2. Convex Sets and Convex Functions

We have encountered convex sets and convex functions on several occasions. Here we would like to discuss these notions in a more systematic way. Among nonlinear functions, the convex ones are the closest ones to the linear, in fact, functions that are convex and concave at the same time are just the linear affine functions.

Although convex figures appear since the beginning of mathematics — Archimedes, for instance, observed and made use of the fact that the perimeter of a convex figure K is larger than the perimeter of any other convex figure contained in K, more recently convexity played a relevant role in the study of the thermodynamic equilibrium by J. Willard Gibbs (1839–1903) — the systematic study of convexity began in the early years of the twentieth century with Hermann Minkowski (1864–1909), continued with the treatise of T. Bonnesen and Werner Fenchel (1905–1986) in 1934 and developed after 1950 both in finite and infinite dimensions due to its relevance in several branches of mathematics. Here we shall deal only with convexity in finite-dimensional spaces.

2.1 Convex Sets

a. Definitions

2.1 Definition. *A set $K \subset \mathbb{R}^n$ is said to be* convex *if either $K = \emptyset$ or, whenever we take two points in K, the segment that connects them is entirely contained in K, i.e.,*

$$\lambda x_1 + (1-\lambda)x_2 \in K \qquad \forall\ \lambda \in [0,1],\ \forall\ x_1, x_2 \in K.$$

The following properties, the proof of which we leave to the reader, follow easily from the definition.

2.2 ¶. Show the following:

(i) A linear subspace of $\mathbb{R}^n$ is convex.

M. Giaquinta and G. Modica, *Mathematical Analysis, Foundations and Advanced Techniques for Functions of Several Variables*, DOI 10.1007/978-0-8176-8310-8_2,

ERGEBNISSE DER MATHEMATIK
UND IHRER GRENZGEBIETE
HERAUSGEGEBEN VON DER SCHRIFTLEITUNG
DES
„ZENTRALBLATT FÜR MATHEMATIK"
DRITTER BAND
1

THEORIE DER
KONVEXEN KÖRPER

VON

T. BONNESEN UND W. FENCHEL

MIT 8 FIGUREN

BERLIN
VERLAG VON JULIUS SPRINGER
1934

Figure 2.1. Hermann Minkowski (1864–1909) and the frontispiece of the treatise by T. Bonnesen and Werner Fenchel (1905–1986) on convexity.

(ii) Let $\ell : \mathbb{R}^n \to \mathbb{R}$ be linear and $\alpha \in \mathbb{R}$. Then the sets
$$\Big\{x \in \mathbb{R}^n \,\Big|\, \ell(x) < \alpha\Big\}, \qquad \Big\{x \in \mathbb{R}^n \,\Big|\, \ell(x) \le \alpha\Big\},$$
$$\Big\{x \in \mathbb{R}^n \,\Big|\, \ell(x) \ge \alpha\Big\}, \qquad \Big\{x \in \mathbb{R}^n \,\Big|\, \ell(x) > \alpha\Big\}$$
are convex.

(iii) The intersection of convex sets is convex; in particular, the intersection of any number of half-spaces is convex.

(iv) The interior and the closure of a convex set are convex.

(v) If K is convex, then $\mathrm{cl}(\mathrm{int}(K)) = \mathrm{cl}(K)$, $\mathrm{int}(\mathrm{cl}(K)) = \mathrm{int}(K)$.

(vi) If K is convex, then for $x_0 \in \mathbb{R}^n$ and $t \in \mathbb{R}$ the set
$$tx_0 + (1-t)K := \Big\{x \in \mathbb{R}^n \,\Big|\, x = tx_0 + (1-t)y,\ y \in K\Big\},$$
i.e., the *cone* with vertex x_0 generated by K, is convex.

A *linear combination* of points $(x_1, x_2, \dots, x_k) \in \mathbb{R}^n$, $\sum_{i=1}^k \lambda_i x_i$, with coefficients $\lambda_1, \lambda_2, \dots, \lambda_k$ such that $\sum_{i=1}^k \lambda_i = 1$ and $\lambda_i \ge 0\ \forall i$, is called a *convex combination* of $x_1, \dots, x_k$. The coefficients $\lambda_1, \lambda_2, \dots, \lambda_k$ are called the *barycentric coordinates* of $x := \sum_{i=1}^k \lambda_i x_i$.

Noticing that
$$\sum_{i=1}^k \lambda_i x_i = (1-\lambda_k) \sum_{i=1}^{k-1} \frac{\lambda_i}{1-\lambda_k} x_i + \lambda_k x_k,$$
whenever $0 < \lambda_k < 1$, we infer at once the following.

2.3 Proposition. *The set K is convex if and only if every convex combination of points in K is contained in K.*

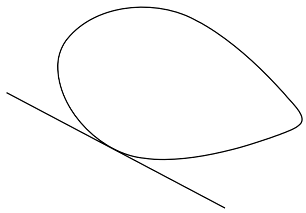


Figure 2.2. A support plane.

2.4 ¶. Show that the representation of a point x as convex combination of points $x_1, x_2, \dots, x_k$ is unique if and only if the vectors $x_2 - x_1$, $x_3 - x_1$, $\dots$, $x_k - x_1$ are linearly independent.

b. The support hyperplanes

We prove that every proper, nonempty, closed and convex subset of $\mathbb{R}^n$, $n \geq 2$, is the intersection of closed half-spaces. To do this, we first introduce the notions of *separating* and *supporting hyperplanes.*

2.5 Definition. *Let $\ell : \mathbb{R}^n \to \mathbb{R}$ be a linear function, $\alpha \in \mathbb{R}$ and $\mathcal{P}$ the hyperplane*

$$\mathcal{P} := \Big\{x \in \mathbb{R}^n \,\Big|\, \ell(x) = \alpha\Big\},$$

and let

$$\mathcal{P}_+ := \Big\{x \in \mathbb{R}^n \,\Big|\, \ell(x) \geq \alpha\Big\}, \qquad \mathcal{P}_- := \Big\{x \in \mathbb{R}^n \,\Big|\, \ell(x) \leq \alpha\Big\}$$

be the corresponding half-spaces that are the closed convex sets of $\mathbb{R}^n$ for which $\mathcal{P}_+ \cup \mathcal{P}_- = \mathbb{R}^n$ and $\mathcal{P}_+ \cap \mathcal{P}_- = \mathcal{P}$. We say that

(i) *two nonempty sets $A, B \subset \mathbb{R}^n$ are* separated *by $\mathcal{P}$ if $A \subset \mathcal{P}_+$ and $B \subset \mathcal{P}_-$;*
(ii) *two nonempty sets $A, B \subset \mathbb{R}^n$ are* strongly separated *by $\mathcal{P}$ if there is $\epsilon > 0$ such that*

$$\ell(x) \leq \alpha - \epsilon \;\forall\; x \in A \qquad \textit{and} \qquad \ell(x) \geq \alpha + \epsilon \;\forall\; x \in B.$$

(iii) *Let $K \subset \mathbb{R}^n$, $n \geq 2$. We say that $\mathcal{P}$ is a* supporting hyperplane *for K if $\mathcal{P} \cap \overline{K} \neq \emptyset$ and K is a subset of one of the two closed half-spaces $\mathcal{P}_+$ and $\mathcal{P}_-$ that is called a* supporting half-space *for K.*

2.6 Theorem. *Let K_1 and K_2 be two nonempty closed and disjoint convex sets. If either K_1 or K_2 is compact, then there exists a hyperplane that strongly separates K_1 and K_2.*

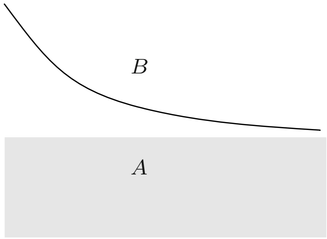


Figure 2.3. Two disjoint and closed convex sets that are not strongly separated.

Proof. Assume for instance that K_1 is compact and let $d := \inf\{|x-y| \,|\, x \in K_1,\ y \in K_2\}$. Clearly d is finite and, for R large,

$$d = \inf\Big\{ |x-y| \,\Big|\, x \in K_1,\ y \in K_2 \cap \overline{B(0,R)} \Big\}.$$

The Weierstrass theorem then yields $x_0 \in K_1$ and $y_0 \in K_2 \cap \overline{B(0,R)}$ such that

$$d = |x_0 - y_0| > 0.$$

The hyperplane through x_0 and perpendicular to $y_0 - x_0$,

$$\mathcal{P}' := \Big\{ x \in \mathbb{R}^n \,\Big|\, (x - x_0) \bullet (y_0 - x_0) = 0 \Big\},$$

is a supporting hyperplane for K_1. In fact, for $x \in K_1$, the function

$$\phi(\lambda) := |y_0 - (x_0 + \lambda(x - x_0))|^2, \qquad \lambda \in [0,1],$$

has a minimum at zero, hence

$$\phi'(0) = 2\,(y_0 - x_0) \bullet (x - x_0) \le 0. \tag{2.1}$$

Similarly, the hyperplane through y_0 and perpendicular to $x_0 - y_0$,

$$\mathcal{P}'' := \Big\{ x \in \mathbb{R}^n \,\Big|\, (x - y_0) \bullet (x_0 - y_0) = 0 \Big\},$$

is a supporting hyperplane for K_2. The conclusion then follows since $\mathrm{dist}(\mathcal{P}', \mathcal{P}'') = d > 0$. □

2.7 Theorem. *We have the following:*

(i) *Every boundary point of a closed and convex set $K \subset \mathbb{R}^n$, $n \ge 2$, is contained in at least a supporting hyperplane.*
(ii) *Every closed convex set $K \ne \emptyset, \mathbb{R}^n$ of $\mathbb{R}^n$ is the intersection of all its supporting half-spaces.*
(iii) *Let $K \subset \mathbb{R}^n$ be a closed set with nonempty interior. Then K is convex if and only if at each of its boundary point there is a supporting hyperplane.*

Proof. (i) Assume $\partial K \ne \emptyset$, i.e., $K \ne \emptyset, \mathbb{R}^n$, let $x_0 \in \partial K$, and let $\{y_k\} \subset \mathbb{R}^n \setminus K$ be a sequence with $y_k \to x_0$ as $k \to \infty$. Let x_k be a point of K nearest to y_k and

$$e_k := \frac{y_k - x_k}{|y_k - x_k|}.$$

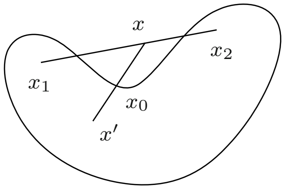


Figure 2.4. Illustration of the proof of (iii) Theorem 2.7.

Then $|e_k| = 1$, $x_k \to x_0$ as $k \to \infty$ and, as in the proof of Theorem 2.6, we see that the hyperplane through x_k and perpendicular to e_k is a supporting hyperplane for K,

$$K \subset \left\{ x \in \mathbb{R}^n \;\middle|\; e_k \bullet (x - x_k) \le 0 \right\}.$$

Possibly passing to a subsequence $\{e_k\}$ and $\{x_k\}$ converge, $e_k \to e$ and $x_k \to x_0$. It follows that

$$K \subset \left\{ x \in \mathbb{R}^n \;\middle|\; e \bullet (x - x_0) \le 0 \right\},$$

i.e., the hyperplane through x_0 perpendicular to e is a supporting hyperplane for K.

(ii) Since $K \neq \emptyset, \mathbb{R}^n$, the boundary of K is nonempty; in particular, the intersection K' of all its supporting half-spaces is closed, nonempty by (i), hence it contains K. Assume by contradiction that there is $x' \in K' \setminus K$. Since K is closed, there is a nearest point $x_0 \in K$ to x' and, as in the proof of Theorem 2.6,

$$K \subset S := \left\{ x \in \mathbb{R}^n \;\middle|\; (x' - x_0) \bullet (x - x_0) \le 0 \right\}.$$

On the other hand, from the definition of K', it follows that $K' \subset S$, hence $x' \in S$, which is a contradiction since $(x' - x_0) \bullet (x' - x_0) > 0$.

(iii) Let K be convex. By assumption $K \neq \emptyset$, if $K = \mathbb{R}^n$, we have $\partial K = \emptyset$ and nothing has to be proved. If $K \neq \mathbb{R}^n$, then through every boundary point there is a supporting hyperplane because of (i).

Conversely, suppose that K is not convex, in particular, $K \neq \emptyset, \mathbb{R}^n$ and $\partial K \neq \emptyset$. It suffices to show that through a point of ∂K there is no supporting hyperplane. Since K is not convex, there exists $x_1, x_2 \in K$ and x on the segment Σ connecting x_1 and x_2 with $x \notin K$. Let x' be a point in the interior of K and Σ' be the segment joining x with x'. At a point $x_0 \in \partial K \cap \Sigma'$ we claim that there is no supporting hyperplane. In fact, let π be such a hyperplane and let H be the corresponding supporting half-space. Since $x' \in \mathrm{int}(K)$, x' does not belong to π, thus Σ' is not contained in π. It follows that $x' \in \mathrm{int}(H)$ and $x \notin H$, hence x_1 and x_2 cannot both be in H since otherwise x also belongs to H. However, this contradicts the fact that H is a supporting half-space. $\square$

2.8 ¶. In (iii) of Theorem 2.7 the assumption that $\mathrm{int}(K) \neq \emptyset$ is essential; think of a curve without inflection points in $\mathbb{R}^2$.

2.9 ¶. Let K be a closed, convex subset of $\mathbb{R}^n$ with $K \neq \emptyset, \mathbb{R}^n$.

(i) Prove that K is the intersection of at most a denumerable supporting half-spaces.

(ii) Moreover, if K is compact, then for any open set $A \supset K$ there exists finitely many supporting half-spaces such that

$$K \subset \bigcap_{k=1}^{N} H_k \subset A.$$

[*Hint.* Remember that $\mathbb{R}^n$ has a denumerable basis.]

2.10 ¶. Using Theorem 2.7, prove the following, compare Proposition 9.126 and Theorem 9.127 of [GM3].

Proposition. *Let $C \subset \mathbb{R}^n$ be an open convex subset and let $\overline{x} \notin C$. Then there exists a linear map $\ell : \mathbb{R}^n \to \mathbb{R}$ such that $\ell(x) < \ell(\overline{x})$ $\forall x \in C$. In particular, C and $\overline{x}$ are separated by the hyperplane $\{x \,|\, \ell(x) = \ell(\overline{x})\}$.*

Consequently,

Theorem. *Let A and B be two nonempty disjoint convex sets. Suppose A is open. Then A and B can be separated by a hyperplane.*

2.11 Definition. *We say that K is* polyhedral *if it is the intersection of a finite number of closed half-spaces. A bounded polyhedral set is called a* polyhedron.

c. Convex hull

2.12 Definition. *The* convex hull *of a set $M \subset \mathbb{R}^n$, $\mathrm{co}(M)$, is the intersection of all convex subsets in $\mathbb{R}^n$ that contain M.*

2.13 Proposition. *The convex hull of $M \subset \mathbb{R}^n$ is convex, indeed the smallest convex set that contains M. Moreover, $\mathrm{co}(M)$ is the set of all convex combinations of points of M,*

$$\mathrm{co}(M) := \Big\{ x \in \mathbb{R}^n \,\Big|\, \exists x_1, x_2, \dots, x_N \in M \text{ such that } x = \sum_{i=1}^N \lambda_i x_i, \\ \text{for some } \lambda_1, \lambda_2, \dots, \lambda_N, \text{ where } \lambda_i \geq 0 \ \forall i, \ \sum_{i=1}^N \lambda_i = 1 \Big\}.$$

2.14 ¶. Prove Proposition 2.13.

2.15 ¶. Prove that
(i) $\mathrm{co}(M)$ is open, if M is open,
(ii) $\mathrm{co}(M)$ is compact, if M is compact.

2.16 ¶. Give examples of sets $M \subset \mathbb{R}^2$ so that
(i) M is closed but $\mathrm{co}(M)$ is not,
(ii) $\mathrm{co}(\overline{M}) \neq \overline{\mathrm{co}(M)}$ although $\mathrm{co}(\overline{M}) \subset \overline{\mathrm{co}(M)}$.

If $M \subset \mathbb{R}^n$, then the convex combinations of at most $n+1$ points in M are sufficient to describe $\mathrm{co}(M)$. In fact, the following holds.

2.17 Theorem (Carathéodory). *Let $M \subset \mathbb{R}^n$. Then*

$$\mathrm{co}(M) := \Big\{ x \in \mathbb{R}^n \,\Big|\, x = \sum_{i=1}^{n+1} \lambda_i x_i, \ x_i \in M, \ \lambda_i \geq 0 \ \forall i, \ \sum_{i=1}^{n+1} \lambda_i = 1 \Big\}.$$

Proof. Let x be a convex combination of m points $x_1, x_2, \dots, x_m$ of M with $m > n+1$,

$$x = \sum_{j=1}^{m} \lambda_j x_j, \qquad \sum_{j=1}^{m} \lambda_j = 1, \ \lambda_j > 0.$$

We want to show that x can be written as convex combinations of $m-1$ points of M.

Since $m - 1 > n$, there are numbers $c_1, c_2, \dots, c_{m-1}$ not all zero such that $\sum_{i=1}^{m-1} c_i(x_i - x_m) = 0$. If $c_m := -\sum_{i=1}^{m-1} c_i$, we have

$$\sum_{i=1}^{m} c_i x_i = 0 \qquad \text{and} \qquad \sum_{i=1}^{m} c_i = 0.$$

Since at least one of the c_i's is positive, we can find $t > 0$ and $k \in \{1, \dots, m\}$ such that

$$\frac{1}{t} = \max\left(\frac{c_1}{\lambda_1}, \frac{c_2}{\lambda_2}, \dots, \frac{c_m}{\lambda_m}\right) = \frac{c_k}{\lambda_k} > 0.$$

The point x is then a convex combination of $x_1, x_2, \dots, x_{k-1}, x_{k+1}, \dots, x_m$; in fact, if

$$\gamma_j := \begin{cases} \lambda_j - tc_j & \text{if } j \neq k, \\ 0 & \text{if } j = k, \end{cases}$$

we have $\sum_{j\neq k} \gamma_j = \sum_{j=1}^m \gamma_j = \sum_{j=1}^m (\lambda_j - tc_j) = \sum_{j=1}^m \lambda_j = 1$ and

$$x = \sum_{j=1}^{m} \lambda_j x_j = \sum_{j=1}^{m} (\gamma_j + tc_j) x_j = \sum_{j\neq k} \gamma_j x_j.$$

We then conclude by backward induction on m. □

2.18 ¶. Prove the following:

(i) In Theorem 2.17 the number $n+1$ is optimal.

(ii) If M is convex, then $\mathrm{co}(M) = M$ and every point in $\mathrm{co}(M)$ is a convex combination of itself.

(iii) If $M = M_1 \cup \dots \cup M_k$, $k \leq n$, where $M_1, \dots, M_k$ are convex sets, then every point of $\mathrm{co}(M)$ is a convex combination of at most k points of M.

d. The distance function from a convex set

We conclude with a characterization of a convex set in terms of its distance function.

Let $C \subset \mathbb{R}^n$ be a nonempty closed set. For every $x \in \mathbb{R}^n$ we define

$$d_C(x) := \mathrm{dist}(x, C) := \inf\Big\{|x-y| \,\Big|\, y \in C\Big\}.$$

It is easily seen that indeed the infimum is a minimum, i.e., there is (at least) a point $y \in C$ of least distance from x. Moreover, the function d_C is Lipschitz-continuous with Lipschitz constant 1,

$$|d_C(x) - d_C(y)| \leq |x - y| \qquad \forall x, y \in \mathbb{R}^n.$$

2.19 Lemma. *If $x \notin C$, then*

$$d_C(x+h) = d_C(x) + L(h;x) + o(|h|) \qquad \text{as } h \to 0, \tag{2.2}$$

where

$$L(h;x) := \min\Big\{ h \bullet \frac{x-z}{|x-z|} \;\Big|\; z \in C, |x-z| = d_C(x)\Big\}$$

is the minimum among the lengths of the projections of h into the lines connecting x to its nearest points $z \in C$. In particular, d_C is differentiable at x if and only if $h \to L(h;x)$ is linear, i.e., if and only if there is a unique minimum point $z \in C$ of least distance from x.

Proof. We prove (2.2), the rest easily follows. We may and do assume that $x = 0$. Moreover, we deal with the function

$$f(h) := d_C^2(h) = \min_{z\in C} |h-z|^2.$$

It suffices to show that

$$f(h) = f(0) + f'(h,0) + o(|h|), \qquad h \to 0, \tag{2.3}$$

where

$$f'(h;0) := \min\Big\{ -2\,h \bullet z \;\Big|\; z \in C, |z| = d_C(0)\Big\}.$$

First, we remark that the functions $q_\epsilon(h)$ defined for $\epsilon \geq 0$ as

$$q_\epsilon(h) := \inf\Big\{ -2\,h \bullet z \;\Big|\; |z| \leq f(0)^{1/2} + \epsilon\Big\}$$

are homogeneous of degree 1 and that $q_\epsilon \to q_0$ increasingly as $\epsilon \to 0$. By Dini's theorem, see [GM3], $\{q_\epsilon\}$ converges uniformly to q_0 in $B(0,1)$. Therefore, for every $\epsilon > 0$ there is c_ϵ such that

$$q_0(h) \geq q_\epsilon(h) \geq q_0(h) - c_\epsilon |h| \qquad \forall h \tag{2.4}$$

and $c_\epsilon \to 0$ as $\epsilon \to 0$.

Now, let us prove (2.3). Since $|y-z|^2 = |z|^2 - 2\,y \bullet z + |y|^2$, we have

$$\begin{aligned} f(h) &\leq \min_{\substack{z\in C \\ |z|=d_C(0)}} |h-z|^2 = |h|^2 + f(0) + \min_{\substack{z\in C \\ |z|=d_C(0)}} (-2\,h \bullet z) \\ &= f(0) + q_0(h) + |h|^2. \end{aligned} \tag{2.5}$$

On the other hand, if $|h| < \epsilon/2$, the minimum of $z \to |z-h|^2$, $z \in C$, is attained at points z_h such that $|z_h| \leq f(0)^{1/2} + \epsilon/2$, hence by (2.4)

$$\begin{aligned} f(h) = \min_{z\in C} |z-h|^2 &= \min_{\substack{z\in C \\ |z|<f(0)^{1/2}+\epsilon}} |z-h|^2 \\ &= \min_{\substack{z\in C \\ |z|<f(0)^{1/2}+\epsilon}} \Big(|z|^2 + |h|^2 - 2\,h \bullet z \Big) \\ &\geq f(0) + |h|^2 + q_\epsilon(h) \geq f(0) + q_0(h) - c_\epsilon |h| + |h|^2. \end{aligned}$$

Therefore

$$f(h) \geq f(0) + q_0(h) + o(|h|) \qquad \text{as } h \to 0,$$

which, together with (2.5), proves (2.3). $\square$

2.20 ¶. Using (2.2), prove that, in general, if there are in C more than one nearest point to x, then

$$\lim_{t\to 0^\pm} \frac{d_C(x+th) - d_C(x)}{t} = \min\Big\{ h \bullet \frac{x-z}{|x-z|} \;\Big|\; z \in C, |x-z| = d_C(x)\Big\}.$$

2.21 Theorem (Motzkin). *Let $C \subset \mathbb{R}^n$ be a nonempty closed set. The following claims are equivalent:*

(i) *C is convex.*
(ii) *For all $x \notin C$ there is a unique nearest point in C to x.*
(iii) *d_C is differentiable at every point in $\mathbb{R}^n \setminus C$.*

Proof. The equivalence of (ii) and (iii) is the content of Lemma 2.19.

(i) $\Rightarrow$ (ii). If z is the nearest point in C to $x \notin C$, then $x - z - \epsilon(y-z) \in C$ if $y \in C$, therefore

$$|x-z|^2 \le |x - z - \epsilon(y-z)|^2 = |x-z|^2 - 2\,\epsilon\,(y-z) \bullet (x-z) + \epsilon^2 |y-z|^2 \tag{2.6}$$

for all $0 \le \epsilon \le 1$. For $\epsilon \to 0$ we get $(x-z) \bullet (y-z) \le 0$ and, because of (2.6) with $\epsilon = 1$

$$|x-y|^2 = |x-z|^2 - 2\,(x-z) \bullet (y-z) + |y-z|^2 > |x-z|^2 \qquad \forall y \in C.$$

(ii) $\Rightarrow$ (i). Suppose C is not convex. It suffices to show that there is a ball B such that $B \cap C = \emptyset$ and $\overline{B} \cap C$ has more than a point. Since C is not convex, there exist $x_1, x_2 \in C$, $x_1 \ne x_2$, such that the open segment connecting x_1 to x_2 is contained in $\mathbb{R}^n \setminus C$. We may suppose that the middle point of this segment is the origin, i.e., $x_2 = -x_1$, and let ρ be such that $\overline{B(0,\rho)} \cap C = \emptyset$. We now consider the family of balls $\{B(w,r)\}$ such that

$$B(w,r) \supset B(0,\rho), \qquad B(w,r) \cap C = \emptyset \tag{2.7}$$

and claim that the corresponding set $\{(w,r)\} \subset \mathbb{R}^{n+1}$ is bounded and closed, hence compact. In fact, since $x_j \notin B(w,r)$, $j = 1,2$, we have $r \ge |w| + \rho$ and $|w \pm x_1|^2 \ge r^2$, hence

$$(|w| + \rho)^2 \le r^2 \le \frac{1}{2}(|w+x_1|^2 + |w-x_1|^2) \le |w|^2 + r^2$$

from which we infer

$$|w| \le \frac{|x_1|^2 - \rho^2}{2\rho}, \qquad r \le \frac{(|x_1|^2 + \rho^2)}{2\rho}.$$

Consider now a ball $B(w_0, r_0)$ with maximal radius r_0 among the family (2.7). We claim that $\partial B(w_0, r_0) \cap C$ contains at least two points. Assuming on the contrary that $\partial B(w_0, r_0) \cap C$ contains only one point y_1, for all θ such that $\theta \bullet (y_1 - w_0) < 0$ and for all $\epsilon > 0$ sufficiently small, $\overline{B(w_0 + \theta\epsilon, r_0)} \cap C = \emptyset$, consequently, by maximality there exists y_ϵ such that

$$y_\epsilon \in \partial B(w_0 + \epsilon\theta, r_0) \cap \partial B(0,\rho). \tag{2.8}$$

From (2.8) we infer, as $\epsilon \to 0$, that there is a point $y_2 \in \partial B(w_0, r_0) \cap \partial B(0,\rho)$, which is unique since $r_0 > \rho$. However, if we choose $\overline{\theta} := y_2 - y_1$, we surely have $\partial B(w_0 + \epsilon\overline{\theta}, r_0) \cap \partial B(0,\rho) = \emptyset$, for sufficiently small ϵ. This contradicts (2.8). □

e. Extreme points

2.22 Definition. *Let $K \subset \mathbb{R}^n$ be a nonempty convex set. A point $x_0 \in K$ is said to be an* extreme point *for K if there are no $x_1, x_2 \in K$ and $\lambda \in]0,1[$ such that $x_0 = \lambda x_1 + (1-\lambda)x_2$.*

The extreme points of a cube are the vertices; the extreme points of a ball are all its boundary points. The extreme points of a set, if any, are boundary points; in particular, an open convex set has no extreme points. Additionally, a closed half-space has no extreme points.

2.23 Theorem. *Let $K \subset \mathbb{R}^n$ be nonempty closed and convex.*

(i) *If K does not contains lines, then K has extreme points.*
(ii) *If K is compact, then K is the convex hull of its extreme points.*

Proof. Let us prove (ii) by induction on the dimension of the smallest affine subspace containing K. We leave then to the reader the task of proving (i), still by induction. If $n = 1$, K is a segment and the claim is trivial. Suppose that the claim holds for convex sets contained in an affine subspace of dimension $n-1$. For $x_0 \in \partial K$, let $\mathcal{P}$ be a supporting hyperplane to K at x_0. The set $K \cap \mathcal{P}$ is compact and convex, hence by the inductive assumption, x_0 is a convex combination of extreme points of $K \cap \mathcal{P}$, that are also extreme points of K. If x_0 is an interior point of K, every line through x_0 cuts K into a segment of extremes x_1 and $x_2 \in \partial K$, hence x_0 is a convex combination of extreme points, since so are $x_1, x_2 \in \partial K$. □

2.2 Proper Convex Functions

a. Definitions

We have already introduced convex functions of one variable, discussed their properties and illustrated a few estimates related to the notion of convexity, see [GM1] and Section 5.3.7 of [GM4]. Here we shall discuss convex functions of several variables.

2.24 Definition. *A function $f : K \subset \mathbb{R}^n \to \mathbb{R}$ defined on a convex set K, is said to be* convex *in K if*

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y) \qquad \forall x, y \in K, \quad \forall \lambda \in [0,1]. \tag{2.9}$$

The function f is said to be strictly convex *if the inequality in (2.9) for $x \neq y$ and $0 < \lambda < 1$ is strict.*

We say that $f : K \to \mathbb{R}$ is concave *if K is convex and $-f : K \to \mathbb{R}$ is convex.*

The convexity of K is needed to ensure that the segment $\{z \in \mathbb{R}^n \,|\, z = \lambda x + (1-\lambda)y,\ \lambda \in [0,1]\}$ belongs to the domain of definition of f. The geometric meaning of the definition is clear: The segment PQ connecting the point $P = (x, f(x))$ to $Q = (y, f(y))$ lies above the graph of the restriction of f to the segment with extreme points $x, y \in K$.

2.25 ¶. Prove the following.

(i) Linear functions are both convex and concave; in fact, they are the only functions that are at the same time convex and concave.
(ii) If f and g are convex, then $f+g, \alpha f,\ \alpha > 0$, $\max(f,g)$ and $\lambda f+(1-\lambda)g$, $\lambda \in [0,1]$, are convex.
(iii) If $f : K \to \mathbb{R}$ is convex and $g : I \supset f(K) \to \mathbb{R}$ is convex and not decreasing, then $g \circ f$ is convex.
(iv) The functions $|x|^p$, $(1+|x|^2)^{p/2}$, $p \geq 1$, $e^{\theta|x|}$, $\theta > 0$, and $x \log x - x$, $x > 0$, are convex.

b. A few characterizations of convexity

We recall that the *epigraph* of a function $f : A \subset \mathbb{R}^n \to \mathbb{R}$ is the subset of $\mathbb{R}^n \times \mathbb{R}$ given by

$$\mathrm{Epi}(f) := \Big\{(x,z)\ \Big|\ x \in A,\ z \in \mathbb{R},\ z \geq f(x)\Big\}.$$

2.26 Proposition. *Let $f : K \subset \mathbb{R}^n \to \mathbb{R}$. The following claims are equivalent:*

(i) *K is convex, and $f : K \to \mathbb{R}$ is convex.*
(ii) *The epigraph of f is a convex set in $\mathbb{R}^{n+1}$.*
(iii) *For every $x_1, x_2 \in K$ the function $\varphi(\lambda) := f(\lambda x_1 + (1-\lambda)x_2)$, $\lambda \in [0,1]$, is well-defined and convex.*
(iv) (Jensen's inequality) *K is convex and for any choice of m points $x_1, x_2, \dots, x_m \in K$, and nonnegative numbers $\alpha_1, \alpha_2, \dots, \alpha_m$ such that $\sum_{i=1}^m \alpha_i = 1$, we have*

$$f\Big(\sum_{i=1}^m \alpha_i x_i\Big) \leq \sum_{i=1}^m \alpha_i f(x_i).$$

Proof. (i) $\Longrightarrow$ (ii) follows at once from the definition of convexity.

(ii) $\Longrightarrow$ (i). Let $\pi : \mathbb{R}^{n+1} \to \mathbb{R}^n$ be the projection map into the first factor, $\pi((x,t)) := x$. Since linear maps map convex sets into convex sets and $K = \pi(\mathrm{Epi}(f))$, we infer that K is a convex set, while the convexity of f follows just by definition.

(i)$\Rightarrow$(iii). For $\lambda, t, s \in [0,1]$ we have

$$\begin{aligned}\varphi(\lambda t + (1-\lambda)s) &= f\Big([\lambda t + (1-\lambda)s]x_1 + [1 - \lambda t - (1-\lambda)s]x_2\Big)\\ &= f\Big(\lambda[tx_1 + (1-t)x_2] + (1-\lambda)[sx_1 + (1-s)x_2]\Big)\\ &\leq \lambda\varphi(t) + (1-\lambda)\varphi(s).\end{aligned}$$

(iii)$\Rightarrow$(i). We have

$$\begin{aligned}f(\lambda x_1 + (1-\lambda)x_2) &= \varphi(\lambda) = \varphi(\lambda \cdot 1 + (1-\lambda)\cdot 0)\\ &\leq \lambda\varphi(1) + (1-\lambda)\varphi(0) = \lambda f(x_1) + (1-\lambda)f(x_2).\end{aligned}$$

(iv)$\Rightarrow$(i). Trivial.

