Extremality and Stationarity of Collections of Sets:

Metric, Slope and Normal Cone Characterisations

Hoa T. Bui
Supervisors: Professor Alex Kruger, Associate Professor David Yost

Submitted to the requirement of a
Doctor of Philosophy

School of Science, Engineering and Information Technology
Federation University Australia

Mount Helen, Victoria, Australia

October 2019

Abstract

Variational analysis, a relatively new area of research in mathematics, has become one of the most powerful tools in nonsmooth optimisation and neighbouring areas. The extremal principle, a tool to substitute the conventional separation theorem in the general nonconvex environment, is a fundamental result in variational analysis. There have seen many attempts to generalise the conventional extremal principle in order to tackle certain optimisation models.

Models involving collections of sets, initiated by the extremal principle, have proved their usefulness in analysis and optimisation, with non-intersection properties (or their absence) being at the core of many applications: recall the ubiquitous convex separation theorem, extremal principle, Dubovitskii Milyutin formalism and various transversality/regularity properties. We study elementary nonintersection properties of collections of sets, making the core of the conventional definitions of extremality and stationarity. In the setting of general Banach/Asplund spaces, we establish nonlinear primal (slope) and linear/nonlinear dual (generalised separation) characterisations of these non-intersection properties. We establish a series of consequences of our main results covering all known formulations of extremality/stationarity and generalised separability properties. This research develops a universal theory, unifying all the current extensions of the extremal principle, providing new results and better understanding for the exquisite theory of variational analysis.

This new study also results in direct solutions for many open questions and new future research directions in the fields of variational analysis and optimisation. Some new nonlinear characterisations of the conventional extremality/stationarity properties are obtained. For the first time, the intrinsic transversality property is characterised in primal space without involving normal cones. This characterisation brings a new perspective on intrinsic transversality. In the process, we thoroughly expose and classify all quantitative geometric and metric characterisations of transversality properties of collections of sets and regularity properties of set-valued mappings.

Statement of Authorship

This is to certify that:

(i) the thesis comprises only my original work towards the PhD except where indicated,

(ii) due acknowledgement has been made in the text to all other material used,

(iii) the thesis is less than $100,000$ words in length, exclusive of tables, maps, bibliographies and appendices.

I have stated clear and fully in this thesis the extent of any collaboration with others. To the best of my knowledge and belief, this thesis contains no material previously published by any other person except where due acknowledgement has been made. The thesis submitted contains no material which has been accepted for an award of any other degree or diploma at any university. The thesis has identified work of others that has been relied upon by providing appropriate acknowledgement, citation, and reference in the text and in the bibliography.

Signed:

Name: Hoa T. Bui

Date: 18/10/2019

Acknowledgements

To my supervisors Professor Alexander Kruger and A/Professor David Yost, I wish to express my deepest gratitude. I must thank you first and foremost for all your encouragement, guidance and for sharing your vast knowledge. You are my greatest supervisors, mentors and friends in Australia. To Alex and Valentine, thank you very much for your sincere care of my health and well-being.

I am also grateful to my other co-authors: Dr. Nguyen Hieu Thao, Prof. Michel Verhaegen, MSc Nguyen Duy Cuong, Dr. Vera Roshchina, Dr. Scott Lindstrom, Dr. Pham Duy Khanh, Tran Thi Tu Trinh, A/Prof Regina Burachik, Dr. Julien Ugon, Dr. Guillermo Pineda-villavicencio. Thank you for your hard work, brilliance, and dedication. Special thanks to Guillermo for organising the reading group, and patiently teaching me about polytopes.

It is a great privilege to meet many great people in the variational analysis and optimisation community, especially the Australian optimisation group. Meeting with these wonderful people inspires me to my research. Special thanks to Prof. Andrew Eberhard, A/Prof. Adil Barigov, A/Prof. Regina Burachik, Dr. Yalcin Kaya, Dr. Vera Roshchina, Dr. Nadezda Sukhorukova, Dr. Julien Ugon, Dr. Minh Dao, Prof. Christiane Tammer, Prof. Russell Luke, Dr. Matthew Tam, and Dr. Pham Duy Khanh. Thank you very much for the opportunities and your support.

I would like to thank two referees for their comments which help to improve significantly the presentation of this thesis.

I would like to thank the people within the School of Science, Engineering and Technology at Federation University, Helen Wade and for her administrative support. I also acknowledge fellow PhD students Charlotte, Cuong, Samuel, Saleem, and Leonard for their support and friendship. Thanks Eunsil, Dada, Brianne, Jae and HyunJae for your special friendship, companionship and for taking good care of me. Thanks Sinh, Mai, Linh, Thao and Dat, my dearest friends, for your understanding and support regardless of many difficult circumstances.

Lastly, I am thankful for the solely love, support, and encouragement of my family.

Hoa T. Bui was supported by an Australian Government Research Training Program (RTP) Stipend and RTP Fee-Offset Scholarship through Federation University Australia. Further support through funding from the Australian Research Council (ARC) - project DP160100854 is also acknowledged.

Publications

This thesis comprises the following manuscripts which are either published, accepted, submitted or in preparation for publication in internationally peer reviewed journals. In every such case, I have been a significant and integral contributor involved in both the work undertaken and the preparation of the manuscript.

Part I:

1. Hoa T. Bui, Alexander Y. Kruger, *About Extensions of the Extremal Principle*, Vietnam Journal of Mathematics **46** (2018) no. 2, 215–242
2. Hoa T. Bui, Alexander Y. Kruger, *Extremality, Stationarity and Generalized Separation of Collections of Sets*, Journal of Optimization Theory and Applications, **182** (2019) no. 1, 211-264
3. Nguyen Hieu Thao, Hoa T. Bui, Nguyen Duy Cuong and Michel Verhaegen, *Some new Characterizations of Intrinsic Transversality in Hilbert Spaces*, Set-Valued and Variational Analysis Theory and Applications, **36** (February 2020), DOI:10.1007/s11228-020-00531-7
4. Hoa T. Bui, Nguyen Duy Cuong, Alexander Y. Kruger, *Transversality of Collections of Sets: Metric Characterizations* Optimization Online: 2019-07-7275, 2019 (accepted to Vietnam Journal of Mathematics)
5. Hoa T. Bui, Alexander Y. Kruger, *Nonlinear Characterizations of Extremality and Stationarity Properties*, arxiv:1909.08995 (submitted to Optimization, A Journal of Mathematical Programming and Operations Research)

Part II:

6. Hoa T. Bui, Pham Duy Khanh, T.T. Trinh Tran, *Characterizations of Nonsmooth Robustly Quasiconvex Functions*, Journal of optimization Theory and Applications, **180** (2019) no. 3, 775-786
7. Hoa T. Bui, Regina S. Burachik, Alexander Y. Kruger, David T. Yost, *Characterizations of Zero Duality Gap via Abstract Convexity*, arxiv:1910.08156

Part III:

8. Hoa T. Bui, Guillermo Pineda-Villavicencio, Julien Ugon, *Connectivity of cubical polytopes*, Journal of Combinatorial Theory, Series A, **169** (January 2020), 105–126
9. Hoa T. Bui, Guillermo Pineda-Villavicencio, Julien Ugon, *On the Linkedness of cubical polytopes*, arXiv:1802.09230 (submitted to Israel Journal of Mathematics)

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Goals of Research

The first goal of our research is to provide a universal tool that covers all the existing extensions of the *extremal principle*. In order to do this, we study geometric non-intersection properties of finite collections of sets, which are the building blocks of the conventional extremality/stationarity properties. This study develops a new efficient tool to study extremality, stationarity and transversality/regularity properties of collections of sets.

The second goal is to expand further the applications of the techniques of variational analysis to optimisation and neighbouring areas. Particularly, in our research, the techniques of variational analysis are employed to study generalised convexity properties. We consider two different properties of generalised convexity: robust quasiconvexity and abstract convexity. They represent two independent generalisations of convexity. The former property aims to preserve certain important optimisation properties of convex functions such as: all the lower level sets are convex; local minimum is global minimum; and stationary points are global minimisers. The latter reflects one of the fundamental concepts of convex analysis, which is every lower semicontinuous convex function f is the upper envelope of affine functions.

The third goal of my study is to broaden the border of my research to other branches of mathematics. I study polytopes, the fundamental objects in optimisation, via their graphs and face lattices. This research potentially creates connections between many different fields in optimisation.

1.2 Motivation, Literature Review and Contributions

1.2.1 Arrangement of Collections of Sets

Models involving pairs, or more generally finite (and even infinite) collections of sets, are pretty common in various fields of mathematics, especially variational analysis, optimisation and neighbouring areas. For instance, the classical *separation theorem* [130,141,168] for convex sets, one of the key results of functional and nonlinear analysis, is instrumental when establishing multiplier rules or subdifferential calculus rules; the ubiquitous *feasibility problems* [100] cover solving systems of (generalised) equations, while appropriate *transversality* (regular intersection) conditions produce *constraint qualifications* and convergence estimates for *alternating projections* (*von Neumann method*) [16, 52, 107, 128], or more generally *cyclic projections*, which in turn provide a convenient model for convergence analysis of computational algorithms. On the other hand, (local) *optimality* in optimisation problems can naturally be interpreted as ‘irregular’ (extremal) intersection of certain sets, and sufficient transversality conditions immediately produce necessary optimality conditions.

In this major part of the thesis, we study a unified theory to thoroughly study metric, slope and normal cone characterisations of certain types of arrangement of collections of sets.

Extremality, Stationarity Properties and Extremal Principle

In the framework of convex analysis, the property of two sets that one set does not meet the interior of the other set (while both are nonempty), which is the key assumption of the conventional separation theorem, provides an example of the absence of regular intersection. Thus, the separation theorem can be interpreted as a characterisation of the absence of regular intersection of convex sets. On the other hand, the opposite property, i.e., when the intersection of one set with the interior of the other set is nonempty, is a typical qualification condition in various convex analysis statements [168].

There have been many successful applications of convex analysis, and particularly the separation theorem in the nonconvex settings by considering appropriate local convex approximations of sets and functions. The most prominent example in the optimisation area is probably given by the *Dubovitskii–Milyutin formalism* [54]. The Clarke tangent cone (2.3) is an important example of a local convex approximation of a set. However, in many situations it is impossible to construct satisfactory local convex approximations of nonconvex sets. For instance, if a set consists of two intersecting lines on the plain, it is easy to check that its Clarke tangent cone at the point of intersection is trivial (contains only the zero vector), and provides no meaningful information about the set. The powerful tools of convex analysis generally fail outside of the comfortable convex setting.

A new tool with the potential to substitute the conventional separation theorem in the general nonconvex environment — currently known as the *Extremal principle* — was suggested in 1979–1980 in [101–103]. It was observed that the separation theorem actually characterises a kind of extremal arrangement of sets. Building on this observation, the key assumption of the conventional separation theorem, that one set does not meet the interior of the other set, was replaced by a more general geometric (local) *extremality* (Definition 2.3.1(i) and (ii)). This property assumes neither convexity of the sets nor that one of the sets has nonempty interior, while still embracing many conventional (and generalised) optimality notions. Moreover, it is applicable to any finite ($n \geq 2$) collections of sets (cf. [101–103]) and, if the space is Banach and the sets are closed, the function measuring the distance between points in different translated sets satisfies the assumptions of the Ekeland variational principle. Employing an appropriate sum rule to the perturbation of the distance function arising from the application of the Ekeland variational principle, one can formulate dual characterisations of extremality in terms of appropriate normals to individual sets. Such dual conditions can be interpreted as a kind of generalised separation; cf. conditions (ii) and (iii) in Theorem 2.7.3.

Since its inception 40 years ago, the extremal principle has indeed impacted strongly on the nonconvex analysis and optimisation substituting the conventional separation theorem. Several examples of its application to proving necessary optimality conditions (multiplier rules) can be found already in the very first publications [102, 103]. Numerous applications of the extremal principle in multiple publications on optimisation and analysis are exposed and commented on in the monograph [120]. The original publications have been followed by several studies of the concept of extremality of collections of sets and the extremal principle, resulting in its additional characterisations and several extensions.

It was established in [88, 89] that, similar to the classical Lagrange multiplier rule, the conclusion of the extremal principle (the generalised separation), being a dual necessary condition of extremality, actually characterises a property which is weaker than local extremality. This property can be interpreted as a kind of stationarity of a collection of sets. The explicit primal space definition of this property was introduced (cf. Definition 2.3.1(iv)), and it was shown (by refining slightly the original proof of the extremal principle) that for this property the conclusion of the extremal principle is not only necessary (in the Asplund space setting), but also sufficient, thus producing the ultimate conventional version of the extremal principle, known as the *extended extremal principle*; cf. Theorem 2.7.3. This stationarity property called *approximate stationarity* [92] happens to be strongly connected to (in fact the negation of) another important geometric property of a collection of sets called *transversality* (also known under other names; cf. Definition 2.5.1), having its roots in the classical differential geometry and instrumental for the convergence analysis of alternating projections; see a discussion of the role of the latter property in [100], while a table illustrating the evolution of the terminology can be found in [99, Section 2]. The connections of the approximate stationarity and (extended) extremal principle can be traced further to the fundamental for variational analysis *metric regularity* property of set-valued mappings and the corresponding coderivative criterion; see [92].

The powerful (extended) extremal principle, despite its recognised universality and wide applicability, has its limitations. The authors in [97] encountered a problem while attempting to extend the extremal principle to infinite collections of sets. The initial idea to consider families of finite subcollections and apply the conventional extremal principle to each of them failed, as well as the belief of the authors in [97] in the unlimited universality of the conventional extremal principle. Uniform estimates were required, holding for all finite subcollections, and the conventional extremal principle was unable to provide such estimates. The solution was found within the conventional (extended) extremal principle, more precisely, in its proof, providing another piece of evidence for the well-known fact that important mathematical results are often deeper than their statements.

The conventional extended extremal principle asserts the equivalence of two properties of a collection of sets: one in the primal space (approximate stationarity) and another one in the dual space (generalised separability); see Theorem 2.7.3, which for completeness gives two (equivalent) versions of the latter property. Both (in the case of Theorem 2.7.3 all three) properties are formulated “for any ε there exist ...”, and as such, the ε's and the other parameters in different properties seem completely independent. However, this cannot be true. To prove that one property of this kind implies another one, there must be a way, given any ε in the second property, to construct another one to be substituted into the first property to produce the needed estimates, i.e., the parameters must be related. The relationship between the parameters in the two parts of the conventional (extended) extremal principle is exactly what was required to establish the uniform estimates needed in [97], and it indeed could be found in the proof of the conventional result. This observation made the authors of [97] carve out the core part of (the proof of) the conventional (extended) extremal principle with ε and other parameters fixed (a kind of ε-extremality), and formulate it as a separate statement in [97, Theorem 3.1] (see Proposition 3.3.14(iii) below), exposing the relationship between the parameters hidden in the proof of the conventional statement. As a result, both the conventional statement and its extension to infinite collections of sets are corollaries of [97, Theorem 3.1].

There are many equivalent formulations of the primal space approximate stationarity property (as well as other extremality and stationarity properties) and the dual space generalised separability properties, all formulated in the form “for any ε there exist ...”; see Section 2.3. The relationships between the parameters involved in these formulations can also be established and are important for identifying and analysing the core arguments in the conventional proofs of metric and dual characterisations of the extremality and stationarity and their extensions. This analysis is performed in Sections 2.3 and 3.3.

Another successful ‘surgical operation’ on the proof of the conventional extremal principle was done earlier by Zheng and Ng in [169], producing an ε-separability characterisation of another kind of ε-extremality property with the explicit relationship between the ε's in both properties. Unlike the conventional extremal principle and its extension in [97, Theorem 3.1] which assume the sets to have a common point, [169, Lemmas 2.2 and 2.2'] assume, on the contrary, that the intersection of the sets is empty. At the same time, their proofs follow the original ideas from [102, 103], utilizing the Ekeland variational principle and either the Asplund space fuzzy or the general Banach space Clarke–Rockafellar subdifferential sum rule; cf. Lemma 2.2.1 below. The statements have been further polished and analysed in a sequence of subsequent papers [112, 113, 170, 172]. In particular, it was proved by Guoyin Li et al in [113, Theorem 3.1] that the conclusion of [169, Lemma 2.2] is actually equivalent to the Ekeland variational principle, and as such implies (the Banach space with Clarke normal cones version of) the extremal principle. This fact confirms the need to move from the conventional "for any ε ..." extremality statements to more subtle ones with ε fixed. The most advanced version of the Zheng and Ng lemma was given in [172, Theorems 3.1 and 3.4] (*unified separation theorems*; see Theorem 4.2.7 below), where, besides other improvements, an additional condition was added to the concluding part, relating the dual vectors involved in the ε-separability characterisation with certain primal space vectors involved in the original ε-extremality property.

Refining again the original proof of the conventional extremal principle, a systematic study of the non-intersection properties involved in all four parts of Definition 2.3.1 was conducted recently in our research (cf. [29, 30]), producing a series of elementary generalised separation statements, clarifying the relationships between them and, particularly, unifying the statements from [170, 172, 174] and [97].

Our research follows this scheme of research.

Contributions of Research

1- In our research, we expose, analyse and refine primal and dual characterisations of approximate stationarity and transversality of collections of sets, leading to a unifying theory, encompassing all existing approaches to obtaining 'extremal' statements. For that, we examine and clarify quantitative relationships between the parameters involved in the respective definitions and statements (cf. Chapters 3& 4).

2- We formulate in Theorem 3.3.1 (and prove) a new general ε-extremality statement which exposes the core arguments and the role of the parameters in the conventional proofs of dual characterisations of the extremality and stationarity and, in particular, implies [97, Theorem 3.1] and [172, Theorems 3.1 and 3.4]. We also establish in Section 3.3 a series of other consequences of Theorem 3.3.1, covering primal and dual space conditions involved in (hopefully) all known formulations of extremality/stationarity and generalised separability properties (cf. Chapter 3).

3- In the setting of general Banach/Asplund spaces, we establish nonlinear primal (slope) and dual (generalised separation) characterisations of these non-intersection properties (cf. Chapter 4).

4. Some new (even in the linear setting) characterisations of the conventional extremality and stationarity properties are obtained. Realisations of the obtained characterisations in the Hölder setting are formulated (cf. Chapter 5).

5- Besides, we clarify the relations between various quantitative geometric and metric characterisations of the transversality properties of collections of sets and the corresponding regularity properties of set-valued mappings (cf. Chapter 2). We expose all the parameters involved in the definitions and characterisations and establish relationships between them. This allows us to classify all quantitative geometric and metric characterisations of transversality and regularity, and subdivide them into two groups with complete exact equivalencies between the parameters within each group and clear relations between the values of the parameters in different groups (cf. Chapter 2).

Intrinsic Transversality

Transversality and subtransversality are two important properties of collections of sets which reflect the mutual arrangement of the sets around the reference point in normed vector spaces. These properties are widely known as *constraint qualification conditions* in optimisation and variational analysis for formulating optimality conditions [75, 120, 125, 167] and calculus rules for subdifferentials, normal cones and coderivatives [70, 72, 73, 75, 97, 98, 120, 125], and as *key ingredients* for establishing sufficient and/or necessary conditions for linear convergence of computational algorithms [13, 53, 65, 100, 109, 110, 114, 135, 159]. A variety of their sufficient and/or necessary conditions in both primal and dual spaces are well-studied in [90–92, 97–100, 104, 106] by Kruger and his collaborators.

Transversality is strictly stronger than subtransversality. It is sufficient for many applications where the latter is not, for example, in proving linear convergence of the alternating projection method for solving nonconvex feasibility problems [109,110], or in establishing error bounds for the Douglas-Rachford algorithm [65, 135] and its modified variants [159]. However, transversality is too restrictive for many applications, and there have been a number of attempts to identify weaker properties, still sufficient for such applications. Of course, one cannot expect a universal transversality-type property that works well for all applications. When formulating necessary optimality conditions for optimisation problems in terms of abstract Lagrange multipliers and establishing intersection rules for tangent cones in Banach spaces, Bivas et al. [20] recently introduced a property called *tangential transversality*, which is a primal space property lying between transversality and subtransversality, but there is no evidence that this new property is actually different from both.

When establishing linear convergence criteria of the alternating projection algorithm for solving nonconvex feasibility problems, a series of meaningful transversality-type properties have been introduced and analysed in the literature: *affine-hull transversality* [135], *inherent transversality* [18], *separable intersection property* [128] and *intrinsic transversality* [52]. In contrast to *tangential transversality*, the latter ones are dual space properties since they are defined in terms of normal vectors. Unlike the transversality property, all the above transversality-type properties are independent on the underlying space, that is, if a property is satisfied in an ambient space X, then so is it in any ambient space containing X. Recall that in Euclidean spaces, a pair of two closed sets $\{A, B\}$ is transversal at a common point $\bar{x}$ if and only if

$$\overline{N}_A(\bar{x}) \cap \left(-\overline{N}_B(\bar{x})\right) = \{0\}, \tag{1.1}$$

where $\overline{N}_A(\bar{x})$ stands for the limiting normal cone to A at $\bar{x}$. This characterisation reveals that transversality is a property that involves all the limiting normals to the sets at the reference point. This fundamentally explains why the property is not invariant with respect to the ambient space and becomes too restrictive for many applications. Indeed, the hidden idea leading to the introduction of the above dual space transversality-type properties in the context of nonconvex alternating projections is simply based on the observation that not all such normal vectors are relevant for analysing convergence of the algorithm. The *affine-hull transversality* is merely transversality but considered only in the affine hull L of the two sets, that is, the pair of translated sets $\{A-\bar{x}, B-\bar{x}\}$ is transversal at 0 in the subspace $L-\bar{x}$. As a consequence, the analysis of this property is straightforwardly obtained from that of transversality [135]. The key feature of the *inherent transversality*[1] [18] is the use of *restricted normal cones* in place of the conventional limiting normal cones in characterisation (1.1) of transversality in Euclidean spaces. As a result, the analysis of this property is reduced to the calculus of the restricted normal vectors as conducted in [18]. The *separable intersection property* [128, Definition 1] was motivated by nonconvex alternating projections and ultimately designed for this algorithm. The *intrinsic transversality* was also introduced in the context of nonconvex alternating projections in Euclidean spaces [52], it turns out to be an important property itself in variational analysis as demonstrated by Ioffe [75, Section 9.2] and [74] and Kruger [96]. On the one hand, a variety of characterisations of intrinsic transversality in various settings (Euclidean, Hilbert, Asplund, Banach and normed linear spaces) have been established by a number of researchers [52, 74, 75, 96, 100, 107, 128]. On the other hand, there are still a number of important open questions about this property, for example, the ones raised by Kruger [96, page 140] or the challenge by Ioffe about primal counterparts of intrinsic transversality [74, page 358]. It is known that for pairs of closed and convex sets, subtransversality admits an equivalent characterisation in terms of normal vectors, and the latter is equivalent to intrinsic transversality in the Euclidean space setting [96]. Another interesting question is whether this equivalence is also valid in the nonconvex setting.

Contributions of Research

Motivated by a number of questions concerning transversality-type properties of pairs of sets recently raised by Ioffe [74] and Kruger [96], this work reports several new characterisations of the intrinsic transversality property in Hilbert spaces. New results in terms of normal vectors clarify the picture of

[1] The property originated in [18, Theorem 2.13] without a name, then was refined and termed as *inherent transversality* in Definition 4.4 of the preprint "Drusvyatskiy, D., Ioffe, A.D., Lewis, A.S.: Alternating projections and coupling slope. Preprint, arXiv:1401.7569, 1–17 (2014)".

intrinsic transversality, its variants and sufficient conditions for subtransversality, and unify several of them. For the first time, intrinsic transversality is characterised by an equivalent condition which does not involve normal vectors. This characterisation offers another perspective on intrinsic transversality. As a consequence, the obtained results allow us to answer a number of open questions about transversality-type properties (cf. Chapter 6).

1.2.2 Generalised Convexity

The functions that are, close to, but not precisely, convex often possess some important properties of convex functions. The idea of 'almost' convex functions, namely generalised convex functions, leads to various generalisations of the concept of convexity. In fact, many real world models are more easily formulated using mathematical models that involve generalised convex functions than using models with conventional convex functions. The topic of generalised convex functions is relatively new, but since getting its name, has received tremendous interests due to its wide range of applications in mechanics, economics, engineering, finance etc. Generalised convexity is now considered as an independent field of mathematics.

Generalised convexity is strongly related to variational analysis. They complement one another. The technique of variational analysis plays a key role in the development of generalised convexity. In turn, generalised convexity enriches the theory of variational analysis with numerous applications in nonconvex optimisation, variational inequalities, and equilibrium problems.

Robustly Quasiconvex Functions

One way to generalise the notion of convexity is to drop the conventional convexity assumption, but still preserve some desired properties. Here, we consider three prominent optimisation properties of convex functions: (1) all the lower level sets are convex; (2) local minimum is global minimum; (3) stationary points are global minimisers. Each of the properties reflects different classes of generalised convex functions: quasiconvex functions, explicitly quasiconvex functions, pseudo convex functions respectively (cf. [64, 67, 115, 152]). However, unlike their forerunner, convexity, these three generalised convexity properties are not at all stable under linear perturbations; counterexamples can be found in [67]. The stability of these properties is significant of practical and computational aspects. The concept of a more stable property was introduced in [10, 67], namely "robustly quasiconvexity".

Robustly quasiconvex functions were first defined in [67] under the name "s-quasiconvex" or "stable quasiconvex", and then renamed "robustly quasiconvex" in [10]. This class of functions holds a notable role, as many important optimisation properties of generalised convex functions are stable when perturbed by a linear functional with a sufficiently small norm (for instance, all lower level sets are convex, each minimum is global minimum, each stationary point is a global minimiser). For interested readers, we refer to [67] again, and further related works [10, 134].

The question of characterising convexity and generalised convexity properties in terms of subdifferentials receives tremendous attention in optimisation theory and variational analysis. For decades, there have been many significant contributions devoted to this question such as [37, 38, 120, 137, 141] for convex functions, [3–5, 11, 47, 131] for quasiconvex functions and [11] for robustly quasiconvex functions.

Significant contributions concerning dual criteria for quasiconvex functions are in [3, 4]. These characterisations are applicable for a wide range of subdifferentials, for instance Rockafellar-Clarke subdifferentials in Banach spaces, and Fréchet subdifferentials in Asplund spaces.

A zero and first order characterisations of robust convexity were given in [11, Proposition 5.3] for finite dimensional spaces. We remark that there is an oversight in the proof given there; although the function f is only assumed to be lower semicontinuous, the existence of z in the second paragraph actually requires continuity.

Contributions of Research

Two criteria for the robust quasiconvexity of lower semicontinuous functions are established in terms of Fréchet subdifferentials in Asplund spaces. The first criterion extends to such spaces a result established in [11]. The second criterion is totally new even if it is applied to lower semicontinuous functions on finite-dimensional spaces (cf. Chapter 7).

Abstract Convexity

The theory of *abstract convexity*, also called *convexity without linearity*, is a powerful tool that allows to extend many facts from classical convex analysis to more general frameworks. It has been the focus of active research for the last fifty years because of its many applications in functional analysis, approximation theory, and nonconvex analysis. Nevertheless, just like convex analysis, the development of abstract convexity has been mainly motivated by applications to optimization.

The works [34, 35, 44, 48, 55–57, 71, 78, 119, 121, 122, 144, 146–149, 154, 166] use abstract convexity for applications to nonconvex optimization. As a recent example, the tools of abstract convexity are used in [19, 156, 157] to derive stronger versions of Lagrangian duality and minimax theorems that are applicable to lower semicontinuous functions which are bounded below by a quadratic function. A deep study on abstract convexity can be found in the seminal book of Alexander Rubinov [144], see also the monograph by Ivan Singer [77]. Abstract convexity reflects one of the fundamental facts of convex analysis: every lower semicontinuous convex function f is the *upper envelope* of affine functions. More precisely, at every point x, we have

$$f(x) = \sup\{h(x) : \ h \text{ is an affine function}, h \leq f\}. \tag{1.2}$$

Most results in convex analysis are consequences of the two important aspects of (1.2): (i) the "supremum" operation, and (ii) the set over which this supremum is taken. Results emanating from aspect (ii) are likely to depend on specific properties of linear/affine functions. How to distinguish which facts from convex analysis follow from the "upper envelope" operation, and which follow from the particular structure of the set of linear/affine functions? Abstract convexity is the fundamental tool that addresses this crucial question: it retains the "upper envelope" operation in aspect (i) of (1.2), but changes the set of functions over which the supremum is taken. These functions are called *abstract linear* and they naturally induce the *abstract affine* sets. Since aspect (i) of (1.2) is retained, global properties of convex analysis may be preserved even when dealing with nonconvex models. Thus, this approach is sometimes called a "non-affine global support function technique" (see for example [34, 78, 144]).

Many tools from convex analysis, such as *subgradients* and *ε-subgradients*, have their "abstract" counterparts, obtained again by using abstract linear functions. For instance, the *abstract* subgradient of an abstract convex function f at a point x collects all the supporting abstract linear functions which are minorants of f (i.e., their graphs stay below the graph of f), and coincide with f at x. This extends the concept of the convex subgradient and provides a valuable tool for studying certain nonconvex optimization problems (see [34, 78, 144, 149]). Another example is the Fenchel-Moreau conjugate f^* of a function f. The definition of f^* uses the set of linear functions, and abstract convexity allows to produce "abstract" types of conjugates.

In [23], zero duality gap is shown to be equivalent to (a) certain properties involving $\varepsilon-$subgradients and (b) other facts involving conjugate functions. One of the aims of the present work is to extend these results to the context of abstract convexity. Additionally, we supplement the sum rule for abstract subdifferentials, improving the result in [78]. To the best of our knowledge, with the exception of [35], this paper is the very first attempt to consider explicitly the *weak* topology* (pointwise convergence topology) on the space of abstract linear functions to deduce calculus rules for subdifferentials. We extend the fundamental Banach–Alaoglu–Bourbaki theorem on the weak* closedness of the dual unit ball of dual space X^* to the general space $\mathcal{L}$ of abstract linear functions.

Contributions of Research

We extend conditions for zero duality gap to the context of nonconvex and nonsmooth optimisation. By using tools provided by the theory of *abstract convexity*, we establish new characterisations of zero duality gap under no assumptions on the topology of $\mathcal{L}$. Moreover, under a mild assumption on the set $\mathcal{L}$ (namely, assuming that $\mathcal{L}$ is equipped with the weak*-topology), we establish and extend several fundamental results of convex analysis. In particular, we prove that the zero duality gap property can be stated in terms of an inclusion involving ε-subdifferentials. The *weak* topology* $C(\mathcal{L}, X)$ is exploited to obtain the sum rule for abstract ε-subdifferential. The Banach–Alaoglu–Bourbaki theorem is extended to the abstract linear functions space $\mathcal{L}$. We extend a fact recently established by Borwein, Burachik and Yao in the conventional convex case (cf. Chapter 8).

1.2.3 Convex Polytopes

Convex polytopes are fundamental geometric objects, especially in optimisation. There has seen a growing interest in study their combinatorial properties.

The k-dimensional *skeleton* of a polytope P is the set of all its faces of dimension of at most k. The 1-skeleton of P is the *graph* $G(P)$ of P. We denote by $V(P)$ the vertex set of P.

This part of the thesis studies the (vertex) connectivity of a cubical polytope, the (vertex) connectivity of the graph of the polytope. A *cubical* d-polytope is a polytope with all its facets being cubes. By a cube we mean any polytope that is combinatorially equivalent to a cube; that is, one whose face lattice is isomorphic to the face lattice of a cube.

Connectivity

In the three-dimensional world, Euler's formula implies that the graph of a cubical 3-polytope P has $2|V(P)| - 4$ edges, and hence its minimum degree is three; the *degree* of a vertex records the number of edges incident with the vertex. Besides, every 3-polytope is 3-connected by Balinski's theorem [9]. Hence the dimension, minimum degree and connectivity of a cubical 3-polytope all coincide.

Theorem 1.2.1 (Balinski [9]). *The graph of a d-polytope is d-connected.*

This equality between dimension, minimum degree and connectivity of a cubical polytope no longer holds in higher dimensions. In Blind and Blind's classification of cubical d-polytopes where every vertex has degree d or $d+1$ [21], the authors exhibited cubical d-polytopes with the same graph as the $(d+1)$-cube; for an explicit example, check [79, Section 4]. More generally, the paper [79, Section 6] exhibited cubical d-polytopes with the same $(\lfloor d/2 \rfloor - 1)$-skeleton as the d'-cube for every $d' > d$, the so-called *neighbourly cubical d-polytopes*. And even more generally, Sanyal and Ziegler [151, page 422], and later Adin, Kalmanovich and Nevo [1, Section 5], produced cubical d-polytopes with the same k-skeleton as the d'-cube for every $1 \leq k \leq \lfloor d/2 \rfloor - 1$ and every $d' > d$, the so-called *k-neighbourly cubical d-polytopes*. Thus the minimum degree or connectivity of a cubical d-polytope for $d \geq 4$ does not necessarily coincide with its dimension; this is what one would expect. However, somewhat surprisingly, we can prove a result connecting the connectivity of a cubical polytope to its minimum degree, regardless of the dimension; this is a vast generalisation of a similar, and well-known, result in the d-cube [138, Proposition 1]; see also Section 10.1.

Define a *separator* of a polytope as a set of vertices disconnecting the graph of the polytope. Let

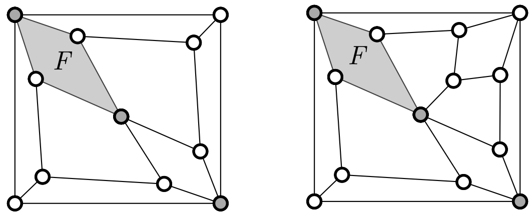
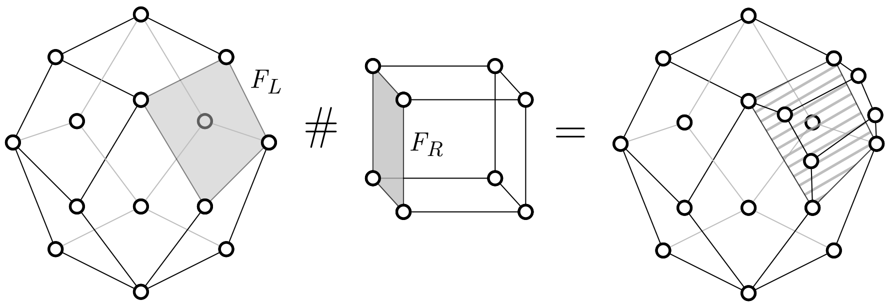


Figure 1.1: Cubical 3-polytopes with minimum separators not consisting of the neighbours of some vertex. The vertices of the separator are coloured in gray. The removal of the vertices of a face F does not leave a 2-connected subgraph: the remaining vertex in gray disconnects the subgraph. Infinitely many more examples can be generated by using well-known expansion operations such as those in [27, Fig. 3].



(a) P_1 (b) P_2 (c) $P_1 \# P_2$

Figure 1.2: Connected sum of two cubical polytopes.

X be a set of vertices in a graph G. Denote by $G[X]$ the subgraph of G induced by X, the subgraph of G that contains all the edges of G with vertices in X. Write $G - X$ for $G[V(G) \setminus X]$; that is, the subgraph $G - X$ is obtained by removing the vertices in X and their incident edges. Our main result is the following.

Theorem (Connectivity Theorem). *A cubical d-polytope P with minimum degree δ is* $\min\{\delta, 2d-2\}$-*connected for every $d \geq 3$.*

Furthermore, for any $d \geq 4$, every minimum separator X of cardinality at most $2d-3$ consists of all the neighbours of some vertex, and the subgraph $G(P) - X$ contains exactly two components, with one of them being the vertex itself.

A *simple* vertex in a d-polytope is a vertex of degree d; otherwise we say that the vertex is *nonsimple*. An immediate corollary of the theorem is the following.

Corollary. *A cubical d-polytope with no simple vertices is $(d+1)$-connected.*

Remark 1.2.2. The connectivity theorem is best possible in the sense that there are infinitely many cubical 3-polytopes with minimum separators not consisting of the neighbours of some vertex (Fig. 1.1).

It is not hard to produce examples of polytopes with differing values of minimum degree and connectivity. The *connected sum* $P_1 \# P_2$ of two d-polytopes P_1 and P_2 with two projectively isomorphic facets $F_1 \subset P_1$ and $F_2 \subset P_2$ is obtained by gluing P_1 and P_2 along F_1 and F_2 [175, Example 8.41]. Projective transformations on the polytopes P_1 and P_2, such as those in [139, Def. 3.2.3], may be required for $P_1 \# P_2$ to be convex. Figure 1.2 depicts this operation. A connected sum of two copies of a cyclic d-polytope with $d \geq 4$ and $n \geq d+1$ vertices ([175, Theorem 0.7]), which is a polytope whose facets are all simplices, results in a d-polytope of minimum degree $n-1$ that is d-connected but not $(d+1)$-connected.

On our way to prove the connectivity theorem we prove results of independent interest, for instance, the following (Corollary 10.2.5 in Section 10.2).

Corollary. *Let P be a cubical d-polytope and let F be a proper face of P. Then the subgraph $G(P)-V(F)$ is $(d-2)$-connected.*

Remark 1.2.3. The examples of Fig. 1.1 also establish that the previous corollary is best possible in the sense that the removal of the vertices of a proper face F of a cubical d-polytope does not always leave a $(d-1)$-connected subgraph of the graph of the polytope.

The corollary gives another unusual property of cubical polytopes. A tight result of Perles and Prabhu [132, Theorem 1] implies that the removal of the vertices of any k-face $(-1 \le k \le d-1)$ from a d-polytope leaves a $\max\{1, d-k-1\}$-connected subgraph of the graph of the polytope.

The connectivity theorem also gives rise to the following corollary and open problem.

Corollary. *There are functions $f : \mathbb{N} \to \mathbb{N}$ and $g : \mathbb{N} \to \mathbb{N}$ such that, for every $d \ge 4$,*

(i) the function $f(d)$ gives the maximum number such that every cubical d-polytope with minimum degree $\delta \le f(d)$ is δ-connected;

(ii) the function $g(d)$ gives the maximum number such that every minimum separator with cardinality at most $g(d)$ of every cubical d-polytope consists of the neighbourhood of some vertex; and

(iii) $2d-3 \le g(d)$ and $g(d) < f(d)$.

An exponential bound in d for $f(d)$ is readily available. The connected sum of two copies of a neighbourly cubical d-polytope with minimum degree $\delta > 2^{d-1}$, which exists by [79, Theorem 16], results in a cubical d-polytope with minimum degree δ and with a minimum separator of cardinality 2^{d-1}, the number of vertices of the facet along which we glued. The cardinality of this separator gives at once the announced upper bound. This exponential bound in conjunction with the connectivity theorem gives that

$$2d-3 \le g(d) < f(d) \le 2^{d-1}. \tag{1.3}$$

The following problem naturally arises.

Problem 1.2.4. *For $d \ge 4$ provide precise values for the functions $f(d)$ and $g(d)$ or improve the lower and upper bounds in* (1.3).

We suspect that both functions are linear in d.

Contributions of Research

We first establish that, for any $d \ge 3$, the graph of a cubical d-polytope with minimum degree δ is $\min\{\delta, 2d-2\}$-connected. Second, we show, for any $d \ge 4$, that every minimum separator of cardinality at most $2d-3$ in such a graph consists of all the neighbours of some vertex and that removing the vertices of the separator from the graph leaves exactly two components, with one of them being the vertex itself (cf. Chapter 10).

Linkedness

A graph with at least $2k$ vertices is k-linked if, for every set of $2k$ distinct vertices organised in arbitrary k unordered pairs of vertices, there are k vertex-disjoint paths joining the vertices in the pairs. If the graph of a polytope is k-linked we say that the polytope is also *k-linked.*

Larman and Mani in 1970 proved that simplicial d-polytopes, d-dimensional polytopes with all their facets being combinatorially equivalent to simplices, are $\lfloor (d+1)/2 \rfloor$-linked; this is the maximum possible linkedness given the facts that a $\lfloor (d+1)/2 \rfloor$-linked graph is at least $(2\lfloor (d+1)/2 \rfloor - 1)$-connected and that some of these graphs are d-vertex-connected but not $(d+1)$-vertex-connected.

Here we establish that d-dimensional cubical polytopes are also $\lfloor (d+1)/2 \rfloor$-linked for every $d \neq 3$; this is again the maximum possible linkedness for such a class of polytopes.

