C. Sublinearity and Support Functions

Introduction In classical real analysis, the simplest functions are *linear*. In convex analysis, the next simplest convex functions (apart from the affine functions, widely used in §B.1.2), are so-called *sublinear*. We give three motivations for their study.

(i) A suitable generalization of linearity. A linear function ℓ from $\mathbb{R}^n$ to $\mathbb{R}$, or a linear form on $\mathbb{R}^n$, is primarily defined as a function satisfying for all $(x_1, x_2) \in \mathbb{R}^n \times \mathbb{R}^n$ and $(t_1, t_2) \in \mathbb{R} \times \mathbb{R}$:

$$\ell(t_1 x_1 + t_2 x_2) = t_1 \ell(x_1) + t_2 \ell(x_2) . \tag{0.1}$$

A corresponding definition for a sublinear function σ from $\mathbb{R}^n$ into $\mathbb{R}$ is: for all $(x_1, x_2) \in \mathbb{R}^n \times \mathbb{R}^n$ and $(t_1, t_2) \in \mathbb{R}^+ \times \mathbb{R}^+$,

$$\sigma(t_1 x_1 + t_2 x_2) \leqslant t_1 \sigma(x_1) + t_2 \sigma(x_2) . \tag{0.2}$$

A first observation is that requiring an inequality in (0.2), rather than an equality, allows infinite values for σ without destroying the essence of the concept of sublinearity. Of course, (0.2) is less stringent than (0.1), but more stringent than the definition of a convex function: the inequality must hold in (0.2) even if $t_1 + t_2 \neq 1$. This confirms that sublinear functions, which generalize linear functions, are particular instances of convex functions.

Remark 0.1 Note that (0.1) and (0.2) can be made more similar by restricting t_1 and t_2 in (0.1) to be positive – this leaves unchanged the definition of a linear function.

The prefix "sub" comes from the inequality-sign "$\leqslant$" in (0.2). It also suggests that sublinearity is less demanding than linearity, but this is a big piece of luck. In fact, draw the graph of a convex and of a concave function and ask a non-mathematician: "which is convex?". He will probably give the wrong answer. Yet, if convex functions were defined the other way round, (0.2) should have the "$\geqslant$" sign. The associated concept would be superlinearity, an unfortunate wording which suggests *more* or *better* than linear. □

In a word, sublinear functions are reasonable candidates for "simplest non-trivial convex functions". Whether they are interesting candidates will be seen in (ii) and (iii). Here, let us just mention that their epigraphs are convex cones, the next simplest convex epigraphs after half-spaces.

(ii) Tangential approximation of convex functions. To say that a function $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at x is to say that there is a linear function ℓ_x which approximates $f(x+h) - f(x)$ to first order, i.e.

$$f(x+h) - f(x) = \ell_x(h) + o(\|h\|) .$$

This fixes the rate of change of f when x is moved along a line d: with $\varepsilon(t) \to 0$ if $t \to 0$,

$$\frac{f(x+td) - f(x)}{t} = \ell_x(d) + \varepsilon(t) \quad \text{for all } t \neq 0 .$$

Geometrically, the graph of f has a tangent hyperplane at $(x, f(x)) \in \mathbb{R}^n \times \mathbb{R}$; and this hyperplane is the graph of the affine function $h \mapsto f(x) + \ell_x(h)$.

When f is merely convex, its graph may have no tangent hyperplane at a given $(x, f(x))$. Nevertheless, under reasonable assumptions, $f(x+h) - f(x)$ can still be approximated to first order by a function which is sublinear: there exists a sublinear function $h \mapsto \sigma_x(h)$ such that

$$f(x+h) - f(x) = \sigma_x(h) + o(\|h\|) .$$

This will be seen in Chap. D.

Geometrically, gr σ_x is no longer a hyperplane but rather a cone, which is therefore tangent to gr f (the word "tangent" should be understood here in its intuitive meaning of a tangent surface, as opposed to tangent cones of Chap. A; neither gr σ_x nor gr f are convex). Thus, one can say that differentiable functions are "tangentially linear", while convex functions are "tangentially sublinear". See Fig. 0.1, which displays the graphs of a differentiable and of a convex function. The graph of ℓ_x is the thick line L, while the graph of σ_x is made up of the two thick half-lines S_1 and S_2.

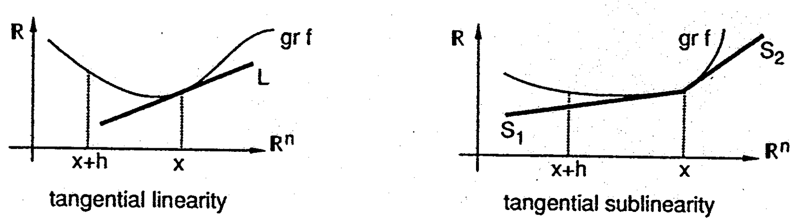

Fig. 0.1. Two concepts of tangency

(iii) Nice correspondence with nonempty closed convex sets. In the Euclidean space $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$, a linear form ℓ can be represented by a vector: there is a unique $s \in \mathbb{R}^n$ such that

$$\ell(x) = \langle s, x \rangle \quad \text{for all } x \in \mathbb{R}^n . \tag{0.3}$$

The definition (0.3) of a linear function is more geometric than (0.1), and just as accurate. A large part of the present chapter will be devoted to generalizing the above representation theorem to sublinear functions.

First observe that, given a nonempty set $S \subset \mathbb{R}^n$, the function $\sigma_S : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ defined by

$$\sigma_S(x) := \sup \{\langle s, x \rangle \,:\, s \in S\} \tag{0.4}$$

is sublinear. It is called the *support function* of S, already encountered in Sects A.4.1 and B.1.3(a). When S is bounded, its support function is finite everywhere; otherwise, σ_S can take on the value $+\infty$ but it remains lower semi-continuous. Furthermore, it is easy to check that σ_S is also the support function of the closure of S, and even of the closed convex hull of S. It is therefore logical to consider support functions of nonempty closed convex sets only.

Now, a key result is that the mapping $S \mapsto \sigma_S$ is then bijective: a lower semi-continuous (i.e. closed) sublinear function is the support function of a *uniquely determined* nonempty closed convex set. Thus, (0.4) establishes the announced representation, just as (0.3) does in the linear case. Note that the linear case is covered: it corresponds to S being a singleton $\{s\}$ in (0.4).

This correspondence between nonempty closed convex sets of $\mathbb{R}^n$ and closed sublinear functions allows fruitful and enlightening geometric interpretations when studying these functions. Vice versa, it provides powerful analytical tools for the study of these sets. In particular, when closed convex sets are combined (intersected, added, etc.) to form new convex sets, we will show how their support functions are correspondingly combined: the mapping (0.4) is an *isomorphism*, with respect to a number of structures.

1 Sublinear Functions

1.1 Definitions and First Properties

Definition 1.1.1 A function $\sigma : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is said to be *sublinear* if it is convex and positively homogeneous (of degree 1): $\sigma \in \operatorname{Conv} \mathbb{R}^n$ and

$$\sigma(tx) = t\sigma(x) \quad \text{for all } x \in \mathbb{R}^n \text{ and } t > 0\,. \tag{1.1.1}$$

□

Remark 1.1.2 Inequality in (1.1.1) would be enough to define positive homogeneity: a function σ is positively homogeneous if and only if it satisfies

$$\sigma(tx) \leqslant t\sigma(x) \quad \text{for all } x \in \mathbb{R}^n \text{ and } t > 0\,. \tag{1.1.2}$$

In fact, (1.1.2) implies ($tx \in \mathbb{R}^n$ and $t^{-1} > 0$!)

$$\sigma(x) = \sigma(t^{-1}tx) \leqslant t^{-1}\sigma(tx)$$

which, together with (1.1.2), shows that σ is positively homogeneous. □

We deduce from (1.1.1) that $\sigma(0) = t\sigma(0)$ for all $t > 0$. This leaves only two possible values for $\sigma(0)$: 0 and $+\infty$. However, most of the sublinear functions to be encountered in the sequel do satisfy $\sigma(0) = 0$. According to our Definition B.1.1.3 of convex functions, σ should be finite somewhere; otherwise $\operatorname{dom} \sigma$ would be empty. Now, if $\sigma(x) < +\infty$, (1.1.1) shows that $\sigma(tx) < +\infty$ for all $t > 0$. In other words, $\operatorname{dom} \sigma$ is a cone, convex because σ is itself convex. Note that, being convex, σ is continuous relatively to ri dom σ, but discontinuities may occur on the boundary-rays of dom σ, including at 0.

The following result is a geometric characterization of sublinear functions.

Proposition 1.1.3 *A function* $\sigma : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *is sublinear if and only if its epigraph* epi σ *is a nonempty convex cone in* $\mathbb{R}^n \times \mathbb{R}$.

Proof. We know that σ is a convex function if and only if epi σ is a nonempty convex set in $\mathbb{R}^n \times \mathbb{R}$ (Proposition B.1.1.6). Therefore, we just have to prove the equivalence between positive homogeneity and epi σ being a cone.

Let σ be positively homogeneous. For $(x, r) \in \operatorname{epi}\sigma$, the relation $\sigma(x) \leqslant r$ gives

$$\sigma(tx) = t\sigma(x) \leqslant tr \quad \text{for all } t > 0\,,$$

so epi σ is a cone. Conversely, if epi σ is a cone in $\mathbb{R}^n \times \mathbb{R}$, the property $(x, \sigma(x)) \in \operatorname{epi}\sigma$ implies $(tx, t\sigma(x)) \in \operatorname{epi}\sigma$, i.e.

$$\sigma(tx) \leqslant t\sigma(x) \quad \text{for all } t > 0\,.$$

From Remark 1.1.2, this is just positive homogeneity. □

Another important concept in analysis is *subadditivity*: a function σ is subadditive when it satisfies

$$\sigma(x_1 + x_2) \leqslant \sigma(x_1) + \sigma(x_2) \quad \text{for all } (x_1, x_2) \in \mathbb{R}^n \times \mathbb{R}^n \tag{1.1.3}$$

– watch the difference with (0.2). Here again, the inequality is understood in $\mathbb{R} \cup \{+\infty\}$. Together with positive homogeneity, the above axiom gives another characterization (analytical, rather than geometrical) of sublinear functions.

Proposition 1.1.4 *A function* $\sigma : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, *not identically equal to* $+\infty$, *is sublinear if and only if one of the following two properties holds:*

$$\sigma(t_1x_1 + t_2x_2) \leqslant t_1\sigma(x_1) + t_2\sigma(x_2) \quad \textit{for all } x_1, x_2 \in \mathbb{R}^n \textit{ and } t_1, t_2 > 0\,, \tag{1.1.4}$$

or

$$\sigma \textit{ is positively homogeneous and subadditive}\,. \tag{1.1.5}$$

Proof. [*sublinearity* $\Rightarrow$ *(1.1.4)*] For $x_1, x_2 \in \mathbb{R}^n$ and $t_1, t_2 > 0$, set $t := t_1 + t_2 > 0$; we have

$$\begin{aligned} \sigma(t_1x_1 + t_2x_2) &= \sigma\left(t[\tfrac{t_1}{t}x_1 + \tfrac{t_2}{t}x_2]\right) \\ &= t\sigma\left(\tfrac{t_1}{t}x_1 + \tfrac{t_2}{t}x_2\right) && \text{[positive homogeneity]} \\ &\leqslant t\left[\tfrac{t_1}{t}\sigma(x_1) + \tfrac{t_2}{t}\sigma(x_2)\right], && \text{[convexity]} \end{aligned}$$

and (1.1.4) is proved.

[*(1.1.4)* $\Rightarrow$ *(1.1.5)*] A function satisfying (1.1.4) is obviously subadditive (take $t_1 = t_2 = 1$) and satisfies (take $x_1 = x_2 = x$, $t_1 = t_2 = 1/2\,t$)

$$\sigma(tx) \leqslant t\sigma(x) \quad \text{for all } x \in \mathbb{R}^n \text{ and } t > 0\,,$$

which is just positive homogeneity because of Remark 1.1.2.

[*(1.1.5)* $\Rightarrow$ *sublinearity*] Take $t_1, t_2 > 0$ with $t_1 + t_2 = 1$ and apply successively subadditivity and positive homogeneity:

$$\sigma(t_1x_1 + t_2x_2) \leqslant \sigma(t_1x_1) + \sigma(t_2x_2) = t_1\sigma(x_1) + t_2\sigma(x_2)\,,$$

hence σ is convex. □

Corollary 1.1.5 *If σ is sublinear, then*

$$\sigma(x) + \sigma(-x) \geqslant 0 \quad \textit{for all } x \in \mathbb{R}^n\,. \tag{1.1.6}$$

Proof. Take $x_2 = -x_1$ in (1.1.3) and remember that $\sigma(0) \geqslant 0$. □

It is worth mentioning that, to become sublinear, a positively homogeneous function just needs to be subadditive as well (rather than convex, as suggested by Definition 1.1.1); then, of course, it becomes convex at the same time. Figure 1.1.1 summarizes the connections between the classes of functions given so far. Note for completeness that a convex and subadditive function need not be sublinear: think of $f(x) \equiv 1$.

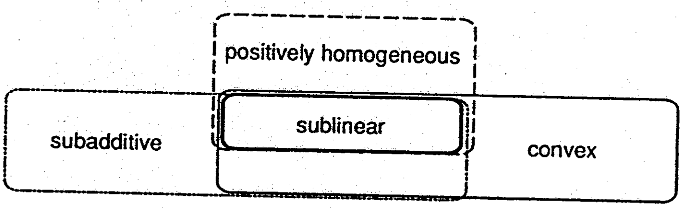

Fig. 1.1.1. Various classes of functions

Similarly, one can ask when a sublinear function becomes linear. For a linear function, (1.1.6) holds as an equality, and the next result implies that this is exactly what makes the difference.

Proposition 1.1.6 *Let σ be sublinear and suppose that there exist $x_1, \dots, x_m$ in* $\operatorname{dom}\sigma$ *such that*

$$\sigma(x_j) + \sigma(-x_j) = 0 \quad \textit{for } j = 1, \dots, m\,. \tag{1.1.7}$$

Then σ is linear on the subspace spanned by $x_1, \dots, x_m$.

Proof. With $x_1, \dots, x_m$ as stated, each $-x_j$ is in $\operatorname{dom}\sigma$. Let $x := \sum_{j=1}^m t_jx_j$ be an arbitrary linear combination of $x_1, \dots, x_m$; we must prove that $\sigma(x) = \sum_{j=1}^m t_j\sigma(x_j)$. Set

$$J_1 := \{j : t_j > 0\}, \quad J_2 := \{j : t_j < 0\}$$

and obtain (as usual, $\sum_\emptyset = 0$):

$$\begin{aligned}
\sigma(x) &= \sigma\left(\sum_{J_1} t_j x_j + \sum_{J_2}(-t_j)(-x_j)\right) \\
&\leqslant \sum_{J_1} t_j\sigma(x_j) + \sum_{J_2}(-t_j)\sigma(-x_j) && \text{[from (1.1.4)]} \\
&= \sum_{J_1} t_j\sigma(x_j) + \sum_{J_2} t_j\sigma(x_j) = \sum_{j=1}^m t_j\sigma(x_j) && \text{[from (1.1.7)]} \\
&= -\sum_{J_1} t_j\sigma(-x_j) - \sum_{J_2}(-t_j)\sigma(x_j) && \text{[from (1.1.7)]} \\
&\leqslant -\sigma\left(-\sum_{j=1}^m t_j x_j\right) && \text{[from (1.1.4)]} \\
&= -\sigma(-x) \leqslant \sigma(x)\,. && \text{[from (1.1.6)]}
\end{aligned}$$

In summary, we have proved $\sigma(x) \leqslant \sum_{j=1}^m t_j\sigma(x_j) \leqslant -\sigma(-x) \leqslant \sigma(x)$. □

Thanks to this result, we are entitled to define

$$U := \{x \in \mathbb{R}^n \, : \, \sigma(x) + \sigma(-x) = 0\} \tag{1.1.8}$$

which is a subspace of $\mathbb{R}^n$: the subspace in which σ is *linear*, sometimes called the *lineality space* of σ. Note that U nonempty corresponds to $\sigma(0) = 0$ (even if U reduces to $\{0\}$).