(i)⇒(iv). We proceed by induction on m. If $m = 1$, the claim is trivial. For $m > 1$, let $\alpha := \alpha_1 + \cdots + \alpha_{m-1}$, so that $\alpha_m = 1 - \alpha$. We have

$$\sum_{i=1}^{m} \alpha_i x_i = \alpha \sum_{i=1}^{m-1} \frac{\alpha_i}{\alpha} x_i + (1-\alpha) x_m,$$

with $0 \le \alpha_i/\alpha \le 1$ and $\sum_{i=1}^{m-1} (\alpha_i/\alpha) = 1$. Therefore we conclude, using the inductive assumption, that

$$\begin{aligned} f\Big(\sum_{i=1}^{m} \alpha_i x_i\Big) &\le \alpha f\Big(\sum_{i=1}^{m-1} \frac{\alpha_i}{\alpha} x_i\Big) + (1-\alpha) f(x_m) \\ &\le \alpha \sum_{i=1}^{m-1} \frac{\alpha_i}{\alpha} f(x_i) + (1-\alpha) f(x_m) = \sum_{i=1}^{m} \alpha_i x_i. \end{aligned}$$

□

From (ii) of Proposition 2.26 and Carathéodory's theorem, Theorem 2.17, we infer at once the following.

2.27 Corollary. *Let $K \subset \mathbb{R}^n$ be a convex set. The function $f : K \subset \mathbb{R}^n \to \mathbb{R}$ is convex if and only if*

$$f(x) := \inf \Big\{ \sum_{i=1}^{n+1} \lambda_i f(x_i) \,\Big|\, \forall x_1, x_2, \dots, x_{n+1} \in K \text{ such that } x = \sum_{i=1}^{n+1} \lambda_i x_i, \\ \text{with } \lambda_i \ge 0,\ \sum_{i=1}^{n+1} \lambda_i = 1 \Big\}.$$

Of course, the level sets $\{x \in K \,|\, f(x) \le c\}$ and $\{x \in K \,|\, f(x) < c\}$ of a convex function $f : K \to \mathbb{R}$ are convex sets; however, there exist nonconvex functions whose level sets are convex; for instance, the function x^3, $x \in \mathbb{R}$, or, more generally, the composition of a convex function $f : K \to \mathbb{R}$ with a monotone function $\varphi : \mathbb{R} \to \mathbb{R}$.

2.28 Definition. *A function with convex level sets is called a* quasiconvex *function.*[1]

c. Support function

Let $f : K \subset \mathbb{R}^n \to \mathbb{R}$ be a function. We say that a linear function $\ell : \mathbb{R}^n \to \mathbb{R}$ is a *support function* for f at $x \in K$ if

$$f(y) \ge f(x) + \ell(y - x) \qquad \forall y \in K.$$

[1] We notice that "quasiconvex" is used with different meanings in different contexts.

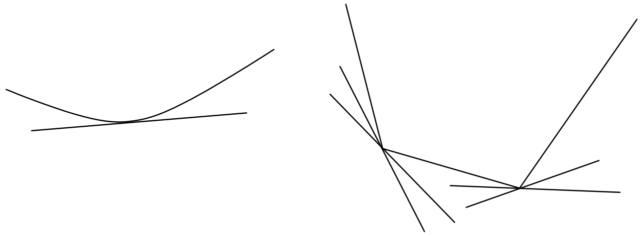


Figure 2.5. Convex functions and supporting affine hyperplanes.

2.29 Definition. *Let $f : K \to \mathbb{R}$ be a convex function. The set of linear maps $\ell : \mathbb{R}^n \to \mathbb{R}$ such that $y \to f(x) + \ell(y-x)$ is a support function for f at x is called the* subdifferential *of f at x and denoted by $\partial f(x)$.*

Trivially, if $\ell \in \partial f(x)$, then the graph of $y \mapsto f(x)+\ell(y-x)$ at $(x, f(x))$ is a supporting hyperplane for the epigraph of f at $(x, f(x))$. Conversely, on account of Proposition 2.30, every affine supporting hyperplane to $\mathrm{Epi}(f)$ is the graph of a linear map belonging to the subdifferential to f at x provided it contains no vertical vectors. This is the case if f is convex on an open set, as shown by the following proposition.

2.30 Proposition. *Let $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ be a function, where Ω is convex and open. Then f is convex if and only if for every $x \in \Omega$ there is a linear support function for f at x.*

Proof. Let f be convex and $\overline{x} \in \Omega$. The epigraph of f is convex and its closure is convex; moreover, $(\overline{x}, f(\overline{x})) \in \partial\,\mathrm{Epi}(f)$. Consequently, there is a supporting hyperplane $\mathcal{P}$ of $\mathrm{Epi}(f)$ at $(\overline{x}, f(\overline{x}))$ that does not contain vertical vectors, otherwise it would divide Ω in two parts and, as a consequence, the epigraph of f. We then conclude that there exist a linear map $\varphi : \mathbb{R}^n \to \mathbb{R}$ and constants $\alpha, \beta \in \mathbb{R}$ such that $\mathcal{P} = \{(x,y)\,|\,\varphi(x)+\alpha y = \beta\}$ and

$$\varphi(x-\overline{x}) + \alpha(y - f(\overline{x})) \geq 0 \ \forall (x,y) \in \mathrm{Epi}(f), \qquad \alpha \neq 0. \tag{2.10}$$

Moreover, we have $\alpha \geq 0$ since in (2.10) we can choose y arbitrarily large. Thus, $\alpha > 0$ and, if we set $\ell(x) := -\varphi(x)/\alpha$, from (2.10) with $y = f(x)$, we infer

$$f(x) \geq f(\overline{x}) + \ell(x-\overline{x}) \qquad \forall x \in \Omega.$$

Conversely, let us prove that $f : \Omega \to \mathbb{R}$ is convex if it has at every point a linear support function. Let $x_1, x_2 \in \Omega$, $x_1 \neq x_2$, and $\lambda \in]0,1[$, set $x_0 := \lambda x_1 + (1-\lambda)x_2$, $h := x_1 - x_0$, so that $x_2 = x_0 - \frac{\lambda}{1-\lambda}h$. Let ℓ be the linear support function for f at x_0. We have

$$f(x_1) \geq f(x_0) + \ell(h), \qquad f(x_2) \geq f(x_0) - \frac{\lambda}{1-\lambda}\ell(h).$$

Multiplying the first inequality by $\lambda/(1-\lambda)$ and summing to the second, we get

$$\frac{\lambda}{1-\lambda} f(x_1) + f(x_2) \geq \left(\frac{\lambda}{1-\lambda} + 1\right) f(x_0),$$

i.e., $f(x_0) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$. $\square$

2.31 Remark. A consequence of the above is the following claim that complements Jensen's inequality. With the same notation of Proposition 2.26, *if f is* strictly convex *and $\alpha_i > 0$ $\forall i$, then the equality*

$$f\Big(\sum_{i=1}^m \alpha_i x_i\Big) = \sum_{i=1}^m \alpha_i f(x_i) \tag{2.11}$$

implies that $x_j = x_0$ $\forall j = 1, \dots, m$ where $x_0 := \sum_{i=1}^m \alpha_i x_i$. In fact, if $\ell(x) := f(x_0) + m \bullet (x - x_0)$ is a linear affine support function for f at x_0, the function

$$\psi(x) := f(x) - f(x_0) - m \bullet (x - x_0)\,, \qquad x \in K$$

is nonnegative and, because of (2.11),

$$\sum_{i=1}^m \psi(x_i) = 0.$$

Hence $\psi(x_j) = 0$ $\forall j = 1 \dots, m$. Since ψ is strictly convex, we conclude that $x_j = x_0$ $\forall j = 1, \dots, m$.

d. Convex functions of class C^1 and C^2

We now present characterizations of smooth convex function in an open set.

2.32 Theorem. *Let Ω be an open and convex set in $\mathbb{R}^n$ and let $f : \Omega \to \mathbb{R}$ be a function of class C^1. The following claims are equivalent:*

(i) *f is convex.*
(ii) *For all $x_0 \in \Omega$, the graph of f lies above the tangent plane to the graph of f at $(x_0, f(x_0))$,*

$$f(x) \geq f(x_0) + \nabla f x_0 \bullet (x - x_0) \qquad \forall x_0, x \in \Omega. \tag{2.12}$$

(iii) *The differential of f is a* monotone operator*, i.e.,*

$$\Big(\nabla f(y) - \nabla f(x)\Big) \bullet (y - x) \geq 0 \qquad \forall x, y \in \Omega. \tag{2.13}$$

Notice that in one dimension the fact that ∇f is monotone means simply that f' is increasing. Actually, we could deduce Theorem 2.32 from the analogous theorem in one dimension, see [GM1], but we prefer giving a self-contained proof.

Proof. (i)$\Longrightarrow$(ii). Let $x_0, x \in \Omega$ and $h := x - x_0$. The function $t \mapsto f(x_0 + th)$, $t \in [0,1]$, is convex, hence $f(x_0 + th) \le tf(x_0 + h) + (1-t)f(x_0)$, i.e.,

$$f(x_0 + th) - f(x_0) \le t[f(x_0 + h) - f(x_0)].$$

We infer

$$\frac{f(x_0 + th) - f(x_0)}{t} - \nabla f(x_0) \bullet h \le f(x_0 + h) - f(x_0) - \nabla f(x_0) \bullet h\,.$$

Since for $t \to 0^+$ the left-hand side converges to zero, we conclude that the right-hand side, which is independent from t, is nonnegative.

(ii)$\Longrightarrow$ (i). Let us repeat the argument in the proof of Proposition 2.30. For $x \in \Omega$ the map $h \to f(x) + \nabla f(x) \bullet h$ is a support function for f at x. Let $x_1, x_2 \in \Omega$, $x_1 \neq x_2$, and let $\lambda \in]0,1[$. We set $x_0 := \lambda x_1 + (1-\lambda)x_2$, $h := x_1 - x_0$, hence $x_2 = x_0 - \frac{\lambda}{1-\lambda}h$. From (2.12) we infer

$$f(x_1) \ge f(x_0) + \nabla f(x_0) \bullet h\,, \qquad f(x_2) \ge f(x_0) - \frac{\lambda}{1-\lambda} \nabla f(x_0) \bullet h\,.$$

Multiplying the first inequality by $\lambda/(1-\lambda)$ and summing to the second we get

$$\frac{\lambda}{1-\lambda} f(x_1) + f(x_2) \ge \left(\frac{\lambda}{1-\lambda} + 1\right) f(x_0),$$

i.e., $f(x_0) \le \lambda f(x_1) + (1-\lambda) f(x_2)$.

(ii)$\Rightarrow$(iii). Trivially, (2.12) yields

$$f(x) - f(y) \le \nabla f(x) \bullet (x-y)\,, \qquad f(x) - f(y) \ge \nabla f(y) \bullet (x-y)\,,$$

hence

$$\nabla f(y) \bullet (x-y) \le f(x) - f(y) \le \nabla f(x) \bullet (x-y)\,,$$

i.e., (2.13).

(iii)$\Rightarrow$(ii). Assume now that (2.13). For $x_0, x \in \Omega$ we have

$$f(x) - f(x_0) = \int_0^1 \frac{d}{dt} f(tx + (1-t)x_0)\, dt = \left(\int_0^1 \nabla f(tx + (1-t)y)\, dt\right) \bullet (x - x_0)$$

and

$$\Big(\nabla f(tx + (1-t)y)\Big) \bullet (x - x_0) \ge \nabla f(x_0) \bullet (x - x_0)\,,$$

hence

$$f(x) - f(x_0) \ge \left(\int_0^1 \nabla f(x_0)\, dt\right) \bullet (x - x_0) = \nabla f(x_0) \bullet (x - x_0)\,.$$

□

Let f belong to $C^2(\Omega)$. Because of (iii) of Proposition 2.26, $f : \Omega \to \mathbb{R}$ is convex if and only if for every $x_1, x_2 \in \Omega$ the function

$$\varphi(\lambda) := f((1-\lambda)x_1 + \lambda x_2) \qquad \lambda \in [0,1]$$

is convex and $C^2([0,1])$. By Theorem 2.32 φ is convex if and only if φ' is increasing in $[0,1]$, i.e., if and only if $\varphi'' \ge 0$. Since

$$\varphi''(0) = \Big(\mathbf{H} f(x_1)(x_2 - x_1)\Big) \bullet (x_2 - x_1)\,,$$

we conclude the following.

2.33 Theorem. *Let $\Omega \subset \mathbb{R}^n$ be an open and convex set of $\mathbb{R}^n$ and let $f : \Omega \to \mathbb{R}$ be a function of class $C^2(\Omega)$. Then f is convex if and only if the Hessian matrix of f is nonnegative at every point in Ω,*

$$\mathbf{H}f(x)h \bullet h \geq 0 \qquad \forall x \in \Omega, \quad \forall h \in \mathbb{R}^n.$$

Similarly, one can prove that *f is strictly convex if the Hessian matrix of f is positive at every point in Ω.*

Notice that $f(x) = x^4$, $x \in \mathbb{R}$, is strictly convex, but $\mathbf{H}f(0) = 0$.

2.34 ¶. Let $f : K \subset \mathbb{R}^n \to \mathbb{R}$ be a convex function, K being convex and bounded. Prove the following:

(i) In general, f has no maximum points.

(ii) If f is not constant, then f has no interior maximum point; in other words, if f is not constant, then

$$f(x) < \sup_{y \in K} f(y) \qquad \forall x \in \operatorname{int}(K);$$

possible maximum points lie on ∂K if K is closed.

(iii) if K has extremal points, possible maximum points lie on the extremal points of K; in the case that K has finite many extremal points, then f has a maximum point and

$$\max_{x \in K} f(x) = \max_{i=1,N} f(x_i).$$

(iv) In general, f has no minimum points.

(v) The set of minimum points is convex and reduces to a point if f is strictly convex.

(vi) Local minimum points are global minimum points.

In particular, from (iii) it follows that if $f : Q \to \mathbb{R}$ is convex, Q being a closed cube in $\mathbb{R}^n$, then f has maximum and the maximum points lie on the vertices of Q.

e. Lipschitz continuity of convex functions

Let $f : Q \subset \mathbb{R}^n \to \mathbb{R}$ be a convex function defined on a closed cube Q. Then it is easy to see that $f(x) \leq \sup_{\partial Q} f$ for every $x \in Q$. Moreover, one sees by downward induction that f has maximum and the maximum points lie on the vertices of Q, see Exercise 2.34.

2.35 Theorem. *Let $\Omega \subset \mathbb{R}^n$ be an open and convex set and let $f : \Omega \to \mathbb{R}$ be convex. Then f is locally Lipschitz in Ω.*

Proof. Let $x_0 \in \Omega$ and let $Q(x_0, r)$ be a sufficiently small closed cube contained in Ω with sides of length $2r$ parallel to the axes. Since f is convex, $f_{|Q(x_0,r)}$ has maximum value at one of the vertices of $Q(x_0, r)$. If

$$L_r := \sup_{x \in \partial B(x_0,r)} f(x),$$

then $L_r < +\infty$ since $\partial B(x_0, r) \subset Q(x_0, r)$. Let us prove that

$$|f(x) - f(x_0)| \leq \frac{L_r - f(x_0)}{r} |x - x_0| \qquad \forall x \in B(x_0, r). \tag{2.14}$$

Without loss in generality, we may assume $x_0 = 0$ and $f(0) = 0$. Let $x \neq 0$ and let $x_1 := \frac{r}{|x|}x$ and $x_2 := -\frac{r}{|x|}x$. Since $x_1 \in \partial B(x_0, r)$ and $x = \lambda x_1 + (1 - \lambda)0$, $\lambda := |x|/r$, the convexity of f yields

$$f(x) \le \frac{|x|}{r} f(x_1) \le \frac{L_r}{r}|x|,$$

whereas, since $x_2 \in \partial B(x_0, r)$ and $0 = \lambda x + (1-\lambda)x_2$, $\lambda := 1/(1+|x|/r)$, we have $0 = f(0) \le \lambda f(x) + (1-\lambda) f(x_2) \le (1-\lambda)L_r$, i.e.,

$$-f(x) \le \frac{1-\lambda}{\lambda} L_r = \frac{L_r}{r}|x|.$$

Therefore, $|f(x)| \le (L_r/r)|x|$ for all $x \in B(0, r)$, and (2.14) is proved.

In particular, (2.14) tells that f *is continuous in* Ω.

Let K and K_1 be two compact sets in Ω with $K \subset\subset K_1 \subset \Omega$ and let $\delta := \operatorname{dist}(K, \partial K_1) > 0$. Let M_1 denote the oscillation of f in K_1,

$$M_1 := \sup_{x,y \in K_1} |f(x) - f(y)|,$$

which is finite by the Weierstrass theorem. For every $x_0 \in K$, the cube centered at x_0 with sides parallel to the axes of length $2r$, $r = \delta/\sqrt{n}$, is contained in K_1. It follows from (2.14) that

$$|f(x) - f(x_0)| \le \frac{L_r - f(x_0)}{r}|x - x_0| \le \frac{M_1}{r}|x - x_0| \qquad \forall x \in K \cap B(x_0, r).$$

On the other hand, for $x \in K \setminus B(x_0, r)$ we have $|x - x_0| \ge r$, hence

$$|f(x) - f(x_0)| \le M_1 \le \frac{M_1}{r}|x - x_0|.$$

In conclusion, for every $x \in K$

$$|f(x) - f(x_0)| \le \frac{M_1}{r}|x - x_0|$$

and, x_0 being arbitrary in K (and M_1 and r independent from r and x_0), we conclude that f is Lipschitz-continuous in K with Lipschitz constant smaller than M_1/r. □

Actually, the above argument shows more: *A locally equibounded family of convex functions is also locally equi-Lipschitz.*

f. Supporting planes and differentiability

2.36 Theorem. *Let $\Omega \subset \mathbb{R}^n$ be open and convex and let $f : \Omega \to \mathbb{R}$ be convex. Then f has a unique support function at x_0 if and only if f is differentiable at x_0.*

In this case, of course, the supporting function is the linear tangent map to f at x_0,

$$y \mapsto \nabla f(x_0) \bullet y\,.$$

As a first step, we prove the following proposition.

2.37 Proposition. *Let $\Omega \subset \mathbb{R}^n$ be open and convex, let $f : \Omega \to \mathbb{R}$ be convex and let $x_0 \in \Omega$. For every $v \in \mathbb{R}^n$ the right and left derivatives defined by*

$$\frac{\partial f}{\partial v^+}(x) := \lim_{t \to 0^+} \frac{f(x+tv) - f(v)}{t},$$

$$\frac{\partial f}{\partial v^-}(x) := \lim_{t \to 0^-} \frac{f(x+tv) - f(v)}{t},$$

exist and $\frac{\partial f}{\partial v^-}(x_0) \le \frac{\partial f}{\partial v^+}(x_0)$. *Moreover, for any* $m \in \mathbb{R}$ *such that* $\frac{\partial f}{\partial v^-}(x) \le m \le \frac{\partial f}{\partial v^-}(x)$, *there exists a linear map* $\ell : \mathbb{R}^n \to \mathbb{R}$ *such that* $f(x) \ge f(x_0) + \ell(x - x_0)$ $\forall x \in \Omega$ *and* $\ell(v) = m$.

Proof. Without loss in generality we assume $x_0 = 0$ and $f(0) = 0$.

The function $\varphi(t) := f(tv)$ is convex in an interval around zero; thus, compare [GM1], φ has right-derivative $\varphi'_+(0)$ and left-derivative $\varphi'_-(0)$ and $\phi'_-(0) \le \varphi'_+(0)$. Since $\frac{\partial f}{\partial v^-}(0) = \varphi'_-(0)$ and $\frac{\partial f}{\partial v^+}(0) = \varphi'_+(0)$, the first part of the claim is proved.

(ii) If $\frac{\partial f}{\partial v^-}(0) \le m \le \frac{\partial f}{\partial v^+}(0)$, the graph of the linear map $t \to mt$ is a supporting line for $\mathrm{Epi}(\varphi)$ at $(0,0)$, i.e., for $\mathrm{Epi}(f) \cap V_0 \times \mathbb{R}$, $V_0 := \mathrm{Span}\{v\}$. We now show that the graph of the linear function $\ell_0 : V_0 \to \mathbb{R}$, $\ell_0(tv) := mt$, *extends to a supporting hyperplane to* $\mathrm{Epi}(f)$ *at* $(0, f(0))$, *which is in turn the graph of a linear map* $\ell : \mathbb{R}^n \to \mathbb{R}$.

Choose a vector $w \in \mathbb{R}^n$ with $w \notin V_0$, and remark that for $x, y \in V_0$ and $r, s > 0$ we have

$$\begin{aligned}\frac{r}{r+s}\ell_0(x) + \frac{s}{r+s}\ell_0(y) &= \ell_0\Big(\frac{r}{r+s}x + \frac{s}{r+s}y\Big)\\ &\le f\Big(\frac{r}{r+s}x + \frac{s}{r+s}y\Big) = f\Big(\frac{r}{r+s}(x - sw) + \frac{s}{r+s}(y + rw)\Big)\\ &\le \frac{r}{r+s}f(x - sw) + \frac{s}{r+s}f(y + rw);\end{aligned}$$

so that multiplying by $r + s$ we get

$$r\ell_0(x) + s\ell_0(y) \le rf(x - sw) + sf(x + rw),$$

i.e.,

$$g(x, s) := \frac{\ell_0(x) - f(x - sw)}{s} \le \frac{f(y + rw) - \ell_0(y)}{r} =: h(y, r).$$

For $\overline{x} \in V_0 \cap \Omega$ and $\overline{s}$ sufficiently small so that $\overline{x} + \overline{s}w$ and $\overline{x} - \overline{s}w$ ly in Ω, the values $g(\overline{x}, \overline{s})$ and $h(\overline{x}, \overline{s})$ are finite, hence

$$-\infty < g(\overline{x}, \overline{s}) \le \sup_{V_0 \times \mathbb{R}} g(x, s) \le \inf_{V_0 \times \mathbb{R}} h(x, s) \le h(\overline{x}, \overline{s}) < +\infty.$$

Consequently, there exists $\alpha \in \mathbb{R}$ such that

$$\frac{\ell_0(x) - f(x - sw)}{s} \le -\alpha \le \frac{f(x + rw) - \ell_0(x)}{r}$$

for all $x \in V_0$, $r, s \ge 0$ with $x - sw,\ x + rw \in \Omega$. This yields

$$\ell_0(x) + \alpha t \le f(x + tw) \qquad \forall x \in V_0,\ \forall t \in \mathbb{R} \text{ with } x + tw \in \Omega.$$

In conclusion, ℓ_0 has been extended to the linear function $\ell_1 : \mathrm{Span}\{v, w\} \to \mathbb{R}$ defined by $\ell_1(v) := \ell_0(v)$, $\ell_1(w) := \alpha$ for which $\ell_1(z) \le f(z)$ for all $z \in \mathrm{Span}\{v, w\}$. Of course, repeating the argument for finite many directions concludes the proof. □

Proof of Theorem 2.36. Without loss in generality, we assume $x_0 = 0$ and $f(0) = 0$.

Suppose that $\mathrm{Epi}(f)$ has a unique supporting hyperplane at 0. The restriction of f to any of the straight lines $\mathrm{Span}\, v$ through 0 has a unique support line since otherwise, as in Proposition 2.37, we could construct two different hyperplanes supporting $\mathrm{Epi}(f)$ at $(0,0)$. In particular, $\frac{\partial f}{\partial v^-}(0) = \frac{\partial f}{\partial v^+}(0)$, i.e., f is differentiable in the direction v at 0. Then, from Proposition 2.38, we conclude that f is differentiable at 0.

Conversely, suppose that f is differentiable in any direction and let $\ell : \mathbb{R}^n \to \mathbb{R}$ be a linear function, the graph of which is a supporting hyperplane for $\mathrm{Epi}(f)$ at $(0,0)$. Then $\ell(x) \le f(x)$ for all $x \in \Omega$ and, for every $v \in \mathbb{R}^n$ and $t > 0$ small,

$$\ell(v) = \frac{\ell(tv)}{t} \le \frac{f(tv)}{t}.$$

For $t \to 0^+$ we get $\ell(v) \le \frac{\partial f}{\partial v}(0)$; replacing v with $-v$ we also have $\ell(-v) \le \frac{\partial f}{\partial(-v)}(0)$, thus $\ell(v) = \frac{\partial f}{\partial v}(0)$, i.e., ℓ is uniquely defined. □

2.38 Proposition. *Let $\Omega \subset \mathbb{R}^n$ be open and convex and let $f : \Omega \to \mathbb{R}$ be convex. Then f is differentiable at $x_0 \in \Omega$ if and only if f has partial derivatives at x_0.*

Proof. We may and do assume that $x_0 = 0$ and $f(0) = 0$. Therefore, assume f is convex and has partial derivatives at 0. Additionally,

$$\phi(h) := f(h) - f(0) - \nabla f(0) \bullet h\,, \qquad h \in \Omega,$$

is convex and has zero partial derivatives at 0. Writing $h = \sum_{i=1}^n h^i e_i$, we have for every $i = 1, \dots, n$

$$\frac{\phi(nh^i e_i)}{nh^i} = o(1), \qquad h^i \to 0;$$

additionally, Jensen's inequality yields

$$\phi(h) = \phi\Big(\frac{1}{n}\sum_{i=1}^n h^i n e_i\Big) \le \frac{1}{n}\sum_{i=1}^n \phi(nh^i e_i).$$

Using Cauchy's inequality we then get

$$\phi(h) \le \sum_{i=1}^n h^i \frac{\phi(h^i n e_i)}{nh^i} \le |h| \bigg(\sum_{i=1}^n \Big|\frac{\phi(h^i n e_i)}{h^i n}\Big|^2\bigg)^{1/2} = |h|\epsilon(h)$$

where

$$\epsilon(h) := \bigg(\sum_{i=1}^n \Big|\frac{\phi(h^i n e_i)}{h^i n}\Big|^2\bigg)^{1/2}.$$

Notice that $\epsilon(h) \ge 0$, and $\epsilon(h) \to 0$ as $h \to 0$. Replacing h with $-h$ we also get

$$\phi(-h) \le |h|\epsilon(-h) \qquad \text{with } \epsilon(-h) \ge 0, \text{ and } \epsilon(-h) \to 0 \text{ as } h \to 0.$$

Since $\phi(h) \ge -\phi(-h)$ (in fact, $0 = \phi((h-h)/2) \le \phi(h)/2 + \phi(-h)/2$) we obtain

$$-|h|\epsilon(-h) \le \phi(-h) \le \phi(h) \le |h|\epsilon(h)$$

and conclude that

$$\Big|\frac{\phi(h)}{h}\Big| \le \max(\epsilon(h), \epsilon(-h)), \qquad \text{therefore} \qquad \lim_{h\to 0}\frac{\phi(h)}{|h|} = 0,$$

i.e., ϕ and, consequently, f, is differentiable at 0. □

2.39 ¶. For $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ and $v \in \mathbb{R}^n$ set

$$\frac{\partial f}{\partial v^+}(x) := \lim_{t\to 0^+}\frac{f(x+tv) - f(v)}{t}.$$

Assuming that Ω is open and convex and $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ is convex, prove the following:

(i) For all $x \in \Omega$ and $v \in \mathbb{R}^n$, $\frac{\partial f}{\partial v^+}(x)$ exists.

(ii) $v \to \frac{\partial f}{\partial v^+}(x)$, $v \in \mathbb{R}^n$, is a convex and positively 1-homogeneous function.

(iii) $f(x+v) - f(x) \ge \frac{\partial f}{\partial v^+}(x)$ for all $x \in \Omega$ and all $v \in \mathbb{R}^n$.

(iv) $v \to \frac{\partial f}{\partial v^+}(x)$ is linear if and only if f is differentiable at x.

g. Extremal points of convex functions

The extremal points of convex functions have special features. In Exercise 2.34, for instance, we saw that a convex function $f : K \to \mathbb{R}$ need not have a minimum point even when K is compact; moreover, minimizers form a convex subset of K. We also saw that local minimizers are in fact global minimizers and that, assuming $f \in C^1(K)$ and x_0 interior to K, the point x_0 is a minimizer for f if and only if $Df(x_0) = 0$. When a minimizer x_0 is not necessarily an interior point, we have the following proposition.

2.40 Proposition. *Let Ω be an open set of $\mathbb{R}^n$, K a convex subset of Ω and $f : \Omega \to \mathbb{R}$ a convex function of class $C^1(\Omega)$. The following claims are equivalent:*

(i) *x_0 is a minimum point of f in K.*
(ii) *$\mathbf{D}f(x_0) \bullet (x - x_0) \geq 0\ \forall x \in K$.*
(iii) *$\mathbf{D}f(x) \bullet (x - x_0) \geq 0\ \forall x \in K$.*

Proof. (i) $\Leftrightarrow$ (ii). If x_0 is a minimizer in K, for all $x \in K$ and $\lambda \in]0,1[$ we have

$$f(x_0) \leq f((1-\lambda)x_0 + \lambda\, x),$$

hence

$$\frac{f(x_0 + \lambda\,(x - x_0)) - f(x_0)}{\lambda} \geq 0.$$

When $\lambda \to 0$, the left-hand side converges to $\mathbf{D}f(x_0) \bullet (x - x_0)$, hence (ii). Conversely, since f is convex and of class $C^1(\Omega)$ we have

$$f(x) \geq f(x_0) + \mathbf{D}f(x_0) \bullet (x - x_0) \geq f(x_0) \qquad \forall x \in K,$$

thus x_0 is a minimizer of f in K.

(ii) $\Leftrightarrow$ (iii). From Theorem 2.32 we know that $\mathbf{D}f$ is a monotone operator

$$(\mathbf{D}f(x) - \mathbf{D}f(x_0)) \bullet (x - x_0) \geq 0.$$

Thus (ii) implies (iii) trivially.

(iii) $\Leftrightarrow$ (ii). For any $x \in K$ and $\lambda \in]0,1[$ (iii) yields

$$\mathbf{D}f(x_0 + \lambda(x - x_0)) \bullet (\lambda(x - x_0)) \geq 0,$$

hence for $\lambda > 0$

$$\mathbf{D}f(x_0 + \lambda(x - x_0)) \bullet (x - x_0) \geq 0.$$

Since $\lambda \to \mathbf{D}f(x_0 + \lambda(x - x_0)) \bullet (x - x_0)$ is continuous at 0, for $\lambda \to 0^+$ we get (ii). $\square$

The analysis of maximum points is slightly more delicate. In the 1-dimensional case a convex function $f : [a,b] \to \mathbb{R}$ has a maximum point in a or b. However, in higher dimensions the situation is more complicated.

2.41 Example. The function

$$f(x,y) := \begin{cases} \dfrac{x^2}{y} & \text{if } y > 0, \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

is convex in $\{(x,y)\,|\,y > 0\} \cup \{(0,0)\}$, as the reader can verify. Notice that f is discontinuous at $(0,0)$ and $(0,0)$ is a minimizer for f.

Consider the closed convex set

$$K_1 := \Big\{(x,y) \,\Big|\, x^4 \le y \le 1\Big\}.$$

We have $\sup_{\partial K_1} f(x,y) = +\infty$ since $f(x,x^4) = 1/x^2 \to \infty$ as $x \to 0$. Hence the function $f : K_1 \to \mathbb{R}$ is convex, K_1 is compact but f is unbounded on K_1.

Consider the compact and convex set

$$K_2 := \Big\{(x,y) \,\Big|\, x^2 + x^4 \le y \le 1\Big\}.$$

We have

$$f(x,y) \le \frac{x^2}{x^2+x^4} < 1 \quad \forall (x,y) \in K_2 \qquad \text{and} \qquad \sup_{(x,y)\in K_2} f(x,y) = 1.$$

Therefore, the function $f : K_2 \to \mathbb{R}$ is convex, defined on a compact convex set, bounded from above, but has no maximum point.

2.42 Proposition. *Let $K \subset \mathbb{R}^n$ be a convex and closed set that does not contain straight lines and let $f : K \to \mathbb{R}$ be a convex function.*

(i) *If f has a maximum point $\overline{x}$, then $\overline{x}$ is an extremal point of K.*
(ii) *If f is bounded from above and K is polyhedral, then f has a maximum point in K.*

Proof. The proof is by induction on the dimension. For $n = 1$, the unique closed convex subsets of $\mathbb{R}$ are the closed and bounded intervals $[a,b]$ or the closed half-lines, and in this case (i) and (ii) hold. We now proceed by induction on n.

(i) If f has a maximizer in K, then there exists $\overline{x} \in \partial K$ where f attains its maximum value. Denoting by L the supporting hyperplane of K at $\overline{x}$, then f attains its maximum in $L \cap K$ that is closed, convex and of dimension $n-1$. By the inductive assumption there exists $\widehat{x} \in L \cap K$ which is both an extremal point of $L \cap K$ and a maximizer for f. Since $\overline{x}$ needs to be also an extremal point for K, (i) holds in dimension n.

(ii) Let

$$M := \sup_{x\in K} f(x) = \sup_{x \in \partial K} f(x).$$

Since K is polyhedral, we have $\partial K = (K \cap L_1) \cup \dots \cup (K \cap L_N)$, where $L_1, L_2, \dots, L_N$ are the hyperplanes that define K. Hence

$$M = \sup_{x \in K\cap L_i} f(x) \qquad \text{for some } i.$$

However, $K \cap L_i$ is polyhedral and $\dim(K \cap L_i) < n$. It follows that there is $\widehat{x} \in K \cap L_i$ such that $f(\widehat{x}) = M$. □

2.3 Convex Duality

a. The polar set of a convex set

A basic construction when dealing with convexity is *convex duality*. Here we see it as the construction of the *polar set*.

