2014

Helly's Theorem and its Equivalences via Convex Analysis

Adam Robinson
Portland State University

Helly's Theorem and Its Equivalences via Convex Analysis

Adam Robinson

Advisor: Dr. Mau Nam Nguyen

MTH 401 Honors Project

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Bachelor of Science

at

Portland State University

2014

Helly's Theorem and Its Equivalences via Convex Analysis

Adam Robinson

Advisor: Dr. Mau Nam Nguyen

Portland State University, 2014

ABSTRACT

Helly's theorem is an important result from Convex Geometry. It gives sufficient conditions for a family of convex sets to have a nonempty intersection. A large variety of proofs as well as applications are known. Helly's theorem also has close connections to two other well-known theorems from Convex Geometry: Radon's theorem and Carathéodory's theorem. In this project we study Helly's theorem and its relations to Radon's theorem and Carathéodory's theorem by using tools of Convex Analysis and Optimization. More precisely, we will give a novel proof of Helly's theorem, and in addition we show in a complete way that these three famous theorems are equivalent in the sense that using one of them allows us to derive the others.

Contents

Introduction

In Convex Geometry, geometric properties of convex sets and functions are investigated. The foundations of this field were developed by many accomplished mathematicians, such as Hermann Brunn, Hermann Minkowski, Werner Fenchel, Constantin Carathéodory, and Eduard Helly. At the beginning of the 1960's, *Convex Analysis* was grown out of Convexity, and this new field was systematically developed by the works of R. Tyrrell Rockafellar, Jean-Jacques Moreau, and others. Convex Analysis is more concerned with the generalized differentiation theory of convex functions and sets rather than only with their geometric properties. The presence of the convexity makes it possible to develop calculus rules for a generalized derivative concept called the *subdifferential*, which can be used to deal with convex functions that are not necessarily differentiable. Convex Analysis then becomes the mathematical foundation for Convex Optimization, a fast growing field with numerous applications to Control Systems, Estimation and Signal Processing, Communications and Networks, Electronic Circuit Design, Data Analysis and Modeling, Statistics, Economics and Finance, etc.

In this project we use Convex Analysis and Optimization to study some basic results of Convex Geometry. We mainly focus on a theorem introduced by Eduard Helly in 1913 which gives sufficient conditions for a family of convex sets to have a nonempty intersection. Using modern tools from convex analysis and optimization, we will study Helly's theorem from both theoretical and numerical viewpoints. In particular, we will give a novel proof, via Convex Analysis, of Helly's theorem and its connections to Radon's theorem and Carathéodory's theorem, which are also important results from Convex Geometry.

It has been mentioned in several references that the above-mentioned theorems of Helly, Radon, and Carathéodory are equivalent in the sense that using one of them allows us to derive the others. However, in [8, p. 47] P. M. Gruber says that "*we were not able to locate in the literature a complete proof in the context of* $\mathbb{R}^n$"; see also the excellent surveys [5] and [6], as well as the monograph [3]. We will use the tools of Convex Analysis to provide a complete treatment for the equivalence of these theorems. The analysis involves the use of generalized differentiation properties of the class of distance functions associated with convex sets. This class of functions, which reflects the connection between convex functions and sets, allows us to give a simple, self-contained, complete proof of the described equivalence. Further on, using the concept of distance function we are able to study effective numerical algorithms for finding a point in the intersection of the given family of convex sets, whose existence is guaranteed by Helly's theorem.

Chapter 1

Elements of Convex Analysis and Optimization

In this chapter we introduce some important concepts and results of Convex Analysis and Optimization that will be used in the subsequent chapters. The materials presented here can be found in many books on Convex Analysis and Optimization; see, e.g., [10, 11] and the references therein.

Throughout the thesis we consider the Euclidean space $\mathbb{R}^n$ equipped with the Euclidean norm of an element $x = (x_1, \ldots, x_n)$ given by

$$\|x\| := \sqrt{x_1^2 + \ldots + x_n^2},$$

and the inner product of any two elements $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ given by

$$\langle x, y \rangle := x_1 y_1 + \cdots + x_n y_n.$$

It follows from the definition that $\|x\|^2 = \langle x, x \rangle$.

A subset Ω of $\mathbb{R}^n$ is said to be *convex* if

$$\lambda x + (1-\lambda)y \in \Omega$$

whenever $x, y \in \Omega$ and $\lambda \in (0,1)$. Geometrically, a subset Ω is convex if for any $x, y \in \Omega$, the line segment connecting x and y belongs to the set.

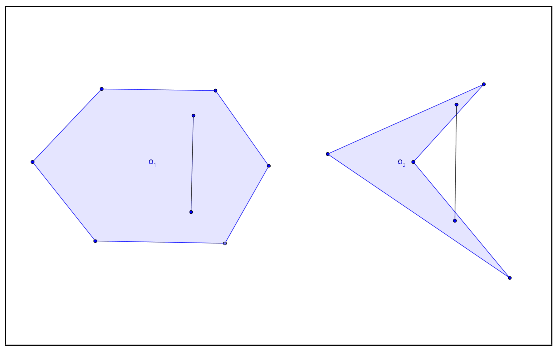

FIGURE 1.0.1: Convex set Ω_1 and nonconvex set Ω_2.

From the definition, it is obvious that if $\{\Omega_i\}_{i\in I}$ is a collection of convex sets in $\mathbb{R}^n$, then the intersection $\bigcap_{i\in I}\Omega_i$ is also convex. In particular, the intersection of any two convex sets is also a convex set. This property motivates the definition of the *convex hull* of an arbitrary subset of $\mathbb{R}^n$. Given a subset $\Omega \subset \mathbb{R}^n$, define the convex hull of Ω by

$$\operatorname{co}\Omega := \bigcap \{C \mid C \text{ is convex and } \Omega \subseteq C\}.$$

Equivalently, the convex hull of a set Ω is the smallest convex set containing Ω.

The following important result is a direct consequence of the definition.

Proposition 1.0.1. *For any convex subset Ω of $\mathbb{R}^n$, its convex hull admits the representation*

$$\operatorname{co}\Omega = \Big\{ \sum_{i=1}^{m} \lambda_i w_i \;\Big|\; \sum_{i=1}^{m} \lambda_i = 1,\ \lambda_i \geq 0,\ w_i \in \Omega,\ m \in \mathbb{N} \Big\}.$$

A function $f\colon \mathbb{R}^n \to \mathbb{R}$ defined on a convex set Ω is called *convex* if

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$$

for all $x, y \in \Omega$ and $\lambda \in (0,1)$. If this inequality becomes strict for whenever $x \neq y$, $x, y \in \Omega$, we say that f is *strictly convex* on Ω.

