{dist}(x, K_2 \cap B_X)| \\ &= \sup_{\|x\| \le 1} |\text{dist}(x, K_1) - \text{dist}(x, K_2)|. \end{aligned}$$

Of course, the last supremum is an equivalent formulation of $\mu(K_1, K_2)$. $\square$

Beyond a Hilbert space setting one cannot rely on the relation (4). The next example shows that the metrics ϱ and μ are not only different, but also unrelated.

Example 2.5. Let $X = \mathbb{R}^2$ be equipped with the norm $\|x\| = \max\{|x_1|, |x_2|\}$. Consider the rays $K_1 = \mathbb{R}_+(1, a)$ and $K_2 = \mathbb{R}_+(1, b)$ with $a, b \in [0, 1]$. One can easily check that $\varrho(K_1, K_2) = |b - a|$. On the other hand, a matter of computation yields

$$\mu(K_1, K_2) = \frac{2\,|b - a|}{(1 + a)(1 + b)}\,.$$

For instance, if $a = 1/4$ and $b = 1/2$, then $\mu(K_1, K_2) = 4/15$ is greater than $\varrho(K_1, K_2) = 1/16$. By contrast, if $a = 2/5$ and $b = 1/2$, then $\mu(K_1, K_2) = 2/21$ is smaller than $\varrho(K_1, K_2) = 1/10$. This shows that ϱ and μ cannot be compared.

2.2.1. An application of μ

Normality is a fundamental concept of the theory of convex cones. One says that $K \in \Xi(X)$ is *normal* if the set $K_\bullet = (B_X + K) \cap (B_X - K)$ is bounded. This is not perhaps the most common way of introducing the concept of normality, but the above definition is equivalent to the original one given by Krein [21]. The normality index

$$\Phi_{\text{nor}}(K) = \sup\{r \geq 0 : \; rK_\bullet \subset B_X\} \tag{5}$$

is a coefficient that measures to which extent the convex cone K is normal.

The next proposition, taken from [17], says that the function $\Phi_{\text{nor}} : \Xi(X) \to \mathbb{R}$ behaves in a nonexpansive manner if perturbations are measured with respect to the metric μ.

Proposition 2.6. *Let X be a normed space. Then, for all $K_1, K_2 \in \Xi(X)$, one has*

$$|\Phi_{\text{nor}}(K_1) - \Phi_{\text{nor}}(K_2)| \leq \mu(K_1, K_2).$$

The proof of Proposition 2.6 becomes easier if one represents the normality index (5) in the equivalent form

$$\Phi_{\text{nor}}(K) = \inf_{\|z\|=1} \max\{\text{dist}(z, K), \text{dist}(-z, K)\}\,.$$

With such a formula at hand, one sees better why μ enters naturally into the picture.

2.3. The spherical metric σ

An element $K \in \Xi(X)$ if fully determined by the trace $K \cap S_X$ left on the unit sphere S_X. In view of this observation, the expression

$$\sigma(K_1, K_2) = \text{haus}(K_1 \cap S_X, K_2 \cap S_X) \tag{6}$$

is a natural tool for quantifying the distance between K_1 and K_2. One refers to σ as the *spherical metric* on $\Xi(X)$.

The expression (6) was originally introduced for closed linear subspaces of a Banach space by Gohberg and Markus [9]. In such a context, it has been extensively studied by Gurariĭ [11], Paraska [24], and many other Soviet mathematicians. The survey paper by Ostrovskiĭ [23] provides interesting bibliographical comments.

Proposition 2.7. *Let X be a normed space. Then,*

(a) *The function $\sigma : \Xi(X) \times \Xi(X) \to \mathbb{R}$ satisfies the axioms of a metric.*
(b) *$\sigma(K_1, K_2) \leq 2$ for all $K_1, K_2 \in \Xi(X)$.*
(c) *For all $U \in \mathrm{Isom}(X)$ and $K_1, K_2 \in \Xi(X)$, one has $\sigma(U(K_1), U(K_2)) = \sigma(K_1, K_2)$.*

Proof. (a) and (b) are immediate. For the proof of (c), observe that

$$\begin{aligned}\sup_{x \in U(K_1)\cap S_X} \mathrm{dist}(x, U(K_2) \cap S_X) &= \sup_{x \in U(K_1 \cap S_X)} \mathrm{dist}(x, U(K_2 \cap S_X)) \\ &= \sup_{U^{-1}x \in K_1 \cap S_X} \mathrm{dist}(U^{-1}x, K_2 \cap S_X) \\ &= \sup_{z \in K_1 \cap S_X} \mathrm{dist}(z, K_2 \cap S_X).\end{aligned}$$

A similar relation holds if one exchanges the role of K_1 and K_2. □

2.3.1. Two applications of σ

The coefficient (5) is not the only one that serve as tool for measuring the degree of normality of a convex cone. One could perfectly well consider instead the expression

$$\Psi_{\mathrm{nor}}(K) = \inf_{u,v \in K \cap S_X} \left\| \frac{u+v}{2} \right\| . \tag{7}$$

Although the functions Φ_{nor} and Ψ_{nor} are different, they do share a number of properties in common. The next result is in the same vein as Proposition 2.6. Since the spherical trace $K \cap S_X$ of K appears in the very definition of the coefficient (7), it is no wonder that the spherical metric σ enters into action. The proof of the Proposition 2.8 can be found in [17, Lemma 1].

Proposition 2.8. *Let X be a normed space. Then, for all $K_1, K_2 \in \Xi(X)$, one has*

$$|\Psi_{\mathrm{nor}}(K_1) - \Psi_{\mathrm{nor}}(K_2)| \leq \sigma(K_1, K_2).$$

The next application of the spherical metric σ has to do with the concept of sharpness. One says that $K \in \Xi(X)$ is *sharp* if there is a nonzero vector $y \in X^*$ such that $\|x\| \leq y(x)$ for all $x \in K$. Convex cones having such a qualitative behavior are very useful in practice. From a quantitative point of view, a suitable way of measuring the degree of sharpness of K is by means of the expression

$$\Phi_{\mathrm{sharp}}(K) = \sup_{\substack{c \geq 0,\, \|y\|_* = 1 \\ K \subset \mathrm{rev}(c,y)}} c\,. \tag{8}$$

Here, $\mathrm{rev}(c, y)$ denotes the revolution-like cone with parameters c and y, i.e.,

$$\mathrm{rev}(c, y) = \{x \in X : c\|x\| \leq y(x)\}.$$

Revolution-like cones, also called Bishop-Phelps cones, are widely used in functional analysis and they do not need further presentation. The reasoning leading to the expression (8) is explained with care in reference [16], from where we take the following result.

Proposition 2.9. *Let X be a normed space. Then, for all $K_1, K_2 \in \Xi(X)$, one has*

$$|\Phi_{\rm sharp}(K_1) - \Phi_{\rm sharp}(K_2)| \le \sigma(K_1, K_2).$$

At first sight, it is not entirely clear why the metric σ is here the right choice. To see this, one must write the sharpness index (8) in the equivalent form

$$\Phi_{\rm sharp}(K) = \sup_{\|y\|_*=1} \inf_{x \in K \cap S_X} y(x)$$

and then recall Propositions 1.2 and 1.4.