Being k-linked imposes a stronger demand on a graph than just being k-vertex-connected, or d-connected for short. A k-linked graph needs to be at least $(2k-1)$-connected, and yet there are $(2k-1)$-connected graphs that are not k-linked. The classification of 2-linked graphs [153,161] contextualised for 3-polytopes readily gives examples of this phenomenon: with the exception of simplicial 3-polytopes, no 3-polytope, despite being 3-connected by Balinski's theorem [9], is 2-linked. However, there is a linear function $f(k)$ such that every $f(k)$-connected graph is k-linked, which follows from works of Bollobás and Thomason [22]; Kawarabayashi, Kostochka and Yu [80]; and Thomas and Wollan [162]. In the case of polytopes, Larman and Mani [108, Theorem 2] proved that every d-polytope is $\lfloor (d+1)/3 \rfloor$-linked, a result that was slightly improved to $\lfloor (d+2)/3 \rfloor$ in [164, Theorem 2.2].

The first edition of the Handbook of Discrete and Computational Geometry [61, Problem 17.2.6] posed the question of whether or not every d-polytope is $\lfloor d/2 \rfloor$-linked. This question had already been answered in the negative by Gallivan in the 1970s with a construction of a d-polytope that is not $\lfloor 2(d+4)/5 \rfloor$-linked; see [116, p. 107] or [60]. A weak positive result however follows from [162]: every d-polytope with minimum degree at least $5d$ is $\lfloor d/2 \rfloor$-linked.

Restricting our attention to particular classes of polytopes gives stronger results. *Simplicial d-polytopes*, polytopes in which every facet is a simplex, are $\lfloor (d+1)/2 \rfloor$-linked [108, Theorem 2]. Since there are simplicial d-polytopes that are d-connected but not $(d+1)$-connected, the bound of $\lfloor (d+1)/2 \rfloor$ is best possible for this class of polytopes. Polytopes with small number of vertices were considered in [164], where it was shown that d-polytopes with $d+\gamma+1$ vertices are $\lfloor (d-\gamma+1)/2 \rfloor$-linked for $0 \leq \gamma \leq (d+2)/5$.

In his PhD thesis [165, Question 5.4.12] Wotzlaw asked whether every cubical d-polytope is $\lfloor d/2 \rfloor$-linked. Here we answer the question in the strongest possible way.

Theorem. *For every $d \neq 3$, a cubical d-polytope is $\lfloor (d+1)/2 \rfloor$-linked.*

Our methodology relies on results on the connectivity of strongly connected subcomplexes of cubical polytopes, whose proof ideas were first developed in [33], and a number of new insights into the structure of d-cube (Section 11.1). One obstacle that forces some tedious analysis is the fact that the 3-cube is not 2-linked.

In line with the main result of Chapter 10 (or in [33]), where it was proved that a cubical d-polytope of minimum degree δ is $\min\{\delta, 2d-2\}$-connected, we wonder if the following is true.

Question 1.2.5. *For every $\delta \neq 3$, is a cubical polytope with minimum degree δ necessarily $\lfloor (\delta+1)/2 \rfloor$-linked?*

Contributions of Research

We establish that d-dimensional cubical polytopes are $\lfloor (d+1)/2 \rfloor$-linked for every $d \neq 3$; this is again the maximum possible linkedness for such a class of polytopes (cf. Chapter 11).

1.3 Organisation of the Thesis

The organisation of the thesis is as follows:

In the first and major part of the thesis (chapter 2-6), following the works by Kruger, we examine extremality and stationarity properties of collections of sets using techniques of variational analysis. We thoroughly study metric, slope and normal cone characterisations of these properties. The core arguments used in various proofs of the extremal principle and its extensions as well as in primal and dual characterisations are fully exposed, analysed and refined, leading to a unifying theory, encompassing all existing approaches to obtaining 'extremal' statements. This work contributes to the development of nonsmooth analysis and optimisation.

In the second part (chapter 7-8), we study some generalised convexity notions using the techniques of variational analysis. In particular, robustly quasiconvex functions, which retain many important optimisation properties of convex functions, are characterised by means of Fréchet subdifferentials; and generalisations of the convex subdifferentials are studied in the framework of abstract convexity, also known as the theory of convexity without linearity. This work develops certain special global tools for solving nonsmooth optimisation problems.

In the third part (chapter 9-11), we study convex polytopes via their graphs and face lattices. Many continuous optimisation problems can be discretised into problems of finding the best solution from a finite set of points, using powerful tools like simplex method. Here, we study properties of graphs of some convex polytopes. In particular, the connectivity and the linkedness properties of graphs of cubical polytopes are explored. This work contributes to the development of combinatorics optimisation.

Part I

Extremal Principle: Characterisations of Non-Intersection Properties

Overview

This research studies a geometric non-intersection properties of finite collections of sets initiated 40 years ago by the *extremal principle* [101–103]. Models involving collections of sets have proved their usefulness in analysis and optimisation, with non-intersection properties (or their absence) being at the core of many applications: recall the ubiquitous convex *separation theorem, Dubovitskii–Milyutin formalism* [54] and various *transversality/regularity* properties [20, 52, 74, 75, 90–92, 96, 99, 100, 127].

In the setting of general Banach/Asplund spaces, we establish metric and linear dual (generalised separation) characterisations (Chapter 3) and nonlinear primal (slope) and dual characterisations (Chapter 4) of these non-intersection properties. As an application, in Chapter 5 some new (even in the linear setting) characterisations of the conventional extremality/stationarity properties are obtained. Realisations of the obtained characterisations in the Hölder setting are formulated. In Chapter 6, some open questions on the intrinsic transversality are answered thoroughly; and its primal space characterisation is established.

Chapter 2

Preliminaries

2.1 Basic Definitions and Notions

Our basic notation is standard; see e.g., [51, 75, 120, 129, 130, 143]. Throughout X is either a metric or (more often) a normed vector space. In the latter case, we often require it to be Banach or Asplund. The distance and the norm are denoted by $d(\cdot,\cdot)$ and $\|\cdot\|$, respectively. We use the same symbols to denote distances and norms in all spaces (primal and dual). When considering products of spaces, we usually assume them equipped with the maximum distance or norm. $B_\delta(x)$ and $\overline{B}_\delta(x)$ denote, respectively, the open and closed balls with center x and radius $\delta > 0$. Given a point x and a set A in X, $d(x, A) := \inf\limits_{a\in A} d(x, a)$ denotes the distance from x to A; in particular $d(x, \emptyset) := +\infty$. Given two subsets $A, B \subset X$, $d(A, B) := \inf\limits_{a\in A} d(a, B)$ denotes the distance between A and B. Given a set A, a point $a \in A$ and a number $\delta > 0$, we call the set $A \cap B_\delta(a)$ a *localisation* of the set A near a. If X is a normed vector space, its topological dual is denoted by X^* while $\langle\cdot,\cdot\rangle$ denotes the bilinear form defining the pairing between the two spaces. The open unit balls in X and X^* are denoted by $\mathbb{B}$ and $\mathbb{B}^*$, respectively. $\mathbb{N}$ stands for the set of all positive integers. We also use the notation $\mathbb{R}_\infty := \mathbb{R} \cup \{+\infty\}$. Given a function $f : X \to \mathbb{R}_\infty$, its domain is the set $\operatorname{dom} f := \{x \in X : f(x) < \infty\}$.

A set-valued mapping $F : X \rightrightarrows Y$ between two sets X and Y is a mapping, which assigns to every $x \in X$ a subset (possibly empty) $F(x)$ of Y. We use the notations

$$\operatorname{gph} F := \{(x, y) \in X \times Y \mid y \in F(x)\}, \quad \operatorname{dom} F := \{x \in X \mid F(x) \neq \emptyset\}$$

for the graph and the domain of F, respectively, and $F^{-1} : Y \rightrightarrows X$ for the inverse of F. This inverse (which always exists with possibly empty values) is defined by

$$F^{-1}(y) := \{x \in X \,|\, y \in F(x)\}, \quad y \in Y,$$

and satisfies

$$(x, y) \in \operatorname{gph} F \quad \Longleftrightarrow \quad (y, x) \in \operatorname{gph} F^{-1}.$$

Obviously, $\operatorname{dom} F^{-1} = F(X)$.

The distance from a point $x \in X$ to a set $\Omega \subset X$ is defined by $\operatorname{dist}(x, \Omega) := \inf\limits_{\omega\in\Omega} \|x - \omega\|$, and we use the convention $\operatorname{dist}(x, \Omega) = +\infty$ when $\Omega = \emptyset$. The set-valued mapping

$$P_\Omega : X \rightrightarrows X : x \mapsto \{\omega \in \Omega, \quad \|x - \omega\| = \operatorname{dist}(x, \Omega)\}$$

is the *projector* on Ω. An element $\omega \in P_\Omega(x)$ is called a *projection.* Note that the projector is not, in general, single-valued and can have empty values. The single-valuedness of P_Ω everywhere in fact defines the *Chebyshev property* of Ω. Every nonempty closed convex set in a Hilbert space is Chebyshev. The *inverse of the projector*, P_Ω^{-1}, is defined by

$$P_\Omega^{-1}(\omega) \equiv \{x \in X |\ \omega \in P_\Omega(x) \quad \forall \omega \in \Omega\}.$$

The proximal normal cone to Ω at a point $\bar{x} \in \Omega$ is defined by

$$N_\Omega^p(\bar{x}) := \text{cone}\left(P_\Omega^{-1}(\bar{x}) - \bar{x}\right),$$

which is a convex cone. Here cone $(\cdot)$ denotes the smallest cone containing the set in the brackets.

2.2 Normal Cones and Subdifferentials

In the thesis, we use dual tools – normal cones and subdifferentials, usually in the Fréchet or Clarke sense. Given a subset A of a normed vector space X and a point $\bar{x} \in A$, the set (cf. [88])

$$N_A^F(\bar{x}) := \left\{x^* \in X^* : \limsup_{x \to \bar{x},\, x \in A \setminus \{\bar{x}\}} \frac{\langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|} \leq 0\right\} \tag{2.1}$$

is the *Fréchet normal cone* to A at $\bar{x}$. It is a nonempty closed convex cone, often trivial (i.e. $N_A^F(\bar{x}) = \{0\}$). The *Clarke normal cone* to A at $\bar{x}$ is defined as the set (cf. [37])

$$N_A^C(\bar{x}) := \left\{x^* \in X^* : \langle x^*, z\rangle \leq 0 \quad \text{for all} \quad z \in T_A^C(\bar{x})\right\}, \tag{2.2}$$

where $T_A^C(\bar{x})$ is the *Clarke tangent cone* to A at $\bar{x}$:

$$T_A^C(\bar{x}) := \Big\{z \in X : \forall x_k \overset{A}{\to} \bar{x},\ \forall t_k \downarrow 0,\ \exists z_k \to z \\ \text{such that} \quad x_k + t_k z_k \in A \quad \text{for all} \quad k \in \mathbb{N}\Big\}. \tag{2.3}$$

The set (2.2) is a nonempty weak* closed convex cone, and $N_A^F(\bar{x}) \subset N_A^C(\bar{x})$. If A is a convex set, then (2.1) and (2.2) reduce to the normal cone in the sense of convex analysis (cf. e.g. [88, Proposition 1.19], [37, Proposition 2.4.4]):

$$N_A(\bar{x}) := \{x^* \in X^* : \langle x^*, x - \bar{x}\rangle \leq 0 \quad \text{for all} \quad x \in A\}.$$

We will often use the generic notation N for both Fréchet and Clarke normal cones, specifying wherever necessary that either $N := N^F$ or $N := N^C$.

The following ε-extension ($\varepsilon \geq 0$) of (2.1) is used in the sequel: the set of *ε-normal elements* to A at $a \in A$:

$$N_\varepsilon(a \mid A) := \left\{x^* \in X^* \mid \limsup_{x \to a,\, x \in A \setminus \{a\}} \frac{\langle x^*, x - a\rangle}{\|x - a\|} \leq \varepsilon\right\}. \tag{2.4}$$

When $\varepsilon = 0$, it reduces to (2.1). It is easy to check that $N_\varepsilon(a \mid A) \supset N_A(a) + \varepsilon\mathbb{B}$ for any $\varepsilon \geq 0$, and if A is not convex, the inclusion can be strict (see [84]).

Several kinds of *subdifferential sum rules* are used throughout when deducing dual space results. They are collected in the next lemma.

Lemma 2.2.1 (Subdifferential sum rules)**.** *Suppose X is a normed vector space, $f_1, f_2 : X \to \mathbb{R}_\infty$, and $\bar{x} \in \text{dom}\, f_1 \cap \text{dom}\, f_2$.*

(i) **Convex sum rule**. *Suppose f_1 and f_2 are convex and f_1 is continuous at a point in* dom f_2. *Then*

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

(ii) **Differentiable sum rule**. *Suppose f_1 is Fréchet differentiable at $\bar{x}$. Then,*

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

(iii) **Fuzzy sum rule**. *Suppose X is Asplund, f_1 is Lipschitz continuous and f_2 is lower semicontinuous in a neighbourhood of $\bar{x}$. Then, for any $\varepsilon > 0$, there exist $x_1, x_2 \in X$ with $\|x_i - \bar{x}\| < \varepsilon$, $|f_i(x_i) - f_i(\bar{x})| < \varepsilon$ $(i = 1, 2)$, such that*

$$\partial^F(f_1+f_2)(\bar{x}) \subset \partial^F f_1(x_1) + \partial^F f_2(x_2) + \varepsilon\mathbb{B}^*.$$

(iv) **Clarke–Rockafellar sum rule**. *Suppose f_1 is Lipschitz continuous and f_2 is lower semicontinuous in a neighbourhood of $\bar{x}$. Then*

$$\partial^C(f_1+f_2)(\bar{x}) \subset \partial^C f_1(\bar{x}) + \partial^C f_2(\bar{x}).$$

The first sum rule in the lemma above is the conventional subdifferential sum rule of convex analysis; see e.g. [76, Theorem 0.3.3] and [168, Theorem 2.8.7]. Together with the second one, theses are examples of *exact* sum rules. The third sum rule is known as the *fuzzy* or *approximate* sum rule (Fabian [59]) for Fréchet subdifferentials in Asplund spaces; cf., e.g., [88, Rule 2.2] and [120, Theorem 2.33]. Note that, unlike the sum rules in parts (i) and (ii) of the lemma, the subdifferentials in the right-hand side of the inclusion are computed not at the reference point, but at some points nearby. This explains the name. The fourth sum rule is formulated in terms of Clarke subdifferentials. It was established in Rockafellar [142, Theorem 2]. Similar to the previous one, it is valid generally only as inclusion. Nevertheless, it is another example of exact sum rule.

Recall that a Banach space is *Asplund* if every continuous convex function on an open convex set is Fréchet differentiable on a dense subset [136], or equivalently, if the dual of each its separable subspace is separable. We refer the reader to [25, 120, 136] for discussions about and characterisations of Asplund spaces. All reflexive, particularly, all finite dimensional Banach spaces are Asplund.

The following facts are immediate consequences of the definition of the Fréchet subdifferential and normal cone (cf. e.g. [88, Propositions 1.10 and 1.29]).

Lemma 2.2.2. *Suppose X is a normed vector space and $f : X \to \mathbb{R}_\infty$. If $\bar{x} \in \operatorname{dom} f$ is a point of local minimum of f, then $0 \in \partial^F f(\bar{x})$.*

Lemma 2.2.3. *Suppose X_1 and X_2 are normed vector spaces, $\bar{x}_1 \in A_1 \subset X_1$ and $\bar{x}_2 \in A_2 \subset X_2$. Then*

$$N^F_{A_1\times A_2}(\bar{x}_1, \bar{x}_2) = N^F_{A_1}(\bar{x}_1) \times N^F_{A_2}(\bar{x}_2).$$

We are going to use a representation of the subdifferential of a special convex function on X^n given in the next lemma; cf. [41, 99].

Lemma 2.2.4. *Let X be a normed vector space and*

$$\psi(u_1,\ldots,u_n) := \max_{1\le i\le n-1} \|u_i - a_i - u_n\|, \quad u_1,\ldots,u_n \in X, \tag{2.5}$$

where $a_i \in X$ $(i = 1,\ldots,n-1)$. Let $x_1,\ldots,x_n \in X$ and $\max\limits_{1\le i\le n-1} \|x_i - a_i - x_n\| > 0$. Then

$$\partial\psi(x_1,\ldots,x_n) = \Big\{ (x_1^*,\ldots,x_n^*) \in (X^*)^n \mid \sum_{i=1}^n x_i^* = 0, \\ \sum_{i=1}^{n-1} \|x_i^*\| = 1,\ \sum_{i=1}^{n-1} \langle x_i^*, x_i - a_i - x_n\rangle = \max_{1\le i\le n-1} \|x_i - a_i - x_n\| \Big\}. \tag{2.6}$$

Remark 2.2.5. (i) It is easy to notice that in the representation (2.6), for any $i = 1,\ldots,n-1$, either $\langle x_i^*, x_i - a_i - x_n\rangle = \max\limits_{1\le j\le n-1} \|x_j - a_j - x_n\|$ or $x_i^* = 0$.

(ii) The maximum norm on X^{n-1} used in (2.5) and (2.6) is a composition of the given norm on X and the maximum norm on $\mathbb{R}^{n-1}$. The corresponding dual norm produces the sum of the norms in (2.6). Any other finite dimensional norm can replace the maximum norm in (2.5) and (2.6) as long as the corresponding dual norm is used to replace the sum in (2.6).

2.3 Arrangement of Collections of sets

Here we consider n sets $\Omega_1, \ldots, \Omega_n$ $(2 \le n < \infty)$ and write $\{\Omega_1, \ldots, \Omega_n\}$ to denote the collection of the sets as a single object.

2.3.1 Extremality, Stationarity and Approximate Stationarity

The next definition collects several extremality and stationarity properties of collections of sets.

Definition 2.3.1. Suppose $\Omega_1, \ldots, \Omega_n$ are subsets of a normed vector space X and $\bar{x} \in \cap_{i=1}^n \Omega_i$. The collection $\{\Omega_1, \ldots, \Omega_n\}$ is

(i) **extremal** if and only if for any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying

$$\bigcap_{i=1}^n (\Omega_i - a_i) = \emptyset \quad \text{and} \quad \max_{1 \le i \le n} \|a_i\| < \varepsilon; \tag{P1}$$

(ii) **locally extremal** at $\bar{x}$ if and only if there exists a number $\rho \in]0, \infty]$ such that, for any $\varepsilon > 0$, there are vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying

$$\bigcap_{i=1}^n (\Omega_i - a_i) \cap B_\rho(\bar{x}) = \emptyset \quad \text{and} \quad \max_{1 \le i \le n} \|a_i\| < \varepsilon; \tag{P2}$$

(iii) **stationary** at $\bar{x}$ if and only if for any $\varepsilon > 0$, there exist a number $\rho \in]0, \varepsilon[$ and vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying

$$\bigcap_{i=1}^n (\Omega_i - a_i) \cap B_\rho(\bar{x}) = \emptyset \quad \text{and} \quad \max_{1 \le i \le n} \|a_i\| < \varepsilon\rho; \tag{P3}$$

(iv) **approximately stationary** at $\bar{x}$ if and only if for any $\varepsilon > 0$, there exist a number $\rho \in]0, \varepsilon[$, points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ and vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying

$$\bigcap_{i=1}^n (\Omega_i - \omega_i - a_i) \cap (\rho\mathbb{B}) = \emptyset \quad \text{and} \quad \max_{1 \le i \le n} \|a_i\| < \varepsilon\rho. \tag{P4}$$

The condition (P1) (conditions (P2), (P4)) means that an appropriate arbitrarily small shift of the sets makes them nonintersecting (in a neighborhood of $\bar{x}$). This is a very general model that embraces many optimality notions. In nonconvex analysis, the extremality properties replaces the crucial assumption on the emptyness of the intersection of one of the sets with the interior of the other set in convex separation theorem. It is easy to observe that if the collection of sets $\{\Omega_1, \ldots, \Omega_n\}$ is extremal, it is locally extremal at any points in $\cap_{i=1}^n \Omega_i$ with any positive parameter $\rho > 0$, when the sets are convex the inverse also holds. On the other hand, the extremality condition can be considered as a special case of the local extremality with $\rho = +\infty$. The other two conditions correspond to more subtle properties of optimisation problems, closer to stationarity.

Unlike the definition of local extremality (Definition 2.3.1(ii)), the definitions of stationarity and approximate stationarity (Definition 2.3.1(iii)–(iv)) relate the size of the 'shifts' of the sets to that of the neighborhood in which the sets become nonintersecting. That makes these stationarity properties have strong connections with important regular/transversal properties of collections of sets (Section 2.3.2), and closed related to the fundamental property of metric regularity of set-valued mappings (see Section 2.5). In condition (P4), instead of the common point $\bar{x}$, the sets Ω_i $(i = 1, \ldots, n)$ are considered near their own points $\omega_i \in \Omega_i$.

The formulas in Definition 2.3.1 and several their modifications discussed in Proposition 2.3.4 are central for our analysis and are going to be extensively referred to throughout the thesis. We will use special P tags: (P1), (P2), . . . for such formulas as well as some other formulas in Chapter 2 involved in primal space of extremality and stationarity. M tags: (M1), (M2),. . . are for the formulas involved in the corresponding metric characterisations. D tags: (D1), (D2), . . . are reserved for the formulas involved in the corresponding dual space characterisations of extremality and stationarity.

For easy consultation by the reader, we present below a table to categorise each group of definitions and characterisations of extremality and transversality properties in this preliminary section.

Extremality and Transversality			
	Definition/Geometric characterisation & reformulation	Metric characterisation	Dual characterisation
Extremality Local extremality Stationarity Approximate stationarity	(P1), (P5) (P2), (P6) (P3), (P7) (P4), (P8)	 (M3) (M4), (M4.1), (M4.2)	(D1), (D2) (D1), (D2) (D1), (D2) (D3), (D4), (D5), (D6)
Semitransversality Subtransversality Transversality	(P9) (P10), (P12) (P11). (P13)	(M9) (M10) (M11), (M11.1), (M11.2)	 Theorem 2.7.15 (ii), (iii)

Table 2.1: List of definitions and characterisations for extremality and transversality properties.

The properties in parts (i) and (ii) of Definition 2.3.1 were introduced in [102] and [83], respectively; see also [88, 120]. The properties in parts (iii) and (iv) first appeared in [89] and [85], respectively; see also [29, 91]. Property (iv) was referred to in [85] as *extremality near $\bar{x}$*. The name *approximate stationarity* was suggested in [92].

Unlike condition (ii), in conditions (iii) and (iv) the magnitudes of the "shifts" of the sets are related to that of the neighbourhood in which the sets become nonintersecting that is the condition $\max\limits_{1\le i\le n} \|a_i\|/\rho < \varepsilon$. Compared to (iii), in condition (iv), instead of the common point $\bar{x}$, each set Ω_i is considered near its own point ω_i.

The relationships between the properties in Definition 2.3.1 are straightforward. The equivalences in part (ii) of the proposition below were proved in [90, Proposition 14].

Proposition 2.3.2. *Suppose $\Omega_1, \ldots, \Omega_n$ are subsets of a normed vector space X and $\bar{x} \in \cap_{i=1}^n \Omega_i$.*

(i) For the properties in Definition 2.3.1, the following implications hold true: (i) ⇒ (ii) ⇒ (iii) ⇒ (iv).

(ii) If the sets are convex, then the implications in the previous item hold as equivalences: (i) ⇔ (ii) ⇔ (iii) ⇔ (iv).

(iii) If the collection $\{\Omega_1, \ldots, \Omega_n\}$ is locally extremal at $\bar{x}$ with $\rho = \infty$, then it is extremal.

(iv) If the collection $\{\Omega_1, \ldots, \Omega_n\}$ is locally extremal at $\bar{x}$ with some $\rho \in]0, \infty]$, then the collection of $n+1$ sets $\Omega_1, \ldots, \Omega_n$, $B_\rho(\bar{x})$ is extremal.

It is easy to check that all the implications in Proposition 2.3.2(i) can be strict; see examples in [29, 90, 92] and Example 2.3.3.

Thanks to Proposition 2.3.2(i), approximate stationarity is the weakest of the four properties in Definition 2.3.1. It happens to be an important type of mutual arrangement of a collection of sets in space.

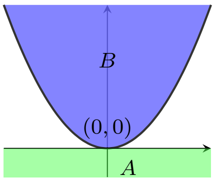


Figure 2.1: Extremality

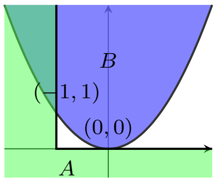


Figure 2.2: Not extremality & local extremality

The next example illustrates the difference between the extremality, the local extremality, the stationarity, and the approximate stationarity.

Example 2.3.3. 1. The sets $A := \{(x_1, x_2) \mid x_2 \le 0\}$ and $B := \{(x_1, x_2) \mid x_1^2 \le x_2\}$ in $\mathbb{R}^2$ (see Fig. 2.1) are obviously extremal.

2. If the set A above is modified slightly: $A := \{(x_1, x_2) \mid x_2 \le 0 \text{ or } x_1 \le -1\}$ (see Fig. 2.2), then $\{A, B\}$ is not extremal any more. At the same time, it is still locally extremal at $(0, 0) \in A \cap B$ (but not at $(-1, 1)$!).

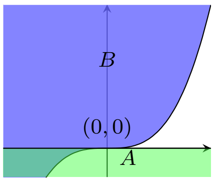


Figure 2.3: Stationarity

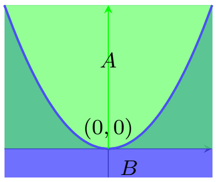


Figure 2.4: Stationarity

3. The set A is as in example 1. (figure 2.1), but the set B is the epigraph of the function x^3, $B := \{(x_1, x_2) : x_2 \le x_1^3\}$ (see figure 2.3), then $\{A, B\}$ is not locally extremal at $(0, 0)$, but stationary at $(0, 0)$.

4. The sets $A := \{(x_1, x_2) : x_2 \ge 0\}$ and $B := \{(x_1, x_2) : x_2 \le x_1^2\}$ in $\mathbb{R}^2$ (see figure 2.4) are also stationary at $(0, 0)$.

5. The sets $A := \{(x_1, x_2) : x_2 \le |x_1|\}$ and $B := \{(x_1, x_2) : x_2 \ge -|x_1|\}$ in $\mathbb{R}^2$ (see figure 2.5) are not stationary at $(0, 0)$ but approximately stationary at $(0, 0)$.

6. The sets $A := \{(x_1, x_2) : x_2 \le 0\}$ and $B := \{(x_1, x_2) : x_2 \ge x_1\}$ in $\mathbb{R}^2$ (see figure 2.6) are not approximately stationary at $(0, 0)$.

In Definition 2.3.1, it is sufficient to consider translations of all but one sets allowing the remaining set unchanged. Based on this observation, the following proposition provides asymmetric conditions which are useful for our analysis in Chapters 3 and (4).

Proposition 2.3.4. *Let $\Omega_1, \ldots, \Omega_n$ be subsets of a normed vector space X and $\bar{x} \in \cap_{i=1}^n \Omega_i$. The collection $\{\Omega_1, \ldots, \Omega_n\}$ is*

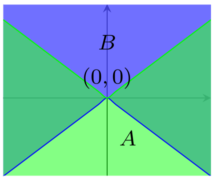


Figure 2.5: Approximate stationarity

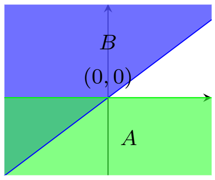


Figure 2.6: Not approximate stationarity

(i) *extremal if and only if, for any* $\varepsilon > 0$*, there exist vectors* $a_i \in X$ $(i = 1, \ldots, n-1)$ *satisfying*

$$\bigcap_{i=1}^{n-1}(\Omega_i - a_i) \cap \Omega_n = \emptyset, \qquad \max_{1\le i\le n-1}\|a_i\| < \varepsilon; \tag{P5}$$

(ii) *locally extremal at* $\bar{x}$ *if and only if there exists a number* $\rho \in]0, +\infty]$ *such that, for any* $\varepsilon > 0$*, there exist vectors* $a_i \in X$ $(i = 1, \ldots, n-1)$ *satisfying*

$$\bigcap_{i=1}^{n-1}(\Omega_i - a_i) \cap \Omega_n \cap B_\rho(\bar{x}) = \emptyset, \qquad \max_{1\le i\le n-1}\|a_i\| < \varepsilon; \tag{P6}$$

moreover, if $\{\Omega_1, \ldots, \Omega_n\}$ *is locally extremal at* $\bar{x}$ *with some* $\rho \in]0, +\infty]$*, then the above condition holds with any* $\rho' \in]0, \rho[$ *in place of* ρ*; if* $\rho = +\infty$*, one can take* $\rho' := +\infty$*;*

(iii) *stationary at* $\bar{x}$ *if and only if, for any* $\varepsilon > 0$*, there exist a number* $\rho \in]0, \varepsilon[$ *and vectors* $a_i \in X$ $(i = 1, \ldots, n-1)$ *satisfying*

$$\bigcap_{i=1}^{n-1}(\Omega_i - a_i) \cap \Omega_n \cap B_\rho(\bar{x}) = \emptyset \quad \text{and} \quad \max_{1\le i\le n-1}\|a_i\| < \varepsilon\rho; \tag{P7}$$

(iv) *approximately stationary at* $\bar{x}$ *if and only if, for any* $\varepsilon > 0$*, there exist a number* $\rho \in]0, \varepsilon[$*, points* $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ $(i = 1, \ldots, n)$ *and vectors* $a_i \in X$ $(i = 1, \ldots, n-1)$ *satisfying*

$$\bigcap_{i=1}^{n-1}(\Omega_i - \omega_i - a_i) \cap (\Omega_n - \omega_n) \cap (\rho\mathbb{B}) = \emptyset \quad \text{and} \quad \max_{1\le i\le n-1}\|a_i\| < \varepsilon\rho. \tag{P8}$$

Proof. The properties above imply the corresponding ones in Definition 2.3.1 with $a_n = 0$.

For the opposite implication, given vectors $a_i \in X$ $(i = 1, \ldots, n)$, it is natural to consider vectors $a'_i := a_i - a_n$ $(i = 1, \ldots, n-1)$. Then

$$\bigcap_{i=1}^{n-1}(\Omega_i - a'_i) \cap \Omega_n = \bigcap_{i=1}^{n}(\Omega_i - a_i) + a_n. \tag{2.7}$$

From condition (P1), the set $\cap_{i=1}^{n}(\Omega_i - a_i)$ is nonempty, so is $\cap_{i=1}^{n-1}(\Omega_i - a'_i)$. Hence, (P5) holds with the collection of vectors a'_i's in place of a_i's. Given an $\varepsilon > 0$, if one of the conditions $\max_{1\le i\le n-1}\|a_i\| < \varepsilon$ or $\max_{1\le i\le n-1}\|a_i\| < \rho\varepsilon$ is satisfied with some $\varepsilon' \in]0, \varepsilon/2]$ in place of ε, then the corresponding condition $\max_{1\le i\le n-1}\|a_i\| < \varepsilon$ or $\max_{1\le i\le n-1}\|a_i\| < \rho\varepsilon$ is satisfied with the collection a'_i's in place of a_i's. Given a $\rho \in]0, +\infty]$ and a $\rho' \in]0, \rho[$, one can take a smaller $\varepsilon' > 0$ to ensure that $\rho' + \varepsilon' < \rho$. Then, in view of (2.7) and assuming $\max_{1\le i\le n-1}\|a_i\| < \varepsilon$ with ε' in place of ε, we have

$$\bigcap_{i=1}^{n-1}(\Omega_i - a'_i) \cap \Omega_n \cap B_{\rho'}(\bar{x}) \subset \bigcap_{i=1}^{n}(\Omega_i - a_i) \cap B_\rho(\bar{x}).$$

The above inclusion obviously holds also with $\rho' = \rho = +\infty$. Thus, condition (P2) implies condition (P6) with a_i''s and ρ' in place of a_i's and ρ, respectively. Observe that conditions (P4) and (P8) are actually conditions (P2) and (P6), respectively, with the sets $\Omega_i - \omega_i$ $(i = 1, \ldots, n)$ in place of Ω_i $(i = 1, \ldots, n)$. Hence, condition (P4) implies condition (P8) with a_i''s and ρ' in place of a_i's and ρ, respectively. It follows that the properties in Definition 2.3.1 imply the corresponding ones in the proposition with a_i''s in place of a_i's, as well as the 'moreover' part in item (ii). □

There exist several modifications of the properties discussed in Section 2.7, scattered in the literature. Below we briefly discuss some of them which are going to be important for our subsequent study.

The properties in Definition 2.3.1 involve translations of all the sets. It is easy to see that in all the properties it is sufficient to consider translations of all but one sets leaving the remaining set unchanged. This simple observation leads to asymmetric conditions in the next proposition which can be useful, especially in the case $n = 2$.

Proposition 2.3.5 (Distance characterisations of extremality)**.** *Suppose X is a normed linear space, $A, B \subset X$ are closed and $A \cap B \neq \emptyset$. The pair $\{A, B\}$ is extremal if and only if for any $\varepsilon > 0$ there exist $u, v \in X$ such that* $\max\{\|u\|, \|v\|\} < \varepsilon$ *and any one of the two following (equivalent) conditions hold:*

(i) $d(a - u, B - v) > 0$ for all $a \in A$;

(ii) $d(b - v, A - u) > 0$ for all $b \in B$.

Proof. It is sufficient to show that each of the conditions (i) or (ii) is equivalent to $(A - u) \cap (B - v) = \emptyset$. Each of these conditions obviously implies $(A - u) \cap (B - v) = \emptyset$. Conversely, if $(A - u) \cap (B - v) = \emptyset$, then $a - u \notin B - v$ for any $a \in A$, and consequently, since B is closed, $d(a - u, B - v) > 0$, i.e. condition (i) is satisfied. Similarly, since A is closed, condition $(A - u) \cap (B - v) = \emptyset$ implies $d(b - v, A - u) > 0$ for all $b \in B$, hence, condition (ii). □

The closedness assumption in Proposition 2.3.5 cannot be dropped.

Example 2.3.6. The pair of sets $A := \mathbb{R}^2 \setminus \{(t, 0) \mid t > 0\}$ and $B := \{(0, 0)\}$ in $\mathbb{R}^2$ is obviously extremal in the sense of Definition 2.3.1(i). At the same time, $d(a - u, B - v) = d(b - v, A - u) = 0$ for all $a \in A$, $b \in B$ and $u, v \in \mathbb{R}^2$.

Remark 2.3.7. Condition $(A - u) \cap (B - v) = \emptyset$ which is crucial for the extremality property in Definition 2.3.1 is obviously implied by the stronger condition $d(A - u, B - v) > 0$, which is also stronger than each of the conditions (i) or (ii) in Proposition 2.3.5. As the next example shows, condition $d(A - u, B - v) > 0$ can be strictly stronger than $(A - u) \cap (B - v) = \emptyset$ even when both A and B are closed.

Example 2.3.8. In this example, we use the superscript notation Kth term to indicate the order of an element in a sequence in ℓ^∞.

Consider two sets in ℓ^∞:

$$A := \left\{ x = (x^k) \mid x^K \in [K, K+1] \cup \left[K + 1 + \frac{1}{K}, K + 2\right] \text{ for some } K \in \mathbb{N}; \right.$$
$$\left. x^k \in [-1, 1] \text{ for all } k \neq K \right\},$$
$$B := \left\{ x = (x^k) \mid x^K = K + 1 \text{ for some } K \in \mathbb{N};\ x^k = 0 \text{ for all } k \neq K \right\}.$$

Observe that $A \cap B = B \neq \emptyset$.

We first show that A is closed. Let $(x_n) \subset A$ and $x_n \to x_0 \in \ell^\infty$. There exist numbers $K, N \in \mathbb{N}$ such that $x_n^K \in [K, K + 1] \cup [K + 1 + 1/K, K + 2]$ for all $n > N$. Indeed, assume on the contrary that for any $N > 0$ there exist $m, n > N$, $m \neq n$ and $K_m, K_n \in \mathbb{N}$, $K_n \neq K_m$ such that $x_m^{K_m} \in [K_m, K_m+1] \cup [K_m+1+1/K_m, K_m+2]$ and $x_n^{K_n} \in [K_n, K_n+1] \cup [K_n+1+1/K_n, K_n+2]$. Then

$$\begin{aligned}\|x_n - x_m\| &\geq \max\left\{|x_n^{K_n} - x_m^{K_n}|, |x_n^{K_m} - x_m^{K_m}|\right\} \\ &\geq \max\left\{|x_n^{K_n}| - |x_m^{K_n}|, |x_m^{K_m}| - |x_n^{K_m}|\right\} \geq \max\{K_n, K_m\} - 1 \geq 1,\end{aligned}$$

which contradicts the assumption that (x_n) is convergent. Hence, $x_0 = (x_0^1, x_0^2, \ldots)$ with $x_0^K \in [K, K+1] \cup [K+1+1/K, K+2]$ and $x_0^k \in [-1, 1]$ for all $k \neq K$. Thus, A is closed.

A similar argument can be used to show that B is closed. Observe that for all $x_1, x_2 \in B$ with $x_1 \neq x_2$ one has $\|x_1 - x_2\| = \max\{K_1, K_2\} + 1$ for some $K_1, K_2 \in \mathbb{N}$, $K_1 \neq K_2$. Hence, $\|x_1 - x_2\| > 1$. It follows that any convergent sequence $(x_n) \subset B$ must be stationary when n is sufficiently large. This immediately yields the closedness of B.

Now we show that $\{A, B\}$ is extremal. Given an $\varepsilon \in (0,1)$, find and an $n \in \mathbb{N}$ such that $1/n < \varepsilon \leq 1/(n-1)$ and define a $u \in \ell^\infty$ as follows: $u^i = 1/n$ if $i < n$, and $u^i = 1/(i+1)$ if $i \geq n$. We have $\|u\| \leq \dfrac{1}{n} < \varepsilon$. Let $b = (b^k) \in B$, i.e. there exists a $K \in \mathbb{N}$ such that $b^K = K+1$ and $b^k = 0$ for all $k \neq K$. Then $0 < (b+u)^k \leq 1/n < 1$ for all $k \neq K$. If $K < n$, then $(b+u)^K = K+1+1/n$. If $K \geq n$, then $(b+u)^K = K+1+1/(K+1)$. In both cases, $K+1 < (b+u)^K < K+1+1/K$. Hence, $b+u \notin A$, and consequently, $(A-u) \cap B = \emptyset$.

Let $u = (u^k), v = (v^k) \in \ell^\infty$ be such that $\max\{\|u\|, \|v\|\} < 1/2$ and $(A-u) \cap (B-v) = \emptyset$. We are going to show that $d(A-u, B-v) = 0$. Obviously $\|u - v\| < 1$. Moreover, $|u^k - v^k| < 1/k$ for all $k \in \mathbb{N}$. Indeed, suppose on the contrary that $1/K \leq |u^K - v^K| < 1$ for some $K \in \mathbb{N}$ and choose a $b = (b^k) \in B$ such that $b^K = K+1$ and $b^k = 0$ for $k \neq K$. Then for any $k \neq K$, we have $\|(b+u-v)^k\| = \|(u-v)^k\| < 1$, and $(b+u-v)^K = K+1+u^K - v^K$, and consequently, either $K < (b+u-v)^K < K+1$ or $K+1+1/K \leq (b+u-v)^K < K+2$. In any case, $b+u-v \in A$, and $b - v \in A - u$, which is a contradiction. Thus, $|u^k - v^k| < 1/k$ for all $k \in \mathbb{N}$. For any $\varepsilon > 0$, we can find a $b \in B$ and a $K \in \mathbb{N}$ such that $b^K \neq 0$, $|u^K - v^K| < \varepsilon$. Set $a^K := b^K$ and $a^k := u^k - v^k$ for all $k \neq K$. Then $a = (a^k) \in A$ and $\|(a-u) - (b-v)\| = |a^K - b^K| < \varepsilon$. Hence, $d(A-u, B-v) = 0$.

2.3.2 Semitransversality, Subtransversality and Transversality

The content of this subsection is mostly taken from [28, Section 2]. In this section, we study 'good arrangements' of collections of sets in normed vector spaces near a point in their intersection, known as *transversality* (*regularity*) properties and playing an important role in optimisation and variational analysis, e.g., as constraint qualifications in optimality conditions, and qualification conditions in subdifferential, normal cone and coderivative calculus, and convergence analysis of computational algorithms [2, 8, 13–15, 28, 39, 40, 42, 43, 52, 75, 90–92, 96–100, 104–107, 109, 111, 123–125, 128, 129, 171, 173].

The next definition is a modification of [106, Definition 3.1].

Definition 2.3.9. Let $\Omega_1, \ldots, \Omega_n$ be subsets of a normed vector space X, $\bar{x} \in \cap_{i=1}^n \Omega_i$ and $\alpha > 0$.