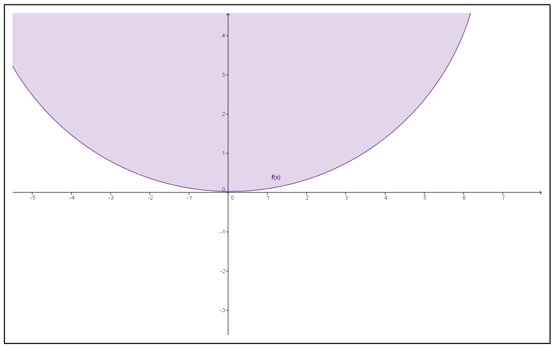

FIGURE 1.0.2: A convex function f(x) and its epigraph, the shaded region above f(x).

A mapping $B\colon \mathbb{R}^n \to \mathbb{R}^m$ is called *affine* if there exist an $m \times n$ matrix A and an element $b \in \mathbb{R}^m$ such that

$$B(x) = Ax + b \text{ , for all } x \in \mathbb{R}^n.$$

Let us present in the proposition below some operations that preserve convexity of functions.

Proposition 1.0.2. **(i)** *Let $f_i\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be convex functions for all $i = 1, \ldots, m$. Then the following functions are convex as well:*

- *The multiplication by scalars λf for any $\lambda > 0$.*
- *The sum function $\sum_{i=1}^m f_i$.*
- *The maximum function $\max_{1 \leq i \leq m} f_i$.*

(ii) *Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex and let $\varphi : \mathbb{R} \to \overline{\mathbb{R}}$ be nondecreasing. Then the composition $\varphi \circ f$ is convex.*

(iii) *Let $B : \mathbb{R}^n \to \mathbb{R}^p$ be an affine mapping and let $f : \mathbb{R}^p \to \overline{\mathbb{R}}$ be a convex function. Then* *the composition $f \circ B$ is convex.*

(iv) *Let $f_i : \mathbb{R}^n \to \overline{\mathbb{R}}$, for $i \in I$, be a collection of convex functions with a nonempty index set I. Then the supremum function $f(x) := \sup_{i \in I} f_i(x)$ is convex.*

The class of distance functions presented in what follows plays a crucial role throughout the thesis. Given a set $\Omega \subset \mathbb{R}^n$, the *distance function* associated with Ω is defined by

$$d(x; \Omega) := \inf \big\{ \|x - \omega\| \;\big|\; \omega \in \Omega \big\}.$$

For each $x \in \mathbb{R}^n$, the *Euclidean projection* from x to Ω is defined by

$$\Pi(x; \Omega) := \big\{ \omega \in \Omega \;\big|\; \|x - \omega\| = d(x; \Omega) \big\}.$$

A function $f : \mathbb{R}^n \to \mathbb{R}$ is called *Lipschitz continuous* on $\mathbb{R}^n$ if there exists a constant $\ell \geq 0$ such that

$$\|f(x) - f(y)\| \leq \ell \|x - y\| \text{ for all } x, y \in \mathbb{R}^n.$$

From the definition, it is obvious that any Lipschitz continuous function is uniformly continuous on $\mathbb{R}^n$. We will prove in the proposition below that the distance function associated with a nonempty set is Lipschitz continuous on $\mathbb{R}^n$.

Proposition 1.0.3. *Given a nonempty set Ω, the distance function $d(\cdot; \Omega)$ is Lipschitz continuous on $\mathbb{R}^n$ with Lipschitz constant $\ell = 1$.*

Proof. We will show that

$$\|d(x; \Omega) - d(y; \Omega)\| \leq \|x - y\| \text{ for all } x, y \in \mathbb{R}^n.$$

For any $\omega \in \Omega$, one has by the triangle inequality that

$$d(x;\Omega) \leq \|x-\omega\| \leq \|x-y\| + \|y-\omega\|.$$

This implies

$$d(x;\Omega) \leq \|x-y\| + \inf\{\|y-\omega\| \mid \omega \in \Omega\} = \|x-y\| + d(y;\Omega),$$

and hence $d(x;\Omega) - d(y;\Omega) \leq \|x-y\|$. Similarly, we have that

$$d(y;\Omega) - d(x;\Omega) \leq \|x-y\|.$$

Therefore, $|d(y;\Omega) - d(x;\Omega)| \leq \|y-x\|$. $\square$

Proposition 1.0.4. *If Ω is a convex set, then the distance function $d(\cdot;\Omega)$ is a convex function.*

Proof. Take any x,$y \in \mathbb{R}^n$ and $\lambda \in (0,1)$. We will show that the distance function given by $f(x) := d(x;\Omega)$ for $x \in \mathbb{R}^n$ satisfies the convex inequality:

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y). \tag{1.0.1}$$

Fix any $\epsilon > 0$. By the properties of the infimum, there exists $u \in \Omega$ such that

$$\|x-u\| < d(x;\Omega) + \frac{\epsilon}{2}.$$

Similarly, there exists $v \in \Omega$ such that

$$\|y-v\| < d(y;\Omega) + \frac{\epsilon}{2}.$$

Since Ω is a convex set, $\lambda u + (1-\lambda)v \in \Omega$. Then

$$\begin{aligned} d(\lambda x + (1-\lambda)y; \Omega) &\le \|\lambda x + (1-\lambda)y - (\lambda u + (1-\lambda)v)\| \\ &\le \|\lambda(x-u)\| + \|(1-\lambda)(y-v)\| \\ &= \lambda\|x-u\| + (1-\lambda)\|y-v\| \\ &\le \lambda\left[d(x;\Omega) + \frac{\epsilon}{2}\right] + (1-\lambda)\left[d(y;\Omega) + \frac{\epsilon}{2}\right] \\ &= \lambda d(x;\Omega) + (1-\lambda)d(y;\Omega) + \epsilon. \end{aligned}$$

Then we obtain (1.0.1) by letting $\epsilon \to 0$. $\square$

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function and let $\bar{x} \in \mathbb{R}^n$. A vector $v \in \mathbb{R}^n$ is called a *subgradient* of f at $\bar{x}$ if

$$\langle v, x - \bar{x}\rangle \le f(x) - f(\bar{x}) \text{ for all } x \in \mathbb{R}^n.$$

The set of all subgradients of f at $\bar{x}$ is called the *subdifferential* of the function at this point and is denoted by $\partial f(\bar{x})$.