2.4. The logarithmic gap metric λ

The *gap* between a pair K_1, K_2 of elements in $\Xi(X)$ is defined as the number

$$\delta(K_1, K_2) = \max\left\{ \sup_{x \in K_1 \cap S_X} \operatorname{dist}(x, K_2), \sup_{x \in K_2 \cap S_X} \operatorname{dist}(x, K_1) \right\}. \tag{9}$$

The above expression was introduced by Krasnoselskiĭ and Krein [20] as a tool for measuring the inclination between two closed linear subspaces of a given Hilbert space. The term "gap" must be handle with care because it admits several meanings in the mathematical literature. Some authors refer to $\delta(K_1, K_2)$ as the opening between K_1 and K_2.

The proposition below is classical and has interesting applications in linear control theory. A full chapter of the book by Feintuch [8] gravitates around this result. The formula (10) can be traced back to the book by Akhiezer and Glazman [1]. Recall that the symbol

$$\|A\|_{\rm oper} = \sup_{\|x\|=1} \|Ax\|$$

indicates the operator or spectral norm of a linear continuous operator $A : X \to X$.

Proposition 2.10. *Let K_1, K_2 be nontrivial closed linear subspaces of a Hilbert space X. Then,*

$$\delta(K_1, K_2) = \|\Pi_{K_1} - \Pi_{K_2}\|_{\rm oper} \tag{10}$$

with Π_{K_1} and Π_{K_2} denoting the orthogonal projectors onto K_1 and K_2, respectively.

For closed convex cones in reflexive Banach spaces, the gap function δ was used for the first time by Walkup and Wets [29]. We shall comment on the contribution of these authors latter in Section 5. In general normal spaces, the gap function δ can be characterized by using the concept of conic neighborhood. By a *conic neighborhood* of $K \in \Xi(X)$ one understands a set of the form

$$V_r(K) = \{x \in X : \operatorname{dist}(x, K) \le r\|x\|\}$$

with $r \in [0, 1]$. In applications, the parameter r is usually positive and small. The name attributed to $V_r(K)$ comes from the fact that this set is a cone, i.e., it is stable under multiplication by positive scalars. However, one should be aware that $V_r(K)$ is not convex in general.

Proposition 2.11. *Let X be a normed space. Then, for all $K_1, K_2 \in \Xi(X)$, one has*

$$\delta(K_1, K_2) = \inf\{r \geq 0 : K_1 \subset V_r(K_2),\ K_2 \subset V_r(K_1)\}. \tag{11}$$

Proof. Let $\nu(K_1, K_2)$ denote the infimum in (11). If $\delta(K_1, K_2) = r$, then

$$K_1 \subset K_2 + (r + \varepsilon)\|x\| B_X,$$
$$K_2 \subset K_1 + (r + \varepsilon)\|x\| B_X$$

for all $\varepsilon > 0$ and $x \in K$. This yields

$$K_1 \subset V_{r+\varepsilon}(K_2), \tag{12}$$
$$K_2 \subset V_{r+\varepsilon}(K_1). \tag{13}$$

One gets $\nu(K_1, K_2) \leq r + \varepsilon$, and then one lets $\varepsilon \to 0$. Conversely, suppose that $\nu(K_1, K_2) = r$. In such a case, one can write (12)–(13) for all $\varepsilon > 0$. Hence,

$$\operatorname{dist}(x, K_2) \leq (r + \varepsilon)\|x\| \text{ for all } x \in K_1, \tag{14}$$
$$\operatorname{dist}(x, K_1) \leq (r + \varepsilon)\|x\| \text{ for all } x \in K_2. \tag{15}$$

By taking the supremum in (14) with respect to $x \in K_1 \cap S_X$, and the supremum in (15) with respect to $x \in K_2 \cap S_X$, one arrives at $\delta(K_1, K_2) \leq r + \varepsilon$. Again, one lets $\varepsilon \to 0$. □

The next proposition provides an alternative characterization of the truncated Pompieu-Hausdorff distance in the context of a Hilbert space. Often, the expression $\delta(K_1, K_2)$ is much easier to compute than $\varrho(K_1, K_2)$.

Proposition 2.12. *If X is a Hilbert space, then $\delta(K_1, K_2) = \varrho(K_1, K_2)$ for all $K_1, K_2 \in \Xi(X)$.*

Proof. Since $\operatorname{dist}[\,\cdot\,, K_1]$ and $\operatorname{dist}[\,\cdot\,, K_2]$ are positively homogeneous functions, the number (9) can be written in the equivalent form

$$\delta(K_1, K_2) = \max\left\{ \sup_{x \in K_1 \cap B_X} \operatorname{dist}(x, K_2),\ \sup_{x \in K_2 \cap B_X} \operatorname{dist}(x, K_1) \right\}.$$

The later expression looks similar to (2), except that now we are using $\operatorname{dist}(\,\cdot\,, K_1)$ and $\operatorname{dist}(\,\cdot\,, K_2)$ instead of $\operatorname{dist}(\,\cdot\,, K_1 \cap B_X)$ and $\operatorname{dist}(\,\cdot, K_2 \cap B_X)$, respectively. For completing the proof of the proposition it suffices to recall the relation (4). □

Beyond a Hilbert space setting the use of δ as tool for measuring distances has a great inconvenience. Indeed, the function δ does not satisfy the triangular inequality! This phenomenon was already observed by Gohberg and Markus [9] while working with closed linear subspaces of a Banach space X.

Example 2.13. Consider the space $\mathbb{R}^2$ equipped with the Manhattan norm $\|x\| = |x_1| + |x_2|$. Let us see what happens with the triangular inequality if one works with the lines

$$K_i = \{(u_1, u_2) \in \mathbb{R}^2 : u_2 = a_i\, u_1\}$$

with $i \in \{1,2,3\}$. A matter of computation shows that

$$\delta(K_1, K_2) = \max\left\{\frac{1}{\max\{1, |a_2|\}}\,\frac{|a_2 - a_1|}{1 + |a_1|}\,,\,\frac{1}{\max\{1, |a_1|\}}\,\frac{|a_2 - a_1|}{1 + |a_2|}\right\}.$$

Similar formulas are obtained for $\delta(K_1, K_3)$ and $\delta(K_2, K_3)$. For simplicity, let $0 = a_1 < a_2 < a_3 \le 1$. This choice yields $\delta(K_1, K_3) = a_3$, $\delta(K_1, K_2) = a_2$, and $\delta(K_2, K_3) = (a_3 - a_2)/(1 + a_2)$. Since $a_2 + (a_3 - a_2)/(1 + a_2) < a_3$, the triangular inequality is violated.

Similar examples can be constructed in other normed spaces as well. Fortunately, this pathological behavior of δ can be successfully remediated. The next lemma gives us a clue on how to proceed.

Lemma 2.14. *Let X be a normed space. Then, the weak triangular inequality*

$$\delta(K_1, K_3) \le \delta(K_1, K_2) + \delta(K_2, K_3) + \delta(K_1, K_2)\,\delta(K_2, K_3) \tag{16}$$

holds for all $K_1, K_2, K_3 \in \Xi(X)$.

In view of (16), the function δ can be converted into a true metric by means of a suitable logarithmic transformation. It is not difficult to guess that an appropriate choice is

$$\lambda(K_1, K_2) = \log\left[1 + \delta(K_1, K_2)\right]. \tag{17}$$

We refer to the above number as the *logarithmic gap distance* between K_1 and K_2. The expression (17) was originally introduced by Gohberg and Marcus [9] for measuring distances between closed linear subspaces of a Banach space.