(i) $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$*semitransversal* at $\bar{x}$ if and only if there exists a number $\delta > 0$ such that

$$\bigcap_{i=1}^n (\Omega_i - a_i) \cap B_\rho(\bar{x}) \neq \emptyset \tag{P9}$$

for all $\rho \in]0, \delta[$ and $a_i \in X$ $(i = 1, \ldots, n)$ with $\max\limits_{1 \leq i \leq n} \|a_i\| < \alpha\rho$.

(ii) $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$*subtransversal* at $\bar{x}$ if and only if there exist numbers $\delta_1 > 0$ and $\delta_2 > 0$ such that

$$\bigcap_{i=1}^n \Omega_i \cap B_\rho(x) \neq \emptyset \tag{P10}$$

for all $\rho \in]0, \delta_1[$ and $x \in B_{\delta_2}(\bar{x})$ with $\max_{1 \le i \le n} d(x, \Omega_i) < \alpha\rho$.

(iii) $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$*transversal* at $\bar{x}$ if and only if there exist numbers $\delta_1 > 0$ and $\delta_2 > 0$ such that

$$\bigcap_{i=1}^{n} (\Omega_i - \omega_i - a_i) \cap (\rho\mathbb{B}) \neq \emptyset \tag{P11}$$

for all $\rho \in]0, \delta_1[$, $\omega_i \in \Omega_i \cap B_{\delta_2}(\bar{x})$ and $a_i \in X$ $(i = 1, \ldots, n)$ with $\max_{1 \le i \le n} \|a_i\| < \alpha\rho$.

Three properties in Defintion 2.3.9 are known under various names. A table illustrating the evolution of the terminology can be found in [99, Section 2].

Each of the properties in Definition 2.3.9 is determined by a number $\alpha > 0$, playing the role of a rate of the respective property, and either a number $\delta > 0$ in item (i) or two numbers $\delta_1 > 0$ and $\delta_2 > 0$ in items (ii) and (iii). The exact upper bound of all $\alpha > 0$ such that the property holds with some $\delta > 0$, or $\delta_1 > 0$ and $\delta_2 > 0$, is called the *modulus* of this property. We use notations $\mathrm{s_e tr}\,[\Omega_1, \ldots, \Omega_n](\bar{x})$, $\mathrm{str}\,[\Omega_1, \ldots, \Omega_n](\bar{x})$ and $\mathrm{tr}\,[\Omega_1, \ldots, \Omega_n](\bar{x})$ for the moduli of the respective properties. If the property does not hold, then by convention the respective modulus equals 0.

If $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$semitransversal (respectively, $\alpha-$subtransversal or $\alpha-$transversal) at $\bar{x}$ with some $\alpha > 0$ and $\delta > 0$ (respectively, $\delta_1 > 0$ and $\delta_2 > 0$), we often simply say that $\{\Omega_1, \ldots, \Omega_n\}$ is semitransversal (respectively, subtransversal or transversal) at $\bar{x}$.

The role of the δ's in the definitions is more technical: they control the size of the interval for the values of ρ and, in the case of subtransversality and transversality in parts (ii) and (iii), the size of the neighbourhoods of $\bar{x}$ involved in the respective definitions. Of course, if a property is satisfied with some $\delta_1 > 0$ and $\delta_2 > 0$, it is satisfied also with the single $\delta := \min\{\delta_1, \delta_2\}$ in place of both δ_1 and δ_2. We use here two different parameters to emphasise their different roles in the definitions and the corresponding characterisations. Moreover, we are going to provide quantitative estimates for the values of these parameters, which can be important in applications.

The negation (or the absence) of the stationarity (see Definition 2.3.1(iii)) implies that the value $\mathrm{s_e tr}\,[\Omega_1, \ldots, \Omega_n](\bar{x}) > 0$ that is $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$semitransversal at $\bar{x}$ with some $\delta > 0$ and some $\alpha > 0$. The negation of the approximate stationarity (Definition 2.3.1) means $\mathrm{tr}\,[\Omega_1, \ldots, \Omega_n](\bar{x}) > 0$, or the collection $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$transversal at $\bar{x}$ with some $\delta_1, \delta_2 > 0$ and some $\alpha > 0$ (cf. [29, 90–92]).

The representation of the subtransversality property in part (ii) of Definition 2.3.9 differs from the corresponding one in [106, Definition 3.1]. In view of the next proposition, the two representations are equivalent.

Proposition 2.3.10. *Let $\Omega_1, \ldots, \Omega_n$ be subsets of a normed vector space X, $\bar{x} \in \cap_{i=1}^n \Omega_i$, and $\alpha > 0$. $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$subtransversal at $\bar{x}$ with some $\delta_1 > 0$ and $\delta_2 > 0$ if and only if*

$$\bigcap_{i=1}^{n} (\Omega_i + (\alpha\rho)\mathbb{B}) \cap B_{\delta_2}(\bar{x}) \subset \bigcap_{i=1}^{n} \Omega_i + \rho\mathbb{B} \tag{P12}$$

for all $\rho \in]0, \delta_1[$.

Proof. Condition (P12) is equivalent to the inclusion $x \in \cap_{i=1}^n \Omega_i + \rho\mathbb{B}$ holding for all $x \in B_{\delta_2}(\bar{x})$ with $\max_{1 \le i \le n} d(x, \Omega_i) < \alpha\rho$. In its turn, this inclusion is equivalent to (P10). □

If all the sets coincide, the subtransversality property is satisfied trivially.

Proposition 2.3.11. *Let Ω be a subset of a normed vector space X, $\bar{x} \in \Omega$, and $\alpha \in]0, 1]$. The collection $\{\Omega, \ldots, \Omega\}$ of $n \ge 2$ copies of Ω is $\alpha-$subtransversal at $\bar{x}$ with any $\delta_1 > 0$ and $\delta_2 > 0$.*

Proof. Condition (P10) in the current setting takes the form $\Omega \cap B_\rho(x) \neq \emptyset$ and is trivially satisfied if $d(x, \Omega) < \alpha\rho$. $\square$

The transversality property in part (iii) of Definition 2.3.9 admits several alternative representations. The three equivalent representations in the next proposition are of independent interest. They differ from the one in Definition 2.3.9(iii) by values of the parameters δ_1 and δ_2. The relationship between the values of the parameters in the two groups of representations can be estimated.

Proposition 2.3.12. *Let $\Omega_1, \ldots, \Omega_n$ be subsets of a normed vector space X, $\bar{x} \in \cap_{i=1}^n \Omega_i$, and $\alpha > 0$. Then, $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$transversal at $\bar{x}$ if and only if there exist numbers $\delta_1 > 0$ and $\delta_2 > 0$ such that the following equivalent conditions hold:*

(i) condition (P11) *is satisfied for all $\rho \in]0, \delta_1[$, $\omega_i \in \Omega_i$ and $a_i \in X$ with $\omega_i + a_i \in B_{\delta_2}(\bar{x})$ and $\|a_i\| < \alpha\rho$ $(i = 1, \ldots, n)$;*

(ii) condition (P9) *is satisfied for all $\rho \in]0, \delta_1[$ and $a_i \in \delta_2\mathbb{B}$ with $d(\bar{x}, \Omega_i - a_i) < \alpha\rho$ $(i = 1, \ldots, n)$;*

(iii) for all $\rho \in]0, \delta_1[$, $x \in X$ and $a_i \in X$ with $x + a_i \in B_{\delta_2}(\bar{x})$ and $d(x, \Omega_i - a_i) < \alpha\rho$ $(i = 1, \ldots, n)$, we have

$$\bigcap_{i=1}^{n} (\Omega_i - a_i) \cap B_\rho(x) \neq \emptyset. \tag{P13}$$

Moreover, if $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$transversal at $\bar{x}$ with some $\delta_1 > 0$ and $\delta_2 > 0$, then conditions (i)–(iii) *hold with any $\delta_1' \in]0, \delta_1]$ and $\delta_2' > 0$ satisfying $\delta_2' + \alpha\delta_1' \leq \delta_2$ in place of δ_1 and δ_2.*

Conversely, if conditions (i)–(iii) *hold with some $\delta_1 > 0$ and $\delta_2 > 0$, then $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$transversal at $\bar{x}$ with any $\delta_1' \in]0, \delta_1]$ and $\delta_2' > 0$ satisfying $\delta_2' + \alpha\delta_1' \leq \delta_2$ in place of δ_1 and δ_2.*

Proof. We first prove the equivalence of conditions (i)–(iii). Let numbers $\delta_1 > 0$ and $\delta_2 > 0$ be given.

(iii) $\Rightarrow$ (ii). This implication is trivial.

(ii) $\Rightarrow$ (i). Let $\rho \in]0, \delta_1[$, $\omega_i \in \Omega_i$, $a_i \in X$, $\omega_i + a_i \in B_{\delta_2}(\bar{x})$ and $\|a_i\| < \alpha\rho$ $(i = 1, \ldots, n)$. Set $a_i' := a_i + \omega_i - \bar{x}$ $(i = 1, \ldots, n)$. We have $a_i \in \delta_2\mathbb{B}$ and $d(\bar{x}, \Omega_i - a_i') \leq \|\bar{x} - \omega_i + a_i'\| = \|a_i\| < \alpha\rho$ $(i = 1, \ldots, n)$. By (ii), (P9) is satisfied with a_i' in place of a_i $(i = 1, \ldots, n)$. This is equivalent to (P11).

(i) $\Rightarrow$ (iii). Let $\rho \in]0, \delta_1[$, $x \in X$, $a_i \in X$, $x + a_i \in B_{\delta_2}(\bar{x})$ and $d(x, \Omega_i - a_i) < \alpha\rho$ $(i = 1, \ldots, n)$. Choose $\omega_i \in \Omega_i$ such that $\|x - \omega_i + a_i\| < \alpha\rho$ and set $a_i' := x - \omega_i + a_i$ $(i = 1, \ldots, n)$. We have $\omega_i + a_i' = x + a_i \in B_{\delta_2}(\bar{x})$ and $\|a_i'\| < \alpha\rho$ $(i = 1, \ldots, n)$. By (i), (P13) is satisfied with a_i' in place of a_i $(i = 1, \ldots, n)$. This is equivalent to (P11).

Suppose $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$transversal at $\bar{x}$ with some $\delta_1 > 0$ and $\delta_2 > 0$, and let $\delta_1' \in]0, \delta_1]$ and $\delta_2' > 0$ be such that $\delta_2' + \alpha\delta_1' \leq \delta_2$. Then, for all $\rho \in]0, \delta_1'[$, $\omega_i \in \Omega_i$ and $a_i \in X$ with $\omega_i + a_i \in B_{\delta_2'}(\bar{x})$ and $\|a_i\| < \alpha\rho$ $(i = 1, \ldots, n)$, we have $\|\omega_i - \bar{x}\| \leq \|\omega_i + a_i - \bar{x}\| + \|a_i\| < \delta_2' + \alpha\delta_1' \leq \delta_2$ $(i = 1, \ldots, n)$. By Definition 2.3.9 (iii), (P11) is satisfied, and consequently condition (i) (as well as conditions (ii) and (iii)) holds with δ_1' and δ_2'.

Conversely, suppose conditions (i)–(iii) hold with some $\delta_1 > 0$ and $\delta_2 > 0$, and let $\delta_1' \in]0, \delta_1]$ and $\delta_2' > 0$ be such that $\delta_2' + \alpha\delta_1' \leq \delta_2$. Then, for all $\rho \in]0, \delta_1'[$, $\omega_i \in \Omega_i \cap B_{\delta_2'}(\bar{x})$ and $a_i \in X$ $(i = 1, \ldots, n)$ with $\max\limits_{1 \leq i \leq n} \|a_i\| < \alpha\rho$, we have $\|\omega_i + a_i - \bar{x}\| \leq \|\omega_i - \bar{x}\| + \|a_i\| < \delta_2' + \alpha\delta_1' \leq \delta_2$, or $\omega_i + a_i \in B_{\delta_2}(\bar{x})$. By (i), (P11) is satisfied, and consequently $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$transversal at $\bar{x}$ with δ_1' and δ_2'. $\square$

Remark 2.3.13. (i) The inequality $\delta_2' + \alpha\delta_1' \leq \delta_2$ in the last part of Proposition 2.3.12 and some statements below can be replaced by the equality $\delta_2' + \alpha\delta_1' = \delta_2$ providing in a sense the best estimate for the values of the parameters δ_1' and δ_2'.

(ii) Each of the conditions (i)–(iii) in Proposition 2.3.12 can serve as an equivalent definition of $\alpha-$transversality. For the estimates derived in the subsequent statements in this thesis, such equivalent definitions can have some advantages compared to the original one in Definition 2.3.9 (iii). In particular, they allow to reduce some conventional proofs to a few lines.

$\alpha-$transversality is the strongest of the three properties in Definition 2.3.9.

Proposition 2.3.14. *Let $\Omega_1, \ldots, \Omega_n$ be subsets of a normed vector space X, $\bar{x} \in \cap_{i=1}^n \Omega_i$, and $\alpha > 0$. If $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$transversal at $\bar{x}$ with some $\delta_1 > 0$ and $\delta_2 > 0$, then it is $\alpha-$semitransversal at $\bar{x}$ with $\delta := \delta_1$ and $\alpha-$subtransversal at $\bar{x}$ with any $\delta_1' \in]0, \delta_1]$ and $\delta_2' > 0$ such that $\delta_2' + \alpha\delta_1' \leq \delta_2$. As a consequence,* $\mathrm{s_e tr}[\Omega_1, \ldots, \Omega_n](\bar{x}) \geq \mathrm{tr}[\Omega_1, \ldots, \Omega_n](\bar{x})$ *and* $\mathrm{str}[\Omega_1, \ldots, \Omega_n](\bar{x}) \geq \mathrm{tr}[\Omega_1, \ldots, \Omega_n](\bar{x})$.

Proof. Let $\{\Omega_1, \ldots, \Omega_n\}$ be $\alpha-$transversal at $\bar{x}$ with some $\delta_1 > 0$ and $\delta_2 > 0$. Since condition (P9) is a particular case of condition (P11) with $\omega_i = \bar{x}$ $(i = 1, \ldots, n)$, we conclude that $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$semitransversal at $\bar{x}$ with $\delta := \delta_1$. Since $d(\bar{x}, \Omega_i - a_i) \leq \|a_i\|$ for any $a_i \in X$ and any $i = 1, \ldots, n$, $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$subtransversal at $\bar{x}$ with any $\delta_1' \in]0, \delta_1]$ and $\delta_2' > 0$ such that $\delta_2' + \alpha\delta_1' \leq \delta_2$ in view of Proposition 2.3.12 (ii). □

Properties $\alpha-$semtransversality and $\alpha-$subtransversality are in general independent; cf. [106, Section 3.2].

All three transversality properties in Definition 2.3.9 are only meaningful when $\bar{x} \in \mathrm{bd} \cap_{i=1}^n \Omega_i$.

Proposition 2.3.15. *Let $\Omega_1, \ldots, \Omega_n$ be subsets of a normed vector space X. If $\bar{x} \in \mathrm{int} \cap_{i=1}^n \Omega_i$, then, for any $\alpha > 0$, all three properties in Definition 2.3.9 hold true (with some $\delta > 0$, or $\delta_1 > 0$ and $\delta_2 > 0$).*

Proof. Let $\bar{x} \in \mathrm{int} \cap_{i=1}^n \Omega_i$ and $\alpha > 0$. In view of Proposition 2.3.14, we only need to prove that $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$transversal at $\bar{x}$. Choose a number $\delta > 0$ such that $B_\delta(\bar{x}) \subset \cap_{i=1}^n \Omega_i$. Then $\bar{x} \in \cap_{i=1}^n (\Omega_i - a_i)$ for all $a_i \in \delta\mathbb{B}$ $(i = 1, \ldots, n)$, and consequently, condition (P9) is satisfied for all $\rho > 0$. Hence, condition (ii) in Proposition 2.3.12 holds with $\delta_2 := \delta$ and any $\delta_1 > 0$, and $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$transversal at $\bar{x}$. □

In view of Proposition 2.3.15, the collection $\{\Omega_1, \ldots, \Omega_n\}$ being approximately stationary at some $\bar{x} \in \cap_{i=1}^n \Omega_i$ implies $\bar{x} \in \mathrm{bd} \cap_{i=1}^n \Omega_i$.

If $\bar{x} \in \mathrm{bd} \cap_{i=1}^n \Omega_i$ and the sets are closed, then the $\alpha-$subtransversality and $\alpha-$transversality properties can only hold with $\alpha \leq 1$.

Proposition 2.3.16. *Let $\Omega_1, \ldots, \Omega_n$ be closed subsets of a normed vector space X, $\bar{x} \in \mathrm{bd} \cap_{i=1}^n \Omega_i$ and $\alpha > 0$. If $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$subtransversal (particularly, if it is $\alpha-$transversal) at $\bar{x}$, then $\alpha \leq 1$.*

Proof. Let $\{\Omega_1, \ldots, \Omega_n\}$ be $\alpha-$subtransversal at $\bar{x}$ with some $\delta_1 > 0$ and $\delta_2 > 0$. Choose a point $x \notin \cap_{i=1}^n \Omega_i$ such that $\|x - \bar{x}\| < \min\{\delta_2, \alpha\delta_1\}$. Then $x \in B_{\delta_2}(\bar{x})$, $0 < \max\limits_{1 \leq i \leq n} d(x, \Omega_i) \leq d(x, \cap_{i=1}^n \Omega_i) \leq \|x - \bar{x}\| < \alpha\delta_1$, and by Definition 2.3.9 (ii), $\cap_{i=1}^n \Omega_i \cap B_\rho(x) \neq \emptyset$ for all $\rho \in]0, \delta_1[$ satisfying $d(x, \cap_{i=1}^n \Omega_i) < \alpha\rho$, which is only possible when $\alpha \leq 1$. If $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$transversal at $\bar{x}$, then by Proposition 2.3.14, it is $\alpha-$subtransversal at $\bar{x}$, and $\alpha \leq 1$. □

In the convex case, the requirements that the relations in parts (i) and (iii) of Definition 2.3.9 hold for all small $\rho > 0$ can be relaxed.

Proposition 2.3.17. *Let $\Omega_1, \ldots, \Omega_n$ be convex subsets of a normed vector space X, $\bar{x} \in \cap_{i=1}^n \Omega_i$ and $\alpha > 0$.*

(i) $\{\Omega_1, \ldots, \Omega_n\}$ *is* $\alpha-$*semitransversal at* $\bar{x}$ *with some* $\delta > 0$ *if and only if*

$$\bigcap_{i=1}^{n}(\Omega_i - a_i) \cap B_\delta(\bar{x}) \neq \emptyset \tag{2.8}$$

for all $a_i \in X$ *with* $\|a_i\| < \alpha\delta$ $(i = 1, \ldots, n)$.

(ii) $\{\Omega_1, \ldots, \Omega_n\}$ *is* $\alpha-$*transversal at* $\bar{x}$ *with some* $\delta_1 > 0$ *and* $\delta_2 > 0$ *if and only if*

$$\bigcap_{i=1}^{n}(\Omega_i - \omega_i - a_i) \cap (\delta_1\mathbb{B}) \neq \emptyset \tag{2.9}$$

for all $\omega_i \in \Omega_i \cap B_{\delta_2}(\bar{x})$ *and* $a_i \in X$ *with* $\|a_i\| < \alpha\delta_1$ $(i = 1, \ldots, n)$.

Proof. (i) If $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$semitransversal at $\bar{x}$ with some $\delta > 0$, then, by Definition 2.3.9(i), for all $a_i \in X$ with $\|a_i\| < \alpha\delta$ $(i = 1, \ldots, n)$, and any number ρ satisfying $\alpha^{-1} \max_{1\le i\le n} \|a_i\| < \rho < \delta$, condition (P9) holds. The latter condition obviously implies (2.8).

Conversely, suppose that, for some $\delta > 0$, condition (2.8) is satisfied for all $a_i \in X$ with $\|a_i\| < \alpha\delta$ $(i = 1, \ldots, n)$. Let ρ be an arbitrary number in $]0, \delta[$ and let $a_i \in X$ with $\|a_i\| < \alpha\rho$ $(i = 1, \ldots, n)$. Set $t := \rho/\delta$ and $a_i' := a_i/t$ $(i = 1, \ldots, n)$. Then $0 < t < 1$ and $\|a_i'\| = \|a_i\|/t < \alpha\delta$ $(i = 1, \ldots, n)$, and consequently, there exists an $x' \in \cap_{i=1}^{n}(\Omega_i - a_i') \cap B_\delta(\bar{x})$, i.e. $x' \in B_\delta(\bar{x})$ and $x' = \omega_i - a_i'$ for some $\omega_i \in \Omega_i$, or equivalently, $a_i = t(\omega_i - x')$ $(i = 1, \ldots, n)$. In view of the convexity of the sets, we have $t\omega_i + (1-t)\bar{x} \in \Omega_i$ $(i = 1, \ldots, n)$. Set $x := \bar{x} + t(x' - \bar{x})$. We have $x = t\omega_i + (1-t)\bar{x} - t(\omega_i - x') \in \Omega_i - a_i$ $(i = 1, \ldots, n)$. Moreover, $\|x - \bar{x}\| = t\|x' - \bar{x}\| < \rho$. Hence, condition (P9) is satisfied. In view of Definition 2.3.9(i), $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$semitransversal at $\bar{x}$ with δ.

(ii) If $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$transversal at $\bar{x}$ with some $\delta_1 > 0$ and $\delta_2 > 0$, then, by Definition 2.3.9(ii), for any $\omega_i \in \Omega_i \cap B_{\delta_2}(\bar{x})$, $a_i \in X$ with $\|a_i\| < \alpha\delta_1$ $(i = 1, \ldots, n)$, and any number ρ satisfying $\alpha^{-1} \max_{1\le i\le n} \|a_i\| < \rho < \delta_1$, condition (P11) holds. The latter condition obviously implies (2.9).

Conversely, suppose that, for some $\delta_1 > 0$ and $\delta_2 > 0$, condition (2.9) is satisfied for all $\omega_i \in \Omega_i \cap B_{\delta_2}(\bar{x})$ and $a_i \in X$ with $\|a_i\| < \alpha\delta_1$ $(i = 1, \ldots, n)$. Then the collection of convex sets $\Omega_i - \omega_i$ $(i = 1, \ldots, n)$, considered near their common point 0, satisfies the conditions in part (i) and is consequently $\alpha-$semitransversal at 0 with δ_1 uniformly over $\omega_i \in \Omega_i \cap B_{\delta_2}(\bar{x})$ $(i = 1, \ldots, n)$. This means that $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$transversal at $\bar{x}$ with δ_1 and δ_2.

□

Employing the same arguments as in the proof of Proposition 2.3.17, it is easy to show that in the convex case the alternative representations of $\alpha-$transversality in Proposition 2.3.12 can also be simplified.

Proposition 2.3.18. *Let* $\Omega_1, \ldots, \Omega_n$ *be convex subsets of a normed vector space* X, $\bar{x} \in \cap_{i=1}^{n}\Omega_i$, $\alpha > 0$, $\delta_1 > 0$ *and* $\delta_2 > 0$. *Conditions* (i)–(iii) *in Proposition 2.3.12 are satisfied if and only if the following equivalent conditions hold:*

(i) condition (2.9) *is satisfied for all* $\omega_i \in \Omega_i$ *and* $a_i \in X$ *with* $\omega_i + a_i \in B_{\delta_2}(\bar{x})$ *and* $\|a_i\| < \alpha\delta_1$ $(i = 1, \ldots, n)$*;*

(ii) for all $a_i \in \delta_2\mathbb{B}$ *with* $d(\bar{x}, \Omega_i - a_i) < \alpha\delta_1$ $(i = 1, \ldots, n)$*, we have*

$$\bigcap_{i=1}^{n}(\Omega_i - a_i) \cap B_{\delta_1}(\bar{x}) \neq \emptyset;$$

(iii) for all $x \in X$ and $a_i \in X$ with $x + a_i \in B_{\delta_2}(\bar{x})$ and $d(x, \Omega_i - a_i) < \alpha\delta_1$ $(i = 1, \ldots, n)$, we have

$$\bigcap_{i=1}^{n} (\Omega_i - a_i) \cap B_{\delta_1}(x) \neq \emptyset.$$

When the sets are convex, the semitransversality and transversality properties are equivalent.

Proposition 2.3.19. *Let $\Omega_1, \ldots, \Omega_n$ be convex subsets of a normed vector space X, $\bar{x} \in \cap_{i=1}^{n} \Omega_i$ and $\alpha > 0$. If $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$semitransversal at $\bar{x}$ with some $\delta > 0$, then, for any $\alpha' \in]0, \alpha[$, it is $\alpha'-$transversal at $\bar{x}$ with $\delta_1 := \delta$ and $\delta_2 := (\alpha - \alpha')\delta$.*

As a consequence, $\{\Omega_1, \ldots, \Omega_n\}$ is semitransversal at $\bar{x}$ if and only if it is transversal at $\bar{x}$, and $\mathrm{s_e tr}[\Omega_1, \ldots, \Omega_n](\bar{x}) = \mathrm{tr}[\Omega_1, \ldots, \Omega_n](\bar{x})$.

Proof. Let $\{\Omega_1, \ldots, \Omega_n\}$ be $\alpha-$semitransversal at $\bar{x}$ with some $\delta > 0$, and let $\alpha' \in]0, \alpha[$, $\delta_1 := \delta$ and $\delta_2 := (\alpha - \alpha')\delta$. Let $\omega_i \in \Omega_i \cap B_{\delta_2}(\bar{x})$ and $a_i \in X$ with $\|a_i\| < \alpha'\delta$ $(i = 1, \ldots, n)$, set $a'_i := \omega_i + a_i - \bar{x}$ $(i = 1, \ldots, n)$. Then $\|a'_i\| \leq \|\omega_i - \bar{x}\| + \|a_i\| < \delta_2 + \alpha'\delta = \alpha\delta$ $(i = 1, \ldots, n)$, and by Proposition 2.3.18(i), $\cap_{i=1}^{n} (\Omega_i - a'_i) \cap B_\delta(\bar{x}) \neq \emptyset$, which is equivalent to condition (2.8). By Proposition 2.3.18(ii), $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha'-$transversal at $\bar{x}$ with δ_1 and δ_2. Since α' can be chosen arbitrarily close to α, we have $\mathrm{s_e tr}[\Omega_1, \ldots, \Omega_n](\bar{x}) \leq \mathrm{tr}[\Omega_1, \ldots, \Omega_n](\bar{x})$. In view of Proposition 2.3.14, this inequality actually holds as equality. □

Remark 2.3.20. (i) Proposition 2.3.19 strengthens [90, Proposition 13(iv)].

(ii) In view of Propositions 2.3.19 and 2.3.14, in the convex case, $\alpha-$semitransversality is in general stronger than $\alpha-$subtransversality.

2.4 Metric Characterisations

The transversality properties of collections of sets in Definition 2.3.9 admit equivalent characterisations in metric terms.

Proposition 2.4.1. *Let $\Omega_1, \ldots, \Omega_n$ be subsets of a normed vector space X, $\bar{x} \in \cap_{i=1}^{n} \Omega_i$ and $\alpha > 0$.*

(i) $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$semitransversal at $\bar{x}$ with some $\delta > 0$ if and only if

$$\alpha d\left(\bar{x}, \bigcap_{i=1}^{n} (\Omega_i - a_i)\right) \leq \max_{1 \leq i \leq n} \|a_i\| \tag{M9}$$

for all $a_i \in X$ $(i = 1, \ldots, n)$ with $\max\limits_{1 \leq i \leq n} \|a_i\| < \alpha\delta$.

(ii) $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$subtransversal at $\bar{x}$ with some $\delta_1 > 0$ and $\delta_2 > 0$ if and only if

$$\alpha d\left(x, \bigcap_{i=1}^{n} \Omega_i\right) \leq \max_{1 \leq i \leq n} d(x, \Omega_i) \tag{M10}$$

for all $x \in B_{\delta_2}(\bar{x})$ with $\max\limits_{1 \leq i \leq n} d(x, \Omega_i) < \alpha\delta_1$.

(iii) $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$transversal at $\bar{x}$ with some $\delta_1 > 0$ and $\delta_2 > 0$ if and only if

$$\alpha d\left(0, \bigcap_{i=1}^{n} (\Omega_i - \omega_i - a_i)\right) \leq \max_{1 \leq i \leq n} \|a_i\| \tag{M11}$$

for all $\omega_i \in \Omega_i \cap B_{\delta_2}(\bar{x})$ and $a_i \in X$ $(i = 1, \ldots, n)$ with $\max\limits_{1 \leq i \leq n} \|a_i\| < \alpha\delta_1$.

Proof. (i) By Definition 2.3.9 (i), $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$semitransversal at $\bar{x}$ with some $\delta > 0$ if and only if $d(\bar{x}, \cap_{i=1}^n (\Omega_i - a_i)) < \rho$ for all $a_i \in X$ $(i = 1, \ldots, n)$ with $\max\limits_{1\le i\le n} \|a_i\| < \alpha\delta$ and all numbers $\rho > \alpha^{-1} \max\limits_{1\le i\le n} \|a_i\|$. The conclusion follows.

(ii) By Definition 2.3.9 (ii), $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$subtransversal at $\bar{x}$ with some $\delta_1 > 0$ and $\delta_2 > 0$ if and only if $d(x, \cap_{i=1}^n \Omega_i) < \rho$ for all $x \in B_{\delta_2}(\bar{x})$ with $\max\limits_{1\le i\le n} d(x, \Omega_i) < \alpha\delta_1$ and all numbers $\rho > \alpha^{-1} \max\limits_{1\le i\le n} d(x, \Omega_i)$. The conclusion follows.

(iii) By the Definition 2.3.9 (iii), $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$transversal at $\bar{x}$ with some $\delta_1 > 0$ and $\delta_2 > 0$ if and only if $d(0, \cap_{i=1}^n (\Omega_i - \omega_i - a_i)) < \rho$ for all $\omega_i \in \Omega_i \cap B_{\delta_2}(\bar{x})$ and $a_i \in X$ $(i = 1, \ldots, n)$ with $\max\limits_{1\le i\le n} \|a_i\| < \alpha\delta_1$, and all numbers $\rho > \alpha^{-1} \max\limits_{1\le i\le n} \|a_i\|$. The conclusion follows.

□

The alternative metric characterisations of $\alpha-$transversality in the next proposition correspond to the respective properties in Proposition 2.3.12.

Proposition 2.4.2. *Let $\Omega_1, \ldots, \Omega_n$ be subsets of a normed vector space X, $\bar{x} \in \cap_{i=1}^n \Omega_i$, and $\alpha > 0$. Then, $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$transversal at $\bar{x}$ if and only if there exist numbers $\delta_1 > 0$ and $\delta_2 > 0$ such that the following equivalent conditions hold:*

(i) inequality (M11) *is satisfied for all $\omega_i \in \Omega_i$ and $a_i \in X$ $(i = 1, \ldots, n)$ with $\omega_i + a_i \in B_{\delta_2}(\bar{x})$ and $\max\limits_{1\le i\le n} \|a_i\| < \alpha\delta_1$;*

(ii) for all $a_i \in \delta_2\mathbb{B}$ $(i = 1, \ldots, n)$ with $\max\limits_{1\le i\le n} d(\bar{x}, \Omega_i - a_i) < \alpha\delta_1$, we have

$$\alpha d\left(\bar{x}, \bigcap_{i=1}^n (\Omega_i - a_i)\right) \le \max_{1\le i\le n} d(\bar{x}, \Omega_i - a_i); \tag{M11.1}$$

(iii) for all $x \in X$ and $a_i \in X$ with $x + a_i \in B_{\delta_2}(\bar{x})$ and $d(x, \Omega_i - a_i) < \alpha\delta_1$ $(i = 1, \ldots, n)$, we have

$$\alpha d\left(x, \bigcap_{i=1}^n (\Omega_i - a_i)\right) \le \max_{1\le i\le n} d(x, \Omega_i - a_i). \tag{M11.2}$$

Moreover, if $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$transversal at $\bar{x}$ with some $\delta_1 > 0$ and $\delta_2 > 0$, then conditions (i)–(iii) *hold with any $\delta_1' \in]0, \delta_1]$ and $\delta_2' > 0$ satisfying $\delta_2' + \alpha\delta_1' \le \delta_2$ in place of δ_1 and δ_2.*

Conversely, if conditions (i)–(iii) *hold with some $\delta_1 > 0$ and $\delta_2 > 0$, then $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$transversal at $\bar{x}$ with any $\delta_1' \in]0, \delta_1]$ and $\delta_2' > 0$ satisfying $\delta_2' + \alpha\delta_1' \le \delta_2$ in place of δ_1 and δ_2.*

Proof. The statement is a consequence of Proposition 2.3.12. It suffices to notice that inequality (M11) is equivalent to condition (P8) satisfied for any $\rho > \alpha^{-1} \max\limits_{1\le i\le n} \|a_i\|$, inequality (M11.1) is equivalent to condition (P5) satisfied for any $\rho > \alpha^{-1} \max\limits_{1\le i\le n} d(\bar{x}, \Omega_i - a_i)$, and inequality (M11.2) is equivalent to condition (P9) satisfied for any positive number ρ with $\rho > \alpha^{-1} \max\limits_{1\le i\le n} d(x, \Omega_i - a_i)$. □

Remark 2.4.3. Condition (M11.2) served as the main metric characterisation of transversality in [90,91] and subsequent publications. Condition (M11.1) has been picked up recently in [30]. This condition seems an important advancement as it replaces an arbitrary point x with the given reference point $\bar{x}$. Condition (M11) seems new. In accordance with Proposition 2.3.12, it is the most straightforward metric counterpart of the original geometric property (P8).

In the convex case, the estimates in parts (i) and (iii) of Proposition 2.4.1 as well as the alternative metric characterisations of $\alpha-$transversality in Proposition 2.4.2 can be simplified. The next two statements are direct consequences of Propositions 2.4.2 and 2.3.18, respectively.

Proposition 2.4.4. *Let $\Omega_1, \ldots, \Omega_n$ be convex subsets of a normed vector space X, $\bar{x} \in \cap_{i=1}^n \Omega_i$ and $\alpha > 0$.*

(i) $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$semitransversal at $\bar{x}$ with some $\delta > 0$ if and only if

$$d\left(\bar{x}, \bigcap_{i=1}^{n} (\Omega_i - a_i)\right) < \delta \tag{2.10}$$

for all $a_i \in X$ with $\|a_i\| < \alpha\delta$ $(i = 1, \ldots, n)$.

(ii) $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$transversal at $\bar{x}$ with some $\delta_1 > 0$ and $\delta_2 > 0$ if and only if

$$d\left(0, \bigcap_{i=1}^{n} (\Omega_i - \omega_i - a_i)\right) < \delta_1 \tag{2.11}$$

for all $\omega_i \in \Omega_i \cap B_{\delta_2}(\bar{x})$ and $a_i \in X$ with $\|a_i\| < \alpha\delta_1$ $(i = 1, \ldots, n)$.

Corollary 2.4.5. *Let $\Omega_1, \ldots, \Omega_n$ be convex subsets of a normed vector space X, $\bar{x} \in \cap_{i=1}^n \Omega_i$, $\alpha > 0$, $\delta_1 > 0$ and $\delta_2 > 0$. Conditions* (i)–(iii) *in Proposition 2.3.18 are satisfied if and only if the following equivalent conditions hold:*

(i) *condition (2.11) is satisfied for all $\omega_i \in \Omega_i$ and $a_i \in X$ with $\omega_i + a_i \in B_{\delta_2}(\bar{x})$ and $\|a_i\| < \alpha\delta_1$ $(i = 1, \ldots, n)$;*

(ii) *for all $a_i \in \delta_2\mathbb{B}$ $(i = 1, \ldots, n)$ with $\max\limits_{1\le i\le n} d(\bar{x}, \Omega_i - a_i) < \alpha\delta_1$, we have*

$$\alpha d\left(\bar{x}, \bigcap_{i=1}^{n} (\Omega_i - a_i)\right) < \delta_1; \tag{2.12}$$

(iii) *for all $x \in X$ and $a_i \in X$ with $x + a_i \in B_{\delta_2}(\bar{x})$ and $d(x, \Omega_i - a_i) < \alpha\delta_1$ $(i = 1, \ldots, n)$, we have*

$$\alpha d\left(x, \bigcap_{i=1}^{n} (\Omega_i - a_i)\right) < \delta_1. \tag{2.13}$$

Moreover, if $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$transversal at $\bar{x}$ with some $\delta_1 > 0$ and $\delta_2 > 0$ then conditions (i)–(iii) *hold with any $\delta_1' \in]0, \delta_1]$ and $\delta_2' > 0$ satisfying $\delta_2' + \alpha\delta_1' \le \delta_2$ in place of δ_1 and δ_2.*

Conversely, if conditions (i)–(iii) *hold with some $\delta_1 > 0$ and $\delta_2 > 0$, then $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$transversal at $\bar{x}$ with any $\delta_1' \in]0, \delta_1]$ and $\delta_2' > 0$ satisfying $\delta_2' + \alpha\delta_1' \le \delta_2$ in place of δ_1 and δ_2.*

As discussed in Section 2.3.2, the absence of semitransversality (transversality) implies stationarity (approximate stationarity). Based on this observation, Proposition 2.4.1 gives us metric characterisations of stationarity and approximate stationarity.

Corollary 2.4.6. *Let $\Omega_1, \ldots, \Omega_n$ be subsets of a normed vector space X, $\bar{x} \in \cap_{i=1}^n \Omega_i$.*

(i) $\{\Omega_1, \ldots, \Omega_n\}$ is stationary at $\bar{x}$ if and only if for any $\varepsilon > 0$, there is an $\rho \in]0, \varepsilon[$ and $a_i \in X$ $(i = 1, \ldots, n)$ with $\max\limits_{1\le i\le n} \|a_i\| < \varepsilon\rho$ such that

$$\varepsilon d\left(\bar{x}, \bigcap_{i=1}^{n} (\Omega_i - a_i)\right) \le \max_{1\le i\le n} \|a_i\|. \tag{M3}$$

(ii) $\{\Omega_1, \ldots, \Omega_n\}$ *is approximately stationary at* $\bar{x}$ *if and only if for any* $\varepsilon > 0$*, there are an* $\rho \in]0, \varepsilon[$*,* $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ *and* $a_i \in X$ $(i = 1, \ldots, n)$ *with* $\max_{1\le i\le n} \|a_i\| < \varepsilon\rho$

$$\varepsilon d\left(0, \bigcap_{i=1}^{n}(\Omega_i - \omega_i - a_i)\right) \le \max_{1\le i\le n} \|a_i\|. \tag{M4}$$

2.5 Regularity of Set-Valued Mappings

In this section, we clarify quantitative relationships between transversality properties of collections of sets and the corresponding regularity properties of set-valued mappings.

The next definition is a modification of [106, Definition 5.1].

Definition 2.5.1. Let $F : X \rightrightarrows Y$ be a set-valued mapping between metric spaces, $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ and $\alpha > 0$.

(i) F is $\alpha-$semiregular at $(\bar{x}, \bar{y})$ if there exists a number $\delta > 0$ such that

$$\alpha d(\bar{x}, F^{-1}(y)) \le d(y, \bar{y})$$

for all $y \in B_{\alpha\delta}(\bar{y})$.

(ii) F is $\alpha-$subregular at $(\bar{x}, \bar{y})$ if there exist numbers $\delta_1 > 0$ and $\delta_2 > 0$ such that

$$\alpha d(x, F^{-1}(\bar{y})) \le d(\bar{y}, F(x))$$

for all $x \in B_{\delta_2}(\bar{x})$ with $d(\bar{y}, F(x)) < \alpha\delta_1$.

(iii) F is $\alpha-$regular at $(\bar{x}, \bar{y})$ if there exist numbers $\delta_1 > 0$ and $\delta_2 > 0$ such that

$$\alpha d(x, F^{-1}(y)) \le d(y, F(x)) \tag{2.14}$$

for all $x \in X$ and $y \in Y$ with $d(x, \bar{x}) + d(y, \bar{y}) < \delta_2$ with $d(y, F(x)) < \alpha\delta_1$.

The number $\alpha > 0$ in each of the properties in Definition 2.5.1 plays the role of a rate of the respective property. The exact upper bound of all $\alpha > 0$ such that the property holds with some $\delta > 0$, or $\delta_1 > 0$ and $\delta_2 > 0$, is called the *modulus* of this property. We use notations $\mathrm{s_e rg}\,[F](\bar{x}, \bar{y})$, $\mathrm{srg}\,[F](\bar{x}, \bar{y})$ and $\mathrm{rg}\,[F](\bar{x}, \bar{y})$ for the moduli of the respective properties. If the property does not hold, then by convention the respective modulus equals 0.