Another important concept of Convex Analysis is called the *normal cone* to a nonempty convex set $\Omega \subset \mathbb{R}^n$ at a point $\bar{x} \in \Omega$ and defined by

$$N(\bar{x};\Omega) := \{v \in \mathbb{R}^n \mid \langle v, x - \bar{x}\rangle \le 0 \text{ for all } x \in \Omega\}.$$

In what follows we study subdifferential formulas for distance functions to convex sets. We first pay attention to the case where the reference point belongs to the set.

Proposition 1.0.5. *Suppose that $\bar{x} \in \Omega$. Then*

$$\partial d(\bar{x};\Omega) = N(\bar{x};\Omega) \cap \mathbb{B},$$

where $\mathbb{B}$ is the closed unit ball of $\mathbb{R}^n$.

Proof. Fix any $v \in \partial d(\bar{x}; \Omega)$. Then

$$\langle v, x - \bar{x} \rangle \le d(x; \Omega) - d(\bar{x}; \Omega) = d(x; \Omega) \text{ for all } x \in \mathbb{R}^n. \tag{1.0.2}$$

Since the distance function $d(\cdot; \Omega)$ satisfies a Lipschitz condition with the Lipschitz constant $\ell = 1$, one has

$$\langle v, x - \bar{x} \rangle \le \|x - \bar{x}\| \text{ for all } x \in \mathbb{R}^n.$$

This implies $\|v\| \le 1$ or $v \in \mathbb{B}$. It also follows from (1.0.2) that

$$\langle v, x - \bar{x} \rangle \le 0 \text{ for all } x \in \Omega.$$

Thus $v \in N(\bar{x}; \Omega)$. Therefore, $v \in N(\bar{x}; \Omega) \cap \mathbb{B}$.

Let us now prove the opposite inclusion. Fix any $v \in N(\bar{x}; \Omega) \cap \mathbb{B}$. Then $\|v\| \le 1$ and

$$\langle v, w - \bar{x} \rangle \le 0 \text{ for all } w \in \Omega.$$

Thus for any $x \in \mathbb{R}^n$ and for any $w \in \Omega$, one has

$$\begin{aligned}
\langle v, x - \bar{x} \rangle &= \langle v, x - w + w - \bar{x} \rangle \\
&= \langle v, x - w \rangle + \langle v, w - \bar{x} \rangle \\
&\le \langle v, x - w \rangle \le \|v\| \|x - w\| \le \|x - w\|.
\end{aligned}$$

This implies

$$\langle v, x - \bar{x} \rangle \le d(x; \Omega) = d(x; \Omega) - d(\bar{x}; \Omega),$$

and hence $v \in \partial d(\bar{x}; \Omega)$. $\square$

Now we pay attention to the case where the reference point does not belong to the set.

Proposition 1.0.6. *Let Ω be a nonempty closed convex set and let $\bar{x} \notin \Omega$. Then*

$$\partial d(\bar{x};\Omega) = \Big\{\frac{\bar{x} - \Pi(\bar{x};\Omega)}{d(\bar{x};\Omega)}\Big\}.$$

Proof. Fix $\bar{z} := \Pi(\bar{x};\Omega) \in \Omega$ and fix any $v \in \partial d(\bar{x};\Omega)$. By the definition of subdifferential,

$$\begin{aligned}\langle v, x - \bar{x}\rangle &\le d(x;\Omega) - d(\bar{x};\Omega) = d(x;\Omega) - \|\bar{x} - \bar{z}\| \\ &\le \|x - \bar{z}\| - \|\bar{x} - \bar{z}\| \ \text{ for all } \ x \in \mathbb{R}^n.\end{aligned}$$

Denoting $p(x) := \|x - \bar{z}\|$, we have

$$\langle v, x - \bar{x}\rangle \le p(x) - p(\bar{x}) \ \text{ for all } \ x \in \mathbb{R}^n,$$

which implies that $v \in \partial p(\bar{x}) = \Big\{\dfrac{\bar{x} - \bar{z}}{\|\bar{x} - \bar{z}\|}\Big\}$. Let us show that $v = \dfrac{\bar{x} - \bar{z}}{\|\bar{x} - \bar{z}\|}$ is a subgradient of $d(\cdot;\Omega)$ at $\bar{x}$. Indeed, for any $x \in \mathbb{R}^n$, denote $p_x := \Pi(x;\Omega)$ and get

$$\begin{aligned}\langle v, x - \bar{x}\rangle &= \langle v, x - \bar{z}\rangle + \langle v, \bar{z} - \bar{x}\rangle = \langle v, x - \bar{z}\rangle - \|\bar{x} - \bar{z}\| \\ &= \langle v, x - p_x\rangle + \langle v, p_x - \bar{z}\rangle - \|\bar{x} - \bar{z}\|.\end{aligned}$$

Since $\bar{z} = \Pi(\bar{x};\Omega)$, it follows that $\langle \bar{x} - \bar{z}, p_x - \bar{z}\rangle \le 0$, and so we have $\langle v, p_x - \bar{z}\rangle \le 0$. Using the fact that $\|v\| = 1$ and the Cauchy-Schwarz inequality gives us

$$\begin{aligned}\langle v, x - \bar{x}\rangle &= \langle v, x - p_x\rangle + \langle v, p_x - \bar{z}\rangle - \|\bar{x} - \bar{z}\| \\ &\le \|v\| \cdot \|x - p_x\| - \|\bar{x} - \bar{z}\| = \|x - p_x\| - \|\bar{x} - \bar{z}\| \\ &= d(x;\Omega) - d(\bar{x};\Omega) \ \text{ for all } \ x \in \mathbb{R}^n.\end{aligned}$$

Thus we arrive at $v \in \partial d(\bar{x};\Omega)$. $\square$

In what follows we present some useful subdifferential rules for convex functions.

Proposition 1.0.7. **(i)** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function and let $\bar{x} \in \mathbb{R}^n$. Then the following holds for any $\alpha > 0$:*

$$\partial(\alpha f)(\bar{x}) = \alpha \partial f(\bar{x}).$$

(ii) *Let $f_i : \mathbb{R}^n \to \mathbb{R}$ be convex functions for $i = 1, 2$ and let $\bar{x} \in \mathbb{R}^n$. Then*

$$\partial(f_1 + f_2)(\bar{x}) = \partial f_1(\bar{x}) + \partial f_2(\bar{x}).$$

(iii) *Let $B : \mathbb{R}^n \to \mathbb{R}^m$ be an affine mapping given by $B(x) = Ax + b$, where A is an $m \times n$ matrix. Consider a convex function $f : \mathbb{R}^m \to \mathbb{R}$. Let $\bar{x} \in \mathbb{R}^n$ and let $\bar{y} = B(\bar{x})$. Then*

$$\partial(f \circ B)(\bar{x}) = A^T(\partial f(\bar{y})) = \{A^T(v) \mid v \in \partial f(\bar{y})\}. \tag{1.0.3}$$

Given convex functions $f_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \ldots, m$, define

$$f(x) := \max\{f_i(x) \mid i = 1, \ldots, m\}.$$

The proposition below gives a formula for computing the subdifferential of the maximum function.