Proposition 2.15. *Let X be a normed space. Then,*

(*a*) *The function $\lambda : \Xi(X) \times \Xi(X) \to \mathbb{R}$ satisfies the axioms of a metric.*
(*b*) *$\lambda(K_1, K_2) \le \log 2$ for all $K_1, K_2 \in \Xi(X)$.*
(*c*) *For all $U \in \mathrm{Isom}(X)$ and $K_1, K_2 \in \Xi(X)$, one has $\lambda(U(K_1), U(K_2)) = \lambda(K_1, K_2)$.*

Proof. Write the weak triangular inequality (16) in the equivalent form

$$[1 + \delta(K_1, K_3)] \le [1 + \delta(K_1, K_2)]\;[1 + \delta(K_2, K_3)]$$

and then take logarithms on both sides. One obtains the usual triangular inequality for λ. We omit the details concerning the other parts of the proposition. □

3. Lipschitz equivalence of metrics

Despite the violation of the triangular inequality, the gap function δ introduces a topology on $\Xi(X)$. It turns out that such topology is the same as the one induced by the metrics λ, ϱ, μ, and σ. On the other hand, these four metrics are not just topologically equivalent, but also Lipschitz equivalent. Before discussing this theme in depth we need to prepare the ground by recalling some essential results concerning the geometry of normed spaces.

3.1. The sphericity coefficient of a normed space

Definition 3.1. The sphericity coefficient of a normed space X, denoted by c_X, is the infimum of all nonnegative numbers c such that

$$\frac{1}{2}\left\|\frac{u}{\|u\|}-\frac{v}{\|v\|}\right\| \leq c\,\frac{\|u-v\|}{\|u\|+\|v\|} \tag{18}$$

for all nonzero vectors $u, v \in X$.

A few comments on Definition 3.1 are in order. Some authors refer to the expression

$$\alpha(u,v)=\left\|\frac{u}{\|u\|}-\frac{v}{\|v\|}\right\|$$

as the angular distance between the nonzero vectors $u, v \in X$. The motivation behind Definition 3.1 is comparing the angular distance $\alpha(u,v)$ with the usual distance $\|u-v\|$. Dividing by $\|u\|+\|v\|$ in (18) is helpful for obtaining the same degree of homogeneity on both sides of the inequality, otherwise a constant c as above would not exist.

Proposition 3.2. *The sphericity coefficient of a normed space X satisfies $1 \leq c_X \leq 2$. Furthermore, $c_X = 1$ if and only if X is a pre-Hilbert space.*

Proof. Let u, v be arbitrary nonzero vectors in a normed space X. By combining the relation

$$(1/2)(\|u\|+\|v\|) \leq \max\{\|u\|,\|v\|\}$$

and the Massera-Schäffer inequality

$$\frac{1}{2}\left\|\frac{u}{\|u\|}-\frac{v}{\|v\|}\right\| \leq \frac{\|u-v\|}{\max\{\|u\|,\|v\|\}}\,, \tag{19}$$

one sees that (18) holds for the particular choice $c = 2$. This yields the upper bound $c_X \leq 2$. On the other hand, a constant c as in (18) cannot be smaller than 1, otherwise the choice $v = -2u$ would lead to a contradiction. This explains why $c_X \geq 1$. Finally, in a pre-Hilbert space X it is possible to write the Dunkl-Williams inequality

$$\frac{1}{2}\left\|\frac{u}{\|u\|}-\frac{v}{\|v\|}\right\| \leq \frac{\|u-v\|}{\|u\|+\|v\|}$$

as can be checked by a usual squaring process. In fact, as observed by Kirk and Smiley [18], the Dunkl-Williams inequality characterizes the norms that derive from an inner product. This takes care of the second part of the proposition. □

Hilbert spaces are the most "spherical" among all normed spaces. The less spherical normed spaces are those for which $c_X = 2$. The example below enters into this category.

Example 3.3. Let $X = \mathbb{R}^2$ be equipped with the Chebyshev norm $\|x\| = \max\{|x_1|,|x_2|\}$. Let

$$u=(1,1), \qquad v=\left(1+\frac{\varepsilon}{2-\varepsilon},1-\frac{\varepsilon}{2-\varepsilon}\right)$$

with $\varepsilon > 0$ standing for a small parameter. A simple computation leads to

$$\frac{1}{2}\left\|\frac{u}{\|u\|} - \frac{v}{\|v\|}\right\| = \frac{\varepsilon}{2} \quad \text{and} \quad \frac{\|u-v\|}{\|u\|+\|v\|} = \frac{\varepsilon}{4-\varepsilon}. \tag{20}$$

Hence, $(4-\varepsilon)/2 \leq c_X \leq 2$. By letting $\varepsilon \to 0$, one sees that $c_X = 2$.

Finite dimensionality plays here no role. What really matters is the total lack of sphericity of Chebyshev-type norms.

Example 3.4. Consider now the space $X = C([a,b],\mathbb{R})$ of continuous functions $x : [0,1] \to \mathbb{R}$ equipped with its usual norm $\|x\| = \max_{t\in[a,b]} |x(t)|$. If one chooses

$$u(t) = 1, \qquad v(t) = 1 + \frac{\varepsilon}{2-\varepsilon} - \left(\frac{2\varepsilon}{2-\varepsilon}\right) t,$$

then one gets the same estimates as in (20). Hence, again $c_X = 2$.

Remark 3.5. Consider the space $\ell_p(\mathbb{R})$ of p-summable real sequences equipped with the norm $\|x\| = \left[\sum_{n\in\mathbb{N}} |x_n|^p\right]^{1/p}$. In principle, one should obtain a sphericity coefficient c_p that goes to 2 when p increases to ∞, and that goes to 1 when p decreases to 2. Needless to say, the explicit computation of c_p is a quite cumbersome task.

3.2. Distance to various truncations of a convex cone

For a closed convex cone K in a normed space X, one trivially has

$$\operatorname{dist}(x,K) \leq \operatorname{dist}(x, K\cap B_X) \leq \operatorname{dist}(x, K\cap S_X) \tag{21}$$

for all $x \in X$. The order of these inequalities can be reversed but this requires incorporating suitable multiplicative constants. The sphericity coefficient c_X of the underlying space X plays here a useful role.

Lemma 3.6. *Let K be a nontrivial closed convex cone in a normed space X. Then,*

(a) $\operatorname{dist}(x, K\cap B_X) \leq c_X \operatorname{dist}(x,K)$ *for all* $x \in B_X$.
(b) $\operatorname{dist}(x, K\cap S_X) \leq 2\operatorname{dist}(x,K)$ *for all* $x \in S_X$.