Let $K \subset \mathbb{R}^n$ be an arbitrary set. The *polar* of K is defined as

$$K^* := \Big\{\xi \,\Big|\, \xi \bullet x \le 1 \ \forall x \in K\Big\}.$$

2.43 Example. (i) If $K = \{x\}$, $x \neq 0$, then its polar

$$K^* = \Big\{\xi \,\Big|\, \xi \bullet x \le 1\Big\},$$

is the closed half-space delimited by the hyperplane $\xi \bullet x = 1$ and containing the origin. Notice that $\xi \bullet x = 1$ is one of the two hyperplanes orthogonal to x at distance $1/|x|$ from the origin.

(ii) If $K := \{0\}$, then trivially $K^* = \mathbb{R}^n$,

(iii) If $K = \overline{B(0,r)}$, then

$$K^* = \overline{B(0,1/r)}.$$

In fact, if $\xi \in B(0,1/r)$, then $\xi \bullet x \le ||\xi||\,||x|| \le \frac{1}{r} r = 1$, i.e., $B(0,1/r) \subset K^*$. On the other hand, $x \bullet y = ||x||\,||y||$ if and only if either $y = 0$ or x is a nonnegative multiple of y. For all $\xi \in K^*$, if $x := r\frac{\xi}{|\xi|} \in \overline{B(0,r)}$, we have $r\,||\xi|| = \xi \bullet x = ||x||\,||\xi|| \le 1$; hence $K^* \subset \overline{B(0,1/r)}$.

Since the polar set is characterized by a family of linear inequalities, we infer the following.

2.44 Proposition. *We have the following:*

(i) *For every nonempty set K, the polar set K^* is convex, closed and contains the origin.*

(ii) *If $\{K_\alpha\}_{\alpha\in\mathcal{A}}$ is a family of nonempty sets of $\mathbb{R}^n$, then*

$$\Big(\bigcup_{\alpha\in\mathcal{A}} K_\alpha\Big)^* = \bigcap_{\alpha\in\mathcal{A}} K_\alpha^*.$$

(iii) *If $K_1 \subset K_2 \subset \mathbb{R}^n$, then $K_1^* \supset K_2^*$.*

(iv) *If $\lambda > 0$ and $K \subset \mathbb{R}^n$, then $(\lambda K)^* = \frac{1}{\lambda}K^*$.*

(v) *If $K \subset \mathbb{R}^n$, then $(\mathrm{co}(K))^* = K^*$.*

(vi) *$(K \cup \{0\})^* = K^*$.*

Proof. (i) By definition K^* is the intersection of a family of closed half-spaces containing 0, hence it is closed, convex and contains the origin.

(ii) From the definition

$$\begin{aligned}\Big(\bigcup_{\alpha\in\mathcal{A}} K_\alpha\Big)^* &= \Big\{\xi \,\Big|\, \xi \bullet x \le 1\ \forall x \in \cup_{\alpha\in\mathcal{A}} K_\alpha\Big\}\\ &= \bigcap_{\alpha\in\mathcal{A}} \Big\{\xi \,\Big|\, \xi \bullet x \le 1\ \forall x \in K_\alpha\Big\} = \bigcap_{\alpha\in\mathcal{A}} K_\alpha^*.\end{aligned}$$

(iii) Writing $K_2 = K_1 \cup (K_2 \setminus K_1)$, it follows from (ii) that $K_2^* \subset K_1^* \cap (K_2 \setminus K_1)^* \subset K_1^*$.

(iv) $\xi \in (\lambda K)^*$ if and only if $\xi \bullet x \le 1$ $\forall x \in \lambda K$, equivalently, if and only if $\xi \bullet \lambda x \le 1$ $\forall x \in K$, i.e., $(\lambda\xi) \bullet x \le 1$ $\forall x \in K$, that is, if and only if $\lambda\xi \in K^*$.

(v) It suffices to notice that ξ satisfies $\xi \bullet x_1 \le 1$ and $\xi \bullet x_2 \le 1$ if and only if $\xi \bullet x \le 1$ for every x that is a convex combination of x_1 and x_2.

(vi) Trivial. □

2.45 Corollary. *Let $K \subset \mathbb{R}^n$. Then the following hold.*

(i) *If $0 \in \mathrm{int}(K)$, then K^* is closed, convex and compact.*

(ii) *If K is bounded, then* $0 \in \operatorname{int}(K^*)$.

Proof. If $0 \in \operatorname{int}(K)$, there is $B(0,r) \subset K$, hence, $K^* \subset B(0,r)^* = \overline{B(0,1/r)}$ and K is bounded. Similarly, if K is bounded, $K \subset B(0,M)$, then $\overline{B(0,1/M)} = B(0,M)^* \subset K^*$ and $0 \in \operatorname{int}(K^*)$. $\square$

A compact convex set with interior points is called a *convex body*. From the above the polar set of a convex body K with $0 \in \operatorname{int}(K)$ is again a convex body with $0 \in \operatorname{int} K^*$.

The following important fact holds.

2.46 Theorem. *Let K be a closed convex set of $\mathbb{R}^n$ with $0 \in K$. Then $K^{**} = K$ where $K^{**} := (K^*)^*$.*

Proof. If $x \in K$, then $\xi \bullet x \le 1\ \forall \xi \in K^*$, hence $x \in K^{**}$ and $K \subset K^{**}$. Conversely, if $x_0 \notin K$, then there is a supporting hyperplane of K

$$\mathcal{P} = \left\{ x \;\middle|\; \eta \bullet x = 1 \right\}$$

that strongly separates K from x, see Theorem 2.6, and, since $0 \in K$,

$$\eta \bullet x < 1 \quad \forall x \in K \qquad \text{and} \qquad \eta \bullet x_0 > 1.$$

The first inequality states that $\eta \in K^*$, whereas the second states that $x_0 \notin K^*$. Consequently, $K^{**} \subset K$. $\square$

Later, in Section 2.4, we shall see a few applications of polarity.

b. The Legendre transform for functions of one variable

In Paragraph a. we introduced the notion of *convex duality* for bodies. We now discuss a similar notion of duality for convex functions: the *Legendre transform*. We begin with functions of one real variable.

Let I be an interval of $\mathbb{R}$ and $f : I \to \mathbb{R}$ be a convex function. Suppose that f is of class C^2 and that $f'' > 0$ in I. Then $f' : I \to \mathbb{R}$ is strictly increasing and we may describe f in terms of the slope p by choosing for every $p \in f'(I)$ the unique $x \in I$ such that $f'(x) = p$ and defining the *Legendre transform* of f as

$$\mathcal{L}_f(p) := xp - f(x), \qquad x := x(p) = (f')^{-1}(p), \qquad p \in f'(I),$$

see Figure 2.6. In this way we have a description of f in terms of the variable p that we say is *dual* of the variable x. It is easy to prove that $\mathcal{L}_f(p)$ is of class C^2 as f and that $\mathcal{L}_f$ is strictly convex. In fact, writing $x = x(p)$ for $x = (f')^{-1}(p)$, we compute

$$(\mathcal{L}_f)'(p) = x(p) + px'(p) - f'(x(p))x'(p) = x(p), \tag{2.15}$$

$$(\mathcal{L}_f)''(p) = D(x(p)) = \frac{1}{D(f')(x(p))} = \frac{1}{f''(x(p))}. \tag{2.16}$$

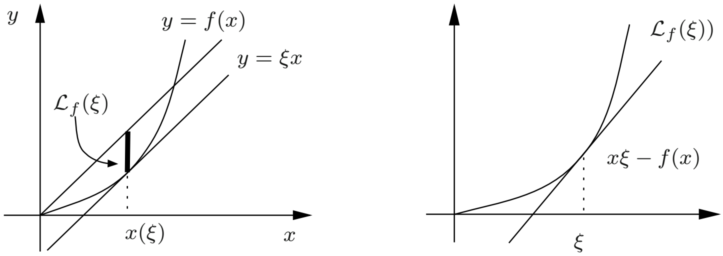


Figure 2.6. A geometric description of the Legendre transform.

c. The Legendre transform for functions of several variables

The previous construction extends to strictly convex functions of several variables giving rise to the *Legendre transform* that is relevant in several fields of mathematics and physics.

Let Ω be an open convex subset of $\mathbb{R}^n$ and let $f : \Omega \to \mathbb{R}$ be a function of class C^s $s \geq 2$ with strictly positive Hessian matrix at every point $x \in \Omega$. Denote by $\mathbf{D}f : \Omega \to \mathbb{R}^n$ the *Jacobian map* of f, with $\Omega^* := \mathbf{D}f(\Omega) \subset \mathbb{R}^n$ and ξ the variable in Ω^*. The Jacobian map, or gradient map, is clearly of class C^{s-1}, and since

$$\det \mathbf{D}(\mathbf{D}f)(x) = \det \mathbf{H}f(x) > 0,$$

the implicit function theorem tells us that Ω^* is open and the gradient map is locally invertible. Actually, the gradient map is a diffeomorphism from Ω onto Ω^* of class C^{s-1}, since it is injective: In fact, if $x_1 \neq x_2 \in \Omega$ and $\gamma(t) := x_1 + tv$, $t \in [0,1]$, $v := x_2 - x_1$, we have

$$\begin{aligned}(\mathbf{D}f(x_2) - \mathbf{D}f(x_1)) \bullet v &= \left(\int_0^1 \frac{d}{ds}(\mathbf{D}f(\gamma(s)))\, ds \right) \bullet v \\ &= \int_0^1 \mathbf{H}f(\gamma(s))v \bullet v \, ds > 0,\end{aligned}$$

i.e., $\mathbf{D}f(x_1) \neq \mathbf{D}f(x_2)$.

Denote by $x(\xi) : \Omega^* \to \Omega$ the inverse of the gradient map

$$x(\xi) := [\mathbf{D}f]^{-1}(\xi) \quad \text{or} \quad \xi = \mathbf{D}f(x(\xi)) \qquad \forall \xi \in \Omega^*.$$

2.47 Definition. *The* Legendre transform *of f is the function $\mathcal{L}_f : \Omega^* \to \mathbb{R}$ given by*

$$\mathcal{L}_f(\xi) := \xi \bullet x(\xi) - f(x(\xi)), \qquad x(\xi) := (\mathbf{D}f)^{-1}(\xi). \tag{2.17}$$

2.48 Theorem. *$\mathcal{L}_f : \Omega^* \to \mathbb{R}$ is of class C^s, and the following formulas hold:*

$$\mathbf{D}\mathcal{L}_f(\xi) = x(\xi) = (\mathbf{D}f)^{-1}(\xi), \qquad \mathbf{H}\mathcal{L}_f(\xi) = \Big(\mathbf{H}f(x(\xi))\Big)^{-1}, \tag{2.18}$$

$$\mathcal{L}_f(\xi) = \xi \bullet x(\xi) - f(x(\xi)), \qquad x(\xi) := \mathbf{D}f^{-1}\xi = \mathbf{D}\mathcal{L}_f(\xi), \tag{2.19}$$

$$f(x) = \xi(x) \bullet x - \mathcal{L}_f(\xi(x)), \qquad \xi(x) = \mathbf{D}f(x). \tag{2.20}$$

In particular, if Ω^* is convex, the Legendre transform $f \to \mathcal{L}_f$ is involutive, i.e., $\mathcal{L}_{\mathcal{L}_f} = f$.

Proof. $\mathcal{L}_f$ is of class C^{s-1}, $s \geq 1$; let us prove that it is of class C^s as f. From $\xi = \mathbf{D}f(x(\xi))$ we infer

$$d\mathcal{L}_f(\xi) = x^\alpha(\xi)\, d\xi_\alpha + \xi_\alpha\, dx^\alpha - \frac{\partial f}{\partial x^\alpha}(x(\xi))\, dx^\alpha = x^\alpha(\xi)\, d\xi_\alpha,$$

i.e., $\frac{\partial \mathcal{L}_f}{\partial \xi_\alpha}(\xi) = x^\alpha(\xi)$. Since $x(\xi)$ is of class C^{s-1}, then $\mathcal{L}_f(\xi)$ is also of class C^s, and $\mathbf{D}\mathcal{L}_f(\xi) = x(\xi)$. Also from $\mathbf{D}f(x(\xi)) = \xi$ for all $\xi \in \Omega^*$ we infer $\mathbf{H}f(x(\xi))\mathbf{D}x(\xi) = \mathrm{Id}$, hence

$$\mathbf{H}\mathcal{L}_f(\xi) = \mathbf{D}x(\xi) = \Big(\mathbf{H}f(x(\xi))\Big)^{-1}.$$

In particular, the Hessian matrix of $\xi \to \mathcal{L}_f(\xi)$ is positive definite. The other claims now follow easily. □

If $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ has a positive definite Hessian matrix and Ω is convex, as previously, then f is strictly convex. However, if $n \geq 2$, the Legendre transform of f, $\mathcal{L}_f : \Omega^* \to \mathbb{R}$, need not be convex since its domain Ω^* in general may not be convex as for the Legendre transform of the function $\exp(|x|^2)$ defined on the unit cube $\Omega := \{x = (x_1, x_2, \dots, x_n) \,|\, \max_i |x_i| \leq 1\}$. However, $\mathcal{L}_f$ has a strictly positive Hessian matrix, in particular, $\mathcal{L}_f$ is locally convex.

Finally, the following characterization of the Legendre transform holds.

2.49 Proposition. *Let $f \in C^s(\Omega)$, Ω be open and convex, $s \geq 2$, and $\mathbf{H}f > 0$ in Ω. Then*

$$\mathcal{L}_f(\xi) = \max\Big\{ \xi \bullet x - f(x) \,\Big|\, x \in \Omega \Big\}. \tag{2.21}$$

Proof. Fix $\xi \in \Omega^*$, and consider the concave function $g(x) := \xi \bullet x - f(x)$, $x \in \Omega$. The function $x \to \mathbf{D}g(x) := \xi - \mathbf{D}f(x)$ vanishes exactly at $\xi = \mathbf{D}f(x)$. It follows that $g(x)$ has an absolute maximum point at $x = \mathbf{D}f^{-1}(\xi)$ and the maximum value is $\mathcal{L}_f(\xi)$. □

Later we shall deal with (2.21).

2.4 Convexity at Work

2.4.1 Inequalities

a. Jensen inequality

Many inequalities find their natural context and can be conveniently treated in terms of convexity. We have already discussed in [GM1] and

Chapter 4 of [GM4] some inequalities as consequences of the convexity of suitable functions of one variable. We recall the *discrete Jensen's inequality.*

2.50 Proposition. *Let $\phi : [a,b] \to \mathbb{R}$ be a convex function, $x_1, \dots, x_m \in [a,b]$ and $\alpha_i \in [0,1]$ $\forall i = 1, \dots, m$ with $\sum_{i=1}^m \alpha_i = 1$. Then*

$$\phi\Big(\sum_{i=1}^m \alpha_i x_i\Big) \le \sum_{i=1}^m \alpha_i \phi(x_i).$$

Moreover, if ϕ is strictly convex *and $\alpha_i > 0$ $\forall i$, then $\phi\Big(\sum_{i=1}^m \alpha_i x_i\Big) = \sum_{i=1}^m \alpha_i \phi(x_i)$ if and only if $x_1 = \dots = x_m$.*

We now list some consequences of Jensen's inequality:

(i) (Young inequality) If $p, q > 1$, $\frac{1}{p} + \frac{1}{q} = 1$, then

$$ab \le \frac{a^p}{p} + \frac{b^q}{q} \qquad \forall a, b \in \mathbb{R}^+$$

with equality if and only if $a^p = b^q$.

(ii) (Geometric and arithmetic means) If $x_1, x_2, \dots, x_n \ge 0$, then

$$\sqrt[n]{x_1 x_2 \dots x_n} \le \frac{1}{n} \sum_{i=1}^n x_i$$

with equality if and only if $x_1 = \dots = x_n = \frac{1}{n}\sum_{i=1}^n x_i$.

(iii) (Hölder inequality) If $p, q > 1$ and $1/p + 1/q = 1$, then for all $x_1, x_2, \dots, x_n \ge 0$ and $y_1, y_2, \dots, y_n \ge 0$ we have

$$\sum_{i=1}^n x_i y_i \le \Big(\sum_{i=1}^n x_i^p\Big)^{1/p} \Big(\sum_{i=1}^n y_i^q\Big)^{1/q},$$

with equality if and only if either $x_i = \lambda y_i$ $\forall i$ for some $\lambda \ge 0$ or $y_1 = \dots = y_n = 0$.

(iv) (Minkowski inequality) If $p, q > 1$ and $1/p + 1/q = 1$, then for all $x_1, x_2, \dots, x_n \ge 0$ and $y_1, y_2, \dots, y_n \ge 0$ we have

$$\Big(\sum_{i=1}^n (x_i + y_i)^p\Big)^{1/p} \le \Big(\sum_{i=1}^n x_i^p\Big)^{1/p} + \Big(\sum_{i=1}^n y_i^p\Big)^{1/p}$$

with equality if and only if either $x_i = \lambda\, y_i$ $\forall i$ for some $\lambda \ge 0$ or $y_1 = \dots = y_n = 0$.

(v) (Entropy inequality) The function $f(p) := \sum_{i=1}^n p_i \log p_i$ defined on $K := \{p \in \mathbb{R}^n \,\big|\, p_i \ge 0, \sum_{i=1}^n p_i = 1\}$ has a unique strict minimum point at $\overline{p} = (1/n, \dots, 1/n)$.

(vi) (Hadamard's inequality) Since the determinant and the trace of a square matrix are respectively the product and the sum of the eigenvalues, the inequality between geometric and arithmetic means yields

$$\det \mathbf{A} \le \left(\frac{\operatorname{tr} \mathbf{A}}{n}\right)^n$$

for every matrix $\mathbf{A}$ that is symmetric and with nonnegative eigenvalues. Moreover, equality holds if and only if $\mathbf{A}$ is a multiple of the identity matrix. A consequence is that for every $\mathbf{A} \in M_{n,n}(\mathbb{R})$ the following *Hadamard's inequality* holds:

$$(\det \mathbf{A})^2 \le \prod_{i=1}^n |A_i|^2$$

where $A_1, A_2, \dots, A_n$ are the columns of $\mathbf{A}$ and $|A_i|$ is the length of the column vector A_i; moreover, equality holds if and only if $\mathbf{A}$ is a multiple of an orthogonal matrix.

b. Inequalities for functions of matrices

Let $\mathbf{A} \in M_{n,n}(\mathbb{R})$ be symmetric and let $\mathbf{A}x = \sum_{i=1}^n \lambda_i (x \bullet u_i)\, u_i$ be its spectral decomposition. Recall that for $f : \mathbb{R} \to \mathbb{R}$, the matrix $f(\mathbf{A})$ is defined as the $n \times n$ symmetric matrix

$$f(\mathbf{A})(x) := \sum_{i=1}^n f(\lambda_i)(x \bullet u_i)\, u_i.$$

Notice that $\mathbf{A}$ and $f(\mathbf{A})$ have the same eigenvectors with corresponding eigenvalues λ and $f(\lambda)$, respectively.

2.51 Proposition. *Let $\mathbf{A} \in M_{n,n}(\mathbb{R})$ be symmetric and let $f : \mathbb{R} \to \mathbb{R}$ be convex. For all $x \in \mathbb{R}^n$ we have*

$$f(x \bullet \mathbf{A}x) \le\ x \bullet f(\mathbf{A})x\,.$$

In particular, if $\{v_1, v_2, \dots, v_n\}$ is an orthonormal basis of $\mathbb{R}^n$, we have

$$\sum_{j=1}^n f(v_j \bullet \mathbf{A}v_j) \le \operatorname{tr}(f(\mathbf{A})).$$

Proof. Let $u_1, u_2, \dots, u_n$ be an orthonormal basis of $\mathbb{R}^n$ of eigenvectors of $\mathbf{A}$ with corresponding eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$. Then

$$x \bullet \mathbf{A}x = \sum_{i=1}^n \lambda_i |\, x \bullet u_i \,|^2, \qquad x \bullet f(\mathbf{A})x = \sum_{i=1}^n f(\lambda_i) |\, x \bullet u_i \,|^2,$$

and, since $\sum_{i=1}^n |\, x \bullet u_i \,|^2 = |x|^2$, the discrete Jensen's inequality yields

$$f(\, x \bullet \mathbf{A}x \,) = f\Big(\sum_{i=1}^{n} \lambda_i |\, x \bullet u_i \,|^2 \Big) \le \sum_{i=1}^{n} f(\lambda_i) |\, x \bullet u_i \,|^2 = \, x \bullet f(\mathbf{A}x)\,.$$

The second part of the claim then follows easily. In fact, from the first part of the claim,

$$\sum_{j=1}^{n} f\Big(v_j \bullet \mathbf{A}v_j \Big) \le \sum_{j=1}^{n} v_j \bullet f(\mathbf{A})v_j \;,$$

and, since $\{v_j\}$ is orthonormal, there exists an orthogonal matrix $\mathbf{R}$ such that $v_j = \mathbf{R}u_j$, and the spectral theorem yields

$$\sum_{j=1}^{n} v_j \bullet f(\mathbf{A})v_j \; = \sum_{j=1}^{n} u_j \bullet \mathbf{R}^T f(\mathbf{A})\mathbf{R}u_j \; = \sum_{j=1}^{n} f(\lambda_j) = \operatorname{tr} f(\mathbf{A}).$$

□

2.52 ¶. Show that

$$\frac{\Big(\prod_{i=1}^{N} x_i\Big)^{1/N} + \Big(\prod_{i=1}^{N} y_i\Big)^{1/N}}{\Big[\prod_{i=1}^{N} (x_i + y_i)\Big]^{1/N}} = \Big(\prod_{i=1}^{N} \frac{x_i}{x_i + y_i}\Big)^{1/N} + \Big(\prod_{i=1}^{N} \frac{y_i}{x_i + y_i}\Big)^{1/N}$$

$$\le \frac{1}{N} \sum_{i=1}^{N} \frac{x_i}{x_i + y_i} + \frac{1}{N} \sum_{i=1}^{N} \frac{y_i}{x_i + y_i} = 1.$$

2.53 ¶. Show that if $p, q > 1$, $1/p + 1/q = 1$, then for all $x_1, x_2, \dots, x_n \ge 0$,

$$\Big(\sum_{i=1}^{n} x_i^p \Big)^{1/p} = \max\Big\{ \sum_{i=1}^{n} x_i y_i \;\Big|\; y_i \ge 0,\; \sum_{i=1}^{n} y_i^q = 1 \Big\}.$$

c. Doubly stochastic matrices

A matrix $\mathbf{A} = (a_{jk}) \in M_{n,n}(\mathbb{R})$ is said to be *doubly stochastic* if

$$a_{jk} \ge 0, \qquad \sum_{i=1}^{n} a_{ik} = 1, \qquad \sum_{i=1}^{n} a_{ji} = 1, \qquad \forall j, k = 1, \dots, n. \tag{2.22}$$

Important examples are given by the matrix that in each row and in each column contains exactly an element equal to 1. They are characterized by a permutation σ of $\{1, \dots, n\}$ such that $a_{jk} = 1$ if $k = \sigma(j)$ and $a_{jk} = 0$ if $k \ne \sigma(j)$; for this reason they are called *permutation matrices.* Clearly, if (a_{jk}) is a permutation matrix, then $a_{jk} x_k = x_{\sigma(j)}$.

Condition (2.22) defines the space Ω_n of doubly stochastic matrices as the intersection of closed half-spaces and affine hyperplanes of $\mathbb{R}^{n^2}$, hence as a closed convex subset of the space $M_{n,n}$ of $n \times n$ matrices.

2.54 Theorem (Birkhoff). *The set Ω_n of doubly stochastic matrices is a compact and convex subset of an affine subspace of dimension $(n-1)^2$, the extremal points of which are the permutation matrices. Consequently, every doubly stochastic matrix is the convex combination of at most $(n-1)^2 + 1$ permutation matrices.*

Proof. Since $a_{jk} \leq 1$, $\forall \mathbf{A} = (a_{jk}) \in \Omega_n$, the set Ω_n is bounded, hence compact being closed. Conditions (2.22) writes as $a_{ij} \geq 0$ and

$$\begin{cases} a_{nk} = 1 - \sum_{j<n} a_{jk} & k < n, \\ a_{jn} = 1 - \sum_{k<n} a_{jk} & j < n, \\ a_{nn} = 2 - n + \sum_{j,k<n} a_{jk}, \end{cases}$$

hence Ω_n is the image of the subset P defined by

$$\begin{cases} a_{jk} \geq 0 & j,k < n, \\ \sum_{j<n} a_{jk} \leq 1 & k < n, \\ \sum_{k<n} a_{jk} \leq 1 & j < n, \\ \sum_{ij} a_{jk} \geq n-2 \end{cases} \tag{2.23}$$

through an affine and *injective* map from $\mathbb{R}^{(n-1)^2}$ into $M_{n,n}$. Moreover, P has interior points as, for instance, $a_{jk} := A/(n-1)$, $1 \leq j,k < n$ with $(n-2)/(n-1) < A < 1$, hence Ω_n has dimension $(n-1)^2$.

Of course, the permutation matrices are extremal points of Ω_n. We now prove that they are the unique extremal points. We first observe that if $\mathbf{A} = (a_{jk})$ is an extremal point of Ω_n, then it has to satisfy at least $(n-1)^2$ equations of the n^2 conditions in (2.22). Otherwise we could find $\mathbf{B} := (b_{jk}) \neq 0$ such that $a_{jk} \pm \epsilon b_{jk}$, ϵ small, still satisfies (2.22); moreover, $a_{jk} = \frac{1}{2}(a_{jk} + \epsilon b_{jk}) + \frac{1}{2}(a_{jk} - \epsilon b_{jk})$ and $\mathbf{A}$ would not be an extremal point. This means that $\mathbf{A} = (a_{jk})$ has at most $n^2 - (n-1)^2 = 2n-1$ nonzero elements implying that at least one column has to have one nonzero element, hence 1, and, of course, the row corresponding to this 1 will have all other elements zero. Deleting this row and this column we still have an extremal point of Ω_{n-1}; by downward induction we then reduce to prove the claim for 2×2 matrices where it is trivially true. □

We shall now discuss an extension of Proposition 2.51.

2.55 Proposition. *Let* $\mathbf{A}$ *be an* $n \times n$ *symmetric matrix, let* $\{u_1, \dots, u_n\}$ *be an orthonormal basis of eigenvectors of* $\mathbf{A}$ *with corresponding eigenvalues* $\lambda_1, \lambda_2, \dots, \lambda_n$ *and let* $v_1, v_2, \dots, v_n$ *be any other orthonormal basis of* $\mathbb{R}^n$. *For* $\lambda \in \mathbb{R}^n$, *set*

$$K_\lambda := \left\{ x \in \mathbb{R}^n \,\middle|\, x = \mathbf{S}\lambda,\ \mathbf{S} \in \Omega_n \right\}.$$

Then K_λ *is convex and we have*

$$(\, v_1 \bullet \mathbf{A}v_1\,,\, v_2 \bullet \mathbf{A}v_2\,, \dots,\, v_n \bullet \mathbf{A}v_n\,) \in K_\lambda.$$

Moreover, for any convex function $f : U \supset K_\lambda \to \mathbb{R}$ *the following inequality holds:*

$$f(\, \mathbf{A}v_1 \bullet v_1\,, \dots,\, \mathbf{A}v_n \bullet v_n\,) \leq f(\lambda_{\sigma_1}, \dots \lambda_{\sigma_n})$$

for some permutation $\sigma \in \mathcal{P}_n$.

Proof. The matrix $\mathbf{S} = (s_{ij})$, $s_{ij} := |\, u_i \bullet v_j\,|^2$ is doubly stochastic. Moreover, on account of the spectral theorem, $v_j \bullet \mathbf{A}v_j = \sum_{i=1}^n \lambda_i |\, v_j \bullet u_i\,|^2$. Hence $\mathbf{A}v_j \bullet v_j = S_j \bullet \lambda$, where S_j is the jth column of the matrix $\mathbf{S}$. We then conclude that

$$(\, v_1 \bullet \mathbf{A}v_1\,,\, v_2 \bullet \mathbf{A}v_2\,, \dots,\, v_n \bullet \mathbf{A}v_n\,) \in K_\lambda.$$

It is easily seen that $g(\mathbf{S}) := f(\mathbf{S}\lambda) : K_\lambda \to \mathbb{R}$ is convex. Therefore g attains its maximum value at the extremal points of K_λ, which are permutation matrices because of Birkhoff's theorem, Theorem 2.54. □

Different choices of f now lead to interesting inequalities.

(i) Choose $f(t_1, t_2, \dots, t_k) := \sum_{i=1}^k t_i$, so that both f and $-f$ are convex, and, as before, let $\mathbf{A}$ be a symmetric $n \times n$ matrix and let $\{v_1, v_2, \dots, v_n\}$ be an orthonormal basis of $\mathbb{R}^n$. Then for $1 \le k \le n$ the following estimates for $\sum_{j=1}^k \mathbf{A}v_j \bullet v_j$ holds:

$$\sum_{j=1}^k \lambda_{n-j+1} \le \sum_{j=1}^k \mathbf{A}v_j \bullet v_j \le \sum_{j=1}^k \lambda_j, \tag{2.24}$$

$\lambda_1, \lambda_2, \dots, \lambda_n$ being the eigenvalues of $\mathbf{A}$ ordered so that $\lambda_1 \ge \lambda_2 \ge \dots \ge \lambda_n$.

(ii) Choose $f(t) := (\prod_{i=1}^k t_i)^{1/k}$, $k \ge 1$, that is concave on $\{t \in \mathbb{R}^n \,|\, t \ge 0\}$, and let $\mathbf{A}$ be a symmetric positively semidefinite $n \times n$ matrix. Applying Proposition 2.55 to $-f$, for every orthonormal basis $\{v_1, v_2, \dots, v_n\}$ we find for every k, $1 \le k \le n$,

$$\Big(\prod_{i=1}^k \lambda_{n-i+1}\Big)^{1/k} \le \Big(\prod_{j=1}^k \mathbf{A}v_j \bullet v_j\Big)^{1/k} \tag{2.25}$$

$\lambda_1, \lambda_2, \dots, \lambda_n$ being the eigenvalues of $\mathbf{A}$ ordered so that $\lambda_1 \ge \lambda_2 \dots \ge \lambda_n \ge 0$.
Using the inequality between the geometric and arithmetic means and (2.24) we also find

$$\Big(\prod_{j=1}^k \mathbf{A}v_j \bullet v_j\Big)^{1/k} \le \frac{1}{k}\sum_{j=1}^k \mathbf{A}v_j \bullet v_j \le \frac{1}{k}\sum_{j=1}^k \lambda_j. \tag{2.26}$$

When $k = n$ we find again

$$\det \mathbf{A} = \prod_{j=1}^n \lambda_j \le \prod_{j=1}^n \mathbf{A}v_j \bullet v_j \le \Big(\frac{\operatorname{tr} \mathbf{A}}{n}\Big)^n. \tag{2.27}$$

2.56 Theorem (Brunn–Minkowski). *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be two symmetric and nonnegative matrices. Then*

$$\Big(\det(\mathbf{A}+\mathbf{B})\Big)^{1/n} \ge (\det \mathbf{A})^{1/n} + (\det \mathbf{B})^{1/n},$$
$$\det(\mathbf{A}+\mathbf{B}) \ge \det \mathbf{A} + \det \mathbf{B}.$$

Proof. Let $\{v_1, v_2, \dots, v_n\}$ be an orthonormal basis of eigenvectors of $\mathbf{A}+\mathbf{B}$. Then

$$\begin{aligned}\Big(\det(\mathbf{A}+\mathbf{B})\Big)^{1/n} &= \Big(\prod_{i=1}^n (\mathbf{A}+\mathbf{B})v_j \bullet v_j\Big)^{1/n} \\ &\ge \Big(\prod_{j=1}^n \mathbf{A}v_j \bullet v_j\Big)^{1/n} + \Big(\prod_{j=1}^n \mathbf{B}v_j \bullet v_j\Big)^{1/n} \\ &\ge (\det \mathbf{A})^{1/n} + (\det \mathbf{B})^{1/n},\end{aligned}$$

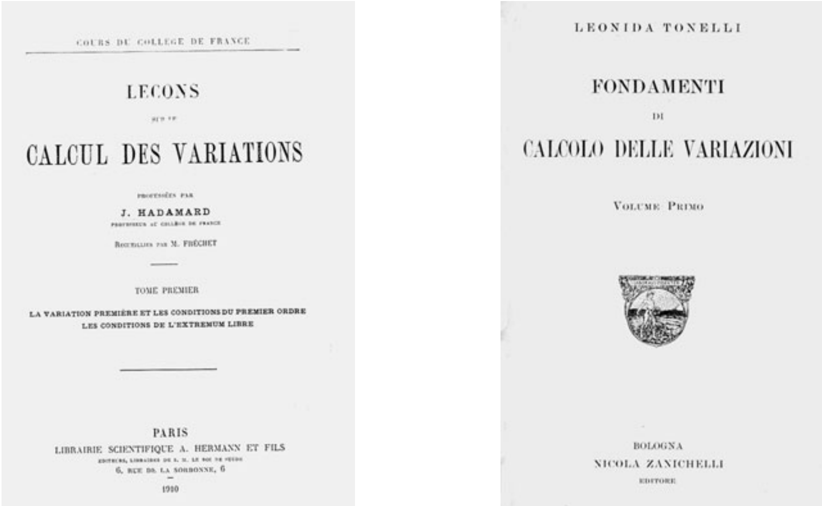


Figure 2.7. Frontispieces of two volumes about calculus of variations.

where we used Exercise 2.52 in the first estimate and (2.27) in the second one. The second inequality follows by taking the power n of the first. $\square$

2.4.2 Dynamics: Action and energy

Legendre's transform has a central role in the dual description of the dynamics of mechanical systems, the *Lagrangian* and the *Hamiltonian* models.