What is interesting in this concept is its geometric interpretation. If V is another subspace such that $U \cap V = \{0\}$, there holds by definition $\sigma(x)+\sigma(-x) > 0$ for all $0 \neq x \in V$. This means that, if $0 \neq d \in V$, σ is "V-shaped" along d: for $t > 0$, $\sigma(td) = \alpha t$ and $\sigma(-td) = \beta t$, for some α and β in $\mathbb{R} \cup \{+\infty\}$ such that $\alpha + \beta > 0$; whereas $\alpha + \beta$ would be 0 if d were in U. See Fig. 1.1.2 for an illustration. For d of norm 1, the number $\alpha + \beta$ above could be called the "lack of linearity" of σ along d: when restricted to the line d, the graph of σ makes an angle; when finite, the number $\alpha + \beta$ measures how acute this angle is.

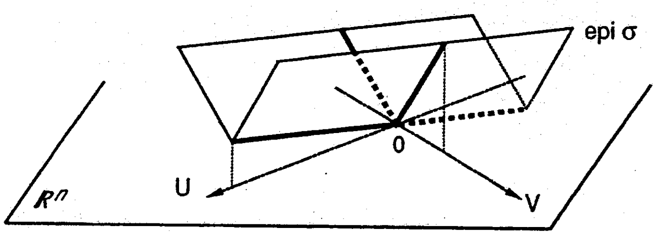

Fig. 1.1.2. Subspace of linearity of a sublinear function

Figure 1.1.2 suggests that $\operatorname{gr} \sigma$ is a hyperplane not only in U, but also in the translations of U: the restriction of σ to $\{y\} + U$ is affine, for any fixed y. This comes from the next result.

Proposition 1.1.7 *Let σ be sublinear. If $x \in U$, i.e. if*

$$\sigma(x) + \sigma(-x) = 0\,, \tag{1.1.9}$$

then there holds

$$\sigma(x + y) = \sigma(x) + \sigma(y) \quad \textit{for all } y \in \mathbb{R}^n\,. \tag{1.1.10}$$

Proof. In view of subadditivity, we just have to prove "$\geqslant$" in (1.1.10). Start from the identity $y = x + y - x$; apply successively subadditivity and (1.1.9) to obtain

$$\sigma(y) \leqslant \sigma(x+y) + \sigma(-x) = \sigma(x+y) - \sigma(x). \qquad \square$$

1.2 Some Examples

We start with some simple situations. If K is a nonempty convex cone, its indicator function

$$\mathrm{i}_K(x) := \begin{cases} 0 & \text{if } x \in K, \\ +\infty & \text{if not} \end{cases}$$

is clearly sublinear. In $\mathbb{R}^n \times \mathbb{R}$, the epigraph $\operatorname{epi} \mathrm{i}_K$ is made up of all the copies of K, shifted upwards. Likewise, its distance function

$$\mathrm{d}_K(x) := \inf \{ \|y - x\| \ : \ y \in K \}$$

is also sublinear: nothing in the picture is essentially changed when both x and y are multiplied by $t > 0$. Another example is the function from $\mathbb{R}^2$ to $\mathbb{R} \cup \{+\infty\}$

$$\sigma(x) = \sigma(\xi, \eta) := \begin{cases} -2\sqrt{\xi\eta} & \text{if } \quad \xi, \eta \geqslant 0 \\ +\infty & \text{if not.} \end{cases}$$

Its positive homogeneity is clear, its convexity is not particularly difficult to check (see Example B.4.3.3), it is therefore sublinear. A good exercise is to try to visualize its epigraph.

Example 1.2.1 Let $f \in \operatorname{Conv} \mathbb{R}^n$; its perspective $\tilde{f}$ of §B.2.2, which is convex, is clearly positively homogeneous (from $\mathbb{R}^{n+1}$ to $\mathbb{R} \cup \{+\infty\}$); it is an important instance of sublinear function. For example, in $\mathbb{R}^2$

$$\tilde{f}(u, \xi) := \begin{cases} \frac{1}{2}\xi^2/u & \text{if } u > 0, \\ +\infty & \text{if not} \end{cases} \tag{1.2.1}$$

is the perspective of $\xi \mapsto f(\xi) = 1/2\, \xi^2$.

Note that $\tilde{f}(0,0) = +\infty$. The closure of $\tilde{f}$ can be computed with the help of Example B.3.2.3: clearly enough, the asymptotic function of f is $\mathrm{i}_{\{0\}}$. Hence $(\mathrm{cl}\, \tilde{f})(0,0) = 0$, while $\tilde{f}$ coincides with its closure everywhere else. $\square$

Example 1.2.2 (Norms) We recall that a norm $\|\cdot\|$ on $\mathbb{R}^n$ is a function from $\mathbb{R}^n$ to $[0, +\infty[$ satisfying the following properties:

(i) $\|x\| = 0$ if and only if $x = 0$;
(ii) $\|tx\| = |t|\, \|x\|$ for all $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$;
(iii) $\|x_1 + x_2\| \leqslant \|x_1\| + \|x_2\|$ for all $(x_1, x_2) \in \mathbb{R}^n \times \mathbb{R}^n$.

Clearly, $\|\cdot\|$ is a positive (except at 0) and finite sublinear function which, moreover, is symmetric i.e. $\|-x\| = \|x\|$ for all x. It is linear on no line: the subspace U of (1.1.8) is reduced to $\{0\}$.

Conversely, if σ is a sublinear function from $\mathbb{R}^n$ into $[0, +\infty[$ which is linear on no line, i.e. such that

$$\sigma(x) + \sigma(-x) > 0 \quad \text{for all } x \neq 0,$$

then $|||x||| := \max\{\sigma(x), \sigma(-x)\}$ is a norm on $\mathbb{R}^n$. □

Example 1.2.3 (Quadratic Semi-Norms) Take a symmetric positive semi-definite operator Q from $\mathbb{R}^n$ to $\mathbb{R}^n$ and define

$$f(x) := \sqrt{\langle Qx, x\rangle} \quad \text{for all } x \in \mathbb{R}^n.$$

Convexity of f (i.e. its subadditivity, i.e. the Cauchy-Schwarz inequality) is rather tedious to prove directly. Consider, however, the convex set

$$E_Q := \{x \in \mathbb{R}^n \, : \, \langle Qx, x\rangle \leqslant 1\}.$$

Then f can be obtained as follows:

$$\begin{aligned} f(x) &= \inf\{\lambda > 0 \, : \, \langle Qx, x\rangle \leqslant \lambda^2\} \\ &= \inf\{\lambda > 0 \, : \, \langle Q\tfrac{x}{\lambda}, \tfrac{x}{\lambda}\rangle \leqslant 1\} \\ &= \inf\{\lambda > 0 \, : \, \tfrac{x}{\lambda} \in E_Q\} \end{aligned}$$

and we will see below that this establishes convexity – hence sublinearity – of f.

Observe in passing that E_Q is the sublevel-set at level 1 of both f and $f^2 = \langle Q\cdot, \cdot\rangle$. Decompose the space as $\mathbb{R}^n = \operatorname{Ker} Q \oplus \operatorname{Im} Q$: the intersection of E_Q with $\operatorname{Im} Q$ is an elliptic set centered at the origin, say $\widetilde{E}_Q$. The entire E_Q is the cylinder $\widetilde{E}_Q + \operatorname{Ker} Q$, whose asymptotic cone is just the subspace $\operatorname{Ker} Q$. If and only if $\operatorname{Ker} Q = \{0\}$, i.e. Q is positive definite, is E_Q compact; it is an elliptic body. On the other hand, f is finite, nonnegative, symmetric because E_Q has center 0; and f is zero on the asymptotic cone $\operatorname{Ker} Q$ of E_Q. Theorem 1.2.5 below establishes the convexity of f, which is therefore a semi-norm, actually a norm if Q is positive definite. □

The mapping $E_Q \mapsto f$, introduced in Example 1.2.3, is important in the context of sublinear functions; let us put it in perspective.

Definition 1.2.4 (Gauge) Let C be a closed convex set containing the origin. The function γ_C defined by

$$\gamma_C(x) := \inf\{\lambda > 0 \, : \, x \in \lambda C\} \tag{1.2.2}$$

is called the *gauge* of C. As usual, we set $\gamma_C(x) := +\infty$ if $x \in \lambda C$ for no $\lambda > 0$. □

Geometrically, γ_C can be obtained as follows: shift C ($\subset \mathbb{R}^n$) in the hyperplane $\mathbb{R}^n \times \{1\}$ of the graph-space $\mathbb{R}^n \times \mathbb{R}$ (by contrast to a perspective-function, the

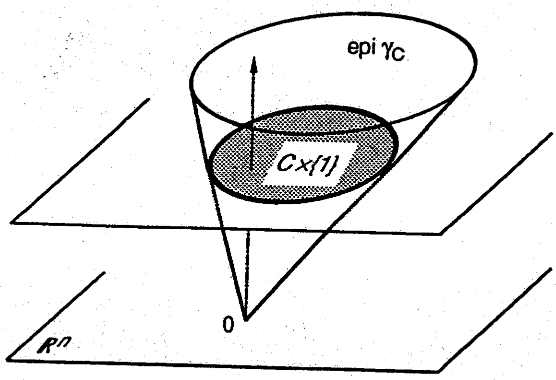

Fig. 1.2.1. The epigraph of a gauge

present shift is vertical, along the axis of function-values). Then the epigraph of γ_C is the cone generated by this shifted copy of C; see Fig. 1.2.1.

The next result summarizes the main properties of a gauge. Each statement should be read with Fig. 1.2.1 in mind, even though the picture is slightly misleading, due to closure problems.

Theorem 1.2.5 *Let C be a closed convex set containing the origin. Then*

(i) *its gauge γ_C is a nonnegative closed sublinear function;*

(ii) *γ_C is finite everywhere if and only if 0 lies in the interior of C;*

(iii) *C_∞ being the asymptotic cone of C,*

$$\{x \in \mathbb{R}^n \,:\, \gamma_C(x) \leqslant r\} = rC \quad \textit{for all } r > 0\,,$$
$$\{x \in \mathbb{R}^n \,:\, \gamma_C(x) = 0\} = C_\infty\,.$$

Proof. [*(i) and (iii)*] Nonnegativity and positive homogeneity are obvious from the definition of γ_C; also, $\gamma_C(0) = 0$ because $0 \in C$. We prove convexity via a geometric interpretation of (1.2.2). Let

$$K_C := \text{cone}\,(C \times \{1\}) = \{(\lambda c, \lambda) \in \mathbb{R}^n \times \mathbb{R} \,:\, c \in C,\, \lambda \geqslant 0\}$$

be the convex conical hull of $C \times \{1\} \subset \mathbb{R}^n \times \mathbb{R}$. It is convex (beware that K_C need not be closed) and γ_C is clearly given by

$$\gamma_C(x) = \inf\{\lambda \,:\, (x, \lambda) \in K_C\}\,.$$

Thus, γ_C is the lower-bound function of §B.1.3(g), constructed on the convex set K_C; this establishes the convexity of γ_C, hence its sublinearity.

Now we prove

$$\{x \in \mathbb{R}^n \,:\, \gamma_C(x) \leqslant 1\} = C\,. \tag{1.2.3}$$

This will imply the first part in (iii), thanks to positive homogeneity. Then the second part will follow because of (A.2.2.2): $C_\infty = \cap\{rC \,:\, r > 0\}$ and closedness of γ_C will also result from (iii) via Proposition B.1.2.2.

So, to prove (1.2.3), observe first that $x \in C$ implies from (1.2.2) that certainly $\gamma_C(x) \leqslant 1$. Conversely, let x be such that $\gamma_C(x) \leqslant 1$; we must prove that $x \in C$. For this we prove that $x_k := (1 - 1/k)x \in C$ for $k = 1, 2, \ldots$ (and then, the desired property will come from the closedness of C). By positive homogeneity, $\gamma_C(x_k) = (1 - 1/k)\gamma_C(x) < 1$, so there is $\lambda_k \in]0, 1[$ such that $x_k \in \lambda_k C$, or equivalently $x_k/\lambda_k \in C$. Because C is convex and contains the origin, $\lambda_k(x_k/\lambda_k) + (1 - \lambda_k)0 = x_k$ is in C, which is what we want.

[*(ii)*] Assume $0 \in \operatorname{int} C$. There is $\varepsilon > 0$ such that for all $x \neq 0$, $x_\varepsilon := \varepsilon x/\|x\| \in C$; hence $\gamma_C(x_\varepsilon) \leqslant 1$ because of (1.2.3). We deduce by positive homogeneity

$$\gamma_C(x) = \frac{\|x\|}{\varepsilon}\gamma_C(x_\varepsilon) \leqslant \frac{\|x\|}{\varepsilon};$$

this inequality actually holds for all $x \in \mathbb{R}^n$ ($\gamma_C(0) = 0$) and γ_C is a finite function.

Conversely, suppose γ_C is finite everywhere. By continuity (Theorem B.3.1.2), γ_C has an upper bound $L > 0$ on the unit ball:

$$\|x\| \leqslant 1 \quad \Longrightarrow \quad \gamma_C(x) \leqslant L \quad \Longrightarrow \quad x \in LC,$$

where the last implication comes from (iii). In other words, $B(0, 1/L) \subset C$. $\square$

Since γ_C is the lower-bound function of the cone K_C $(= K_C + \{0\} \times \mathbb{R}^+)$ of Fig. 1.2.1, we know from (B.1.3.6) that $K_C \subset \operatorname{epi}\gamma_C \subset \operatorname{cl} K_C$; but γ_C has a closed epigraph, therefore

$$\operatorname{epi}\gamma_C = \operatorname{cl} K_C = \overline{\operatorname{cone}}(C \times \{1\}). \tag{1.2.4}$$

Since $C_\infty = \{0\}$ if and only if C is compact (Proposition A.2.2.3), we obtain another consequence of (iii):

Corollary 1.2.6 *C is compact if and only if $\gamma_C(x) > 0$ for all $x \neq 0$.* $\square$

Example 1.2.7 The quadratic semi-norms of Example 1.2.3 can be generalized: let $f \in \overline{\operatorname{Conv}}\,\mathbb{R}^n$ have nonnegative values and be positively homogeneous of degree 2, i.e.

$$0 \leqslant f(tx) = t^2 f(x) \quad \text{for all } x \in \mathbb{R}^n \text{ and all } t > 0.$$

Then, $\sqrt{f}$ is convex; in fact

$$\begin{aligned}\sqrt{f}(x) &= \inf\{\lambda > 0 : \sqrt{f(x)} \leqslant \lambda\}\\ &= \inf\{\lambda > 0 : f(x) \leqslant \lambda^2\}\\ &= \inf\{\lambda > 0 : \tfrac{x}{\lambda} \in S_1(f)\},\end{aligned}$$

which reveals the sublevel-set

$$S_1(f) = \{x \in \mathbb{R}^n : f(x) \leqslant 1\} =: C.$$

In other words, $\sqrt{f}$ is the gauge of a closed convex set C containing the origin. $\square$

Gauges are examples of sublinear functions which are closed. This is not the case of all sublinear functions: see the function $\tilde{f}$ of (1.2.1); another example in $\mathbb{R}^2$ is

$$h(\xi, \eta) := \begin{cases} 0 & \text{if } \eta > 0\,, \\ |\xi| & \text{if } \eta = 0\,, \\ +\infty & \text{if } \eta < 0\,. \end{cases}$$

By taking the closure, or lower semi-continuous hull, of a sublinear function σ, we get a new function defined by

$$\operatorname{cl} \sigma(x) := \liminf_{x' \to x} \sigma(x') \tag{1.2.5}$$

which is (i) closed by construction, (ii) convex (Proposition B.1.2.6) and (iii) positively homogeneous, as is immediately seen from (1.2.5). For example, to close the above h, one must set $h(\xi, 0) = 0$ for all ξ. We retain from this observation that, when we close a sublinear function, we obtain a new function which is closed, of course, but which inherits sublinearity. The subclass of sublinear functions that are also closed is extremely important; in fact most of our study will be restricted to these.