Proof. Part (a). Consider first the case $\|x\| = 1$. Pick $\varepsilon > 0$ and a point $x_\varepsilon \in K$ such that

$$\|x - x_\varepsilon\| \leq \operatorname{dist}(x,K) + \varepsilon. \tag{22}$$

If $\|x_\varepsilon\| \leq 1$, then

$$\operatorname{dist}(x, K\cap B_X) \leq \|x - x_\varepsilon\| \leq \operatorname{dist}(x,K) + \varepsilon \leq c_X \operatorname{dist}(x,K) + \varepsilon.$$

By letting $\varepsilon \to 0$ one obtains the announced inequality. If $\|x_\varepsilon\| > 1$, then

$$\operatorname{dist}(x, K\cap B_X) \leq \left\| x - \frac{x_\varepsilon}{\|x_\varepsilon\|}\right\| \leq 2c\,\frac{\|x - x_\varepsilon\|}{1+\|x_\varepsilon\|}$$

with c being a nonnegative constant as in (18). By combining (22) and $2/(1+\|x_\varepsilon\|) \leq 1$, one gets

$$\operatorname{dist}(x, K\cap B_X) \leq c\,(\operatorname{dist}(x,K) + \varepsilon).$$

We now let $\varepsilon \to 0$ and $c \to c_X$. The general case $x \in B_X$ is treated as follows. Suppose that $x \neq 0$, otherwise we are done. From the first part of the proof, one has

$$\operatorname{dist}\left(\frac{x}{\|x\|}, K \cap B_X\right) \leq c_X \operatorname{dist}\left(\frac{x}{\|x\|}, K\right).$$

One gets in this way

$$\operatorname{dist}(x, K \cap B_X) \leq \operatorname{dist}(x, K \cap \|x\| B_X) \leq c_X \operatorname{dist}(x, K).$$

Part (b). Let $x \in S_X$. Pick $\varepsilon > 0$ and $x_\varepsilon \in K$ as in (22). Slightly perturbing x_ε if necessary, one may suppose that $x_\varepsilon \neq 0$. The Massera-Schäffer inequality (19) yields

$$\operatorname{dist}(x, K \cap S_X) \leq \left\| x - \frac{x_\varepsilon}{\|x_\varepsilon\|} \right\| \leq \frac{2\,\|x - x_\varepsilon\|}{\max\{1, \|x_\varepsilon\|\}}.$$

Due to (22) and $\max\{1, \|x_\varepsilon\|\} \geq 1$, one ends up with

$$\operatorname{dist}(x, K \cap S_X) \leq 2(\operatorname{dist}(x, K) + \varepsilon).$$

The announced result follows by letting $\varepsilon \to 0$. $\square$

The constant 2 in Lemma 3.6(b) cannot be replaced by something smaller, even if the norm of X derives from an inner product.

Lemma 3.7. *If K be a nontrivial closed convex cone in a normed space X, then*

$$\begin{aligned} \operatorname{dist}(x, K \cap B_X) &= \operatorname{dist}(x, K) \quad \textit{for all } x \in (1/2)B_X, \qquad (23)\\ \operatorname{dist}(x, K \cap 2B_X) &= \operatorname{dist}(x, K) \quad \textit{for all } x \in B_X. \end{aligned}$$

Proof. Both equalities are equivalent. This can be seen by exploiting the positive homogeneity of $\operatorname{dist}(\,\cdot\,, K)$ and the general relation

$$\operatorname{dist}(x, K \cap tB_X) = t \operatorname{dist}\left(t^{-1}x\,, K \cap B_X\right) \tag{24}$$

for all $x \in X$ and $t > 0$. So, we concentrate the attention in (23). Let $x \in (1/2)B_X$. In view of (21), we just need to check that $\operatorname{dist}(x, K \cap B_X) \leq \operatorname{dist}(x, K)$. Consider first the case $\|x\| < 1/2$. Take $\varepsilon > 0$ such that $2\|x\| + \varepsilon \leq 1$ and find a point $x_\varepsilon \in K$ as in (22). One has

$$\begin{aligned} \|x_\varepsilon\| &\leq \|x\| + \|x - x_\varepsilon\| \\ &\leq \|x\| + \operatorname{dist}(x, K) + \varepsilon \\ &\leq 2\|x\| + \varepsilon, \end{aligned}$$

and therefore $x_\varepsilon \in K \cap B_X$. This shows that

$$\operatorname{dist}(x, K \cap B_X) \leq \|x - x_\varepsilon\| \leq \operatorname{dist}(x, K) + \varepsilon.$$

One lets finally $\varepsilon \to 0$. The case $\|x\| = 1/2$ is derived from the previous one by relying on a continuity argument. Indeed, one passes to the limit on both sides of

$$\operatorname{dist}(x_n, K \cap B_X) \leq \operatorname{dist}(x_n, K)$$

with $\{x_n\}_{n \in \mathbb{N}}$ being any sequence converging to x such that $\|x_n\| < 1/2$. $\square$

The radius 1/2 appearing in (23) seems a bit curious, but this is the best result in a general normed space. A larger radius is acceptable only if X has a special structure. For instance, if X is a Hilbert space, then the equality (23) holds not just on $(1/2)B_X$, but on the whole ball B_X.

3.3. Comparing metrics

Everthing is now ready to compare the gap function δ and the four different metrics introduced in Section 2.

Theorem 3.8. *Let X be a normed space. Then, for all K_1, K_2 in $\Xi(X)$, one has*

$$\delta(K_1,K_2) \leq \varrho(K_1,K_2) \leq c_X\, \delta(K_1,K_2) \tag{25}$$

$$\delta(K_1,K_2) \leq \mu(K_1,K_2) \leq 2\, \varrho(K_1,K_2) \tag{26}$$

$$\delta(K_1,K_2) \leq \sigma(K_1,K_2) \leq 2\, \delta(K_1,K_2) \tag{27}$$

$$(\log 2)\, \delta(K_1,K_2) \leq \lambda(K_1,K_2) \leq \delta(K_1,K_2). \tag{28}$$

In particular, the metrics ρ, μ, σ, and λ are Lipschitz equivalent.

Proof. The right inequality in (25) is due to Lemma 3.6(a). In view of Lemma 3.7 and the general relation (24), one has

$$\begin{aligned}\mu(K_1,K_2) &= \sup_{\|x\|\leq 1} |\mathrm{dist}(x, K_1 \cap 2B_X) - \mathrm{dist}(x, K_2 \cap 2B_X)| \\ &= 2 \sup_{\|x\|\leq 1/2} |\mathrm{dist}(x, K_1 \cap B_X) - \mathrm{dist}(x, K_2 \cap B_X)| \\ &\leq 2\varrho(K_1,K_2).\end{aligned}$$

This proves the right inequality in (26). The right inequality in (27) is due to Lemma 3.6(b). The remaining inequalities are obvious. The constant $\log 2$ in (28) is the best possible. □

The next corollary is one among the many conclusions that can be drawn from Theorem 3.8.

Corollary 3.9. *Let X be a normed space. Then, the metrics ϱ, μ, σ and λ induce the same topology on $\Xi(X)$. Such topology admits as basis the collection of sets of the form*

$$\mathcal{O}_r(K) = \{Q \in \Xi(X) : \delta(Q,K) < r\}$$

with $K \in \Xi(X)$ and $r > 0$.

The set $\mathcal{O}_r(K)$ is said to be an open "gap-ball" with center K and radius r. Note that $\mathcal{O}_r(K)$ is not a ball in the usual sense because δ is not a metric, unless, of course, the underlying space X is Hilbert.

4. Comparison with Kuratowski convergence

If not stated otherwise, convergence in $\Xi(X)$ is understood in the sense of the truncated Pompeiu-Hausdorff metric. In view of Corollary 3.9, this is equivalent to convergence with respect to the metrics μ, σ or λ. It is also equivalent to convergence with respect to the topology induced by the gap-balls.

Definition 4.1. A sequence $\{K_n\}_{n\in\mathbb{N}}$ in $\Xi(X)$ is convergent if there is some $K \in \Xi(X)$ such that $\varrho(K_n, K) \to 0$ as $n \to \infty$. In such a case, the set K is called the limit of $\{K_n\}_{n\in\mathbb{N}}$.