According to the *Hamilton* or *minimal action principle*, see Chapter 3, a mechanical system is characterized by a function $L(t, x, v)$, $L : \mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ called its *Lagrangian*, and its motion $t \to x(t) \in \mathbb{R}^N$ satisfies the following condition: If at times t_1 and t_2, $t_1 < t_2$, the system is at positions $x(t_1)$ and $x(t_2)$ respectively, then the motion in the interval of time $[t_1, t_2]$ happens in such a way as to make the *action*

$$\mathcal{A}(x) := \int_{t_1}^{t_2} L(t, x(t), x'(t))\, dt$$

stationary. More precisely, $x(t)$ is the actual motion from $x(t_1)$ to $x(t_2)$ if and only if for any arbitrary path $\gamma(t)$ with values in $\mathbb{R}^N$ such that $\gamma(t_1) = \gamma(t_2) = 0$, we have

$$0 = \frac{d}{d\epsilon}\mathcal{A}(x + \epsilon\gamma)\Big|_{\epsilon=0} = \frac{d}{d\epsilon}\int_{t_1}^{t_2} L\Big(t, x(t) + \epsilon\gamma(t), x'(t) + \epsilon\gamma'(t)\Big)\, dt\Big|_{\epsilon=0}.$$

Differentiating under the integral sign, we find

$$
\begin{aligned}
0 &= \int_{t_1}^{t_2} \sum_{i=1}^N \Big(L_{x^i} \gamma^i(t) + L_{v^i} {\gamma^i}'(t) \Big)\, dt \\
&= \int_{t_1}^{t_2} \sum_{i=1}^N \Big(L_{x^i} - \frac{d}{dt} L_{v^i} \Big) \gamma^i(t)\, dt + \sum_{i=1}^N L_{v^i} \gamma^i(t) \Big|_{t_1}^{t_2} \\
&= \int_{t_1}^{t_2} \sum_{i=1}^N \Big(L_{x^i} - \frac{d}{dt} L_{v^i} \Big) \gamma^i(t)\, dt
\end{aligned}
$$

for all $\gamma : [t_1, t_2] \to \mathbb{R}^N$, $\gamma(t_1) = \gamma(t_2) = 0$, where

$$
L_{x^i} := \frac{\partial L}{\partial x^i}(t, x(t), x'(t)), \qquad L_{v^i} := \frac{\partial L}{\partial v^i}(t, x(t), x'(t)).
$$

As a consequence of the fundamental lemma of the Calculus of Variations, see Lemma 1.51, the motion of the system is a solution of the *Euler–Lagrange* equations

$$
\frac{d}{dt} L_{v^i}(t, x(t), x'(t)) = L_{x^i}(t, x(t), x'(t)) \qquad \forall i = 1, \dots, N. \tag{2.28}
$$

This is an invariant way (with respect to changes of coordinates) of expressing Newton's law of dynamics. We notice that (2.28) are N ordinary differential equations of second order in the unknown $x(t)$.

There is another equivalent way of describing the law of dynamics at least when the Lagrangian L is of class C^2 and $\det \frac{\partial^2 L}{\partial v^2} > 0$, i.e., $L \in C^2(\mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^N)$ and $v \to L(t, x, v)$ is strictly convex for all (t, x). As we have seen, in this case the function

$$
v \longrightarrow p := L_v(t, x, v) = \frac{\partial}{\partial v} L(t, x, v)
$$

is globally invertible with inverse function $v = \psi(t, x, p)$ of class C^2 and we may form the Legendre transform of L with respect to v

$$
H(t, x, p) := p \bullet v - L(t, x, v), \qquad v := \psi(t, x, p),
$$

called the *Hamiltonian* or the *energy* of the system. For all (t, x, p) we have

$$
\begin{cases}
p = \dfrac{\partial L}{\partial v}(t, x, v), \\
L(t, x, v) + H(t, x, p) = p \bullet v\,, & \qquad v = \psi(t, x, p) \\
H_t(t, x, p) + L_t(t, x, v) = 0, \\
H_x(t, x, p) + L_x(t, x, v) = 0,
\end{cases}
$$

and, as we saw in (2.18),

$$
H_p(t, x, p) = v = \psi(t, x, p).
$$

For a curve $t \to x(t)$, if we set $v(t) = x'(t)$ and $p(t) := L_v(t, x(t), x'(t))$, we have

$$\begin{cases} v(t) = x'(t) = \psi(t, x(t), p(t)), \\ L(t, x(t), v(t)) + H(t, x(t), p(t)) = p(t) \bullet v(t), \\ H_t(t, x(t), p(t)) + L_t(t, x(t), v(t)) = 0, \\ H_x(t, x(t), p(t)) + L_x(t, x(t), v(t)) = 0. \end{cases}$$

Consequently, $t \to x(t)$ solves Euler–Lagrange equations (2.28), that can be written as

$$\begin{cases} \dfrac{dx}{dt} = v(t), \\ \dfrac{d}{dt} L_v(t, x(t), v(t)) = L_x(t, x(t), v(t)) \end{cases}$$

if and only if

$$\begin{cases} x'(t) = H_p(t, x(t), p(t)), \\ p'(t) = \dfrac{d}{dt} L_v(t, x(t), v(t)) = L_x(t, x(t), v(t)) = -H_x(t, x(t), p(t)). \end{cases}$$

Summing up, $t \to x(t)$ solves the Euler–Lagrange equations if and only if $t \to (x(t), p(t)) \in \mathbb{R}^{2N}$ solves the system of $2N$ first order differential equations, called the *canonical Hamilton system*

$$\begin{cases} x'(t) = \ \ H_p(t, x(t), p(t)), \\ p'(t) = -H_x(t, x(t), p(t)). \end{cases}$$

We emphasize the fact that, if the Hamiltonian does not depend explicitly on time (*autonomous Hamiltonians*), $H = H(x, p)$, then H is constant along the motion,

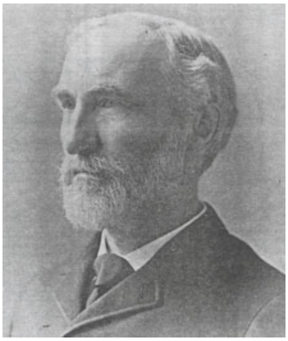
$$\frac{d}{dt} H(x(t), p(t)) = \frac{\partial H}{\partial x} \bullet x' + \frac{\partial H}{\partial p} \bullet p' = p' \bullet x' - x' \bullet p' = 0.$$

We shall return to the Lagrangian and Hamiltonian models of mechanics in Chapter 3.

2.4.3 The thermodynamic equilibrium

Here we briefly hint at the use of convexity in the discussion of the thermodynamic equilibrium by J. Willard Gibbs (1839–1903).

For the sake of simplicity we consider a quantity of N moles of a *simple fluid*, i.e., of a fluid in which equilibrium points may be described in terms of the following six thermodynamic variables:



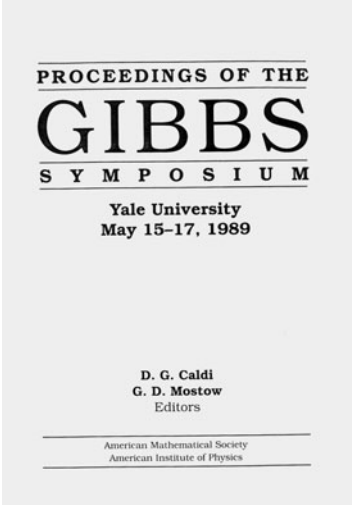


Figure 2.8. J. Willard Gibbs (1839–1903) and the frontispiece of Gibbs Symposium at Yale.

(i) V, the volume,
(ii) p, the pressure,
(iii) T, the absolute temperature,
(iv) U, the internal energy,
(v) S, the entropy,
(vi) μ, the chemical potential,
(vii) N, the number of moles.

For simple fluids, Gibbs provided a description of the thermodynamic equilibrium which is compatible with the thermodynamic laws established a few years earlier by Rudolf Clausius (1822–1888). In modern terms and freeing our presentation from experimental discussions, Gibbs assumed the following:

(i) The balance law, called the *fundamental equation*,

$$T\,dS = dU + p\,dV + \mu\,dN \tag{2.29}$$

in the variable domains $T > 0$, $V > 0$, $U > 0$, $p > 0$, $N > 0$, $\mu \in \mathbb{R}$ and $S \in \mathbb{R}$.

(ii) The equilibrium configurations can be parametrized either by the independent variables S, V and N or by the independent variables U, V and N, and, at equilibrium, the other thermodynamic quantities are functions of the chosen independent variables.

(iii) The entropy function $S = S(U, V, N)$ is of class C^1 and positively homogeneous of degree 1,

$$S(\lambda U, \lambda V, \lambda N) = \lambda S(U, V, N), \qquad \forall \lambda > 0.$$

(iv) The entropy function $S = S(U, V, N)$ is concave.
(v) The free energy function $U = U(S, V, N)$ is of class C^1, convex and positively homogeneous of degree 1.

A few comments on (i), (ii), ..., (v) are appropriate:

(i) The fundamental equation (2.29) contains the *first principle of thermodynamics*: *the elementary mechanic work done on a system plus the differential of the heat furnished to the system plus the variation of moles is an exact differential* $p\,dV - T\,dS + \mu\,dN = -dU$.
(ii) The homogeneity of S amounts, via (2.29), to the invariance at equilibrium of temperature, pressure and chemical potential when moles change.
(iii) The assumption of C^1-regularity of the entropy function, in addition to being useful, is essential in order to deduce the Gibbs necessary condition for the existence of coexisting phases.
(iv) If we choose as independent variables the internal energy U, the volume V and the number of moles N, then S, T and V are functions of (U, V, N). The fundamental equation then allows us to compute the absolute temperature and the chemical potential as partial derivatives of the entropy function $S = S(U, V, N)$, that thus describes the whole system, finding[2]

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)_{V,N}, \qquad \frac{p}{T} = \left(\frac{\partial S}{\partial V}\right)_{U,N}, \qquad \frac{\mu}{T} = \left(\frac{\partial S}{\partial N}\right)_{U,V}. \tag{2.30}$$

(v) The function $U \to S(U, V, N)$ is strictly increasing. Therefore, we can replace the independent variables (U, V, N) with the variables (S, V, N) and obtain an equivalent description of the equilibrium of the fluid in terms of the internal energy function $U = U(S, V, N)$, concluding that

$$T = \left(\frac{\partial U}{\partial S}\right)_{V,N}, \qquad -p = \left(\frac{\partial U}{\partial V}\right)_{S,N}, \qquad \mu = \left(\frac{\partial U}{\partial N}\right)_{S,V}.$$

(vi) The concavity of the entropy function is a way to formulate the second principle of thermodynamics. Consider, in fact, two quantities of the same fluid with parameters at the equilibrium $x_1 := (U_1, V_1, N_1)$ and $x_2 := (U_2, V_2, N_2)$, and a quantity of $N_1 + N_2$ moles of the same fluid with volume $V_1 + V_2$ and internal energy $U_1 + U_2$. The second principle of thermodynamics states that the entropy has to increase

$$S(x_1 + x_2) \geq S(x_1) + S(x_2).$$

Because of the arbitrariness of x_1 and x_2 and the homogeneity of S, we may infer

[2] Here we use the symbolism of physicists. For instance, by $\left(\frac{\partial S}{\partial U}\right)_{V,N}$ we mean that the function S is seen as a function of the independent variables (U, V, N) and that it is differentiated with respect to U and, consequently, the resulting function is a function of (U, V, N).

$$S((1-\alpha)x_1+\alpha x_2) \geq (1-\alpha)S(x_1)+\alpha S(x_2) \qquad \forall x_1, x_2,\ \forall \alpha \in [0,1],$$

i.e., $S(x) = S(U,V,N)$ *is a concave function.*

(vii) Similar arguments may justify the homogeneity and convexity of the internal energy function.

Gibbs's conclusion is that a simple fluid is described by a 3-dimensional surface which is at the same time the graph of $S(x)$, $x = (U,V,N) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+$ (concave, positively homogeneous of degree one and of class C^1) and the graph of the function $U(y)$, $y = (S,V,N) \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_+$, convex, positively homogeneous of degree one and of class C^1.

Since S is positively homogeneous, it is determined by its values when restricted to a "section", i.e., by its values when the energy, the volume or the number of moles is prescribed. For instance, assuming $N = 1$ and denoting by (u,v) the internal energy and the volume per mole, the entropy function per mole

$$s(u,v) := S(u,v,1),$$

describes the equilibrium of a mole of the matter under scrutiny and from (2.30)

$$\frac{1}{T(u,v)} = \Big(\frac{\partial s}{\partial u}\Big)_v, \qquad \frac{p(u,v)}{T(u,v)} = \Big(\frac{\partial s}{\partial v}\Big)_u. \tag{2.31}$$

Clearly, $s(u,v)$ is concave and the entropy S for N moles by homogeneity is given by

$$S(U,V,N) = NS\Big(\frac{U}{N},\frac{V}{N},1\Big) = N\,s\Big(\frac{U}{N},\frac{V}{N}\Big).$$

In particular, differentiating we get

$$\begin{aligned}\frac{1}{T(U,V,N)} &= \frac{\partial s}{\partial u}\Big(\frac{U}{N},\frac{V}{N}\Big),\\ p(U,V,N) &= \frac{\partial s}{\partial v}\Big(\frac{U}{N},\frac{V}{N}\Big),\\ \mu(U,V,N) &= s\Big(\frac{U}{N},\frac{V}{N}\Big) - \frac{1}{T}\,\frac{U}{N} - p\,\frac{V}{N},\end{aligned}$$

and (2.29) transforms into

$$T\,ds = du + p\,dv.$$

a. Pure and mixed phases

Gibbs also provided a description of the coexistence of different phases in terms of an analysis of the graph of a convex function. Let $s(x)$, $x \in \mathbb{R}_+ \times \mathbb{R}_+$, be a convex function in the variables $x := (u,v)$. We say that the phase x is *pure* for a liquid if $(x, s(x))$ is an extreme point of the epigraph of f. The other points are called points of *coexistent phases*: These are points x for which $(x, f(x))$ is a convex combination of the extreme points

$(x_i, f(x_i))$ of the epigraph Epi(f) of f. Since Epi(f) has dimension 3, Corollary 2.27 tells us that the boundary of Epi(f) splits into three sets

$$\begin{aligned}\Sigma_0 &:= \Big\{\text{extreme points of Epi}(f)\Big\},\\ \Sigma_1 &:= \Big\{\text{linear combinations of two points in } \Sigma_0\Big\},\\ \Sigma_2 &:= \Big\{\text{linear combinations of three points of } \Sigma_0\Big\}\end{aligned}$$

corresponding to equilibrium with pure phases, with two mixed phases and three mixed phases, respectively.

A typical situation is the one in which the pure phases are of three different types, as for water: solid, liquid and gaseous states. Then Σ_1 corresponds to the situation in which two states of the liquid coexist, and Σ_3 corresponds to states in which the three states are present at the same time.

2.57 Proposition. *Let $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ be a convex function of class C^1 and let $x_1, x_2, \dots, x_k$ be k points in Ω. A necessary and sufficient condition for the existence of $x \in \Omega$, $x \neq x_i$ $\forall i$, such that*

$$(x, f(x)) = \sum_{i=1}^k \alpha_i (x_i, f(x_i)) \qquad \textit{with} \qquad \sum_{i=1}^k \alpha_i = 1,\ \alpha_i \in [0,1] \quad (2.32)$$

is that the supporting hyperplanes to f at the points $x_1, x_2, \dots, x_k$ are the same plane. In particular, $Df(x)$ is then constant in the convex envelope of $x_1, x_2, \dots, x_k$.

Proof. Let $M := \mathrm{co}(\{x_1, x_2, \dots, x_k\})$. The convexity of $f(x)$ implies that f is linear affine in M,

$$(x, f(x)) = \sum_{i=1}^k \alpha_i (x_i, f(x_i)), \qquad \sum_{i=1}^k \alpha_i = 1,\ \alpha_i \in]0,1[,$$

for all $x \in M$ if and only if (2.32) holds. In this case the segment joining any two points $a, b \in M$ is contained in the support hyperplanes of f at a and at b. On the other hand, a support hyperplane to f at b that contains the segment joining $(a, f(a))$ with $(b, f(b))$ is also a supporting hyperplane to f at a. Since f is of class C^1, f has a unique support hyperplane at a, $z = \nabla f(a)(x-a) + f(a)$, hence the support hyperplanes to f at a and b must coincide, and $\nabla f(x)$ is constant in M. □

In the context of thermodynamics of simple fluids, the previous proposition when applied to the entropy function, see (2.31), yields the following statement.

2.58 Proposition (Gibbs). *In a simple fluid with entropy function of class C^1 two or three phases may coexists at the equilibrium only if they are at the same temperature and the same pressure.*

In principle, we may describe the thermodynamic equilibrium in terms of entropy function in the dual variables of the energy and volume, i.e., in terms of the absolute temperature and pressure. However, first we need to write $s = s(T, p)$ and $V = V(T, p)$. The Legendre duality formula turns out to be useful. In fact, starting from the internal energy $U := U(S, V, N)$ that can be obtained inverting the entropy function $S = S(U, V, N)$, we consider the internal energy per mole, $u(s, v) := U(s, v, 1)$, for which we have

$$du = T\,ds - p\,dv.$$

The dual variables of (u, v) are then $(T, -p)$: the absolute temperature T and minus the pressure p. At this point, we introduce *Gibbs's energy* as

$$G(T, p) := \sup_{s,v}\Big\{u(s, v) + pv - Ts\Big\}$$

and observe that $G(-T, p)$ is the Legendre transform of the concave function $-u$,

$$G(T, p) = \mathcal{L}_u(T, -p).$$

Therefore, at least in the case where u is strictly convex, we infer

$$s = -\Big(\frac{\partial G}{\partial T}\Big)_p, \qquad v = \Big(\frac{\partial G}{\partial p}\Big)_T.$$

2.4.4 Polyhedral sets

a. Regular polyhedra

We recall that a set K is said to be *polyhedral* if it is the intersection of finitely many closed half-spaces. A bounded polyhedral set is called a *polyhedron.*

Consider a convex polygon K containing the origin with vertices $A_1, A_2, \dots, A_N$. The vertices are the extreme points of $K \subset \mathbb{R}^n$ and $K = \mathrm{co}(\{A_1, A_2, \dots, A_N\})$, hence, compare Proposition 2.44,

$$K^* = \{A_1, A_2, \dots, A_N\}^* = \bigcap_{i=1}^{N} \{A_i\}^*$$

and, compare Theorem 2.46, $K = (K^*)^*$. Accordingly, K^* is a polyhedron, the intersection of the N half-spaces containing the origin and delimited by the hyperplanes $\{\xi \mid \xi \bullet A_i = 1\}$ in $\mathbb{R}^n$, see Figure 2.9.

2.59 ¶. The reader is invited to compute the polar sets of various convex sets of the plane.

The construction works in the same way in all $\mathbb{R}^n$'s, $n \geq 2$. Though difficult to visualize, and cumbersome to check, in $\mathbb{R}^3$, the polar set of a regular tetrahedron centered at the origin is a regular tetrahedron centered at the origin, the polar set of a cube centered at the origin is an octahedron centered at the origin, and the polar set of a dodecahedron centered at the origin is an icosahedron centered at the origin.

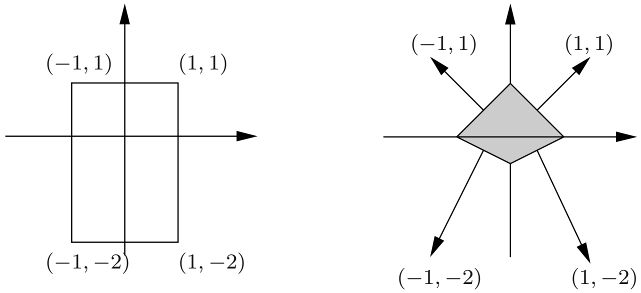


Figure 2.9. The polar set of a rectangle that contains the origin.

b. Implicit convex cones

Polyhedral sets that are cones play an important role. Let us start with cones defined implicitly by a matrix $\mathbf{A} \in M_{n,N}(\mathbb{R})$ and a vector $b \in \mathbb{R}^n$ as

$$K := \left\{ x \in \mathbb{R}^N \,\middle|\, x \geq 0,\ \mathbf{A}x = b \right\} \tag{2.33}$$

where if $x = (x^1, x^2, \dots, x^N)$, $x \geq 0$ stands for $x^i \geq 0\ \forall i = 1, \dots, N$. In this case, K is a convex polyhedral closed set of $\mathbb{R}^n$ that does not contain straight lines, hence, see Theorem 2.23, K does have extreme points. They are characterized as follows.

2.60 Definition. *Let K be as in* (2.33)*. We say that $x \in K$ is a* base point *of K if either $x = 0$ (in this case $0 \in K$) or, if $\alpha_1, \alpha_2, \dots, \alpha_k$ are the indices of the nonzero components of x, the columns $A_{\alpha_1}, \dots A_{\alpha_k}$ of $\mathbf{A}$ are linearly independent.*

2.61 Theorem. *Let K be as in* (2.33)*. Extreme points of K are all and only the base points of K.*

Proof. Clearly, if $0 \in K$, then 0 is an extreme point of K. Suppose that $x = (x^1 \dots, x^k, 0, \dots, 0) \in K$, $x^i > 0\ \forall i = 1, \dots, k$, is a base point for K, and, contrary to the claim, x is not an extreme point for K. Then there are $y, z \in K$, $y \neq z$, such that $x = (y + z)/2$. Since $x, y, z \in K$, it would follow that $y = (y^1, y^2, \dots, y^k, 0, \dots, 0)$, $z = (z^1, z^2, \dots, z^k, 0, \dots, 0)$ and $b = \sum_{i=1}^k y^i A_i = \sum_{i=1}^k z^i A_i$. Since $A_1, A_2, \dots, A_k$ are linearly independent, we would then have $y = z$, a contradiction.

Conversely, suppose that x is a nonzero extreme point of K and that $x = (x^1, x^2, \dots, x^k, 0, \dots, 0)$ with $x^i > 0\ \forall i = 1, \dots, k$. Then

$$x^1 A_1 + \dots + x^k A_k = b.$$

We now infer that $A_1, A_2, \dots, A_k$ are linearly independent. Suppose they are not independent, i.e., there is a nonzero $y = (y^1, y^2, \dots, y^k, 0, \dots, 0)$ such that

$$y^1 A_1 + \dots + y^k A_k = 0.$$

Now we choose $\theta > 0$ in such a way that $u := x + \theta y$ and $v := x - \theta y$ still have nonnegative coordinates and $u, v \in K$. Then $x = (u + v)/2$, $u \neq v$, and x would not be an extreme point. $\square$

2.62 Remark. Actually, Theorem 2.61 provides us with an algorithm for computing the extreme points of a polyhedral convex set as base points. Since base points correspond to a choice of linearly independent columns, Theorem 2.61 shows that K has finitely many extreme points.

The next proposition shows the existence of a base point without any reference to the convex set theory. We include it for the reader's convenience.

2.63 Proposition. *Let $K \neq \emptyset$ be as in* (2.33). *Then K has at least one base point.*

Proof. Of course, there is a point x with minimum, say k, nonzero components such that $\mathbf{A}x = b$ and no $x' \geq 0$ with $\mathbf{A}x' = b$ and number of components nonzero $< k$.

Let $\alpha_1, \dots, \alpha_k$ be the indices of nonzero components of x. We now prove that *the columns $A_{\alpha_1}, \dots, A_{\alpha_k}$ are linearly independent*, i.e., that x is a base point of K. Suppose they are not independent, i.e.,

$$\sum_{i=1}^{k} \theta_i A_{\alpha_i} = 0$$

where $\theta_1, \theta_2, \dots, \theta_k$ are not all zero. We may assume that at least one of the θ_i is positive. Then

$$b = \sum_{i=1}^{k} A_{\alpha_i} x_i = \sum_{i=1}^{k} A_{\alpha_i}(x_i - \lambda\theta_i)$$

for all $\lambda \in \mathbb{R}$. However, for

$$\lambda := \min\Big\{ \frac{x_i}{\theta_i} \,\Big|\, \theta_i > 0 \Big\} =: \frac{x_{i_0}}{\theta_{i_0}}$$

we have $x_{i_0} - \lambda\theta_{i_0} = 0$. It follows that $x' := x - \lambda\theta \geq 0$, $b = \mathbf{A}'x'$ and x' has a number of nonzero components less than k, a contradiction. □

c. Parametrized convex cones

Particularly useful are the *finite cones*, i.e., cones generated by finitely many points, $A_1, A_2, \dots, A_N \in \mathbb{R}^n$. They have the form

$$C := \Big\{ \sum_{i=1}^{N} x^i A_i \,\Big|\, x^i \geq 0,\ i = 1, \dots, N \Big\}$$

and with the notation rows by columns, they can be written in a compact form as

$$C := \{ y \in \mathbb{R}^n \,|\, y = \mathbf{A}x,\ x \geq 0 \}$$

where $\mathbf{A} \in M_{n,N}$ is the $n \times N$ matrix

$$\mathbf{A} = [A_1 \,|\, A_2 \,|\, \dots \,|\, A_N].$$

Trivially, a finite cone is a polyhedral set that does not contain straight lines, hence has extreme points. We say that a finite cone is a *base cone* if it is generated by linearly independent vectors.

2.64 Proposition. *Every finite cone C is convex, closed and contains the origin.*

Proof. Trivially, C is convex and contains the origin. so it remains to prove that C is closed. Let $\mathbf{A} \in M_{n,N}$ be such that $C = \Big\{y = \mathbf{A}x \,|\, x \geq 0\Big\}$. C is surely closed if $\mathbf{A}$ has linearly independent columns, i.e., if $\mathbf{A}$ is injective. In fact, in this case the map $x \to \mathbf{A}x$ has a linear inverse, hence it is a closed map and $C = \mathbf{A}(\{x \geq 0\})$. For the general case, consider the cones $C_1, \dots, C_k$ associated to the submatrices of $\mathbf{A}$ that have linearly independent columns. As we have already remarked $C_1, \dots, C_k$ are closed sets. We claim that

$$C = C_1 \cup C_2 \cup \cdots \cup C_k, \tag{2.34}$$

hence C is closed, too. In order to prove (2.34), observe that since every cone generated by a submatrix of $\mathbf{A}$ is contained in C, we have $C_i \subset C\ \forall i$. On the other hand, if $b \in C$, Proposition 2.63 yields a submatrix $\mathbf{A}'$ of $\mathbf{A}$ with linearly independent columns such that $b = \mathbf{A}'x'$ for some $x' \geq 0$, i.e., $b \in \cup_i C_i$. □

The following claims readily follow from the results of Paragraph a.

2.65 Corollary. *Let C_1 and C_2 be two finite cones in $\mathbb{R}^n$. Then*

(i) *if $C_1 \subset C_2$, then $C_2^* \subset C_1^*$,*
(ii) *$C_1^* \cup C_2^* = (C_1 \cap C_2)^*$,*
(iii) *$C_1 = C_1^{**}$.*

Finally, let us compute the polar set of a finite cone.

2.66 Proposition. *Let $C = \{\mathbf{A}x \,|\, x \geq 0\}$, $\mathbf{A} \in M_{n,N}(\mathbb{R})$. Then*

$$C^* = \Big\{\xi \,\Big|\, \mathbf{A}^T\xi \leq 0\Big\} \tag{2.35}$$

and

$$C^{**} := \Big\{x \,\Big|\, x \bullet \xi \leq 0\ \forall \xi \text{ such that } \mathbf{A}^T\xi \leq 0\Big\}. \tag{2.36}$$

Proof. Since C is a cone, we have

$$C^* = \Big\{\xi \,\Big|\, \xi \bullet b \leq 1\ \forall b \in C\Big\} = \Big\{\xi \,\Big|\, \xi \bullet b \leq 0\ \forall b \in C\Big\}.$$

Consequently,

$$C^* = \Big\{\xi \,\Big|\, \xi \bullet \mathbf{A}x \leq 0\ \forall x \geq 0\Big\} = \Big\{\xi \,\Big|\, \mathbf{A}^T\xi \bullet x \leq 0\ \forall x \geq 0\Big\} = \Big\{\xi \,\Big|\, \mathbf{A}^T\xi \leq 0\Big\}$$

and

$$\begin{aligned} C^{**} &= \Big\{x \,|\, x \bullet \xi \leq 1\ \forall \xi \in C^*\Big\} = \Big\{x \,|\, x \bullet \xi \leq 0\ \forall \xi \in C^*\Big\} \\ &= \Big\{x \,|\, x \bullet \xi \leq 0\ \forall \xi \text{ such that } \mathbf{A}^T\xi \leq 0\Big\}. \end{aligned}$$

□

d. Farkas–Minkowski's lemma

2.67 Theorem (Farkas–Minkowski). *Let $\mathbf{A} \in M_{n,N}(\mathbb{R})$ and $b \in \mathbb{R}^n$. One and only one of the following claims holds:*

(i) *$\mathbf{A}x = b$ has a nonnegative solution.*
(ii) *There exists a vector $y \in \mathbb{R}^n$ such that $\mathbf{A}^T y \geq 0$ and $y \bullet b < 0$.*

In other words, using the same notations as in Theorem 2.67, the claims

(i) x is a nonnegative solution of $\mathbf{A}x = b$,
(ii) if $\mathbf{A}^T y \leq 0$, then $y \bullet b \leq 0$

are equivalent.

Proof. The claim is a rewriting of the equality $C = C^{**}$ in the case of finite cones, and, ultimately, a direct consequence of the separation property of convex sets. Let $C := \{\mathbf{A}x \,|\, x > 0\}$. Claim (i) rewrites as $b \in C$, while, according to (2.36), claim (ii) rewrites as $b \notin C^{**}$. □

2.68 Example (Fredholm alternative theorem). The Farkas–Minkowski lemma, equivalently the equality $C = C^{**}$ for finite cones, can be also seen as a generalization of the Fredholm alternative theorem for linear maps: $\mathrm{Im}(\mathbf{A}) = (\ker \mathbf{A}^T)^\perp$. In fact, if $b = \mathbf{A}x$, $\mathbf{A} \in M_{n,N}$, and if we write $x = u - v$ with $u, v \geq 0$, the equation $\mathbf{A}x = b$ rewrites as

$$b = \begin{pmatrix} \boxed{\mathbf{A}} & \boxed{-\mathbf{A}} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}, \qquad u, v \geq 0.$$

Therefore, $b \in \mathrm{Im}\,\mathbf{A}$ if and only if the previous system has a nonnegative solution. This is equivalent to saying that the alternative provided by the Farkas lemma is not true; consequently,

$$\text{if} \quad \begin{pmatrix} \boxed{\mathbf{A}^T} \\ \boxed{-\mathbf{A}^T} \end{pmatrix} \xi \leq 0, \ \text{then} \ b \bullet \xi \leq 0$$

i.e.,

$$b \bullet \xi \leq 0 \text{ for all } \xi \text{ such that } \mathbf{A}^T \xi = 0$$

and, in conclusion,

$$b \bullet \xi = 0 \text{ for all } \xi \text{ such that } \mathbf{A}^T \xi = 0,$$

i.e., $b \in (\ker \mathbf{A}^T)^\perp$.

2.69 ¶. Let $\mathbf{A} \in M_{m,n}(\mathbb{R})$ and $b \in \mathbb{R}^m$ and let K be the closed convex set

$$K := \{x \in \mathbb{R}^n \,\big|\, \mathbf{A}x \geq b, \ x \geq 0\}.$$

Characterize the extreme points of K.
[*Hint.* Introduce the new variables, called *slack variables* $x' \geq 0$, so that the constraints $\mathbf{A}x \geq b$ become

$$\mathbf{A}' \begin{pmatrix} x \\ x' \end{pmatrix} = b, \qquad \mathbf{A}' := \begin{pmatrix} \boxed{\mathbf{A}} & \boxed{-\,\mathrm{Id}} \end{pmatrix}.$$

Set $K' := \{z \,\big|\, \mathbf{A}'z \geq b, \ z \geq 0\}$. Show that *$x$ is an extreme point for K if and only if $z := (x, x')$ with $x' := \mathbf{A}x - b$ is an extreme point for K'*.]

Figure 2.10. Gaspard Monge (1746–1818) and the frontispiece of the *Principes de la théorie des richesses* di Antoine Cournot (1801–1877).