The last two regularity properties in Definition 2.5.1 have been very well studied for decades due to their numerous important applications; see, e.g., monographs [51, 75, 82, 120]. Note the notations $\mathrm{subreg}\,(F; \bar{x} \mid \bar{y})$ and $\mathrm{reg}\,(F; \bar{x} \mid \bar{y})$ often used for the respective moduli (cf., e.g., [51]), as well as an important difference in their definitions which is reflected in the relations:

$$\mathrm{srg}\,[F](\bar{x}, \bar{y}) = \frac{1}{\mathrm{subreg}\,(F; \bar{x} \mid \bar{y})}, \quad \mathrm{rg}\,[F](\bar{x}, \bar{y}) = \frac{1}{\mathrm{reg}\,(F; \bar{x} \mid \bar{y})}.$$

(In fact, $\mathrm{rg}\,[F](\bar{x}, \bar{y})$ coincides with the *modulus of surjection* [75].) Unlike its more famous siblings, the first property in Definition 2.5.1 has only recently started attracting attention of researchers; see [36, 92].

As in the case of Definition 2.3.9 of the transversality properties, the role of the δ's in the above definitions is more technical: they control the size of the neighbourhoods involved in the respective definitions. Of course, if a property in part (ii) or (iii) is satisfied with some $\delta_1 > 0$ and $\delta_2 > 0$, it is satisfied also with the single $\delta := \min\{\delta_1, \delta_2\}$ in place of both δ_1 and δ_2. Moreover, in this case (or more generally, when $\delta_1 \geq \delta_2$) the inequality $d(\bar{y}, F(x)) < \alpha\delta_1$ in part (ii) and the inequality $d(y, F(x)) < \alpha\delta_1$ in part (iii) can be dropped. We use here two different parameters to emphasise their different roles in the definitions and the corresponding characterisations, and expose connections with the transversality properties in Definition 2.3.9.

It is immediate from the definition that, if a mapping is $\alpha-$regular at a point in its graph, it is also $\alpha-$semiregular and $\alpha-$subregular at this point; hence, $\mathrm{s_e rg}\,[F](\bar{x}, \bar{y}) \geq \mathrm{rg}\,[F](\bar{x}, \bar{y})$ and $\mathrm{srg}\,[F](\bar{x}, \bar{y}) \geq \mathrm{rg}\,[F](\bar{x}, \bar{y})$.

Note the combined inequality $d(x, \bar{x}) + d(y, \bar{y}) < \delta_2$ employed in (2.14) instead of the more traditional separate conditions $x \in B_{\delta_2}(\bar{x})$ and $y \in B_{\delta_2}(\bar{y})$. This replacement does not affect the property of metric $\alpha-$regularity itself, but can have an effect on the value of δ_2 which ensures the property. Employing this inequality in (2.14) is convenient for establishing the relationship between the regularity and transversality properties.

Now we return to our main setting of a collection of $n \geq 2$ subsets $\Omega_1, \ldots, \Omega_n$ of a normed vector space X, having a common point $\bar{x} \in \cap_{i=1}^n \Omega_i$, and consider a set-valued mapping $F : X \rightrightarrows X^n$ given by

$$F(x) := (\Omega_1 - x) \times \ldots \times (\Omega_n - x), \quad x \in X. \tag{2.15}$$

which is going to play the key role in establishing relationships between the regularity and transversality properties. It was most likely first used by Ioffe in [70]. Observe that $\bar{y} := (0, \ldots, 0) \in F(\bar{x})$ and

$$F^{-1}(x_1, \ldots, x_n) = (\Omega_1 - x_1) \cap \ldots \cap (\Omega_n - x_n), \quad x_1, \ldots, x_n \in X.$$

Let the space $Y := X^n$ be equipped with the maximum norm.

The next proposition is a reformulation of Proposition 2.4.1.

Proposition 2.5.2. *Let $\Omega_1, \ldots, \Omega_n$ be subsets of a normed vector space X, $\bar{x} \in \cap_{i=1}^n \Omega_i$, $\alpha > 0$, $F : X \rightrightarrows X^n$ be defined by (2.15), and $\bar{y} := (0, \ldots, 0) \in X^n$.*

(i) $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$semitransversal at $\bar{x}$ with some $\delta > 0$ if and only if F is $\alpha-$semiregular at $(\bar{x}, \bar{y})$ with the same δ.

(ii) $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$subtransversal at $\bar{x}$ with some $\delta_1 > 0$ and $\delta_2 > 0$ if and only if F is $\alpha-$semiregular at $(\bar{x}, \bar{y})$ with the same δ_1 and δ_2.

(iii) $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$transversal at $\bar{x}$ with some $\delta_1 > 0$ and $\delta_2 > 0$ if and only if

$$\alpha d\left(0, F^{-1}(\omega_1 + a_1, \ldots, \omega_n + a_n)\right) \leq \|y\| \tag{2.16}$$

for all $\omega_i \in \Omega_i \cap B_{\delta_2}(\bar{x})$ $(i = 1, \ldots, n)$ and $y := (a_1, \ldots, a_n) \in X^n$ with $\|y\| < \alpha\delta_1$.

Thanks to parts (i) and (ii) of Proposition 2.5.2, we have the exact equivalences between the $\alpha-$semitransversality and $\alpha-$subtransversality properties of $\{\Omega_1, \ldots, \Omega_n\}$ on one hand, and the respective $\alpha-$semiregularity and $\alpha-$subregularity properties of F on the other hand. However, we do not seem to have the exact equivalence between the remaining two properties, at least quantitatively, as condition (2.16) is not exactly of the form (2.14). Fortunately, Proposition 2.5.2 resolves the issue. Here is its reformulation in terms of F. It shows that the conventional $\alpha-$regularity property in Definition 2.5.1(iii) is not a direct counterpart of the conventional $\alpha-$transversality property in Definition 2.3.9(iii), but rather of its alternative representation in Proposition 2.3.12(iii).

Proposition 2.5.3. *Let $\Omega_1, \ldots, \Omega_n$ be subsets of a normed vector space X, $\bar{x} \in \cap_{i=1}^n \Omega_i$, $\alpha > 0$, $F : X \rightrightarrows X^n$ be defined by (2.15), and $\bar{y} := (0, \ldots, 0) \in X^n$. Then, $\{\Omega_1, \ldots, \Omega_n\}$ is $\alpha-$transversal at $\bar{x}$ if and only if there exist numbers $\delta_1 > 0$ and $\delta_2 > 0$ such that the following equivalent conditions hold:*

(i) inequality (2.16) is satisfied for all $\omega_i \in \Omega_i$ $(i = 1, \ldots, n)$ and $y := (a_1, \ldots, a_n) \in X^n$ with $\omega_i + a_i \in B_{\delta_2}(\bar{x})$ $(i = 1, \ldots, n)$ and $\|y\| < \alpha\delta_1$;

(ii) for all $y \in \delta_2\mathbb{B}$ *with* $d(y, F(\bar{x})) < \alpha\delta_1$*, we have*

$$\alpha d\left(\bar{x}, F^{-1}(y)\right) \le d(y, F(\bar{x}));$$

(iii) inequality (2.14) *is satisfied for all* $x \in X$ *and* $y := (a_1, \ldots, a_n) \in X^n$ *with* $x + a_i \in B_{\delta_2}(\bar{x})$ $(i = 1, \ldots, n)$ *and* $d(y, F(x)) < \alpha\delta_1$*, i.e.* F *is* $\alpha-$*regular at* $(\bar{x}, \bar{y})$ *with the same* δ_1 *and* δ_2*.*

Moreover, if $\{\Omega_1, \ldots, \Omega_n\}$ *is* $\alpha-$*transversal at* $\bar{x}$ *with some* $\delta_1 > 0$ *and* $\delta_2 > 0$*, then conditions* (i)–(iii) *hold with any* $\delta_1' \in]0, \delta_1]$ *and* $\delta_2' > 0$ *satisfying* $\delta_2' + \alpha\delta_1' \le \delta_2$ *in place of* δ_1 *and* δ_2*.*

Conversely, if conditions (i)–(iii) *hold with some* $\delta_1 > 0$ *and* $\delta_2 > 0$*, then* $\{\Omega_1, \ldots, \Omega_n\}$ *is* $\alpha-$*transversal at* $\bar{x}$ *with any* $\delta_1' \in]0, \delta_1]$ *and* $\delta_2' > 0$ *satisfying* $\delta_2' + \alpha\delta_1' \le \delta_2$ *in place of* δ_1 *and* δ_2*.*

Thanks to Proposition 2.5.3, we have equivalence between the $\alpha-$transversality property of $\{\Omega_1, \ldots, \Omega_n\}$ and the $\alpha-$regularity property of F, although not necessarily with the same δ_1 and δ_2. The three conditions in Proposition 2.3.12 together with condition (iii) in Proposition 2.5.3 provide a series of metric characterisations of both equivalent properties in terms of the set-valued mapping F. Observe also that, thanks to Proposition 2.3.12, the set-valued mapping F given by (2.15) provides an important case when the point x in the inequality (2.14) defining metric $\alpha-$regularity can be replaced by the fixed reference point $\bar{x}$.

The next corollary of Propositions 2.5.3 and 2.5.2 collects 'δ-free' versions of the discussed equivalences. It recaptures [106, Proposition 5.1].

Corollary 2.5.4. *Let* $\Omega_1, \ldots, \Omega_n$ *be subsets of a normed vector space* X*,* $\bar{x} \in \cap_{i=1}^n \Omega_i$*,* $\alpha > 0$*,* $F : X \rightrightarrows X^n$ *be defined by* (2.15)*, and* $\bar{y} := (0, \ldots, 0) \in X^n$*.*

(i) $\{\Omega_1, \ldots, \Omega_n\}$ *is* $\alpha-$*semitransversal at* $\bar{x}$ *if and only if* F *is* $\alpha-$*semiregular at* $(\bar{x}, \bar{y})$*. Hence,* $\mathrm{s_e tr}\,[\Omega_1, \ldots, \Omega_n](\bar{x}) = \mathrm{s_e rg}\,[F](\bar{x}, \bar{y})$*.*

(ii) $\{\Omega_1, \ldots, \Omega_n\}$ *is* $\alpha-$*subtransversal at* $\bar{x}$ *if and only if* F *is* $\alpha-$*subregular at* $(\bar{x}, \bar{y})$*.*
Hence, $\mathrm{str}\,[\Omega_1, \ldots, \Omega_n](\bar{x}) = \mathrm{srg}\,[F](\bar{x}, \bar{y})$*.*

(iii) $\{\Omega_1, \ldots, \Omega_n\}$ *is* $\alpha-$*transversal at* $\bar{x}$ *if and only if* F *is* $\alpha-$*regular at* $(\bar{x}, \bar{y})$*. Hence,* $\mathrm{tr}\,[\Omega_1, \ldots, \Omega_n](\bar{x}) = \mathrm{rg}\,[F](\bar{x}, \bar{y})$*.*

In view of Proposition 2.4.4, it follows from the above corollary that, when the sets $\Omega_1, \ldots, \Omega_n$ are convex, the semitransversality and transversality properties for the mapping F defined by (2.15) are equivalent, and $\mathrm{s_e rg}\,[F](\bar{x}, \bar{y}) = \mathrm{rg}\,[F](\bar{x}, \bar{y})$.

In view of Propositions 2.5.3 and 2.5.2, the $\alpha-$transversality properties of collections of sets can be viewed as particular cases of the corresponding $\alpha-$regularity properties of set-valued mappings. Now we are going to show that the two popular models are in a sense equivalent.

Given an arbitrary set-valued mapping $F : X \rightrightarrows Y$ between metric spaces and a point $(\bar{x}, \bar{y}) \in \mathrm{gph}\, F$, we can construct the two sets:

$$\Omega_1 := \mathrm{gph}\, F, \quad \Omega_2 := X \times \{\bar{y}\} \tag{2.17}$$

in the product space $X \times Y$. To establish the relationship between the regularity properties of F and the transversality properties of the collection of sets (2.17), we have to assume that X and Y are normed vector spaces.

The next proposition translates the metric characterisations of the transversality properties of collection of sets in Proposition 2.5.2 into certain metric properties of the set-valued mapping F.

Proposition 2.5.5. *Let X and Y be normed vector spaces, $F : X \rightrightarrows Y$, $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ and $\alpha > 0$. Let Ω_1 and Ω_2 be defined by (2.17).*

(i) If $\{\Omega_1, \Omega_2\}$ is $\alpha-$semitransversal at $(\bar{x}, \bar{y})$ with some $\delta > 0$, then

$$\alpha d\left(\bar{x} + u, F^{-1}(\bar{y} + v)\right) \leq \max\left\{\|u\|, \frac{1}{2}\|v\|\right\} \tag{2.18}$$

for all $u \in (\alpha\delta)\mathbb{B}_X$ and $v \in (2\alpha\delta)\mathbb{B}_Y$.

Moreover, if $\alpha \leq 1$, then both conditions are equivalent.

(ii) If $\{\Omega_1, \Omega_2\}$ is $\alpha-$subtransversal at $(\bar{x}, \bar{y})$ with some $\delta_1 > 0$ and $\delta_2 > 0$, then

$$\alpha d\left(x, F^{-1}(\bar{y})\right) \leq \max\left\{d((x, y), \operatorname{gph} F), \|y - \bar{y}\|\right\} \tag{2.19}$$

for all $x \in B_{\delta_2}(\bar{x})$ and $y \in B_{\min\{\alpha\delta_1, \delta_2\}}(\bar{y})$ with $d((x, y), \operatorname{gph} F) < \alpha\delta_1$.

Moreover, if $\alpha \leq 1$, then both conditions are equivalent.

(iii) If $\{\Omega_1, \Omega_2\}$ is $\alpha-$transversal at $(\bar{x}, \bar{y})$ with some $\delta_1 > 0$ and $\delta_2 > 0$, then

$$\alpha d\left(x + u, F^{-1}(y + v)\right) \leq \max\left\{\|u\|, \frac{1}{2}\|v\|\right\} \tag{2.20}$$

for all $(x, y) \in \operatorname{gph} F \cap B_{\delta_2}(\bar{x}, \bar{y})$, $u \in (\alpha\delta_1)\mathbb{B}_X$ and $v \in (2\alpha\delta_1)\mathbb{B}_Y$.

Moreover, if $\alpha \leq 1$, then both conditions are equivalent.

Proof. First observe from (2.17) that

$$\Omega_1 \cap \Omega_2 = F^{-1}(\bar{y}) \times \{\bar{y}\}, \tag{2.21}$$

and consequently, $(\bar{x}, \bar{y}) \in \Omega_1 \cap \Omega_2$.

(i) Given $a_1 := (u_1, v_1)$ and $a_2 := (u_2, v_2) \in X \times Y$, we have

$$\begin{aligned} (\Omega_1 - a_1) \cap (\Omega_2 - a_2) &= \left(F^{-1}(\bar{y} + v_1 - v_2) - u_1\right) \times \{\bar{y} - v_2\}, \\ d\left((\bar{x}, \bar{y}), (\Omega_1 - a_1) \cap (\Omega_2 - a_2)\right) &= \max\left\{d\left(\bar{x} + u_1, F^{-1}(\bar{y} + v_1 - v_2)\right), \|v_2\|\right\}. \end{aligned} \tag{2.22}$$

Thus, inequality

$$\alpha d\left((\bar{x}, \bar{y}), (\Omega_1 - a_1) \cap (\Omega_2 - a_2)\right) \leq \max\{\|a_1\|, \|a_2\|\} \tag{2.23}$$

implies

$$\alpha d\left(\bar{x} + u_1, F^{-1}(\bar{y} + v_1 - v_2)\right) \leq \max\{\|u_1\|, \|u_2\|, \|v_1\|, \|v_2\|\}, \tag{2.24}$$

and the converse implication is true if $\alpha \leq 1$.

We claim that the following conditions are equivalent:

(a) condition (2.24) holds for all $u_1, u_2 \in (\alpha\delta)\mathbb{B}_X$ and $v_1, v_2 \in (\alpha\delta)\mathbb{B}_Y$;

(b)

$$\alpha d\left(\bar{x} + u, F^{-1}(\bar{y} + v_1 - v_2)\right) \leq \max\{\|u\|, \|v_1\|, \|v_2\|\} \tag{2.25}$$

for all $u \in (\alpha\delta)\mathbb{B}_X$ and $v_1, v_2 \in (\alpha\delta)\mathbb{B}_Y$;

(c) condition (2.18) holds for all $u \in (\alpha\delta)\mathbb{B}_X$ and $v \in (2\alpha\delta)\mathbb{B}_Y$.

(a) $\Rightarrow$ (b). Given a $u \in (\alpha\delta)\mathbb{B}_X$ and $v_1, v_2 \in (\alpha\delta)\mathbb{B}_Y$, condition (2.25) is a consequence of (2.24) with $u_1 := u$ and $u_2 := 0$.

(b) $\Rightarrow$ (a). Given $u_1, u_2 \in (\alpha\delta)\mathbb{B}_X$ and $v_1, v_2 \in (\alpha\delta)\mathbb{B}_Y$, condition (2.24) is a consequence of (2.25) with $u := u_1$:

$$\begin{aligned}\alpha d\left(\bar{x}+u_1, F^{-1}(\bar{y}+v_1-v_2)\right) &\le \max\{\|u_1\|, \|v_1\|, \|v_2\|\} \\ &\le \max\{\|u_1\|, \|u_2\|, \|v_1\|, \|v_2\|\}.\end{aligned}$$

(b) $\Rightarrow$ (c). Given a $u \in (\alpha\delta)\mathbb{B}_X$ and a $v \in (2\alpha\delta)\mathbb{B}_Y$, condition (2.18) is a consequence of (2.25) with $v_1 := v/2$ and $v_2 := -v/2$.

(c) $\Rightarrow$ (b). Given a $u \in (\alpha\delta)\mathbb{B}_X$ and $v_1, v_2 \in (\alpha\delta)\mathbb{B}_X$, condition (2.25) is a consequence of (2.18) with $v := v_1 - v_2$:

$$\begin{aligned}\alpha d\left(\bar{x}+u, F^{-1}(\bar{y}+v_1-v_2)\right) &\le \max\left\{\|u\|, \frac{1}{2}\|v_1-v_2\|\right\} \\ &\le \max\left\{\|u\|, \frac{1}{2}(\|v_1\|+\|v_2\|)\right\} \\ &\le \max\{\|u\|, \|v_1\|, \|v_2\|\}.\end{aligned}$$

Hence, (a) $\Leftrightarrow$ (c), which, in view of Proposition 2.4.1(i), completes the proof.

(ii) Given an $(x, y) \in X \times Y$, thanks to (2.17) and (2.21), we have

$$d((x,y), \Omega_2) = \|y-\bar{y}\|, \quad d((x,y), \Omega_1 \cap \Omega_2) = \max\left\{d(x, F^{-1}(\bar{y})), \|y-\bar{y}\|\right\}.$$

Thus, inequality

$$\alpha d\left((x,y), \Omega_1 \cap \Omega_2\right) \le \max\left\{d((x,y), \Omega_1), d((x,y), \Omega_2)\right\}$$

implies condition (2.19) and the converse implication is true if $\alpha \le 1$. In view of Proposition 2.4.1(ii), this proves the assertion.

(iii) Given $a_1 := (u_1, v_1)$, $a_2 := (u_2, v_2)$, $(x, y) \in X \times Y$ and $x_2 \in X$, we have

$$\begin{aligned}&(\Omega_1 - (x,y) - a_1) \cap (\Omega_2 - (x_2, \bar{y}) - a_2) \\ &\qquad = \left(F^{-1}(y+v_1-v_2) - x - u_1\right) \times \{-v_2\}, \\ &d\left((0,0), (\Omega_1 - (x,y) - a_1) \cap (\Omega_2 - (x_2,\bar{y}) - a_2)\right) \\ &\qquad = \max\left\{d\left(x+u_1, F^{-1}(y+v_1-v_2)\right), \|v_2\|\right\}.\end{aligned}$$

Thus, inequality

$$\alpha d\left((0,0), (\Omega_1 - (x,y) - a_1) \cap (\Omega_2 - (x_2,\bar{y}) - a_2)\right) \le \max\{\|a_1\|, \|a_2\|\} \tag{2.26}$$

implies

$$\alpha d\left(x+u_1, F^{-1}(y+v_1-v_2)\right) \le \max\{\|u_1\|, \|u_2\|, \|v_1\|, \|v_2\|\}, \tag{2.27}$$

and the converse implication is true if $\alpha \le 1$. The same arguments as in the proof of (i) show that the last inequality holds for all $(x, y) \in \Omega_1 \cap B_{\delta_2}(\bar{x}, \bar{y})$, $u_1, u_2 \in (\alpha\delta_1)\mathbb{B}_X$ and $v_1, v_2 \in (\alpha\delta_1)\mathbb{B}_Y$ if and only if inequality (2.20) holds for all $(x, y) \in \operatorname{gph} F \cap B_{\delta_2}(\bar{x}, \bar{y})$, $u \in (\alpha\delta_1)\mathbb{B}_X$ and $v \in (2\alpha\delta_1)\mathbb{B}_Y$. In view of Proposition 2.4.1(iii), this proves the assertion.

□

Employing the estimates established in the proof of Proposition 2.5.3, we can also translate the metric characterisations of the $\alpha-$transversality in Proposition 2.4.2 into corresponding properties of the set-valued mapping F.

Proposition 2.5.6. *Let X and Y be normed vector spaces, $F : X \rightrightarrows Y$, $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ and $\alpha > 0$. Let Ω_1 and Ω_2 be defined by* (2.17)*. If $\{\Omega_1, \Omega_2\}$ is $\alpha-$transversal at $\bar{x}$, then there exist numbers $\delta_1 > 0$ and $\delta_2 > 0$ such that the following equivalent conditions hold:*

(i) for all $(x, y) \in \operatorname{gph} F$, $(u, v_1) \in (\alpha\delta_1)\mathbb{B}$ and $v_2 \in \min\{\alpha\delta_1, \delta_2\}\mathbb{B}$ with $x+u \in B_{\delta_2}(\bar{x})$, $y+v_1 \in B_{\delta_2}(\bar{y})$, we have

$$\alpha d\left(x + u, F^{-1}(y + v_1 - v_2)\right) \leq \max\{\|u\|, \|v_1\|, \|v_2\|\}; \tag{2.28}$$

(ii) for all $(u, v_1) \in \delta_2\mathbb{B}$ and $v_2 \in \min\{\alpha\delta_1, \delta_2\}\mathbb{B}$ with $d((\bar{x}, \bar{y}) + (u, v_1), \operatorname{gph} F) < \alpha\delta_1$, condition (2.25) *is satisfied;*

(iii) for all (x, y), $(u, v_1) \in X \times Y$ and $v_2 \in Y$ with $(x, y) + (u, v_1) \in B_{\delta_2}(\bar{x}, \bar{y})$, $d((x, y) + (u, v_1), \operatorname{gph} F) < \alpha\delta_1$ and $\|y + v_2 - \bar{y}\| < \min\{\alpha\delta_1, \delta_2\}$, we have

$$d\left(x + u, F^{-1}(y + v_1 - v_2)\right) \leq \max\left\{d\left((x, y) + (u, v_1), \operatorname{gph} F\right), \|y + v_2 - \bar{y}\|\right\}. \tag{2.29}$$

Moreover, if $\{\Omega_1, \Omega_2\}$ is $\alpha-$transversal at $\bar{x}$ with some $\delta_1 > 0$ and $\delta_2 > 0$, then conditions (i)–(iii) *hold with any $\delta_1' \in]0, \delta_1]$ and $\delta_2' > 0$ satisfying $\delta_2' + \alpha\delta_1' \leq \delta_2$ in place of δ_1 and δ_2.*

Conversely, if $\alpha \leq 1$ and conditions (i)–(iii) *hold with some $\delta_1 > 0$ and $\delta_2 > 0$, then $\{\Omega_1, \Omega_2\}$ is $\alpha-$transversal at $\bar{x}$ with any $\delta_1' \in]0, \delta_1]$ and $\delta_2' > 0$ satisfying $\delta_2' + \alpha\delta_1' \leq \delta_2$ in place of δ_1 and δ_2.*

Proof. Definitions (2.17) yield representation (2.21), and consequently, $(\bar{x}, \bar{y}) \in \Omega_1 \cap \Omega_2$.

(i) Given $a_1 := (u_1, v_1)$, $a_2 := (u_2, v_2)$, $(x, y) \in X \times Y$ and $x_2 \in X$, condition (2.26) implies (2.27), and the conditions are equivalent if $\alpha \leq 1$. Moreover, given $x, u \in X$ and $y, v_1, v_2 \in Y$, condition (2.27) holds with $u_1 := u$ for all $u_2 \in X$ with $\|u_2\| < \alpha\delta_1$ if and only if condition (2.28) is satisfied. Hence, condition (i) is equivalent to Proposition 2.4.2(i).

(ii) Given $a_1 := (u_1, v_1)$ and $a_2 := (u_2, v_2)$, condition (2.23) implies (2.24), and the conditions are equivalent if $\alpha \leq 1$. Moreover, given a $u \in X$ and $v_1, v_2 \in Y$, condition (2.24) holds with $u_1 := u$ for all $u_2 \in X$ with $\|u_2\| < \alpha\delta_1$ if and only if condition (2.25) is satisfied. Hence, condition (ii) is equivalent to Proposition 2.4.2(ii).

(iii) Given $a_1 := (u_1, v_1)$, $a_2 := (u_2, v_2)$ and $(x, y) \in X \times Y$, we have representation (2.22), and consequently,

$$d\left((x, y), (\Omega_1 - a_1) \cap (\Omega_2 - a_2)\right) = \max\left\{d\left(x + u_1, F^{-1}(y + v_1 - v_2)\right), \|y + v_2 - \bar{y}\|\right\}.$$

Thus, inequality

$$d\left((x, y), (\Omega_1 - a_1) \cap (\Omega_2 - a_2)\right) \leq \max\left\{d\left((x, y), \Omega_1 - a_1\right), d\left((x, y), \Omega_2 - a_2\right)\right\}$$

implies condition (2.29), and the converse implication is true if $\alpha \leq 1$. Hence, condition (iii) is equivalent to Proposition 2.4.2(iii).

The conclusions follow from Proposition 2.4.2. □

Next we apply the metric estimates in Proposition 2.4.1 to establish relations between regularity properties of set-valued mappings in Definition 2.5.1 and the corresponding transversality properties of the collection of sets (2.17).

Theorem 2.5.7. *Let X and Y be normed vector spaces, $F : X \rightrightarrows Y$, $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ and $\alpha > 0$. Let Ω_1 and Ω_2 be defined by* (2.17)*, $\alpha_1 := (2\alpha^{-1} + 1)^{-1}$ and $\alpha_2 := 2\alpha$.*

(i) If F is $\alpha-$semiregular at $(\bar{x},\bar{y})$ with some $\delta>0$, then $\{\Omega_1,\Omega_2\}$ is α_1-semitransversal at $(\bar{x},\bar{y})$ with $\delta':=(1+\alpha/2)\delta$.

Conversely, if $\{\Omega_1,\Omega_2\}$ is $\alpha-$semitransversal at $(\bar{x},\bar{y})$ with some $\delta>0$, then F is α_2-semiregular at $(\bar{x},\bar{y})$ with the same δ.

(ii) If F is $\alpha-$subregular at $(\bar{x},\bar{y})$ with some $\delta_1>0$ and $\delta_2>0$, then $\{\Omega_1,\Omega_2\}$ is α_1-subtransversal at $(\bar{x},\bar{y})$ with any $\delta_1'>0$ and $\delta_2'>0$ such that $\delta_1'\le(1+\alpha/2)\delta_1$ and $\alpha_1\delta_1'+\delta_2'\le\delta_2$.

Conversely, if $\{\Omega_1,\Omega_2\}$ is $\alpha-$subtransversal at $(\bar{x},\bar{y})$ with some $\delta_1>0$ and $\delta_2>0$, then F is α_2-subregular at $(\bar{x},\bar{y})$ with $\delta_1':=\min\{\delta_1,\alpha^{-1}\delta_2\}$ and δ_2.

(iii) If F is $\alpha-$regular at $(\bar{x},\bar{y})$ with some $\delta_1>0$ and $\delta_2>0$, then $\{\Omega_1,\Omega_2\}$ is α_1-transversal at $(\bar{x},\bar{y})$ with any $\delta_1'>0$ and $\delta_2'>0$ such that $\delta_1'\le(1+\alpha/2)\delta_1$ and $\alpha_1\delta_1'+\delta_2'\le\delta_2/2$.

Conversely, if $\{\Omega_1,\Omega_2\}$ is $\alpha-$transversal at $(\bar{x},\bar{y})$ with some $\delta_1>0$ and $\delta_2>0$, then F is α_2-regular at $(\bar{x},\bar{y})$ with any $\delta_1'\in]0,\delta_1]$ and $\delta_2'>0$ such that $\alpha_2\delta_1'+\delta_2'\le\delta_2$.

Proof. Note that $\alpha_1\in]0,1[$.

(i) Let F be $\alpha-$semiregular at $(\bar{x},\bar{y})$ with some $\delta>0$. Set $\delta':=(1+\alpha/2)\delta$. Let $u\in(\alpha_1\delta')\mathbb{B}_X$ and $v\in(2\alpha_1\delta')\mathbb{B}_Y$. Observe that
$$2\alpha_1\delta'=\frac{2(1+\alpha/2)}{2\alpha^{-1}+1}\delta=\alpha\delta.$$
Thus, $v\in(\delta\mathbb{B}_Y)$. In view of Definition 2.5.1(i),
$$\begin{aligned}d(\bar{x}+u,F^{-1}(\bar{y}+v))&\le d(\bar{x},F^{-1}(\bar{y}+v))+\|u\|\le\alpha^{-1}\|v\|+\|u\|\\&\le(2\alpha^{-1}+1)\max\left\{\|u\|,\frac{1}{2}\|v\|\right\}=\alpha_1^{-1}\max\left\{\|u\|,\frac{1}{2}\|v\|\right\}.\end{aligned}$$
By Proposition 2.4.1(i), $\{\Omega_1,\Omega_2\}$ is α_1-semitransversal at $(\bar{x},\bar{y})$ with δ'.

Conversely, let $\{\Omega_1,\Omega_2\}$ be $\alpha-$semitransversal at $(\bar{x},\bar{y})$ with some $\delta>0$. Let $v\in(\alpha_2\delta)\mathbb{B}_Y$. Thus, $v\in(2\alpha\delta)\mathbb{B}_Y$. By Proposition 2.4.1(i), $\alpha d\left(\bar{x},F^{-1}(\bar{y}+v)\right)\le\|v\|/2$, and consequently, $\alpha_2 d\left(\bar{x},F^{-1}(\bar{y}+v)\right)\le\|v\|$. Hence, by Definition 2.5.1(i), F is $\alpha-$semiregular at $(\bar{x},\bar{y})$ with the same δ.

(ii) Let F be $\alpha-$subregular at $(\bar{x},\bar{y})$ with some $\delta_1>0$ and $\delta_2>0$. Choose numbers $\delta_1'>0$ and $\delta_2'>0$ such that $\delta_1'\le(1+\alpha/2)\delta_1$ and $\delta_2'+\alpha_1\delta_1'\le\delta_2$. Observe that
$$2\alpha_1\delta_1'\le\frac{2(1+\alpha/2)}{2\alpha^{-1}+1}\delta_1=\alpha\delta_1.\tag{2.30}$$
Let $x\in B_{\delta_2'}(\bar{x})$, $y\in B_{\min\{\alpha_1\delta_1',\delta_2'\}}(\bar{y})$ and $d((x,y),\operatorname{gph}F)<\alpha_1\delta_1'$. Then, for any number ε with $d((x,y),\operatorname{gph}F)<\varepsilon<\alpha_1\delta_1'$, there exists a point $(x',y')\in\operatorname{gph}F$ such that $\|(x,y)-(x',y')\|<\varepsilon$. Thus,
$$\begin{aligned}\|x'-\bar{x}\|&\le\|x-\bar{x}\|+\|x'-x\|<\delta_2'+\alpha_1\delta_1'\le\delta_2,\\ d(\bar{y},F(x'))&\le\|y'-\bar{y}\|\le\|y-\bar{y}\|+\|y'-y\|<2\alpha_1\delta_1'\le\alpha\delta_1,\end{aligned}$$
and, in view of Definition 2.5.1(ii),
$$\begin{aligned}d\left(x,F^{-1}(\bar{y})\right)&\le d\left(x',F^{-1}(\bar{y})\right)+\|x'-x\|\\&\le\alpha^{-1}d(\bar{y},F(x'))+\|x'-x\|\\&\le\alpha^{-1}(\|y-\bar{y}\|+\|y'-y\|)+\|x'-x\|\\&<\alpha^{-1}\|y-\bar{y}\|+(\alpha^{-1}+1)\varepsilon\\&\le(2\alpha^{-1}+1)\max\{\|y-\bar{y}\|,\varepsilon\}=\alpha_1^{-1}\max\{\|y-\bar{y}\|,\varepsilon\}.\end{aligned}$$

Letting $\varepsilon \downarrow d((x,y), \operatorname{gph} F)$, we arrive at (2.19). By Proposition 2.4.1(ii), $\{\Omega_1, \Omega_2\}$ is α_1-subtransversal at $(\bar{x}, \bar{y})$ with δ_1' and δ_2'.

Conversely, let $\{\Omega_1, \Omega_2\}$ be $\alpha-$subtransversal at $(\bar{x}, \bar{y})$ with some $\delta_1 > 0$ and $\delta_2 > 0$. Set $\delta_1' := \min\{\delta_1, \alpha^{-1}\delta_2\}$. Let $x \in B_{\delta_2}(\bar{x})$ and $d(\bar{y}, F(x)) < \alpha_2\delta_1'$. Then, for any number ε with $d(\bar{y}, F(x)) < \varepsilon < \alpha_2\delta_1'$, there exists a point $y' \in F(x)$ such that $\|y' - \bar{y}\| < \varepsilon$. Set $y := \dfrac{y' + \bar{y}}{2}$. Observe that

$$\|y - y'\| = \|y - \bar{y}\| = \frac{\|y' - \bar{y}\|}{2} < \frac{\alpha_2\delta_1'}{2} = \alpha\delta_1',$$
$$\|y - \bar{y}\| < \alpha\delta_1' \le \min\{\alpha\delta_1, \delta_2\}, \quad d((x,y), \operatorname{gph} F) \le \|y - y'\| < \alpha\delta_1' \le \alpha\delta_1.$$

By Proposition 2.4.1(ii),

$$\alpha d\left(x, F^{-1}(\bar{y})\right) \le \max\left\{d((x,y), \operatorname{gph} F), \|y - \bar{y}\|\right\} = \|y - \bar{y}\| = \frac{\|y' - \bar{y}\|}{2} < \frac{\varepsilon}{2}.$$

Thus, $\alpha_2 d\left(x, F^{-1}(\bar{y})\right) < \varepsilon$. Letting $\varepsilon \downarrow d((x,y), \operatorname{gph} F)$, we obtain

$$\alpha_2 d\left(x, F^{-1}(\bar{y})\right) \le d((x,y), \operatorname{gph} F).$$

By Definition 2.5.1(ii), F is α_2-subregular at $(\bar{x}, \bar{y})$ with δ_1' and δ_2.

(iii) Let F be $\alpha-$regular at $(\bar{x}, \bar{y})$ with some $\delta_1 > 0$ and $\delta_2 > 0$. Choose numbers $\delta_1' > 0$ and $\delta_2' > 0$ such that $\delta_1' \le (1 + \alpha/2)\delta_1$ and $\delta_2' + \alpha_1\delta_1' \le \delta_2/2$. Then we have (2.24). Let $(x,y) \in \operatorname{gph} F \cap B_{\delta_2'}(\bar{x}, \bar{y})$, $u \in (\alpha_1\delta_1')\mathbb{B}_X$ and $v \in (2\alpha_1\delta_1')\mathbb{B}_Y$. Set $y' := y + v$. Then

$$\|x - \bar{x}\| + \|y' - \bar{y}\| \le \|x - \bar{x}\| + \|y - \bar{y}\| + \|v\| < 2\delta_2' + 2\alpha_1\delta_1' \le \delta_2,$$
$$d(y', F(x))) \le \|y' - y\| = \|v\| < 2\alpha_1\delta_1' \le \alpha\delta_1,$$

and, in view of Definition 2.5.1(iii),

$$\begin{aligned} d\left(x + u, F^{-1}(y + v)\right) &\le d\left(x, F^{-1}(y')\right) + \|u\| \le \alpha^{-1} d\left(y', F(x)\right) + \|u\| \\ &\le \alpha^{-1}\|v\| + \|u\| \le (2\alpha^{-1} + 1)\max\left\{\|u\|, \frac{1}{2}\|v\|\right\} \\ &= \alpha_1^{-1}\max\left\{\|u\|, \frac{1}{2}\|v\|\right\}. \end{aligned}$$

By Proposition 2.4.1(iii), $\{\Omega_1, \Omega_2\}$ is α_1-transversal at $(\bar{x}, \bar{y})$ with δ_1' and δ_2'.

Conversely, let $\{\Omega_1, \Omega_2\}$ be $\alpha-$transversal at $(\bar{x}, \bar{y})$ with some $\delta_1 > 0$ and $\delta_2 > 0$. Choose numbers $\delta_1' \in]0, \delta_1]$ and $\delta_2' > 0$ such that $\alpha_2\delta_1' + \delta_2' \le \delta_2$. Let $x \in X$ and $y \in Y$ be such that $\|x - \bar{x}\| + \|y - \bar{y}\| < \delta_2'$ and $d(y, F(x)) < \alpha_2\delta_1'$. Then, for any number ε with $d(y, F(x)) < \varepsilon < \alpha_2\delta_1'$, there exists a point $y' \in F(x)$ such that $\|y - y'\| < \varepsilon$. Then $(x, y') \in \operatorname{gph} F$,

$$\|x - \bar{x}\| < \delta_2' < \delta_2, \quad \|y' - \bar{y}\| \le \|y' - y\| + \|y - \bar{y}\| < \alpha_2\delta_1' + \delta_2' \le \delta_2,$$
$$\|y - y'\| < \alpha_2\delta_1' \le \alpha_2\delta_1 = 2\alpha\delta_1.$$

By Proposition 2.4.1(iii),

$$\alpha d\left(x, F^{-1}(y)\right) \le \frac{\|y - y'\|}{2} < \frac{\varepsilon}{2}.$$

Thus, $\alpha_2 d\left(x, F^{-1}(y)\right) < \varepsilon$. Letting $\varepsilon \downarrow d(y, F(x))$, we obtain

$$\alpha_2 d\left(x, F^{-1}(y)\right) \le d(y, F(x)).$$

By Definition 2.5.1(iii), F is α_2-regular at $(\bar{x}, \bar{y})$ with δ_1' and δ_2'.

□

The next corollary collects 'δ-free' versions of the relations in Theorem 2.5.7. It recaptures [106, Theorem 5.1].

Corollary 2.5.8. *Let X and Y be normed vector spaces, $F : X \rightrightarrows Y$, $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ and $\alpha > 0$. Let Ω_1 and Ω_2 be defined by* (2.17).

(i) F is semiregular at $(\bar{x}, \bar{y})$ if and only if $\{\Omega_1, \Omega_2\}$ is semitransversal at $(\bar{x}, \bar{y})$. Moreover,

$$\frac{1}{2\mathrm{s_e rg}\,[F](\bar{x}, \bar{y})^{-1} + 1} \leq \mathrm{s_e tr}\ [\Omega_1, \Omega_2](\bar{x}) \leq \frac{\mathrm{s_e rg}\,[F](\bar{x}, \bar{y})}{2}.$$

(ii) F is subregular at $(\bar{x}, \bar{y})$ if and only if $\{\Omega_1, \Omega_2\}$ is subtransversal at $(\bar{x}, \bar{y})$. Moreover,

$$\frac{1}{2\mathrm{srg}\,[F](\bar{x}, \bar{y})^{-1} + 1} \leq \mathrm{str}\ [\Omega_1, \Omega_2](\bar{x}) \leq \frac{\mathrm{srg}\,[F](\bar{x}, \bar{y})}{2}.$$

(iii) F is regular at $(\bar{x}, \bar{y})$ if and only if $\{\Omega_1, \Omega_2\}$ is transversal at $(\bar{x}, \bar{y})$. Moreover,

$$\frac{1}{2\mathrm{rg}\,[F](\bar{x}, \bar{y})^{-1} + 1} \leq \mathrm{tr}\ [\Omega_1, \Omega_2](\bar{x}) \leq \frac{\mathrm{rg}\,[F](\bar{x}, \bar{y})}{2}.$$

2.6 Distances between n Sets

When studying mutual arrangement of collections of sets in space, particularly their extremality, stationarity and regularity/transversality properties, we need to be able to estimate the 'distance' between the sets, i.e. how 'far apart' they are, or, at least, whether they have a common point. In the case of two sets, the conventional distance

$$d(\Omega_1, \Omega_2) := \inf_{\omega_1 \in \Omega_1,\ \omega_2 \in \Omega_2} d(\omega_1, \omega_2) \tag{2.31}$$

does the job. The general case of $n \geq 2$ sets is not that straightforward. In this section we discuss several candidates for the role of 'distance'.