There are also other modes of set-convergence disseminated in the literature. Perhaps the most popular one is the so-called Kuratowski convergence. For a sequence $\{C_n\}_{n\in\mathbb{N}}$ of sets in a normed space X, one introduces the notation

$$\liminf_{n\to\infty} C_n = \{x \in X : \lim_{n\to\infty} \operatorname{dist}(x, C_n) = 0\}$$
$$\limsup_{n\to\infty} C_n = \{x \in X : \liminf_{n\to\infty} \operatorname{dist}(x, C_n) = 0\}$$

to indicate, respectively, its lower and its upper Kuratowski limit. Such limits were introduced at the beginning of the 20th century by Paul Painlevé (1863-1933), but the Polish mathematician Kazimierz Kuratowski (1896-1980) popularized their use in his famous textbook on Set Theory and Topology [22].

Definition 4.2. A sequence $\{C_n\}_{n\in\mathbb{N}}$ of sets in a normed space X is Kuratowski convergent if there exists a set C such that $\limsup_{n\to\infty} C_n \subset C \subset \liminf_{n\to\infty} C_n$. In such a case, one writes $C = \lim_{n\to\infty} C_n$ and refers to this set as the Kuratowski limit of $\{C_n\}_{n\in\mathbb{N}}$.

Some basic facts concerning the Kuratowski convergence of closed convex cones are recorded in the next proposition. All these facts are well known (cf. [2, 3, 27]).

Proposition 4.3. *Let $\{K_n\}_{n\in\mathbb{N}}$ be closed convex cones in a normed space X. Then,*

(a) $\liminf_{n\to\infty} K_n$ *is a closed convex cone.*

(b) $\limsup_{n\to\infty} K_n$ *is not necessarily convex. However, it is closed and stable under multiplication by a positive scalar.*

(c) *If $\{K_n\}_{n\in\mathbb{N}}$ is nondecreasing, then it is Kuratowski convergent and* $\lim_{n\to\infty} K_n = \operatorname{cl}[\cup_{n\in\mathbb{N}} K_n]$.

(d) *If $\{K_n\}_{n\in\mathbb{N}}$ is nonincreasing, then it is Kuratowski convergent and* $\lim_{n\to\infty} K_n = \cap_{n\in\mathbb{N}} K_n$.

Below we compare convergence in the sense of Kuratowski and convergence with respect to the metric ϱ. Proposition 4.4 is an adaptation to the conic setting of a more general result on convergence of nonempty closed sets (cf. [27, Section 4C]). Finite dimensionality is an essential assumption in Proposition 4.4. This fact will be illustrated with the help of two examples.

Proposition 4.4. *Let X be a finite dimensional normed space. Then, a sequence $\{K_n\}_{n\in\mathbb{N}}$ converges in the metric space $(\Xi(X), \varrho)$ if and only if it is Kuratowski convergent.*

Example 4.5. In the Hilbert space $\ell_2(\mathbb{R})$, consider the sequence $\{K_n\}_{n\geq 2}$ of closed convex cones given by

$$K_n = \{x \in \ell_2(\mathbb{R}) : x_1 \geq \ldots \geq x_n\}.$$

No constraint is imposed on the tail $\{x_k\}_{k\geq n+1}$ of x. The inclusion $K_{n+1} \subset K_n$ implies that $\{K_n\}_{n\geq 2}$ admits

$$\cap_{n\geq 2} K_n = \{x \in \ell_2(\mathbb{R}) : x \text{ is nonincreasing}\}$$

as Kuratowski limit. We now evaluate the distance between two successive terms of $\{K_n\}_{n\geq 2}$. Since we are working in a Hilbert space, one has

$$\varrho(K_n, K_{n+1}) = \delta(K_n, K_{n+1}) = \sup_{\substack{x\in K_n\\ \|x\|=1}} \operatorname{dist}(x, K_{n+1}).$$

By taking x as the $(n+1)$-th canonical vector e_{n+1} of $\ell_2(\mathbb{R})$, one gets

$$\varrho(K_n, K_{n+1}) \geq \operatorname{dist}[e_{n+1}, K_{n+1}] = \sqrt{n/(n+1)} \geq \sqrt{2/3}\,.$$

This prevents the convergence of $\{K_n\}_{n\geq 2}$ with respect to the metric ϱ.

Example 4.6. Again in the Hilbert space $\ell_2(\mathbb{R})$, let $\{K_n\}_{n\geq 1}$ be given by

$$K_n = \{x \in \ell_2(\mathbb{R}) : x_1 \geq 0, \ldots, x_n \geq 0 \text{ and } x_k = 0 \text{ for all } k \geq n+1\}.$$

This time $K_n \subset K_{n+1}$, and therefore $\{K_n\}_{n\geq 1}$ admits

$$\operatorname{cl}\,[\cup_{n\geq 1} K_n] = \{x \in \ell_2(\mathbb{R}) : x_n \geq 0\ \forall n \geq 1\}$$

as Kuratowski limit. However, for all n and m with $m \geq n+1$, one has

$$\varrho(K_n, K_m) = \delta(K_n, K_m) = \sup_{\substack{x\in K_m\\ \|x\|=1}} \operatorname{dist}(x, K_n) = 1.$$

Hence, $\{K_n\}_{n\geq 1}$ does not converge with respect to the metric ϱ.

Examples 4.5 and 4.6 are quite astonishing from at least one point of view. The metric ϱ actually does not care if we are dealing with a sequence that is upward or downward monotonic. Monotonicity alone does not ensure convergence in $(\Xi(X), \varrho)$, and, of course, this is a striking difference with respect to Kuratowski convergence.

5. Walkup-Wets Isometry Theorem

Hilbert space analysts and linear control theorists are well aware of the many applications of Proposition 2.10. Perhaps the most important consequence of that proposition is the isometry relation

$$\delta(K_1^\perp, K_2^\perp) = \delta(K_1, K_2) \tag{29}$$

involving a pair (K_1, K_2) of closed linear subspaces of a Hilbert space, and the corresponding pair $(K_1^\perp, K_2^\perp)$ of orthogonal linear subspaces. In fact, the equality (29) can be formulated in a much broader context. To do this, we need to introduce first some notation. Recall that

$$K^+ = \{y \in X^* : y(x) \geq 0 \text{ for all } x \in K\} \tag{30}$$

stands for the *dual cone* of $K \in \Xi(X)$. By construction, the set (30) belongs to $\Xi(X^*)$. As usual, one endows the topological dual space X^* with the dual norm $\|\cdot\|_*$. The gap function on $\Xi(X^*)$ will be denoted by δ_* in order to emphasize duality. Thus, for any pair Q_1, Q_2 in $\Xi(X^*)$, one sets

$$\delta_*(Q_1, Q_2) = \max\left\{ \sup_{x \in Q_1 \cap S_{X^*}} \operatorname{dist}(x, Q_2), \sup_{x \in Q_2 \cap S_{X^*}} \operatorname{dist}(x, Q_1) \right\}.$$

The next result was originally established by Walkup and Wets [29] for reflexive Banach spaces. Its extension to arbitrary normed spaces can be found in the book by Beer [4].