Note that, for a closed sublinear function σ,

$$\sigma(0) \leqslant \lim_{t \downarrow 0} \sigma(tx) = 0 \quad \text{for all } x \in \operatorname{dom} \sigma\,,$$

so certainly $\sigma(0) = 0$; otherwise, $\operatorname{dom} \sigma$ would be empty, a situation that we reject from our definitions. Another observation is that a closed sublinear function σ coincides with its asymptotic function:

$$\sigma'_\infty = \sigma \quad \text{if } \sigma \text{ is closed and sublinear}$$

(take $x_0 = 0$ in the definition of Proposition B.3.2.1). In particular, if σ is finite everywhere, then Proposition B.3.2.6 tells us that it is Lipschitzian, and its best Lipschitz constant is

$$\sup \{\sigma(d) \; : \; \|d\| = 1\}\,. \tag{1.2.6}$$

1.3 The Convex Cone of All Closed Sublinear Functions

Similarly to convex functions, sublinear functions, closed or not, can be combined to give new sublinear functions.

Proposition 1.3.1 (i) *If σ_1 and σ_2 are [closed] sublinear and t_1, t_2 are positive numbers, then $\sigma := t_1\sigma_1 + t_2\sigma_2$ is [closed] sublinear, if not identically $+\infty$.*

(ii) *If $\{\sigma_j\}_{j \in J}$ is a family of [closed] sublinear functions, then $\sigma := \sup_{j \in J} \sigma_j$ is [closed] sublinear, if not identically $+\infty$.*

Proof. Concerning convexity and closedness, everything is known from §B.2. Note in passing that a closed sublinear function is zero (hence finite) at zero. As for positive homogeneity, it is straightforward. □

Proposition 1.3.2 *Let $\{\sigma_j\}_{j\in J}$ be a family of sublinear functions all minorized by some linear function. Then*

(i) $\sigma := \operatorname{co}(\inf_{j\in J}\sigma_j)$ *is sublinear.*

(ii) *If $J = \{1,\dots,m\}$ is a finite set, we obtain the infimal convolution*

$$\operatorname{co}\min\{\sigma_1,\dots,\sigma_m\} = \sigma_1 \mathbin{\dot{+}} \cdots \mathbin{\dot{+}} \sigma_m\,.$$

Proof. [*(i)*] Once again, the only thing to prove for (i) is positive homogeneity. Actually, it suffices to multiply x and each x_j by $t > 0$ in a formula giving $\operatorname{co}(\inf_j \sigma_j)(x)$, say (B.2.5.3).

[*(ii)*] By definition, computing $\operatorname{co}(\min_j \sigma_j)(x)$ amounts to solving the minimization problem in the m couples of variables $(x_j, \alpha_j) \in \operatorname{dom}\sigma_j \times \mathbb{R}$

$$\left|\begin{array}{ll} \inf \sum_{j=1}^m \alpha_j\sigma_j(x_j) & \alpha_j \geqslant 0 \\ \sum_{j=1}^m \alpha_j = 1\,,\ \sum_{j=1}^m \alpha_j x_j = x\,. & \end{array}\right. \tag{1.3.1}$$

In view of positive homogeneity, the variables α_j play no role by themselves: the relevant variables are actually the products $\alpha_j x_j$ and (1.3.1) can be written – denoting $\alpha_j x_j$ again by x_j:

$$\operatorname{co}\big(\min_j \sigma_j\big)(x) = \inf\Big\{\sum_{j=1}^m \sigma_j(x_j) \;:\; \sum_{j=1}^m x_j = x\Big\}\,.$$

We recognize the infimal convolution of the σ_j's. □

From Proposition 1.3.1(i), the collection of all closed sublinear functions has an *algebraic* structure: it is a convex cone contained in $\overline{\operatorname{Conv}}\,\mathbb{R}^n$. It contains another convex cone, namely the collection of finite sublinear functions.

A *topological* structure can be defined on the latter cone. In linear analysis, one defines the Euclidean distance between two linear forms $\ell_1 = \langle s_1,\cdot\rangle$ and $\ell_2 = \langle s_2,\cdot\rangle$:

$$\|\ell_1 - \ell_2\| := \|s_1 - s_2\| = \max_{\|x\|\leqslant 1} |\ell_1(x) - \ell_2(x)|\,.$$

A distance can also be defined on the convex cone of *everywhere finite* sublinear functions (the extended-valued case is somewhat more delicate, just as with unbounded sets; see some explanations in §1.5.2), which of course contains the vector space of linear forms.

Theorem 1.3.3 *For σ_1 and σ_2 in the set Φ of sublinear functions that are finite everywhere, define*

$$\Delta(\sigma_1,\sigma_2) := \max_{\|x\|\leqslant 1} |\sigma_1(x) - \sigma_2(x)|\,. \tag{1.3.2}$$

Then Δ is a distance on Φ.

Proof. Clearly $\Delta(\sigma_1,\sigma_2) < +\infty$ and $\Delta(\sigma_1,\sigma_2) = \Delta(\sigma_2,\sigma_1)$. Now positive homogeneity of σ_1 and σ_2 gives for all $x \neq 0$

$$\begin{aligned}|\sigma_1(x) - \sigma_2(x)| &= \|x\|\left|\sigma_1\left(\tfrac{x}{\|x\|}\right) - \sigma_2\left(\tfrac{x}{\|x\|}\right)\right| \\ &\leqslant \|x\| \max_{\|u\|=1} |\sigma_1(u) - \sigma_2(u)| \\ &\leqslant \|x\| \, \Delta(\sigma_1,\sigma_2)\,.\end{aligned}$$

In addition, $\sigma_1(0) = \sigma_2(0) = 0$, so

$$|\sigma_1(x) - \sigma_2(x)| \leqslant \|x\| \, \Delta(\sigma_1,\sigma_2) \quad \text{for all } x \in \mathbb{R}^n$$

and $\Delta(\sigma_1,\sigma_2) = 0$ if and only if $\sigma_1 = \sigma_2$.

As for the triangle inequality, we have for arbitrary $\sigma_1, \sigma_2, \sigma_3$ in Φ

$$|\sigma_1(x) - \sigma_3(x)| \leqslant |\sigma_1(x) - \sigma_2(x)| + |\sigma_2(x) - \sigma_3(x)| \quad \text{for all } x \in \mathbb{R}^n\,,$$

so there holds

$$\begin{aligned}\Delta(\sigma_1,\sigma_3) &\leqslant \max_{\|x\|\leqslant 1} \left[|\sigma_1(x) - \sigma_2(x)| + |\sigma_2(x) - \sigma_3(x)|\right] \\ &\leqslant \max_{\|x\|\leqslant 1} |\sigma_1(x) - \sigma_2(x)| + \max_{\|x\|\leqslant 1} |\sigma_2(x) - \sigma_3(x)|\,,\end{aligned}$$

which is the required inequality. □

The index-set in (1.3.2) can be replaced by the unit sphere $\|x\| = 1$, just as in (1.2.6); and the distance between an arbitrary $\sigma \in \Phi$ and the zero-function (which is in Φ) is just the value (1.2.6). The function $\Delta(\cdot, 0)$ acts like a *norm* on the convex cone Φ.

Example 1.3.4 Consider $\|\cdot\|_1$ and $\|\cdot\|_\infty$, the ℓ_1- and ℓ_∞-norms on $\mathbb{R}^n$. They are finite sublinear (Example 1.2.2) and there holds

$$\Delta(\|\cdot\|_1, \|\cdot\|_\infty) = \frac{n-1}{\sqrt{n}}\,.$$

To accept this formula, consider that, for symmetry reasons, the maximum in the definition (1.3.2) of Δ is achieved at $x = (1/\sqrt{n}, \dots, 1/\sqrt{n})$. □

The convergence associated with this new distance function turns out to be the natural one:

Theorem 1.3.5 *Let (σ_k) be a sequence of finite sublinear functions and let σ be a finite function. Then the following are equivalent when $k \to +\infty$:*

(i) *(σ_k) converges pointwise to σ;*

(ii) *(σ_k) converges to σ uniformly on each compact set of $\mathbb{R}^n$;*

(iii) $\Delta(\sigma_k, \sigma) \to 0$.

Proof. First, the (finite) function σ is of course sublinear whenever it is the pointwise limit of sublinear functions. The equivalence between (i) and (ii) comes from the general Theorem B.3.1.4 on the convergence of convex functions.

Now, (ii) clearly implies (iii). Conversely $\Delta(\sigma_k, \sigma) \to 0$ is the uniform convergence on the unit ball, hence on any ball of radius $L > 0$ (the maximand in (1.3.2) is positively homogeneous), hence on any compact set. □

2 The Support Function of a Nonempty Set

2.1 Definitions, Interpretations

Definition 2.1.1 (Support Function) Let S be a nonempty set in $\mathbb{R}^n$. The function $\sigma_S : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ defined by

$$\mathbb{R}^n \ni x \mapsto \sigma_S(x) := \sup \{\langle s, x\rangle \ : \ s \in S\} \tag{2.1.1}$$

is called the *support function* of S. □

For a given S, the support function therefore depends on the scalar product: changing $\langle \cdot, \cdot \rangle$ changes σ_S. In (2.1.1), the space where s runs and the space where σ_S acts are dual to each other.

The supremum in (2.1.1) may be finite or infinite, achieved on S or not. In this context, S can be interpreted as an index set: $\sigma_S(\cdot)$ is the supremum of the collection of linear forms $\langle s, \cdot\rangle$ over S. We obtain immediately:

Proposition 2.1.2 *A support function is closed and sublinear.*

Proof. This results from Proposition 1.3.1(ii) (a linear form is closed and convex!). Observe in particular that a support function is null (hence $< +\infty$) at the origin. □

The domain of σ_S is a convex cone, closed or not. Actually, $x \in \operatorname{dom} \sigma_S$ means that, for some $r := \sigma_S(x)$:

$$S \subset \{s \in \mathbb{R}^n \ : \ \langle s, x\rangle \leqslant r\} \tag{2.1.2}$$

i.e. S is contained in a closed half-space "opposite" to x.

Proposition 2.1.3 *The support function of S is finite everywhere if and only if S is bounded.*

Proof. Let S be bounded, say $S \subset B(0, L)$ for some $L > 0$. Then

$$\langle s, x\rangle \leqslant \|s\| \, \|x\| \leqslant L\|x\| \quad \text{for all } s \in S\,,$$

which implies $\sigma_S(x) \leqslant L\|x\|$ for all $x \in \mathbb{R}^n$.

Conversely, finiteness of the convex σ_S implies its continuity on the whole space (Theorem B.3.1.2), hence its local boundedness: for some L,

$$\langle s, x\rangle \leqslant \sigma_S(x) \leqslant L \quad \text{for all } (s, x) \in S \times B(0,1)\,.$$

If $s \neq 0$, we can take $x = s/\|s\|$ in the above relation, which implies $\|s\| \leqslant L$. □

Observing that

$$-\sigma_S(-x) = -\sup_{s\in S}\,[-\langle s, x\rangle] = \inf_{s\in S}\, \langle s, x\rangle\,,$$

the number $\sigma_S(x) + \sigma_S(-x)$ of (1.1.6) is particularly interesting here:

Definition 2.1.4 (Breadth of a Set) The *breadth* of a nonempty set S along $x \neq 0$ is

$$\sigma_S(x) + \sigma_S(-x) = \sup_{s \in S} \langle s, x \rangle - \inf_{s \in S} \langle s, x \rangle ,$$

a number in $[0, +\infty]$. It is 0 if and only if S lies entirely in some affine hyperplane orthogonal to x; such a hyperplane is expressed as

$$\{y \in \mathbb{R}^n \; : \; \langle y, x \rangle = \sigma_S(x)\} ,$$

which in particular contains S. The intersection of all these hyperplanes is just the affine hull of S. □

If x has norm 1 and is interpreted as a *direction*, the breadth of S measures how "thick" S is along x: it is the distance between the two hyperplanes orthogonal to x and "squeezing" S. This observation calls for a more general comment: a sublinear function $x \mapsto \sigma(x)$ being positively homogeneous, the norm of its argument x has little importance. This argument should always be thought of as an *oriented direction*, i.e. a normalized vector of $\mathbb{R}^n$. Accordingly, we will generally use from now on the notation $\sigma(d)$, more suggestive for a support function than $\sigma(x)$.

Here, we give two geometric constructions which help interpreting a support function.

Interpretation 2.1.5 (Construction in $\mathbb{R}^n$) Given $S \subset \mathbb{R}^n$ and $d \neq 0$, consider for each $r \in \mathbb{R}$ the closed half-space alluded to in (2.1.2):

$$H^-_{d,r} := \{z \in \mathbb{R}^n \; : \; \langle z, d \rangle \leqslant r\} . \tag{2.1.3}$$

If (2.1.2) holds, we can find r large enough so that $S \subset H^-_{d,r}$. The value $\sigma_S(d)$ is the smallest of those r: decreasing r as much as possible while keeping S in $H^-_{d,r}$ means "leaning" onto S the affine hyperplane $H_{d,r} := \{z \in \mathbb{R}^n : \langle z, d \rangle = r\}$. See Fig. 2.1.1 for an illustration.

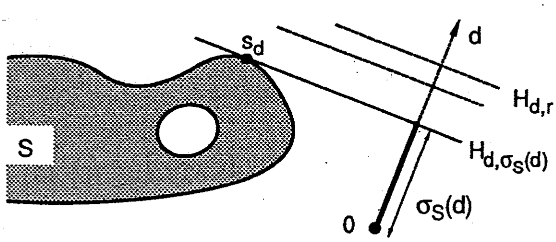

Fig. 2.1.1. Supporting hyperplanes and support functions

If (2.1.2) does not hold, however, this operation is impossible: S is "unbounded in the oriented direction" d and $\sigma_S(d) = +\infty$. Take for example $S := \mathbb{R}^+ \times \{0\}$ in $\mathbb{R}^2$. For $d = (1, 1)$ say (and assuming that $\langle \cdot, \cdot \rangle$ is the usual dot-product), no closed half-space of the form (2.1.3) can contain S, even if r is increased to $+\infty$.

If S is compact, the supremum of the (continuous) function $\langle \cdot, d\rangle$ is achieved on S, no matter how d is chosen. This means that, somewhere on the hyperplane $H_{d,\sigma_S(d)}$ there is some s_d which is also in S, actually a boundary point of S; more accurately, s_d lies in the face of S exposed by d (assuming S convex). □

Figure 2.1.1 suggests (and Proposition 2.2.1 below confirms) that the support functions of S and of $\overline{\text{co}}\, S$ coincide. Note also that the distance from the origin 0 to the "optimal" hyperplane $H_{d,\sigma_S(d)}$ is $|\sigma_S(d/\|d\|)|$. This is easily confirmed: project the origin onto $H_{d,\sigma_S(d)}$ to obtain the vector $t^* d$ such that $\langle d, t^* d\rangle = \sigma_S(d)$. Then the distance from 0 to $H_{d,\sigma_S(d)}$ is $\|t^* d\|$.

Interpretation 2.1.6 (Construction in $\mathbb{R}^{n+1}$) In the graph-space $\mathbb{R}^n \times \mathbb{R}$, we shift S down to $\mathbb{R}^n \times \{-1\}$ and consider the convex conical hull K_S of this shifted copy of S. Then the polar cone $(K_S)^\circ$ of K_S is nothing else than the epigraph of σ_S. Indeed

$$K_S = \mathbb{R}^+ \operatorname{co}(S \times \{-1\}) = \operatorname{co}[\mathbb{R}^+(S \times \{-1\})],$$

so that

$$\begin{aligned}(K_S)^\circ &= \{(d,r) : t\langle s,d\rangle - tr \leqslant 0 \text{ for all } s \in S \text{ and } t > 0\} \\ &= \{(d,r) \ :\ \langle s,d\rangle \leqslant r \text{ for all } s \in S\} \\ &= \{(d,r) \ :\ \sup_{s\in S}\langle s,d\rangle \leqslant r\} = \operatorname{epi}\sigma_S\,.\end{aligned}$$

This is illustrated on Fig. 2.1.2. On this picture, $0 \in S$; this implies $\sigma_S(d) \geqslant 0$ for all d, which is obvious just from the definition (2.1.1) of σ_S. □

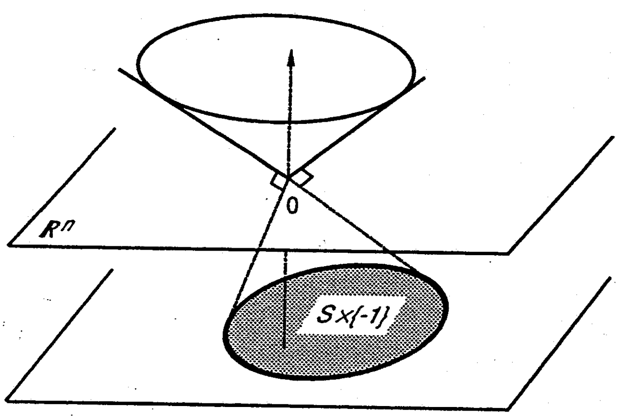

Fig. 2.1.2. The epigraph of a support function

2.2 Basic Properties

First, we list some properties of support functions that are directly derived from their definition.