Theorem 1.0.8. *Suppose that $f_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \ldots, m$ are convex functions. Then*

$$\partial f(\bar{x}) = \operatorname{co} \bigcup_{i \in I(\bar{x})} \partial f_i(\bar{x}), ,$$

where $I(\bar{x}) := \{i = 1, \ldots, m \mid f_i(\bar{x}) = f(\bar{x})\}$.

Chapter 2

The Theorems of Carathéodory, Radon, and Helly: Their Statements and Proofs

In this chapter we give a survey of the statements and proofs of the theorems of Carathéodory, Radon, and Helly; see, e.g., [10].

2.1 Carathéodory Theorem

Given a set Ω, the *convex cone generated by* Ω, denoted by K_Ω, is the smallest convex cone containing Ω.

Lemma 2.1.1. *Consider a nonempty set* $\Omega \subset \mathbb{R}^n$. *The convex cone generated by* Ω *has the following representation:*

$$K_\Omega = \left\{ \sum_{i=1}^{m} \lambda_i a_i \;\middle|\; \lambda_i \geq 0,\ a_i \in \Omega,\ m \in \mathbb{N} \right\}.$$

When Ω is convex, $K_\Omega = \mathbb{R}_+\Omega$.

Proof. Denote by C the set on the right-hand side of the equality above. Since C is clearly a convex cone containing Ω, $C \supset K_\Omega$. It remains to prove the opposite inclusion. Consider an element of C given by $x = \sum_{i=1}^m \lambda_i a_i$, where $\lambda_i \geq 0$, $a_i \in \Omega$, and $m \in \mathbb{N}$. If $\lambda_i = 0$ for all i, then $x = 0 \in K_\Omega$. If $\lambda_i > 0$ for some i, define $\lambda := \sum_{i=1}^m \lambda_i > 0$. Thus

$$x = \lambda(\sum_{i=1}^m \frac{\lambda_i}{\lambda} a_i) \in K_\Omega.$$

Hence $C \subset K_\Omega$. We have proved that $C = K_\Omega$. □

Lemma 2.1.2. *Consider a set $\Omega \subset \mathbb{R}^n, \Omega \neq \emptyset$, and any element $x \in K_\Omega \backslash \{\emptyset\}$. The following holds:*

$$x = \sum_{i=1}^k \lambda_i a_i,\ \lambda_i > 0,\ a_i \in \Omega \textit{ and } k \leq n.$$

Proof. For an element $x \in K_\Omega \backslash \{\emptyset\}$, we have, by 2.1.1, the representation $x = \sum_{i=1}^m \mu_i a_i$, with $\mu_i > 0$, $a_i \in \Omega$ and $m \in \mathbb{N}$. If the elements $a_1, \ldots, a_m$ are linearly dependent, then there exist $\gamma_i \in \mathbb{R}$ for $i = 1, \ldots, m$, not all zeros, such that

$$\sum_{i=1}^m \gamma_i a_i = 0.$$

Define the nonempty set $I := \{i = 1, \ldots, m \mid \gamma_i > 0\}$. For any $\epsilon > 0$ we have

$$x = \sum_{i=1}^m \mu_i a_i = \sum_{i=1}^m \mu_i a_i - \epsilon(\sum_{i=1}^m \gamma_i a_i) = \sum_{i=1}^m (\mu_i - \epsilon\gamma_i) a_i.$$

Choose $\epsilon := \min\{\frac{\mu_i}{\gamma_i} \mid i \in I\} = \frac{\mu_{i_0}}{\gamma_{i_0}}$, with $i_0 \in I$, and define $\beta_i := \mu_i - \epsilon\gamma_i$ for $i = 1, \ldots, m$. We now have the following definition for an element $x \in K_\Omega \backslash \{\emptyset\}$:

$$x = \sum_{i=1}^m \beta_i a_i.$$

Then $\beta_{i_0} = 0$, and $\beta_i \geq 0$, for $i = 1, \ldots, m$. Continuing this process, we can represent x with a positive linear combination of the linearly independent elements $\{a_j \mid j \in J\}$, with $J \subset \{1, \ldots, m\}$. Linear independence in an n-dimensional space requires that $|J| \leq n$. $\square$

Theorem 2.1.3. (Carathéodory's Theorem) *For $\Omega \subset \mathbb{R}^n, \Omega \neq \emptyset$, any element $x \in \operatorname{co}\Omega$ may be expressed as a combination of, at most, n+1 elements of Ω.*

Proof. Define $B := \{1\} \times \Omega \in \mathbb{R}^{n+1}$. We observe that $\operatorname{co} B = \{1\} \times \operatorname{co}\Omega$, and that $\operatorname{co} B \in K_B$. Consider any $x \in \operatorname{co}\Omega$. Take an element $(1, x) \in \operatorname{co} B$; there exist $\lambda_i \geq 0$ and $(1, a_i) \in B$, $i = 0, \ldots, m$ with $m \leq n$ so that

$$(1, x) = \sum_{i=0}^{m} \lambda_i (1, a_i).$$

We thus conclude that $x = \sum_{i=0}^{m} \lambda_i a_i$, $\sum_{i=0}^{m} \lambda_i = 1$ with $\lambda_i \geq 0$ and $m \leq n$. $\square$

2.2 Radon Theorem

Next we shall prove a lemma, Radon's and then Helly's theorem.