Theorem 5.1. *Let X be a normed space. Then, for all $K_1, K_2 \in \Xi(X)$, one has*

$$\delta_*(K_1^+, K_2^+) = \delta(K_1, K_2).$$

Theorem 5.1 has known an impressive amount of applications, including theoretical questions pertaining to the realm of convex analysis and optimization: continuity of the Legendre-Fenchel transformation [6, 25], comparison of sharpness and solidity indices [14, 16], and so on. An equivalent formulation of Theorem 5.1 reads as follows. The notation λ_* stands of course for the logarithmic gap metric on $\Xi(X^*)$.

Theorem 5.2. *Let X be a normed space. Then, the duality operation $K \mapsto K^+$ is an isometry between the metric spaces $(\Xi(X), \lambda)$ and $(\Xi(X^*), \lambda_*)$.*

Whenever we refer to the Walkup-Wets Isometry Theorem, it is the version stated in Theorem 5.2 what we have in mind, and not the original formulation given in Theorem 5.1. Recall that δ is not even a metric. What makes the logarithmic gap metric so special with respect to dualization? We confess that this is something rather misterious to us. It should be mentioned that the Walkup-Wets Isometry Theorem cannot be formulated in terms of the spherical metric. Such a "default" in the spherical metric was already noticed by Ostrovskii [23] while dealing with closed subspaces of a Banach space. Ostrovskii's counter-example is as follows.

Example 5.3. Let $X = \mathbb{R}^2$ be equipped with the norm $\|x\| = |x_1| + |x_2|$. Hence, the dual space $X^* = \mathbb{R}^2$ is equipped with $\|y\|_* = \max\{|y_1|, |y_2|\}$. Consider a parameter $a \in]0, 1[$. A matter of computation shows that the spherical distance between the closed linear subspaces

$$K_1 = \{(u_1, u_2) \in \mathbb{R}^2 : au_1 - u_2 = 0\},$$
$$K_2 = \{(u_1, u_2) \in \mathbb{R}^2 : u_2 = 0\}$$

is $\sigma(K_1, K_2) = 2a/(1 + a)$. On the other hand, the spherical distance between the corresponding orthogonal linear subspaces

$$K_1^\perp = \{(w_1, w_2) \in \mathbb{R}^2 : w_1 + aw_2 = 0\},$$
$$K_2^\perp = \{(w_1, w_2) \in \mathbb{R}^2 : w_1 = 0\}$$

is $\sigma_*(K_1^\perp, K_2^\perp) = a$. Hence, $\sigma_*(K_1^\perp, K_2^\perp) \neq \sigma(K_1, K_2)$.

Remark 5.4. The same default is observed in the truncated Pompeiu-Hausdorff metric and in the uniform metric. However, by working out some easy examples (rays, half-spaces, etc), we noticed that $\varrho(K_1, K_2)$ was always equal to $\mu_*(K_1^+, K_2^+)$. Is this a mere coincidence? We leave as conjecture the following statement: if X is a normed space, then the duality operation $K \mapsto K^+$ is an isometry between the metric spaces $(\Xi(X), \varrho)$ and $(\Xi(X^*), \mu_*)$.

6. Particular classes of convex cones

Sometimes one must deal with convex cones posssesing a special structure. If that is the case, then it is natural to consider an ad hoc metric that takes into account the specific context.

6.1. Fitted cones

In this section one equips the product space $X \times \mathbb{R}$ with the norm $\|(x,t)\| = [\|x\|^2 + t^2]^{1/2}$. A common way of constructing convex cones in $X \times \mathbb{R}$ is by choosing a convex set C in $\mathcal{B}(X)$ and forming

$$F(C) = \mathbb{R}_+(C \times \{1\}) \tag{31}$$

$$= \{(x,t) \in X \times \mathbb{R} : \; t \geq 0, \; x \in tC\}. \tag{32}$$

The equality (31) is the definition of $F(C)$, whereas (32) is a useful characterization. Clearly, $F(C)$ belongs to $\Xi(X \times \mathbb{R})$. One says that $F(C)$ is the convex cone *fitted* by C. This terminology, although not widely spread in the literature, is used by a number of authors [10, 28]. Be aware, however, that not everyone asks the same properties to the ingredient set C.

Lemma 6.1. *Let X be a normed space. Then, F is a bijection between the sets*

$$\mathcal{C}(X) = \{C \in \mathcal{B}(X) : C \textit{ is convex}\},$$
$$\mathcal{F}(X) = \{K \in \Xi(X \times \mathbb{R}) : K \textit{ is fitted}\}.$$

Its inverse $F^{-1} : \mathcal{F}(X) \to \mathcal{C}(X)$ is given by $F^{-1}(K) = \{x \in X : (x,1) \in K\}$.

Note that $\mathcal{F}(X)$ is a proper subset of $\Xi(X \times \mathbb{R})$. In view of the above lemma, a natural way of measuring the distance between two fitted cones is by mean of the expression

$$\sigma_{\text{fit}}(K_1, K_2) = \text{haus}(F^{-1}(K_1), F^{-1}(K_2)),$$

i.e., one evaluates the Pompeiu-Hausdorf distance between the sets C_1 and C_2 that serve to fit K_1 and K_2, respectively.

Example 6.2. Let $X = \mathbb{R}^n$ be equipped with its usual norm. Let $K_r = F(rB_X)$ be the convex cone fitted by a ball of radius r. In particular, K_0 corresponds to a ray. One gets

$$\sigma_{\text{fit}}(K_r, K_0) = r,$$
$$\delta(K_r, K_0) = r/\sqrt{1 + r^2}.$$

Note that $\sigma_{\text{fit}}(K_r, K_0)$ and $\delta(K_r, K_0)$ behave in a different way when $r \to \infty$. The latter expression approaches 1, where as the former diverges to ∞.

The next proposition compares the "fitted" metric σ_{fit} with the spherical metric. Being consistent with Definition 3.1, we denote by $c_{X\times\mathbb{R}}$ the sphericity coefficient of the normed space $X\times\mathbb{R}$. We know already that $c_{X\times\mathbb{R}}\in[1,2]$ if X is a general normed space and that $c_{X\times\mathbb{R}}=1$ if X is pre-Hilbert. Without further ado, we state:

Proposition 6.3. *Let X be a normed space. Then, σ_{fit} is a metric on $\mathcal{F}(X)$ and*

$$\sigma(K_1,K_2)\le c_{X\times\mathbb{R}}\,\sigma_{\mathrm{fit}}(K_1,K_2) \tag{33}$$

for all $K_1,K_2\in\mathcal{F}(X)$. However, σ_{fit} and σ are not Lipschitz equivalent on $\mathcal{F}(X)$.