Proposition 2.2.1 *For $S \subset \mathbb{R}^n$ nonempty, there holds $\sigma_S = \sigma_{\text{cl}\, S} = \sigma_{\text{co}\, S}$; whence*

$$\sigma_S = \sigma_{\overline{\text{co}}\, S} \,. \tag{2.2.1}$$

Proof. The continuity [resp. linearity, hence convexity] of the function $\langle s, \cdot \rangle$, which is maximized over S, implies that $\sigma_S = \sigma_{\text{cl}\, S}$ [resp. $\sigma_S = \sigma_{\text{co}\, S}$]. Knowing that $\overline{\text{co}}\, S = \text{cl co}\, S$ (Proposition A.1.4.2), (2.2.1) follows immediately. □

This result is of utmost importance: it says that the concept of support function *does not distinguish* a set S from its closed convex hull. Thus, when dealing with support functions, it makes no difference if we restrict ourselves to the case of closed convex sets.

As a result of (2.1.1) and (2.2.1), we can write

$$s \in \overline{\text{co}}\, S \quad \Longrightarrow \quad [\langle s, d \rangle \leqslant \sigma_S(d) \quad \text{for all } d \in \mathbb{R}^n] \,.$$

Now, what about the converse? Can it be that the above (infinite) set of inequalities still holds if s is not in $\overline{\text{co}}\, S$? The answer is no:

Theorem 2.2.2 *For the nonempty $S \subset \mathbb{R}^n$ and its support function σ_S, there holds*

$$s \in \overline{\text{co}}\, S \quad \Longleftrightarrow \quad [\langle s, d \rangle \leqslant \sigma_S(d) \quad \textit{for all } d \in X] \,, \tag{2.2.2}$$

where the set X can be indifferently taken as: the whole of $\mathbb{R}^n$, the unit ball $B(0,1)$ or its boundary the unit sphere $\widetilde{B}$, or dom σ_S.

Proof. First, the equivalence between all the choices for X is clear enough; in particular due to positive homogeneity. Because "$\Rightarrow$" is Proposition 2.2.1, we have to prove "$\Leftarrow$" only, with $X = \mathbb{R}^n$ say.

So suppose that $s \notin \overline{\text{co}}\, S$. Then $\{s\}$ and $\overline{\text{co}}\, S$ can be strictly separated (Theorem A.4.1.1): there exists $d_0 \in \mathbb{R}^n$ such that

$$\langle s, d_0 \rangle > \sup \{\langle s', d_0 \rangle \;:\; s' \in \overline{\text{co}}\, S\} = \sigma_S(d_0) \,,$$

where the last equality is (2.2.1). Our result is proved by contradiction. □

As a result, a closed convex set is completely determined by its support function: between the classes of closed convex sets and of support functions, there is a correspondence which is bijective, as illustrated on Fig. 2.2.1.

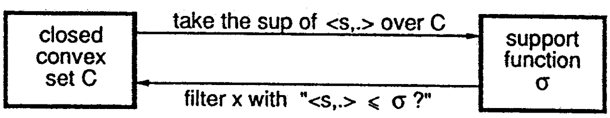

Fig. 2.2.1. Correspondence between closed convex sets and support functions

Thus, whether a given point s belongs to a given closed convex set S can be checked with the help of (2.2.2), which holds as an equivalence. Actually, more can be said: the support function *filters* the interior, the relative interior and the affine hull of a closed convex set.

This property is best understood with Fig. 1.1.2 in mind. Let V be the subspace parallel to aff S, and $U := V^\perp$. Indeed, U is just given in (1.1.8) with $\sigma = \sigma_S$: U can be viewed either as the subspace where the sublinear function σ_S is linear, or where the supported set S is flat; by contrast, σ_S is V-shaped in V, while S is thick along V. When drawn in the geometric space of convex sets, Fig. 1.1.2 becomes Fig. 2.2.2, which is very helpful to follow the next proof.

Theorem 2.2.3 *Let S be a nonempty closed convex set in $\mathbb{R}^n$. Then*

(i) $s \in \operatorname{aff} S$ *if and only if*

$$\langle s, d\rangle = \sigma_S(d) \quad \textit{for all } d \textit{ with } \sigma_S(d) + \sigma_S(-d) = 0\,; \tag{2.2.3}$$

(ii) $s \in \operatorname{ri} S$ *if and only if*

$$\langle s, d\rangle < \sigma_S(d) \quad \textit{for all } d \textit{ with } \sigma_S(d) + \sigma_S(-d) > 0\,; \tag{2.2.4}$$

(iii) *in particular,* $s \in \operatorname{int} S$ *if and only if*

$$\langle s, d\rangle < \sigma_S(d) \quad \textit{for all } d \neq 0\,. \tag{2.2.5}$$

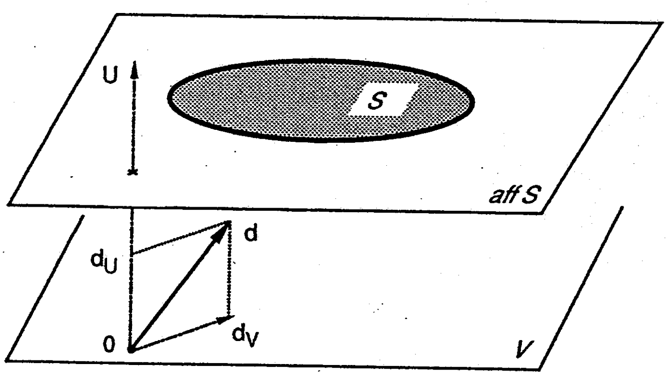

Fig. 2.2.2. Affine hulls and orthogonal spaces

Proof. [(*i*)] Let first $s \in S$. We have already seen in Definition 2.1.4 that

$$-\sigma_S(-d) \leqslant \langle s, d\rangle \leqslant \sigma_S(d) \quad \text{for all } d \in \mathbb{R}^n\,.$$

If the breadth of S along d is zero, we obtain a pair of equalities: for such d, there holds

$$\langle s, d\rangle = \sigma_S(d)\,,$$

an equality which extends by affine combination to any element $s \in \operatorname{aff} S$.

Conversely, let s satisfy (2.2.3). A first case is when the only d described in (2.2.3) is $d = 0$; as a consequence of our observations in Definition 2.1.4, there is no affine hyperplane containing S, i.e. $\operatorname{aff} S = \mathbb{R}^n$ and there is nothing to prove. Otherwise, there does exist a hyperplane H containing S; it is defined by

$$H := \{p \in \mathbb{R}^n \; : \; \langle p, d_H \rangle = \sigma_S(d_H)\}, \tag{2.2.6}$$

for some $d_H \neq 0$. We proceed to prove $\langle s, \cdot \rangle \leqslant \sigma_H$.

In fact, the breadth of S along d_H is certainly 0, hence $\langle s, d_H \rangle = \sigma_S(d_H)$ because of (2.2.3), while (2.2.6) shows that $\sigma_S(d_H) = \sigma_H(d_H)$. On the other hand, it is obvious that $\sigma_H(d) = +\infty$ if d is not collinear to d_H. In summary, we have proved $\langle s, d \rangle \leqslant \sigma_H(d)$ for all d, i.e. $s \in H$. We conclude that our s is in any affine manifold containing S: $s \in \operatorname{aff} S$.

[*(iii)*] In view of positive homogeneity, we can normalize d in (2.2.5). For $s \in \operatorname{int} S$, there exists $\varepsilon > 0$ such that $s + \varepsilon d \in S$ for all d in the unit sphere $\widetilde{B}$. Then, from the very definition (2.1.1),

$$\sigma_S(d) \geqslant \langle s + \varepsilon d, d \rangle = \langle s, d \rangle + \varepsilon \quad \text{for all } d \in \widetilde{B}\,.$$

Conversely, let $s \in \mathbb{R}^n$ be such that

$$\sigma_S(d) - \langle s, d \rangle > 0 \quad \text{for all } d \in \widetilde{B}$$

which implies, because σ_S is closed and the unit sphere is compact:

$$0 < \varepsilon := \inf \{\sigma_S(d) - \langle s, d \rangle \; : \; d \in \widetilde{B}\} \leqslant +\infty\,.$$

Thus

$$\langle s, d \rangle + \varepsilon \leqslant \sigma_S(d) \quad \text{for all } d \in \widetilde{B}\,.$$

Now take u with $\|u\| < \varepsilon$. From the Cauchy-Schwarz inequality, we have for all $d \in \widetilde{B}$

$$\langle s + u, d \rangle = \langle s, d \rangle + \langle u, d \rangle \leqslant \langle s, d \rangle + \varepsilon \leqslant \sigma_S(d)$$

and this implies $s + u \in S$ because of Theorem 2.2.2: $s \in \operatorname{int} S$ and (iii) is proved.

[*(ii)*] Look at Fig. 2.2.2 again: decompose $\mathbb{R}^n = V \oplus U$, where V is the subspace parallel to $\operatorname{aff} S$ and $U = V^\perp$. In the decomposition $d = d_V + d_U$, $\langle \cdot, d_U \rangle$ is constant over S, so S has 0-breadth along d_U and

$$\sigma_S(d) = \sup_{s \in S} \langle s, d_V + d_U \rangle = \langle s, d_U \rangle + \sup_{s \in S} \langle s, d_V \rangle$$

for any $s \in S$. With these notations, a direction described as in (2.2.4) is a d such that

$$\sigma_S(d) + \sigma_S(-d) = \sigma_S(d_V) + \sigma_S(-d_V) > 0\,.$$

Then, (ii) is just (iii) written in the subspace V. □

We already know that the domain of σ_S is a convex cone, which consists of all oriented directions "along which S is bounded" (remember Interpretation 2.1.5). This can be made more explicit.

Proposition 2.2.4 *Let S be a nonempty closed convex set in $\mathbb{R}^n$. Then* $\operatorname{cl}\operatorname{dom}\sigma_S$ *and the asymptotic cone S_∞ of S are mutually polar cones.*

Proof. Recall from §A.3.2 that, if K_1 and K_2 are two closed convex cones, then $K_1 \subset K_2$ if and only if $(K_1)^\circ \supset (K_2)^\circ$.

Let $p \in S_\infty$. Fix s_0 arbitrary in S and use the fact that $S_\infty = \cap_{t>0} t(S - s_0)$ (§A.2.2): for all $t > 0$, we can find $s_t \in S$ such that $p = t(s_t - s_0)$. Now, for $q \in \operatorname{dom}\sigma_S$, there holds

$$\langle p, q\rangle = t\langle s_t - s_0, q\rangle \leqslant t\,[\sigma_S(q) - \langle s_0, q\rangle] < +\infty$$

and letting $t \downarrow 0$ shows that $\langle p, q\rangle \leqslant 0$. In other words, $\operatorname{dom}\sigma_S \subset (S_\infty)^\circ$; then $\operatorname{cl}\operatorname{dom}\sigma_S \subset (S_\infty)^\circ$ since the latter is closed.

Conversely, let $q \in (\operatorname{dom}\sigma_S)^\circ$, which is a cone, hence $tq \in (\operatorname{dom}\sigma_S)^\circ$ for any $t > 0$. Thus, given $s_0 \in S$, we have for arbitrary $p \in \operatorname{dom}\sigma_S$

$$\langle s_0 + tq, p\rangle = \langle s_0, p\rangle + t\langle q, p\rangle \leqslant \langle s_0, p\rangle \leqslant \sigma_S(p)\,,$$

so $s_0 + tq \in S$ by virtue of Theorem 2.2.2. In other words: $q \in (S - s_0)/t$ for all $t > 0$ and $q \in S_\infty$. □

2.3 Examples

Let us start with elementary situations. The simplest example of a support function is that of a singleton $\{s\}$. Then $\sigma_{\{s\}}$ is merely $\langle s, \cdot\rangle$, we have a first illustration of the introduction (iii) to this chapter: the concept of a linear form $\langle s, \cdot\rangle$ can be generalized to s not being a singleton, which amounts to generalizing linearity to closed sublinearity (more details will be given in §3). The case when S is the unit ball $B(0, 1)$ is also rather simple:

$$\sigma_{B(0,1)}(d) \geqslant \left\langle \frac{d}{\|d\|}, d\right\rangle = \|d\| \quad (\text{if } d \neq 0)$$

and, for $s \in B(0, 1)$, the Cauchy-Schwarz inequality implies $\langle s, d\rangle \leqslant \|d\|$. Altogether,

$$\sigma_{B(0,1)}(d) = \|d\|\,. \tag{2.3.1}$$

Our next example is the simplest possible illustration of Proposition 2.2.4, namely when S_∞ is S itself:

Example 2.3.1 (Cones, Half-Spaces, Subspaces) Let K be a closed convex cone of $\mathbb{R}^n$. Then

$$\sigma_K(d) = \begin{cases} 0 & \text{if } \langle s, d\rangle \leqslant 0 \text{ for all } s \in K, \\ +\infty & \text{otherwise}\,. \end{cases}$$

In other words, σ_K is the indicator function of the polar cone K°. Note the symmetry: since $K^{\circ\circ} = K$, the support function of K° is the indicator of K. In summary:

$$\mathrm{i}_K = \sigma_{K^\circ} \quad \text{and} \quad \sigma_K = \mathrm{i}_{K^\circ}\,.$$

Two particular cases are of interest. One is when K is a half-space:

$$K := \{s \in \mathbb{R}^n \;:\; \langle s, v\rangle \leqslant 0\}\,;$$

then it is clear enough that

$$\sigma_K(d) = \begin{cases} 0 & \text{if } d = tv \text{ with } t \geqslant 0\,, \\ +\infty & \text{otherwise}\,. \end{cases} \tag{2.3.2}$$

Needless to say, the support function of the half-line $\mathbb{R}^+ v$ (the polar of K) is in turn the indicator of K.

The other interesting case is that of a subspace. Let $A : \mathbb{R}^n \to \mathbb{R}^m$ be a linear operator and H be defined by

$$H := \operatorname{Ker} A = \{s \in \mathbb{R}^n \;:\; As = 0\}\,.$$

Then the support function of H is the indicator of the orthogonal subspace $H^\perp$:

$$\sigma_H(d) = \mathrm{i}_{H^\perp}(d) = \begin{cases} 0 & \text{if } \langle s, d\rangle = 0 \text{ for all } s \in H\,, \\ +\infty & \text{otherwise}\,. \end{cases}$$

The subspace $H^\perp$ can be defined with the help of the adjoint of A:

$$H^\perp = (\operatorname{Ker} A)^\perp = \operatorname{Im} A^* = \{A^*\lambda \;:\; \lambda \in \mathbb{R}^m\}\,.$$

If A or H are defined in terms of linear constraints

$$H := \{s \in \mathbb{R}^n \;:\; \langle s, a_j\rangle = 0 \text{ for } j = 1, \dots, m\}\,,$$

then $H^\perp = \{\sum_{j=1}^m \lambda_j a_j : \lambda \in \mathbb{R}^m\}$.

All these calculations are useful when dealing with closed convex polyhedra, expressed as intersections of half-spaces and subspaces.

Figure 2.3.1 illustrates a modification in which our cone K is modified to $K' := K \cap B(0,1)$. The calculus rules of §3.3 will prove what is suggested by the picture: the support function of K' is the distance function to K° (check the similarity of the appropriate triangles, and note that $\sigma_{K'}(d) = 0$ when $d \in K^\circ$). □

Example 2.3.2 The asymptotic cone of the set

$$S := \{s = (\rho, \tau) \in \mathbb{R}^2 \;:\; \rho > 0,\, \tau \geqslant 1/\rho\} \tag{2.3.3}$$

is $S_\infty = \{(\rho,\tau) \in \mathbb{R}^2 : \rho \geqslant 0,\ \tau \geqslant 0\}$ and, from Proposition 2.2.4, the closure of $\operatorname{dom} \sigma_S$ is $\{(\xi,\eta) : \xi \leqslant 0,\ \eta \leqslant 0\}$.