Lemma 2.2.1. *For a set of vectors in $\mathbb{R}^n$, $\Upsilon = \{\upsilon_1, \ldots, \upsilon_m\}$, if $|\Upsilon| \geq n + 2$, then the vectors are affine dependent.*

Proof. Consider the set $\Upsilon' = \{\upsilon_2 - \upsilon_1, \ldots, \upsilon_m - \upsilon_1\}$. Since $|\Upsilon'| \geq n + 1$, those vectors are linearly dependent, and it follows that the vectors $\{\upsilon_1, \ldots, \upsilon_m\}$ are affine dependent. $\square$

Theorem 2.2.2. (Radon's Theorem)

Define the set $\Omega := \{\omega_1, \ldots, \omega_m\}$, with $|\Omega| \geq n + 2$, in $\mathbb{R}^n$. There exist two nonempty, disjoint subsets $\Omega_1 \subset \Omega$ and $\Omega_2 \subset \Omega$ so that

$$\Omega_1 \cup \Omega_2 = \Omega \text{ and } \operatorname{co}\Omega_1 \cap \operatorname{co}\Omega_2 \neq \emptyset.$$

Proof. Since we have that $|\Omega| = m \geq n+2$, the vectors $\{\omega_1, \ldots, \omega_m\}$ are affine dependent by 2.2.1. So there exist real numbers $\lambda_1, \ldots, \lambda_m$, some of which are positive, so that

$$\sum_{i=1}^{m} \lambda_i \omega_i = 0, \ \sum_{i=1}^{m} \lambda_i = 0.$$

We may define the index sets $I_1 := \{1, \ldots, m \mid \lambda_i \geq 0\}$ and $I_2 := \{1, \ldots, m \mid \lambda_i < 0\}$. Furthermore, I_1, $I_2 \neq \emptyset$ and $\sum_{i \in I_1} \lambda_i = -\sum_{i \in I_2} \lambda_i$. Define $\lambda := \sum_{i \in I_1} \lambda_i$, so we have

$$\sum_{i \in I_1} \lambda_i \omega_i = -\sum_{i \in I_2} \lambda_i \omega_i \text{ and } \sum_{i \in I_1} \frac{\lambda_i}{\lambda} \omega_i = -\sum_{i \in I_2} \frac{\lambda_i}{\lambda} \omega_i.$$

Define sets $\Omega_1 := \{\omega_i \mid i \in I_1\}$ and $\Omega_2 := \{\omega_i \mid i \in I_2\}$. These sets are nonempty and disjoint subsets of Ω with $\Omega_1 \cup \Omega_2 = \Omega$ and $\operatorname{co} \Omega_1 \cap \operatorname{co} \Omega_2 \neq \emptyset$. Thus we have

$$\sum_{i \in I_1} \frac{\lambda_i}{\lambda} \omega_i = -\sum_{i \in I_2} \frac{\lambda_i}{\lambda} \omega_i \in \operatorname{co} \Omega_1 \cap \operatorname{co} \Omega_2.$$

The proof is now complete. □

2.3 Helly Theorem

In this section we present Helly's theorem and its proof. The proof presented below is based on induction and Radon's theorem.

Theorem 2.3.1. (Helly's Theorem) *Consider $O := \{\Omega_1, \ldots, \Omega_m\}$, a collection of convex sets in $\mathbb{R}^n$, with $|O| \geq n+1$. If the intersection of any n+1 of these sets is nonempty, then all the sets in the collection O have a nonempty intersection; more formally $\bigcap_{i=1}^{m} \Omega_i \neq \emptyset$.*

Proof. We shall prove the theorem by induction on $|O|$. For $m = n+1$, the result of the theorem is trivial. For the induction step, suppose that the theorem holds for $|O| = m \geq n+1$. Consider a collection of convex sets in $\mathbb{R}^n$, $\{\Omega_1, \ldots, \Omega_m, \Omega_{m+1}\}$, with the property that any $n+1$ of the sets have a nonempty intersection. For $i = 1, \ldots, m+1$, define the following:

$$\Theta_i := \bigcap_{j=1, j\neq i}^{m+1} \Omega_j,$$

which has the property that $\Theta_i \subset \Omega_j$ whenever $j \neq i$ and $i, j \in \{1, \ldots, m+1\}$. By hypothesis, $\Theta_i \neq \emptyset$, so there exists an element $\omega_i \in \Theta_i$ for each $i \in \{1, \ldots, m+1\}$. Define the set $W := \{\omega_1, \ldots, \omega_{m+1}\}$. Applying Radon's theorem on the set W, we may find two nonempty, disjoint subsets $W_1 := \{\omega_i \mid i \in I_1\}$, and $W_2 := \{\omega_i \mid i \in I_2\}$. These subsets have the following properties: $W_1 \cup W_2 = W$, and $\operatorname{co} W_1 \cap \operatorname{co} W_2 \neq \emptyset$. We may select an element $w \in \operatorname{co} W_1 \cap \operatorname{co} W_2$, and show that $w \in \bigcap_{i=1}^{m+1} \Omega_i \neq \emptyset$. In order to do this, we first note that $i \neq j$ for $i \in I_1$ and $j \in I_2$, which implies that $\theta_i \subset \Omega_j$ when $i \in I_1$ and $j \in I_2$. For a particular $i \in \{1, \ldots, m+1\}$ and $i \in I_1$, $\omega_i \in \Theta_i \subset \Omega_j$ for any $j \in I_2$. Because Ω_i is convex, $\omega \in \operatorname{co} W_2 = \operatorname{co} \{\omega_j \mid j \in I_2\} \subset \Omega_i$. Thus we have that $\omega \in \Omega_i$ for any $i \in I_1$. We may, by a similar argument, show that $\omega \in \Omega_i$ for every $i \in I_2$. $\square$

Chapter 3

The Theorems of Carathéodory, Radon, and Helly: Their Equivalence

3.1 Carathéodory's and Helly's Theorem

In this section we study the equivalence of the theorems of Carathéodory and Helly based on subdifferential properties of distance functions. Let us start with two useful lemmas.

Lemma 3.1.1. *Let Ω_i, $i = 1, \ldots, m$, be nonempty closed, convex sets. If at least one of the sets Ω_i is bounded, then the maximum function*

$$f(x) := \max\{d(x;\Omega_i) \mid i = 1, \ldots, m\}, x \in \mathbb{R}^n \,, \tag{3.1.1}$$

has an absolute minimum on $\mathbb{R}^n$.