2.70 ¶. Prove the following variants of the Farkas lemma.

Theorem. *Let $\mathbf{A} \in M_{n,N}(\mathbb{R})$ and $b \in \mathbb{R}^n$. One and only one of the following alternatives holds:*

- $\mathbf{A}x \geq b$ *has a solution* $x \geq 0$.
- *There exists* $y \leq 0$ *such that* $\mathbf{A}^T y \geq 0$ *and* $b \bullet y < 0$.

Theorem. *Let $\mathbf{A} \in M_{n,N}(\mathbb{R})$ and $b \in \mathbb{R}^n$. One and only one of the following alternatives holds:*

- $\mathbf{A}x \leq b$ *has a solution* $x \geq 0$.
- *There exists* $y \geq 0$ *such that* $\mathbf{A}^T y \geq 0$ *and* $b \bullet y < 0$.

[*Hint.* Introduce the slack variables, as in Example 2.68.]

2.4.5 Convex optimization

Let f and $\varphi^1, \varphi^2, \dots, \varphi^m : \mathbb{R}^n \to \mathbb{R}$ be functions of class C^1. Here we discuss the constrained *minimum problem*

$$f(x) \to \min \qquad \text{in} \qquad \mathcal{F} := \Big\{ x \in \mathbb{R}^n \,\Big|\, \varphi^j(x) \leq 0,\ j = 1, \dots, m \Big\} \tag{2.37}$$

and, in particular, we present necessary and sufficient conditions for its solvability, compare also Section 4.

Let $\varphi := (\varphi^1, \dots, \varphi^m) : \mathbb{R}^n \to \mathbb{R}^m$ and let x_0 be a minimum point for f in $\mathcal{F}$. If $\varphi^j(x_0) < 0\ \forall j$, $\varphi(x_0) < 0$ for short, then x_0 is interior to $\mathcal{F}$ and Fermat's theorem implies $\mathbf{D}f(x_0) = 0$. If $\varphi(x_0) = 0$, then x_0 is a minimum point constrained to $\partial\mathcal{F} := \{x \in \mathbb{R}^n \,|\, \varphi(x) = 0\}$. Consequently, if the Jacobian matrix $\mathbf{D}\varphi(x_0)$ has maximal rank so that $\partial\mathcal{F}$ is a regular submanifold in a neighborhood of x_0, we have

$$\mathbf{D}f(x_0)(v) = 0 \qquad \forall v \in \mathrm{Tan}_{x_0}\,\partial\mathcal{F},$$

i.e.,

$$\nabla f(x_0) \perp \mathrm{Tan}_{x_0}\,\partial\mathcal{F},$$

and, from Lagrange's multiplier theorem (or Fredholm's alternative theorem) we infer the existence of a vector $\lambda^0 = (\lambda^0_1, \dots, \lambda^0_m) \in \mathbb{R}^m$ such that

$$\mathbf{D}f(x^0) = \sum_{j=1}^m \lambda^0_j \mathbf{D}\varphi^j(x^0).$$

In general, it may happen that $\varphi^j(x^0) = 0$ for some j and $\varphi^j(x^0) < 0$ for the others. For $x \in \mathcal{F}$, denote by $J(x)$ the set of indices j such that $\varphi^j(x) = 0$. We say that the constraint φ^j is *active* at x if $j \in J(x)$.

2.71 Definition. *We say that a vector $h \in \mathbb{R}^n$ is an* admissible direction *for $\mathcal{F}$ at $x \in \mathcal{F}$ if there exists a sequence $\{x^k\} \subset \mathcal{F}$ such that*

$$x_k \neq x \;\forall k,\; x_k \to x \text{ as } k \to \infty \qquad \text{and} \qquad \frac{x_k - x}{|x_k - x|} \to \frac{h}{|h|}.$$

The set of the admissible directions for $\mathcal{F}$ at x is denoted by $\Gamma(x)$. It is easily seen that $\Gamma(x)$ is a closed cone not necessarily convex. Additionally, it is easy to see that $\Gamma(x)$ is the set of directions $h \in \mathbb{R}^n$ for which there is a regular curve $r(t)$ in $\mathcal{F}$ with $r(0) = x$ and $r'(0) = h$.

Denote by $\widetilde{\Gamma}(x)$ the cone with vertex at zero, this time convex, of the directions that "point to $\mathcal{F}$",

$$\widetilde{\Gamma}(x) := \Big\{ h \in \mathbb{R}^n \,\Big|\, \nabla\varphi^j(x) \bullet h \le 0 \;\forall j \in J(x) \Big\};$$

it is not difficult to prove that $\Gamma(x) \subset \widetilde{\Gamma}(x)$.

2.72 Definition. *We say that the constraints are* qualified *at $x \in \mathcal{F}$ if $\Gamma(x) = \widetilde{\Gamma}(x)$.*

Not always are the constraints qualified, see Example 2.76. The following proposition gives a sufficient condition which ensures that the constraints are qualified.

2.73 Proposition. *Let $\varphi = (\varphi^1, \varphi^2, \dots, \varphi^m) : \mathbb{R}^n \to \mathbb{R}^m$ be of class C^1, $\mathcal{F} := \{x \in \mathbb{R}^n \,|\, \varphi(x) \le 0\}$ and $x_0 \in \mathcal{F}$. If there exists $\overline{h} \in \mathbb{R}^n$ such that for all $j \in J(x_0)$ we have*

(i) *either* $\nabla\varphi^j(x_0) \bullet \overline{h} < 0$

(ii) *or* φ^j *is affine and* $\nabla\varphi^j(x_0)\bullet\overline{h} \le 0$,

then the constraints $\{\varphi^j\}$ *are qualified at* x_0. *Consequently, the constraints are qualified at* x_0 *if one of the following conditions holds:*

(i) *There exists* $\overline{x} \in \mathcal{F}$ *such that* $\forall j \in J(x_0)$, *either* φ^j *is convex and* $\varphi^j(\overline{x}) < 0$, *or* φ^j *is affine and* $\varphi^j(\overline{x}) \le 0$.

(ii) *The vectors* $\nabla\varphi^j(x_0)$, $j \in J(x_0)$, *are linearly independent.*

Proof. Step 1. Let us prove that $\widetilde{\Gamma}(x_0) \subset \Gamma(x_0)$. Let h be such that $\nabla\varphi^j(x_0)\bullet h \le 0$. We claim that for every $\delta > 0$ we have $h + \delta\overline{h} \in \Gamma(x_0)$, thus concluding that $h \in \Gamma(x_0)$, $\Gamma(x_0)$ being closed.

Choose a positive sequence $\{e_k\}$ such that $\epsilon_k \to 0$ and consider the sequence $\{x_k\}$ defined by $x_k := x_0 + \epsilon_k(h + \delta\overline{h})$. Trivially $x_k \to x_0$ and $\frac{x_k - x_0}{|x_k - x_0|} = \frac{h+\delta\overline{h}}{|h+\delta\overline{h}|}$, thus $h + \delta\overline{h} \in \Gamma(x_0)$ if we prove that $x_k \in \mathcal{F}$ for k large. Let $j \in J(x_0)$. If $\nabla\varphi^j(x_0)\bullet\overline{h} < 0$, then

$$\nabla\varphi^j(x_0)\bullet(h + \delta\overline{h}) < 0$$

and, since

$$\varphi^j(x_k) = \varphi^j(x_0) + \epsilon_k\,\nabla\varphi^j(x_0)\bullet(h + \delta\overline{h}) + o(\epsilon_k),$$

we conclude that $\varphi^j(x_k) < 0$ for k large. If φ^j is affine and $\nabla\varphi^j(x_0)\bullet\overline{h} \le 0$, then

$$\varphi^j(x_k) = \varphi^j(x_0) + \epsilon_k\,\nabla\varphi^j(x_0)\bullet h + \delta\overline{h} \le 0.$$

Step 2. Let us now prove the second part of the claim. Let $\overline{h} := \overline{x} - x_0$ and $j \in J(x_0)$. If φ^j is convex, we have

$$\nabla\varphi^j(x_0)\bullet\overline{h} \le \varphi(\overline{x}) < 0,$$

whereas if φ^j is affine, we have

$$\nabla\varphi^j(x_0)\bullet\overline{h} = \varphi(\overline{x}) \le 0.$$

Therefore, (i) follows from Step 1.

We now assume that $J(x_0) = \{1, 2\dots, p\}$, $1 \le p \le n$, and let $\varphi := (\varphi^1, \dots, \varphi^p)$. Let $b := (-1, -1, \dots, -1) \in \mathbb{R}^p$. Then the linear system $\mathbf{D}\varphi(x_0)\overline{x} = b$, $x \in \mathbb{R}^n$ is solvable since Rank $\mathbf{D}\varphi(x_0) = p$. If $\overline{h}$ is any such solution, then $\nabla\varphi^j(x_0)\bullet\overline{h} = \mathbf{D}\varphi(x_0)\overline{h} = -1$ for all $j \in J(x_0)$, and (ii) follows from Step 1. □

2.74 Theorem (Kuhn–Tucker). *Let* x_0 *be a solution of* (2.37). *Suppose that the constraints are qualified at* x_0. *Then the following* Kuhn–Tucker equilibrium condition *holds: For all* $j \in J(x_0)$ *there exists* $\lambda_j^0 \ge 0$ *such that*

$$\nabla f(x_0) + \sum_{j\in J(x_0)} \lambda_j^0 \nabla\varphi^j(x_0) = 0. \tag{2.38}$$

Theorem 2.74 is a simple application of the following version of the Farkas lemma.

2.75 Lemma (Farkas). *Let* v *and* $v_1, v_2, \dots, v_p$ *be vectors of* $\mathbb{R}^n$. *There exist* $\lambda_j \ge 0$ *such that*

$$v = \sum_{j=1}^p \lambda_j v_j \tag{2.39}$$

if and only if

$$\Big\{h \in \mathbb{R}^n \,\Big|\, h\bullet v_j \le 0,\ \forall j = 1, \dots, p\Big\} \subset \Big\{h \in \mathbb{R}^n \,\Big|\, h\bullet v \le 0\Big\}. \tag{2.40}$$

Proof. In fact, if $\mathbf{A} := [v_1|v_2|\dots|v_n]$, (2.39) states that $\mathbf{A}\lambda = v$ has a nonnegative solution $\lambda \geq 0$. This is equivalent to saying that the second alternative of the Farkas lemma is false, i.e., $\forall h \in \mathbb{R}^n$ such that $\mathbf{A}^T h \geq 0$, we have $h \bullet v \geq 0$, that is, if $h \in \mathbb{R}^n$ satisfies $h \bullet v_j \geq 0$ for all j, then $h \bullet v \leq 0$. This is precisely (2.40). □

Proof of Theorem 2.74. For any $h \in \Gamma(x_0)$, let $r : [0,1] \to \mathcal{F}$ be a regular curve with $r(0) = x_0$ and $r'(0) = h$. Since 0 is a minimum point for $f(r(t))$, we have $\frac{d}{dt} f(r(t))_{|t=0} \geq 0$, i.e.,

$$-\mathbf{D}f(x_0) \bullet h \leq 0 \qquad \forall h \in \Gamma(x_0),$$

i.e., $h \in \left\{ h \in \mathbb{R}^n \,\middle|\, h \bullet v \leq 0 \right\}$. Recalling the definition of $\Gamma(x_0)$, the claim follows by applying Lemma 2.75 with $v := -\nabla f(x^0)$ and $v_j = \nabla \varphi^j(x^0)$. □

2.76 Example. Let $\mathcal{P}$ be the problem of minimizing $-x_1$ with the constraints $x_1 \geq 0$ and $x_2 \geq 0$, $(1-x_1)^3 - x_2 \geq 0$. Clearly the unique solution is $x^0 = (1,0)$. Show that the constraints are not qualified at x^0 and that the Kuhn–Tucker theorem does not hold.

2.77 Remark. In analogy with Lagrange's multiplier theorem we may rewrite the Kuhn–Tucker equilibrium conditions (2.38) as

$$\begin{cases} \mathbf{D}f(x^0) + \sum_{j=1}^m \lambda_j^0 \mathbf{D}\varphi^j(x^0) = 0, \\ \lambda_j^0 \geq 0 \qquad \forall j = 1, \dots, m, \\ \sum_{j=1}^m \lambda_j^0 \varphi^j(x^0) = 0, \end{cases}$$

or, using the vectorial notation,

$$\begin{cases} \mathbf{D}f(x^0) + \lambda^0 \bullet \mathbf{D}\varphi(x_0) = 0, \\ \lambda^0 \geq 0, \\ \lambda^0 \bullet \varphi(x_0) = 0, \end{cases} \tag{2.41}$$

where $\lambda^0 = (\lambda_1^0, \dots, \lambda_m^0) \in \mathbb{R}^m$ and $\varphi = (\varphi^1, \dots, \varphi^m) : \mathbb{R}^n \to \mathbb{R}^m$. In fact, the equation $\sum_{j=1}^m \lambda_j^0 \varphi^j(x^0) = 0$ implies $\lambda_h^0 = 0$ if the corresponding constraint φ^h is not active. If (2.41) holds for some λ^0, we call it a *Lagrange multiplier* of (2.37) at x_0.

2.4.6 Stationary states for discrete-time Markov processes

Suppose that a system can be in one of n possible states, denote by $p_j^{(k)}$ the probability that it is in the state j at the discrete time k and set $p^{(k)} := (p_1^{(k)}, p_2^{(k)}, \dots, p_n^{(k)})$. A homogeneous Markov chain with values in a finite set is characterized by the fact that the probabilities of the states at time $k+1$ are a linear function of the probabilities at time k and that such a function does not depend on k, that is, there is a $n \times n$ matrix $\mathbf{P} \in M_{n,n}(\mathbb{R})$ such that

$$p^{(k+1)} = p^{(k)}\mathbf{P} \qquad \forall k, \tag{2.42}$$

where the product is the usual row by column product of linear algebra.

The matrix $\mathbf{P} = (p_{ij})$ is called the *transition matrix*, or *Markov matrix* of the system.

Since $\sum_{j=1}^n p_j^{(k)} = 1$ for every k, the matrix $\mathbf{P}$ has to be *stochastic* or *Markovian*, meaning that

$$\mathbf{P} = (p_{ij}), \qquad \sum_{j=1}^n p_{ij} = 1, \qquad p_{ij} \geq 0.$$

According to (2.42), the evolution of the system is then described by the powers of $\mathbf{P}$,

$$p^{(k)} = p^{(0)}\,\mathbf{P}^k \qquad \forall k. \tag{2.43}$$

A *stationary state* is a fixed point of $\mathbf{P}$ i.e., $x \in \mathbb{R}^n$ such that

$$x = \mathbf{P}^T x, \qquad \sum_{j=1}^n x^j = 1, \qquad x \geq 0. \tag{2.44}$$

The Perron–Frobenius theorem, see [GM3], ensures the existence of a stationary state.

2.78 Theorem (Perron–Frobenius). *Every Markov matrix has a stationary state.*

Proof. This is just a special case of the fact that every continuous map from a compact convex set into itself has a fixed point, see [GM3]. However, since here we deal with a linear map $x \to \mathbf{P}x$, we give a direct proof which uses compactness.

Let $S := \{x \in \mathbb{R}^n \,|\, x \geq 0, \sum_{j=1}^n x^j = 1\}$. S is a convex closed and bounded set of $\mathbb{R}^n$, and $\mathbf{P}$ maps S into S and is stochastic. Fix $x_0 \in S$ and consider the sequence $\{x_k\}$ given by

$$x_k := \frac{1}{k}\sum_{i=0}^{k-1} x_0\mathbf{P}^i.$$

x_k is a convex combination of points in S and therefore $x_k \in S$. The sequence $\{x_k\}$ is then bounded and, by the Bolzano–Weierstrass theorem, there exists a subsequence $\{x_{n_k}\}$ of $\{x_k\}$ and $x \in S$ such that $x_{n_k} \to x$. On the other hand, for any k we have

$$x_k - x_k\mathbf{P} = \frac{1}{k}\Big(\sum_{i=0}^{k-1} x_0\mathbf{P}^i - \sum_{i=0}^{k-1} x_0\mathbf{P}^{i+1}\Big) = \frac{1}{k}(x_0 - x_0\mathbf{P}^{k+1})$$

so that

$$|x_k - x_k\mathbf{P}| \leq \frac{1}{k}.$$

Passing to the limit along the subsequence $\{x_{k_n}\}$, we then get $x - x\mathbf{P} = 0$. □

Another proof of Theorem 2.78. We give another proof of this claim which uses only convexity arguments, in particular, the Farkas–Minkowski theorem. Let $\mathbf{P}$ be a stochastic $n \times n$ matrix. Define

$$u := (1, 1, \dots, 1) \in \mathbb{R}^n, \qquad b := (0, 0, \dots, 0, 1) \in \mathbb{R}^{n+1}$$

and

$$\mathbf{A} = \begin{pmatrix} \boxed{\mathbf{P}^T - \mathrm{Id}} \\ u^T \end{pmatrix} \qquad \text{in } M_{(n+1),n}(\mathbb{R}).$$

The existence of a stationary point x for $\mathbf{P}$ is then equivalent to

$$\mathbf{A}x = b \qquad \text{has a nonnegative solution } x \geq 0. \tag{2.45}$$

Now, we show that Farkas's alternative does not hold, i.e., the system $\mathbf{A}^T y \geq 0$, $b \bullet y < 0$ has no solution. Suppose it holds; then there is a y such that $b \bullet y = y_{n+1} < 0$. If we write y as $y = (z^1, z^2, \dots, z^n, -\lambda) =: (z, -\lambda)$, $\lambda > 0$, we then have

$$0 \leq \mathbf{A}^T y = y^T \mathbf{A} = (z, -\lambda) \begin{pmatrix} \boxed{\mathbf{P}^T - \mathrm{Id}} \\ u^T \end{pmatrix} = z(\mathbf{P}^T - \mathrm{Id}) - \lambda u^T,$$

i.e.,

$$z^T(\mathbf{P}^T - \mathrm{Id}) \geq \lambda u^T.$$

Thus

$$\sum_{j=1}^n z^j p_{ji} - z^i \geq \lambda > 0 \qquad \forall i = 1, \dots, n. \tag{2.46}$$

On the other hand, if m is the index such that $z^m = \max_j z^j$, we have

$$\sum_{j=1}^n z^j p_{jm} \leq \max_j z^j = z^m,$$

hence

$$\sum_{j=1}^n z^j p_j^m - z^m \leq 0,$$

and this contradicts (2.46). □

2.4.7 Linear programming

We shall begin by illustrating some classical examples.

2.79 Example (Investment management). A bank has 100 million dollars to invest: a part L in loans at a rate, say, of 10% and a part S in bonds, say at 5%, with the aim of maximizing its profits $0.1L + 0.05S$. Of course, the bank has trivial restrictions, $L \geq 0$, $S \geq 0$ and $L + S \leq 100$, but also needs some cash of at least 25% of the total amount, $S \geq 0.25(L + S)$, i.e., $3S \geq L$ and needs to satisfy requests for important clients which on average require 30 million dollars, i.e., $L \geq 30$. The problem is then

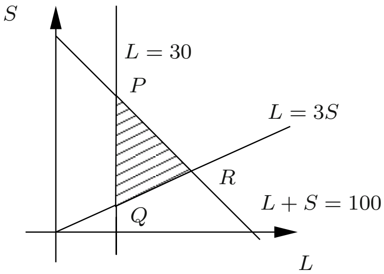


Figure 2.11. Illustration for Example 2.79.

$$\begin{cases} 0.10L + 0.05S \to \max, \\ L + S \leq 100, \; L \leq 3S, \; L \geq 30, \\ L \geq 0, \; S \geq 0. \end{cases}$$

With reference to Figure 2.11, the shaded triangle represent the admissible values (L, S); on the other hand, the gradient of the objective function $C = 0, 1L + 0.05S$ is constant $\nabla C = (0.1, 0.05)$ and the level lines of C are straight lines. Consequently, the optimal portfolio is to be found among the extreme points P, Q and R of the triangle, and, as it is easy to verify, the optimal configuration is in R.

2.80 Example (The diet problem). The daily diet of a person is composed of a number of components $j = 1, \dots, n$. Suppose that component j has a unitary cost c_j and contains a quantity a_{ij} of the nourishing i, $i = 1, \dots, m$, that is required in a daily quantity b_i. We want to minimize the cost of the diet. With standard vectorial notation the problem is

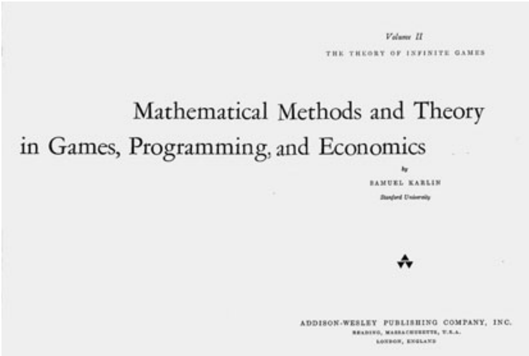
$$c \bullet x \to \min \qquad \text{in} \qquad \Big\{x \,\Big|\, \mathbf{A}x \geq b, x \geq 0\Big\}.$$

2.81 Example (The transportation problem). Suppose that a product (say oil) is produced in quantity s_i at places $i = 1, 2, \dots, n$ (Arabia, Venezuela, Alaska, etc.) and is requested at the different markets j, $j = 1, 2, \dots, m$ (New York, Tokyo, etc.) in quantity d_j. If c_{ij} is the transportation cost from i to j, we want to minimize the cost of transportation taking into account the constraints. The problem is then finding $x = (x_{ij}) \in \mathbb{R}^{nm}$ such that

$$\begin{cases} \sum_{i,j} c_{ij} x_{ij} \to \min, \\ \sum_{i=1}^{n} x_{ij} = d_j, \; \forall j, \\ \sum_{j=1}^{m} x_{ij} \leq s_i \; \forall i, \\ x \geq 0. \end{cases}$$

Here x is a vector with real-valued components, but for other products, for instance cars, the unknown would be a vector with integral components.

2.82 Example (Maximum profit). Suppose we are given $s_1, \dots, s_n$ quantities of basic products (resources) from which we may produce goods that sell at prices $p_1, p_2, \dots, p_m$. If a_{ij} is the quantity of product i, $i = 1, \dots, n$, to produce j, $j = 1, \dots, m$, our problem is finding the quantities x_j of goods j in order to maximize profits, i.e.,


Volume II

THE THEORY OF INFINITE GAMES

Mathematical Methods and Theory in Games, Programming, and Economics

by

SAMUEL KARLIN

Stanford University

ADDISON-WESLEY PUBLISHING COMPANY, INC.

READING, MASSACHUSETTS, U.S.A.

LONDON, ENGLAND

Figure 2.12. A classical textbook on linear programming and economics.

$$\begin{cases} \sum_{j=1}^m p_j x_j \to \max, \\ \sum_{j=1}^n a_{ij} x_j \le s_i, \\ x \ge 0. \end{cases}$$

In the previous examples, one wants to minimize or maximize a function, called the *objective function*, which is linear, in a set of *admissible* or *feasible* solutions, defined by a finite number of constraints defined by linear equalities or inequalities: This is the generic problem of *linear programming*. By possibly changing the sign of the objective function and/or of the inequalities constraints, observing that an equality constraint is equivalent to two inequalities constraints and replacing the variable x whose components are not necessarily nonnegative with $x = u - v$, $u, v \ge 0$, the linear programming problem can always be transformed into

$$f(x) := c \bullet x \to \min \qquad \text{in} \qquad \mathcal{P} := \Big\{ x \,\Big|\, \mathbf{A}x \ge b,\ x \ge 0 \Big\}, \tag{2.47}$$

where $c, x \in \mathbb{R}^n$, $\mathbf{A} \in M_{m,n}$ and $b \in \mathbb{R}^m$.

One of the following situations may, in principle, happen to hold:

(i) $\mathcal{P}$ is empty,

(ii) $\mathcal{P}$ is nonempty and the objective function in not bounded from below on $\mathcal{P}$,

(iii) $\mathcal{P}$ is nonempty and f is bounded from below.

In the last case, f has (at least) a minimizer and all the minimizers are extreme points of the convex set $\mathcal{P}$ by Proposition 2.42. We say that problem (2.47) has an *optimal solution.*

The problem transforms then into the problem of deciding in which of the previous cases we find ourselves and of possibly finding the optimal extreme points. In the real applications, where the number of constraints may be quite high, the effectiveness of the algorithm is also a further problem. Giving up efficiency, we approach the first two problems as follows.

We introduce the slack variables $x' := \mathbf{A}x - b \geq 0$ and transform the constraint $\mathbf{A}x \geq b$ into

$$\mathbf{A}' \begin{pmatrix} x \\ x' \end{pmatrix} = b, \qquad \mathbf{A}' := \left(\boxed{\ \mathbf{A}\ } \quad \boxed{-\operatorname{Id}} \right).$$

Writing $z = (x, x')$ and $F(z) := \sum_{i=1}^n c^i x^i + \sum_{i=1}^m 0 \cdot x'_i$, problem (2.47) transforms into

$$F(z) \to \min \qquad \text{in} \qquad \mathcal{F} := \Big\{ z \,\Big|\, \mathbf{A}'z = b,\ z \geq 0 \Big\}. \tag{2.48}$$

It is easily seen that $\mathcal{F}$ is nonempty if $\mathcal{P}$ is nonempty and that F is bounded from below on $\mathcal{F}$ if and only if f is bounded from below on $\mathcal{P}$. Therefore, F attains its minimum in one of the extreme points of $\mathcal{F}$ if and only if f has a minimizer in $\mathcal{P}$. All extreme points of $\mathcal{F}$ can be found by means of Theorem 2.61; the minimizers are then detected by comparison.

a. The primal and dual problem

Problem (2.47) is called the *primal* problem of linear programming, since one also introduces the *dual problem* of linear programming as

$$g(y) := b \bullet y \to \max \qquad \text{in} \qquad \mathcal{P}^* = \Big\{ y \,\Big|\, \mathbf{A}^T y \leq c,\ y \geq 0 \Big\}. \tag{2.49}$$

Of course, (2.49) can be rephrased as the minimum problem

$$h(y) := -\, b \bullet y \to \min \qquad \text{in} \qquad \mathcal{P}^* = \Big\{ y \,\Big|\, -\mathbf{A}^T y \geq -c,\ y \geq 0 \Big\} \tag{2.50}$$

which is similar to (2.47): Just exchange $-b$ and c, and replace $\mathbf{A}$ with $-\mathbf{A}^T$, and the following holds.

2.83 Proposition. *The dual problem of linear programming* (2.49) *has a solution if and only if $\mathcal{P}^* \neq \emptyset$ and g is bounded from above.*

The next theorem motivates the notation primal and dual problems of linear programming.

2.84 Theorem (Kuhn–Tucker equilibrium conditions). *Let f and $\mathcal{P}$ be as in* (2.47) *and let g and $\mathcal{P}^*$ be as in* (2.49)*. We have the following:*

(i) *$g(y) \leq f(x)$ for all $x \in \mathcal{P}$ and all $y \in \mathcal{P}^*$.*
(ii) *f has a minimizer $\overline{x} \in \mathcal{P}$ if and only if g has a maximizer $\overline{y} \in \mathcal{P}^*$ and, in this case, $f(\overline{x}) = g(\overline{y})$.*
(iii) *Let $x \in \mathcal{P}$ and $y \in \mathcal{P}^*$. The following claims are equivalent:*
   a) $(c - \mathbf{A}^T y) \bullet x = 0$.
   b) $(\mathbf{A}x - b) \bullet y = 0$.
   c) $f(x) = g(y)$.
   d) x is a minimizer for f and $y \in \mathcal{P}^*$ is a maximizer for g.

Proof. If $x \in \mathcal{P}$, then $x \geq 0$ and $\mathbf{A}x \geq b$. For $y \in \mathcal{P}^*$ we then get

$$f(x) = x \bullet c \geq x \bullet \mathbf{A}^T y = \mathbf{A}x \bullet y \geq b \bullet y = g(y),$$

i.e., (i).

(ii) Let $\overline{x}$ be a minimizer for the primal problem. Then f is bounded from below. We introduce the slack variables $x' = \mathbf{A}x - b \geq 0$ and set $z = (x, x')$. Then $\overline{x}$ is a solution of the primal problem (2.47) if and only if $\overline{z} := (\overline{x}, \overline{x'})^T$ minimizes

$$F(z) := c \bullet x \qquad \text{in} \qquad \mathcal{F} := \Big\{ z \,\Big|\, \mathbf{A}'z = b,\ z \geq 0 \Big\}$$

where

$$\mathbf{A}' := \Big(\boxed{\quad \mathbf{A} \quad} \quad \boxed{\quad \mathrm{Id} \quad} \Big).$$

We may also assume that $\overline{z}$ is an extreme point of $\mathcal{F}$. As we saw in the proof of Theorem 2.61, if $\alpha_1, \alpha_2, \dots, \alpha_k$ are the indices of the nonzero components of $\overline{z}$, the submatrix $\mathbf{B}$ of $\mathbf{A}'$ made of the columns of indices $\alpha_1, \alpha_2, \dots, \alpha_k$ has maximal rank. If x_B denotes the vector with components the nonzero components of x, then $\mathbf{B}x_B = b$, and if we set $c_B := (c_{\alpha_1}, c_{\alpha_2}, \dots, c_{\alpha_k})$ and choose $\overline{y}$ such that $\mathbf{B}^T\overline{y} = c_B$, we have

$$g(\overline{y}) = \overline{y} \bullet b = \overline{y} \bullet \mathbf{B}x_B = \overline{\mathbf{B}^T\overline{y}}x_B = c_B \bullet x_B = f(\overline{x}).$$

Then (i) yields that $\overline{y}$ is a maximizer of the dual problem.

(iii) (a) or (b) $\Rightarrow$ (c). If $(c - \mathbf{A}^T y) \bullet x = 0$ with $x \in \mathcal{P}$ and $y \in \mathcal{P}^*$, then

$$f(x) = c \bullet x = \mathbf{A}^T y \bullet x = y \bullet \mathbf{A}x \leq b \bullet y = g(y),$$

thus $f(x) = g(y)$ because of (i).

(c) $\Rightarrow$ (a) and (b). If $f(x) = g(y)$ and we set $\gamma := b - \mathbf{A}x$, we have

$$0 = f(x) - g(y) = c \bullet x - b \bullet y = c \bullet x - \mathbf{A}x \bullet y + \gamma \bullet y = (c - \mathbf{A}^T y) \bullet x + \gamma \bullet y.$$

Since the addenda are nonnegative, we conclude

$$(c - \mathbf{A}^T x) \bullet x = 0 \qquad \text{and} \qquad (\mathbf{A}x - b) \bullet y = 0.$$

(c) $\Rightarrow$ (d). If $f(x) = g(y)$, then (i) yields $f(x') \geq g(y) = f(x)$ for all $x' \in \mathcal{P}$, hence x is a minimizer of f. Similarly y is a maximizer of g in $\mathcal{P}^*$.

(d) $\Rightarrow$ (c). This follows trivially from (ii). □

A consequence of the previous theorem is the following *duality theorem of linear programming.*

2.85 Corollary (Duality theorem). *Let* (2.47) *and* (2.49) *be the primal and the dual problems of linear programming. One and only one of the following alternatives arises:*

(i) *There exist a minimizer $\overline{x} \in \mathcal{P}$ for f and a maximizer $\overline{y} \in \mathcal{P}^*$ for g and $f(\overline{x}) = g(\overline{y})$. This arises if and only if $\mathcal{P}$ and $\mathcal{P}^*$ are both nonempty.*

(ii) *$\mathcal{P} \neq \emptyset$ and f is not bounded from below in $\mathcal{P}$.*

(iii) *$\mathcal{P}^* \neq \emptyset$ and g is not bounded from above in $\mathcal{P}^*$.*

(iv) *$\mathcal{P}$ and $\mathcal{P}^*$ are both empty.*

Proof. Trivially, (iv) is inconsistent with any of (i), (ii) or (iii); (iii) is inconsistent with (ii) because of (i) of Theorem 2.84, and (iii) is inconsistent with (i). Similarly (ii) is inconsistent with (i). Therefore, the four alternatives are disjoint. If (ii), (iii) and (iv) do not hold, we therefore have

$$\begin{cases} \mathcal{P} = \emptyset \text{ or } (\mathcal{P} \neq \emptyset \text{ and } f \text{ is bounded from below}),\\ \mathcal{P}^* = \emptyset \text{ or } (\mathcal{P}^* \neq \emptyset \text{ and } g \text{ is bounded from above}),\\ \mathcal{P} \text{ or } \mathcal{P}^* \text{ are nonempty,} \end{cases}$$

that is, one of the following alternatives holds:

$$\mathcal{P} \neq \emptyset \text{ and } f \text{ is bounded from below,}$$
$$\mathcal{P}^* \neq \emptyset \text{ and } g \text{ is bounded from above,}$$
$$\mathcal{P} \neq \emptyset,\ \mathcal{P}^* \neq \emptyset,\ f \text{ is bounded from below and } g \text{ is bounded from above.}$$

In any case, both the primal and the dual problem of linear programming have solutions and, according to (iii) of Theorem 2.84, the alternative (i) holds. $\square$

Corollary 2.85 is actually a *convex duality* theorem: Here we supply a direct proof by duality, using Farkas's alternative.