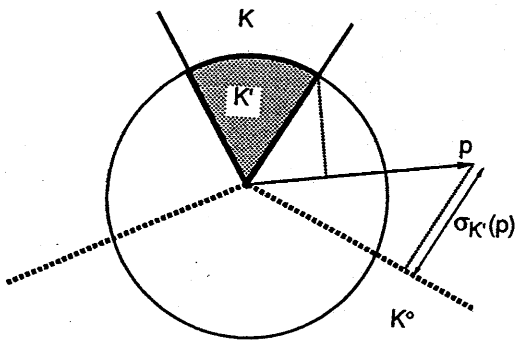

Fig. 2.3.1. Support function of a truncated cone

The exact status of the boundary of $\operatorname{dom}\sigma_S$ (i.e. when $\xi\eta = 0$) is not specified by Proposition 2.2.4: is σ_S finite there? The computation of σ_S can be done directly from the definitions (2.1.1) and (2.3.3). However the following geometric argument yields simpler calculations (see Fig. 2.3.2): for given $d = (\xi, \eta) \neq (0,0)$, consider the hyperplane $H_{d,\sigma_S(d)} = \{(\alpha, \beta) : \xi\alpha + \eta\beta = \sigma_S(d)\}$. It has to be tangent to the boundary of S, defined by the equation $\alpha\beta = 1$. So, the discriminant $\sigma_S^2(d) - 4\xi\eta$ of the equation in α

$$\xi\alpha + \eta\frac{1}{\alpha} = \sigma_S(d)$$

must be 0. We obtain directly $\sigma_S(\xi, \eta) = -2\sqrt{\xi\eta}$ for $\xi < 0, \eta < 0$ (the sign is "$-$" because $0 \notin S$; remember Theorem 2.2.2). Finally, Proposition 2.1.2 tells us that the closed function $(\xi, \eta) \mapsto \sigma_S(\xi, \eta)$ has to be 0 when $\xi\eta = 0$. All this is confirmed by Fig. 2.3.2. □

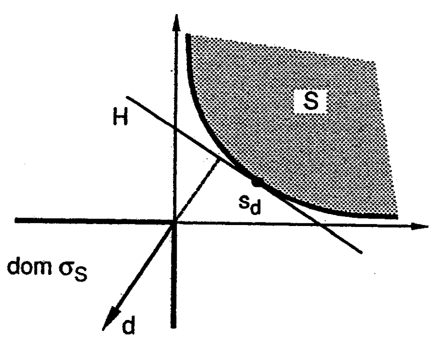

Fig. 2.3.2. A support function

Remark 2.3.3 Two features concerning the boundary of $\operatorname{dom}\sigma_S$ are worth mentioning on the above example: the supremum in (2.1.1) is not attained when $d \in \operatorname{bd}\operatorname{dom}\sigma_S$ (the point s_d of Fig. 2.1.1 is sent to infinity when d approaches $\operatorname{bd}\operatorname{dom}\sigma_S$), and $\operatorname{dom}\sigma_S$ is closed.

These are not the only possible cases: Example 2.3.1 shows that the supremum in (2.1.1) can well be attained for all $d \in \operatorname{dom} \sigma_S$; and in the example

$$S := \{(\rho, \tau) \,:\, \tau \geqslant \tfrac{1}{2}\rho^2\},$$

$\operatorname{dom} \sigma_S$ is not closed. The difference is that, now, S has no asymptote "at finite distance". □

Example 2.3.4 Just as in Example 1.2.3, let Q be a symmetric positive definite operator from $\mathbb{R}^n$ to $\mathbb{R}^n$. Its sublevel-set

$$E_Q := \{s \in \mathbb{R}^n \,:\, \langle Qs, s\rangle \leqslant 1\}$$

is elliptic, with support function

$$d \mapsto \sigma_{E_Q}(d) := \max\{\langle s, d\rangle \,:\, \langle Qs, s\rangle \leqslant 1\}. \tag{2.3.4}$$

Calling $Q^{1/2}$ the square root of Q, the change of variable $p = Q^{1/2}s$ in (2.3.4) gives

$$\sigma_{E_Q}(d) = \max\{\langle p, Q^{-1/2}d\rangle \,:\, \|p\|^2 \leqslant 1\}$$

whose unique solution for $d \neq 0$ (again Cauchy-Schwarz!) is $p = \frac{Q^{-1/2}d}{\|Q^{-1/2}d\|}$ and finally

$$\sigma_{E_Q}(d) = \|Q^{-1/2}d\| = \sqrt{\langle d, Q^{-1}d\rangle}. \tag{2.3.5}$$

Observe in this example the "duality" between the gauge $x \mapsto \sqrt{\langle Qx, x\rangle}$ of E_Q and its support function (2.3.5).

When Q is merely symmetric positive semi-definite, E_Q becomes an elliptic cylinder, whose asymptotic cone is $\operatorname{Ker} Q$ (remember Example 1.2.3). Then Proposition 2.2.4 tells us that

$$\operatorname{cl} \operatorname{dom} \sigma_{E_Q} = (\operatorname{Ker} Q)^\circ = (\operatorname{Ker} Q)^\perp = \operatorname{Im} Q.$$

When $d \in \operatorname{Im} Q$, $\sigma_{E_Q}(d)$ is finite indeed and (2.3.5) does hold, $Q^{-1}d$ denoting now any element p such that $Qp = d$. We leave this as an exercise. □

3 The Isomorphism Between Closed Convex Sets and Closed Sublinear Functions

3.1 The Fundamental Correspondence

We have seen in Proposition 2.1.2 that a support function is closed and sublinear. What about the converse? Are there closed sublinear functions which support no set in $\mathbb{R}^n$? The answer is no: any closed sublinear function *can be viewed* as a support function. The key lies in the representation of a closed convex function f via affine functions minorizing it: when the starting f is also positively homogeneous, the underlying affine functions can be assumed linear.

Theorem 3.1.1 *Let σ be a closed sublinear function; then there is a linear function minorizing σ. In fact, σ is the supremum of the linear functions minorizing it. In other words, σ is the support function of the nonempty closed convex set*

$$S_\sigma := \{s \in \mathbb{R}^n \, : \, \langle s, d \rangle \leqslant \sigma(d) \text{ for all } d \in \mathbb{R}^n\} . \tag{3.1.1}$$

Proof. Being convex, σ is minorized by some affine function (Proposition B.1.2.1): for some $(s, r) \in \mathbb{R}^n \times \mathbb{R}$,

$$\langle s, d \rangle - r \leqslant \sigma(d) \quad \text{for all } d \in \mathbb{R}^n . \tag{3.1.2}$$

Because $\sigma(0) = 0$, the above r is nonnegative. Also, by positive homogeneity,

$$\langle s, d \rangle - \tfrac{1}{t} r \leqslant \sigma(d) \quad \text{for all } d \in \mathbb{R}^n \text{ and all } t > 0 .$$

Letting $t \to +\infty$, we see that σ is actually minorized by a linear function:

$$\langle s, d \rangle \leqslant \sigma(d) \quad \text{for all } d \in \mathbb{R}^n . \tag{3.1.3}$$

Now observe that the minorization (3.1.3) is sharper than (3.1.2): when expressing the closed convex σ as the supremum of all the affine functions minorizing it (Proposition B.1.2.8), we can restrict ourselves to linear functions. In other words

$$\sigma(d) = \sup \{\langle s, d \rangle \, : \, \text{the linear } \langle s, \cdot \rangle \text{ minorizes } \sigma\} ;$$

in the above index-set, we just recognize S_σ. □

One of the important points in this result is the *nonemptiness* of S_σ in (3.1.1); we have here the analytical form of Hahn-Banach theorem: there *exists* a linear function minorizing the closed sublinear function σ; compare this with the geometric form given in Theorem A.4.1.1.

Another way of expressing Theorem 3.1.1 is that the closed convex set $\operatorname{epi} \sigma$ is the intersection of the closed half-spaces containing it; but since $\operatorname{epi} \sigma$ is actually a cone, these half-spaces can be assumed to have *linear* hyperplanes as boundaries (remember Corollary A.4.2.7). A connection between S_σ and the cone polar to $\operatorname{epi} \sigma$ is thus introduced; Chap. D will exploit this remark.

The main consequence of this important theorem is an assessment of closed sublinear functions. Section 2.2 has established a bijection from closed convex sets onto support functions. Thanks to Theorem 3.1.1, this bijection is actually onto *closed sublinear functions*, which is of course much more satisfactory: the latter class of functions is defined in abstracto, while the former class was ad hoc, as far as this bijection was concerned.

Thus, the wording "support function" in Fig. 2.2.1 can everywhere be replaced by "closed sublinear". This replacement can be done in Theorem 2.2.2 as well:

Corollary 3.1.2 *For a nonempty closed convex set S and a closed sublinear function σ, the following are equivalent:*

(i) *σ is the support function of S,*

(ii) *$S = \{s : \langle s, d\rangle \leqslant \sigma(d)$ for all $d \in X\}$, where the set X can be indifferently taken as: the whole of $\mathbb{R}^n$, the unit ball $B(0, 1)$ or its boundary, or* $\operatorname{dom}\sigma$.

Proof. The case $X = \mathbb{R}^n$ is just Theorem 3.1.1. The other cases are then clear. □

Remember the outer construction of §A.4.2(b): a closed convex set S is geometrically characterized as an intersection of half-spaces, which in turn can be characterized in terms of the support function of S. Each $(d, r) \in \mathbb{R}^n \times \mathbb{R}$ defines (for $d \neq 0$) the half-space $H^-_{d,r}$ via (2.1.3). This half-space contains S if and only if $r \geqslant \sigma(d)$, and Corollary 3.1.2 expresses that

$$S = \cap\{s : \langle s, d\rangle \leqslant r \text{ for all } d \in \mathbb{R}^n \text{ and } r \geqslant \sigma(d)\},$$

in which the couple (d, r) plays the role of an index, running in the index-set $\operatorname{epi}\sigma \subset \mathbb{R}^n \times \mathbb{R}$ (compare with the discussion after Definition 2.1.1). Of course, this index-set can be reduced down to $\mathbb{R}^n$: the above formula can be simplified to

$$S = \cap\{s : \langle s, d\rangle \leqslant \sigma(d) \text{ for all } d \in X\}$$

where X can be taken as in Corollary 3.1.2.

Recall from §A.2.4 that an exposed face of a convex set S is defined as the set of points of S which maximize some (nonzero) linear form. This concept appears as particularly welcome in the context of support functions:

Definition 3.1.3 (Direction Exposing a Face) Let C be a nonempty closed convex set, with support function σ. For given $d \neq 0$, the set

$$\mathrm{F}_C(d) := \{x \in C : \langle s, d\rangle = \sigma(d)\}$$

is called the exposed face of C associated with d, or the *face exposed* by d. □

For a unified notation, the entire C can be considered as the face exposed by 0. On the other hand, a given d may expose no face at all (when C is unbounded).

Symmetrically to Definition 3.1.3, one can ask what are those $d \in \mathbb{R}^n$ such that $\langle \cdot, d\rangle$ is maximized at a given $x \in C$. We obtain nothing other than the normal cone $\mathrm{N}_C(x)$ to C at x, as is obvious from its Definition A.5.2.3. The following result is simply a restatement of Proposition A.5.3.3.

Proposition 3.1.4 *For x in a nonempty closed convex set C, it holds*

$$x \in \mathrm{F}_C(d) \quad \Longleftrightarrow \quad d \in \mathrm{N}_C(x). \qquad \square$$

When d describes the set of normalized directions, the corresponding exposed faces exactly describe the boundary of C:

Proposition 3.1.5 *For a nonempty closed convex set C, it holds*

$$\operatorname{bd} C = \cup\{\mathrm{F}_C(d) : d \in X\}$$

where X can be indifferently taken as: $\mathbb{R}^n \backslash \{0\}$, the unit sphere $\widetilde{B}$, or $\operatorname{dom}\sigma_C \backslash \{0\}$.

Proof. Observe from Definition 3.1.3 that the face exposed by $d \neq 0$ does not depend on $\|d\|$. This establishes the equivalence between the first two choices for X. As for the third choice, it is due to the fact that $F_C(d) = \emptyset$ if $d \notin \operatorname{dom} \sigma_C$.

Now, if x is interior to C and $d \neq 0$, then $x + \varepsilon d \in C$ and x cannot be a maximizer of $\langle \cdot, d \rangle$: x is not in the face exposed by d. Conversely, take x on the boundary of C. Then $N_C(x)$ contains a nonzero vector d; by Proposition 3.1.4, $x \in F_C(d)$. □

3.2 Example: Norms and Their Duals, Polarity

Let $|\!|\!| \cdot |\!|\!|$ be an arbitrary norm on $\mathbb{R}^n$. It is a positive (except at 0) closed sublinear function and its sublevel-set

$$B := \{x \in \mathbb{R}^n : |\!|\!| x |\!|\!| \leqslant 1\} \tag{3.2.1}$$

is particularly interesting. It is the unit ball associated with the norm, a symmetric, convex, compact set containing the origin as an interior point; $|\!|\!| \cdot |\!|\!|$ is the gauge of B (§1.2). On the other hand, why not take the set whose support function is $|\!|\!| \cdot |\!|\!|$? In view of Corollary 3.1.2, it is defined by

$$\{s \in \mathbb{R}^n : \langle s, x \rangle \leqslant |\!|\!| x |\!|\!| \text{ for all } x \in \mathbb{R}^n\} =: B^* . \tag{3.2.2}$$

It is an easy exercise to check that B^* is also symmetric, convex, compact; and it contains the origin as an interior point (Theorem 2.2.3(iii)).

Now, we have two closed convex sets B and B^*. We can generate two more closed sublinear functions: take the support function σ_B of B and the gauge γ_{B^*} of B^*. It turns out that we then obtain the same function, which actually is a norm, denoted by $|\!|\!| \cdot |\!|\!|^*$: the so-called *dual norm* of $|\!|\!| \cdot |\!|\!|$. The game finishes there: the two sets that $|\!|\!| \cdot |\!|\!|^*$ supports and is the gauge of, respectively, are B and B^*.

Proposition 3.2.1 *Let B and B^* be defined by* (3.2.1) *and* (3.2.2), *where $|\!|\!| \cdot |\!|\!|$ is a norm on $\mathbb{R}^n$. The support function of B and the gauge of B^* are the same function $|\!|\!| \cdot |\!|\!|^*$ defined by*

$$|\!|\!| s |\!|\!|^* := \max\{\langle s, x \rangle : |\!|\!| x |\!|\!| \leqslant 1\} . \tag{3.2.3}$$

Furthermore, $|\!|\!| \cdot |\!|\!|^*$ is a norm on $\mathbb{R}^n$. The support function of its unit ball B^* and the gauge of its supported set B are the same function $|\!|\!| \cdot |\!|\!|$: there holds

$$|\!|\!| x |\!|\!| = \max\{\langle s, x \rangle : |\!|\!| s |\!|\!|^* \leqslant 1\} . \tag{3.2.4}$$

Proof. It is a particular case of the results 3.2.4 and 3.2.5 below. □

Note the following symmetric relation ("Cauchy-Schwarz")

$$\langle s, x \rangle \leqslant |\!|\!| s |\!|\!|^* \, |\!|\!| x |\!|\!| \quad \text{for all } (s, x) \in \mathbb{R}^n \times \mathbb{R}^n , \tag{3.2.5}$$

which comes directly from (3.2.3), using positive homogeneity. It expresses the duality correspondence between the two Banach spaces $(\mathbb{R}^n, |\!|\!| \cdot |\!|\!|)$ and $(\mathbb{R}^n, |\!|\!| \cdot |\!|\!|^*)$.