Proof. Let $\gamma := \inf\{f(x) \mid x \in \mathbb{R}^n\}$. Since $f(x) \geq 0$ for all $x \in \mathbb{R}^n$, it is obvious that γ is a real number. Let $\{x_k\}$ be a sequence in $\mathbb{R}^n$ such that $\lim_{k\to\infty} f(x_k) = \gamma$. Without loss of generality, assume that Ω_1 is bounded. By the definition of limit, one finds $k_0 \in I\!N$ such that

$$0 \le d(x_k;\Omega_1) \le f(x_k) < \gamma + 1 \text{ for all } k \ge k_0.$$

Choosing $w_k \in \Omega_1$ with $\|x_k - w_k\| = d(x_k;\Omega)$, we get $\|x_k - w_k\| \le \gamma + 1$, and hence $\|x_k\| \le \|w_k\| + 1$ for all $k \ge k_0$. Since Ω_1 is bounded, the sequence $\{x_k\}$ is bounded as well, and so we can assume that it has a subsequence $\{x_{k_\ell}\}$ that converges to $\bar{x}$. By the continuity of f,

$$\gamma = \lim_{\ell\to\infty} f(x_{k_\ell}) = f(\bar{x}) \le f(x) \text{ for all } x \in \mathbb{R}^n.$$

Therefore, f has an absolute minimum at $\bar{x}$. $\square$

Given any $u \in \mathbb{R}^n$, define the active index set at u, associated with the function f given in (3.1.1), by

$$I(u) := \{i = 1,\ldots,m \mid d(u;\Omega_i) = f(u)\}.$$

Lemma 3.1.2. *Let Ω_i, $i = 1,\ldots,m$, be nonempty closed, convex sets with $\bigcap_{i=1}^m \Omega_i = \emptyset$. Consider the function f defined in* (3.1.1)*. Then f has an absolute minimum at $\bar{x}$ if and only if*

$$\bar{x} \in \operatorname{co}\{w_i \mid i \in I(\bar{x})\},$$

where $w_i := \Pi(\bar{x};\Omega_i)$.

Proof. The assumption $\bigcap_{i=1}^m \Omega_i = \emptyset$ implies that $f(x) = \max\{d(x;\Omega_i) \mid i = 1,\ldots,m\} > 0$ for all $x \in \mathbb{R}^n$ and $x \notin \Omega_i$ for every $i \in I(x)$. It follows from Theorem 1.0.8 that the function f has an absolute minimum at $\bar{x}$ if and only if

$$0 \in \partial f(\bar{x}) = \operatorname{co}\{\partial d(\bar{x};\Omega_i) \mid i \in I(\bar{x})\} = \operatorname{co}\left\{\frac{\bar{x} - w_i}{d(\bar{x};\Omega_i)} \mid i \in I(\bar{x})\right\}.$$

Since $f(\bar{x}) = d(\bar{x};\Omega_i) > 0$ for all $i \in I(\bar{x})$, we can denote this common value by r. By

Proposition 1.0.1, there exist $\lambda_i \geq 0$ for $i \in I(\bar{x})$ with $\sum_{i\in I(\bar{x})} \lambda_i = 1$ such that

$$0 = \sum_{i\in I(\bar{x})} \lambda_i \frac{\bar{x} - w_i}{r},$$

which is equivalent to $\bar{x} \in \operatorname{co}\{w_i \mid i \in I(\bar{x})\}$. $\square$

Now we are ready to prove the main theorem of this section, namely showing that Carathéodory's theorem and Helly's theorem are equivalent in the described sense.

Theorem 3.1.3. *Consider the following statements:*

(i) *Carathéodory's theorem: For a nonempty convex set $\Omega \subset \mathbb{R}^n$ with $\bar{x} \in \operatorname{co} A$ there exist $\lambda_i \geq 0$ and $w_i \in \Omega$ for $i = 1, \ldots, n+1$ such that*

$$\bar{x} = \sum_{i=1}^{n+1} \lambda_i \omega_i \,.$$

(ii) *Helly's theorem: For any collection of nonempty closed, convex sets $\{\Omega_1, \ldots, \Omega_m\}$, $m \geq n+2$, in $\mathbb{R}^n$ with the property that the intersection of any $n+1$ sets from this collection is nonempty, one has*

$$\bigcap_{i=1}^{m} \Omega_i \neq \emptyset \,.$$

Then statement **(i)** *implies statement* **(ii)**, *and vice versa.*

Proof. Let us first prove **(ii)** assuming that **(i)** holds. Consider the case where Ω_i, for $i = 1, \ldots, m$, are nonempty closed, convex sets such that at least one of them is bounded. The assumption implies that the intersection of any collection of k sets from the whole collection, where $k \leq n+1$, is nonempty. Suppose that $\bigcap_{i=1}^{m} \Omega_i = \emptyset$. Consider the function f defined by (3.1.1) and let $\bar{x}$ be a point in $\mathbb{R}^n$ at which f has an absolute minimum. Note that this point exists by Lemma 3.1.1. From Lemma 3.1.2 one has

$$\bar{x} \in \operatorname{co}\{w_i \mid i \in I(\bar{x})\}\,,$$

where $w_i := \Pi(\bar{x}; \Omega_i)$. Applying Carathéodory's theorem, we get $J \subset I$, $|J| \leq n+1$, such that

$$\bar{x} = \sum_{j \in J} \lambda_j w_i.$$

Defining the function

$$g(x) := \max\{d(x; \Omega_j) \mid j \in J\},$$

by the given assumption, $\bigcap_{j \in J} \Omega_j \neq \emptyset$, there exists a point u in this intersection with $g(u) = 0$. Since $d(\bar{x}; \Omega_j) = r > 0$ for all $j \in J$, where $r := f(\bar{x}) = g(\bar{x}) > 0$, the active index set at $\bar{x}$ associated with the function g is J. By Lemma 3.1.2, one has that g has an absolute minimum at $\bar{x}$, and hence $0 < r = g(\bar{x}) \leq g(u) = 0$. This contradiction implies the statement **(ii)**.

In the case where we do not assume that at least one of the sets is bounded, we choose an element in each intersection of $n+1$ sets. Let $t > 0$ be sufficiently large such that the closed ball $\mathbb{B}(0; t)$ covers all such points. Then we only need to apply the previous case for the collection $\{\Theta_i \mid i = 1, \ldots, m\}$, where $\Theta_i := \Omega_i \cap \mathbb{B}(0; t)$.

Now we show how to derive **(i)** from **(ii)**. Fix any element $\bar{y} \in \operatorname{co} \Omega$. If $\bar{y} \in \Omega$, then it is obviously a convex combination of itself, so the conclusion is obvious. Suppose that $\bar{y} \notin \Omega$. We follow and simplify significantly the proof in [7], pp. 40-41. In particular, we do not assume that $\Omega \in \mathbb{R}^n$ is closed and bounded as in [7], pp. 40-41. For each point $\bar{x} \in \Omega$, denote by $\mathcal{L}_1(\bar{x})$ and $\mathcal{L}_2(\bar{x})$ the closed half-spaces with bounding hyperplane passing through $\bar{x}$ and being perpendicular to the line connecting $\bar{x}$ and $\bar{y}$, where $\mathcal{L}_2(\bar{x})$ contains $\bar{y}$, i.e.,

$$\mathcal{L}_1(\bar{x}) := \{w \in \mathbb{R}^n \mid \langle \bar{x} - \bar{y}, w - \bar{x} \rangle \geq 0\},$$

$$\mathcal{L}_2(\bar{x}) := \{w \in \mathbb{R}^n \mid \langle \bar{x} - \bar{y}, w - \bar{x} \rangle \leq 0\}.$$