We start with discussing distances between $n \geq 2$ points.

2.6.1 Distances between n Points

Our aim in this subsection is to consider ways of defining *n-point distances* estimating quantitatively the overall 'closeness' of a collection of $n \geq 2$ points in a metric space (X, d).

Given $n \geq 2$ points $\omega_1, \ldots, \omega_n$ in X, one can use one of the following two quantities:

$$d_1(\omega_1, \ldots, \omega_n) := \max_{1 \leq i \leq n-1} d(\omega_i, \omega_n), \tag{2.32}$$

$$d_2(\omega_1, \ldots, \omega_n) := \inf_{x \in X} \max_{1 \leq i \leq n} d(\omega_i, x). \tag{2.33}$$

When $n = 2$, (2.32) reduces to the conventional distance $d(\omega_1, \omega_2)$. However, when $n > 2$ this distance is not symmetric: the last point in the list plays a special role, and the quantity itself depends on the choice of the last point. In contrast to (2.32), in definition (2.33) all points play the same role, and it involves minimization over an additional parameter x. Formulas (2.32) and (2.33) produce in general different numbers, whatever the choice of the last point in (2.32) is.

In the setting of a normed vector space, the following symmetric distance can be of interest:

$$d_3(\omega_1, \dots, \omega_n) := \max_{1 \le i \le n} \left\| \omega_i - \frac{1}{n} \sum_{j=1}^{n} \omega_j \right\|. \tag{2.34}$$

Example 2.6.1. Consider three points in $\mathbb{R}$: 0, 1 and 5. By (2.32), we have $d_1(0,5,1) = \max\{|0-1|, |5-1|\} = 4$, while $d_1(0,1,5) = d_1(1,5,0) = 5$. The infimum in definition (2.33) is achieved at $x = 2.5$ and equals $d_2(0,1,5) = \max\{|0-2.5|, |1-2.5|, |5-2.5|\} = 2.5$. The average $\frac{1}{3}(0+1+5)$ equals 2, and formula (2.34) gives $d_3(0,1,5) = \max\{|0-2|, |1-2|, |5-2|\} = 3$. Thus, for these three points all three definitions (2.32), (2.33) and (2.34) give different numbers.

Note the obvious connection between the distances d_1 and d_2:

$$d_2(\omega_1, \dots, \omega_n) := \inf_{x \in X} d_1(\omega_1, \dots, \omega_n, x).$$

Observe also that $d_1(\omega_1, \dots, \omega_n)$ is actually the maximum distance between the points $(\omega_1, \dots, \omega_{n-1})$ and $(\omega_n, \dots, \omega_n)$ in X^{n-1}, while $d_2(\omega_1, \dots, \omega_n)$ represents the distance from the points $(\omega_1, \dots, \omega_n) \in X^n$ to the 'diagonal' subspace $\{(\omega, \dots, \omega) \in X^n : \omega \in X\}$. If the points $\omega_1, \dots, \omega_n$ do not all coincide, then the quantity computed in accordance with formula (2.33) will remain strictly positive if the infimum there is taken not over the whole space, but over any its subset. This can be a way of defining 'localised distances'.

Some properties of the quantities (2.32) and (2.33) are collected in the next proposition.

Proposition 2.6.2. *Suppose $\omega_1, \dots, \omega_n$ $(n \ge 2)$ are points in a metric space X.*

(i) $d_2(\omega_1, \dots, \omega_n) \le d_1(\omega_1, \dots, \omega_n) \le 2d_2(\omega_1, \dots, \omega_n)$.

(ii) If X is a normed vector space with the distance induced by the norm and $n = 2$, then $d_1(\omega_1, \omega_2) = 2d_2(\omega_1, \omega_2)$, i.e. the last inequality in (i) *holds as equality.*

Suppose additionally that X is a normed vector space.

(iii) $d_2(\omega_1, \dots, \omega_n) \le d_3(\omega_1, \dots, \omega_n) \le 2d_2(\omega_1, \dots, \omega_n)$.

(iv) If $n = 2$, then $d_2(\omega_1, \omega_2) = d_3(\omega_1, \omega_2)$, i.e. the first inequality in (iii) *holds as equality.*

Proof. (i) The first inequality follows immediately from the definitions:

$$d_2(\omega_1, \dots, \omega_n) = \inf_{x \in X} \max_{1 \le i \le n} d(\omega_i, x) \le \max_{1 \le i \le n} d(\omega_i, \omega_n) = \max_{1 \le i \le n-1} d(\omega_i, \omega_n) = d_1(\omega_1, \dots, \omega_n).$$

To prove the second inequality, first fix an $x \in X$.

$$d_1(\omega_1, \dots, \omega_n) = \max_{1 \le i \le n-1} d(\omega_i, \omega_n) \le \max_{1 \le i \le n-1} (d(\omega_i, x) + d(\omega_n, x)) \le 2 \max_{1 \le i \le n} d(\omega_i, x).$$

Taking the infimum over $x \in X$ in the right-hand side of the above inequality, we arrive at the second inequality in (i).

(ii) Let X be a normed vector space with the distance induced by the norm and $n = 2$. Then

$$\begin{aligned} d_1(\omega_1, \omega_2) = \|\omega_1 - \omega_2\| &= 2 \max \left\{ \left\| \frac{\omega_1 + \omega_2}{2} - \omega_1 \right\|, \left\| \frac{\omega_1 + \omega_2}{2} - \omega_2 \right\| \right\} \\ &\ge 2 \inf_{x \in X} \max\{\|x - \omega_1\|, \|x - \omega_2\|\} = 2d_2(\omega_1, \omega_2). \end{aligned}$$

Combining this with the second inequality in (i) proves (ii).

(iii) Suppose that X is a normed vector space. The first inequality follows immediately from the definitions. To prove the second inequality, first fix an $x \in X$. Then

$$d_3(\omega_1, \dots, \omega_n) = \max_{1 \le i \le n} \left\| \omega_i - \frac{1}{n} \sum_{j=1}^{n} \omega_j \right\| \le \max_{1 \le i \le n} \|\omega_i - x\| + \left\| \frac{1}{n} \sum_{j=1}^{n} \omega_j - x \right\|$$
$$\le \max_{1 \le i \le n} \|\omega_i - x\| + \frac{1}{n} \sum_{j=1}^{n} \|\omega_j - x\| \le 2 \max_{1 \le i \le n} \|\omega_i - x\| .$$

Taking the infimum over $x \in X$ in the right-hand side of the second inequality, we arrive at the last inequality in (iii).

(iv) Let $n = 2$. Using definitions (2.34) and (2.32), and the equality in (ii), we obtain:

$$\begin{aligned} d_3(\omega_1, \omega_2) &= \max \left\{ \left\| \omega_1 - \frac{\omega_1 + \omega_2}{2} \right\|, \left\| \omega_2 - \frac{\omega_1 + \omega_2}{2} \right\| \right\} \\ &= \frac{1}{2} \|\omega_1 - \omega_2\| \\ &= \frac{1}{2} d_1(\omega_1, \omega_2) = d_2(\omega_1, \omega_2). \end{aligned}$$

□

It follows from part (ii) of Proposition 2.6.2 that quantity (2.33) does not reduce to the conventional distance when $n = 2$: in the setting of a normed vector space it equals $\frac{1}{2} d(\omega_1, \omega_2)$. In part (ii) of Proposition 2.6.2, the infimum in formula (2.33) is computed explicitly. Unfortunately, when $n > 2$, this seems impossible in general even in the setting of a normed vector space; see the discussion in [31, Section 6].

Remark 2.6.3. In the setting of a normed vector space, if $n = 2$ and $\omega_1 \neq \omega_2$, then, in view of Proposition 2.6.2(ii) and (iv), the first inequality in Proposition 2.6.2(i) and the second inequality in Proposition 2.6.2(iii) are strict. If $n > 2$, then all the inequalities in Proposition 2.6.2(i) and (iii) can be strict. This fact is illustrated by Example 2.6.1.

2.6.2 Distances between n Sets

Now we employ the distances between collections of points discussed in the previous subsection to quantify 'closeness' of collections of sets. Given $n \ge 2$ subsets $\Omega_1, \dots, \Omega_n$ of a metric space X and an n-point distance d, the distance between $\Omega_1, \dots, \Omega_n$ is defined in the usual way:

$$d(\Omega_1, \dots, \Omega_n) := \inf_{\omega_1 \in \Omega_1, \dots, \omega_n \in \Omega_n} d(\omega_1, \dots, \omega_n). \tag{2.35}$$

Applying construction (2.35) to the n-point distances (2.32), (2.33) and (2.34), we obtain the following definitions of particular distances between n sets, respectively:

$$d_1(\Omega_1, \dots, \Omega_n) := \inf_{\omega_1 \in \Omega_1, \dots, \omega_n \in \Omega_n} \max_{1 \le i \le n-1} d(\omega_i, \omega_n), \tag{2.36}$$

$$d_2(\Omega_1, \dots, \Omega_n) := \inf_{\omega_1 \in \Omega_1, \dots, \omega_n \in \Omega_n, x \in X} \max_{1 \le i \le n} d(\omega_i, x), \tag{2.37}$$

$$d_3(\Omega_1, \dots, \Omega_n) := \inf_{\omega_1 \in \Omega_1, \dots, \omega_n \in \Omega_n} \max_{1 \le i \le n} \left\| \omega_i - \frac{1}{n} \sum_{j=1}^{n} \omega_j \right\| . \tag{2.38}$$

When $n = 2$, definition (2.36) reduces to the conventional distance (2.31). Definition (2.38) is meaningful in the setting of a normed vector space only.

Proposition 2.6.2 leads to similar relations for the distances between n sets.

Proposition 2.6.4. *Suppose $\Omega_1, \ldots, \Omega_n$ are subsets of a metric space X.*

(i) $d_2(\Omega_1, \ldots, \Omega_n) \le d_1(\Omega_1, \ldots, \Omega_n) \le 2d_2(\Omega_1, \ldots, \Omega_n)$.

(ii) If $n = 2$, then $d_1(\Omega_1, \Omega_2) = 2d_2(\Omega_1, \Omega_2)$.

(iii) $d_1(\Omega_1, \ldots, \Omega_n) > 0 \iff d_2(\Omega_1, \ldots, \Omega_n) > 0$.

Suppose additionally that X is a normed vector space.

(iv) $d_2(\Omega_1, \ldots, \Omega_n) \le d_3(\Omega_1, \ldots, \Omega_n) \le 2d_2(\Omega_1, \ldots, \Omega_n)$.

(v) If $n = 2$, then $d_2(\Omega_1, \Omega_2) = d_3(\Omega_1, \Omega_2)$.

(vi) $d_1(\Omega_1, \ldots, \Omega_n) > 0 \iff d_2(\Omega_1, \ldots, \Omega_n) > 0 \iff d_3(\Omega_1, \ldots, \Omega_n) > 0$.

Proof. Conditions (i), (ii) and (iv), (v) are consequences of the corresponding conditions in Proposition 2.6.2. Condition (iii) is a consequence of condition (i). Condition (vi) is a consequence of conditions (iii) and (iv). □

Another distance can be of interest and is going to be used in the sequel. Given a subset Ω_{n+1} of X, define

$$d_{\Omega_{n+1}}(\Omega_1, \ldots, \Omega_n) := \inf_{\omega_1 \in \Omega_1, \ldots, \omega_n \in \Omega_n, x \in \Omega_{n+1}} \max_{1 \le i \le n} d(\omega_i, x).$$

Now observe that $d_{\Omega_{n+1}}$ as well as the symmetric distance d_2 (2.37) are particular cases of the asymmetric distance d_1 (2.36) applied to $n+1$ sets:

$$d_{\Omega_{n+1}}(\Omega_1, \ldots, \Omega_n) = d_1(\Omega_1, \ldots, \Omega_n, \Omega_{n+1}) \quad \text{and} \quad d_2(\Omega_1, \ldots, \Omega_n) = d_1(\Omega_1, \ldots, \Omega_n, X). \tag{2.39}$$

This observation makes the asymmetric distance d_1 a rather general quantitative measure of closeness of collections of sets. Apart from the most straightforward case $\Omega_{n+1} := X$ used in the above example, another useful particular case is given by $\Omega_{n+1} := B_\rho(\bar{x})$ where $\bar{x} \in X$ is a fixed point (related to the sets $\Omega_1, \ldots, \Omega_n$) and $\rho > 0$. This allows one to examine closeness of sets in a neighbourhood of the given point.

Remark 2.6.5. (i) As it was observed earlier, the distance d_1 (2.36) depends in general on the order of the sets. However, thanks to Proposition 2.6.4(iii), if it is strictly positive for some permutation of the sets, it remains strictly positive for any other permutation.

(ii) The maximum operation in all the above definitions of distances between n points and n sets corresponds to the maximum norm in either $\mathbb{R}^{n-1}$ or $\mathbb{R}^n$. It can be replaced in all these definitions and subsequent statements by any other finite dimensional norm producing different but in a sense equivalent 'distances'. The p-norm version of the quantity (2.36) was considered in [172] under the name *(p-weighted) nonintersect index*. In the current thesis, for the sake of simplicity of presentation only the maximum norm is considered.

The next proposition and its corollary characterise a set of points in the given collection of sets, which are almost closest (up to ε) points of these sets with respect to the chosen n-point distance. (When $n = 2$ such points play a key role in the geometric versions of the Ekeland variational principle considered in Section 3.) It also introduces a two-step procedure, which is going to be used in the sequel. Given a collection of sets $\Omega_1, \ldots, \Omega_n$ with empty intersection and a collection of points $\omega_i \in \Omega_i$ $(i = 1, \ldots, n)$, we

1) consider another collection of sets $\Omega_1 - \omega_1, \ldots, \Omega_n - \omega_n$, whose intersection is obviously nonempty, and

2) construct 'small' (up to ε) translation vectors $a_1, \ldots, a_n$ such that the translated sets $\Omega_1 - \omega_1 - a_1, \ldots, \Omega_n - \omega_n - a_n$ have empty intersection again.

Thus, the proposition translates ε-closeness of a collection of points, which served as the key assumption when proving unified separation theorems in [172], into the language of ε-translations of the sets employed in Definition 2.3.1 of extremality/stationarity properties, the corresponding Proposition 2.3.4.

For simplicity and in view of the observed above universality of the distance d_1 (2.36), we first consider this distance.

Proposition 2.6.6. *Suppose $\Omega_1, \ldots, \Omega_n$ are subsets of a normed vector space X, $\cap_{i=1}^n \Omega_i = \emptyset$, $\omega_i \in \Omega_i$ $(i = 1, \ldots, n)$ and $\varepsilon > 0$. If*

$$d_1(\omega_1, \ldots, \omega_n) < d_1(\Omega_1, \ldots, \Omega_n) + \varepsilon, \tag{2.40}$$

then $M := d_1(\omega_1, \ldots, \omega_n) > 0$ and condition (P8) *is satisfied, where $a_i := \frac{\varepsilon'}{M}(\omega_n - \omega_i)$ $(i = 1, \ldots, n-1)$ and ε' is either any number in $[0, \varepsilon[$ satisfying $M - d_1(\Omega_1, \ldots, \Omega_n) < \varepsilon' \leq M$ if $d_1(\Omega_1, \ldots, \Omega_n) > 0$, or $\varepsilon' = M$ if $d_1(\Omega_1, \ldots, \Omega_n) = 0$.*

Proof. In view of $\cap_{i=1}^n \Omega_i = \emptyset$, we have $M > 0$. Let condition (2.40) be satisfied.

Suppose first that $d_1(\Omega_1, \ldots, \Omega_n) > 0$. Choose a positive number $\varepsilon' < \varepsilon$ such that

$$M - d_1(\Omega_1, \ldots, \Omega_n) < \varepsilon' \leq M.$$

Then $\max\limits_{1 \leq i \leq n-1} \|a_i\| = \varepsilon' < \varepsilon$. Suppose that the first condition in (P8) does not hold. Then there exists a point $x \in \cap_{i=1}^{n-1} (\Omega_i - \omega_i - a_i) \cap (\Omega_n - \omega_n)$, and consequently, $\hat{\omega}_i := \omega_i + a_i + x \in \Omega_i$ $(i = 1, \ldots, n-1)$ and $\hat{\omega}_n := \omega_n + x \in \Omega_n$. Thus,

$$\|\hat{\omega}_i - \hat{\omega}_n\| = \left\| \omega_i - \omega_n - \frac{\varepsilon'}{M}(\omega_i - \omega_n) \right\| = \left(1 - \frac{\varepsilon'}{M}\right) \|\omega_i - \omega_n\| \quad (i = 1, \ldots, n-1),$$

and consequently,

$$d_1(\Omega_1, \ldots, \Omega_n) \leq d_1(\hat{\omega}_1, \ldots, \hat{\omega}_n) = \left(1 - \frac{\varepsilon'}{M}\right) d_1(\omega_1, \ldots, \omega_n) = M - \varepsilon'.$$

This contradicts the choice of ε'. Hence, conditions (P8) hold true.

In the case $d_1(\Omega_1, \ldots, \Omega_n) = 0$, set $a_i := \omega_n - \omega_i$ $(i = 1, \ldots, n-1)$. Then, by (2.40), $\max\limits_{1 \leq i \leq n-1} \|a_i\| = d_1(\omega_1, \ldots, \omega_n) < \varepsilon$ and

$$\bigcap_{i=1}^{n-1} (\Omega_i - \omega_i - a_i) \cap (\Omega_n - \omega_n) = \bigcap_{i=1}^{n} (\Omega_i - \omega_n) = \bigcap_{i=1}^{n} \Omega_i - \omega_n = \emptyset,$$

i.e., conditions (P8) hold true. ☐

Applying Proposition 2.6.6 to the collection of $n+1$ sets $\Omega_1, \ldots, \Omega_n, X$, we arrive at the following statement in terms of the distance d_2 (2.37).

Corollary 2.6.7. *Suppose $\Omega_1, \ldots, \Omega_n$ are subsets of a normed vector space X, $\cap_{i=1}^n \Omega_i = \emptyset$, $\omega_i \in \Omega_i$ $(i = 1, \ldots, n)$, $x \in X$ and $\varepsilon > 0$. If*

$$\max_{1 \leq i \leq n} d(\omega_i, x) < d_2(\Omega_1, \ldots, \Omega_n) + \varepsilon, \tag{2.41}$$

then $M := \max_{1\le i\le n} \|\omega_i - x\| > 0$ *and conditions* (P1) *hold true with* $\Omega_i' := \Omega_i - \omega_i$ *in place of* Ω_i *and* $a_i := \frac{\varepsilon'}{M}(x-\omega_i)$ $(i = 1,\ldots,n)$*, where* ε' *is either any number in* $]0,\varepsilon[$ *satisfying* $M - d_2(\Omega_1,\ldots,\Omega_n) < \varepsilon' \le M$ *if* $d_2(\Omega_1,\ldots,\Omega_n) > 0$*, or* $\varepsilon' = M$ *if* $d_2(\Omega_1,\ldots,\Omega_n) = 0$.

2.7 Extended Extremal Principle

The next well-known theorem (see Remark 2.7.2 below) gives approximate dual necessary conditions of local extremality in terms of Fréchet normals. It can be considered as a generalisation of the classical convex separation theorem to pairs of nonconvex sets.

Theorem 2.7.1. *Suppose* $\Omega_1,\ldots,\Omega_n$ *are closed subsets of an Asplund space* X *and* $\bar x \in \cap_{i=1}^n \Omega_i$*. If the collection* $\{\Omega_1,\ldots,\Omega_n\}$ *is local extremal at* $\bar x$ *then the following equivalent conditions hold with* N *standing for the Fréchet normal cone* ($N := N^F$)*:*

(i) *for any* $\varepsilon > 0$*, there exist points* $\omega_i \in \Omega_i \cap B_\varepsilon(\bar x)$ *and vectors* $x_i^* \in X^*$ $(i = 1,\ldots,n)$ *such that*

$$\left\|\sum_{i=1}^n x_i^*\right\| < \varepsilon, \quad x_i^* \in N_{\Omega_i}(\omega_i)\ (i = 1,\ldots,n) \quad \textit{and} \quad \sum_{i=1}^n \|x_i^*\| = 1; \tag{D1}$$

(ii) *for any* $\varepsilon > 0$*, there exist points* $\omega_i \in \Omega_i \cap B_\varepsilon(\bar x)$ *and vectors* $x_i^* \in X^*$ $(i = 1,\ldots,n)$ *such that*

$$\sum_{i=1}^n x_i^* = 0, \quad \sum_{i=1}^n d(x_i^*, N_{\Omega_i}(\omega_i)) < \varepsilon \quad \textit{and} \quad \sum_{i=1}^n \|x_i^*\| = 1. \tag{D2}$$

Remark 2.7.2. The inequalities in (D1) and (D2) are only meaningful when $\varepsilon \le 1$, because otherwise they are direct consequences of the corresponding equalities. Indeed, when $\varepsilon > 1$, condition (i) in Theorem 2.7.1 is satisfied automatically while condition (ii) guarantees only the existence of nontrivial normals in the ε-neighbourhood of $\bar x$ to at least one of the sets Ω_i $(i = 1,\ldots,n)$, which is trivial as long as $\bar x$ is a boundary point of one of the sets (which is the case when $\{\Omega_1,\ldots,\Omega_n\}$ is locally extremal at $\bar x$). If conditions (D1) and (D2) hold with $\varepsilon = 1$, they also hold with some $\varepsilon < 1$. Thanks to these observations, when applying Theorem 2.7.1 or its extensions, one can always assume that $\varepsilon < 1$.

Both conclusions in the above theorem are pretty common dual space properties used in many contemporary formulations of the extremal principle and its extensions. Properties (i) and (ii) can be found e.g. in, respectively, [120, Definition 2.5] (the *approximate extremal principle*) and [88, Definition 2.3] (the *generalised Euler equation*); cf. [92, property $(\mathrm{SP})_S$]. Condition (i) guarantees the existence of a pair of vectors x_i^* $(i = 1,\ldots,n)$ in the dual space, which are 'almost normal' (up to ε) to the corresponding sets at certain points with $\sum_{i=1}^n x_i^* = 0$ and $\sum_{i=1}^n \|x_i^*\| = 1$, while condition (ii) guarantees the existence of a set of vectors x_i^* which are exactly normal (in the Fréchet sense) to the corresponding sets at certain points with their sum $\sum_{i=1}^n x_i^*$ being small (up to ε) and $\sum_{i=1}^n \|x_i^*\| = 1$.

The extremal principle in Theorem 2.7.1 gives necessary conditions of (local) extremality which are in general not sufficient. Just like in the classical analysis and optimisation theory, it actually characterises a weaker than extremality property which can be interpreted as a kind of stationarity. The properties in the next proposition came to life as a result of a search for the weakest assumptions on the sets $\{\Omega_1,\ldots,\Omega_n\}$ (approximate stationarity) which still ensure the conclusions of the extremal principle.

Theorem 2.7.3 (Extended extremal principle)**.** *Suppose* $\Omega_1,\ldots,\Omega_n$ *are closed subsets of an Asplund space* X *and* $\bar x \in \cap_{i=1}^n \Omega_i$*. Then,* $\{\Omega_1,\ldots,\Omega_n\}$ *is approximately stationary at* $\bar x$ *if and only if* (D1) *and* (D2) *holds with* N *standing for the Fréchet normal cone* ($N := N^F$)*.*

The fact that the generalised separation actually characterises a weaker than (local) extremality property of approximate stationarity, and this characterisation is necessary and sufficient was first established (in a slightly different form) in [85] in the setting of a Fréchet smooth Banach space and extended to Asplund spaces in [86]. The full proof of the extended extremal principle appeared in [87], while the name *Extended extremal principle* was introduced in [88]. It is worth noting that the proof of the necessity (of either condition (ii) or condition (iii)) in the extended extremal principle follows that of the conventional extremal principle and only refines some estimates, while the sufficiency is almost straightforward and is valid in arbitrary normed vector spaces.

Remark 2.7.4. (i) Conditions (D1) and (D2) in Theorem 2.7.3 represent two kinds of widely used *generalised (approximate) separation* of a collection of sets: they claim the existence of n vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$, satisfying the normalisation condition $\sum_{i=1}^{n} \|x_i^*\| = 1$, and being either normal to the respective sets at some points close (up to ε) to $\bar{x}$ with their sum being almost (up to ε) 0 (condition (ii)), or almost (up to ε) normal with their sum equal exactly 0 (condition (iii)).

The equivalence of the conditions (D1) and (D2) in Theorem 2.7.3 is not difficult to check directly (even in the setting of an arbitrary normed vector space and without assuming the closedness of the sets) using elementary arguments which involve small perturbations of the vectors and then scaling the perturbed vectors to ensure the normalisation condition. Such arguments have been used in many proofs and are scattered across a number of publications. They were made explicit (in the case of two sets) in [29, Lemma 1]. In the next subsection, we formulate a more general statement, which will be used also in Section 3.3.

(ii) The inequality in (D2) is only meaningful when $\varepsilon \leq 1$, because otherwise it is a direct consequence of the equalities. If condition (D2) holds with $\varepsilon = 1$, it also holds with some $\varepsilon < 1$. Thanks to these observations, when applying Theorem 2.7.3 or its extensions, one can always assume that $\varepsilon < 1$.

(iii) Theorem 2.7.3 (as well as the conventional extremal principle) with conditions (ii) and (iii) in their current form cannot be extended beyond Asplund spaces. However, replacing in these conditions Fréchet normal cones with normal cones corresponding to other subdifferentials possessing reasonable (approximate or exact) sum rules in the respective *trustworthy* [69] spaces, one can employ basically the same routine to show the necessity of the amended conditions in these spaces; see e.g. [24, 69, 97, 120] and the discussion in the Introduction. Thanks to Lemma 2.2.1(iv) and (i), in general Banach spaces one can use Clarke normal cones or even conventional normal cones in the sense of convex analysis if the sets are convex. Note that the sufficiency of the conditions (D1) and (D2) in Theorem 2.7.3 is only valid for the Fréchet normal cones, though in general normed vector spaces. Thus, with such extensions, we only have (i) ⇒ (ii) ⇔ (iii), unless the sets are convex.

2.7.1 Perturbations and Scaling of Vectors

Below we present several assertions containing elementary arguments which are used in proving the equivalence of the conditions(D1) and (D2) in Theorem 2.7.3 and similar facts. We start with a general statement, which will be used also in Section 3.3.

Lemma 2.7.5. *Suppose $K_1, \ldots, K_n$ are cones in a normed vector space, $\varepsilon > 0$, $\rho > 0$ and $\lambda > 0$. Suppose also that vectors $z_1, \ldots, z_n$ satisfy*

$$\lambda \sum_{i=1}^{n} d(z_i, K_i) + \rho \left\| \sum_{i=1}^{n} z_i \right\| < \varepsilon, \quad \sum_{i=1}^{n} \|z_i\| = 1. \tag{2.42}$$

(i) *If $\varepsilon < \rho$ and $\lambda \leq \rho - \varepsilon$, then there exist vectors $\hat{z}_i$ $(i = 1, \ldots, n)$ satisfying the following conditions:*

$$\sum_{i=1}^{n} \hat{z}_i = 0, \quad \sum_{i=1}^{n} d(\hat{z}_i, K_i) < \frac{\varepsilon}{\lambda}, \quad \sum_{i=1}^{n} \|\hat{z}_i\| = 1. \tag{2.43}$$

(ii) *If $\rho < \lambda$ and $\varepsilon \le \lambda - \rho$ then there exist vectors $\hat{z}_i$ $(i = 1, \dots, n)$ satisfying the following conditions:*

$$\left\| \sum_{i=1}^n \hat{z}_i \right\| < \frac{\varepsilon}{\rho}, \quad \hat{z}_i \in K_i \ (i = 1, \dots, n), \quad \sum_{i=1}^n \|\hat{z}_i\| = 1. \tag{2.44}$$

(iii) *Moreover, if the underlying space is dual to a normed vector space X, and*

$$\sum_{i=1}^n \langle z_i, x_i \rangle \ge \tau \max_{1 \le i \le n} \|x_i\| \tag{2.45}$$

for some vectors $x_i \in X$ $(i = 1, \dots, n)$, not all zero, and a number $\tau \in]0, 1]$, then the vectors $\hat{z}_i$ $(i = 1, \dots, n)$ in parts (i) *or* (ii) *satisfy*

$$\sum_{i=1}^n \langle \hat{z}_i, x_i \rangle > \hat{\tau} \max_{1 \le i \le n} \|x_i\|, \tag{2.46}$$

where $\hat{\tau} := \dfrac{\tau\rho - \varepsilon}{\rho + \varepsilon}$ under the assumptions in part (i)*, and $\hat{\tau} := \dfrac{\tau\lambda - \varepsilon}{\lambda + \varepsilon}$ under the assumptions in part* (ii).

Proof. (i) Let $\varepsilon < \rho$ and $\lambda \le \rho - \varepsilon$. Set $z := \sum_{i=1}^n z_i$ and $v_i := z_i - \frac{1}{n} z$ $(i = 1, \dots, n)$. Then $\sum_{i=1}^n v_i = 0$ and, by (2.42), $\|z\| < \frac{\varepsilon}{\rho} < 1$,

$$\sum_{i=1}^n \|v_i\| \le \sum_{i=1}^n \|z_i\| + \|z\| < 1 + \frac{\varepsilon}{\rho} \quad \text{and} \quad \sum_{i=1}^n \|v_i\| \ge \sum_{i=1}^n \|z_i\| - \|z\| = 1 - \|z\| > 0.$$

Set $\hat{z}_i := v_i / \sum_{i=1}^n \|v_i\|$. Then $\sum_{i=1}^n \hat{z}_i = 0$, $\sum_{i=1}^n \|\hat{z}_i\| = 1$ and

$$\sum_{i=1}^n d(\hat{z}_i, K_i) \le \frac{\sum_{i=1}^n d(z_i, K_i) + \|z\|}{\sum_{i=1}^n \|v_i\|} < \frac{\varepsilon - (\rho - \lambda)\|z\|}{\lambda(1 - \|z\|)} \le \frac{\varepsilon - \varepsilon\|z\|}{\lambda(1 - \|z\|)} = \frac{\varepsilon}{\lambda},$$

i.e., all conditions in (2.43) are satisfied.

(ii) Let $\rho < \lambda$ and $\varepsilon \le \lambda - \rho$. By (2.42), there exist vectors $v_i \in K_i$ $(i = 1, \dots, n)$ such that

$$\lambda \sum_{i=1}^n \|z_i - v_i\| + \rho \left\| \sum_{i=1}^n z_i \right\| < \varepsilon.$$

In particular, $\sum_{i=1}^n \|z_i - v_i\| < \frac{\varepsilon}{\lambda} < 1$. Hence,

$$\sum_{i=1}^n \|v_i\| \le \sum_{i=1}^n \|z_i\| + \sum_{i=1}^n \|z_i - v_i\| = 1 + \sum_{i=1}^n \|z_i - v_i\|,$$
$$\sum_{i=1}^n \|v_i\| \ge \sum_{i=1}^n \|z_i\| - \sum_{i=1}^n \|z_i - v_i\| = 1 - \sum_{i=1}^n \|z_i - v_i\| > 0,$$
$$\left\| \sum_{i=1}^n v_i \right\| \le \sum_{i=1}^n \|z_i - v_i\| + \left\| \sum_{i=1}^n z_i \right\| < \frac{1}{\rho} \left(\varepsilon - (\lambda - \rho) \sum_{i=1}^n \|z_i - v_i\| \right).$$

Set $\hat{z}_i := v_i/\sum_{i=1}^{n}\|v_i\|$. Then $\hat{z}_i \in K_i$ $(i = 1,\ldots,n)$, $\sum_{i=1}^{n}\|\hat{z}_i\| = 1$ and

$$\left\|\sum_{i=1}^{n}\hat{z}_i\right\| = \frac{\|\sum_{i=1}^{n} v_i\|}{\sum_{i=1}^{n}\|v_i\|} < \frac{\varepsilon - (\lambda-\rho)\sum_{i=1}^{n}\|z_i - v_i\|}{\rho(1-\sum_{i=1}^{n}\|z_i - v_i\|)} \le \frac{\varepsilon - \varepsilon\sum_{i=1}^{n}\|z_i - v_i\|}{\rho(1-\sum_{i=1}^{n}\|z_i - v_i\|)} = \frac{\varepsilon}{\rho},$$

i.e., all conditions in (2.44) are satisfied.

(iii) Suppose that the underlying space is dual to a normed vector space X and condition (2.45) is satisfied for some vectors $x_i \in X$ $(i = 1,\ldots,n)$, not all zero, and a number $\tau \in]0,1]$. Then, using the notations introduced above, we have:

$$\sum_{i=1}^{n}\langle\hat{z}_i, x_i\rangle \ge \frac{\sum_{i=1}^{n}\langle z_i, x_i\rangle - (\sum_{i=1}^{n}\|z_i - v_i\|)\max_{1\le i\le n}\|x_i\|}{\sum_{i=1}^{n}\|v_i\|}$$
$$\ge \frac{\tau - \sum_{i=1}^{n}\|z_i - v_i\|}{\sum_{i=1}^{n}\|v_i\|}\max_{1\le i\le n}\|x_i\|.$$

Employing the estimates in part (i), we obtain

$$\frac{\tau - \sum_{i=1}^{n}\|z_i - v_i\|}{\sum_{i=1}^{n}\|v_i\|} = \frac{\tau - \|z\|}{\sum_{i=1}^{n}\|v_i\|} > \frac{\tau - \frac{\varepsilon}{\rho}}{1+\frac{\varepsilon}{\rho}} = \frac{\tau\rho - \varepsilon}{\rho+\varepsilon},$$

while the estimates in part (ii) give

$$\frac{\tau - \sum_{i=1}^{n}\|z_i - v_i\|}{\sum_{i=1}^{n}\|v_i\|} \ge \frac{\tau - \sum_{i=1}^{n}\|z_i - v_i\|}{1+\sum_{i=1}^{n}\|z_i - v_i\|} > \frac{\tau - \frac{\varepsilon}{\lambda}}{1+\frac{\varepsilon}{\lambda}} = \frac{\tau\lambda - \varepsilon}{\rho+\lambda}.$$

Thus, in both cases we arrive at (2.46).

□

The next two corollaries present two important special cases of Lemma 2.7.5.

Corollary 2.7.6. *Suppose $K_1,\ldots,K_n$ are cones in a normed vector space, $\varepsilon \in]0,1[$ and vectors $z_1,\ldots,z_n$ satisfy*

$$\left\|\sum_{i=1}^{n} z_i\right\| < \varepsilon, \quad z_i \in K_i\ (i = 1,\ldots,n), \quad \sum_{i=1}^{n}\|z_i\| = 1.$$

Then there exist vectors $\hat{z}_i$ $(i = 1,\ldots,n)$ satisfying the following conditions:

$$\sum_{i=1}^{n}\hat{z}_i = 0, \quad \sum_{i=1}^{n} d(\hat{z}_i, K_i) < \frac{\varepsilon}{1-\varepsilon}, \quad \sum_{i=1}^{n}\|\hat{z}_i\| = 1.$$

Moreover, if the underlying space is dual to a normed vector space X, and condition (2.45) is satisfied for some vectors $x_i \in X$ $(i = 1,\ldots,n)$, not all zero, and a number $\tau \in]0,1]$, then the vectors $\hat{z}_i$ $(i = 1,\ldots,n)$ satisfy condition (2.46) with $\hat{\tau} := \dfrac{\tau-\varepsilon}{1+\varepsilon}$.

Proof. Apply Lemma 2.7.5(i) with $\rho = 1$ and $\lambda = 1-\varepsilon$. □

Corollary 2.7.7. *Suppose $K_1,\ldots,K_n$ are cones in a normed vector space, $\varepsilon \in]0,1[$ and vectors $z_1,\ldots,z_n$ satisfy*

$$\sum_{i=1}^{n} z_i = 0, \quad \sum_{i=1}^{n} d(z_i, K_i) < \varepsilon, \quad \sum_{i=1}^{n}\|z_i\| = 1.$$

Then there exist vectors $\hat{z}_i$ $(i=1,\ldots,n)$ *satisfying the following conditions:*

$$\left\|\sum_{i=1}^{n}\hat{z}_i\right\| < \frac{\varepsilon}{1-\varepsilon}, \quad \hat{z}_i \in K_i\ (i=1,\ldots,n), \quad \sum_{i=1}^{n}\|\hat{z}_i\| = 1. \tag{2.47}$$

Moreover, if the underlying space is dual to a normed vector space X*, and condition* (2.45) *is satisfied for some vectors* $x_i \in X$ $(i=1,\ldots,n)$*, not all zero, and a number* $\tau \in]0,1]$*, then the vectors* $\hat{z}_i$ $(i=1,\ldots,n)$ *satisfy condition* (2.46) *with* $\hat{\tau} := \dfrac{\tau-\varepsilon}{1+\varepsilon}$.

Proof. Apply Lemma 2.7.5(ii) with $\rho = 1-\varepsilon$ and $\lambda = 1$. □

As an immediate consequence of Corollaries 2.7.6 and 2.7.7, we obtain the following important assertion.

Corollary 2.7.8. *Suppose* $\Omega_1,\ldots,\Omega_n$ *are subsets of a normed vector space* X *and* $\bar{x} \in \cap_{i=1}^{n}\Omega_i$*. Conditions* (D1) *and* (D2) *in Theorem 2.7.3 are equivalent.*

When $n=2$, the main estimate in Corollary 2.7.6 can be improved.

Proposition 2.7.9. *Suppose* K_1 *and* K_2 *are cones in a normed vector space,* $\varepsilon > 0$*, and vectors* z_1 *and* z_2 *satisfy*

$$\|z_1+z_2\| < \varepsilon, \quad z_1 \in K_1,\ z_2 \in K_2, \quad \|z_1\|+\|z_2\| = 1.$$

Then, there exist vectors $\hat{z}_1$ *and* $\hat{z}_2$ *satisfying the following conditions:*

$$\hat{z}_1+\hat{z}_2 = 0, \quad \hat{z}_1 \in K_1, \quad d(\hat{z}_2,K_2) < \varepsilon, \quad \|\hat{z}_1\|+\|\hat{z}_2\| = 1.$$

Proof. Without loss of generality, we can assume that $\|z_2\| \le \dfrac{1}{2} \le \|z_1\|$. Set $\hat{z}_1 := \dfrac{z_1}{2\,\|z_1\|}$ and $\hat{z}_2 := -\hat{z}_1$. Then $\hat{z}_1+\hat{z}_2 = 0$, $\|\hat{z}_1\| = \|\hat{z}_2\| = \dfrac{1}{2}$, $\hat{z}_1 \in K_1$ and

$$d(\hat{z}_2,K_2) \le \left\|\hat{z}_2 - \frac{z_2}{2\,\|z_1\|}\right\| = \frac{\|z_1+z_2\|}{2\,\|z_1\|} < \frac{\varepsilon}{2\,\|z_1\|} \le \varepsilon.$$

This completes the proof. □

Next we formulate an asymmetric modification of Corollary 2.7.7 which will be used in Section 3.3.

Proposition 2.7.10. *Suppose* $K_1,\ldots,K_n$ *are cones in a normed vector space,* $\varepsilon \in]0,1[$ *and vectors* $z_1,\ldots,z_n$ *satisfy conditions*

$$\sum_{i=1}^{n} z_i = 0, \quad \sum_{i=1}^{n} d(z_i,K_i) < \varepsilon, \quad \sum_{i=1}^{n-1}\|z_i\| = 1.$$

Then, there exist vectors $\hat{z}_i$ $(i=1,\ldots,n)$ *satisfying the following conditions:*

$$\sum_{i=1}^{n}\hat{z}_i = 0, \quad \hat{z}_i \in K_i\ (i=1,\ldots,n-1), \quad d(\hat{z}_n,K_n) < \frac{\varepsilon}{1-\varepsilon} \quad \text{and} \quad \sum_{i=1}^{n-1}\|\hat{z}_i\| = 1. \tag{2.48}$$

Proof. Take $y_i \in K_i$ $(i = 1, \ldots, n)$ such that $\sum_{i=1}^{n} \|z_i - y_i\| < \varepsilon$. Then $\left\|\sum_{i=1}^{n} y_i\right\| < \varepsilon$ and $\sum_{i=1}^{n-1} \|y_i\| > 1 - \varepsilon$. It follow that $d(-\sum_{i=1}^{n-1} y_i, K_n) \leq d(-\sum_{i=1}^{n-1} y_i, y_n) < \varepsilon$. Hence, vectors $\hat{z}_i := y_i / \sum_{i=1}^{n-1} \|y_i\|$ $(i = 1, \ldots, n-1)$ and $\hat{z}_n := -\sum_{i=1}^{n-1} \hat{z}_i$. satisfy all the conditions in (2.48). $\square$

Observe that, by scaling the vectors, the normalisation condition $\sum_{i=1}^{n} \|x_i^*\| = 1$ in the dual generalised separation properties (D1) and (D2) can be dropped if the inequalities there are amended to

$$\sum_{i=1}^{n} d(x_i^*, N_{\Omega_i}(\omega_i)) < \varepsilon \sum_{i=1}^{n} \|x_i^*\| \quad \text{and} \quad \left\|\sum_{i=1}^{n} x_i^*\right\| < \varepsilon \sum_{i=1}^{n} \|x_i^*\|,$$

respectively. Note that each of the amended inequalities still implies that $\sum_{i=1}^{n} \|x_i^*\| > 0$. Moreover, the normalisation condition $\sum_{i=1}^{n} \|x_i^*\| = 1$ involving n vectors can be replaced by the similar asymmetric condition involving $n-1$ vectors: $\sum_{i=1}^{n-1} \|x_i^*\| = 1$. This observation can be especially useful in the case $n = 2$.