Furthermore, equality holds in (3.2.5) when $s \neq 0$ and $x \neq 0$ form an associated pair via Proposition 3.1.4:

$$\frac{s}{\|s\|_*} \in F_{B^*}(x) \quad \text{or equivalently} \quad \frac{x}{\|x\|} \in F_B(s)\,.$$

Thus, a norm automatically defines another norm (its dual); and the operation is symmetric: the dual of the dual norm is the norm itself.

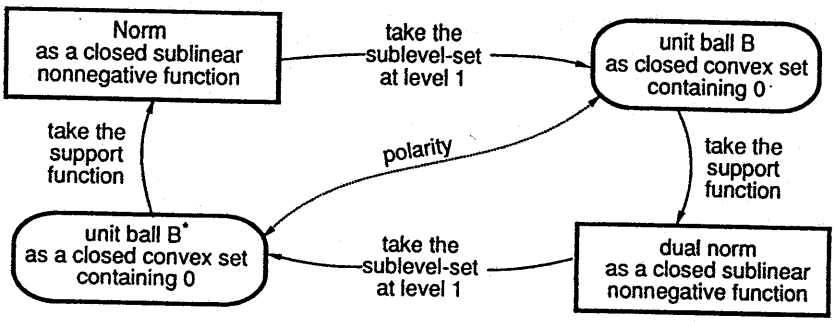

Fig. 3.2.1. Dual norms and polar sets

Remark 3.2.2 The operation (3.2.3) – (3.2.4) establishes a "duality" correspondence within a subclass of closed sublinear functions: those that are symmetric, finite everywhere, and positive (except at 0) – in short, norms.

This analytic operation has its counterpart in the geometric world: starting from a closed convex set which is symmetric, bounded and contains the origin as an interior point – in short, a "unit ball" – such as B, one constructs via gauges and support functions another closed convex set B^* which has the same properties. This correspondence is called *polarity*, demonstrated by Fig. 3.2.1: the polar (set) of B is

$$B^* := \{s \, : \, \langle s, x\rangle \leqslant 1 \text{ for all } x \in B\}\,. \tag{3.2.6}$$

As can be seen with a separation argument, the polar of B^* is symmetrically (proceed as for Theorem A.4.2.6 and remember that $0 \in B$)

$$(B^*)^* := \{x \, : \, \langle s, x\rangle \leqslant 1 \text{ for all } s \in B^*\} = B\,. \tag{3.2.7}$$

□

We leave it as an exercise to draw the unit balls of the ℓ_1- and ℓ_∞-norms on $\mathbb{R}^n$:

$$\|x\|_1 := \sum_{i=1}^{n} |x^i| \quad \text{and} \quad \|x\|_\infty := \max\{|x^1|, \dots, |x^n|\}$$

(proceed as in Interpretation 2.1.5: a picture in $\mathbb{R}^n$ will do). Observe on the picture thus obtained that they are in polarity correspondence if the scalar product is the usual dot-product $\langle x, y\rangle = x^\top y$.

Another situation is illustrated by the "hexagonal norm" of Fig. 3.2.2. Observe how elongation in one direction corresponds to contraction for the polar. Also: a facet of one of the sets is exposed by a vertex in the polar.

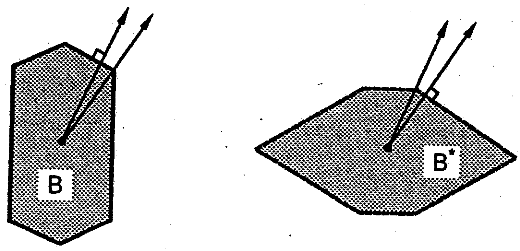

Fig. 3.2.2. Hexagonal unit-balls

Example 3.2.3 Other important norms are the quadratic norms, defined by

$$\|x\|_Q := \sqrt{\langle Qx, x\rangle}$$

where Q is a symmetric positive definite linear operator. They are important because they derive from a scalar product on $\mathbb{R}^n$, namely:

$$\langle x, y\rangle_Q := \langle Qx, y\rangle .$$

We refer to Example 2.3.4, more precisely formula (2.3.5), to compute the corresponding dual norm

$$(\|s\|_Q)^* = \sqrt{\langle s, Q^{-1}s\rangle} = \|s\|_{Q^{-1}} .$$

When $Q = I_n$, we get back the Euclidean norm $\langle\cdot,\cdot\rangle^{1/2}$. A comparison of (2.3.1) and (3.2.3) shows that it is self-dual: $\|\cdot\|^* = \|\cdot\|$. Among all the possible norms on $\mathbb{R}^n$, it is the only one having this property (once the scalar product is chosen!). □

Actually, polarity neither relies upon symmetry, nor boundedness, nor on having 0 as an interior point. To take gauges and support functions resulting in (3.2.6), (3.2.7), the only important property is after all that 0 be in the closed convex set under consideration (B or B^*). In other words, the polarity relations (3.2.6), (3.2.7) establish an involution between sets that are merely closed convex, and contain the origin. More precisely, we have the following result:

Proposition 3.2.4 *Let C be a closed convex set containing the origin. Its gauge γ_C is the support function of a closed convex set containing the origin, namely*

$$C^\circ := \{s \in \mathbb{R}^n \;:\; \langle s, d\rangle \leqslant 1 \text{ for all } d \in C\}, \tag{3.2.8}$$

which defines the polar *(set) of C.*

Proof. We know that γ_C (which, by Theorem 1.2.5(i), is closed, sublinear and nonnegative) is the support function of some closed convex set containing the origin, say D; from (3.1.1),

$$D = \{s \in \mathbb{R}^n \ : \ \langle s, d \rangle \leqslant r \text{ for all } (d, r) \in \operatorname{epi} \gamma_C\}\,.$$

As seen in (1.2.4), epi γ_C is the closed convex conical hull of $C \times \{1\}$; we can use positive homogeneity to write

$$D = \{s \in \mathbb{R}^n \ : \ \langle s, d \rangle \leqslant 1 \text{ for all } d \text{ such that } \gamma_C(d) \leqslant 1\}\,.$$

In view of Theorem 1.2.5(iii), the above index-set is just C; in other words, $D = C^\circ$. □

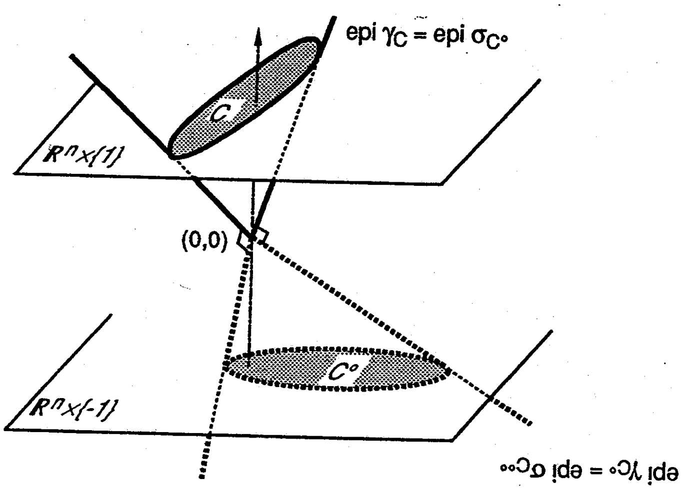

Fig. 3.2.3. Gauges and supports

Geometrically, the above proof is illustrated by Fig. 3.2.3, in which dual elements are drawn in dashed lines: $D = C^\circ$ is obtained by cutting the polar cone $(\operatorname{epi} \gamma_C)^\circ$ at the level -1. Turn the picture upside down: cutting the polar cone $(\operatorname{epi} \gamma_{C^\circ})^\circ$ at the level which has now become -1, we obtain $(C^\circ)^\circ$. But the polarity between closed convex cones is involutive: the picture shows that $(\operatorname{epi} \gamma_{C^\circ})^\circ$ is our original cone epi γ_C. In other words, $C^{\circ\circ} = C$, Proposition 3.2.4 has its dual version:

Corollary 3.2.5 *Let C be a closed convex set containing the origin. Its support function σ_C is the gauge of C°.* □

Remark 3.2.6 The elementary operation making up polarity is a one-to-one mapping between nonzero vectors and affine hyperplanes not containing the origin, via the equation inspired from (3.2.8):

$$s \mapsto H(s) := H_{s,1} = \{y \in \mathbb{R}^n : \langle s, y\rangle = 1\}. \tag{3.2.9}$$

Direct calculations show for example that the polar of the half-space

$$H^- := \{y = (\xi, \eta) \in \mathbb{R}^2 : \xi \leqslant 2\}$$

is the segment

$$(H^-)^\circ = \{(\rho, 0) : 0 \leqslant \rho \leqslant 1/2\}.$$

This simple example suggests the following comment: if σ is a given nonnegative closed sublinear function, it is the gauge of a set G which can be immediately constructed: along $0 \neq s \in \mathbb{R}^n$, plot the point $g(s) = s/\sigma(s) \in [0, +\infty]s$. Then G is the union of the segments $[0, g(s)]$, with s describing the unit sphere. If, along the same s, we plot the point $\sigma(s)s$, we likewise get a description of the set S supported by σ, but in a much less direct way: G is now *enveloped* by the affine hyperplane orthogonal to s and containing the point $\sigma(s)s$; now, *differentiation* is involved.

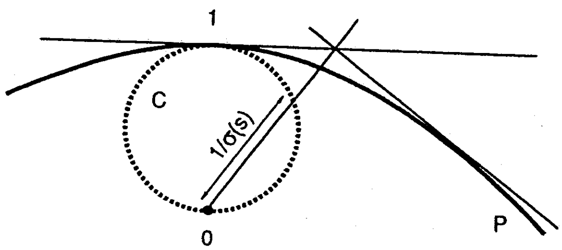

Fig. 3.2.4. Description of mutually polar sets

An expert in geometry will for example see on Fig. 3.2.4 that the polar of the circle

$$C = \{(\rho, \tau) : \rho^2 + (\tau - 1/2)^2 \leqslant 1/4\}$$

has a parabolic boundary. We leave it as an exercise to compute the gauge of C, and to realize that it is the support function of

$$P = \{(\xi, \eta) : \xi^2 \leqslant 1 - \eta\}.$$

Constructing a set from its gauge thus appears to be substantially easier than it is from its support function. Furthermore, to make a support function, we need a scalar product, while a gauge just needs an origin in $\mathbb{R}^n$. These advantages, however, are balanced by the rich calculus which can be developed with support functions, and which will be the subject of §3.3. □

It is clear from (3.2.8) that $\langle s, d\rangle \leqslant 1$ for all $(d, s) \in C \times C^\circ$; this implies in particular that no nonzero $s \in C^\circ$ can be in the asymptotic cone of C. Furthermore, the property $\langle s, d\rangle = 1$ means that d exposes in C° a face $F_{C^\circ}(d)$ containing s; and s exposes likewise in $C^{\circ\circ} = C$ a face $F_C(s)$ containing d. Because the boundary of a closed convex set is described by its exposed faces (Proposition 3.1.5), the following result is then natural; compare it with Fig. 3.2.2.

Proposition 3.2.7 *Let C be a nonempty compact convex set having 0 in its interior, so that C° enjoys the same properties. Then, for all d and s in $\mathbb{R}^n$, the following statements are equivalent (the notation* (3.2.9) *is used)*

(i) *$H(s)$ is a supporting hyperplane to C at d;*
(ii) *$H(d)$ is a supporting hyperplane to C° at s;*
(iii) *$d \in \operatorname{bd} C$, $s \in \operatorname{bd} C^\circ$ and $\langle s, d\rangle = 1$;*
(iv) *$d \in C$, $s \in C^\circ$ and $\langle s, d\rangle = 1$.*

Proof. Left as an exercise; the assumptions are present to make sure that every nonzero vector in $\mathbb{R}^n$ does expose a face in each set. □

Finally, suppose that C in (3.2.8) is a (closed convex) cone. By positive homogeneity, the number "1" can be replaced by any positive number, and even by "0" (remember the proof of Theorem 3.1.1). We recognize the definition of polarity between closed convex cones. Remembering Example 2.3.1, we see that, for a closed convex cone K, $\sigma_{K^\circ} = \gamma_K$, hence $\gamma_K = \mathrm{i}_K$, which could be checked directly from Definition 1.2.4.

3.3 Calculus with Support Functions

From §1.3, the set of sublinear functions has a structure allowing calculus. Likewise, a calculus exists with subsets of $\mathbb{R}^n$. Then a natural question is: to what extent are these structures in correspondence via the supporting operation? In other words, to what extent is the supporting operation an isomorphism? The answer turns out to be very rich indeed.

We start with the order relation

Theorem 3.3.1 *Let S_1 and S_2 be nonempty closed convex sets; call σ_1 and σ_2 their support functions. Then*

$$S_1 \subset S_2 \quad \Longleftrightarrow \quad \sigma_1(d) \leqslant \sigma_2(d) \text{ for all } d \in \mathbb{R}^n .$$

Proof. Apply the equivalence stated in Corollary 3.1.2:

$$\begin{aligned} S_1 \subset S_2 &\Longleftrightarrow s \in S_2 \text{ for all } s \in S_1 \\ &\Longleftrightarrow \sigma_2(d) \geqslant \langle s, d\rangle \text{ for all } s \in S_1 \text{ and all } d \in \mathbb{R}^n \\ &\Longleftrightarrow \sigma_2(d) \geqslant \sup_{s \in S_1} \langle s, d\rangle \text{ for all } d \in \mathbb{R}^n . \end{aligned}$$

□

In a way, the above result generalizes Theorem 2.2.2. The next statement goes with Propositions 1.3.1 and 1.3.2.

Theorem 3.3.2

(i) *Let σ_1 and σ_2 be the support functions of the nonempty closed convex sets S_1 and S_2. If t_1 and t_2 are positive, then*

$$t_1\sigma_1 + t_2\sigma_2 \text{ is the support function of } \operatorname{cl}(t_1 S_1 + t_2 S_2) .$$

(ii) *Let* $\{\sigma_j\}_{j\in J}$ *be the support functions of the family of nonempty closed convex sets* $\{S_j\}_{j\in J}$. *Then*

$$\sup_{j\in J}\sigma_j \text{ is the support function of } \overline{\text{co}}\,\{\cup S_j \,:\, j\in J\}\,.$$

(iii) *Let* $\{\sigma_j\}_{j\in J}$ *be the support functions of the family of closed convex sets* $\{S_j\}_{j\in J}$. *If*

$$S := \bigcap_{j\in J} S_j \neq \emptyset\,,$$

then

$$\sigma_S = \overline{\text{co}}\,\{\inf \sigma_j \,:\, j\in J\}\,.$$

Proof. [*(i)*] Call S the closed convex set $\text{cl}(t_1S_1+t_2S_2)$. By definition, its support function is

$$\sigma_S(d) = \sup\{\langle t_1s_1+t_2s_2, d\rangle \,:\, s_1\in S_1,\ s_2\in S_2\}\,.$$

In the above expression, s_1 and s_2 run independently in their index sets S_1 and S_2, t_1 and t_2 are positive, so

$$\sigma_S(d) = t_1\sup_{s\in S_1}\langle s,d\rangle + t_2\sup_{s\in S_2}\langle s,d\rangle\,.$$

[*(ii)*] The support function of $S := \cup_{j\in J}S_j$ is

$$\sup_{s\in\cup S_j}\langle s,d\rangle = \sup_{j\in J}\left[\sup_{s_j\in S_j}\langle s_j,d\rangle\right] = \sup_{j\in J}\sigma_j(d)\,.$$

This implies (ii) since $\sigma_S = \sigma_{\overline{\text{co}}\,S}$.

[*(iii)*] The set $S := \cap S_j$ being nonempty, it has a support function σ_S. Now, from Corollary 3.1.2,

$$\begin{aligned} s\in S &\iff s\in S_j \quad \text{for all } j\in J\\ &\iff \langle s,\cdot\rangle \leqslant \sigma_j \quad \text{for all } j\in J\\ &\iff \langle s,\cdot\rangle \leqslant \inf_{j\in J}\sigma_j \quad \iff \quad \langle s,\cdot\rangle \leqslant \overline{\text{co}}(\inf_{j\in J}\sigma_j)\end{aligned}$$

where the last equivalence comes directly from the Definition B.2.5.3 of a closed convex hull. Again Corollary 3.1.2 tells us that the closed sublinear function $\overline{\text{co}}(\inf\sigma_j)$ is just the support function of S. □

Observe in (i) that, if S_2 is bounded, then $t_1S_1+t_2S_2$ is automatically closed.