A proof of Corollary 2.85 which uses convex duality. Set

$$\widehat{\mathbf{A}} := \begin{pmatrix} \boxed{-\mathbf{A}} & \boxed{0} \\ \boxed{0} & \boxed{\mathbf{A}^T} \\ c^T & -b \end{pmatrix}$$

and

$$\widehat{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \qquad \widehat{b} = \begin{pmatrix} -b \\ c \\ 0 \end{pmatrix}.$$

Then (i) is equivalent to

$$\widehat{\mathbf{A}}\widehat{x} \leq \widehat{b} \qquad \text{has a solution } \widehat{x} \geq 0.$$

Farkas's alternative then yields the following: If (i) does not hold, then there exists $\widehat{y} = (u, v, \lambda)$ such that

$$\begin{cases} \begin{pmatrix} u^T & v^T & \lambda \end{pmatrix} \begin{pmatrix} \boxed{-\mathbf{A}} & \boxed{0} \\ \boxed{0} & \boxed{\mathbf{A}^T} \\ c^T & -b^T \end{pmatrix} \geq 0, \\ \begin{pmatrix} u^T & v^T & \lambda \end{pmatrix} \begin{pmatrix} -b \\ c \\ 0 \end{pmatrix} < 0, \\ \begin{pmatrix} u^T & v^T & \lambda \end{pmatrix} \geq 0, \end{cases}$$

or, after a simple computation, the problem

$$\mathbf{A}u \geq \lambda\, b, \qquad \mathbf{A}^T v \leq \lambda\, c, \qquad c \bullet u \leq b \bullet v \tag{2.51}$$

has a solution (u, v, λ) with $u \geq 0$, $v \geq 0$ and $\lambda \geq 0$.

Now, we claim that $\lambda = 0$. In fact, if $\lambda \neq 0$, then $u/\lambda \in \mathcal{P}$, $v/\lambda \in \mathcal{P}^*$, consequently, $c \bullet u/\lambda < b \bullet v/\lambda$: a contradiction because of (i) of Theorem 2.84. Thus, (2.51) reduces to the following claim: The problem

$$\mathbf{A}u \geq 0, \qquad \mathbf{A}^T v \leq 0, \qquad c \bullet u < b \bullet v$$

has a solution (u, v) with $u \geq 0$ and $v \geq 0$.

We notice that the inequality $c \bullet u < b \bullet v$ implies that either $c \bullet u < 0$ or $b \bullet v > 0$ or both. In the case $c \bullet u < 0$, we have $\mathcal{P}^* = \emptyset$, since otherwise if $y \geq 0$ and $\mathbf{A}^T y \leq c$, then from $\mathbf{A}u \geq 0$, $u \geq 0$ we would infer $0 \leq y \bullet \mathbf{A}u = A^T y \bullet u \leq c \bullet u$, a contradiction. If, moreover, $\mathcal{P} = \emptyset$, the alternative (iv) holds; otherwise, if $x \in \mathcal{P}$, then $\mathbf{A}(x + \theta u) \geq b + \theta 0 = b$, $x + \theta u \geq 0$ for some $\theta \geq 0$, and $c \bullet x + \theta u = c \bullet x + \theta\, c \bullet u \to -\infty$ as $\theta \to +\infty$, that is, the alternative (ii) holds.

In the case $b \bullet v > 0$, as in the case $c \bullet u < 0$, we see that $\mathcal{P} = \emptyset$. If also $\mathcal{P}^* = \emptyset$, then (iv) holds; while, if there exists $y \in \mathcal{P}^*$, then $v + \theta y \in \mathcal{P}^*$ and $v + \theta y \to +\infty$ as $\theta \to +\infty$, and (iii) holds. □

2.86 Example. Let us illustrate the above discussing the dual of the transportation problem. Suppose that crude oil is extracted in quantities s_i, $i = 1, \dots, n$ in places $i = 1, \dots, n$ and is requested in the markets $j = 1, \dots, m$ in quantity d_j. Let c_{ij} be the transportation cost from i to j. The optimal transportation problem consists in determining the quantities of oil to be transported from i to j minimizing the overall transportation cost

$$\sum_{i,j} c_{ij} x_{ij} \to \min, \tag{2.52}$$

and satisfying the constraints, in our case, the markets requests and the capability of production

$$\begin{cases} \sum_{j=1}^m x_{ij} \leq s_i & \forall i, \\ \sum_{i=1}^n x_{ij} = d_j & \forall j, \\ x \geq 0. \end{cases} \tag{2.53}$$

Of course, a necessary condition for the solvability is that the production be larger than the markets requests

$$\sum_{j=1}^m d_j = \sum_{\substack{i=1,n \\ j=1,m}} x_{ij} \leq \sum_{i=1}^n s_i.$$

Introducing the matrix notation

$$\begin{cases} x := (x_{11}, \dots, x_{1m}, x_{21}, \dots, x_{2m}, \dots, x_{n1}, \dots, x_{nm}) \in \mathbb{R}^{nm}, \\ c := (c_{11}, \dots, c_{1m}, c_{21}, \dots, c_{2m}, \dots, c_{n1}, \dots, c_{nm}) \in \mathbb{R}^{nm}, \\ b := (s_1, s_2, \dots, s_n, d_1, \dots, d_m) \end{cases}$$

and setting $\mathbf{A} \in M_{n+m,nm}(\mathbb{R})$,

$$\mathbf{A} := \begin{pmatrix} u & 0 & 0 & \dots & 0 \\ 0 & u & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & u \\ \dots & & & & \\ e_1 & e_1 & e_1 & \dots & e_1 \\ e_2 & e_2 & e_2 & \dots & e_2 \\ \dots & & & & \\ e_m & e_m & e_m & \dots & e_m \end{pmatrix},$$

where $u := (1, 1, \dots, 1) \in \mathbb{R}^m$ and $0 = (0, 0, \dots, 0) \in \mathbb{R}^m$, we may formulate our problem as

$$\begin{cases} c \bullet x \to \min, \\ \mathbf{A}x \le b, \\ x \ge 0. \end{cases}$$

The dual problem is then

$$\begin{cases} b \bullet y \to \max, \\ \mathbf{A}^T y \le c, \\ y \ge 0, \end{cases}$$

that is, because of the form of $\mathbf{A}$ and setting

$$y := (u_1, u_2, \dots, u_n, v_1, v_2, \dots, v_m),$$

the maximum problem

$$\begin{cases} \sum_{i=1}^n s_i u_i + \sum_{j=1}^m d_j v_j \to \max, \\ u_i + v_j \le c_{ij} \qquad \forall i, j, \\ u \ge 0, \ v \ge 0. \end{cases}$$

If we interpret u_i as the toll at departure and v_i as the toll at the arrival requested by the shipping agent, the dual problem may be regarded as the problem of maximizing the profit of the shipping agent. Therefore, the quantities $\overline{u}_i$ and $\overline{v}_i$ which solve the dual problem represent the maximum tolls one may apply in order not to be out of the market.

2.87 Example. In the primal problem of linear programming one minimizes a linear function on a polyhedral set

$$\begin{cases} c \bullet x \to \min, \\ \mathbf{A}x \le b, \ x \ge 0, \end{cases}$$

or, equivalently,

$$\begin{cases} -c \bullet x \to \max, \\ \mathbf{A}x \le b, \ x \ge 0. \end{cases}$$

Since the constraint is qualified at all points, the primal problem has a minimum $x \ge 0$ if and only if the Kuhn–Tucker equilibrium condition holds, i.e., there exists $\lambda \ge 0$ such that

$$(c - \mathbf{A}^T \lambda)x = 0.$$

This way we find again the optimality conditions of linear programming.

2.4.8 Minimax theorems and the theory of games

The theory of games consists in mathematical models used in the study of processes of decisions that involve conflict or cooperation. The modern origin of the theory dates back to a famous paper by John von Neumann (1903–1957) published in German in 1928 with the title "On the Theory of Social Games"[3] and to the very well-known book by von Neumann and the

[3] J. von Neumann, Theorie der Gesellschaftsspiele, Math. Ann. **100** (1928) 295–320.

Figure 2.13. John von Neumann (1903–1957) and Oskar Morgenstern (1902–1976).

economist Oskar Morgenstern, *Theory of Games and Economic Behavior* published in 1944. There one can find several types of games with one or more players, with zero or nonzero sum, cooperative or non-cooperative, ... For its relevance in economy, social sciences or biology the theory has greatly developed[4]. Here we confine ourselves to illustrating only a few basic facts.

a. The minimax theorem of von Neumann

In a game with two players P and Q, each of them relays on a set of possible strategies, say respectively A and B; also, two *utility functions* $U_P(x, y)$ and $U_Q(x, y)$ are given, representing for each choice of the strategy $x \in A$ of P and $y \in B$ of Q the gain for P and Q resulting from the choices of the strategies x and y.

Let us consider the simplest case of a *zero sum game* in which the common value $K(x, y) := U_P(x, y) = -U_Q(x, y)$ is at the same time the gain for P and minus the gain for Q resulting from the choices of the strategies x and y.

[4] The interested reader is referred for classical literature to

- J. von Neumann, O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ, 1944, that follows a work of Ernst Zermelo (1871–1951), *Über eine Anwendung der Mengenlehre auf die Theorie des Schachspiels*, 1913 and a work of Emile Borel (1871–1956) *La théorie du jeu et les équations intégrales à noyau symmétrique*, 1921.
- R. Luce, H. Raiffa, *Games and Decisions: Introduction and Critical Survey*, Wiley, New York, 1957.
- S. Karlin, *Mathematical Methods and Theory in Games, Programming and Economics*, 2 vols., Addison–Wesley, Reading, MA, 1959.
- W. Lucas, An overview of the mathematical theory of games, Manage. Sci. **18** (1972), 3–19.
- M. Shubik, *Game Theory in the Social Sciences: Concepts and Solutions*, MIT Press, Boston, MA, 1982.

Each player tries to do his best against every strategy of the other player. In doing that, the *expected payoff* or, simply, *payoff*, i.e., the remuneration that P and Q can expect not taking into account the strategy of the other player, are

$$\begin{aligned}\mathrm{Payoff}(P) &:= \inf_{y\in B}\sup_{x\in A} U_P(x,y) = \inf_{y\in B}\sup_{x\in A} K(x,y),\\ \mathrm{Payoff}(Q) &:= \inf_{x\in A}\sup_{y\in B} U_Q(x,y) = \inf_{x\in A}\sup_{y\in B} -K(x,y) = -\sup_{x\in A}\inf_{y\in B} K(x,y).\end{aligned}$$

Although the game has zero sum, the payoffs of the two players are not related, in general, we trivially only have

$$\sup_{x\in A}\inf_{y\in B} K(x,y) \le \inf_{y\in B}\sup_{x\in A} K(x,y), \tag{2.54}$$

i.e.,

$$\mathrm{Payoff}(P) + \mathrm{Payoff}(Q) \ge 0.$$

Of course, if the previous inequality is strict, there are no choices of strategies that allow both players to reach their payoff.

The next proposition provides a condition for the existence of a couple of *optimal strategies*, i.e., of strategies that allow each players to reach their payoff.

2.88 Proposition. *Let A and B be arbitrary sets and $K : A\times B \to \mathbb{R}$. Define $f : A \to \mathbb{R}$ and $g : B \to \mathbb{R}$ respectively as*

$$f(x) := \inf_{y\in B} K(x,y), \qquad g(y) := \sup_{x\in A} K(x,y).$$

Then there exists $(\overline{x},\overline{y}) \in A\times B$ such that

$$K(x,\overline{y}) \le K(\overline{x},\overline{y}) \le K(\overline{x},y) \qquad \forall x,y \in A\times B \tag{2.55}$$

if and only if f attains its maximum in A, g attains its minimum in B and $\sup_{x\in A} f(x) = \inf_{y\in B} g(y)$. In this case,

$$\sup_{x\in A}\inf_{y\in B} K(x,y) = K(\overline{x},\overline{y}) = \inf_{y\in B}\sup_{x\in A} K(x,y).$$

Proof. If $(\overline{x},\overline{y})$ satisfies (2.55), then

$$\begin{aligned}K(\overline{x},\overline{y}) &= \inf_{y\in B} K(\overline{x},y) = f(\overline{x}) \le \sup_{x\in A} f(x),\\ K(\overline{x},\overline{y}) &= \sup_{x\in A} K(x,\overline{y}) = g(\overline{y}) \ge \inf_{y\in B} g(y),\end{aligned}$$

hence $\sup_{x\in A} f(x) = \inf_{y\in B} g(y)$ if we take into account (2.54). We leave the rest of the proof to the reader. □

A point $(\overline{x}, \overline{y})$ with property (2.55) is a *saddle point* for K. Therefore, in the context of games with zero sum, the saddle points of K yield couples of optimal strategies. The value of K on a couple of optimal strategies is called the *value of the play*. Answering the question of when there exists a saddle point is more difficult and is the content of the next theorem.

We recall that a function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *quasiconvex* if its sublevel sets are convex, and *quasiconcave* if $-f$ is quasiconvex.

2.89 Theorem (Minimax theorem of von Neumann). *Let $A \subset \mathbb{R}^n$ and $B \subset \mathbb{R}^n$ be two compact convex sets and let $K : A \times B \to \mathbb{R}$ be a function such that*

(i) *$x \to K(x, y)$ is quasiconvex and lower semicontinuous $\forall y \in B$,*
(ii) *$y \to K(x, y)$ is quasiconcave and upper semicontinuous $\forall x \in A$.*

Then K has a saddle point in $A \times B$.

Proof. According to Proposition 2.88 it suffices to prove that numbers

$$a := \min_{x\in A} \max_{y\in B} K(x, y) \qquad \text{and} \qquad b := \max_{y\in B} \min_{x\in A} K(x, y)$$

exist and are equal. Fix $y \in B$, the function $x \to K(x, y)$ attains its minimum at $z(y) \in A$ being A compact, and $K(z(y), y) = \min_{x\in A} K(x, y)$. Set

$$h(y) := -K(z(y), y), \qquad y \in B.$$

We now show that h is quasiconvex and lower semicontinuous, thus there is

$$b := -\min_{y\in B} \Big(-\min_{x\in A} K(x, y)\Big) = \max_{y\in B} \min_{y\in A} K(x, y).$$

Similarly, one proves the existence of a.

Let us show that h is quasiconvex and lower semicontinuous, that is, that for all $t \in \mathbb{R}$ the set

$$H := \Big\{y \in B \,\Big|\, h(y) \le t\Big\}$$

is convex and closed. First we will show that H is convex. For any $w \in B$, consider

$$G(w) := \Big\{y \in B \,\Big|\, -K(z(w), y) \le t\Big\}.$$

Because of (ii), $G(w)$ is convex and closed; moreover, $H \subset G(w)$ $\forall w$, since $K(z(y), y) \le K((z(w), y)$ $\forall w, y \in B$. In particular, for $x, y \in H$ and $\lambda \in]0, 1[$ we have $u \in G(w)$ $\forall w \in B$ if $u := (1-\lambda)y + \lambda x$, hence $u \in G(u)$, i.e., $u \in H$. This proves that H is convex. Let us prove now that H is closed. Let $\{y_n\} \subset H$, $y_n \to y$ in B, then $y \in G(w)$ $\forall w \in B$, in particular, $y \in G(y)$, i.e., $y \in H$. Therefore, H is closed.

Let us prove that $a = b$. Since $b \le a$ trivially, it remains to show that $a \le b$. Fix $\epsilon > 0$ and consider the function $T : A \times B \to \mathcal{P}(A \times B)$ given by

$$T(x, y) := \Big\{(u, v) \in A \times B \,\Big|\, K(u, y) < b + \epsilon,\ K(x, v) > a - \epsilon\Big\}.$$

We have $T(x, y) \ne \emptyset$ since $\min_{u\in A} K(u, y) \le b$ and $\max_{v\in B} K(x, v) \ge a$; moreover, $T(x, y)$ is convex. Since

$$\begin{aligned} T^{-1}(\{(u, v)\}) : &= \Big\{(x, y) \in A \times B \,\Big|\, (u, v) \in T(x, y)\Big\} \\ &= \Big\{(x, y) \in A \times B \,\Big|\, K(u, y) < b + \epsilon,\ K(x, v) > a - \epsilon\Big\} \\ &= \Big\{x \in A \,\Big|\, K(x, v) > a - \epsilon\Big\} \times \Big\{y \in B \,\Big|\, K(u, y) < b - \epsilon\Big\}, \end{aligned}$$

$T^{-1}(\{(u,v)\})$ is also open. We now claim, compare Theorem 2.90, that there is a fixed point for T, i.e., that there exists $(\overline{x},\overline{y}) \in A \times B$ such that $(\overline{x},\overline{y}) \in T(\overline{x},\overline{y})$, i.e., $a - \epsilon < k(\overline{x},\overline{y}) < b + \epsilon$. ϵ being arbitrary, we conclude $a \leq b$. □

For its relevance, we now state and prove the fixed point theorem we have used in the proof of the previous theorem.

2.90 Theorem (Kakutani). *Let K be a nonempty, convex and compact set, and let $F : K \to \mathcal{P}(K)$ be a function such that*

(i) *$F(x)$ is nonempty and convex for each $x \in K$,*
(ii) *$F^{-1}(y)$ is open in K for every $y \in \mathcal{P}(K)$.*

Then F has at least a fixed point, i.e., there exists $\overline{x}$ such that $\overline{x} \in F(\overline{x})$.

Proof. Clearly, the family of open sets $\{F^{-1}(y)\}_y$ is an open covering of K, consequently, there exist $y_1, y_2, \dots, y_n \in \mathcal{P}(K)$ such that $K \subset \cup_{i=1}^n F^{-1}(y_i)$. Let $\{\varphi_i\}$ be a partition of unity associated to $\{F^{-1}(y_i)\}_{i=1,\dots,n}$ and set

$$p(x) := \sum_{i=1}^n \varphi_i(x) y_i \qquad \forall x \in K_0 := \mathrm{co}(\{y_1, y_2, \dots, y_n\}) \subset K.$$

Obviously, p is continuous and $p(K_0) \subset K_0$. According to Brouwer's theorem, see [GM3], p has a fixed point $\overline{x} \in K_0$. To conclude, we now prove that $p(x) \in F(x)$ $\forall x \in K_0$, from which we infer that $\overline{x} = p(\overline{x}) \in F(\overline{x})$, i.e., $\overline{x}$ is a fixed point for F. Let $x \in K_0$. For each index j such that $\varphi_j(x) \neq 0$ we have trivially $x \in F^{-1}(y_j)$, thus $y_j \in F(x)$. Since $F(x)$ is convex, we see that

$$p(x) = \sum_{i=1}^n \varphi_i(x) y_i = \sum_{\{j \,|\, \varphi_j(x) \neq 0\}} \varphi_j(x) y_j,$$

hence $p(x) \in F(x)$. □

We now present a variant of Theorem 2.89.

2.91 Theorem. *Let $K : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $K = K(x,y)$, be a function convex in x for any fixed y and concave in y for any fixed x. Assume that there exist $\overline{x} \in \mathbb{R}^n$ and $\overline{y} \in \mathbb{R}^m$ such that*

$$\begin{aligned} K(x,\overline{y}) &\to +\infty \qquad \text{as } x \to +\infty, \\ K(\overline{x},y) &\to -\infty \qquad \text{as } y \to -\infty. \end{aligned}$$

Then K has a saddle point (x_0, y_0).

Observe that $K(x,y)$ is continuous in each variable. Let us start with a special case of Theorem 2.89 for which we present a more direct proof.

2.92 Proposition. *Let A and B be compact subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, and let $K : A \times B \to \mathbb{R}$, $K = K(x,y)$ be a function that is convex and lower semicontinuous in x for any fixed y and concave and upper semicontinuous in y for any fixed x. Then K has a a saddle point $(x_0, y_0) \in A \times B$.*

Proof. Step 1. Since $x \to K(x,y)$ is lower semicontinuous and A is compact, then for every $y \in B$ there exists at least one $x = x(y)$ such that

$$K(x(y),y) = \inf_{x\in A} K(x,y). \tag{2.56}$$

Let

$$g(y) := \inf_{x\in A} K(x,y) = K(x(y),y), \qquad y \in B. \tag{2.57}$$

The function g is upper semicontinuous, because $\forall y_0$ and $\forall \epsilon > 0$ there exists $\overline{x}$ such that

$$g(y_0) + \epsilon \geq K(\overline{x}, y_0) \geq \limsup_{y\to y_0} K(\overline{x}, y) \geq \limsup_{y\to y_0} g(y).$$

Consequently, there exists $y_0 \in B$ such that

$$g(y_0) := \max_{y\in B} g(y), \tag{2.58}$$

and, therefore,

$$g(y_0) \leq K(x, y_0) \qquad \forall x \in A. \tag{2.59}$$

Step 2. We now prove that for every $y \in B$ there exists $\widetilde{x}(y) \in A$ such that

$$K(\widetilde{x}(y), y) \leq g(y_0) \qquad \forall y \in B. \tag{2.60}$$

Fix $y \in B$. For $n = 1, 2, \dots$, let $y_n := (1-1/n)y_0 + (1/n)y$. Denote by $x_n := x(y_n)$, a minimizer of $x \mapsto K(x, y_n)$, i.e., $K(x_n, y_n) = \min_{x\in A} K(x, y_n) = g(y_n)$. Since $y \mapsto K(x,y)$ is concave, by (2.58)

$$\Big(1 - \frac{1}{n}\Big)K(x_n, y_0) + \frac{1}{n}K(x_n, y) \leq K(x_n, y_n) = g(y_n) \leq g(y_0)$$

and, since $g(y_0) = K(x(y_0), y_0) \leq K(x_n, y_0)$, we conclude that

$$K(x(y_n), y) \leq g(y_0) \qquad \forall n,\ \forall y \in B. \tag{2.61}$$

Since A is compact, there exist $\widetilde{x}(y) \in A$ and a subsequence $\{k_n\}$ such that $x_{k_n} \to \widetilde{x}(y)$ and $K(\widetilde{x}(y), y) = \min_n K(x(y_n), y)$, and, in turn,

$$K(\widetilde{x}(y), y) \leq \liminf_{n\to\infty} K(x_{k_n}, y) \leq g(y_0) \qquad \forall y \in B.$$

Step 3. Let us prove that

$$K(\widetilde{x}(y), y_0) = g(y_0) \qquad \forall y \in B. \tag{2.62}$$

We need to prove that $K(\widetilde{x}(y), y_0) \leq g(y_0)$, as the opposite inequality is trivial. With the notations of Step 2, from the concavity of $y \mapsto K(x,y)$

$$\Big(1 - \frac{1}{n}\Big)K(x_n, y_0) + \frac{1}{n}K(x_n, y) \leq K(x_n, y_n) = g(y_n) \leq g(y_0).$$

Consequently,

$$K(\widetilde{x}(y), y_0) \leq \liminf_{n\to\infty} K(x(y_n), y_0) \leq g(y_0).$$

Step 4. Let us prove the claim when $x \to K(x,y)$ is strictly convex. By Step 3, $\widetilde{x}(y)$ is a minimizer of the map $x \to K(x, y_0)$ as x_0 is. Since $x \mapsto K(x, y_0)$ is strictly convex, the minimizer is *unique*, thus concluding $\widetilde{x}(y) = x_0\ \forall y \in B$. The claim then follows from (2.59), (2.60) and (2.62).

Step 5. In case $x \to K(x,y)$ is merely convex, we introduce for every $\epsilon > 0$ the perturbed Lagrangian K_ϵ

$$K_\epsilon(x,y) := K(x,y) + \epsilon||x||, \qquad x \in A,\ y \in B$$

which is strictly convex. From Step 4 we infer the existence of a saddle point (x_ϵ, y_ϵ) for K_ϵ, i.e.,

$$K(x_\epsilon, y) + \epsilon||x_e|| \leq K(x_\epsilon, y_\epsilon) + \epsilon||x_\epsilon|| \leq K(x, y_\epsilon) + \epsilon||x|| \qquad \forall x \in A,\ y \in B.$$

Passing to subsequences, $x_\epsilon \to x_0 \in A$, $y_\epsilon \to y_0 \in B$, and from the above

$$K(x_0, y) \leq K(x, y_0) \qquad \forall x \in A,\ y \in B,$$

that is, (x_0, y_0) is a saddle point for K. $\square$

Proof of Theorem 2.91. For $k = 1, 2, \dots$, let $A_k := \{x \,|\, |x| \le k\}$, $B_k := \{y \,|\, |y| \le k\}$. By Proposition 2.92, $K(x, y)$ has a saddle point (x_k, y_k) on $A_k \times B_k$, i.e.,

$$K(x_k, y) \le K(x_k, y_k) \le K(x, y_k) \qquad \forall x \in A_k,\ y \in B_k. \tag{2.63}$$

Choosing $x = \overline{x}$, $y := \overline{y}$ in (2.63) we then have

$$K(x_k, \overline{y}) \le K(x_k, y_k) \le K(\overline{x}, y_k) \qquad \forall k$$

which implies trivially that $\{x_k\}$ and $\{y_k\}$ are both bounded. Therefore, passing eventually to subsequences, $x_k \to x_0$, $y_k \to y_0$, and from (2.63)

$$K(x_0, y) \le K(x_0, y_0) \le K(x, y_0) \qquad \forall x \in A_k,\ y \in B_k.$$

Since k is arbitrary, (x_0, y_0) is a saddle point for K on the whole $\mathbb{R}^n \times \mathbb{R}^m$. □

b. Optimal mixed strategies

An interesting case in which the previous theory applies is the case of finite strategies. We assume that the game (with zero sum) is played many times and that players P and Q choose their strategies, which are finitely many, on the basis of the frequency of success or of the probability: If the strategies of P and Q are respectively $\{E_1, E_2, \dots, E_m\}$ and $\{F_1, F_2, \dots, F_n\}$ and if $U(E_i, F_j)$ is the utility function resulting from the choices of E_i by P and F_j by Q, we assume that P chooses E_i with probability x_i and Q chooses F_j with probability y_j. Define now

$$A := \Big\{x \in \mathbb{R}^m \,\Big|\, 0 \le x_i \le 1,\ \sum_{i=1}^m x_i = 1\Big\},$$
$$B := \Big\{y \in \mathbb{R}^n \,\Big|\, 0 \le y_j \le 1,\ \sum_{j=1}^n y_j = 1\Big\};$$

then the payoff functions of the two players are given by

$$U_P(x, y) = -U_Q(x, y) = K(x, y) := \sum_{i,j} U(E_i, E_j) x_i y_j. \tag{2.64}$$

Since $K(x, y)$ is a homogeneous polynomial of degree 2, von Neumann's theorem applies to get the following result.

2.93 Theorem. *In a game with zero sum, there exist optimal mixed strategies* $(\overline{x}, \overline{y})$. *They are given by saddle points of the expected payoff function* (2.64)*, and for them we have*

$$\max_{x \in A} \min_{y \in B} K(x, y) = K(\overline{x}, \overline{y}) = \min_{y \in B} \max_{x \in A} K(x, y).$$

2.94 A linear programming approach. Theorem 2.93, although ensuring the existence of optimal mixed strategies, gives no method to find them, which, of course, is quite important. Notice that A and B are compact and convex sets with the vectors of the standard basis $e_1, e_2, \dots, e_m$

of $\mathbb{R}^m$ and $e_1, e_2, \dots, e_n$ of $\mathbb{R}^n$ as extreme points, respectively. Since $x \to K(x,y)$ and $y \to K(x,y)$ are linear, they attain their maximum and minimum at extreme points, hence

$$f(x) := \min_{y\in B} K(x,y) = \min_{1\le j\le n} K(x,e_i),$$
$$g(y) := \max_{x\in A} K(x,y) = \max_{1\le i\le m} K(e_i,y).$$

Notice that $f(x)$ and $g(y)$ are affine maps. Set $\mathbf{U} := (U_{ij})$, $U_{ij} := U(E_i,E_j)$; then maximizing f in A is equivalent to maximize a real number z subject to the constraints $z \le K(z,e_i)$ $\forall i$ and $x \in A$, that is, to solve

$$\begin{cases} F(x,z) := z \to \max, \\ z\begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix} \le \mathbf{U}x, \\ \sum_{i=1}^m x_i = 1, \\ x \ge 0. \end{cases}$$

Similarly, minimizing g in B is equivalent to solving

$$\begin{cases} G(y,w) := w \to \min, \\ w \ge \mathbf{U}^T y \le 0, \\ \sum_{i=1}^n y_i = 1, \\ y \ge 0. \end{cases}$$

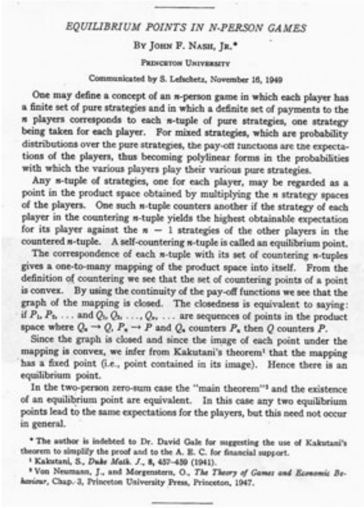
These are two problems of linear programming, one the dual of the other, and they can be solved with the methods of linear programming, see Section 2.4.7.

c. Nash equilibria

2.95 Example (The prisoner dilemma). Two prisoners have to serve a one-year prison sentence for a minor crime, but they are suspected of a major crime. Each of them receives separately the following proposal: If he accuses the other of the major crime, he will not have to serve the one-year sentence for the minor crime and, if the other does not accuse him of the major crime (in which case he will have to serve the relative 5-year prison sentence), he will be freed. The possible strategies are two: (a) accusing the other and (b) not accusing the other; the corresponding utility functions for the two prisoners P and Q (in years of prison to serve, with negative sign, so that we have to maximize) are

$$\begin{array}{llll} U_P(a,a) = -5, & U_P(a,n) = 0, & U_P(n,a) = -6, & U_P(n,n) = -1, \\ U_Q(a,a) = -5, & U_Q(a,n) = -6, & U_Q(n,a) = 0, & U_Q(n,n) = -1. \end{array}$$

We see at once that the strategy of accusing each other gives the worst result with respect to the choice of not accusing the other. Nevertheless, the choice of accusing the



EQUILIBRIUM POINTS IN N-PERSON GAMES

BY JOHN F. NASH, JR.*

PRINCETON UNIVERSITY

Communicated by S. Lefschetz, November 16, 1949

One may define a concept of an n-person game in which each player has a finite set of pure strategies and in which a definite set of payments to the n players corresponds to each n-tuple of pure strategies, one strategy being taken for each player. For mixed strategies, which are probability distributions over the pure strategies, the pay-off functions are the expectations of the players, thus becoming polylinear forms in the probabilities with which the various players play their various pure strategies.

Any n-tuple of strategies, one for each player, may be regarded as a point in the product space obtained by multiplying the n strategy spaces of the players. One such n-tuple counters another if the strategy of each player in the countering n-tuple yields the highest obtainable expectation for its player against the $n-1$ strategies of the other players in the countered n-tuple. A self-countering n-tuple is called an equilibrium point.

The correspondence of each n-tuple with its set of countering n-tuples gives a one-to-many mapping of the product space into itself. From the definition of countering we see that the set of countering points of a point is convex. By using the continuity of the pay-off functions we see that the graph of the mapping is closed. The closedness is equivalent to saying: if $P_1, P_2, \ldots$ and $Q_1, Q_2, \ldots, Q_n, \ldots$ are sequences of points in the product space where $Q_n \to Q$, $P_n \to P$ and Q_n counters P_n then Q counters P.

Since the graph is closed and since the image of each point under the mapping is convex, we infer from Kakutani's theorem[1] that the mapping has a fixed point (i.e., point contained in its image). Hence there is an equilibrium point.

In the two-person zero-sum case the "main theorem"[2] and the existence of an equilibrium point are equivalent. In this case any two equilibrium points lead to the same expectations for the players, but this need not occur in general.

* The author is indebted to Dr. David Gale for suggesting the use of Kakutani's theorem to simplify the proof and to the A. E. C. for financial support.

[1] Kakutani, S., *Duke Math. J.*, 8, 457–459 (1941).

[2] Von Neumann, J., and Morgenstern, O., *The Theory of Games and Economic Behaviour*, Chap. 3, Princeton University Press, Princeton, 1947.

NON-COOPERATIVE GAMES

JOHN NASH

(Received October 11, 1950)

Introduction

Von Neumann and Morgenstern have developed a very fruitful theory of two-person zero-sum games in their book *Theory of Games and Economic Behavior*. This book also contains a theory of n-person games of a type which we would call cooperative. This theory is based on an analysis of the interrelationships of the various coalitions which can be formed by the players of the game.

Our theory, in contradistinction, is based on the *absence* of coalitions in that it is assumed that each participant acts independently, without collaboration or communication with any of the others.

The notion of an *equilibrium point* is the basic ingredient in our theory. This notion yields a generalization of the concept of the solution of a two-person zero-sum game. It turns out that the set of equilibrium points of a two-person zero-sum game is simply the set of all pairs of opposing "good strategies."