Let us now show that

$$\bigcap_{\bar{x}\in\Omega} \mathcal{L}_1(\bar{x}) = \emptyset.$$

Indeed, suppose that this set is nonempty. Then there exists $\bar{z} \in \bigcap_{\bar{x}\in\Omega} \mathcal{L}_1(\bar{x})$, implying that

$$\langle x - \bar{y}, \bar{z} - x\rangle \geq 0 \text{ for all } x \in \Omega\,.$$

It follows that

$$\begin{aligned}&\langle \bar{z} - \bar{y}, \bar{y} - x\rangle = \langle \bar{z} - x + x - \bar{y}, \bar{y} - x\rangle = \langle \bar{z} - x, \bar{y} - x\rangle + \langle x - \bar{y}, \bar{y} - x\rangle\\ &= \langle \bar{z} - x, \bar{y} - x\rangle - \|x - \bar{y}\|^2 < 0\end{aligned}$$

for all $x \in \Omega$ (note that $\|x - \bar{y}\|^2 > 0$, since $\bar{y} \notin \Omega$). Since $\bar{y} \in \operatorname{co}\Omega$, there exist $x_i \in \Omega$ and $\lambda_i \geq 0$ for $i = 1, \ldots, m$ such that

$$\bar{y} = \sum_{i=1}^{m} \lambda_i x_i \text{ and } \sum_{i=1}^{m} \lambda_i = 1.$$

Then

$$0 = \langle \bar{z} - \bar{y}, \bar{y} - \bar{y}\rangle = \langle \bar{z} - \bar{y}, \bar{y} - \sum_{i=1}^{m} \lambda_i x_i\rangle = \sum_{i=1}^{m} \lambda_i \langle \bar{z} - \bar{y}, \bar{y} - x_i\rangle < 0\,.$$

This contradiction shows that

$$\bigcap_{\bar{x}\in\Omega} \mathcal{L}_1(\bar{x}) = \emptyset.$$

Since $\mathcal{L}_1(x)$ is a nonempty closed, convex set for every x, by **(ii)** there exist $x_1, \ldots, x_{n+1} \in \Omega$ such that

$$\bigcap_{i=1}^{n+1} \mathcal{L}_1(x_i) = \emptyset.$$

However, this implies $\bar{y} \in \operatorname{co}\{x_1, \ldots, x_{n+1}\}$. Let us prove this by contradiction. Assuming that this is not the case, there exist $a, b \in \mathbb{R}^n$ such that

$$\langle a, \bar{y}\rangle > b \text{ and } \langle a, x_i\rangle \leq b \text{ for all } i = 1, \ldots, n + 1.$$

Define

$$\ell := \{\bar{y} - ta \mid t \geq 0\}.$$

Then

$$\langle x_i - \bar{y}, \bar{y} - ta - x_i \rangle = \langle x_i - \bar{y}, \bar{y} - x_i \rangle - t\langle a, x_i - \bar{y} \rangle \,.$$

Thus we can find a value t_0 such that this expression is positive for all $t > t_0$ and for all i, but this implies $\bigcap_{i=1}^{n+1} \mathcal{L}_1(x_i) \neq \emptyset$. This contradiction shows that $\bar{y} \in \operatorname{co}\{x_1, \ldots, x_{n+1}\}$. $\square$

3.2 Carathéodory's and Radon's Theorem

In this section we study the equivalence of the theorems of Carathéodory and Radon, thus establishing the equivalence of all three theorems.

Theorem 3.2.1. *Consider the following statements:*

(i) *Carathéodory's theorem: For a nonempty convex set $A \subset \mathbb{R}^n$, with $\bar{x} \in \operatorname{co} A$, there exist $\lambda_i \geq 0$ with $\sum_{i=1}^{n+1} \lambda_i = 1$ and $w_i \in A$ for $i = 1, \ldots, n+1$ such that*

$$\bar{x} = \sum_{i=1}^{n+1} \lambda_i w_i \,.$$

(ii) *Radon's theorem: Given a set $\Omega := \{\omega_1, \ldots, \omega_m\}$, $m \geq n+2$, in $\mathbb{R}^n$, there exist two nonempty, disjoint subsets $\Omega_1 \subset \Omega$ and $\Omega_2 \subset \Omega$ such that*

$$\Omega_1 \cup \Omega_2 = \Omega \text{ and } \operatorname{co}\Omega_1 \cap \operatorname{co}\Omega_2 \neq \emptyset \,.$$

The statements **(i)** *and* **(ii)** *are equivalent.*

Proof. (i)$\Longrightarrow$ (ii): Consider the set $\Omega \subset \mathbb{R}^n$ defined by $\Omega = \{w_1, \ldots, w_m\}$ with $m \geq n+2$. Then

$$\frac{w_1 + \cdots + w_m}{m} \in \operatorname{co}\{w_1, \ldots, w_m\}\,.$$

By Carathéodory's theorem, we have the representation

$$\frac{w_1 + \cdots + w_m}{m} = \sum_{i=1}^{n+1} \lambda_i w_i,$$

with $\lambda_i \geq 0$ and $\sum_{i=1}^{n+1} \lambda_i = 1$. It follows that

$$\sum_{i=1}^{m} \beta_i w_i = \sum_{i=1}^{m} \lambda_i w_i\,,$$

where $\beta_i = \frac{1}{m}, i = 1, \ldots, m$, $\lambda_i \geq 0, \sum_{i=1}^{n+1} \lambda_i = 1$, and $\lambda_i = 0$ for all $i = n+2, \ldots, m$. So

$$\sum_{i=1}^{m} (\beta_i - \lambda_i) w_i = 0\,.$$

We rewrite the above as

$$\sum_{i=1}^{m} \gamma_i w_i = 0\,,$$

with $\gamma_i = \beta_i - \lambda_i$, $\sum_{i=1}^{m} \gamma_i = 0$, where not all $\gamma_i's$ are zeros. Then

$$\sum_{i\in I} \gamma_i x_i + \sum_{j \in J} \gamma_j x_j = 0\,,$$

where $I := \{i \in \{1, \ldots, m\} \mid \gamma_i > 0\}$ and $J := \{i \in \{1, \ldots, m\} \mid \gamma_i \leq 0\}$. Observe that I and J are both nonempty. Define $\gamma := \sum_{i\in I} \gamma_i$, $\Omega_1 := \{w_i \mid i \in I\}$, and $\Omega_2 := \{w_j \mid j \in J\}$. Then $\gamma = -\sum_{j\in J} \gamma_j$ and

$$\frac{1}{\gamma} \sum_{i\in I} \gamma_i w_i = -\frac{1}{\gamma} \sum_{j\in J} \gamma_j w_j \in \operatorname{co}\Omega_1 \cap \operatorname{co}\Omega_2.$$

At this point, we can see that Ω_1 and Ω_2 satisfy the requirements of Radon's theorem.