As for (iii), we have seen in Proposition 1.3.2(ii) that, if $J=\{1,\dots,m\}$ is a finite set, then the "co" operation can be replaced by the infimal convolution: there holds

$$\sigma_{S_1\cap\ldots\cap S_m} = \text{cl}\,(\sigma_1 \mathbin{\dot{+}} \cdots \mathbin{\dot{+}} \sigma_m)\,. \tag{3.3.1}$$

This last formula is a simplification of (iii), but the closure operation should not be forgotten, and it is something really complicated; these issues will be addressed more thoroughly in §E.2.3.

Returning to the end of Example 2.3.1, let K be a closed convex cone and take $K' := K \cap B(0,1)$. In view of the above observation, the support function of K' is given by an inf-convolution:

$$\sigma_{K'}(d) = \operatorname{cl}\{\inf_y[\sigma_K(y) + \sigma_B(d-y)]\}.$$

Since $\sigma_K = \mathrm{i}_{K^\circ}$, the infimum forces y to be in K°, in which case σ_K vanishes; knowing that $\sigma_{B(0,1)} = \|\cdot\|$, the infimum is

$$\inf\{\|d-y\| \;:\; y \in K^\circ\}.$$

Here, we are in a favourable case: this infimum is actually a minimum – achieved at the projection $\mathrm{p}_{K^\circ}(d)$ – and the result is a finite convex function, hence continuous; the closure operation is useless and can be omitted. In a word,

$$\sigma_{K\cap B(0,1)} = \mathrm{d}_{K^\circ}. \tag{3.3.2}$$

Positive homogeneity can also be exploited in Theorem 3.3.2(i) to write

$$\sigma_{tS}(d) = \sigma_S(td) \quad \text{for all } d \in \mathbb{R}^n \text{ and } t > 0,$$

a formula which also holds for negative t (just write the definition). More generally:

Proposition 3.3.3 *Let $A : \mathbb{R}^n \to \mathbb{R}^m$ be a linear operator, with adjoint A^* (for some scalar product $\langle\langle\cdot,\cdot\rangle\rangle$ in $\mathbb{R}^m$). For $S \subset \mathbb{R}^n$ nonempty, we have*

$$\sigma_{\operatorname{cl} A(S)}(y) = \sigma_S(A^*y) \quad \textit{for all } y \in \mathbb{R}^m.$$

Proof. Just write the definitions

$$\sigma_{A(S)}(y) = \sup_{s\in S}\langle\langle As, y\rangle\rangle = \sup_{s\in S}\langle s, A^*y\rangle$$

and use Proposition 2.2.1 to obtain the result. □

Taking an image-function (see §B.2.4) is another operation involving a linear operator. Its status is slightly more delicate.

Proposition 3.3.4 *Let $A : \mathbb{R}^m \to \mathbb{R}^n$ be a linear operator, with adjoint A^* (for some scalar product $\langle\langle\cdot,\cdot\rangle\rangle$ in $\mathbb{R}^m$). Let σ be the support function of a nonempty closed convex set $S \subset \mathbb{R}^m$. If σ is minorized on the inverse image*

$$\overset{-1}{A}(d) = \{p \in \mathbb{R}^m \;:\; Ap = d\} \tag{3.3.3}$$

of each $d \in \mathbb{R}^n$, then the support function of the set $(\overset{-1}{A^*})(S)$ is the closure of the image-function $A\sigma$.

Proof. Our assumption is tailored to guarantee $A\sigma \in \operatorname{Conv}\mathbb{R}^n$ (Theorem B.2.4.2). The positive homogeneity of $A\sigma$ is clear: for $d \in \mathbb{R}^n$ and $t > 0$,

$$(A\sigma)(td) = \inf_{Ap=td}\sigma(p) = \inf_{A(p/t)=d} t\sigma(p/t) = t\inf_{Aq=d}\sigma(q) = t(A\sigma)(d).$$

Thus, the closed sublinear function $\mathrm{cl}\,(A\sigma)$ supports some set S'; by definition, $s \in S'$ if and only if

$$\langle s, d\rangle \leqslant \inf\{\sigma(p) : Ap = d\} \quad \text{for all } d \in \mathbb{R}^n ;$$

but this just means

$$\langle s, Ap\rangle \leqslant \sigma(p) \quad \text{for all } p \in \mathbb{R}^m ,$$

i.e. $A^*s \in S$, because $\langle s, Ap\rangle = \langle\!\langle A^*s, p\rangle\!\rangle$. □

The inverse image $(\overset{-1}{A^*})(S)$ of the closed set S under the continuous mapping A^* is closed. By contrast, $A\sigma$ need not be a closed function. As a particular case, suppose that S is bounded (σ_S is finite everywhere) and that A is surjective; then $A\sigma$ is finite everywhere as well, which means that $(\overset{-1}{A^*})(S)$ is compact.

Remark 3.3.5 The assumption made in Proposition 3.3.4 means exactly that the function $A\sigma$ is nowhere $-\infty$; in other words, its closure $\mathrm{cl}\,(A\sigma)$ is the support function of a *nonempty* set: $(\overset{-1}{A^*})(S) \neq \emptyset$. This last property can be rewritten as

$$S \cap \operatorname{Im} A^* \neq \emptyset \quad \text{or} \quad 0 \in S - \operatorname{Im} A^* = S + (\operatorname{Ker} A)^{\perp} . \tag{3.3.4}$$

□

It has already been mentioned that taking an image-function is an important operation, from which several other operations can be constructed. We give two examples inspired from those at the end of §B.2.4:

- Let S_1 and S_2 be two nonempty closed convex sets of $\mathbb{R}^n$, with support functions σ_1 and σ_2 respectively. With $\mathbb{R}^m = \mathbb{R}^n \times \mathbb{R}^n$, take $A(x, y) := x + y$ and $\sigma(d_1, d_2) := \sigma_1(d_1) + \sigma_2(d_2)$; observe that σ is the support function of $S := S_1 \times S_2$, associated with the scalar product

$$\langle\!\langle (s_1, s_2), (d_1, d_2)\rangle\!\rangle := \langle s_1, d_1\rangle + \langle s_2, d_2\rangle .$$

  Then we obtain $A\sigma = \sigma_1 \overset{+}{\vee} \sigma_2$. On the other hand, the adjoint of A is clearly given by

$$A^*x = (x, x) \in \mathbb{R}^n \times \mathbb{R}^n \quad \text{for all } x \in \mathbb{R}^n ,$$

  so that the inverse image of S under A^* is nothing but $S_1 \cap S_2$: we recover (3.3.1).
- Let σ be the support function of some nonempty closed convex set $S \subset \mathbb{R}^n \times \mathbb{R}^p$ and let $A(x, y) := x$, so that the image of σ under A is defined by

$$\mathbb{R}^n \ni x \mapsto (A\sigma)(x) = \inf\{\sigma(x, y) : y \in \mathbb{R}^p\} .$$

  Now A^* is

$$\mathbb{R}^n \ni x \mapsto A^*x = (x, 0) \in \mathbb{R}^n \times \mathbb{R}^p$$

  and $\mathrm{cl}\,(A\sigma)$ is the support function of the slice $\{x \in \mathbb{R}^n : (x, 0) \in S\}$. This last set must not be confused with the projection of S onto $\mathbb{R}^n$, whose support function is $x \mapsto \sigma_S(x, 0)$ (Proposition 3.3.3).

Having studied the isomorphism with respect to order and algebraic structures, we pass to topologies. Theorem 1.3.3 has defined a distance Δ on the set of finite sublinear functions. Likewise, the *Hausdorff distance* Δ_H can be defined for nonempty closed sets (see §1.5). When restricted to nonempty compact convex sets, Δ_H plays the role of the distance introduced in Theorem 1.3.3:

Theorem 3.3.6 *Let S and S' be two nonempty compact convex sets of $\mathbb{R}^n$. Then*

$$\Delta(\sigma_S, \sigma_{S'}) := \max_{\|d\| \leqslant 1} |\sigma_S(d) - \sigma_{S'}(d)| = \Delta_H(S, S'). \qquad (3.3.5)$$

Proof. As mentioned in §1.5.1, for all $r \geqslant 0$, the property

$$\max \{ \mathrm{d}_S(d) \ : \ d \in S' \} \leqslant r \qquad (3.3.6)$$

simply means $S' \subset S + B(0, r)$.

Now, the support function of $B(0,1)$ is $\| \cdot \|$ – see (2.3.1). Calculus rules on support functions therefore tell us that (3.3.6) is also equivalent to

$$\sigma_{S'}(d) \leqslant \sigma_S(d) + r\|d\| \quad \text{for all } d \in \mathbb{R}^n ,$$

which in turn can be written

$$\max_{\|d\| \leqslant 1} [\sigma_{S'}(d) - \sigma_S(d)] \leqslant r .$$

In summary, we have proved

$$\max_{d \in S'} \mathrm{d}_S(d) = \max_{\|d\| \leqslant 1} [\sigma_{S'}(d) - \sigma_S(d)]$$

and symmetrically

$$\max_{d \in S} \mathrm{d}_{S'}(d) = \max_{\|d\| \leqslant 1} [\sigma_S(d) - \sigma_{S'}(d)] ;$$

the result follows. □

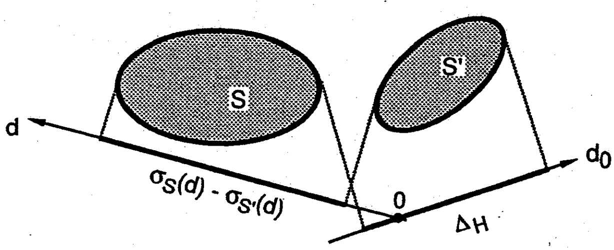

Fig. 3.3.1. Hausdorff distances

Naturally, the max in (3.3.5) is attained at some d_0: for S and S' convex compact, there exists d_0 of norm 1 such that

$$\Delta_H(S, S') = \Delta(\sigma_S, \sigma_{S'}) = |\sigma_S(d_0) - \sigma_{S'}(d_0)| .$$

Figure 3.3.1 illustrates a typical situation. When $S' = \{0\}$, we obtain the number

$$\Delta_H(\{0\}, S) = \max_{s \in S} \|s\| = \max_{\|d\|=1} \sigma_S(d) ,$$

already seen in (1.2.6); it is simply the distance from 0 to the most remote hyperplane $H_{d,\sigma_S(d)}$ touching S (see again the end of Interpretation 2.1.5).

Using (3.3.5), it becomes rather easy to compute the distance in Example 1.3.4, which becomes the Hausdorff distance (in fact an excess) between the corresponding unit balls.

When speaking of limits of nonempty convex compact sets to a nonempty convex compact set, the following result is a further illustration of our isomorphism.

Proposition 3.3.7 *A convex-compact-valued and locally bounded multifunction $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is outer [resp. inner] semi-continuous at $x_0 \in \operatorname{int} \operatorname{dom} F$ if and only if its support function $x \mapsto \sigma_{F(x)}(d)$ is upper [resp. lower] semi-continuous at x_0 for all d of norm 1.*

Proof. Calculus with support functions tells us that our definition (1.5.2) of outer semi-continuity is equivalent to

$$\forall \varepsilon > 0, \exists \delta > 0 : y \in B(x_0, \delta) \implies \sigma_{F(y)}(d) \leqslant \sigma_{F(x_0)}(d) + \varepsilon\|d\| \text{ for all } d \in \mathbb{R}^n$$

and division by $\|d\|$ shows that this is exactly upper semi-continuity of the support function for $\|d\| = 1$. Same proof for inner/lower semi-continuity. □

Thus, a convex-compact-valued, locally bounded mapping F is both outer and inner semi-continuous at x_0 if and only if its support function $\sigma_{F(\cdot)}(d)$ is continuous at x_0 for all d. In view of Theorem 1.3.5, $\sigma_{F(\cdot)}(d)$ is then continuous at x_0 uniformly for $d \in B(0,1)$; and Theorem 3.3.6 tells us that this property in turn means:

$$\Delta_H(F(x), F(x_0)) \to 0 \quad \text{when} \quad x \to x_0 .$$

The following interpretation in terms of sequences is useful.

Corollary 3.3.8 *Let (S_k) be a sequence of nonempty convex compact sets and S a nonempty convex compact set. When $k \to +\infty$, the following are equivalent*

(i) *$S_k \to S$ in the Hausdorff sense, i.e. $\Delta_H(S_k, S) \to 0$;*
(ii) *$\sigma_{S_k} \to \sigma_S$ pointwise;*
(iii) *$\sigma_{S_k} \to \sigma_S$ uniformly on each compact set of $\mathbb{R}^n$.* □

Let us sum up this Section 3.3: when combining/comparing closed convex sets, one knows what happens to their support functions (apply the results 3.3.1 – 3.3.3). Conversely, when closed sublinear functions are combined/compared, one knows what happens to the sets they support. The various rules involved are summarized in Table 3.3.1. Each S_j is a nonempty closed convex set, with support function σ_j.

This table deserves some comments.

– Generally speaking, it helps to remember that when a set increases, its support function increases (first line); hence the "crossing" of closed convex hulls in the last two lines.
– The rule of the last line comes directly from the definition (2.1.1) of a support function, if each S_j is thought of as a singleton.
– Most of these rules are still applicable without closed convexity of each S_j (remembering that $\sigma_S = \sigma_{\overline{\text{co}}\, S}$). For example, the equivalence in the first line requires closed convexity of S_2 only. We mention one trap, however: when intersecting sets, each set must be closed and convex. A counter-example in one dimension is $S_1 := \{0, 1\}$, $S_2 := \{0, 2\}$; the support functions do not see the difference between $S_1 \cap S_2 = \{0\}$ and $\text{co}\, S_1 \cap \text{co}\, S_2 = [0, 1]$.

Table 3.3.1. Calculus rules for support functions

Closed convex sets	Closed sublinear functions
$S_1 \subset S_2$	$\sigma_1 \leqslant \sigma_2$
$\Delta_H(S_1, S_2)$ (S_i bounded)	$\Delta(\sigma_1, \sigma_2)$ (σ_i finite)
Hausdorff convergence (on bounded sets)	uniform/compact or pointwise convergence (on finite functions)
tS $(t > 0)$	$t\sigma$
$\text{cl}\,(S_1 + S_2)$	$\sigma_1 + \sigma_2$
$\text{cl}\, A(S)$ (A linear)	$\sigma \circ A^*$
$(\overset{-1}{A^*})(S)$ (A linear)	$\text{cl}\,(A\sigma)$
$\cap_{j \in J} S_j$ (nonempty)	$\overline{\text{co}}\,(\inf_{j \in J} \sigma_j)$ (minorized)
$\overline{\text{co}}(\cup_{j \in J} S_j)$	$\sup_{j \in J} \sigma_j$

Example 3.3.9 (Maximal Eigenvalues) Recall from §B.1.3(e) that, if the eigenvalues of a symmetric matrix A are denoted by $\lambda_1(A) \geqslant \cdots \geqslant \lambda_n(A)$, the function

$$\text{S}_n(\mathbb{R}) \ni A \mapsto f_m(A) := \sum_{j=1}^{m} \lambda_j(A)$$

is convex – and finite everywhere. Its positive homogeneity is obvious, therefore it is the support function of a certain convex compact set C_m of symmetric matrices. Let us compute the set C_1 when the scalar product in $\text{S}_n(\mathbb{R})$ is the standard dot-product of $\mathbb{R}^{n \times n}$:

$$\langle\!\langle A, B \rangle\!\rangle := \text{tr}\, AB = \sum_{i,j=1}^{n} A_{ij} B_{ij}\,.$$

Indeed, we know that

$$\lambda_1(A) = \sup_{x^\top x = 1} x^\top A x = \sup_{x^\top x = 1} \langle\!\langle x x^\top, A \rangle\!\rangle\,.$$

Hence C_1 is the closed convex hull of the set of matrices $\{xx^\top : x^\top x = 1\}$, which is clearly compact. Actually, its Hausdorff distance to $\{0\}$ is

$$\Delta_H(\{0\}, C_1) = \max_{x^\top x=1} \sqrt{\langle\langle xx^\top, xx^\top \rangle\rangle} = 1\,.$$

Incidentally, $\lambda_1(\cdot)$ is therefore nonexpansive in $S_n(\mathbb{R})$.