Observe further that parameter ε in the Definition 2.3.1(iv) of approximate stationarity and its reformulation in Proposition 2.3.4(iv) as well as the metric and dual characterisations in Proposition 2.4.2(ii) and (iii) and 2.7.3(ii) and (iii) plays multiple roles. To get a deeper insight into the approximate stationarity property, it makes sense to split the parameter ε into two components. From now on, we will use the letter α to denote the component controlling the size of the shifts of the sets. This parameter is going to be crucial for quantifying the corresponding transversality property, playing the role of the rate/modulus of the property.

Based on the above observations, we now formulate a list of primal and dual equivalent characterisations of approximate stationarity, complementing Definition 2.3.1(iv), Proposition 2.3.4(iv), Proposition 2.4.2(ii) and (iii) and Theorem 2.7.3, which will be used in the sequel.

Proposition 2.7.11. *Suppose $\Omega_1, \ldots, \Omega_n$ are subsets of a normed vector space X and $\bar{x} \in \cap_{i=1}^{n} \Omega_i$. The following conditions are equivalent:*

(i) *the collection $\{\Omega_1, \ldots, \Omega_n\}$ is approximately stationary at $\bar{x}$;*

(ii) *for any $\varepsilon > 0$ and $\alpha > 0$, there exist a number $\rho \in]0, \varepsilon[$, points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ and vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying*

$$\bigcap_{i=1}^{n} (\Omega_i - \omega_i - a_i) \cap (\rho \mathbb{B}) = \emptyset \quad \textit{and} \quad \max_{1 \leq i \leq n} \|a_i\| < \alpha\rho;$$

(iii) *for any $\varepsilon > 0$ and $\alpha > 0$, there exist a number $\rho \in]0, \varepsilon[$, points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ $(i = 1, \ldots, n)$ and vectors $a_i \in X$ $(i = 1, \ldots, n-1)$ satisfying*

$$\bigcap_{i=1}^{n-1} (\Omega_i - \omega_i - a_i) \cap (\Omega_n - \omega_n) \cap (\rho \mathbb{B}) = \emptyset \quad \textit{and} \quad \max_{1 \leq i \leq n-1} \|a_i\| < \alpha\rho;$$

(iv) *for any $\varepsilon > 0$ and $\alpha > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n)$ such that*

$$\alpha d\left(\bar{x}, \bigcap_{i=1}^{n} (\Omega_i - a_i)\right) > \max_{1 \le i \le n} d\left(\bar{x}, \Omega_i - a_i\right) \quad \text{and} \quad \max_{1 \le i \le n} \|a_i\| < \varepsilon; \tag{M4.1}$$

(v) *for any $\varepsilon > 0$ and $\alpha > 0$, there exist a point $x \in B_\varepsilon(\bar{x})$ and vectors $a_i \in X$ $(i = 1, \ldots, n)$ such that*

$$\alpha d\left(x, \bigcap_{i=1}^{n} (\Omega_i - a_i)\right) > \max_{1 \le i \le n} d(x, \Omega_i - a_i) \quad \text{and} \quad \max_{1 \le i \le n} \|a_i\| < \varepsilon. \tag{M4.2}$$

With N standing for either Clarke ($N := N^C$) or Fréchet ($N := N^F$) normal cone, the following conditions are equivalent to conditions (D1) *and* (D2) *in Theorem 2.7.3:*

(vi) *for any $\varepsilon > 0$ and $\alpha > 0$, there exist points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ such that*

$$\left\|\sum_{i=1}^{n} x_i^*\right\| < \alpha, \quad x_i^* \in N_{\Omega_i}(\omega_i) \ (i = 1, \ldots, n) \quad \text{and} \quad \sum_{i=1}^{n} \|x_i^*\| = 1; \tag{D3}$$

(vii) *for any $\varepsilon > 0$ and $\alpha > 0$, there exist points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ such that*

$$\sum_{i=1}^{n} x_i^* = 0, \quad \sum_{i=1}^{n} d(x_i^*, N_{\Omega_i}(\omega_i)) < \alpha \quad \text{and} \quad \sum_{i=1}^{n} \|x_i^*\| = 1; \tag{D4}$$

(viii) *for any $\varepsilon > 0$ and $\alpha > 0$, there exist points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ such that*

$$\left\|\sum_{i=1}^{n} x_i^*\right\| < \alpha, \quad x_i^* \in N_{\Omega_i}(\omega_i) \ (i = 1, \ldots, n) \quad \text{and} \quad \sum_{i=1}^{n-1} \|x_i^*\| = 1; \tag{D5}$$

(ix) *for any $\varepsilon > 0$ and $\alpha > 0$, there exist points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ such that*

$$\sum_{i=1}^{n} x_i^* = 0, \quad \sum_{i=1}^{n} d(x_i^*, N_{\Omega_i}(\omega_i)) < \alpha \quad \text{and} \quad \sum_{i=1}^{n-1} \|x_i^*\| = 1. \tag{D6}$$

If X is Asplund and $N = N^F$, then all conditions (i)–(ix) *are equivalent.*

Remark 2.7.12. The maximum in each of the conditions in Definition 2.3.1, Proposition 2.4.2, Proposition 2.3.4 and the first (primal space) part of Proposition 2.7.11 can be replaced by the sum. Moreover, any norm on $\mathbb{R}^n$ (or $\mathbb{R}^{n-1}$ in the case of Proposition 2.3.4 and Proposition 2.7.11(iii)) can be used instead. The sum of the norms in the normalisation conditions in (D1) and (D2) and the second (dual space) part of Proposition 2.7.11 stands for the corresponding dual norm and can be replaced by the maximum, or any other norm on $\mathbb{R}^n$ or $\mathbb{R}^{n-1}$.

The "α-version" of the approximate stationarity based on its equivalent representation in Proposition 2.7.11(i) is going to be used in the subsequent study.

Definition 2.7.13 (Approximate α-stationarity). Suppose $\Omega_1, \ldots, \Omega_n$ are subsets of a normed vector space X, $\bar{x} \in \cap_{i=1}^n \Omega_i$ and $\alpha > 0$. The collection $\{\Omega_1, \ldots, \Omega_n\}$ is approximately α-stationary at $\bar{x}$ if and only if for any $\varepsilon > 0$, there exist a number $\rho \in]0, \varepsilon[$, points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ and vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying conditions (P4) with α in place of ε.

In view of the above discussion, the next assertion is straightforward.

Proposition 2.7.14. *Suppose $\Omega_1,\ldots,\Omega_n$ are subsets of a normed vector space X and $\bar{x}\in\cap_{i=1}^n\Omega_i$. The collection $\{\Omega_1,\ldots,\Omega_n\}$ is approximately stationary at $\bar{x}$ if and only if it is approximately α-stationary at $\bar{x}$ for all $\alpha>0$.*

In our deeper analysis of the core arguments in proofs of metric and dual characterisations of the extremality and stationarity in Sections 3.2 and 3.3, we will study properties which correspond to fixing, besides α, also other parameters involved in the primal and dual properties discussed above.

Theorems 2.4.2 and 2.7.3 as well as the equivalent characterisations in Proposition 2.7.11 can be 'reversed' into statements providing primal and dual space criteria for the absence of the approximate stationarity, which turns out to be an important *regularity/transversality* property of collections of sets (Definition 2.3.9), which plays an important role in constraint qualifications, qualification conditions in subdifferential/coderivative calculus and convergence analysis of computational algorithms [17,52,65,75, 90–92,100,104,106,109].

Adding condition (ii) to the list of equivalent conditions sheds additional light on the transversality property.

Theorem 2.7.15 (Transversality: dual criteria)**.** *Suppose $\Omega_1,\ldots,\Omega_n$ are closed subsets of an Asplund space X and $\bar{x}\in\cap_{i=1}^n\Omega_i$. With N standing for the Fréchet normal cone ($N:=N^F$), the following conditions are equivalent:*

(i) *the collection $\{\Omega_1,\ldots,\Omega_n\}$ is transversal at $\bar{x}$;*

(ii) *there exist numbers $\alpha>0$ and $\varepsilon>0$ such that $\left\|\sum_{i=1}^n x_i^*\right\|>\alpha$ for all points $\omega_i\in\Omega_i\cap B_\varepsilon(\bar{x})$ and vectors $x_i^*\in N_{\Omega_i}(\omega_i)$ $(i=1,\ldots,n)$ satisfying $\sum_{i=1}^n\|x_i^*\|=1$;*

(iii) *there exist numbers $\alpha>0$ and $\varepsilon>0$ such that $\sum_{i=1}^n d(x_i^*,N_{\Omega_i}(\omega_i))>\alpha$ for all points $\omega_i\in\Omega_i\cap B_\varepsilon(\bar{x})$ and vectors $x_i^*\in X^*$ $(i=1,\ldots,n)$ satisfying $\sum_{i=1}^n x_i^*=0$ and $\sum_{i=1}^n\|x_i^*\|=1$.*

In view of Theorem 2.7.15, transversality of a collection of sets is equivalent to the absence of the generalised separation.

Remark 2.7.16. (i) In view of Corollaries 2.7.6 and 2.7.7, conditions (ii) and (iii) in Theorem 2.7.15 are equivalent.

(ii) The supremum of all numbers α in part (ii) of Theorem 2.7.15 equal the constant $\mathrm{tr}[\Omega_1,\ldots,\Omega_n](\bar{x})$; see [90, Theorem 1], [92, Theorem 4(vi)] and Corollary 3.2.10 below. The supremum of all numbers α in part (iii) of Theorem 2.7.15 can be different from the constant $\mathrm{tr}[\Omega_1,\ldots,\Omega_n](\bar{x})$, but its relationship with $\mathrm{tr}[\Omega_1,\ldots,\Omega_n](\bar{x})$ can be easily established using elementary arguments discussed in the next subsection.

Example 2.7.17. In the space $\mathbb{R}^2$ equipped with the maximum norm (hence, the dual norm is the sum norm), consider the two perpendicular lines: $\Omega_1:=\{(t,0): t\in\mathbb{R}\}$ and $\Omega_2:=\{(0,t): t\in\mathbb{R}\}$. Then we have $\bar{x}:=(0,0)\in\Omega_1\cap\Omega_2$, $N^F_{\Omega_1}(x)=\{(0,t): t\in\mathbb{R}\}$ for any $x\in\Omega_1$ and $N^F_{\Omega_1}(x)=\{(t,0): t\in\mathbb{R}\}$ for any $x\in\Omega_2$.

If $x_1^*=(0,t_1)$ and $x_2^*=(t_2,0)$ are normal vectors to Ω_1 and Ω_2, respectively, then $\|x_1^*\|+\|x_2^*\|=\|x_1^*+x_2^*\|=|t_1|+|t_2|$. Hence, the supremum of all α in part (ii) of Theorem 2.7.15 is 1 (and is equal to $\mathrm{tr}[\Omega_1,\Omega_2]$).

If $x_1^*, x_2^* \in (\mathbb{R}^2)^*$, $x_1^* + x_2^* = 0$ and $\|x_1^*\| + \|x_2^*\| = 1$, then $x_1^* = -x_2^* = (t_1, t_2)$ for some numbers t_1 and t_2 satisfying $|t_1| + |t_2| = \frac{1}{2}$, and the distances from x_1^* and x_2^* to the corresponding normal cones equal $|t_1|$ and $|t_2|$, respectively. Hence, the supremum of all α in part (iii) of Theorem 2.7.15 is $\frac{1}{2}$.

Remark 2.7.18. Since the approximate stationarity and transversality properties are complementary to each other, it would be natural to refer to the negation of the approximate α-stationarity property, i.e., the existence of a number $\varepsilon > 0$ such that condition (P11) for all $\rho \in]0, \varepsilon[$, points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ and vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying $\max\limits_{1 \le i \le n} \|a_i\| < \alpha\rho$, as *$\alpha$-transversality* at $\bar{x}$.

2.7.2 More Extensions

As it was pointed out in the overview, the key tool used in the proof of Theorem 2.7.3 is the Ekeland variational principle. The next natural step in the extremal principle refinement process is to single out the core part of the conventional proof of the extremal principle around the application of the Ekeland variational principle, identify the minimal assumptions on the sets and the immediate conclusions and formulate it as a separate statement. Such a result (results) would expose the core arguments behind the extremal principle and could serve as a key building block when constructing other generalised separation statements, applicable in situations where the conventional (extended) extremal principle fails.

We are aware of two recent attempts of this kind: [97, Theorem 3.1] which served as a tool when extending Theorems 2.7.3 to infinite collections of sets, and [170, Lemmas 2.1 and 2.2] used when proving fuzzy multiplier rules in set-valued optimisation problems. The last couple of lemmas have been further refined and strengthened in [172, Theorems 3.1 and 3.4] and [174, Theorem 1.1].

The next two theorems are reformulations for the setting adopted in the current paper of [97, Theorem 3.1] and [172, Theorem 3.4], respectively.

Theorem 2.7.19 (Kruger and López, 2012)**.** *Suppose $\Omega_1, \ldots, \Omega_n$ are closed sets of an Asplund space, and $\bar{x} \in \cap_{i=1}^n \Omega_i$, $\alpha > 0$.*

(i) *If the numbers $\varepsilon > 0, \varepsilon_1, \varepsilon_2 > 0, \varepsilon_1 + \varepsilon_2 \le \varepsilon$; $\rho \in]0, \varepsilon_2/(\alpha+1)[$, and points $\omega_i \in \Omega_i$ and $a_i \in X$ $(i = 1, \ldots, n)$ are given such that $\Omega_i \cap B_\varepsilon(\bar{x})$ and $\max\limits_{1 \le i \le n} \|a_i\| < \alpha\rho$, and* (P4) *holds, then there exist points $\omega_i' \in \Omega_i \cap B_{\varepsilon_2}(\bar{x})$ and $x_i^* \in N_{\Omega_i}(\omega_i')$ satisfying* (D3)*.*

(ii) *If $\omega_i \in \Omega_i$ and $x_i^* \in N_{\Omega_i}(\omega_i)$ $(i = 1, \ldots, n)$ satisfy conditions* (D3)*, then, for any $\delta > 0$, there exists a $\rho \in]0, \delta[$ and points $a_i \in X$ $(i = 1, \ldots, n)$ satisfying conditions* (P4)*.*

Theorem 2.7.20 (Zheng and Ng, 2011)**.** *Suppose $\Omega_1, \ldots, \Omega_n$ are closed sets of an Asplund space, and $\cap_{i=1}^n \Omega_i = \emptyset$, $\varepsilon > 0$. If points $\omega_i \in \Omega_i$ $(i = 1, \ldots, n)$ satisfy condition* (2.40)*, then, for any $\lambda > 0$ and $\tau \in (0, 1)$, there exist points $\omega_i' \in \Omega_i \cap \mathbb{B}_\lambda(\omega_i)$ $(i = 1, \ldots, n)$ and $x_i^* \in X^*$ $(i = 1, \ldots, n)$ such that*

$$\sum_{i=1}^{n-1} \|x_i^*\| = 1, \quad \sum_{i=1}^{n} x_i^* = 0, \quad \sum_{i=1}^{n} d(x_i^*, N_{\Omega_i}(\omega_i')) < \varepsilon/\lambda, \tag{2.49}$$

$$\tau \sum_{i=1}^{n-1} \|\omega_i' - \omega_n'\| \le \sum_{i=1}^{n-1} \langle x_i^*, \omega_i' - \omega_n' \rangle. \tag{2.50}$$

First observe that the extremal principle in Theorem 2.7.3 is a direct corollary of Theorem 2.7.19.

Theorem 2.7.3 *from Theorem* 2.7.19. Let the collection $\{\Omega_1, \ldots, \Omega_n\}$ be approximately stationary at $\bar{x}$. We are going to show that condition (D1) in Theorem 2.7.3 holds true. Given an $\varepsilon > 0$, find an $\varepsilon' > 0$ such that $\varepsilon'(\varepsilon' + 2) < \varepsilon$. By Definition 2.3.1(iv), there exist $\rho \in]0, \varepsilon'[$, $\omega_i \in \Omega_i \cap \mathbb{B}_{\varepsilon'}(\bar{x})$ and $a_i \in X$

$(i=1,\ldots,n)$ such that conditions (P4) are satisfied with ε' in place of ε. Then $\max\limits_{1\le i\le n}\|\omega_i-\bar{x}\|+\rho(\varepsilon'+1) < \varepsilon'+\varepsilon'(\varepsilon'+1) < \varepsilon$, and it follows from Theorem 2.7.19(i) that there exist points $\omega_i' \in \Omega_i \cap \mathbb{B}_\varepsilon(\bar{x})$ and $x_i^* \in N_{\Omega_i}(\omega_i')$ $(i=1,\ldots,n)$, satisfying (D1).

Conversely, let condition (D1) in Theorem 2.7.3 holds true and an $\varepsilon > 0$ be given. Then there exist points $\omega_i \in \Omega_i \cap \mathbb{B}_\varepsilon(\bar{x})$, $x_i^* \in N_{\Omega_i}(\omega_i)$ $(i=1,\ldots,n)$ satisfying conditions (D1). By Theorem 2.7.19(ii), there exists a $\rho \in]0,\varepsilon[$ and points $a_i \in X$ $(i=1,\ldots,n)$ satisfying conditions (P4), i.e. the collection $\{\Omega_1,\ldots,\Omega_n\}$ is approximately stationary at $\bar{x}$. □

Next we compare the statements of Theorem 2.7.19(i) and Theorem 2.7.20. There are important similarities between them: both establish a kind of generalised separation of the two sets, related somehow to the given pair of points $a \in A$ and $b \in B$ possessing a certain approximate 'extremality' property. There are also essential differences.

We start with comparing the assumptions in the two statements. On the first glance, they look mutually exclusive: the first one assumes the existence of a point $\bar{x} \in \cap_{i=1}^n \Omega_i$, while in the second theorem, it is assumed on the contrary that $\cap_{i=1}^n \Omega_i = \emptyset$. However, this distinction is easy to overcome. Given points $\omega_i \in \Omega_i$ $(i=1,\ldots,n)$ in Theorem 2.7.20, one can set $\Omega_i' := \Omega_i - \omega_i$ $(i=1,\ldots,n)$; then $\bar{x} := 0 \in \cap_{i=1}^n(\Omega_i - \omega_i')$ (this trick is used in the proof of Theorem 2.7.20$'$ below). This observation exposes also the different roles played by the pairs $a \in A$ and $b \in B$ in Theorem 2.7.19(i) and Theorem 2.7.20. In the first one, these are actually additional parameters having no analogues in Theorem 2.7.20, which corresponds to $\omega_1 = \ldots = \omega_n = \bar{x}$ in Theorem 2.7.19(i).

The second distinction is related to the main approximate 'extremality' assumptions on the pair of sets: conditions (P4) in Theorem 2.7.19(i) and condition (2.40) in Theorem 2.7.20. The Proposition 2.6.6 shows that condition (2.40) implies a stronger version of conditions (P4).

Proposition 2.6.6 is not reversible: condition (P4) being satisfied with some small $a_i \in X$ $(i=1,\ldots,n)$ does not imply that $d_1(\omega_1,\ldots,\omega_n)$ is close to the distance $d_1(\Omega_1,\ldots,\Omega_n)$ between the n sets.

Example 2.7.21. Let $A := \{(x_1,x_2) \in \mathbb{R}^2 \mid x_2 \le 0\}$ and $B := \{(x_1,x_2) \in \mathbb{R}^2 \mid x_2 \ge 1\}$. Then, assuming that $\mathbb{R}^2$ is equipped with e.g. the sum norm, $d(A,B) = 1$. If $a := (\alpha,0) \in A$ and $b := (\beta,1) \in B$ with some $\alpha,\beta \in \mathbb{R}$, then condition

$$(A - a - u) \cap (B - b - v) \cap (\rho\mathbb{B}) = \emptyset$$

is satisfied with $u := (0,\varepsilon)$, $v := (0,-\varepsilon)$ and any $\varepsilon > 0$. At the same time, $\|a-b\| = |\alpha-\beta|+1$ can be arbitrarily large when the numbers α and β are far apart.

Thus, condition (P4) with small a_i is less restrictive than condition (2.40). Moreover, the first condition in (2.3.1) with $\rho < \infty$ is weaker than (P4) and allows for local versions of the corresponding properties.

The next assertion is immediate from Proposition 2.6.6.

Corollary 2.7.22. *Suppose X is a normed linear space, $\Omega_i \subset X$ $(i=1,\ldots,n)$, $\cap_{i=1}^n \Omega_i = \emptyset$. If sequences $\{\omega_{i,k}\} \subset \Omega_i$ $(i=1,\ldots,n)$ are such that $\|\omega_{i,k}-\omega_{n,k}\| \to d(\Omega_i,\Omega_n)$, then there exist sequences $\{a_{i,k}\} \subset X$ $(i=1,\ldots,n)$ converging to 0, such that*

$$\bigcap_{i=1}^n (\Omega_i - \omega_{i,k} - a_{i,k}) = \emptyset.$$

With the sets in Example 2.7.21, one can easily see that the statement of Corollary 2.7.22 is not reversible.

Now we are going to compare the conclusions of the two theorems. Similarly to the two conditions in Theorem 2.7.3, they represent two different ways of formulating dual extremality/separation conditions:

in terms of normal (in Theorem 2.7.19(i)) or ‘almost normal’ (in Theorem 2.7.20) vectors, with the connection between the two formulations provided by Lemma 2.7.5. However, unlike the two equivalent conditions in Theorem 2.7.3 formulated ‘for any $\varepsilon > 0$’, in both Theorem 2.7.19(i) and Theorem 2.7.20 the number $\varepsilon > 0$ is a given quantitative parameter. Corollary 2.7.6 and (2.7.7) used in the proof of the equivalence of the two conditions in Theorem 2.7.3 cannot provide one-to-one translation between the two settings with the given $\varepsilon > 0$; it only gives estimates, and its application leads to some ‘loss of accuracy’. Note that the proof of [97, Theorem 3.1], where Theorem 2.7.19 is taken from, contains estimates in terms of ‘almost normal’ vectors and then employs the arguments used in the proof of Corollary (2.7.7) to ensure that the vectors belong to the normal cones. To make a fair comparison, one needs to either reformulate Theorem 2.7.20 in terms of normal vectors using Corollary 2.7.7, or extract the pre-Corollary 2.7.7 statement from the proof of [97, Theorem 3.1]. Below for simplicity we follow the first approach. The next statement is a consequence of Theorem 2.7.20 and Corollary 2.7.7.

Theorem 2.7.20′ *Suppose X is an Asplund space, $\Omega_1,\ldots,\Omega_n \subset X$ $(i = 1,\ldots,n)$ are closed, $\cap_{i=1}^n \Omega_i = \emptyset$. If points $\omega_i \in \Omega_i$ $(i = 1,\ldots,n)$ satisfy condition (2.40) with some $\varepsilon \in]0,1[$, then, for any $\lambda > 0$, there exist points $\omega_i' \in \Omega_i \cap \mathbb{B}_{\lambda(1-\varepsilon)}(\omega_i)$ and $x_i^* \in N_{\Omega_i}(\omega_i')$ $(i = 1,\ldots,n)$ such that*

$$\sum_{i=1}^n \|x_i^*\| = 1 \quad \text{and} \quad \left\|\sum_{i=1}^n x_i^*\right\| < \varepsilon/\lambda. \tag{2.51}$$

Proof. Given $\varepsilon > 0$ and $\lambda > 0$, set $\lambda' := \lambda(1-\varepsilon)$. Thanks to Theorem 2.7.20, there exist points $\omega_i' \in \Omega_i \cap \mathbb{B}_{\lambda'}(\omega_i)$ and $x_i^* \in X^*$ $(i = 1,\ldots,n)$ satisfying conditions (2.49) with λ' in place of λ. Then, (2.51) follows thanks to Corollary 2.7.7. $\square$

Now the comparison is straightforward.

Proposition 2.7.23. *Theorem* 2.7.19(i) *implies Theorem* 2.7.20′.

Proof. Under the conditions of Theorem 2.7.20′, set $\Omega_i' := \Omega_i - \omega_i$ $(i = 1,\ldots,n)$. Then $0 \in \cap_{i=1}^n \Omega_i'$. By Proposition 2.6.6, there exist $a_i \in X$ $(i = 1,\ldots,n-1)$ such that $\max\limits_{1\le i\le n-1} \|a_i\| < \varepsilon(1-\varepsilon)$ and $\cap_{i=1}^n (\Omega_i' - a_i) = \emptyset$. Choose an $\varepsilon' \in (2\|u\|,\varepsilon)$ and set $\rho := (\varepsilon-1)\lambda$, $\hat{\varepsilon} := \varepsilon'/\lambda$ and $\delta := (\varepsilon+\lambda)/2$. Then $\hat{\varepsilon} < \varepsilon/\lambda$ and $\delta > \rho(\hat{\varepsilon}+1)$. By Theorem 2.7.19(i) applied to the sets Ω_i' $(i = 1,\ldots,n)$ at 0, there exist points $\omega_i' \in \Omega_i \cap \mathbb{B}_\delta(\omega_i)$, $x_i^* \in N_{\Omega_i}(\omega_i')$ $(i = 1,\ldots,n)$ satisfying conditions (2.51). $\square$

Thus, Theorem 2.7.20′ is a special case of Theorem 2.7.19(i). On the other hand, as demonstrated in [170,172], Theorem 2.7.20 (as well as its version formulated above as Theorem 2.7.20′) is sufficient for many important applications. Next we show that Theorem 2.7.20′ implies the nonlocal version of the extremal principle.

Corollary 2.7.24 (Nonlocal extremal principle)**.** *Suppose X is an Asplund space, $\Omega_1,\ldots,\Omega_n \subset X$ $(n \ge 2)$ are closed and $\bar{x} \in \cap_{i=1}^n \Omega_i$. If the collection $\{\Omega_1,\ldots,\Omega_n\}$ is extremal at $\bar{x}$, then the two equivalent conditions in Theorem 2.7.3 hold true.*

Proof. Let the collection $\{\Omega_1,\ldots,\Omega_n\}$ be extremal and a number $\varepsilon > 0$ be given. Choose an $\varepsilon' \in \left]0, \dfrac{\varepsilon^2}{\varepsilon+1}\right[$. Then $\dfrac{\varepsilon'}{\varepsilon} < \varepsilon - \varepsilon'$ and we can choose a λ such that $\dfrac{2\varepsilon'}{\varepsilon} < \lambda < 2(\varepsilon-\varepsilon')$. There exist vectors $a_i \in X$ $(i = 1,\ldots,n)$ satisfying conditions (P1) with ε' in place of ε. Define $\Omega_i' := \Omega_i - a_i$ and $\omega_i := \bar{x} - a_i$ $(i = 1,\ldots,n)$. Then $\cap_{i=1}^n \Omega_i' = \emptyset$ and $d_1(\omega_1,\ldots,\omega_n) < 2\varepsilon'$. Applying Theorem 2.7.20′, we find points $\omega_i' \in \Omega_i$ $(i = 1,\ldots,n)$ and $x_i^* \in N_{\Omega_i'}(\omega_i' - a_i) = N_{\Omega_i}(\omega_i')$ $(i = 1,\ldots,n)$ such that

$\max\limits_{1\le i\le n} \|\omega_i' - \bar{x}\| < \varepsilon' + \lambda/2 < \varepsilon$, $\sum\limits_{i=1}^{n} \|x_i^*\| = 1$ and $\left\|\sum\limits_{i=1}^{n} x_i^*\right\| < 2\varepsilon'/\lambda < \varepsilon$. Thus, condition (D2) in Theorem 2.7.3 is satisfied. □

Remark 2.7.25. 1. It is not difficult to modify the proof of Corollary 2.7.24 to cater for the relaxed version of nonlocal extremality without the assumption $\cap_{i=1}^{n} \Omega_i \neq \emptyset$.

2. Theorem 2.7.20 does not seem to be able to recapture the full local extremal principle (as in Theorem 2.7.1), not to say the extended extremal principle (as in Theorem 2.7.3).

3. Condition (2.50) in Theorem 2.7.20 determining the 'direction' of the vector x_i^* does not have a direct analogue in the statement of Theorem 2.7.19(i). Together with the first condition in (2.49), it comes from subdifferentiating a norm at a nonzero point in the proof of [172, Theorem 3.4]. Subdifferentiating a norm is an essential component also in the proofs of the conventional extremal principle and all its modifications, including the one in [97, Theorem 3.1]; so analogues of (2.50) are implicitly present in all such proofs. Zheng and Ng [172] seem to be the first to notice the importance of conditions like (2.50) for recapturing the classical convex separation theorem, and make (2.50) explicit in the statement of [172, Theorem 3.4]. In the current paper, keeping in line with the conventional formulations and for the sake of simplicity of the presentation, we will not formulate analogues of the condition (2.50) in the subsequent statements.

Chapter 3

Linear Characterisations

The proof of the conventional extremal principle and all its subsequent extensions is based on the two fundamental results of variational analysis:

- Ekeland variational principle,
- a sum rule for the appropriate subdifferential.

The definitions of all four extremality/stationarity properties involve non-intersection of certain collections of sets being small (in some sense) translations of the original sets (see Definition 2.3.1), and the Ekeland variational principle and the subdifferential sum rule in the proof of each version of the extremal principle are applied to certain functions constructed on these non-intersecting sets.

In this chapter, we refine again the original proof of the conventional extremal principle, and conduct a systematic study of the non-intersection properties involved in all four parts of Definition 2.3.1, producing a series of elementary generalised separation statements, clarifying the relationships between them and, particularly, unifying the statements from [170, 172, 174] and [97].

3.1 Geometric Ekeland Variational Principle

In this section, we are going to use the next two geometric versions of the Ekeland variational principle. They characterise the mutual arrangement of a pair of sets in a complete metric space with respect to a pair of points, being almost (up to ε) closest points of these sets and, similarly to the conventional Ekeland variational principle, establish the existence of another pair of points arbitrarily close (up to an additional parameter or a pair of parameters) to the given one and minimising a certain perturbed function. The perturbed functions in both assertions involve the distance between localisations of the sets near these points, i.e. intersections of the sets with neighbourhoods of the points.

The following classical result due to Ekeland [58] (see also [51, 75, 120, 129, 130]) plays the key role in the subsequent studies.

Lemma 3.1.1 (Ekeland Variational Principle **(EVP)**). *Suppose X is a complete metric space, $f : X \to \mathbb{R}_\infty$ is lower semicontinuous, $\bar{x} \in X$ and $\varepsilon > 0$. If*

$$f(\bar{x}) < \inf_X f + \varepsilon,$$

then, for any $\lambda > 0$, there exists an $\hat{x} \in X$ such that

(i) $d(\hat{x}, \bar{x}) < \lambda$;

(ii) $f(\hat{x}) \leq f(\bar{x})$;

(iii) $f(x) + (\varepsilon/\lambda)d(x, \hat{x}) > f(\hat{x})$ *for all* $x \in X \setminus \{\bar{x}\}$.

Theorem 3.1.2 (Geometric Ekeland Variational Principle **(GEVP)**)**.** *Suppose* X *is a complete metric space,* A *and* B *are closed subsets of* X*,* $a \in A$*,* $b \in B$*, and* $\varepsilon > 0$*. If*

$$d(a, b) < d(A, B) + \varepsilon, \tag{3.1}$$

then, for any $\lambda > 0$*, there exist* $\hat{a} \in A \cap B_\lambda(a)$ *and* $\hat{b} \in B \cap B_\lambda(b)$ *such that*

(i) $d(\hat{a}, \hat{b}) \leq d(a, b)$;

(ii) $d(A \cap B_\xi(\hat{a}), B \cap B_\xi(\hat{b})) + \dfrac{\xi\varepsilon}{\lambda} > d(\hat{a}, \hat{b})$ *for all* $\xi > 0$.

Theorem 3.1.3 (Asymmetric Geometric Ekeland Variational Principle **(AGEVP)**)**.** *Suppose* X *is a complete metric space,* A *and* B *are closed subsets of* X*,* $a \in A$*,* $b \in B$*, and* $\varepsilon > 0$*. If condition* (3.1) *is satisfied, then, for any* $\lambda, \rho > 0$*, there exist* $\hat{a} \in A \cap B_\lambda(a)$ *and* $\hat{b} \in B \cap B_\rho(b)$ *such that*

(i) $d(\hat{a}, \hat{b}) \leq d(a, b)$;

(ii) $d(A \cap B_{\xi\lambda}(\hat{a}), B \cap B_{\xi\rho}(\hat{b})) + \xi\varepsilon > d(\hat{a}, \hat{b})$ *for all* $\xi > 0$.

Unlike GEVP, its asymmetric version AGEVP allows for balls with different radii to be used in the localisations of the sets. This feature is going to play an important role in our subsequent analysis. It is easy to see that GEVP is a particular case of AGEVP with $\rho = \lambda$. We now show that the two geometric versions of the Ekeland variational principle formulated above are both equivalent to the conventional one.

Proposition 3.1.4. *EVP* $\Leftrightarrow$ *GEVP* $\Leftrightarrow$ *AGEVP.*

Proof. We first show that EVP $\Rightarrow$ AGEVP. Let the assumptions of AGEVP be satisfied. Given numbers $\lambda, \rho > 0$, we consider the space $X \times X$ with a metric defined as follows:

$$d_{\lambda,\rho}((x, y), (u, v)) := \max\left\{\frac{1}{\lambda}d(x, u), \frac{1}{\rho}d(y, v)\right\} \quad (x, y, u, v \in X). \tag{3.2}$$

Since X is complete, $(A \times B, d_{\lambda,\rho})$ is a complete space. Choose an $\varepsilon' \in]0, \varepsilon[$ such that (3.1) is satisfied with ε' in place of ε. EVP applied to the function d on $(A \times B, d_{\lambda,\rho})$ gives the existence of points $\hat{a} \in A$ and $\hat{b} \in B$ such that

$$d_{\lambda,\rho}((\hat{a}, \hat{b}), (a, b)) < 1, \quad d(\hat{a}, \hat{b}) \leq d(a, b), \tag{3.3}$$

$$d(x, y) - d(\hat{a}, \hat{b}) + \varepsilon' d_{\lambda,\rho}((x, y), (\hat{a}, \hat{b})) \geq 0 \quad \text{for all} \quad (x, y) \in A \times B. \tag{3.4}$$

In view of the definition (3.2), the first inequality in (3.3) is equivalent to the following two: $d(\hat{a}, a) < \lambda$ and $d(\hat{b}, b) < \rho$. Given any $\xi > 0$, $x \in A \cap B_{\xi\lambda}(\hat{a})$ and $y \in B \cap B_{\xi\rho}(\hat{b})$, by (3.2) and (3.4), we have, respectively, $d_{\lambda,\rho}((x, y), (\hat{a}, \hat{b})) < \xi$ and

$$d(x, y) \geq d(\hat{a}, \hat{b}) - \varepsilon' d_{\lambda,\rho}((x, y), (\hat{a}, \hat{b})) > d(\hat{a}, \hat{b}) - \xi\varepsilon'.$$

It follows that $d(A \cap B_{\xi\lambda}(\hat{a}), B \cap B_{\xi\rho}(\hat{b})) \geq d(\hat{a}, \hat{b}) - \xi\varepsilon' > d(\hat{a}, \hat{b}) - \xi\varepsilon$. This proves AGEVP.

The implication AGEVP $\Rightarrow$ GEVP is straightforward.

To complete the proof, we next show that GEVP $\Rightarrow$ EVP. Let the assumptions of EVP be satisfied. Choose a positive number $\varepsilon' < \varepsilon$ such that $f(\bar{x}) < \inf_X f + \varepsilon'$ and another number α such that $0 < \alpha < \lambda(\varepsilon'^{-1} - \varepsilon^{-1})$. We are going to consider the space $X \times \mathbb{R}$ with the metric $d := d_X + \alpha|\cdot|$, which makes $X \times \mathbb{R}$ a complete metric space, two closed subsets of $X \times \mathbb{R}$: $A := \{(x, y) : x \in X,\ y \geq f(x)\}$ and $B := X \times \{M\}$, where $M := \inf_X f$, and two points $a := (\bar{x}, f(\bar{x})) \in A$ and $b := (\bar{x}, M) \in B$. We have $d(A, B) = 0$ and

$$d(a, b) = \alpha(f(\bar{x}) - M) < \alpha\varepsilon' = d(A, B) + \alpha\varepsilon'.$$

GEVP gives the existence of points $\hat{a} = (\hat{x}, \hat{y}) \in A \cap B_\lambda(a)$ and $\hat{b} = (\hat{x}', M) \in B \cap B_\lambda(b)$ satisfying

$$d(\hat{a}, \hat{b}) \leq d(a, b), \tag{3.5}$$

$$d(A \cap B_\xi(\hat{a}), B \cap B_\xi(\hat{b})) + \frac{\xi\alpha\varepsilon'}{\lambda} > d(\hat{a}, \hat{b}) \quad \text{for all} \quad \xi > 0. \tag{3.6}$$

Observe that $d(\hat{x}, \bar{x}) < \lambda$, and consequently, $(\hat{x}, M) \in B \cap B_\lambda(b)$. Moreover, $\hat{x}' = \hat{x}$. Indeed, if $\hat{x} \neq \hat{x}'$, then we can take $\xi := d(\hat{x}, \hat{x}')$. Then $(\hat{x}, \hat{y}) \in A \cap B_\xi(\hat{a})$, $(\hat{x}, M) \in B \cap \overline{B}_\xi(\hat{b})$ and

$$d((\hat{x}, \hat{y}), (\hat{x}, M)) + \frac{\xi\alpha\varepsilon'}{\lambda} < \alpha(\hat{y} - M) + \left(1 - \frac{\varepsilon'}{\varepsilon}\right) d(\hat{x}, \hat{x}') < \alpha(\hat{y} - M) + d(\hat{x}, \hat{x}') = d(\hat{a}, \hat{b}),$$

which contradicts (3.6); hence $\hat{x}' = \hat{x}$ and $\hat{b} = (\hat{x}, M)$. Consequently, condition (3.5) reduces to $\alpha(\hat{y} - M) \leq \alpha(f(\bar{x}) - M)$, which implies that $f(\hat{x}) \leq \hat{y} \leq f(\bar{x})$.

It remains to prove condition (iii) in EVP. Let $x \neq \hat{x}$. If $f(x) \geq f(\hat{x})$, the condition holds trivially. Let $f(x) < f(\hat{x})$, and take $\xi := d(x, \hat{x}) + \alpha(\hat{y} - f(x))$. Then $(x, f(x)) \in A \cap \overline{B}_\xi(\hat{a})$, $(x, M) \in B \cap B_\xi(\hat{b})$, and by (3.6),

$$\alpha(f(x) - M) + \frac{\alpha\varepsilon'}{\lambda}(d(x, \hat{x}) + \alpha(\hat{y} - f(x))) = d((x, f(x)), (x, M)) + \frac{\xi\alpha\varepsilon'}{\lambda} \geq d(\hat{a}, \hat{b}) = \alpha(\hat{y} - M).$$

Thus,

$$f(x) + \frac{\varepsilon'}{\lambda}(d(x, \hat{x}) + \alpha(\hat{y} - f(x))) \geq \hat{y},$$

or equivalently,

$$f(x) + \frac{\varepsilon'}{\lambda - \alpha\varepsilon'} d(x, \hat{x}) \geq \hat{y}.$$

By the definition of α, we have $\lambda - \alpha\varepsilon' > \lambda\varepsilon'\varepsilon^{-1}$. Hence,

$$f(x) + \frac{\varepsilon}{\lambda} d(x, \hat{x}) > f(x) + \frac{\varepsilon'}{\lambda - \alpha\varepsilon'} d(x, \hat{x}) \geq \hat{y} \geq f(\hat{x}).$$

The proof is complete. □

Condition (ii) in GEVP and AGEVP corresponds to condition (iii) in the conventional EVP, while condition (i) corresponds to the pair of conditions (i) and (ii). The ε-closeness condition (3.1) is going to play an important role in our analysis. It is discussed in Sections 2.7.2 and further in 3.3.3.