To prove that Carathéodory's theorem follows from Radon's theorem, we complete the approach to [8, Theorem 3.2].

(ii) $\Longrightarrow$ (i): Fix any $\bar{x} \in \operatorname{co} A$. Then there exist $\lambda_i > 0$ for $i = 1, \ldots, m$, with $\sum_{i=1}^{m} \lambda_i = 1$ and

$$\bar{x} = \sum_{i=1}^{m} \lambda_i w_i \,,$$

where $w_i \in A$ for all $i = 1, \ldots, m$ and the representation is chosen such that m is minimal. We will show that $m \leq n+1$. Assume the contrary, i.e., that $m > n+1$. By Radon's theorem applied to the set $\Omega := \{w_1, \ldots, w_m\}$ we can assume that, without loss of generality,

$$\sum_{i=1}^{k} \gamma_i w_i = \sum_{i=k+1}^{m} \gamma_i w_i \,,$$

where $\gamma_i \geq 0$ for $i = 1, \ldots, m$ and $\sum_{i=1}^{k} \gamma_i = \sum_{i=k+1}^{m} \gamma_i = 1$, $1 \leq k < m$. Then we have the representation

$$\sum_{i=1}^{m} \beta_i w_i = 0 \,,$$

where $\beta_i := \gamma_i$ for $i = 1, \ldots, k$, and $\beta_i = -\gamma_i$ for $i = k+1, \ldots, m$. Observe that $\sum_{i=1}^{m} \beta_i = 0$. Choose an index $i_0 \in \{1, \ldots, k\}$ such that

$$\epsilon := \frac{\lambda_{i_0}}{\beta_{i_0}} = \min\{\frac{\lambda_i}{\beta_i} \mid i = 1, \ldots, k \text{with } \beta_i > 0\} \,.$$

Define $\alpha_i := \lambda_i - \epsilon\beta_i$ for $i = 1, \ldots, m$. Then $\alpha_i \geq 0$ for $i = 1, \ldots, m$, $\alpha_{i_0} = 0$, $\sum_{i=1}^{m} \alpha_i = 1$, and

$$\bar{x} = \sum_{i=1}^{m} \alpha_i w_i \,.$$

This contradicts the minimal property of m. The proof is now complete. $\square$

Remark 3.2.2. **(i)** Note that the closedness property assumed in Radon's theorem presented in Theorem 3.1.3(ii) can be relaxed, i.e., the statements in Theorem 3.1.3 and Theorem 3.2.1 are equivalent to the following statement:

For any collection of nonempty convex sets $\{\Omega_1, \ldots, \Omega_m\}$, $m \geq n+2$, *in* $\mathbb{R}^n$ *with the property that the intersection of any* $n+1$ *sets from this collection is nonempty, one has*

$$\bigcap_{i=1}^{m} \Omega_i \neq \emptyset\,.$$

Indeed, this statement obviously implies Theorem 3.1.3(ii), and it follows from Theorem 3.2.1(ii); see, e.g., [10, Theorem 3.13].

(ii) As a next step, it would be natural to extend the present investigations to *generalized convexity notions*, like for example H-convexity (see, e.g., [1] and [2]) or d-convexity (cf. [3, Chapter II]).

Summarizing chapters 3 and 4,we have demonstrated the equivalence of Carathéodory's, Radon's and Helly's theorem in convex geometry by demonstrating the following implications:

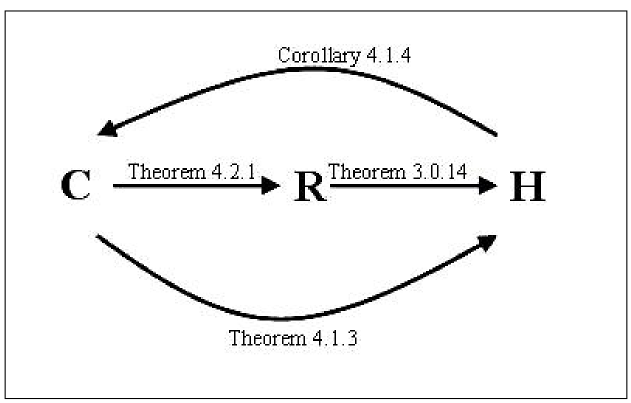

FIGURE 3.2.1: C=Carathéodory's theorem R=Radon's Theorem H=Helly's Theorem
Diagram summarizing the equavalence of the three main theorems.

Bibliography

[1] V. Boltyanski, H. Martini: Carathéodory's theorem and H-convexity. *J. Combin. Theory*, Ser. A **93** (2001), 292-309.

[2] V. Boltyanski, H. Martini: Minkowski addition of H-convex sets and related Helly-type theorems. *J. Combin. Theory*, Ser. A **103** (2003), 323-336.

[3] V. Boltyanski, H. Martini, and P. S. Soltan: *Excursions into Combinatorial Geometry*. Springer, Berlin and Heidelberg, 1997.

[4] S. Boyd, L. Vandenberghe: *Convex Optimization*. Cambridge University Press, Cambridge, England, 2004.

[5] L. Danzer, B. Grünbaum, and V. Klee: Helly's theorem and its relatives. Proceedings of Symposia in Pure Mathematics, Vol. 7 *Convexity*, American Mathematical Society, 1963.

[6] J. Eckhoff: Helly, Radon and Carathéodory type theorems. *Handbook of Convex Geometry*, Vol. 1, P.M. Gruber and J.M. Wills Editors. North-Holland, 1993, pp.

[7] H. G. Eggleston: Convexity. Cambridge Tracts in Mathematics and Mathematical Physics, No. 47, Cambridge University Press, 1958.

[8] P. M. Gruber: *Convex and Discrete Geometry*. Springer-Verlag, 2007.

[9] H. Martini, N.M. Nam, and A. Robinson: On the equivalence of the theorems of Helly, Radon, and Carathéodory via convex analysis, preprint, 2014.

[10] B. Mordukhovich, N. M. Nam: *An Easy Path to Convex Analysis and Applications*, Morgan and Claypool Publishers, 2014.

[11] R. T. Rockafellar: *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.