Proof. Let K_1 and K_2 be fitted by C_1 and C_2, respectively. Let us examine the expression

$$e_{1,2}=\sup_{(x,t)\in K_1\cap S_{X\times\mathbb{R}}}\mathrm{dist}\left((x,t),K_2\cap S_{X\times\mathbb{R}}\right).$$

By using the characterization (32) of a fitted cone, one gets

$$e_{1,2}=\sup_{\substack{u\in C_1,\,t\ge0\\ \|t(u,1)\|=1}}\ \inf_{\substack{v\in C_2,\,s\ge0\\ \|s(v,1)\|=1}}\|t(u,1)-s(v,1)\|.$$

After getting rid of the variables t and s, one arrives at

$$e_{1,2}=\sup_{u\in C_1}\inf_{v\in C_2}\vartheta(u,v)$$

with $\vartheta(u,v)=\alpha\left((u,1),(v,1)\right)$ standing for the angular distance between $(u,1)$ and $(v,1)$. Given the definition of $c_{X\times\mathbb{R}}$, it is clear that

$$\vartheta(u,v)\le c_{X\times\mathbb{R}}\,\|u-v\|,$$

so one ends up with the inequality

$$e_{1,2}\le c_{X\times\mathbb{R}}\sup_{u\in C_1}\mathrm{dist}(u,C_2).$$

A similar estimate is obtained for the symmetric term $e_{2,1}$. This proves (33). Finally, as seen in Example 6.2, the diameter of the metric space $(\mathcal{F}(X),\sigma_{\mathrm{fit}})$ is infinite. Since $(\mathcal{F}(X),\sigma)$ has a finite diameter, the metrics σ_{fit} and σ cannot be Lipschitz equivalent. $\square$

6.2. Spectral convex cones

This section concerns a special class of convex cones in the space $\mathbb{S}_n$ of real symmetric matrices of order n. As usual, $\mathbb{S}_n$ is equipped with the trace inner product $\langle A,B\rangle=\mathrm{trace}(AB)$ and the associated norm. A convex cone $\mathcal{K}$ in $\mathbb{S}_n$ is called *spectral* (or weakly unitarily invariant) if

$$A\in\mathcal{K}\quad\Longrightarrow\quad U^TAU\in\mathcal{K}\ \text{ for all } U\in\mathcal{O}_n \tag{34}$$

with $\mathcal{O}_n$ denoting the group of orthogonal matrices of size $n\times n$. In fact, the concept of spectrality applies to an arbitrary set in $\mathbb{S}_n$ and not just to a convex cone. The next lemma is part of the folklore on spectral cones (cf. [15]). The notation $\lambda(A)=(\lambda_1(A),\ldots,\lambda_n(A))^T$ stands for the vector of eigenvalues of A arranged in nondecreasing order, and $\mathrm{diag}(x)$ stands for the diagonal matrix whose entries on the diagonal are the components of the vector x.

Lemma 6.4. *A convex cone $\mathcal{K}$ in $\mathbb{S}_n$ is spectral if and only if there is a permutation invariant convex cone Q in $\mathbb{R}^n$ such that*

$$\mathcal{K} = \{A \in \mathbb{S}_n : \lambda(A) \in Q\}.$$

Furthermore, such Q is unique and given by

$$Q_{\mathcal{K}} = \{x \in \mathbb{R}^n : \operatorname{diag}(x) \in \mathcal{K}\}.$$

One refers to $Q_{\mathcal{K}}$ as the permutation invariant convex cone induced by $\mathcal{K}$. Recall that a set Q in $\mathbb{R}^n$ is called *permutation invariant* if $\Pi(Q) = Q$ for all permutation matrix Π of order n. A list of interesting examples of spectral convex cones is provided in [15]. What makes spectral convex cones so attractive is that everything boils down to working with their corresponding permutation invariant convex cones.

As shown in the next proposition, measuring the distance between two spectral convex cones, say $\mathcal{K}_1$ and $\mathcal{K}_2$, is the same thing as measuring the distance between $Q_{\mathcal{K}_1}$ and $Q_{\mathcal{K}_2}$. We start by recalling a commutation principle for optimization problems with spectral data. As mentioned before, a spectral set in $\mathbb{S}_n$ is defined by means of the relation (34). Similarly, a function g on $\mathbb{S}_n$ is said to be spectral if $g(U^TAU) = g(A)$ for all $U \in \mathcal{O}_n$.

Lemma 6.5. *Let $\mathcal{N} \subset \mathbb{S}_n$ be a spectral set and $g : \mathbb{S}_n \to \mathbb{R}$ be a spectral function. Let $A_0 \in \mathbb{S}_n$. If B_0 is a local minimum (or a local maximum) of the function $B \in \mathcal{N} \mapsto g(B) + \langle A_0, B\rangle$, then A_0 and B_0 commute.*

Everything is now ready to state:

Proposition 6.6. *Suppose that $\mathcal{K}_1, \mathcal{K}_2 \in \Xi(\mathbb{S}_n)$ are spectral. Then,*

$$\mu(\mathcal{K}_1, \mathcal{K}_2) = \mu(Q_{\mathcal{K}_1}, Q_{\mathcal{K}_2}), \tag{35}$$

$$\sigma(\mathcal{K}_1, \mathcal{K}_2) = \sigma(Q_{\mathcal{K}_1}, Q_{\mathcal{K}_2}). \tag{36}$$

Proof. To start with, we claim that

$$\operatorname{dist}(A, \mathcal{K}) = \operatorname{dist}(\lambda(A), Q_{\mathcal{K}}). \tag{37}$$

for any $A \in \mathbb{S}_n$ and any $\mathcal{K} \in \Xi(\mathbb{S}_n)$ that is spectral. As a consequence of Lemma 6.5, the matrix $B \in \mathcal{K}$ achieving the minimal distance

$$\operatorname{dist}(A, \mathcal{K}) = \inf_{Z \in \mathcal{K}} \|A - Z\| = \|A - B\| \tag{38}$$

commutes with the matrix A. In such a case, A and B can be simultaneously diagonalized by means of a matrix $U \in \mathcal{O}_n$, i.e.,

$$U^TAU = \operatorname{diag}(u),$$
$$U^TBU = \operatorname{diag}(v)$$

with $u, v \in \mathbb{R}^n$. By permuting the columns of U if necessary, one may suppose that $u = \lambda(A)$. By plugging this information in (38) and simplifying, one arrives at (37).

The proof of (35) is now easy:

$$\begin{aligned}\mu(\mathcal{K}_1,\mathcal{K}_2) &= \sup_{\|A\|=1} |\mathrm{dist}(A,\mathcal{K}_1) - \mathrm{dist}(A,\mathcal{K}_2)| \\ &= \sup_{\|A\|=1} |\mathrm{dist}(\lambda(A),Q_{\mathcal{K}_1}) - \mathrm{dist}(\lambda(A),Q_{\mathcal{K}_2})| \\ &= \sup_{\|u\|=1} |\mathrm{dist}(u,Q_{\mathcal{K}_1}) - \mathrm{dist}(u,Q_{\mathcal{K}_2})| \\ &= \mu(Q_{\mathcal{K}_1},Q_{\mathcal{K}_2}).\end{aligned}$$

The equality (36) can be proven along the same lines. □

6.3. Simplicial cones

Let $\Xi_{\mathrm{simpl}}(\mathbb{R}^n)$ be the collection of all simplicial cones in $\mathbb{R}^n$. That K belongs to $\Xi_{\mathrm{simpl}}(\mathbb{R}^n)$ means that K is a polyhedral convex cone generated by n linearly independent unit vectors of $\mathbb{R}^n$. Such basis is necessarily unique and is denoted by $\mathrm{gen}(K)$. For economy of words, one refers to $\mathrm{gen}(K)$ as the generator set of K.