3.2 Localisations and Translations of the Sets

In this section, we continue studying mutual arrangement of n sets in space started in Section 3.1.

The next lemma extends the Asymmetric Geometric Ekeland Variational Principle (AGEVP) from Section 3.1 to the case of $n \geq 2$ sets, and as such it is also equivalent to the Ekeland Variational Principle. Just like in AGEVP, balls of two different radii are used in the concluding part of the lemma, one of them employed in the localisations of the first $n-1$ sets and the other one for the remaining single set. The latter set is going to play a special role in the subsequent analysis.

The lemma translates ε-closeness of a given collection of points into $\xi\varepsilon$-closeness (with an additional parameter ξ) of another collection of points with respect to a collection of certain localisations of the sets (depending on ξ).

Lemma 3.2.1. *Suppose $\Omega_1,\ldots,\Omega_n$ are closed subsets of a complete metric space X, $\omega_i \in \Omega_i$ $(i = 1,\ldots,n)$, and $\varepsilon > 0$. If condition (2.40) is satisfied, then, for all numbers $\lambda,\rho > 0$, there exist $\hat{\omega}_i \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1,\ldots,n-1)$ and $\hat{\omega}_n \in \Omega_n \cap B_\rho(\omega_n)$ such that*

(i) $d_1(\hat{\omega}_1,\ldots,\hat{\omega}_n) \le d_1(\omega_1,\ldots,\omega_n)$;

(ii) $d_1(\Omega_1 \cap B_{\xi\lambda}(\hat{\omega}_1),\ldots,\Omega_{n-1} \cap B_{\xi\lambda}(\hat{\omega}_{n-1}),\Omega_n \cap B_{\xi\rho}(\hat{\omega}_n)) + \xi\varepsilon > d_1(\hat{\omega}_1,\ldots,\hat{\omega}_n)$ *for all* $\xi > 0$.

Proof. The product space X^{n-1} considered with the maximum metric is complete. Set $A := \Omega_1 \times \ldots \times \Omega_{n-1}$, and $B := \{(x,\ldots,x) : x \in \Omega_n\} \subset X^{n-1}$, $a := (\omega_1,\ldots,\omega_{n-1}) \in A$, $b := (\omega_n,\ldots,\omega_n) \in B$. Then $d_1(\Omega_1,\ldots,\Omega_n) = d(A,B)$ and $d_1(\omega_1,\ldots,\omega_n) = d(a,b)$.

Applying AGEVP, we find points

$$\hat{a} = (\hat{\omega}_1,\ldots,\hat{\omega}_{n-1}) \in A \cap B_\lambda(a), \quad \text{and} \quad \hat{b} = (\hat{\omega}_n,\ldots,\hat{\omega}_n) \in B \cap B_\rho(b)$$

satisfying conditions (i) and (ii) in AGEVP. Recalling that the maximum metric is used in X^{n-1}, it follows that $\hat{\omega}_i \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1,\ldots,n-1)$, $\hat{\omega}_n \in B \cap B_\rho(\omega_n)$ and conditions (i) and (ii) above are satisfied. □

In view of the equalities in (2.39), the following corollary involving the symmetric distance d_2 is immediate.

Corollary 3.2.2. *Suppose $\Omega_1,\ldots,\Omega_n$ are closed subsets of a complete metric space X, $\omega_i \in \Omega_i$ $(i = 1,\ldots,n)$, $x \in X$ and $\varepsilon > 0$. If condition (2.40) is satisfied, then, for all $\lambda,\rho > 0$, there exist $\hat{\omega}_i \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1,\ldots,n)$ and $\hat{x} \in B_\rho(x)$ such that*

(i) $\max\limits_{1\le i\le n} d(\hat{\omega}_i,\hat{x}) \le \max\limits_{1\le i\le n} d(\omega_i,x)$;

(ii) $d_{B_{\xi\rho}(\hat{x})}(\Omega_1 \cap B_{\xi\lambda}(\hat{\omega}_1),\ldots,\Omega_n \cap B_{\xi\lambda}(\hat{\omega}_n)) + \xi\varepsilon > \max\limits_{1\le i\le n} d(\hat{\omega}_i,\hat{x})$ *for all* $\xi > 0$.

The next proposition is a consequence of Lemma 3.2.1 in the Banach space setting. It characterises ε-closest points of a collection of sets with empty intersection and involves localisations of the sets and small (up to $\xi\varepsilon$) translations of the first $n-1$ localisations. The proposition transforms the nonlocal non-intersection condition $\cap_{i=1}^n \Omega_i = \emptyset$ into a non-intersection condition of translated localisations of the sets. It also exposes the special role played by the last set in the list.

Proposition 3.2.3. *Suppose $\Omega_1,\ldots,\Omega_n$ are closed subsets of a Banach space X, $\omega_i \in \Omega_i$ $(i = 1,\ldots,n)$, and $\varepsilon > 0$. If $\cap_{i=1}^n \Omega_i = \emptyset$ and condition (2.40) is satisfied, then, for any $\lambda,\rho > 0$, there are points $\hat{\omega}_i \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1,\ldots,n-1)$, $\hat{\omega}_n \in \Omega_n \cap B_\rho(\omega_n)$ such that, for any $\xi > 0$, there exist vectors $a_i \in X$ $(i = 1,\ldots,n-1)$ satisfying*

$$\bigcap_{i=1}^{n-1} \Big(\big((\Omega_i - \hat{\omega}_i) \cap (\xi\lambda)\mathbb{B} \big) - a_i \Big) \cap (\Omega_n - \hat{\omega}_n) \cap (\xi\rho)\mathbb{B} = \emptyset \quad \textit{and} \quad \max_{1\le i\le n-1} \|a_i\| < \xi\varepsilon. \tag{3.7}$$

Proof. Applying Lemma 3.2.1, we find points $\hat{\omega}_i \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1,\ldots,n-1)$ and $\hat{\omega}_n \in \Omega_n \cap B_\rho(\omega_n)$ satisfying condition (ii) in that lemma. Given any number $\xi > 0$, we can now apply Proposition 2.6.6

with the function d_1 and sets $\Omega_1 \cap B_{\xi\lambda}(\hat{\omega}_1)$,..., $\Omega_{n-1} \cap B_{\xi\lambda}(\hat{\omega}_{n-1})$, $\Omega_n \cap B_{\xi\rho}(\hat{\omega}_n)$, points $\hat{\omega}_i$ $(i = 1, \ldots, n)$ and number $\xi\varepsilon$ in place of sets Ω_1,..., Ω_n, points ω_i $(i = 1, \ldots, n)$ and number ε, respectively, to find vectors $a_i \in X$ $(i = 1, \ldots, n-1)$ such that condition (3.7) is satisfied. □

Remark 3.2.4. (i) Under the conditions of Proposition 3.2.3, $0 \in \cap_{i=1}^n (\Omega_i - \omega_i)$, $0 \in \cap_{i=1}^n (\Omega_i - \hat{\omega}_i)$, and the expressions involved in the first condition in (3.7) correspond to localisations of the sets $\Omega_i - \hat{\omega}_i$ $(i = 1, \ldots, n)$ near 0, or equivalently, localisations of the original sets Ω_i near $\hat{\omega}_i$ $(i = 1, \ldots, n)$.

(ii) There are certain similarities between Propositions 2.6.6 and 3.2.3 in terms of both assumptions and conclusions. There are also important differences. The assumptions of Propositions 2.6.6 are weaker: the space is not assumed to be complete and the sets are not assumed to be closed. The concluding non-intersection condition in (P8) (Propositions 2.3.4) is formulated for the given sets and in terms of the given collection of points, while Proposition 3.2.3 establishes existence of another collection of points, and the corresponding condition in (3.7) is formulated for localisations of the sets near these points. The translations of the sets are constructed in Propositions 2.6.6 explicitly and are entirely determined by the given collection of points. At the same time, in Proposition 3.2.3 the size of the translations and localisations of the sets as well as the distance of the new points from the given ones are controlled by additional parameters. These parameters, which appear in Proposition 3.2.3, represent the major feature of this statement compared to Propositions 2.6.6. They provide an additional degree (degrees) of freedom for the applications of the result.

Applying Proposition 3.2.3 to the collection of $n+1$ sets $\Omega_1, \ldots, \Omega_n, X$, we arrive at the following statement.

Corollary 3.2.5. *Suppose $\Omega_1, \ldots, \Omega_n$ are closed subsets of a Banach space X, $\omega_i \in \Omega_i$ $(i = 1, \ldots, n)$, $x \in X$ and $\varepsilon > 0$. If $\cap_{i=1}^n \Omega_i = \emptyset$ and condition* (2.41) *is satisfied, then, for any numbers $\lambda, \rho > 0$, there are points $\hat{\omega}_i \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1, \ldots, n)$ such that, for any $\xi > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying*

$$\bigcap_{i=1}^n \Big(((\Omega_i - \hat{\omega}_i) \cap (\xi\lambda)\mathbb{B}) - a_i \Big) \cap (\xi\rho)\mathbb{B} = \emptyset \quad \textit{and} \quad \max_{1 \le i \le n} \|a_i\| < \xi\varepsilon. \tag{3.8}$$

Observe that conditions (3.7) and (3.8) are exactly conditions (P6) and (P2), respectively, applied to the localisations of the sets $\Omega_1 - \hat{\omega}_1, \ldots, \Omega_n - \hat{\omega}_n$ with $\bar{x} := 0$, and $\xi\varepsilon$ and $\xi\rho$ in place of ε and ρ, respectively.

Proposition 3.2.3 assumes that the sets have empty intersection. Next we demonstrate that it can be also applied to collections of sets having a common point. Specifically, we consider the special case (P2) in the Definition 2.3.1(ii) of local extremality, where the last set in the list (of $n+1$ sets) is a ball centered at this point. Since the radius of the ball is allowed to be infinite, the global setting is covered too.

Proposition 3.2.6. *Suppose $\Omega_1, \ldots, \Omega_n$ are closed subsets of a Banach space X, $\bar{x} \in \cap_{i=1}^n \Omega_i$, $\varepsilon > 0$ and $\rho \in]0, \infty]$. If conditions* (P2) *are satisfied for some vectors $a_i \in X$ $(i = 1, \ldots, n)$, then, for any $\lambda > 0$, there are points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ and a number $\delta \in]0, 1[$ such that, for all $\xi \in]0, \delta[$,*

$$\bigcap_{i=1}^n \Big(((\Omega_i - \omega_i) \cap (\xi\lambda)\mathbb{B}) - a_i' \Big) \cap (\xi\rho)\mathbb{B} = \emptyset \quad \textit{and} \quad \max_{1 \le i \le n} \|a_i'\| < \xi\varepsilon \tag{3.9}$$

for some vectors $a_i' \in X$ $(i = 1, \ldots, n)$.

Moreover, if $\lambda \ge \rho + \varepsilon$, then

$$\bigcap_{i=1}^n (\Omega_i - \omega_i - a_i') \cap (\xi\rho)\mathbb{B} = \emptyset; \tag{3.10}$$

if $\lambda+\varepsilon\le\rho$, *then*

$$\bigcap_{i=1}^{n}\Big(((\Omega_i-\omega_i)\cap(\xi\lambda)\mathbb{B})-a_i'\Big)=\emptyset. \tag{3.11}$$

Proof. Let $\rho'\in]0,\rho[$. By (P2), $\cap_{i=1}^{n+1}\Omega_i'=\emptyset$, where $\Omega_i':=\Omega_i-a_i$ $(i=1,\ldots,n)$ and $\Omega_{n+1}':=\overline{B}_{\rho'}(\bar{x})$, and

$$d_1(\bar{x}-a_1,\ldots,\bar{x}-a_n,\bar{x})=\max_{1\le i\le n}\|a_i\|<\varepsilon.$$

Applying Proposition 3.2.3, we can find points $\omega_i\in\Omega_i\cap B_\lambda(\bar{x})$ $(i=1,\ldots,n)$ and $x\in X$ with $\|x\|<\rho'$ such that, for any $\xi>0$, there exist vectors $a_i'\in X$ $(i=1,\ldots,n)$ satisfying

$$\bigcap_{i=1}^{n}\Big(((\Omega_i-\omega_i)\cap(\xi\lambda)\mathbb{B})-a_i'\Big)\cap\overline{B}_{\rho'}(x)\cap(\xi\rho)\mathbb{B}=\emptyset\quad\text{and}\quad\max_{1\le i\le n}\|a_i'\|<\xi\varepsilon.$$

Set $\delta:=(\rho'-\|x\|)/\rho$. Then, $\delta\in]0,1[$ and, for any $\xi\in]0,\delta[$, we have $(\xi\rho)\mathbb{B}\subset\overline{B}_{\rho'}(x)$, and consequently, (3.9) holds true. If $\lambda\ge\rho+\varepsilon$, then $(\xi\rho)\mathbb{B}\subset(\xi\lambda)\,\mathbb{B}-a_i'$ for all $i=1,\ldots,n$, and consequently, (3.9) implies (3.10). Similarly, if $\lambda+\varepsilon\le\rho$, then $(\xi\lambda)\,\mathbb{B}-a_i'\subset(\xi\rho)\mathbb{B}$ for all $i=1,\ldots,n$, and consequently, (3.9) implies (3.11). □

Remark 3.2.7. Conditions (P2) in Proposition 3.2.6 are formulated for a fixed point $\bar{x}\in\cap_{i=1}^n\Omega_i$ and fixed $\varepsilon>0$ and $\rho\in]0,\infty]$. It presumes a certain balance between the values of ε and ρ: the larger the value of ρ is, the larger value of ε is needed to ensure the existence of vectors $a_i'\in X$ $(i=1,\ldots,n)$ satisfying (P2). In contrast, condition (3.9) involves two additional parameters: an arbitrary $\lambda>0$ and a sufficiently small $\xi\in]0,1[$. Fixed ε and ρ are replaced by $\xi\varepsilon$ and $\xi\rho$, respectively, preserving their ratio, while the sets are replaced by their localisations controlled by $\xi\lambda$. This advancement comes at a price: instead of a single common fixed point $\bar{x}$, we now have to deal with a collection of individual points $\omega_i\in\Omega_i$ $(i=1,\ldots,n)$, whose distance from $\bar{x}$ is controlled by λ. Now the balance between this distance and the size of the localisations of the sets becomes important: choosing a smaller λ ensures that the individual points ω_i are closer to $\bar{x}$ while at the same time reducing the size of the localisations of the sets and, thus, weakening condition (3.9).

Given a collection of sets $\Omega_1,\ldots,\Omega_n$ $(n\ge2)$, a point $\bar{x}\in\cap_{i=1}^n\Omega_i$ and a $\rho\in]0,\infty]$, define (cf. [89,90]):

$$\theta_\rho[\Omega_1,\ldots,\Omega_n](\bar{x})=\sup\left\{r>0:\bigcap_{i=1}^{n}(\Omega_i-a_i)\cap B_\rho(\bar{x})\ne\emptyset\quad\text{for all}\quad a_i\in r\mathbb{B}\right\}.$$

This nonnegative quantity tells us how far the sets can be pushed apart until their intersection becomes empty with respect to the fixed ρ-neighbourhood of $\bar{x}$.

Corollary 3.2.8. *Suppose $\Omega_1,\ldots,\Omega_n$ are closed subsets of a Banach space X and $\bar{x}\in\cap_{i=1}^n\Omega_i$. If $\rho>0$ and $\varepsilon>\theta_\rho[\Omega_1,\ldots,\Omega_n](\bar{x})$, then*

$$\inf_{\substack{\omega_i\in\Omega_i\cap B_{\rho+\varepsilon}(\bar{x})\\(i=1,\ldots,n)}}\limsup_{\alpha\downarrow0}\alpha^{-1}\theta_\alpha[\Omega_1-\omega_1,\ldots,\Omega_n-\omega_n](0)\le\rho^{-1}\theta_\rho[\Omega_1,\ldots,\Omega_n](\bar{x}).$$

Moreover, if $\bar{x}\in\operatorname{bd}\cap_{i=1}^n\Omega_i$, then

$$\liminf_{\substack{\omega_i\to\bar{x},\,\omega_i\in\Omega_i\\(i=1,\ldots,n)}}\limsup_{\alpha\downarrow0}\alpha^{-1}\theta_\alpha[\Omega_1-\omega_1,\ldots,\Omega_n-\omega_n](0)\le\liminf_{\rho\downarrow0}\rho^{-1}\theta_\rho[\Omega_1,\ldots,\Omega_n](\bar{x}).$$

Proof. The first assertion is a direct consequence of Proposition 3.2.6. The second assertion is a consequence of the first one since $\bar{x}\in\operatorname{bd}\cap_{i=1}^n\Omega_i$ implies that $\theta_\rho[\Omega_1,\ldots,\Omega_n](\bar{x})\to0$ as $\rho\downarrow0$; cf. [90, Proposition 3]. □

The next proposition presents a metric counterpart of the conditions (P2). It contains the key ingredients of the metric criteria of approximate stationarity and transversality in Theorems 2.4.1 and 2.4.2, respectively.

Proposition 3.2.9. *Suppose $\Omega_1, \ldots, \Omega_n$ are subsets of a normed vector space X, $\bar{x} \in \cap_{i=1}^n \Omega_i$, $a_i \in X$ $(i = 1, \ldots, n)$, $\varepsilon > 0$, $\rho > 0$ and $\alpha := \dfrac{\varepsilon}{\rho}$. Then*

(i) conditions (P2) *imply condition* (M4.1).

(ii) if conditions (M4.1) *is satisfied, then there exist a number $\rho' \in]0, \rho[$, points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ and vectors $a'_i \in X$ $(i = 1, \ldots, n)$ such that conditions* (P4) *are satisfied with ρ' and a'_i in place of ρ and a_i.*

Proof. (i) If conditions (P2) are satisfied, then $\max\limits_{1 \le i \le n} \|a_i\| < \varepsilon$ and

$$\alpha d\left(\bar{x}, \bigcap_{i=1}^n (\Omega_i - a_i)\right) > \alpha\rho = \varepsilon > \max_{1 \le i \le n} \|a_i\| \ge \max_{1 \le i \le n} d\left(\bar{x}, \Omega_i - a_i\right).$$

(ii) Let conditions (M4.1) be satisfied. Then

$$\max_{1 \le i \le n} d\left(\bar{x}, \Omega_i - a_i\right) \le \max_{1 \le i \le n} \|a_i\| < \varepsilon = \alpha\rho,$$

and there exists a number $\rho' \in]0, \rho[$ such that

$$\alpha d\left(\bar{x}, \bigcap_{i=1}^n (\Omega_i - a_i)\right) > \alpha\rho' > \max_{1 \le i \le n} d\left(\bar{x}, \Omega_i - a_i\right).$$

It follows from the first inequality above that $\cap_{i=1}^n (\Omega_i - a_i) \cap B_{\rho'}(\bar{x}) = \emptyset$, while due to the second inequality, there exist $\omega_i \in \Omega_i$ $(i = 1, \ldots, n)$ such that $\max\limits_{1 \le i \le n} \|a'_i\| < \alpha\rho'$, where $a'_i := a_i + \bar{x} - \omega_i$ $(i = 1, \ldots, n)$.

□

Corollary 3.2.10. *Suppose $\Omega_1, \ldots, \Omega_n$ are subsets of a normed vector space X, $\bar{x} \in \cap_{i=1}^n \Omega_i$ and $\alpha > 0$. The following conditions are equivalent:*

(i) the collection $\{\Omega_1, \ldots, \Omega_n\}$ is approximately $\alpha-$stationary at $\bar{x}$;

(ii) for any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n)$ such that conditions (M4.1) *are satisfied;*

(iii) for any $\varepsilon > 0$, there exist vectors $x \in B_\varepsilon(\bar{x})$ and $a_i \in X$ $(i = 1, \ldots, n)$ such that conditions (M4.2) *are satisfied.*

Proof. (i) ⇒ (ii). Let condition (i) be satisfied and a number $\varepsilon > 0$ be given. Set $\varepsilon' := \varepsilon/(\alpha + 1)$. Then conditions (P4) hold with some $\rho \in]0, \varepsilon'[$, $\omega_i \in \Omega_i \cap B_{\varepsilon'}(\bar{x})$, and $a_i \in X$ $(i = 1, \ldots, n)$. By Proposition 3.2.9(i), applied to the sets $\Omega_i - \omega_i$ $(i = 1, \ldots, n)$ having the common point 0, we have

$$\alpha d\left(0, \bigcap_{i=1}^n (\Omega_i - \omega_i - a_i)\right) > \max_{1 \le i \le n} d(0, \Omega_i - \omega_i - a_i).$$

It follows that the first inequality in (M4.1) is satisfied with $a'_i := a_i + \omega_i - \bar{x}$ in place of a_i $(i = 1, \ldots, n)$, and $\max\limits_{1 \le i \le n} \|a'_i\| < \alpha\rho + \varepsilon' < (\alpha + 1)\varepsilon' = \varepsilon$. Hence, condition (ii) is satisfied.

(ii) ⇒ (i). Let condition (ii) be satisfied and a number $\varepsilon > 0$ be given. Set $\varepsilon' := \min\{\alpha, 1\}\varepsilon$, and find vectors $a_i \in X$ $(i = 1, \ldots, n)$ such that conditions (M4.1) are satisfied with ε' in place of ε. By

Proposition 3.2.9(ii), there are $\rho \in]0, \varepsilon'/\alpha[$, $\omega_i \in \Omega_i \cap B_{\varepsilon'}(\bar{x})$, $a_i' \in X$ $(i = 1, \ldots, n)$ such that conditions (P4) hold with a_i' in place of a_i. Since $\varepsilon' \leq \varepsilon$ and $\varepsilon'/\alpha \leq \varepsilon$, condition (i) is satisfied.

(ii) $\Rightarrow$ (iii) is obvious.

(iii) $\Rightarrow$ (ii). Let condition (iii) be satisfied and a number $\varepsilon > 0$ be given. Set $\varepsilon' := \varepsilon/2$. Then conditions (M4.2) hold with some $x \in B_{\varepsilon'}(\bar{x})$, $a_i \in X$ $(i = 1, \ldots, n)$ and ε' in place of ε. Set $a_i' := a_i + x - \bar{x}$ $(i = 1, \ldots, n)$. Then conditions (M4.2) hold true with a_i' in place of a_i. Hence, condition (ii) is satisfied. □

In view of Proposition 2.7.14, Corollary 3.2.10 immediately yields Proposition 2.4.2.

3.3 Dual Characterisations

This section presents a series of 'generalised separation' statements, providing dual characterisations of certain typical 'extremal' arrangements of collections of sets, discussed in the preceding sections, and traces the relationships between them. These statements contain core arguments, which can be found in various existing versions of the (extended) extremal principle, as well as some new extensions.

The definition of the approximate stationarity (Definition 2.3.1(iv)) and its conventional dual characterisations in the extended extremal principle (Theorem 2.7.3) are all formulated "for any $\varepsilon > 0$ there exist . . . ". At the same time, the proofs of the extremal principle and its extensions establish connections between the values of ε and other parameters involved in the assumptions and the conclusions. These connections, usually hidden in the proofs, are of importance for more subtle extremality statements. We expose them in the statements below.

All the separation statements in this section are consequences of the next general theorem, providing dual characterisations of the slightly weakened version of the asymmetric extremality property contained in Proposition 2.3.4(i). It involves the d_1 distance between n sets defined in (2.36) and actually combines two statements: for closed sets in general Banach spaces and specifically in Asplund spaces. So far it has been common to formulate (and prove!) such statements separately; cf. [120, 172].

Theorem 3.3.1. *Suppose $\Omega_1, \ldots, \Omega_n$ are closed subsets of a Banach space X, $\bar{x} \in \cap_{i=1}^n \Omega_i$, $a_i \in X$ $(i = 1, \ldots, n-1)$, and $\varepsilon > 0$. Suppose also that*

$$\bigcap_{i=1}^{n-1} (\Omega_i - a_i) \cap \Omega_n = \emptyset, \tag{3.12}$$

$$\max_{1 \leq i \leq n-1} \|a_i\| < d_1(\Omega_1 - a_1, \ldots, \Omega_{n-1} - a_{n-1}, \Omega_n) + \varepsilon \tag{3.13}$$

(or simply $\max_{1 \leq i \leq n-1} \|a_i\| < \varepsilon$). Then,

(i) *for any $\lambda > 0$ and $\rho > 0$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n-1)$, $\omega_n \in \Omega_n \cap B_\rho(\bar{x})$, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ such that*

$$\sum_{i=1}^{n} x_i^* = 0, \quad \sum_{i=1}^{n-1} \|x_i^*\| = 1, \tag{3.14}$$

$$\lambda \sum_{i=1}^{n-1} d\left(x_i^*, N_{\Omega_i}(\omega_i)\right) + \rho d\left(x_n^*, N_{\Omega_n}(\omega_n)\right) < \varepsilon, \tag{3.15}$$

$$\sum_{i=1}^{n-1} \langle x_i^*, \omega_n + a_i - \omega_i \rangle = \max_{1 \leq i \leq n-1} \|\omega_n + a_i - \omega_i\|, \tag{3.16}$$

where N in (3.15) stands for the Clarke normal cone ($N := N^C$);

(ii) *if X is Asplund, then, for any $\lambda > 0$, $\rho > 0$ and $\tau \in]0,1[$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1,\ldots,n-1)$, $\omega_n \in \Omega_n \cap B_\rho(\bar{x})$, and vectors $x_i^* \in X^*$ $(i = 1,\ldots,n)$ satisfying conditions (3.14), (3.15), where N in (3.15) stands for the Fréchet normal cone ($N := N^F$), and*

$$\sum_{i=1}^{n-1} \langle x_i^*, \omega_n + a_i - \omega_i \rangle > \tau \max_{1\le i\le n-1} \|\omega_n + a_i - \omega_i\|. \tag{3.17}$$

Proof. Let $\lambda > 0$ and $\rho > 0$. In view of (3.13), we can choose positive numbers ε_1 and ε_2 satisfying

$$\max_{1\le i\le n-1} \|a_i\| - d_1(\Omega_1 - a_1, \ldots, \Omega_{n-1} - a_{n-1}, \Omega_n) < \varepsilon_1 < \varepsilon_2 < \varepsilon \tag{3.18}$$

Note that $\max_{1\le i\le n-1} \|a_i\| = d_1(\bar{x} - a_1, \ldots, \bar{x} - a_{n-1}, \bar{x})$. We can apply Lemma 3.2.1 to find points $\hat{\omega}_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1,\ldots,n-1)$ and $\hat{\omega}_n \in \Omega_n \cap B_\rho(\bar{x})$ such that

$$\max_{1\le i\le n-1} \|\hat{\omega}_i - a_i - \hat{\omega}_n\| < \\ d_1(\Omega_1 \cap B_{\alpha\lambda}(\hat{\omega}_1) - a_1, \ldots, \Omega_{n-1} \cap B_{\alpha\lambda}(\hat{\omega}_{n-1}) - a_{n-1}, \Omega_n \cap B_{\alpha\rho}(\hat{\omega}_n)) + \alpha\varepsilon_1 \tag{3.19}$$

for all $\alpha > 0$. Consider the three functions $f_1, f_2, f_3 : X^n \to \mathbb{R}_+ \cup \{+\infty\}$:

$$f_1(u_1,\ldots,u_n) := \max_{1\le i\le n-1} \|u_i - a_i - u_n\|, \tag{3.20}$$

$$f_2(u_1,\ldots,u_n) := \varepsilon_2 \max\left\{\lambda^{-1} \max_{1\le i\le n-1} \|u_i - \hat{\omega}_i\|, \rho^{-1}\|u_n - \hat{\omega}_n\|\right\}, \tag{3.21}$$

$$f_3(u_1,\ldots,u_n) := \begin{cases} 0 & \text{if } u_i \in \Omega_i \ (i = 1,\ldots,n), \\ \infty & \text{otherwise.} \end{cases} \tag{3.22}$$

Observe that, in view of (3.12), $f_1(\hat{\omega}_1,\ldots,\hat{\omega}_n) = \max_{1\le i\le n-1} \|\hat{\omega}_i - a_i - \hat{\omega}_n\| > 0$. Moreover,

$$f_1(u_1,\ldots,u_n) - f_1(\hat{\omega}_1,\ldots,\hat{\omega}_n) + f_2(u_1,\ldots,u_n) \ge 0 \quad \text{for all} \quad u_i \in \Omega_i \ (i = 1,\ldots,n).$$

Indeed, assume that there are $u_i \in \Omega_i$ $(i = 1,\ldots,n)$ such that the inequality does not hold. Then $(u_1,\ldots,u_n) \ne (\hat{\omega}_i,\ldots,\hat{\omega}_n)$, and consequently, $f_2(u_1,\ldots,u_n) > 0$. Set $\alpha := f_2(u_1,\ldots,u_n)/\varepsilon_1$. We have

$$\|u_i - \hat{\omega}_i\| \le \frac{\varepsilon_1}{\varepsilon_2}\alpha\lambda < \alpha\lambda (i = 1,\ldots,n-1), \quad \|u_n - \hat{\omega}_n\| \le \frac{\varepsilon_1}{\varepsilon_2}\alpha\rho < \alpha\rho$$
$$\max_{1\le i\le n-1} \|u_i - a_i - u_n\| - \max_{1\le i\le n-1} \|\hat{\omega}_i - a_i - \hat{\omega}_n\| + \alpha\varepsilon_1 < 0,$$

which contradicts (3.19). Thus, $(\hat{\omega}_1,\ldots,\hat{\omega}_n)$ is a point of minimum of the sum $f_1 + f_2 + f_3$, and consequently (Lemma 2.2.2)

$$0 \in \partial(f_1 + f_2 + f_3)(\hat{\omega}_1,\ldots,\hat{\omega}_n).$$

Functions f_1 and f_2 are convex and Lipschitz continuous. It is easy to check that the subdifferentials of f_1, f_2 and f_3 possess the following properties:

1) The function f_1 (3.20) is a composition function: $f_1(u_1,\ldots,u_n) = g\big(A(u_1,\ldots,u_n) - (a_1,\ldots,a_{n-1})\big)$, where A is the linear operator from X^n to X^{n-1}: $A(u_1,\ldots,u_n) := (u_1 - u_n, \ldots, u_{n-1} - u_n)$, and g is the maximum norm on X^{n-1}: $g(u_1,\ldots,u_{n-1}) := \max_{1\le i\le n-1} \|u_i\|$. The corresponding dual norm has the form $(v_1^*,\ldots,v_{n-1}^*) \mapsto \sum_{i=1}^{n-1} \|v_i^*\|$. It is easy to check that the adjoint operator $A^* : (X^*)^{n-1} \to (X^*)^n$ is of the form $A^*(v_1^*,\ldots,v_{n-1}^*) = \left(v_1^*,\ldots,v_{n-1}^*, -\sum_{j=1}^{n-1} v_j^*\right)$.

When $f_1(u_1,\ldots,u_n) > 0$, the subdifferential $\partial g\,(u_1-a_1-u_n,\ldots,u_{n-1}-a_{n-1}-u_n)$ is the set of $(v^*_{11},\ldots,v^*_{1,n-1}) \in (X^*)^{n-1}$ satisfying (see e.g. [168, Corollary 2.4.16])

$$\sum_{i=1}^{n-1} \|v^*_{1i}\| = 1 \quad \text{and} \quad \sum_{i=1}^{n-1} \langle v^*_{1i}, u_i - a_i - u_n\rangle = \max_{1\le i\le n-1} \|u_i - a_i - u_n\|\,. \tag{3.23}$$

Thus, in view of the convex chain rule (see e.g. [168, Theorem 2.8.3]), if $f_1(u_1,\ldots,u_n) > 0$, then the subdifferential $\partial f_1(u_1,\ldots,u_n)$ is the set of all vectors $\left(v^*_{11},\ldots,v^*_{1,n-1},-\sum_{j=1}^{n-1} v^*_{1j}\right) \in (X^*)^n$, where vectors $v^*_{1i} \in X^*$ $(i=1,\ldots,n-1)$ satisfy (3.23).

2) The function f_2 (3.21) is a positive multiple of the norm on X^n (translated by $(\hat\omega_1,\ldots,\hat\omega_n)$). Its subgradients $(v^*_{21},\ldots,v^*_{2n})$ at any point satisfy

$$\lambda \sum_{i=1}^{n-1} \|v^*_{2i}\| + \rho\,\|v^*_{2n}\| \le \varepsilon_2. \tag{3.24}$$

3) The function f_3 (3.22) is the indicator function of the set $\Omega_1 \times \ldots \times \Omega_n$. Its subdifferential has a simple representation: $\partial f_3(u_1,\ldots,u_n) = \prod_{i=1}^{n} N_{\Omega_i}(u_i)$ for all $u_i \in \Omega_i$ $(i=1,\ldots,n)$ (Lemma 2.2.3).

From this point the proof splits into two cases.

(i) We apply the Clarke–Rockafellar subdifferential sum rule (Lemma 2.2.1(iv)) to find elements of the three subdifferentials: $(v^*_{11},\ldots,v^*_{1,n-1}) \in \partial g(\hat\omega_1 - a_1 - \hat\omega_n,\ldots,\hat\omega_{n-1} - a_n - \hat\omega_n)$, $(v^*_{21},\ldots,v^*_{2n}) \in \partial f_2(\hat\omega_1,\ldots,\hat\omega_n)$ and $(v^*_{31},\ldots,v^*_{3n}) \in \partial f_3(\hat\omega_1,\ldots,\hat\omega_n)$ such that

$$v^*_{1i} + v^*_{2i} + v^*_{3i} = 0 \;(i=1,\ldots,n-1) \quad \text{and} \quad -\left(\sum_{j=1}^{n-1} v^*_{1j}\right) + v^*_{2n} + v^*_{3n} = 0.$$

Then $v^*_{3i} \in N_{\Omega_i}(\hat\omega_i)$ $(i=1,\ldots,n)$ and conditions (3.23) and (3.24) are satisfied. The conclusion of part (i) of the theorem holds true with $\hat\omega_i$ in place of ω_i $(i=1,\ldots,n)$, $x^*_i := -v^*_{1i}$ $(i=1,\ldots,n-1)$ and $x^*_n := \sum_{j=1}^{n-1} v^*_{1j}$.

(ii) Let X be Asplund and $\tau \in]0,1[$. We can apply the fuzzy sum rule (Lemma 2.2.1(iii)) to the sum $(f_1+f_2)+f_3$ followed by the conventional convex sum rule (Lemma 2.2.1(i)) applied to f_1+f_2. Choose a $\xi > 0$ satisfying the following conditions:

$$\xi < \varepsilon - \varepsilon_2, \quad \xi < \lambda - \max_{1\le i\le n-1} \|\hat\omega_i - \bar x\|, \quad \xi < \rho - \|\hat\omega_n - \bar x\|, \tag{3.25}$$

$$(10-2\tau)\xi < (1-\tau) \max_{1\le i\le n-1} \|\hat\omega_i - a_i - \hat\omega_n\|. \tag{3.26}$$

Since $\tau \in]0,1[$, the last inequality implies in particular that

$$2\xi < \max_{1\le i\le n-1} \|\hat\omega_i - a_i - \hat\omega_n\|. \tag{3.27}$$

Applying the fuzzy sum rule, we find two points $(x_1,\ldots,x_n)$, $(\omega_1,\ldots,\omega_n) \in X^n$ such that $\omega_i \in \Omega_i$ $(i=1,\ldots,n)$ and

$$\max_{1\le i\le n} \|x_i - \hat\omega_i\| < \xi, \quad \max_{1\le i\le n} \|\omega_i - \hat\omega_i\| < \xi, \tag{3.28}$$

and elements of the three subdifferentials: $(v^*_{11},\ldots,v^*_{1,n-1}) \in \partial g(x_1-a_1-x_n,\ldots,x_{n-1}-a_n-x_n)$, $(v^*_{21},\ldots,v^*_{2n}) \in \partial f_2(x_1,\ldots,x_n)$ and $(v^*_{31},\ldots,v^*_{3n}) \in \partial f_3(\omega_1,\ldots,\omega_n)$ such that

$$\sum_{i=1}^{n-1} \|v^*_{1i}+v^*_{2i}+v^*_{3i}\| + \left\| -\left(\sum_{j=1}^{n-1} v^*_{1j}\right) + v^*_{2n}+v^*_{3n}\right\| < \frac{\xi}{\max\{\lambda,\rho\}}.$$

The last condition implies

$$\lambda\sum_{i=1}^{n-1} \|v^*_{1i}+v^*_{2i}+v^*_{3i}\| + \rho\left\| -\left(\sum_{j=1}^{n-1} v^*_{1j}\right) + v^*_{2n}+v^*_{3n}\right\| < \xi. \tag{3.29}$$

In view of (3.25), (3.26), (3.27) and (3.28), we have the following estimates:

$$\|\omega_i-\bar x\| \le \|\hat\omega_i-\bar x\|+\|\omega_i-\hat\omega_i\| < \lambda \quad (i=1,\ldots,n-1), \quad \|\omega_n-\bar x\| < \|\hat\omega_n-\bar x\|+\|\omega_n-\hat\omega_n\| < \rho,$$

(hence, $\omega_i \in \Omega_i \cap B_\lambda(\bar x)$ $(i=1,\ldots,n-1)$ and $\omega_n \in \Omega_n \cap B_\rho(\bar x)$)

$$\|\omega_i-x_i\| \le \|\omega_i-\hat\omega_i\|+\|x_i-\hat\omega_i\| < 2\xi \quad (i=1,\ldots,n), \tag{3.30}$$

$$\begin{aligned}\max_{1\le i\le n-1}\|x_i-a_i-x_n\| &\ge \max_{1\le i\le n-1}\left(\|\hat\omega_i-a_i-\hat\omega_n\|-\|x_i-\hat\omega_i\|-\|x_n-\hat\omega_n\|\right)\\ &> \max_{1\le i\le n-1}\|\hat\omega_i-a_i-\hat\omega_n\|-2\xi>0,\\ \max_{1\le i\le n-1}\|\omega_i-a_i-\omega_n\| &\ge \max_{1\le i\le n-1}\left(\|\hat\omega_i-a_i-\hat\omega_n\|-\|\omega_i-\hat\omega_i\|-\|\omega_n-\hat\omega_n\|\right)\\ &> \max_{1\le i\le n-1}\|\hat\omega_i-a_i-\hat\omega_n\|-2\xi>0,\end{aligned}$$

$$\begin{aligned}(1-\tau)\max_{1\le i\le n-1}\|\omega_i-a_i-\omega_n\| &> (1-\tau)\left(\max_{1\le i\le n-1}\|\hat\omega_i-a_i-\hat\omega_n\|-2\xi\right)\\ &> (10-2\tau)\xi-2(1-\tau)\xi=8\xi,\end{aligned} \tag{3.31}$$

conditions (3.23) and (3.24) are satisfied and $v^*_{3i} \in N_{\Omega_i}(\omega_i)$ $(i=1,\ldots,n)$. Denote $x^*_i := -v^*_{1i}$ $(i=1,\ldots,n-1)$ and $x^*_n := \sum_{j=1}^{n-1} v^*_{1j}$. Then $\sum_{i=1}^{n} x^*_i = 0$ and $\sum_{i=1}^{n}\|x^*_i\| = 1$. Using (3.23) and (3.30), we obtain the following inequalities:

$$\begin{aligned}\sum_{i=1}^{n-1}\langle x^*_i,\omega_n+a_i-\omega_i\rangle &\ge \sum_{i=1}^{n-1}\langle v^*_{1i},x_i-a_i-x_n\rangle-\sum_{i=1}^{n-1}\|v^*_{1i}\|\left(\|\omega_i-x_i\|+\|\omega_n-x_n\|\right)\\ &> \max_{1\le i\le n-1}\|x_i-a_i-x_n\|-4\xi\\ &\ge \max_{1\le i\le n-1}\left(\|\omega_i-a_i-\omega_n\|-\|\omega_i-x_i\|-\|\omega_n-x_n\|\right)-4\xi\\ &> \max_{1\le i\le n-1}\|\omega_i-a_i-\omega_n\|-8\xi.\end{aligned} \tag{3.32}$$

Adding (3.31) and (3.32), we arrive at (3.17). Making use of (3.24), (3.29) and (3.25), we obtain the following estimates:

$$\lambda\sum_{i=1}^{n-1}\|x^*_i-v^*_{3i}\|+\rho\,\|x^*_n-v^*_{3n}\| < \lambda\sum_{i=1}^{n-1}\|v^*_{2i}\|+\rho\,\|v^*_{2n}\|+\xi < \varepsilon_2+\xi<\varepsilon.$$

The last inequality yields (3.15). □

Remark 3.3.2. (i) Conditions (3.12) and (3.13) are implied by conditions (P5). Hence, in view of Proposition 2.3.4(i), Theorem 3.3.1 provides dual necessary characterisations of extremality.