In the immediately following sections we shall define equilibrium points and prove that a finite non-cooperative game always has at least one equilibrium point. We shall also introduce the notions of solvability and strong solvability of a non-cooperative game and prove a theorem on the geometrical structure of the set of equilibrium points of a solvable game.

As an example of the application of our theory we include a solution of a simplified three person poker game.

Formal Definitions and Terminology

In this section we define the basic concepts of this paper and set up standard terminology and notation. Important definitions will be preceded by a subtitle indicating the concept defined. The non-cooperative idea will be implicit, rather than explicit, below.

Finite Game:

For us an *n-person game* will be a set of n *players*, or *positions*, each with an associated finite set of *pure strategies*; and corresponding to each player, i, a *payoff function*, p_i, which maps the set of all n-tuples of pure strategies into the real numbers. When we use the term *n-tuple* we shall always mean a set of n items, with each item associated with a different player.

Mixed Strategy, s_i:

A *mixed strategy* of player i will be a collection of non-negative numbers which have unit sum and are in one to one correspondence with his pure strategies.

We write $s_i = \sum_\alpha c_{i\alpha}\pi_{i\alpha}$ with $c_{i\alpha} \geq 0$ and $\sum_\alpha c_{i\alpha} = 1$ to represent such a mixed strategy, where the $\pi_{i\alpha}$'s are the pure strategies of player i. We regard the s_i's as points in a simplex whose vertices are the $\pi_{i\alpha}$'s. This simplex may be re-

Figure 2.14. The initial pages of two papers by John Nash (1928–).

other brings the advantage of serving one year less in any case: The strategy of not accusing, which, from a *cooperative* point of view is the best, is not the *individual* point of view (even in the presence of a reciprocal agreement; in fact, neither of the two may ensure that the other will not accuse him). This paradox arises quite frequently.

The idea that individual rationality, typical of *noncooperative games* (in which there is no possibility of agreement among the players), precedes collective rationality is at the basis of the notion of the *Nash equilibrium.*

2.96 Definition. *Let A and B be two sets and let f and g be two maps from $A \times B$ into $\mathbb{R}$. The couple of points $(x_0, y_0) \in A \times B$ is called a* Nash point *for f and g if for all $(x, y) \in A \times B$ we have*

$$f(x_0, y_0) \geq f(x, y_0), \qquad g(x_0, y_0) \geq g(x_0, y).$$

In the prisoner's dilemma, the unique Nash point is the strategy of the reciprocal accusation. In a game with zero sum, i.e., $U_P(x, y) = -U_Q(x, y) =: K(x, y)$, clearly (x_0, y_0) is a Nash point if and only if (x_0, y_0) is a *saddle point* for K.

2.97 Theorem (of Nash for two players). *Let A and B be two nonempty, convex and compact sets. Let $f, g : A \times B \to \mathbb{R}$ be two continuous functions such that $x \to f(x, y)$ is concave for all $y \in B$ and $y \to g(x, y)$ is concave for all $x \in A$. Then there exists a Nash equilibrium point for f and g.*

Proof. Introduce the function $F : (A \times B) \times (A \times B) \to \mathbb{R}$ defined by

$$F(p, q) = f(p_1, q_2) + g(q_1, p_2), \qquad \forall p = (p_1, p_2),\ q = (q_1, q_2) \in A \times B.$$

Clearly, F is continuous and concave in p for every chosen q. We claim that there is $q_0 \in A \times B$ such that

$$\max_{p \in A\times B} F(p, q_0) = F(p_0, q_0). \tag{2.65}$$

Before proving the claim, let us complete the proof of the theorem on the basis of (2.65). If we set $(x_0, y_0) := q_0$, we have

$$f(x, y_0) + g(x_0, y) \le f(x_0, y_0) + g(x_0, y_0) \qquad \forall (x, y) \in A \times B.$$

Choosing $x = x_0$, we infer $g(x_0, y) \le g(x_0, y_0)$ $\forall y \in B$, while, by choosing $y = y_0$, we find $f(x, y_0) \le f(x_0, y_0)$ $\forall x \in A$, hence (x_0, y_0) is a Nash point.

Let us prove (2.65). Since the inequality $\ge$ is trivial, for all $q_0 \in A \times B$, we need to prove only the opposite inequality. By contradiction, suppose that $\forall q \in A \times B$ there is $p \in A \times B$ such that $F(p, q) > F(q, q)$ and, then, set

$$G_q := \Big\{ q \in A \times B \,\Big|\, F(p, q) > F(q, q) \Big\}, \qquad p \in A \times B.$$

The family $\{G_p\}_{p \in A\times B}$ is an open covering of $A \times B$; consequently, there are finitely many points $p_1, p_2, \dots, p_k \in A \times B$ such that $A \times B \subset \cup_{i=1}^k G_{p_i}$. Set

$$\varphi_i(q) := \max\Big(F(p_i, q) - F(q, q), 0 \Big), \qquad q \in A \times B,\ i = 1, \dots, k.$$

The functions $\{\varphi_i\}$ are continuous, nonnegative and, for every q, at least one of them does not vanish at q; we then set

$$\psi_i(q) := \frac{\varphi_i(q)}{\sum_{j=1}^k \varphi_j(q)}$$

and define the new map $\psi : A \times B \to A \times B$ by

$$\psi(q) := \sum_{i=1}^k \psi_i(q) p_i.$$

The map ψ maps the convex and compact set $A \times B$ into itself, consequently, it has a fixed point $q' \in A \times B$, $q' = \sum_i \psi(q') p_i$. F being concave,

$$F(q', q') = F\Big(\sum_i \psi_i(q') p_i, q' \Big) \ge \sum_{i=1}^k \psi_i(q') F(q_i, q').$$

On the other hand, $F(p_i, q') > F(q', q')$ if $\psi_i(q') > 0$, hence

$$F(q', q') \ge \sum_{i=1}^k \psi(q') F(q_i, q') > \sum_{i=1}^k \psi(q') F(q', q') = F(q', q'),$$

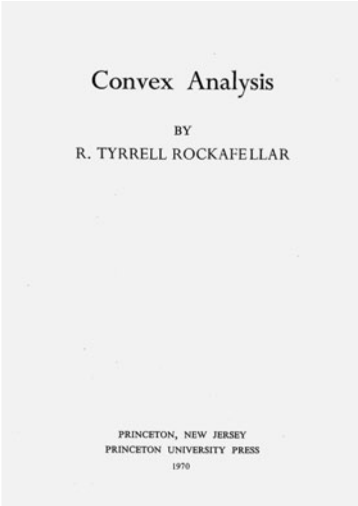
which is a contradiction. □

d. Convex duality

Let $f, \varphi^1, \dots, \varphi^m : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ be convex functions defined on a convex open set Ω. We assume for simplicity that f and $\varphi := (\varphi^1, \varphi^2, \dots, \varphi^m)$ are differentiable. Let

$$\mathcal{F} := \Big\{ x \in \mathbb{R}^n \,\Big|\, \varphi^j(x) \le 0,\ \forall j = 1, \dots, m \Big\}.$$

The *primal problem* of convex optimization is the *minimum problem*



Convex Analysis

BY

R. TYRRELL ROCKAFELLAR

PRINCETON, NEW JERSEY
PRINCETON UNIVERSITY PRESS
1970

ÉTUDES MATHÉMATIQUES
Collection dirigée par P. LELONG
Professeur à la Faculté des Sciences de Paris

analyse convexe
et problèmes
variationnels

Ivar EKELAND
Maître de conférences à l'Université Paris IX
et à l'École Polytechnique

Roger TEMAM
Professeur à l'Université Paris XI
Maître de conférence à l'École Polytechnique

DUNOD GAUTHIER-VILLARS
PARIS - BRUXELLES - MONTRÉAL

Figure 2.15. Two classical monographs on convexity.

$$\textit{Assuming } \mathcal{F} \neq \emptyset, \textit{ minimize } f \textit{ in } \mathcal{F}. \tag{2.66}$$

The associated *Lagrangian* $\mathcal{L} : \Omega \times \mathbb{R}^m_+$ to (2.66), defined by

$$\mathcal{L}(x, \lambda) := f(x) + \lambda \bullet \varphi(x), \qquad x \in \Omega, \ \lambda \geq 0, \tag{2.67}$$

is convex in x for any fixed λ and linear in λ for every fixed x. Therefore, it is not surprising that the Kuhn–Tucker conditions (2.41)

$$\begin{cases} \mathbf{D}f(x^0) + \lambda^0 \bullet \mathbf{D}\varphi(x_0) = 0, \\ \lambda^0 \geq 0, \ x_0 \in \mathcal{F}, \\ \lambda^0 \bullet \varphi(x_0) = 0 \end{cases} \tag{2.68}$$

are also sufficient to characterize minimum points for f on $\mathcal{F}$. Actually, the Kuhn–Tucker equilibrium conditions (2.41) are strongly related to saddle points for the associated Lagrangian $\mathcal{L}(x, \lambda)$.

2.98 Theorem. *Consider the primal problem* (2.66). *Then* (x_0, λ^0) *fulfills* (2.68) *if and only if* (x_0, λ^0) *is a saddle point for* $\mathcal{L}(x, \lambda)$ *on* $\Omega \times \mathbb{R}^m_+$, *i.e.,*

$$\mathcal{L}(x_0, \lambda) \leq \mathcal{L}(x_0, \lambda^0) \leq cL(x, \lambda^0)$$

for all $x \in \mathcal{F}$ *and* $\lambda \in \mathbb{R}^m$, $\lambda \geq 0$. *In particular, if the Kuhn–Tucker equilibrium conditions are satisfied at* $(x_0, \lambda^0) \in \mathcal{F} \times \mathbb{R}^m_+$, *then* x_0 *is a minimizer for* f *on* $\mathcal{F}$.

Proof. From the convexity of $x \mapsto \mathcal{L}(x, \lambda^0)$ and (2.41) we infer

$$\mathcal{L}(x,\lambda^0) \geq \mathcal{L}(x_0,\lambda_0) + \sum_{i=1}^n \Big(\nabla f(x_0) + scd\lambda^0 \mathbf{D}\varphi(x_0)\Big)^i (x-x_0)^i = \mathcal{L}(x_0,\lambda^0) = f(x_0)$$

for all $x \in \Omega$. In particular,

$$f(x) \geq f(x) + \lambda^0 \bullet \varphi(x^0) = \mathcal{L}(x,\lambda^0) \geq f(x_0),$$
$$\mathcal{L}(x_0,\lambda^0) \geq f(x_0) + \lambda \bullet \varphi(x_0) = \mathcal{L}(x,\lambda^0).$$

Conversely, suppose that (x_0, λ^0) is a saddle point for $\mathcal{L}(x,\lambda)$ on $\Omega \times \mathbb{R}^m_+$, i.e.,

$$f(x_0) + \lambda \bullet \varphi(x_0) \leq f(x_0) + \lambda^0 \bullet \varphi(x_0) \leq f(x) + \lambda^0 \bullet \varphi(x)$$

for every $x \in \Omega$ and $\lambda \geq 0$. From the first inequality we infer

$$\lambda \bullet \varphi(x_0) \geq \lambda^0 \varphi(x_0) \tag{2.69}$$

for any $\lambda \geq 0$. This implies that $\varphi(x_0) \leq 0$ and, in turn, $\lambda_0 \bullet \varphi(x_0) \leq 0$. Using again (2.69) with $\lambda = 0$, we get the opposite inequality, thus concluding that $\lambda_0 \bullet \varphi(x_0) = 0$. Finally, from the first inequality, Fermat's theorem yields

$$\nabla f(x_0) + \lambda^0 \bullet \nabla\varphi(x_0) = 0.$$

□

Let us now introduce the *dual problem* of convex optimization. For $\lambda \in \mathbb{R}^m_+$, set

$$g(\lambda) := \inf_{x \in \mathcal{F}} \mathcal{L}(x,\lambda),$$

where $\mathcal{L}(x,\lambda)$ is the Lagrangian in (2.67).

Since $g(\lambda)$ is the infimum of a family of affine functions, $-g$ is convex and proper on

$$\mathcal{G} := \{\lambda \in \mathbb{R}^m \,|\, \lambda \geq 0, g(\lambda) > -\infty\}.$$

The dual problem of convex programming is

$$\textit{Assuming } \mathcal{G} \neq \emptyset, \textit{ maximize } g(\lambda) \textit{ on } \mathcal{G} \tag{2.70}$$

or, equivalently,

$$\textit{Assuming } \mathcal{G} \neq \emptyset, \textit{ maximize } g(\lambda) \textit{ on } \{\lambda \in \mathbb{R}^m \,|\, \lambda \geq 0\}. \tag{2.71}$$

2.99 Theorem. *If $(x_0, \lambda^0) \in \mathcal{F} \times \mathbb{R}^m$ satisfies the Kuhn–Tucker equilibrium conditions* (2.41)*, then x_0 maximizes the primal problem, λ_0 minimizes the dual problem and $f(x_0) = g(\lambda^0) = \mathcal{L}(x_0, \lambda^0)$.*

Proof. By definition, $g(\lambda) = \sup_{x \in \mathcal{F}} \mathcal{L}(x,\lambda)$, and, trivially, $f(x) := \inf_{\lambda \geq 0} \mathcal{L}(x,\lambda)$. Therefore $g(y) \leq f(x)$ for all $x \in \mathcal{F}$ and $\lambda \geq 0$, so that

$$\sup_{\lambda \geq 0} g(\lambda) \leq \inf_{x \in \mathcal{F}} f(x).$$

Since (x_0, λ^0) is a saddle point for $\mathcal{L}$ on $\Omega \times \mathbb{R}_+$, Proposition 2.88 yields the result. □

2.5 A General Approach to Convexity

As we have seen, every closed convex set K is the intersection of all closed half-spaces in which it is contained; in fact, K is the *envelope* of its *supporting hyperplanes*. In other words, a closed convex body is given in a dual way by the supporting hyperplanes. This remark, when applied to closed epigraphs of convex functions, yields a number of interesting correspondences. Here we discuss the so-called *polarity* correspondence.

a. Definitions

It is convenient to allow that convex functions take the value $+\infty$ with the convention $t + (+\infty) = +\infty$ for all $t \in \mathbb{R}$ and $t \cdot (+\infty) = +\infty$ for all $t > 0$. For technical reasons, it is also convenient to allow that convex functions take the value $-\infty$.

2.100 Definition. *$f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is convex if*

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y) \qquad \forall x, y \in \mathbb{R}^n, \quad \forall \lambda \in [0,1]$$

unless $f(x) = -f(y) = \pm\infty$. The effective domain *of f is then defined by*

$$\mathrm{dom}(f) := \Big\{ x \in \mathbb{R}^n \,\Big|\, f(x) < \infty \Big\}.$$

We say that f is proper *if f is nowhere $-\infty$ and* $\mathrm{dom}(f) \neq \emptyset$.

Let $K \subset \mathbb{R}^n$ be a convex set and $f : K \subset \mathbb{R}^n \to \mathbb{R}$ be a convex function. It is readily seen that the function $\overline{f} : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ defined as

$$\overline{f}(x) = \begin{cases} f(x) & \text{if } x \in K, \\ +\infty & \text{if } x \notin K \end{cases}$$

is convex according to Definition 2.100 with effective domain given by K.

One of the advantages of Definition 2.100 is that convex sets and convex functions are essentially the same object.

From one side, $K \subset \mathbb{R}^n$ is convex if and only if its *indicatrix function*

$$I_K(x) := \begin{cases} 0 & \text{if } x \in K, \\ +\infty & \text{if } x \notin K \end{cases} \tag{2.72}$$

is convex in the sense of Definition 2.100. On the other hand, $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is convex if and only if its *epigraph*, defined as usual by

$$\mathrm{Epi}(f) := \Big\{ (x,t) \in \mathbb{R}^n \times \mathbb{R} \,\Big|\, x \in \mathbb{R}^n,\ t \in \mathbb{R},\ t \ge f(x) \Big\}$$

is a convex set in $\mathbb{R}^n \times \mathbb{R}$.

Observe that the constrained minimization problem

$$\begin{cases} f(x) \to \min, \\ x \in K, \end{cases}$$

where f is a convex function and K is a convex set, transforms into the unconstrained minimization problem for the convex function

$$f(x) + I_K(x), \qquad x \in \mathbb{R}^n$$

which is defined by adding to f the indicatrix I_K of K as *penalty function.*

One easily verifies that

(i) f is convex if and only if its epigraph is convex,
(ii) the effective domain of a convex function is convex,
(iii) if f is convex, then $\mathrm{dom}(f) = \pi(\mathrm{Epi}(f))$ where $\pi : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ is the linear projection on the first factor.

We have also proved, compare Theorem 2.35, that *every proper convex function is locally Lipschitz in the interior of its effective domain.* However, in general, a convex function need not be continuous or semicontinuous at the boundary of its effective domain, as, for instance, for the functions f defined as $f(x) = 0$ if $x \in]-1, 1[$, $f(-1) = f(1) = 1$ and $f(x) = +\infty$ if $x \notin [0, 1]$.

b. Lower semicontinuous functions and closed epigraphs

We recall that $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is said to be *lower semicontinuous*, see [GM3], in short *l.s.c.*, if $f(x) \le \liminf_{y \to x} f(y)$. If $f(x) \in \mathbb{R}$, this means the following:

(i) For all $\epsilon > 0$ there is $\delta > 0$ such that for all $y \in B(x, \delta) \setminus \{x\}$ we have $f(x) - \epsilon \le f(y)$.
(ii) There is a sequence $\{x_k\}$ with values in $\mathbb{R}^n \setminus \{x\}$ that converges to x such that $f(x_k) \to f(x)$.

Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. We already know that f is l.s.c. if and only if for every $t \in \mathbb{R}$ the sublevel set $\{x \,|\, f(x) \le t\}$ is closed. Moreover, the following holds.

2.101 Proposition. *The epigraph of a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is closed if and only if f is lower semicontinuous.*

Proof. Let f be l.s.c. and $\{(x_k, t_k)\} \subset \mathrm{Epi}(f)$ a sequence that converges to (x, t). Then $x_\kappa \to x$, $t_k \to t$ and $f(x_k) \le t_k$. It follows that $f(x) \le \liminf_{k\to\infty} f(x_k) \le \liminf_{k\to\infty} t_k = t$, i.e., $(x, t) \in \mathrm{Epi}(f)$.

Conversely, suppose that $\mathrm{Epi}(f)$ is closed. Consider a sequence $\{x_k\}$ with $x_k \to x$ and let $L := \liminf_{k\to\infty} f(x_k)$. If $L = +\infty$, then $f(x) \le L$. If $L < +\infty$, we find a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $f(x_{n_k}) \to L$. Since $(x_{n_k}, f(x_{n_k})) \in \mathrm{Epi}(f)$ and $L < +\infty$, we infer that $(x, L) \in \mathrm{Epi}(f)$, i.e., $f(x) \le L = \liminf_{k\to\infty} f(x_k)$. Since the sequence $\{x_k\}$ is arbitrary, we finally conclude that $f(x) \le \liminf_{y\to x} f(y)$. □

Finally, let us observe that if $f_\alpha : \mathbb{R}^n \to \overline{\mathbb{R}}$, $\alpha \in \mathcal{A}$, is a family of l.s.c. functions, then

$$f(x) := \sup\Big\{ f_\alpha(x) \,\Big|\, \alpha \in \mathcal{A} \Big\}, \qquad x \in \mathbb{R}^n,$$

is lower semicontinuous.

2.102 Definition. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a function. The* closure *of f or its* lower semicontinuous regularization, *in short its* l.s.c. regularization, *is the function*

$$\Gamma f(x) := \sup\Big\{ g(x) \,\Big|\, g : \mathbb{R}^n \to \overline{\mathbb{R}},\ g \text{ is l.s.c.},\ g(y) \le f(y)\ \forall y \Big\}.$$

Clearly, $\Gamma f(x) \le f(x)$ for every x, and, as the pointwise supremum of a family of l.s.c. functions, Γf is lower semicontinuous. Therefore, it is the *greatest lower semicontinuous minorant* of f.

2.103 Proposition. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. Then* $\mathrm{Epi}(\Gamma f) = \mathrm{cl}(\mathrm{Epi}(f))$ *and* $\Gamma f(x) = \liminf_{y \to x} f(y)$ *for every $x \in \mathbb{R}^n$.*

Consequently, $\Gamma f(x) = f(x)$ *if and only if f is l.s.c. at x.*

Proof. (i) First, let us prove that $\mathrm{cl}(\mathrm{Epi}(f))$ is the epigraph of a function $g \le f$, by proving that if $(x,t) \in \mathrm{cl}(\mathrm{Epi}(f))$, then for all $s > t$ we have $(x,s) \in \mathrm{cl}(\mathrm{Epi}(f))$. If $(x_k, t_k) \in \mathrm{Epi}(f)$ converges to (x,t) and $s > t$, for some large k we have $t_k < s$, hence $f(x_k) \le t_k < s$. It follows that definitively $(x_k, s) \in \mathrm{Epi}(f)$, hence $(x,s) \in \mathrm{cl}(\mathrm{Epi}(f))$.

By Proposition 2.101, g is l.s.c. and Γf is closed; therefore, we have $g \le \Gamma f$ and

$$\mathrm{Epi}(\Gamma f) \subset \mathrm{Epi}(g) = \mathrm{cl}(\mathrm{Epi}(f)) \subset \mathrm{Epi}(\Gamma f).$$

(ii) Let $x \in \mathbb{R}^n$. If $\Gamma f(x) = +\infty$, $\Gamma f = +\infty$ in a neighborhood of x, hence $\liminf_{y \to x} f(y) = +\infty$, too.

If $\Gamma f(x) < +\infty$, then for any $t \ge \mathbf{f}(x)$, $(x,t) \in \mathrm{Epi}(\Gamma f)$. (i) yields a sequence $\{(x_k, t_k)\} \subset \mathrm{Epi}(f)$ such that $x_k \to x$ and $y_k \to t$. Therefore

$$\liminf_{k \to \infty} f(x_k) \le \liminf_{k \to \infty} t_k = t,$$

hence

$$\liminf_{k \to \infty} f(x_k) \le \Gamma f(x).$$

On the other hand, since Γf is ls.c. and $\Gamma f \le f$,

$$\Gamma f(x) \le \liminf_{y \to x} \Gamma f(y) \le \liminf_{y \to x} f(y),$$

thus concluding that $\Gamma f(x) = \liminf_{y \to x} f(y)$. It is then easy to check that $f(x) = \Gamma f(x)$ if and only if f is l.s.c. at x. □

Since closed convex sets can be represented as intersections of their supporting half-spaces, of particular relevance are the convex functions with closed epigraphs. According to the above, we have the following.

2.104 Corollary. *$f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is convex and l.s.c. if and only if its epigraph is convex and closed.*

The l.s.c. regularization Γf of a convex function f is a convex and l.s.c. function.

According to the above, $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is l.s.c. and convex if and only if its epigraph $\mathrm{Epi}(f)$ is closed and convex. In particular, $\mathrm{Epi}(f)$ is the intersection of all its supporting half-spaces. The next theorem states that $\mathrm{Epi}(f)$ is actually the intersection of all half-spaces associated to graphs of linear affine functions, i.e., to hyperplanes that do not contain vertical vectors.

We first state a proposition that contains the relevant property.

2.105 Proposition. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be convex and l.s.c. and let $\overline{x} \in \mathbb{R}^n$ be such that $f(\overline{x}) > -\infty$. Then the following hold:*

(i) *For every $\overline{y} < f(\overline{x})$ there exists an affine map $\ell : \mathbb{R}^n \to \mathbb{R}$ such that $f(x) > \ell(x)$ for every $x \in \mathbb{R}^n$ and $\overline{y} < \ell(\overline{x})$.*
(ii) *If $\overline{x} \in \mathrm{int}(\mathrm{dom}(f))$, then there exists an affine map $\ell : \mathbb{R}^n \to \mathbb{R}$ such that $f(x) > \ell(x)$ for every $x \in \mathbb{R}^n$ and $\ell(\overline{x}) = f(\overline{x})$.*

Proof. Since f is lower semicontinuous at $\overline{x}$, there exist $\epsilon > 0$ and $\delta > 0$ such that $\overline{y} \le f(x) - \epsilon\ \forall x \in B(\overline{x}, \delta)$, in particular, $(\overline{x}, \overline{y}) \notin \mathrm{cl}(\mathrm{Epi}(f))$. Therefore, there exists a hyperplane $\mathcal{P} \subset \mathbb{R}^{n+1}$ that strongly separates $\mathrm{Epi}(f)$ from $(\overline{x}, \overline{y})$, i.e., there are a linear map $m : \mathbb{R}^n \to \mathbb{R}$ and numbers $\alpha, \beta \in \mathbb{R}$ such that

$$\mathcal{P} := \Big\{(x, y) \,\Big|\, m(x) + \alpha y = \beta\Big\} \tag{2.73}$$

with

$$m(x) + \alpha y > \beta \ \ \forall (x, y) \in \mathrm{Epi}(f) \qquad \text{and} \qquad m(\overline{x}) + \alpha\overline{y} < \beta. \tag{2.74}$$

Since y may be chosen arbitrarily large in the first inequality, we also have $\alpha \ge 0$. We now distinguish four cases.

(i) If $f(\overline{x}) < +\infty$, then $\alpha \neq 0$ since, otherwise, choosing $(\overline{x}, y)$ with $y > f(\overline{x})$ in (2.74), we get $m(\overline{x}) > \beta$ and $m(\overline{x}) < \beta$, a contradiction. By choosing ℓ as the linear affine map $\ell(x) := (\beta - m(x))/\alpha$, from the first of (2.74) with $y = f(x)$ it follows $\ell(x) < f(x)$ for all x, while from the second we get $\overline{y} \le \ell(\overline{x})$.

(ii) If $f(\overline{x}) = +\infty$ and the function takes value $+\infty$ everywhere, the claim is trivial.

(iii) If $f(\overline{x}) = +\infty$ and $\alpha > 0$ in (2.74), then one chooses ℓ as the linear affine map $\ell(x) := (\beta - m(x))/\alpha$, as in (i).

(iv) Let us consider the remaining case where $f(\overline{x}) = +\infty$. There exists x_0 such that $f(x_0) \in \mathbb{R}$ and $\alpha = 0$ in (2.74). By applying (i) at x_0, we find an affine linear map ϕ such that

$$f(x) \ge \phi(x) \qquad \forall x \in \mathbb{R}^n.$$

For all $c > 0$ the function $\ell(x) := \phi(x) + c(\beta - m(x))$ is then a linear affine minorant of $f(x)$ and, by choosing c sufficiently large, we can make $\ell(\overline{x}) = \phi(\overline{x}) + c(\beta - m(\overline{x})) > \overline{y}$. This concludes the proof of the first claim.

Let us now prove the last claim. Since $\overline{x} \in \mathrm{int}(\mathrm{dom}(f))$ and $f(\overline{x}) > -\infty$, a support hyperplane $\mathcal{P}'$ of $\mathrm{Epi}(f)$ at $(\overline{x}, f(\overline{x}))$ does not contain vertical vectors: otherwise none of the two subspaces associated to $\mathcal{P}'$ could contain $\mathrm{Epi}(f)$. Hence $\mathcal{P}' := \{(x, y) \,|\, m(x) + \alpha y = \beta\}$ for some linear map m and numbers $\alpha, \beta \in \mathbb{R}$ with

$$m(x) + \alpha y \ge \beta \ \ \forall (x, y) \in \mathrm{Epi}(f), \qquad m(\overline{x}) + \alpha f(\overline{x}) = \beta,$$

and $\alpha > 0$. If $\ell(x) := -m(x)/\alpha$, we see at once that

$$f(x) \ge f(\overline{x}) + \ell(x - \overline{x}) \qquad \forall x \in \mathbb{R}^n.$$

$\square$

2.106 Remark. The previous proof yields the existence of a nontrivial lower affine minorant for f which is arbitrarily close to $f(\overline{x})$ at $\overline{x}$ when f is l.s.c. at $\overline{x} \in \mathbb{R}^n$, $f(\overline{x}) > -\infty$ and one the following conditions hold:

- $f(\overline{x}) \in \mathbb{R}$,
- $f = +\infty$ everywhere,
- $f(\overline{x}) = +\infty$ and there exists a further point $x_0 \in \mathbb{R}^n$ such that $f(x_0) \in \mathbb{R}$ and f is l.s.c. at x_0.

Notice also that if f is convex, then $f(x) > -\infty$ and $x \in \mathrm{int}(\mathrm{dom}(f))$ if and only if f is continuous at x, see Theorem 2.35.

2.107 Corollary. *If $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is convex and l.s.c. and $f(\overline{x}) > -\infty$ at some point $\overline{x}$, then $f > -\infty$ everywhere.*

2.108 Definition. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a function. Its* linear l.s.c. envelope $\Gamma\mathrm{L}\, f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *is defined by*

$$\Gamma\mathrm{L}\, f(x) := \sup\Big\{\ell(x) \,\Big|\, \ell : \mathbb{R}^n \to \mathbb{R},\ \ell \text{ affine},\ \ell \leq f\Big\}. \tag{2.75}$$

and, of course, $\Gamma\mathrm{L}\, f(x) = -\infty$ $\forall x$ if no affine linear map ℓ below f exists.

2.109 Theorem. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$.*

(i) $\Gamma\mathrm{L}$ *is convex and l.s.c.*
(ii) *f is convex and l.s.c. if and only if $f(x) = \Gamma\mathrm{L}\, f(x)$ $\forall x \in \mathbb{R}^n$.*
(iii) *Assume f is convex. If at some point $x \in \mathbb{R}^n$ we have $f(x) < +\infty$, then $f(x) = \Gamma\mathrm{L}\, f(x)$ if and only if f is l.s.c. at x.*
(iv) *If $\overline{x}$ is an interior point of the effective domain of f and $f(\overline{x}) > -\infty$, then the supremum in (2.75) is a maximum, i.e., there exists $\xi \in \mathbb{R}^n$ such that*
$$f(y) \geq f(\overline{x}) + \xi \bullet (y - \overline{x}) \qquad \forall y.$$

Proof. Since the supremum of a family of convex and l.s.c. functions is convex and l.s.c., (2.75) implies that $\Gamma\mathrm{L}\, f(x)$ is convex and l.s.c.. If $\Gamma\mathrm{L}\, f(x) = -\infty$ for all x, then, trivially, $\Gamma\mathrm{L}$ is convex and l.s.c.. This proves (i), (ii) and (iii) are trivial if f is identically $-\infty$, and easily follow from the above and (i) of Proposition 2.105, taking also into account Remark 2.106. Finally, (iv) rephrases (ii) of Proposition 2.105. □

The following observation is sometimes useful.

2.110 Proposition. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be convex and l.s.c. and let $r(t) = (1-t)x + t\overline{x}$, $t \in [0,1]$, be the segment joining x to $\overline{x}$. Suppose $f(\overline{x}) < +\infty$. Then*
$$f(x) = \lim_{t \to 0^+} f(r(t)).$$

Proof. Since $f(\overline{x}) < +\infty$,
$$f(x) \leq \liminf_{t \to 0^+} f(t\overline{x} + (1-t)x) \leq \lim_{t \to 0} \Big(t\, f(\overline{x}) + (1-t) f(x)\Big) = f(x).$$
□

c. The Fenchel transform

2.111 Definition. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. The* polar *or* Fenchel transform *of f is the function $f^* : \mathbb{R}^n \to \overline{\mathbb{R}}$ defined by*

$$f^*(\xi) := \sup_{x\in\mathbb{R}^n} \Big(\xi \bullet x - f(x) \Big) = - \inf_{x\in\mathbb{R}^n} \Big(f(x) - \xi \bullet x \Big). \tag{2.76}$$

As we will see, the Fenchel transform rules the entire mechanism of convex duality.

2.112 Proposition. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a function and $f^* : \mathbb{R}^n \to \overline{\mathbb{R}}$ its polar. Then we have the following:*

(i) *$f(x) \geq \xi \bullet x - \eta$ $\forall x$ if and only if $f^*(\xi) \leq \eta$;*
(ii) *$f^*(\xi) = -\infty$ for some ξ if and only if $f(x) = +\infty$ for all x;*
(iii) *if $f \leq g$, then $g^* \leq f^*$;*
(iv) *$f^*(0) = -\inf_{x\in\mathbb{R}^n} f(x)$;*
(v) *the* Fenchel inequality *holds*

$$\xi \bullet x \leq f^*(\xi) + f(x) \qquad \forall x \in \mathbb{R}^n \ \forall \xi \in \mathbb{R}^n,$$

with equality at $(\overline{x}, \overline{\xi})$ if and only if $f(x) \geq f(\overline{x}) + \overline{\xi} \bullet (x - \overline{x})$.
(vi) *f^* is l.s.c. and convex.*

Proof. All of the claims follow immediately from the definition of f^*. □

The polar transform generalizes Legendre's transform.