A natural way of measuring the distance between two elements K_1, K_2 in $\Xi_{\mathrm{simpl}}(\mathbb{R}^n)$ is by evaluating the optimal matching distance

$$d_{\mathrm{gen}}(K_1,K_2) = \min_{\pi\in \mathrm{perm}(n)} \max_{1\leq i\leq n} \|a_i - b_{\pi(i)}\| \tag{39}$$

between the corresponding generator sets

$$\begin{aligned}\mathrm{gen}(K_1) &= \{a_1,\ldots,a_n\},\\ \mathrm{gen}(K_2) &= \{b_1,\ldots,b_n\}.\end{aligned}$$

Here, the symbol $\mathrm{perm}(n)$ stands for the set of all permutations of $\{1,\ldots,n\}$ and $\|\cdot\|$ refers to the usual norm of $\mathbb{R}^n$. Note that the evaluation of a minimum like (39) is a straightforward task. However, the use of d_{gen} makes sense only if we have an efficient algorithm for identifying the generator set of a simplicial cone. One must also bear in mind that the cardinality of $\mathrm{perm}(n)$ grows very rapidly with n.

The optimal matching distance is used in many areas of mathematics, including statistics, linear algebra, and graph theory (cf. [5, 7, 12]). The proposition below exploits the identification of a simplicial cone with its generator set.

Proposition 6.7. *Let $n \geq 2$. Then,*

(a) *d_{gen} is a metric on $\Xi_{\mathrm{simpl}}(\mathbb{R}^n)$.*
(b) *$d_{\mathrm{gen}}(K_1,K_2) \leq 2$ for all $K_1,K_2 \in \Xi_{\mathrm{simpl}}(\mathbb{R}^n)$.*
(c) *For all $U \in \mathcal{O}_n$ and K_1, $K_2 \in \Xi_{\mathrm{simpl}}(\mathbb{R}^n)$, one has $d_{\mathrm{gen}}(U(K_1),\ U(K_2)) = d_{\mathrm{gen}}(K_1,K_2)$.*

6.4. By way of conclusion

Here are the main lessons that can be drawn from this work:

- There are several ways of mesuring the distance between a pair of nontrivial closed convex cones in a normed space. Specially tailored metrics are also availlable if one needs to deal with highly structured cones (like, for instance, fitted, simplicial, or spectral cones).

- The truncated Pompeiu-Hausdorff metric, the uniform metric, the spherical metric, and the logarithmic gap metric are the most important examples. These four metrics are bounded and isometrically invariant. The four of them are not just topologically equivalent, but also Lipschitz equivalent.
- One of the Lipschitz constants relating the uniform metric and the truncated Pompeiu-Hausdorff metric has a very interesting geometrical interpretation: it is the sphericity coefficient of the underlying normed space.
- Thanks to the Walkup-Wets Isometry Theorem, the logarithmic gap metric is the best suited for dealing with duality issues.

Set	Metric	Diameter of set	Isometry invariance?	Walkup-Wets formula?
$\Xi(X)$	ϱ	1	yes	no
$\Xi(X)$	μ	1	yes	no
$\Xi(X)$	σ	2	yes	no
$\Xi(X)$	λ	$\log 2$	yes	yes
$\mathcal{F}(X)$	$\sigma_{\rm fit}$	∞	–	–
$\Xi_{\rm simpl}(\mathbb{R}^n)$	$d_{\rm gen}$	2	yes	–

Table 6.1: Possible ways of measuring distances between closed convex cones.

References

[1] N. I. Akhiezer, I. M. Glazman: Theory of Linear Operators in Hilbert Space. Vol. I, Pitman, Boston (1981).

[2] H. Attouch: Variational Convergence for Functions and Operators, Pitman, Boston (1984).

[3] J. P. Aubin, H. Frankowska: Set-Valued Analysis, Birkhäuser, Boston (1990).

[4] G. Beer: Topologies on Closed and Closed Convex Sets, Kluwer, Dordrecht (1993).

[5] R. Bhatia: Matrix Analysis, Springer, New York (1996).

[6] L. Contesse, J. P. Penot: Continuity of the Fenchel correspondence and continuity of polarities, J. Math. Anal. Appl. 156 (1991) 305–328.

[7] B. Ćurgus, V. Mascioni: Perturbations of roots under linear transformations of polynomials, Constr. Approx. 25 (2007) 255–277.

[8] A. Feintuch: Robust Control Theory in Hilbert Space, Springer, New York (1998).

[9] I. C. Gohberg, A. S. Marcus: Two theorems on the gap between subspaces of a Banach space, Usp. Mat. Nauk 14(5)(89) (1959) 135–140.

[10] O. Güler: Barrier functions in interior point methods, Math. Oper. Res. 21 (1996) 860–885.

[11] V. I. Gurariĭ: Openings and inclinations of subspaces of a Banach space, Teor. Funkts. Funkts. Anal. Prilozh. 1 (1965) 194–204.

[12] F. Hiai, Y. Nakamura: Distance between unitary orbits in von Neumann algebras, Pacific J. Math. 138 (1989) 259–294.

[13] L. Hörmander: Sur la fonction d'appui des ensembles convexes dans un espace localement convexe, Ark. Mat. 3 (1955) 181–186.

[14] A. Iusem, A. Seeger: Axiomatization of the index of pointedness for closed convex cones, Comput. Appl. Math. 24 (2005) 245–283.

[15] A. Iusem, A. Seeger: Angular analysis of two classes of non-polyhedral convex cones: the point of view of optimization theory, Comput. Appl. Math. 26 (2007) 191–214.

[16] A. Iusem, A. Seeger: Normality and modulability indices. Part I: convex cones in normed spaces, J. Math. Anal. Appl. 338 (2008) 365–391.

[17] A. Iusem, A. Seeger: Normality and modulability indices. Part II: convex cones in Hilbert spaces, J. Math. Anal. Appl. 338 (2008) 392–406.

[18] W. A. Kirk, M. F. Smiley: Another characterization of inner product spaces, Amer. Math. Monthly 71 (1964) 890–891.

[19] M. Kisielewicz: Differential Inclusions and Optimal Control, PWN, Warsaw (1991).

[20] M. A. Krasnoselskiĭ, M. G. Krein: Fundamental theorems on the extension of Hermitian operators and certain of their applications to the theory of orthogonal polynomials and the problem of moments, Usp. Mat. Nauk 2(3)(19) (1947) 60–106.

[21] M. G. Krein: Propriétés fondamentales des ensembles coniques normaux dans l'espace de Banach, C. R. (Dokl.) Acad. Sci. URSS, n. Ser. 28 (1940) 13–17.

[22] K. Kuratowski: Introduction to Set Theory and Topology, Addison-Wesley, Reading (1962).

[23] M. I. Ostrovskiĭ: Topologies on the set of all subspaces of a Banach space and related questions of Banach space geometry, Quaest. Math. 17(3) (1994) 259–319.

[24] V. I. Paraska: A remark on the distance between linear operators in the gap metric, Mat. Issled. 10(2)(36) (1975) 271–274.

[25] J. P. Penot, C. Zălinescu: Continuity of the Legendre-Fenchel transform for some variational convergences, Optimization 53 (2004) 549–562.

[26] D. Pompeiu: Sur la Continuité des Fonctions des Variables Complexes, Ph.D. Thesis, Gauthier-Villars, Paris (1905); Toul. Ann. 7 (1905) 265–315.

[27] R. T. Rockafellar, R. J.-B. Wets: Variational Analysis, Springer, Berlin (1998).

[28] S. Shimizu: A remark on homogeneous convex domains, Nagoya Math. J. 105 (1987) 1–7.

[29] D. W. Walkup, R. J.-B. Wets: Continuity of some convex-cone-valued mappings, Proc. Amer. Math. Soc. 18 (1967) 229–235.

[30] A. Seeger: Stabilization of the point spectral mapping, J. Nonlinear Convex Anal. 4 (2003) 277–285.