We leave it as an exercise to prove the following nicer representation of C_1:

$$C_1 = \operatorname{co}\{xx^\top \,:\, x^\top x = 1\} = \{M \in S_n(\mathbb{R}) \,:\, \lambda_n(M) \geqslant 0,\ \operatorname{tr} M = 1\}\,,$$

generalizing to $S_n(\mathbb{R})$ the expression of the unit-simplex in $\mathbb{R}^n$: C_1 could be called the unit *spectraplex*. □

3.4 Example: Support Functions of Closed Convex Polyhedra

Polyhedral sets are encountered all the time, and thus deserve special study. They are often defined by finitely many affine constraints, i.e. obtained as intersections of closed half-spaces; in view of Table 3.3.1, this explains that the infimal convolution encountered in Proposition 1.3.2 is fairly important.

Example 3.4.1 (Compact Convex Polyhedra) First of all, the support function of a polyhedron defined as

$$P := \operatorname{co}\{p_1, \ldots, p_m\} \tag{3.4.1}$$

is trivially

$$d \mapsto \sigma_P(d) = \max\{\langle p_i, d\rangle \,:\, i = 1, \ldots, m\}\,.$$

There is no need to invoke Theorem 3.3.2 for this: a linear function $\langle \cdot, d\rangle$ attains its maximum on an extreme point of P (Proposition A.2.4.6), even if this extreme point is not the entire face exposed by d. □

Example 3.4.2 (Closed Convex Polyhedral Cones) Taking again Example 2.3.1, suppose that the cone K is given as a finite intersection of half-spaces:

$$K = \cap\{K_j \,:\, j = 1, \ldots, m\}\,, \tag{3.4.2}$$

where

$$K_j := H^-_{a_j,0} := \{s \in \mathbb{R}^n \,:\, \langle a_j, s\rangle \leqslant 0\} \tag{3.4.3}$$

(the a_j's are assumed nonzero). We use Proposition 1.3.2:

$$\sigma_K(d) = \operatorname{cl}\inf\Big\{\sum_{j=1}^m \sigma_{K_j}(d_j) \,:\, \sum_{j=1}^m d_j = d\Big\}\,.$$

Only those d_j in K_j° – namely nonnegative multiples of a_j, see (2.3.2) – count to yield the infimum; their corresponding support vanishes and we obtain

$$\sigma_K(d) = \begin{cases} 0 & \text{if } d = \sum_{j=1}^m t_j a_j,\ t_j \geqslant 0 \text{ for } j = 1, \ldots, m\,, \\ +\infty & \text{otherwise}\,. \end{cases}$$

Here, we are lucky: the closure operation is useless because the righthand side is already a closed convex function. Note that we recognize Farkas' Lemma A.4.3.3: $K^\circ = \operatorname{dom} \sigma_K$ is the conical hull of the a_j's, which is closed thanks to the fact that there are *finitely many* generators. □

Example 3.4.3 (Extreme Points and Directions) Suppose our polyhedron is defined in the spirit of 3.4.1, but unbounded:

$$S := \operatorname{co}\{p_1, \ldots, p_m\} + \operatorname{cone}\{a_1, \ldots, a_\ell\}.$$

Then it suffices to observe that $S = P + K^\circ$, with P of (3.4.1) and K of (3.4.2), (3.4.3). Using Table 3.3.1 and knowing that $K^{\circ\circ} = K$ – hence $\sigma_{K^\circ} = \mathrm{i}_K$:

$$\sigma_S(d) = \begin{cases} \max\limits_{i=1,\ldots,m} \langle p_i, d\rangle & \text{if } \langle a_j, d\rangle \leqslant 0 \text{ for } j = 1, \ldots, \ell, \\ +\infty & \text{otherwise}. \end{cases}$$ □

The representations of Examples 3.4.1 and 3.4.3 are not encountered so frequently. Our next examples, dealing with half-spaces, represent the vast majority of situations.

Example 3.4.4 (Inequality Constraints) Perturb Example 3.4.2 to express the support function of $S := \cap H^-_{a_j, b_j}$, with

$$H^-_{a,b} := \{s \in \mathbb{R}^n \; : \; \langle s, a\rangle \leqslant b\} \quad (a \neq 0).$$

Here, we deal with translations of the K_j's: $H^-_{a_j,b_j} = \frac{b_j}{\|a_j\|^2} a_j + K_j$ so, with the help of Table 3.3.1:

$$\sigma_{H^-_{a_j,b_j}}(d) = \begin{cases} t b_j & \text{if } d = t a_j,\ t \geqslant 0, \\ +\infty & \text{otherwise}. \end{cases}$$

Provided that $S \neq \emptyset$, our support function σ_S is therefore the closure of the function

$$d \mapsto \begin{cases} \inf\left\{\sum\limits_{j=1}^m t_j b_j \; : \; \sum\limits_{j=1}^m t_j a_j = d,\ t_j \geqslant 0\right\} & \text{if } d \in \operatorname{cone}\{a_1, \ldots, a_m\}, \\ +\infty & \text{otherwise}. \end{cases}$$ □

Now we have a sudden complication: the domain of σ is still the closed convex cone K°, but the status of the closure operation is no longer quite clear. Also, it is not even clear whether the above infimum is attained. Actually, all this results from Farkas' Lemma of §A.4.3; before giving the details, let us adopt different notation.

Example 3.4.5 (Closed Convex Polyhedra in Standard Form) Even though Example 3.4.4 uses the Definition A.4.2.5 of a closed convex polyhedron, the following "standard" description is often used. Let A be a linear operator from $\mathbb{R}^n$ to $\mathbb{R}^m$, $b \in \operatorname{Im} A \subset \mathbb{R}^m$, $K \subset \mathbb{R}^n$ a closed convex polyhedral cone (K is usually characterized as in Example 3.4.2). Then S is given by

$$S := \{s \in \mathbb{R}^n \,:\, As = b, s \in K\} = (\{s_0\} + H) \cap K\,, \tag{3.4.4}$$

where s_0 is some point in $\mathbb{R}^n$ satisfying $As_0 = b$, and $H := \operatorname{Ker} A$.

With the expression of σ_H given in Example 2.3.1, we see that the support function of $\{s_0\} + H$ is finite only on $\operatorname{Im} A^*$, where it is equal to

$$\sigma_{\{s_0\}}(d) + \sigma_H(d) = \langle s_0, d\rangle = \langle\!\langle b, z\rangle\!\rangle \quad \text{for } d = A^*z,\ z \in \mathbb{R}^m$$

(here, $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ denotes the scalar product in $\mathbb{R}^m$). Thus, σ_S is the closure of the infimal convolution

$$\left(\sigma_{\{s_0\}} + \sigma_H\right) \mathbin{\dot{+}} \sigma_K = \left(\sigma_{\{s_0\}} + \sigma_H\right) \mathbin{\dot{+}} \mathrm{i}_{K^\circ}\,, \tag{3.4.5}$$

which can be made explicit as the function

$$d \mapsto \inf\{\langle\!\langle b, z\rangle\!\rangle \,:\, (z, y) \in \mathbb{R}^m \times K^\circ,\ A^*z + y = d\}\,.$$

Of course, this formula clearly reveals

$$\operatorname{dom}\sigma_S = \operatorname{dom}\sigma_H + \operatorname{dom}\mathrm{i}_{K^\circ} = \operatorname{Im} A^* + K^\circ\,.$$

In the pure *standard form*, $\mathbb{R}^n$ and $\mathbb{R}^m$ are both equipped with the standard dot-product – A being a matrix with m rows and n columns – and K is the nonnegative orthant; K° is therefore the nonpositive orthant. Our "standard" S of (3.4.4) is now

$$\{s \in \mathbb{R}^n \,:\, As = b,\ s \geqslant 0\}\,, \tag{3.4.6}$$

assumed nonempty. Then (3.4.5) becomes

$$\inf\{b^\top z \,:\, A^\top z \geqslant d\}\,, \tag{3.4.7}$$

a function of d which is by no means simpler than in 3.4.4 – only the notation is different. In summary, the support function

$$\sigma_S(d) = \sup\{s^\top d \,:\, As = b,\ s \geqslant 0\} \tag{3.4.8}$$

of the set (3.4.6) is the closure of (3.4.7), considered as a function of $d \in \mathbb{R}^n$.

Now, invoke Farkas' Lemma: write the equivalent statements (i)" and (ii)" from the end of §A.4, with (x, p, α, r) changed to $(-z, -d, s, -\sigma)$:

$$\{z \in \mathbb{R}^n \,:\, A^\top z \geqslant d\} \subset \{z \in \mathbb{R}^n \,:\, b^\top z \geqslant \sigma\} \tag{3.4.9}$$

is equivalent to

$$\exists s \geqslant 0 \quad \text{such that} \quad As = b,\ s^\top d \geqslant \sigma\,. \tag{3.4.10}$$

In other words: the largest σ for which (3.4.9) holds – i.e. the value (3.4.7) – is also the largest σ for which (3.4.10) holds – i.e. $\sigma_S(d)$. The closure operation can be omitted and we do have

$$\sigma_S(d) = \inf\{b^\top z \,:\, A^\top z \geqslant d\} \quad \text{for all } d \in \mathbb{R}^n\,.$$

Another interesting consequence can be noted. Take d such that $\sigma_S(d) < +\infty$: if we put $\sigma = \sigma_S(d)$ in (3.4.9), we obtain a true statement, i.e. (3.4.10) is also true. This means that the supremum in (3.4.8) is attained when it is finite. □

It is worth noting that Example 3.4.5 describes general polyhedral functions, up to notational changes. As such, it discloses results of general interest, namely:

– A linear function which is bounded from below on a closed convex polyhedron attains its minimum on this polyhedron.
– The infimum of a linear function under affine constraints is a closed sublinear function of the righthand side; said otherwise, an image of a polyhedral function is closed: in Example 3.4.5, the polyhedral function in question is

$$\mathbb{R}^m \times \mathbb{R}^n \ni (y, z) \mapsto b^\top z + \mathrm{i}_K(y),$$

and (3.4.7) gives its image under the linear mapping $[\,A^\top \mid 0\,]$.

Exercises

*1**. Let C_1 and C_2 be two nonempty closed convex sets in $\mathbb{R}^n$ and let S be bounded. Show that $C_1 + S = C_2 + S$ implies $C_1 = C_2$.

*2***. Let P be a compact convex polyhedron on $\mathbb{R}^n$ with nonempty interior. Show that P has at least $n+1$ facets.

*3**. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be positively homogeneous, not identically $+\infty$ and minorized by a linear function. Show that $\overline{\mathrm{co}}\, f(x)$ is the supremum of $\sigma(x)$ over all closed sublinear functions σ minorizing f.

*4**. Let f and g be two gauges. Show that $f \overset{+}{\vee} g$ is still a gauge and that

$$\{x : (f \overset{+}{\vee} g)(x) < 1\} = \mathrm{co}\left(\{x : f(x) < 1\} \cup \{x : g(x) < 1\}\right).$$

*5***. Let C be a compact convex set with $0 \in \mathrm{int}\, C$, so that its polar set C° enjoys the same properties. Show that the following relations are equivalent:

(i) the hyperplane $H_{s,1}$ supports C at $x \in C$;
(ii) the hyperplane $H_{x,1}$ supports C° at $s \in C^\circ$;
(iii) $x \in \mathrm{bd}\, C$, $s \in \mathrm{bd}\, C^\circ$, $\langle s, x \rangle = 1$.

6. Draw a picture to compute the support function σ_P of the parabolic set $P := \{(\xi, \eta) \in \mathbb{R}^2 : \eta \geqslant \frac{1}{2}\eta^2\}$. What is $\mathrm{dom}\, \sigma_P$? Show that σ_P is not upper semi-continuous on its domain.

7. Let $|\!|\!|\cdot|\!|\!|$ be a norm on $\mathbb{R}^n$ and denote by $S := \{x \in \mathbb{R}^n : |\!|\!|x|\!|\!| = 1\}$ the associated unit sphere. What is the convex hull of S?

8. Let H be a hyperplane in $\mathbb{R}^n$ and suppose the set S is contained in one of the corresponding half-spaces: $S \subset H_-$. Show that $\overline{\mathrm{co}}\,(S \cap H) = (\overline{\mathrm{co}}\, S) \cap H$. Compare with Exercise III.21.

9. Let $x \in C$, where C is a closed convex set in $\mathbb{R}^n$. Show that $x \in \mathrm{ri}\, C$ if and only if the normal cone $\mathrm{N}_C(x)$ is a subspace.

10. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and $A \subset \mathbb{R}^n$ be given (none of which is assumed convex). Assume f is minorized by some affine function and set $f_A := f + \mathrm{i}_A$.

– Show that $\overline{\mathrm{co}}\,\mathrm{epi}\, f_A \subset \{(x, r) \in \mathbb{R}^n \times \mathbb{R} : x \in \overline{\mathrm{co}}\, A\,,\ r \geqslant \overline{\mathrm{co}}\, f(x)\}$.

– Construct an example for which equality need not hold, even if A is closed convex.

– Show that equality does hold if f is affine.

11.* Recall that $M \succcurlyeq 0$ means: $M \in \mathrm{S}_n(\mathbb{R})$ is positive semidefinite. Show that the *unit spectraplex* $\{M \in \mathrm{S}_n(\mathbb{R}) : M \succcurlyeq 0, \mathrm{tr}\, M = 1\}$ is a convex compact set, whose support function is $M \mapsto \lambda_{\max}(M)$.

12.* In the Euclidean space of $n \times n$ real matrices (equipped with the scalar product $\langle M, P\rangle = \mathrm{tr}\,(M^\top P) = \sum_{ij} M_{ij}P_{ij}$ and associated norme $\|\cdot\|$), denote by S the set of orthogonal matrices ($M^\top M = MM^\top = \mathrm{I}_n$).

– Show that S is contained in the sphere $\widetilde{B}(0, \sqrt{n}) = \{M : \|M\| = \sqrt{n}\}$; deduce that S is compact.

– Show that the support function of S is $M \mapsto \sigma_S(M) = \mathrm{tr}\,(M^\top M)^{1/2}$ [hint: use the polar decomposition of M].

– Express the values of the support functions $\sigma_S(M)$ and $\sigma_{B(0,\sqrt{n})}(M)$ in terms of the eigenvalues of $M^\top M$.

13. Let $K := \{M \in \mathrm{S}_n(\mathbb{R}) : M \succcurlyeq 0\}$ be the cone of symmetric positive semidefinite matrices. What are its interior int K and boundary bd K?

Show that the distance of a symmetric positive semidefinite matrix to the boundary of K is its smallest eigenvalue: $\mathrm{d}_{\mathrm{bd}\,K}(M) = \lambda_{\min}(M)$ for $M \in K$.

14.* Let $(S_k)_k$ be a nested family of compact sets in $\mathbb{R}^n$ ($S_{k+1} \subset S_k$ for all k). Show that $\mathrm{co}\,(\cap_k S_k) = \cap_k \mathrm{co}\, S_k$. Exhibit an example showing that the result becomes wrong if the S_k's are unbounded.

15.* Let $F : \mathbb{R} \longrightarrow \mathbb{R}^n$ be a multifunction such that $F(t)$ is closed convex for each t. Show that the graph of F is a convex set in $\mathbb{R} \times \mathbb{R}^n$ if and only if the function $t \mapsto \sigma_{F(t)}(d)$ is concave for all $d \in \mathbb{R}^n$.

Let C be closed convex in $\mathbb{R}^n$ and $\sigma' : \mathbb{R}^n \to \mathbb{R}$ be nonnegative and positively homogeneous: $+\infty > \sigma'(\lambda d) = \lambda\sigma'(d) \geqslant 0$ for all $(\lambda, d) \in \mathbb{R}_+ \times \mathbb{R}^n$. Show that, for any $t \geqslant 0$, the set $F(t) := \{x \in \mathbb{R}^n : \langle \cdot, x\rangle \leqslant \sigma_C(\cdot) + t\sigma'(\cdot)\}$ is closed convex and contains C. What is its support function? Show that this defines a multifunction $F : \mathbb{R}_+ \longrightarrow \mathbb{R}^n$ whose graph is convex.