2.113 Proposition. *Let Ω be an open set in $\mathbb{R}^n$, let $f : \Omega \to \mathbb{R}$ be a convex function of class C^2 with positive definite Hessian matrix and let $\Gamma\mathrm{L}\, f$ be the l.s.c linear envelope of f. Then*

$$\mathcal{L}_f(\xi) = (\Gamma\mathrm{L}\, f)^*(\xi) \qquad \forall \xi \in \mathbf{D}f(\Omega).$$

Proof. According to Theorem 2.109, $f(x) = \Gamma\mathrm{L}\, f(x)$ for all $x \in \Omega$, while Theorem 2.109 yields for all $\xi \in \mathbf{D}f(\Omega)$

$$\mathcal{L}_f(\xi) = \max_{\Omega}(x \bullet \xi - f(x)) \leq \sup_{x\in\mathbb{R}^n} (x \bullet \xi - \Gamma\mathrm{L}\, f(x)) = (\Gamma\mathrm{L}\, f)^*(\xi).$$

On the other hand,

$$(\Gamma\mathrm{L}\, f)^*(\xi) = \sup_{x\in\overline{\Omega}} (x \bullet \xi - \Gamma\mathrm{L}\, f(x)) =: L.$$

Given $\epsilon > 0$, let $x \in \overline{\Omega}$ be such that $L < \overline{x} \bullet \xi - \Gamma\mathrm{L}\, f(\overline{x}) + \epsilon$. There exists $\{x_k\} \subset \Omega$ such that $f(x_k) = \Gamma\mathrm{L}\, f(x_k) \to \Gamma\mathrm{L}\, f(\overline{x})$, hence for $k > \overline{k}$

$$L \leq x_k \bullet \xi - f(x_k) + 2\epsilon \leq \sup_{x\in\Omega} (x \bullet \xi - f(x)) + 2\epsilon.$$

Since $\epsilon > 0$ is arbitrary, $L \leq \sup_{x\in\Omega}(x \bullet \xi - f(x))$ and the proof is completed. □

The polar of a closed convex set is subsumed to the Fenchel transform, too. In fact, if K is a closed convex set, its indicatrix function, see (2.72), is l.s.c. and convex; hence

$$(I_K)^*(\xi) := \sup_{x\in\mathbb{R}^n} \Big(\xi \bullet x - I_K(x)\Big) = \sup_{x\in K} \xi \bullet x\,. \tag{2.77}$$

Therefore,

$$K^* = \Big\{\xi \,\Big|\, x \bullet \xi \le 1 \ \forall x \in K\Big\} = \Big\{\xi \,\Big|\, (I_K)^*(\xi) \le 1\Big\}.$$

2.114 Definition. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a function. Its* bipolar *is defined as the function $f^{**}(x) := (f^*)^*(x) : \mathbb{R}^n \to \mathbb{R}$, i.e.,*

$$f^{**}(x) := \sup\Big\{\xi \bullet x - f^*(\xi) \,\Big|\, \forall \xi \in \mathbb{R}^n\Big\}.$$

2.115 ¶. Let $\ell(x) := \eta \bullet x + \beta$ be a linear affine map on $\mathbb{R}^n$. Prove that

$$\ell^*(\xi) = \begin{cases} +\infty & \text{if } \xi \ne \eta, \\ -\beta & \text{if } \xi = \eta, \end{cases}$$

and that $(\ell^*)^*(x) = \eta \bullet x + \beta = \ell(x)$.

2.116 Proposition. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a function. Then*

(i) $f^{**} \le f$,
(ii) $f^{**} \le g^{**}$ *if* $f \le g$,
(iii) f^{**} *is the largest l.s.c. convex minorant of f,*

$$f^{**}(x) = \Gamma\mathrm{L}\, f(x) = \sup\Big\{\ell(x) \,\Big|\, \ell : \mathbb{R}^n \to \mathbb{R},\ \ell \text{ affine},\ \ell \le f\Big\}.$$

Proof. (i) From the definition of f^* we have $\xi \bullet x - f^*(\xi) \le f(x)$, hence $f^{**}(x) = \sup_{\xi\in\mathbb{R}^n}(\xi \bullet x - f^*(\xi)) \le f(x)$.

(ii) if $f \le g$, then $g^* \le f^*$ hence $(f^*)^* \le (g^*)^*$.

(iii) f^{**} is convex and l.s.c. , hence $f^{**} = \Gamma\mathrm{L}\, f^{**}$. On the other hand, every linear affine minorant ℓ of f is also an affine linear minorant for f^{**}, since $\ell = \ell^{**} \le f^{**}$. Therefore $\Gamma\mathrm{L}\, f^{**} = \Gamma\mathrm{L}\, f$. □

The following theorem is an easy consequence of Proposition 2.116.

2.117 Theorem. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. Then we have the following:*

(i) *f is convex and l.s.c. if and only if $f = f^{**}$.*
(ii) *Assume that f is convex and $f(x) < +\infty$ at some $x \in \mathbb{R}^n$. Then $f(x) = f^{**}(x)$ if and only if f is l.s.c. at x.*
(iii) *f^* is an involution on the class of proper, convex and l.s.c. functions.*

Proof. Since $f^{**}(x) = \Gamma\mathrm{L}\, f(x)$, (i) and (ii) are a rewriting of (ii) and (iii) of Theorem 2.109.

(iii) Let f be convex, l.s.c, and proper. By (ii) of Proposition 2.112 $f^*(\xi) > -\infty$ for every ξ if and only if $f(x) < +\infty$ at some x, and $f^{**}(x) > -\infty$ for every x if and only if $f^*(\xi) < +\infty$ at some ξ. Since $f^{**} = f$ by (i), we conclude that f^* is proper. Similarly one proves that $f = f^{**}$ is proper if f^* is convex, l.s.c and proper. □

d. Convex duality revisited

Fenchel duality resumes the mystery of convex duality. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a function and consider the *primal problem*

$$(\mathcal{P}) \qquad f(x) \to \min$$

and let

$$p := \inf_x f(x).$$

Introduce a function $\phi(x, b) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ such that $\phi(x, 0) = f(x)$ and consider the *value function* of problem $(\mathcal{P})$ (associated to the "perturbation ϕ")

$$v(b) := \inf_x \phi(x, b). \tag{2.78}$$

We have $v(0) = p$.

Compute now the polar $v^*(\xi)$, $\xi \in \mathbb{R}^m$, of the *value function* $v(b)$. The *dual problem* of problem $(\mathcal{P})$ by means of the chosen perturbation $\phi(x, b)$ is the problem

$$(\mathcal{P}^*) \qquad -v^*(\xi) \to \max.$$

Let $d := \sup_\xi -v^*(\xi)$. Then $v^{**}(0) = d$, in fact,

$$v^{**}(0) = \sup_\xi \Big\{ 0 \bullet \xi \, - v^*(\xi) \Big\} = d.$$

The following theorem connects the existence of a maximizer of the dual problem $(\mathcal{P}^*)$ with the regularity properties of the value function v of the primal problem $(\mathcal{P})$. This is the true content of convex duality.

2.118 Theorem. *With the previous notations we have the following:*

(i) $p \geq d$.
(ii) *Assume v convex and $v(0) < +\infty$. Then $p = d$ if and only if v is l.s.c. at* 0.
(iii) *Assume v convex and $v(0) \in \mathbb{R}$. Then $v(b) \geq \eta \bullet b + v(0)$ $\forall b$ if and only if v is l.s.c. at 0 (equivalently $p = d$ by (ii)) and η is a maximizer for problem $(\mathcal{P}^*)$.*

In particular, if v is convex and continuous at 0, then $p = d$ and $(\mathcal{P}^*)$ has a maximizer.

Proof. (i) Since $v^{**} \leq v$ from Proposition 2.116, we get $d = v^{**}(0) \leq v(0) = p$.

(ii) Since $p = d$ means $v(0) = v^{**}(0)$, (ii) follows from (ii) of Theorem 2.117.

(iii) Assume v convex and $v(0) \in \mathbb{R}$. If $v(b) \geq \eta \bullet b + v(0)$ $\forall b$, we infer $v(0) = v^{**}(0)$, hence by (ii), we conclude that v is l.s.c. at 0. Moreover, the inequality $v(b) \geq \eta \bullet b + v(0)$ $\forall b$ is equivalent to $v(0) + v^*(\eta) = 0$ by the Fenchel inequality. Consequently, $-v^*(\eta) = v(0) = v^{**}(0) = d$, i.e., η is a maximizer for $(\mathcal{P}^*)$.

Conversely, if η maximizes $(\mathcal{P}^*)$ and v is l.s.c. at 0, then we have $-v^*(\eta) = d = v^{**}(0)$ and $v(0) = v^{**}(0)$ by (ii). Therefore $v(0) + v^*(\eta) = 0$, which is equivalent to $v(b) \geq \eta \bullet b + v(0)$ $\forall b$ by the Fenchel inequality. $\square$

The following proposition yields a sufficient condition for applying Theorem 2.118.

2.119 Proposition. *With the previous notations, assume that ϕ is convex and that there exists x_0 such that $p \mapsto \phi(x_0, p)$ is continuous at 0. Then v is convex and $0 \in \text{int}(\text{dom}(v))$. If, moreover, $v(0) > -\infty$, then v is continuous at 0.*

Proof. Let us prove that v is convex since ϕ is convex. Choose $p, q \in \mathbb{R}^n$ and $\lambda \in [0,1]$. We have to prove that $v(\lambda p + (1-\lambda)q) \le \lambda v(p) + (1-\lambda)v(q)$. It is enough to assume $v(p), v(q) < +\infty$. For $a > v(p)$ and $b > v(q)$, let $\overline{x}$ and $\overline{y}$ be such that

$$v(p) \le \phi(\overline{x}, p) \le a, \qquad v(q) \le \phi(\overline{y}, q) \le b.$$

Then we have

$$\begin{aligned} v(\lambda p + (1-\lambda)q) &= \inf_z \phi(z, \lambda p + (1-\lambda)q) \le \phi(\lambda\overline{x} + (1-\lambda)\overline{y}, \lambda p + (1-\lambda)q) \\ &\le \lambda\phi(\overline{x}, p) + (1-\lambda)\phi(\overline{y}, q) \le \lambda a + (1-\lambda)b. \end{aligned}$$

Letting $a \to v(p)$ and $b \to v(q)$ we prove the convexity inequality.

(ii) Since $\phi(x_0, b)$ is continuous at 0, $\phi(x_0, b)$ is bounded near 0; i.e., for some $\delta, M > 0$, $\phi(x_0, b) \le M$ $\forall b \in B(0, \delta)$. Therefore

$$v(b) = \inf_x \phi(x, b) \le M \qquad \forall b \in B(0, \delta),$$

i.e., $0 \in \text{int}(\text{dom}(v))$. If, moreover, $v(0) > -\infty$, then v is never $-\infty$. We then conclude that v takes only real values near 0, consequently, v is continuous at 0. □

A more symmetrical description of convex duality follows assuming that the perturbed functional $\phi(x, b)$ is convex and l.s.c.. In this case, we observe that

$$v^*(\xi) = \phi^*(0, \xi),$$

where $\phi^*(p, \xi)$ is the polar of ϕ on $\mathbb{R}^n \times \mathbb{R}^m$. In fact,

$$\begin{aligned} \phi^*(0,\xi) &= \sup_{x,b}\Big\{ 0 \bullet x + b \bullet \xi - \phi(x,b) \Big\} = \sup_{x,b}\Big\{ b \bullet \xi - \phi(x,b) \Big\} \\ &= \sup_b \Big\{ b \bullet \xi - \inf_x \phi(x,b) \Big\} = v^*(\xi). \end{aligned}$$

The dual problem $(\mathcal{P}^*)$ then rewrites as

$$(\mathcal{P}^*) \qquad\qquad -\phi^*(0, \xi) \to \max,$$

and the corresponding value function is then $-w(p)$, $p \in \mathbb{R}^m$,

$$w(p) := \inf_\xi \phi^*(p, \xi).$$

Since $\phi^{**} = \phi$, the dual problem of $(\mathcal{P}^*)$, namely

$$(\mathcal{P}^{**}) \qquad\qquad \phi^{**}(x, 0) \to \min$$

is again $(\mathcal{P})$. We say that $(\mathcal{P})$ and $(\mathcal{P}^*)$ are dual to each other. Therefore convex duality links the equality $\inf_x \phi(x,0) = \sup_\xi \phi^*(0,\xi)$ and the existence of solutions of one problem to the regularity properties of the value function of the dual problem.

There is also a connection between convex duality and min-max properties typical in game theory. Assume for simplicity that $\phi(x,b)$ is convex and l.s.c. The *Lagrangian* associated to ϕ is the function $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$ defined by

$$-L(x,\xi) := \sup_{b\in\mathbb{R}^m} \Big\{ b \bullet \xi - \phi(x,b) \Big\},$$

i.e.,

$$L(x,\xi) = -\phi_x^*(\xi)$$

where $\phi_x(b) := \phi(x,b)$ for every x and b.

2.120 Proposition. *Let ϕ be convex. Then the following hold:*

(i) *For any $x \in \mathbb{R}^n$, $\xi \to L(x,\xi)$ is concave and upper semicontinuous.*
(ii) *For any $\xi \in \mathbb{R}^n$, $x \to L_x(x,\xi)$ is convex.*

Proof. (i) is trivial since $-L$ is the supremum of a family of linear affine functions. For (ii) observe that $L(x,\xi) = \inf_b\{\phi(x,b) - \xi \bullet b\}$. Let $u, v \in \mathbb{R}^n$ and let $\lambda \in [0,1]$. We want to prove that

$$L(\lambda u + (1-\lambda)v) \le \lambda L(u,\xi) + (1-\lambda)L(v,\xi). \tag{2.79}$$

It is enough to assume that $L(u,\xi) < +\infty$ and $L(v,\xi) < +\infty$. For $a > L(u,\xi)$ and $b > L(v,\xi)$ let $b, c \in \mathbb{R}^m$ be such that

$$\begin{aligned} L(u,\xi) &\le \phi(u,b) - \xi \bullet b \le \alpha, \\ L(v,\xi) &\le \phi(v,c) - \xi \bullet c \le \beta. \end{aligned}$$

Then we have

$$\begin{aligned} L(\lambda u + (1-\lambda)v, \xi) &\le \phi(\lambda u + (1-\lambda)v, \lambda b + (1-\lambda)c) - \xi \bullet \lambda b + (1-\lambda)c \\ &\le \lambda\phi(u,b) + (1-\lambda)\phi(v,c) - \lambda\,\xi \bullet b - (1-\lambda)\,\xi \bullet c \\ &\le \lambda\alpha + (1-\lambda)\beta. \end{aligned}$$

Letting $\alpha \downarrow L(u,b)$ and $\beta \downarrow L(v,c)$, (2.79) follows. □

Observe that

$$\begin{aligned} \phi^*(p,\xi) &= \sup_{x,b}\{ p \bullet x + b \bullet \xi - \phi(x,b)\} \\ &= \sup_x\{ p \bullet x + \sup_b\{ b \bullet \xi - \phi(x,b)\}\} \\ &= \sup_x\{ p \bullet x - L(x,\xi)\}. \end{aligned} \tag{2.80}$$

Consequently,

$$d = \sup_\xi -\phi^*(0,\xi) = \sup_\xi \inf_x L(x,\xi). \tag{2.81}$$

On the other hand, for every x, $b \to \phi_x(b)$ is convex and l.s.c., hence

$$\phi(x,b) = \phi_x(b) = \phi_x^{**}(b) = \sup_\xi \{\, b \bullet \xi \; - \phi_x^*(\xi)\} = \sup_\xi \{\, b \bullet \xi \; + L(x,\xi)\}.$$

Consequently,

$$p = \inf_x \phi(x,0) = \inf_x \sup_\xi L(x,\xi).$$

Therefore, the inequality $d \le p$ is a min-max inequality $\sup_\xi \inf_x L(x,\xi) \le \inf_x \sup_\xi L(x,\xi)$ for the Lagrangian, see Section 2.4.8. In particular, the existence of solutions for both $(\mathcal{P})$ and $(\mathcal{P}^*)$ is related to the existence of *saddle points* for the Lagrangian, see Proposition 2.88.

The above applies surprisingly well in quite a number of cases.

2.121 Example. Let φ be convex and l.s.c. Consider the perturbed function $\phi(x,b) := \varphi(x+b)$. The value function $v(b)$ is then constant, $v(b) = v(0)$ $\forall b$, hence convex and l.s.c. Its polar is

$$v^*(\xi) := \sup_x \{\, \xi \bullet b \; - v(0)\} = \begin{cases} +\infty & \text{if } \xi \ne 0, \\ -v(0) & \text{if } \xi = 0. \end{cases}$$

The dual problem has then a maximum point at $\xi = 0$ with maximum value $d = v(0)$. Finally, we compute its Lagrangian: Changing variable $c := x + b$,

$$L(x,\xi) = -\sup_b \{\, \xi \bullet b \; - \varphi(x+b)\} = -\sup_c \{\, \xi \bullet c \; - \; \xi \bullet x \; - \varphi(c)\} = \xi \bullet x \; - \varphi^*(\xi).$$

Let $\varphi, \psi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be two convex functions and consider the *primal problem*

$$\text{Minimize } \varphi(x) + \psi(x), \qquad x \in \mathbb{R}^n. \tag{2.82}$$

Introduce the perturbed function

$$\phi(x,b) = \varphi(x+b) + \psi(x), \qquad (x,b) \in \mathbb{R}^n \times \mathbb{R}^n, \tag{2.83}$$

for which $\phi(x,0) = \varphi(x) + \psi(x)$, and the corresponding *value function*

$$v(b) := \inf_x (\varphi(x) + \psi(x)). \tag{2.84}$$

Since ϕ is convex, then the value function v is convex, whereas the Lagrangian $L(x,\xi)$ is convex in x and concave in ξ. Let us first compute the Lagrangian. We have

$$-L(x,\xi) := \sup_b \{\, \xi \bullet b \; - \varphi(x+b) - \psi(x)\} = \varphi^*(\xi) - \; \xi \bullet x \; - \psi(x)$$

so that

$$L(x,\xi) = \psi(x) + \; \xi \bullet x \; - \varphi^*(\xi).$$

Now we compute the polar of ϕ. We have

$$\begin{aligned} \phi^*(p,\xi) &= \sup_x \{\, p \bullet x \; - L(x,\xi)\} = \sup_x \{\, p \bullet x \; - \; \xi \bullet x \; - \psi(x) + \varphi^*(\xi)\} \\ &= \sup_x \{\, (p-\xi) \bullet x \; - \psi(x)\} + \varphi^*(\xi) \\ &= \psi^*(p-\xi) + \varphi^*(\xi). \end{aligned}$$

Therefore, the polar of (2.84) is

$$v^*(\xi) = \phi^*(0,\xi) = \varphi^*(\xi) + \psi^*(-\xi) \qquad \forall \xi \in \mathbb{R}^n.$$

As an application of the above we have the following.

2.122 Theorem. *Let φ and ψ be as before, and let ϕ and v be defined by* (2.83) *and* (2.84). *Assume that we have φ continuous at x_0, $\psi(x_0) < +\infty$ at some point x_0 and that $v(0) > -\infty$. Let p and d be defined by the primal and dual optimization problems respectively, through $(x,b) \to \varphi(x+b) + \psi(x)$, given by*

$$p := \inf_x (\varphi(x) + \psi(x)), \tag{2.85}$$

$$d := \sup_\xi (-\varphi^*(\xi) - \psi^*(-\xi)). \tag{2.86}$$

Then $p = d \in \mathbb{R}$ and problem (2.86) *has a maximizer.*

Proof. $\phi(x,b) := \varphi(x+b) + \psi(x)$ is convex. Moreover, since φ is continuous at x_0, then $b \to \phi(x_0, b)$ is continuous at 0. From Proposition 2.119 we then infer that v is convex and continuous at 0. Then the conclusions follow from Theorem 2.118. □

2.123 Example. Let φ be convex. Choose as perturbed functional

$$\phi(x,b) = \varphi(x+b) + \varphi(x)$$

for which $\phi(x,0) = 2\varphi(x)$. Then, by the above,

$$v^*(\xi) = \varphi^*(\xi) + \varphi^*(-\xi)$$

and the Lagrangian is

$$L(x,\xi) = \varphi(x) + \xi \bullet x - \varphi^*(\xi).$$

Let us consider the convex minimization problem already discussed in Paragraph d. Here we extend it a little further.

Let $f, g^1, \dots, g^m : \mathbb{R}^n \subset \mathbb{R}^n \to \mathbb{R}\cup_{+\infty}$ be convex functions defined on $\mathbb{R}^n$. We assume for simplicity that either f or $g := (g^1, g^2, \dots, g^m)$ are continuous. Consider the *primal* minimization problem

$$\text{Minimize } f(x) \text{ with the constraints } g(x) \le 0. \tag{2.87}$$

Let I_K be the indicatrix of the closed convex set $K := \{x = (x_i) \in \mathbb{R}^n \,|\, x_i \le 0 \ \forall i\}$. Problem (2.87) amounts to

$$(\mathcal{P}) \qquad\qquad f(x) + I_K(g(x)) \to \min.$$

Let us introduce the perturbed function

$$\phi(x,b) := f(x) + I_K(g(x) - b)$$

which is convex. Consequently, the associated value function

$$v(b) := \sup_x (f(x) + I_K(\varphi(x) - b)), \qquad b \in \mathbb{R}^m, \tag{2.88}$$

is convex by Proposition 2.119. Now, compute the polar of the value function. First we compute the polar of $I_K(y)$. We have

$$I_K^*(\xi) = \sup_b \{\, \xi \bullet b \,-\, I_K(b)\} = \begin{cases} 0 & \text{if } \xi \geq 0, \\ +\infty & \text{if } \xi < 0. \end{cases}$$

Therefore, changing variables, $c = g(x) - b$,

$$\begin{aligned} -L(x,\xi) &= \sup_b \{\, \xi \bullet b \,-\, f(x) - I_K(g(x) - b)\} \\ &= -f(x) + \sup_c \{\, \xi \bullet g(x) \,-\, \xi \bullet c \,-\, I_k(c)\} \\ &= -f(x) + \, g(x) \bullet \xi \,+ (I_K)^*(-\xi), \end{aligned}$$

hence

$$L(x,\xi) = \begin{cases} f(x) - \, \xi \bullet g(x) & \text{if } \xi \leq 0, \\ -\infty & \text{if } \xi > 0. \end{cases}$$

Notice that $\sup_\xi L(x,\xi) = f(x) + I_K(g(x)) = \phi(x,0)$. Consequently,

$$\phi^*(p,\xi) = \inf_x \, p \bullet x \,-\, L(x,\xi) = \begin{cases} +\infty & \text{if } \xi > 0 \\ \sup_x \{\, p \bullet x \,-\, f(x) + \, g(x) \bullet \xi \,\} & \text{if } \xi \leq 0, \end{cases}$$

and the polar of the value function is

$$v^*(\xi) = \phi^*(0,\xi) = \sup_x \{\, g(x) \bullet \xi \,-\, f(x)\}.$$

Consequently, the dual problem through the perturbation ϕ is

$$(\mathcal{P}^*) \qquad -v^*(\xi) := \inf_x \{ f(x) - \, \xi \bullet \varphi(x) \,\} \to \max \text{ on } \{\xi \geq 0\}.$$

2.124 Theorem. *Let $f, g^1, \dots, g^m : \mathbb{R}^n \subset \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be convex functions defined on $\mathbb{R}^n$. Let p and d be defined by the primal and dual optimization problems*

$$p : = \inf_x (f(x) + I_K(g(x))), \tag{2.89}$$

$$d := \sup_\xi \inf_x L(x,\xi). \tag{2.90}$$

Assume that $p > -\infty$ and that the Slater condition *holds (namely there exists $x_0 \in \mathbb{R}^n$ such that $f(x_0) < +\infty$, $g(x_0) < 0$ and g is continuous at x_0). Then the dual problem has a maximizer.*

Proof. The function $\phi(x,b) = f(x) + I_K(g(x) - b)$ is convex. Moreover the Slater condition implies that $\phi(x_0, b)$ is continuous at 0. We then infer from Proposition 2.119 that the value function v is convex and continuous at 0. The claims then follow from Theorem 2.118. □

2.6 Exercises

2.125 ¶. Prove that the *n-parallelepiped* of $\mathbb{R}^n$ generated by the vectors $e_1, \dots, e_n$ with vertex at 0,

$$K := \{x = \lambda_1 e_1 + \cdots + \lambda_n e_n \;\;|\;\; 0 \le \lambda_i \le 1,\; i = 1, \dots, n\},$$

is convex.

2.126 ¶. $K_1 + K_2,\; \alpha K_1,\; \lambda K_1 + (1-\lambda)K_2,\; \lambda \in [0,1]$, are all convex sets if K_1 and K_2 are convex.

2.127 ¶. Show that the convex hull of a closed set is not necessarily closed.

2.128 ¶. Find out which of the following functions is convex:

$$3x^2 + y^y - 4z^2, \qquad x + x^2 + y^2, \qquad (x+y+1)^p \text{ in } x+y+1>0,$$
$$\exp(xy), \qquad \log(1+x^2+y^2), \qquad \sin(x^2+y^2).$$

2.129 ¶. Let K be a convex set. Prove that the following are convex functions:

(i) The *support function* $\delta(x) := \sup\{\, x \bullet y \,|\, y \in K\}$.
(ii) The *gauge function* $\gamma(x) := \inf\{\lambda \ge 0 \,|\, x \in \lambda K\}$.
(iii) The *distance function* $d(x) := \inf\{|x-y| \,|\, y \in K\}$.

2.130 ¶. Prove that $K \subset \mathbb{R}^n$ is a convex body with $0 \in \mathrm{int}(K)$ if and only if there is a gauge function $F : \mathbb{R}^n \to \mathbb{R}$ such that $K = \{x \in \mathbb{R}^n \,|\, F(x) \le 1\}$.

2.131 ¶. Let $K \subset \mathbb{R}^n$ with $0 \in K$, and for every $\xi \in \mathbb{R}^n$ set

$$d(\xi) := \inf\Big\{ d \in \mathbb{R} \,\Big|\, \xi \bullet x \le d \;\forall x \in K \Big\}.$$

Prove that if K is convex with $0 \in \mathrm{int}(K)$, then $d(\xi)$ is a gauge function, i.e.,

$$d(\xi) := \min\Big\{ \xi \bullet x \,\Big|\, x \in K \Big\}$$

and

$$K^* := \Big\{ \xi \in \mathbb{R}^n \,\Big|\, d(\xi) \le 1 \Big\}.$$

2.132 ¶. Let $f : \mathbb{R}_+ \to \mathbb{R}$ be strictly convex with $f(0) = 0$ and $f'(0) = 0$. Write $\alpha(s) := (f')^{-1}(s)$ and prove that

$$f(x) := \int_0^x f'(s)\, ds, \qquad \mathcal{L}_f(y) := \int_0^y \alpha(s)\, ds, \quad y \ge 0.$$

2.133 ¶. Let $f : [0,1] \times [0,1] \to \mathbb{R}$ be a function which is continuous with respect to each variable separately. As we know, f need not be continuous. Prove that $f(t,x)$ is continuous if it is convex in x for every t.

2.134 ¶. Let $C \subset \mathbb{R}^n$ be a closed convex set. Prove that $x_0 \in C$ is an extreme point if and only if $C \setminus \{x_0\}$ is convex.

2.135 ¶. Let $C \subset \mathbb{R}^n$ be a closed convex set and let $f : C \to \mathbb{R}$ be a continuous, convex and bounded function. Prove that $\sup_C f = \sup_{\partial C} f$.

2.136 ¶. Let S be a set and $C = \mathrm{co}(S)$ its convex hull. Prove that $\sup_C f = \sup_S f$ if f is convex on C.

2.137 ¶. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function and let f_ϵ be its ϵ-mollified where k is a regularizing kernel. Prove that f_ϵ is convex.

2.138 ¶. Let $\varphi : \mathbb{R} \to \mathbb{R}$, $\varphi \geq 0$. Then $f(x,t) := \frac{\varphi(x)}{t}$ is convex in $\mathbb{R}\times]0,\infty[$ if and only if $\sqrt{\varphi}$ is convex.

2.139 ¶. Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a convex function. Prove the following:

(i) If $\Gamma f(x) \neq f(x)$, then $x \in \partial\,\mathrm{dom}(f)$.
(ii) If $\mathrm{dom}(f)$ is closed and f is l.s.c. in $\mathrm{dom}(f)$, then $\Gamma f = f$ everywhere.
(iii) $\inf f = \inf \Gamma f$.
(iv) For all $\alpha \in \mathbb{R}$ we have $\{x \in \mathbb{R}^n \,|\, \Gamma f(x) \leq \alpha\} = \cap_{\beta > \alpha} \mathrm{cl}(\{x \in \mathbb{R}^n \,|\, f(x) \leq \beta\})$.
(v) If f_1 and f_2 are convex functions with $f_1 \leq f_2$, then $\Gamma f_1 \leq \Gamma f_2$.

2.140 ¶. Let f be a l.s.c. convex function and denote by $\mathcal{F}$ the class of affine functions $\ell : \mathbb{R}^n \to \mathbb{R}$ with $\ell(y) \leq f(y)$ $\forall y \in \mathbb{R}^n$. From Theorem 2.109

$$f(x) = \sup\Big\{\ell(x) \,\Big|\, \ell \in \mathcal{F}\Big\}.$$

Prove that there exists an at most denumerable subfamily $\{\ell_n\} \subset \mathcal{F}$ such that $f(x) = \sup_n \ell_n(x)$.
[*Hint.* Recall that every covering has a denumerable subcovering.]

2.141 ¶. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a function. Its *convex l.s.c. envelope* $\Gamma\mathrm{C}\, f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is defined by

$$\Gamma\mathrm{C}\, f(x) := \sup\Big\{g(x) \,\Big|\, g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\},\ g \text{ convex and l.s.c., } g \leq f\Big\}.$$

Prove that $\Gamma\mathrm{C}\, f = \Gamma\mathrm{L}\, f$.
[*Hint.* Apply Theorem 2.109 to the convex and l.s.c. minorants of f.]

2.142 ¶. Prove the following: If $\{f_i\}_{i\in I}$ is a family of convex and l.s.c. functions $f_i : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, then

$$\Big(\inf_{i\in I} f_i\Big)^* = \sup_{i\in I} f_i^*, \qquad \Big(\sup_{i\in I} f_i\Big)^* \leq \inf_{i\in I} f_i^*.$$

2.143 ¶. Prove the following claims:

(i) Let $f(x) := \frac{1}{p}|x|^p$, $p > 1$. Then $f^*(\xi) = \frac{1}{q}|\xi|^q$, $1/p + 1/q = 1$.
(ii) Let $f(x) := |x|$, $x \in \mathbb{R}^n$. Then

$$f^*(\xi) = \begin{cases} 0 & \text{if } |\xi| \leq 1, \\ +\infty & \text{if } |\xi| > 1. \end{cases}$$

(iii) Let $f(t) := e^t$, $t \in \mathbb{R}$. Then

$$f^*(\xi) = \begin{cases} +\infty & \text{if } y \leq 0, \\ 0 & \text{if } y = 0, \\ \xi(\log \xi - 1) & \text{if } y > 0. \end{cases}$$

(iv) Let $f(x) := \sqrt{1+|x|^2}$. Then $\mathcal{L}_f$ is defined on $\Omega^* := \{\xi \mid |\xi| < 1\}$ and

$$\mathcal{L}_f(\xi) = -\sqrt{1-|\xi|^2},$$

consequently,

$$f^*(\xi) = \Gamma\mathrm{L}\,\mathcal{L}_f(\xi) = \begin{cases} -\sqrt{1-|\xi|^2} & \text{if } |\xi| \le 1, \\ +\infty & \text{if } |\xi| > 1. \end{cases}$$

(v) The function $f(x) = \frac{1}{2}|x|^2$ is the unique function for which $f^*(x) = f(x)$.

2.144 ¶. Show that the following computation rules hold.

Proposition. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a function. Then the following hold:*
(i) $(\lambda f)^*(\xi) = \lambda f^*(\xi/\lambda)$ $\forall \xi \in \mathbb{R}^n$ *and* $\forall \lambda > 0$.
(ii) *If we set* $f_y(x) := f(x-y)$, *then we have* $f_y^*(\xi) = f^*(\xi) + \xi \bullet y$ $\forall \xi \in \mathbb{R}^n$ *and* $\forall y \in \mathbb{R}^n$.
(iii) *Let* $\mathbf{A} \in M_{N,n}(\mathbb{R})$, $N \le n$, *be of maximal rank and let* $g(x) := f(\mathbf{A}x)$. *Then*

$$g^*(\xi) = \begin{cases} +\infty & \textit{if } \xi \notin \ker \mathbf{A}^\perp, \\ f^*(\mathbf{A}^{-T}\xi) & \textit{if } \xi \in \ker \mathbf{A}^\perp = \operatorname{Im} \mathbf{A}^T. \end{cases}$$

2.145 ¶. Let $A \subset \mathbb{R}^n$ and $I_A(x)$ be its indicatrix, see (2.72). Prove the following:
(i) If L is a linear subspace if $\mathbb{R}^n$, then $(I_L)^* = I_{L^\perp}$.
(ii) If C is a closed cone with the origin as vertex, then $(I_C)^*$ is the indicatrix function of the cone generated by the vectors through the origin that are orthogonal to C.

http://www.springer.com/978-0-8176-8309-2

Mathematical Analysis
Foundations and Advanced Techniques for Functions of Several Variables
Giaquinta, M.; Modica, G.
2012, XIII, 405 p. 66 illus., Hardcover
ISBN: 978-0-8176-8309-2
A product of Birkhäuser Basel