(ii) Conditions (3.16) and (3.17) relate the dual vectors x^*_i and the primal space vectors $\omega_n+a_i-\omega_i$ $(i=1,\ldots,n-1)$. Such conditions, though not common in the conventional formulations of the extremal/generalised separation statements, seem to provide important additional characterisations of the properties. Conditions of this kind first appeared explicitly in the generalised separation theorems in [172], where the authors also provided motivations for employing such conditions. As one can see from the proof above, conditions (3.17) and (3.16) originate in computing the convex subdifferential of the norm in X^{n-1} at a nonzero point; see (3.23). Subdifferentiating a norm (in either X^{n-1} or X^n) at a nonzero point is a necessary step in the proofs of all existing versions of the extremal principle and its extensions, starting with the very first one in [103, Theorem 6.1], with conditions like (3.23) hidden in the proofs. In several statements in the rest of this section, following [172], we make such conditions exposed.

(iii) Lemma 3.2.1 substitutes in the proof of Theorem 3.3.1 the conventional Ekeland variational principle.

(iv) If condition (3.13) in Theorem 3.3.1 is replaced by a stronger one:

$$\max_{1\le i\le n-1}\|a_i\| = d_1(\Omega_1 - a_1,\ldots,\Omega_{n-1} - a_{n-1},\Omega_n),$$

which means that the infimum of $\max_{1\le i\le n-1}\|\omega_i - a_i - \omega_n\|$ over $\omega_i\in\Omega_i$ $(i=1,\ldots,n)$ is attained at $\omega_1=\ldots=\omega_n=\bar x$, then the application of Lemma 3.2.1 in the proof can be dropped, leading to an improvement in part (i): conditions (3.15) and (3.16) can be replaced by $x_i^*\in N^C_{\Omega_i}(\bar x)$ $(i=1,\ldots,n)$ and $\sum_{i=1}^{n-1}\langle x_i^*,a_i\rangle = \max_{1\le i\le n-1}\|a_i\|$, respectively. A similar fact was observed in [172, Theorem 3.1′].

The assumption $\cap_{i=1}^n\Omega_i\ne\emptyset$ in Theorem 3.3.1 is not restrictive. The common point $\bar x\in\cap_{i=1}^n\Omega_i$ of the collection of sets can be replaced by a collection of individual points $\omega_i\in\Omega_i$ $(i=1,\ldots,n)$. The next statement provides dual characterisations of the slightly weakened version of the asymmetric extremality property contained in Proposition 2.3.4.

Corollary 3.3.3. *Suppose $\Omega_1,\ldots,\Omega_n$ are closed subsets of a Banach space X, $\omega_i\in\Omega_i$ $(i=1,\ldots,n)$, $a_i\in X$ $(i=1,\ldots,n-1)$, and $\varepsilon>0$. Suppose also that*

$$\bigcap_{i=1}^{n-1}(\Omega_i-\omega_i-a_i)\cap(\Omega_n-\omega_n)=\emptyset, \tag{3.33}$$

$$\max_{1\le i\le n-1}\|a_i\| < d_1(\Omega_1-\omega_1-a_1,\ldots,\Omega_{n-1}-\omega_{n-1}-a_{n-1},\Omega_n-\omega_n)+\varepsilon \tag{3.34}$$

(or simply $\max_{1\le i\le n-1}\|a_i\|<\varepsilon$). Then,

(i) *for any $\lambda>0$ and $\rho>0$, there exist points $\omega_i'\in\Omega_i\cap B_\lambda(\omega_i)$ $(i=1,\ldots,n-1)$, $\omega_n'\in\Omega_n\cap B_\rho(\omega_n)$, and vectors $x_i^*\in X^*$ $(i=1,\ldots,n)$ satisfying conditions* (3.14), (3.15) *and* (3.16) *with $N:=N^C$ and ω_i' in place of ω_i $(i=1,\ldots,n)$;*

(ii) *if X is Asplund, then, for any $\lambda>0$, $\rho>0$ and $\tau\in]0,1[$, there exist points $\omega_i'\in\Omega_i\cap B_\lambda(\omega_i)$ $(i=1,\ldots,n-1)$, $\omega_n'\in\Omega_n\cap B_\rho(\omega_n)$, and vectors $x_i^*\in X^*$ $(i=1,\ldots,n)$ satisfying conditions* (3.14), (3.15) *and* (3.17) *with $N:=N^F$ and ω_i' in place of ω_i $(i=1,\ldots,n)$.*

Proof. The sets $\Omega_i':=\Omega_i-\omega_i$ $(i=1,\ldots,n)$ satisfy $0\in\cap_{i=1}^n\Omega_i'$ and conditions (3.12) and (3.13). The conclusion follows from Theorem 3.3.1 after noticing that $N_{\Omega_i'}(\omega_i'-\omega_i)=N_{\Omega_i}(\omega_i')$ $(i=1,\ldots,n)$. □

Theorem 3.3.1 is a particular case of Corollary 3.3.3 with $\omega_i=\bar x$ $(i=1,\ldots,n)$.

The following theorem is an immediate consequence of Theorem 3.3.1. It combines two *unified separation theorems* due to Zheng and Ng [172, Theorems 3.1 and 3.4].

Theorem 3.3.4. *Suppose $\Omega_1, \ldots, \Omega_n$ are closed subsets of a Banach space X, $\cap_{i=1}^n \Omega_i = \emptyset$, $\omega_i \in \Omega_i$ $(i = 1, \ldots, n)$, $\varepsilon > 0$ and condition* (2.40) *is satisfied. Then,*

(i) *for any $\lambda > 0$, there exist points $\omega_i' \in \Omega_i \cap B_\lambda(\omega_i)$ and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions*

$$\sum_{i=1}^n d(x_i^*, N_{\Omega_i}(\omega_i')) < \frac{\varepsilon}{\lambda}, \tag{ZN1}$$

with N standing for the Clarke normal cone ($N := N^C$), and

$$\sum_{i=1}^{n-1} \langle x_i^*, \omega_n' - \omega_i' \rangle = \max_{1 \le i \le n-1} \|\omega_i' - \omega_n'\|; \tag{ZN2}$$

(ii) *if X is Asplund, then, for any numbers $\lambda > 0$ and $\tau \in]0, 1[$, there exist points $\omega_i' \in \Omega_i \cap B_\lambda(\omega_i)$ and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (ZN1) *with N standing for the Fréchet normal cone ($N := N^F$), and*

$$\sum_{i=1}^{n-1} \langle x_i^*, \omega_n' - \omega_i' \rangle > \tau \max_{1 \le i \le n-1} \|\omega_i' - \omega_n'\|. \tag{NZ3}$$

Proof. Observe that the sets $\Omega_i' := \Omega_i - \omega_i$ $(i = 1, \ldots, n)$ and vectors $a_i := \omega_n - \omega_i$ $(i = 1, \ldots, n-1)$ satisfy $0 \in \bigcap_{i=1}^n \Omega_i'$ and

$$\bigcap_{i=1}^{n-1} (\Omega_i' - a_i) \cap \Omega_n' = \bigcap_{i=1}^n (\Omega_i - \omega_n) = \bigcap_{i=1}^n \Omega_i - \omega_n = \emptyset,$$

$$\max_{1 \le i \le n-1} \|a_i\| = d_1(\omega_1, \ldots, \omega_n) < d_1(\Omega_1, \ldots, \Omega_n) + \varepsilon = d_1(\Omega_1' - a_1, \ldots, \Omega_{n-1}' - a_{n-1}, \Omega_n') + \varepsilon.$$

Applying Theorem 3.3.1 with $\rho = \lambda$, we arrive at the conclusions. □

Remark 3.3.5. (i) In [172], instead of the d_1 distance in condition (2.40), a slightly more general p-weighted nonintersect index was used with the corresponding q-weighted sums replacing the usual ones in (ZN1), (ZN2) and (NZ3). This corresponds to considering ℓ_p norms on product spaces and the corresponding ℓ_q dual norms. In the thesis, for simplicity only the maximum norm on product spaces is considered together with the corresponding sum norm in the dual space; (cf. Remark 2.6.5(ii).)

(ii) Theorem 4.2.7 is a consequence of Theorem 3.3.1, which in turn is a consequence of the Ekeland variational principle. Thanks to [113, Theorem 3.1], part (i) of Theorem 4.2.7 is equivalent to the Ekeland variational principle. Hence, the conclusion of Theorem 3.3.1 is also equivalent to the Ekeland variational principle (and to completeness of the space X).

The next theorem is a 'symmetric' version of Theorem 3.3.1.

Theorem 3.3.6. *Suppose $\Omega_1, \ldots, \Omega_n$ are closed subsets of a Banach space X, $\bar{x} \in \cap_{i=1}^n \Omega_i$, $a_i \in X$ $(i = 1, \ldots, n)$, $\rho > 0$ and $\varepsilon > 0$. Suppose also that*

$$\bigcap_{i=1}^n (\Omega_i - a_i) \cap B_\rho(\bar{x}) = \emptyset, \tag{3.35}$$

$$\max_{1 \le i \le n} \|a_i\| < d_{B_\rho(\bar{x})}(\Omega_1 - a_1, \ldots, \Omega_n - a_n) + \varepsilon \tag{3.36}$$

(or simply $\max_{1 \le i \le n} \|a_i\| < \varepsilon$). Then,

(i) *for any* $\lambda > 0$, *there exist points* $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ *and* $x \in B_\rho(\bar{x})$, *and vectors* $x_i^* \in X^*$ $(i = 1, \ldots, n)$ *such that*

$$\lambda \sum_{i=1}^{n} d(x_i^*, N_{\Omega_i}(\omega_i)) + \rho \left\| \sum_{i=1}^{n} x_i^* \right\| < \varepsilon, \quad \sum_{i=1}^{n} \|x_i^*\| = 1, \tag{3.37}$$

$$\sum_{i=1}^{n} \langle x_i^*, x + a_i - \omega_i \rangle = \max_{1 \le i \le n} \|x + a_i - \omega_i\|, \tag{3.38}$$

where N *in* (3.37) *stands for the Clarke normal cone* ($N := N^C$);

(ii) *if* X *is Asplund, then, for any* $\lambda > 0$ *and* $\tau \in]0, 1[$, *there exist points* $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ *and* $x \in B_\rho(\bar{x})$, *and vectors* $x_i^* \in X^*$ $(i = 1, \ldots, n)$ *satisfying conditions* (3.37), *where* N *stands for the Fréchet normal cone* ($N := N^F$), *and*

$$\sum_{i=1}^{n} \langle x_i^*, x + a_i - \omega_i \rangle > \tau \max_{1 \le i \le n} \|x + a_i - \omega_i\|. \tag{3.39}$$

Proof. Choose an $\varepsilon' \in]0, \varepsilon[$ such that condition (3.36) holds true with ε' in place of ε, and then choose a $\rho' \in]0, \rho[$ such that $\rho - \rho' < \varepsilon - \varepsilon'$. It is sufficient to apply Theorem 3.3.1 to the collection of $n + 1$ closed sets $\Omega_1, \ldots, \Omega_n$ and $\Omega_{n+1} := \overline{B}_{\rho'}(\bar{x})$ with ε' and ρ' in place of ε and ρ, respectively. Notice that $\Omega_{n+1} \cap B_{\rho'}(\bar{x}) = B_{\rho'}(\bar{x})$ and $N_{\Omega_{n+1}(\bar{x})}(x) = 0$ for any $x \in B_{\rho'}(x)$. One only needs to check the inequality in (3.37), which is straightforward:

$$\lambda \sum_{i=1}^{n} d(x_i^*, N_{\Omega_i}(\omega_i)) + \rho \left\| \sum_{i=1}^{n} x_i^* \right\| < \varepsilon' + (\rho - \rho') \left\| \sum_{i=1}^{n} x_i^* \right\| \le \varepsilon' + (\rho - \rho') < \varepsilon.$$

The proof is complete. □

Remark 3.3.7. (i) Conditions (3.35) and (3.36) are implied by conditions (P2). Hence, Theorem 3.3.6 provides dual necessary characterisations of the local extremality.

(ii) The inequality in (3.37) combines two constraints on the vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$: they must be close to the respective normal cones and their sum must be small. This inequality obviously implies two separate inequalities:

$$\sum_{i=1}^{n} d(x_i^*, N_{\Omega_i}(\omega_i)) < \frac{\varepsilon}{\lambda} \quad \text{and} \quad \left\| \sum_{i=1}^{n} x_i^* \right\| < \frac{\varepsilon}{\rho},$$

while the converse implication is not true in general. Both these constraints are involved in each of the two generalised separation conditions in the (extended) extremal principle discussed in Section 2.7.2 (conditions (D1) and (D2) of Theorem 2.7.3). However, each of the generalised separation conditions in Theorem 2.7.3 requires actually a stronger version of one of the constraints: either the vectors must belong to the respective normal cones (condition (D1) of Theorem 2.7.3) or their sum must be exactly zero (condition (D2) of Theorem 2.7.3). Fortunately, as the next two corollaries show, the required stronger versions of (one of) the constraints are consequences of the combined condition (3.37) and Lemma 2.7.5.

Corollary 3.3.8. *Suppose* $\Omega_1, \ldots, \Omega_n$ *are closed subsets of a Banach space* X, $\bar{x} \in \cap_{i=1}^n \Omega_i$, $a_i \in X$ $(i = 1, \ldots, n)$, $0 < \varepsilon < \rho$, $0 < \lambda \le \rho - \varepsilon$, *and conditions* (3.35) *and* (3.36) *are satisfied (the latter condition can be replaced by the simpler and stronger one:* $\max\limits_{1 \le i \le n} \|a_i\| < \varepsilon$*). Then,*

(i) *there exist points* $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ *and* $x \in B_\rho(\bar{x})$, *and vectors* $x_i^* \in X^*$ $(i = 1, \ldots, n)$, *satisfying conditions* (D4) *with* N *standing for the Clarke normal cone* ($N := N^C$) *and* $\alpha := \dfrac{\varepsilon}{\lambda}$, *and condition* (3.39) *with* $\tau := \dfrac{\rho - \varepsilon}{\rho + \varepsilon}$;

(ii) *if X is Asplund, then, for any $\tau \in]0, \frac{\rho-\varepsilon}{\rho+\varepsilon}[$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ and $x \in B_\rho(\bar{x})$, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying condition* (D4) *with N standing for the Fréchet normal cone ($N := N^F$) and $\alpha := \frac{\varepsilon}{\lambda}$, and condition* (3.39).

Proof. Set $\tau' := \tau(1 + \frac{\varepsilon}{\rho}) + \frac{\varepsilon}{\rho}$, and observe that $\tau' = 1$ in part (i), and $\tau' \in]0, 1[$ in part (ii). Applying Theorem 3.3.1, we find points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ and $x \in B_\rho(\bar{x})$, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ such that condition (3.37) with the respective normal cone is satisfied and

$$\sum_{i=1}^{n} \langle x_i^*, x + a_i - \omega_i \rangle \geq \tau' \max_{1 \leq i \leq n} \|x + a_i - \omega_i\|. \tag{3.40}$$

The assertion follows from Lemma 2.7.5(i) and (iii) after noticing that $\frac{\tau'\rho - \varepsilon}{\rho + \varepsilon} = \tau$. □

Corollary 3.3.9. *Suppose $\Omega_1, \ldots, \Omega_n$ are closed subsets of a Banach space X, $\bar{x} \in \cap_{i=1}^n \Omega_i$, $a_i \in X$ $(i = 1, \ldots, n)$, $0 < \rho < \lambda$, $0 < \varepsilon \leq \lambda - \rho$, and conditions* (3.35) *and* (3.36) *are satisfied (the latter condition can be replaced by the simpler and stronger one: $\max_{1 \leq i \leq n} \|a_i\| < \varepsilon$). Then,*

(i) *there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ and $x \in B_\rho(\bar{x})$, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$, satisfying conditions* (D3) *with N standing for the Clarke normal cone ($N := N^C$) and $\alpha := \frac{\varepsilon}{\rho}$, and condition* (3.39) *with $\tau := \frac{\lambda - \varepsilon}{\lambda + \varepsilon}$;*

(ii) *if X is Asplund, then, for any $\tau \in]0, \frac{\lambda-\varepsilon}{\lambda+\varepsilon}[$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ and $x \in B_\rho(\bar{x})$, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (D3) *with N standing for the Fréchet normal cone ($N := N^F$) and $\alpha := \frac{\varepsilon}{\rho}$, and condition* (3.39).

Proof. Set $\tau' := \tau(1 + \frac{\varepsilon}{\lambda}) + \frac{\varepsilon}{\lambda}$, and observe that $\tau' = 1$ in part (i), and $\tau' \in]0, 1[$ in part (ii). Applying Theorem 3.3.1, we find points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ and $x \in B_\rho(\bar{x})$, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ such that conditions (3.37) (with the respective normal cone) and (3.40) are satisfied. The assertion follows from Lemma 2.7.5(ii) and (iii) after noticing that $\frac{\tau'\lambda - \varepsilon}{\lambda + \varepsilon} = \tau$. □

The versions of Theorem 3.3.6 and Corollaries 3.3.8 and 3.3.9 with a common point $\bar{x} \in \cap_{i=1}^n \Omega_i$ of a collection of sets replaced by a collection of individual points $\omega_i \in \Omega_i$ $(i = 1, \ldots, n)$, presented in the next three corollaries, follow immediately.

Corollary 3.3.10. *Suppose $\Omega_1, \ldots, \Omega_n$ are closed subsets of a Banach space X, $\omega_i \in \Omega_i$, $a_i \in X$ $(i = 1, \ldots, n)$, $\rho > 0$ and $\varepsilon > 0$. Suppose also that*

$$\bigcap_{i=1}^{n} (\Omega_i - \omega_i - a_i) \cap (\rho\mathbb{B}) = \emptyset, \tag{3.41}$$

$$\max_{1 \leq i \leq n} \|a_i\| < d_{\rho\mathbb{B}}(\Omega_1 - \omega_1 - a_1, \ldots, \Omega_n - \omega_n - a_n) + \varepsilon \tag{3.42}$$

(or simply $\max_{1 \leq i \leq n} \|a_i\| < \varepsilon$). Then,

(i) *for any $\lambda > 0$, there exist points $\omega_i' \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1, \ldots, n)$ and $x \in \rho\mathbb{B}$, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (3.37) *and* (3.38) *with N standing for the Clarke normal cone ($N := N^C$), and ω_i' in place of ω_i $(i = 1, \ldots, n)$;*

(ii) *if X is Asplund, then, for any $\lambda > 0$ and $\tau \in]0,1[$, there exist points $\omega_i' \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1,\ldots,n)$ and $x \in \rho\mathbb{B}$, and vectors $x_i^* \in X^*$ $(i = 1,\ldots,n)$ satisfying conditions* (3.37) *and* (3.39) *with N standing for the Fréchet normal cone ($N := N^F$), and ω_i' in place of ω_i $(i = 1,\ldots,n)$.*

Corollary 3.3.11. *Suppose $\Omega_1,\ldots,\Omega_n$ are closed subsets of a Banach space X, $\omega_i \in \Omega_i$, $a_i \in X$ $(i = 1,\ldots,n)$, $0 < \varepsilon < \rho$, $0 < \lambda \leq \rho - \varepsilon$, and conditions* (3.41) *and* (3.42) *are satisfied (the latter condition can be replaced by the simpler and stronger one: $\max\limits_{1\leq i\leq n} \|a_i\| < \varepsilon$). Then,*

(i) *there exist points $\omega_i' \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1,\ldots,n)$ and $x \in \rho\mathbb{B}$, and vectors $x_i^* \in X^*$ $(i = 1,\ldots,n)$ such that conditions* (D4) *and* (3.39) *are satisfied with N standing for the Clarke normal cone ($N := N^C$), $\alpha := \dfrac{\varepsilon}{\lambda}$, $\tau := \dfrac{\rho-\varepsilon}{\rho+\varepsilon}$, and ω_i' in place of ω_i $(i = 1,\ldots,n)$;*

(ii) *if X is Asplund, then, for any $\tau \in]0, \dfrac{\rho-\varepsilon}{\rho+\varepsilon}[$, there exist points $\omega_i' \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1,\ldots,n)$ and $x \in \rho\mathbb{B}$, and vectors $x_i^* \in X^*$ $(i = 1,\ldots,n)$ such that conditions* (D4) *and* (3.39) *are satisfied with N standing for the Fréchet normal cone ($N := N^F$), $\alpha := \dfrac{\varepsilon}{\lambda}$, and ω_i' in place of ω_i $(i = 1,\ldots,n)$.*

Corollary 3.3.12. *Suppose $\Omega_1,\ldots,\Omega_n$ are closed subsets of a Banach space X, $\omega_i \in \Omega_i$, $a_i \in X$ $(i = 1,\ldots,n)$, $0 < \rho < \lambda$, $0 < \varepsilon \leq \lambda - \rho$, and conditions* (3.41) *and* (3.42) *are satisfied (the latter condition can be replaced by the simpler and stronger one: $\max\limits_{1\leq i\leq n} \|a_i\| < \varepsilon$). Then,*

(i) *there exist points $\omega_i' \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1,\ldots,n)$ and $x \in \rho\mathbb{B}$, and vectors $x_i^* \in X^*$ $(i = 1,\ldots,n)$ such that conditions* (D3) *and* (3.39) *are satisfied with N standing for the Clarke normal cone ($N := N^C$), $\alpha := \dfrac{\varepsilon}{\rho}$, $\tau := \dfrac{\lambda-\varepsilon}{\lambda+\varepsilon}$, and ω_i' in place of ω_i $(i = 1,\ldots,n)$;*

(ii) *if X is Asplund, then, for any $\tau \in]0, \dfrac{\lambda-\varepsilon}{\lambda+\varepsilon}[$, there exist points $\omega_i' \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1,\ldots,n)$ and $x \in \rho\mathbb{B}$, and vectors $x_i^* \in X^*$ $(i = 1,\ldots,n)$ such that conditions* (D3) *and* (3.39) *are satisfied with N standing for the Fréchet normal cone ($N := N^F$), $\alpha := \dfrac{\varepsilon}{\rho}$, and ω_i' in place of ω_i $(i = 1,\ldots,n)$.*

Remark 3.3.13. In the above three corollaries, the assumption of the existence of a common point $\bar{x} \in \cap_{i=1}^n \Omega_i$ of a collection of sets, used in Theorem 3.3.6 and Corollaries 3.3.8 and 3.3.9, is relaxed to that of a collection of individual points $\omega_i \in \Omega_i$ $(i = 1,\ldots,n)$ (which always exist as long as all the sets are nonempty). On the other hand, if such a point $\bar{x} \in \cap_{i=1}^n \Omega_i$ exists, it can be used along with the collection $\omega_i \in \Omega_i$ $(i = 1,\ldots,n)$ to provide additional useful estimates. Indeed, if $\|\omega_i - \bar{x}\| \leq \xi$ $(i = 1,\ldots,n)$ for some $\xi \geq 0$ (one can take, e.g., $\xi := \max\limits_{1\leq i\leq n} \|\omega_i - \bar{x}\|$), then each of the above three corollaries immediately gives $\|\omega_i' - \bar{x}\| < \xi + \lambda$ $(i = 1,\ldots,n)$. This simple observation plays an important role in the proof of the extended extremal principle. It is used also in the next statement, which is a consequence of Corollaries 3.3.11 and 3.3.12.

Proposition 3.3.14. *Suppose $\Omega_1,\ldots,\Omega_n$ are closed subsets of a Banach space X, $\bar{x} \in \cap_{i=1}^n \Omega_i$, $\xi \geq 0$, $\omega_i \in \Omega_i \cap B_\xi(\bar{x})$, $a_i \in X$ $(i = 1,\ldots,n)$, $\alpha \in]0,1[$, $\rho > 0$, $\xi + \rho(1-\alpha) < \delta$, and conditions* (P8) *are satisfied. Then,*

(i) *there exist points $\omega_i' \in B_\delta(\bar{x})$ and vectors $x_i^* \in X^*$ $(i = 1,\ldots,n)$ such that conditions* (D3) *and* (3.39) *are satisfied with N standing for the Clarke normal cone ($N := N^C$), $\tau := \dfrac{1}{1+2\alpha}$, and ω_i' in place of ω_i;*

(ii) *there exist points $\omega_i' \in B_\delta(\bar{x})$ and vectors $x_i^* \in X^*$ $(i = 1,\ldots,n)$ such that conditions* (D4) *and* (3.39) *are satisfied with N standing for the Clarke normal cone ($N := N^C$), $\tau := \dfrac{1-\alpha}{1+\alpha}$, and $\alpha' := \dfrac{\alpha}{1-\alpha}$ and ω_i' in place of α and ω_i, respectively.*

Suppose X is Asplund and N stands for the Fréchet normal cone ($N := N^F$). Then,

(iii) *for any $\tau \in]0, \dfrac{1}{1+2\alpha}[$, there exist points $\omega'_i \in B_\delta(\bar{x})$ and vectors $x_i^* \in X^*$ ($i = 1, \ldots, n$) such that conditions* (D3) *and* (3.39) *are satisfied with ω'_i in place of ω_i;*

(iv) *for any $\tau \in]0, \dfrac{1-\alpha}{1+\alpha}[$, there exist points $\omega'_i \in B_\delta(\bar{x})$ and vectors $x_i^* \in X^*$ ($i = 1, \ldots, n$) such that conditions* (D4) *and* (3.39) *are satisfied with $\alpha' := \dfrac{\alpha}{1-\alpha}$ and ω'_i in place of α and ω_i, respectively.*

Proof. In view of Remark 3.3.13, assertions (i) and (iii) are consequences of Corollary 3.3.11 with $\varepsilon := \alpha\rho$ and $\lambda := (1-\alpha)\rho$, while assertions (ii) and (iv) are consequences of Corollary 3.3.12 with $\varepsilon := \alpha\rho$ and $\lambda := (1+\alpha)\rho$. □

Remark 3.3.15. The assertion in Proposition 3.3.14(iii) improves [97, Theorem 3.1], which was used in [97] as the main tool when extending the extremal principle to infinite collections of sets.

The above proposition yields dual characterisations of approximate $\alpha-$stationarity defined in Definition 2.7.13.

Corollary 3.3.16. *Suppose $\Omega_1, \ldots, \Omega_n$ are closed subsets of a Banach space X, $\bar{x} \in \cap_{i=1}^n \Omega_i$ and $\alpha > 0$. Suppose also that the collection $\{\Omega_1, \ldots, \Omega_n\}$ is approximately α-stationary at $\bar{x}$. Then,*

(i) *for any $\varepsilon > 0$, there exist points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ and vectors $x_i^* \in X^*$ ($i = 1, \ldots, n$) such that conditions* (D3) *and* (3.39) *are satisfied with N standing for the Clarke normal cone ($N := N^C$) and $\tau := \dfrac{1}{1+2\alpha}$;*

(ii) *if additionally $\alpha \in]0,1[$, then, for any $\varepsilon > 0$, there exist points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ and vectors $x_i^* \in X^*$ ($i = 1, \ldots, n$) such that conditions* (D4) *and* (3.39) *are satisfied with N standing for the Clarke normal cone ($N := N^C$), $\tau := \dfrac{1-\alpha}{1+\alpha}$, and $\alpha' := \dfrac{\alpha}{1-\alpha}$ in place of α.*

Suppose X is Asplund and N stands for the Fréchet normal cone ($N := N^F$). Then,

(iii) *for any $\varepsilon > 0$ and $\tau \in]0, \dfrac{1}{1+2\alpha}[$, there exist points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ and vectors $x_i^* \in X^*$ ($i = 1, \ldots, n$) such that conditions* (D3) *and* (3.39) *are satisfied;*

(iv) *if additionally $\alpha \in]0,1[$, then, for any $\varepsilon > 0$ and $\tau \in]0, \dfrac{1-\alpha}{1+\alpha}[$, there exist points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ and vectors $x_i^* \in X^*$ ($i = 1, \ldots, n$) such that conditions* (D4) *and* (3.39) *are satisfied with $\alpha' := \dfrac{\alpha}{1-\alpha}$ in place of α.*

Proof. In each of the assertions, let numbers $\varepsilon > 0$ and τ, satisfying the respective conditions, be given. For parts (i) and (iii), choose a number $\xi \in]0, \dfrac{\varepsilon}{2+\alpha}[$, while for parts (ii) and (iv), choose a number $\xi \in]0, \dfrac{\varepsilon}{2-\alpha}[$. By Definition 2.7.13, there exist a number $\rho \in]0, \xi[$, points $\omega_i \in \Omega_i \cap B_\xi(\bar{x})$ and vectors $a_i \in X$ ($i = 1, \ldots, n$) such that conditions (P4) are satisfied. In parts (i) and (iii), we have $\xi + \rho(1+\alpha) < \xi(2+\alpha) < \varepsilon$, and in parts (ii) and (iv), we have $\xi + \rho(1-\alpha) < \xi(2-\alpha) < \varepsilon$. The conclusions follow from Proposition 3.3.14. □

Remark 3.3.17. The infinitesimal statements in Corollary 3.3.16 are crucial for the extended extremal principle and its extensions to infinite collections of sets. For instance, in view of Proposition 2.7.14, Corollary 3.3.16(iii) and (iv) immediately yield the implications, respectively, (i) ⇒ (ii) and (i) ⇒ (iii) in Theorem 2.7.3.

The dual characterisations of the 'extremal' arrangements of collections of sets given in the statements in the first part of this section are themselves in a sense extremal properties of collections of sets. They can be partially reversed in the setting of a general normed vector space. We start with an 'asymmetric' statement where the last set on the list plays a special role.

Proposition 3.3.18. *Suppose $\Omega_1, \ldots, \Omega_n$ are subsets of a normed vector space X, $\omega_i \in \Omega_i$ $(i = 1, \ldots, n)$ and $\varepsilon > 0$. If vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfy*

$$x_i^* \in N^F_{\Omega_i}(\omega_i) \ \ (i = 1, \ldots, n-1), \quad d(x_n^*, N^F_{\Omega_n}(\omega_n)) < \varepsilon \tag{3.43}$$

and conditions (3.14)*, then there exists a $\delta > 0$ such that, for any $\rho \in]0, \delta[$ and $\tau \in]0, 1[$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n-1)$ satisfying conditions* (P8) *and*

$$\sum_{i=1}^{n-1} \langle x_i^*, a_i \rangle > \tau\varepsilon\rho. \tag{3.44}$$

Proof. By (3.43), we can choose a vector $x^* \in N^F_{\Omega_n}(\omega_n)$ and a positive number $\varepsilon' < \varepsilon$ such that

$$\|x_n^* - x^*\| < \varepsilon'. \tag{3.45}$$

Choose also numbers $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$ such that

$$n\varepsilon_1 + (n-1)\varepsilon_2 < \varepsilon - \varepsilon' \quad \text{and} \quad (n-1)\varepsilon_2 < \varepsilon(1-\tau). \tag{3.46}$$

By the definition of the Fréchet normal cone, there exists a number $\delta > 0$ such that

$$\langle x_i^*, \omega - \omega_i \rangle \le \frac{\varepsilon_1}{\varepsilon + 1} \|\omega - \omega_i\| \quad \text{for all} \quad \omega \in \Omega_i \cap B_{(\varepsilon+1)\delta}(\omega_i) \ (i = 1, \ldots, n-1), \tag{3.47}$$

$$\langle x^*, \omega - \omega_n \rangle \le \varepsilon_1 \|\omega - \omega_n\| \quad \text{for all} \quad \omega \in \Omega_n \cap B_\delta(\omega_n). \tag{3.48}$$

Let $\rho \in]0, \delta[$. Choose vectors $a_i \in X$ $(i = 1, \ldots, n-1)$ such that

$$\|a_i\| < \varepsilon\rho \quad \text{and} \quad \langle x_i^*, a_i \rangle > \varepsilon\rho \|x_i^*\| - \varepsilon_2\rho \quad (i = 1, \ldots, n-1). \tag{3.49}$$

Then, by (3.14) and (3.46),

$$\sum_{i=1}^{n-1} \langle x_i^*, a_i \rangle > \varepsilon\rho - (n-1)\varepsilon_2\rho > \tau\varepsilon\rho.$$

Hence, the inequality in (P8) and condition (3.44) are satisfied. Suppose that the equality in (P8) does not hold. Then there exist $\omega_i' \in \Omega_i$ $(i = 1, \ldots, n)$ and an $x \in \rho\mathbb{B}$ such that

$$\omega_1' - \omega_1 - a_1 = \ldots = \omega_{n-1}' - \omega_{n-1} - a_{n-1} = \omega_n' - \omega_n = x.$$

By (3.49), $\|\omega_i' - \omega_i\| = \|x + a_i\| \le \|x\| + \|a_i\| < (\varepsilon + 1)\rho < (\varepsilon + 1)\delta$ $(i = 1, \ldots, n-1)$ and $\|\omega_n' - \omega_n\| = \|x\| \le \rho < \delta$. Hence, by (3.47), (3.48) and (3.49),

$$\langle x_i^*, x \rangle = \langle x_i^*, \omega_i' - \omega_i \rangle - \langle x_i^*, a_i \rangle < -\varepsilon\rho \|x_i^*\| + (\varepsilon_1 + \varepsilon_2)\rho,$$
$$\langle x^*, x \rangle = \langle x^*, \omega_n' - \omega_n \rangle < \varepsilon_1\rho,$$

and consequently, using (3.14) and (3.46),

$$\langle x^* - x_n^*, x \rangle = \sum_{i=1}^{n-1} \langle x_i^*, x \rangle + \langle x^*, x \rangle < -\varepsilon\rho + n\varepsilon_1\rho + (n-1)\varepsilon_2\rho < -\varepsilon'\rho.$$

On the other hand, by (3.45), $\langle x^* - x_n^*, x \rangle > -\varepsilon'\rho$. A contradiction. ☐

The corresponding 'symmetric' statement follows immediately after applying Proposition 3.3.18 to the collection of $n+1$ sets $\Omega_1, \ldots, \Omega_n$ and X.

Corollary 3.3.19. *Suppose $\Omega_1,\ldots,\Omega_n$ are subsets of a normed vector space X, $\omega_i\in\Omega_i$ $(i=1,\ldots,n)$ and $\varepsilon>0$. If vectors $x_i^*\in X^*$ $(i=1,\ldots,n)$ satisfy conditions* (D1)*, then there is a $\delta>0$ such that, for any $\rho\in]0,\delta[$ and $\tau\in]0,1[$, there exist vectors $a_i\in X$ $(i=1,\ldots,n)$ satisfying condition* (P4) *and*

$$\sum_{i=1}^{n}\langle x_i^*,a_i\rangle>\tau\varepsilon\rho. \tag{3.50}$$

In view of Proposition 2.7.11(iv), the above two statements produce dual sufficient characterisations for approximate stationarity.

Corollary 3.3.20. *Suppose $\Omega_1,\ldots,\Omega_n$ are subsets of a normed vector space X and $\bar x\in\cap_{i=1}^n\Omega_i$. If for any $\varepsilon>0$ there exist points $\omega_i\in\Omega_i\cap B_\varepsilon(\bar x)$ $(i=1,\ldots,n)$ and vectors $x_i^*\in X^*$ $(i=1,\ldots,n)$ satisfying either conditions* (3.43) *or conditions* (D1)*, then the collection $\{\Omega_1,\ldots,\Omega_n\}$ is approximately stationary at $\bar x$.*

Remark 3.3.21. Similar to the dual necessary characterisations of extremality/stationarity properties discussed in the first part of this section, the sufficient conditions in Proposition 3.3.18 and Corollary 3.3.19 contain conditions (3.44) and (3.50), respectively, relating the given dual vectors x_i^* $(i=1,\ldots,n)$ and the primal space translation vectors a_i $(i=1,\ldots,n)$ guaranteed by the statements. In view of Remark 3.3.2(ii), such conditions seem to be an intrinsic feature of the extremality/stationarity properties, independently on whether one goes from primal space conditions to dual space ones or the other way round.

Combining Corollaries 3.3.16 and 3.3.19, we can formulate a full dual characterisation of approximate α-stationarity when either the space is Asplund or the sets are convex.

Corollary 3.3.22. *Suppose $\Omega_1,\ldots,\Omega_n$ are closed subsets of a Banach space X, $\bar x\in\cap_{i=1}^n\Omega_i$ and $\alpha>0$. Suppose also that either X is Asplund or $\Omega_1,\ldots,\Omega_n$ are convex. The collection $\{\Omega_1,\ldots,\Omega_n\}$ is approximately α-stationary at $\bar x$ if and only if, for any $\varepsilon>0$, there exist points $\omega_i\in B_\varepsilon(\bar x)$ and vectors $x_i^*\in X^*$ $(i=1,\ldots,n)$ such that conditions* (D3) *are satisfied with N standing for the Fréchet normal cone ($N:=N^F$).*

Moreover, under the above conditions, if X is Asplund and $\tau\in]0,\dfrac{1}{1+2\alpha}[$, or $\Omega_1,\ldots,\Omega_n$ are convex and $\tau:=\dfrac{1}{1+2\alpha}$, then, for any $\varepsilon>0$, points $\omega_i\in B_\varepsilon(\bar x)$ and vectors $x_i^*\in X^*$ $(i=1,\ldots,n)$ can be chosen to satisfy also condition (3.39)*, while, for any $\hat\tau\in]0,1[$, vectors $a_i\in X$ $(i=1,\ldots,n)$ in Definition 2.7.13 of the approximate α-stationarity can be chosen to satisfy additionally $\sum_{i=1}^{n}\langle x_i^*,a_i\rangle>\hat\tau\alpha\rho$.*

Proof. The 'only if' part together with the condition (3.39) in the 'moreover' part follow from parts (i) and (iii) of Corollary 3.3.16, taking into account that for convex sets the Clarke and Fréchet normal cones coincide. Conversely, given any $\varepsilon>0$, $\hat\tau\in]0,1[$, points $\omega_i\in B_\varepsilon(\bar x)$ and vectors $x_i^*\in X^*$ $(i=1,\ldots,n)$ satisfying conditions (D3), Corollary 3.3.19 with α and $\hat\tau$ in place of ε and τ, respectively, yields the approximate α-stationarity and condition $\sum_{i=1}^{n}\langle x_i^*,a_i\rangle>\hat\tau\alpha\rho$. □

Remark 3.3.23. In view of Corollary 2.7.8 and Remark 2.7.16, the first part of Corollary 3.3.22 yields the statements of Theorems 2.7.3 and 2.7.15.

The assumption $x_i^*\in N_{\Omega_i}^F(\omega_i)$ $(i=1,\ldots,n-1)$ in Proposition 3.3.18 can be relaxed (at the expense of weakening the estimates in (P8) and (3.44)). The next statement is a consequence of Propositions 3.3.18 and 2.7.10.

Corollary 3.3.24. *Suppose $\Omega_1,\ldots,\Omega_n$ are subsets of a normed vector space X, $\omega_i\in\Omega_i$ $(i=1,\ldots,n)$ and $\varepsilon\in]0,1[$. If vectors $x_i^*\in X^*$ $(i=1,\ldots,n)$ satisfy conditions* (D6) *with N standing for the Fréchet* *normal cone ($N := N^F$), then there is a $\delta > 0$ such that, for any $\rho \in]0, \delta[$ and $\tau \in]0, 1[$, there exist vectors $a_i \in X$ $(i = 1, \dots, n-1)$ satisfying conditions* (P8) *and* (3.44) *with $\varepsilon' := \varepsilon/(1-\varepsilon)$ in place of ε.*

Since $\varepsilon/(1-\varepsilon) < \varepsilon$, the conclusions of Corollary 3.3.24 are weaker than those of Proposition 3.3.18. Observe that conditions (P5) involve a localisation of the n-th set (near $\omega_n \in \Omega_n$). The estimates can be improved by considering localisations of all the sets.

Proposition 3.3.25. *Suppose $\Omega_1, \dots, \Omega_n$ are subsets of a normed vector space X, $\omega_i \in \Omega_i$ $(i = 1, \dots, n)$ and $\varepsilon > 0$. If vectors $x_i^* \in X^*$ $(i = 1, \dots, n)$ satisfy conditions* (D6) *with N standing for the Fréchet normal cone ($N := N^F$), then there is a $\delta > 0$ such that, for any $\rho \in]0, \delta[$ and $\tau \in]0, 1[$, there are vectors $a_i \in X$ $(i = 1, \dots, n-1)$ satisfying*

$$\bigcap_{i=1}^{n-1} \Big((\Omega_i - \omega_i) \cap (\rho\mathbb{B}) - a_i \Big) \cap (\Omega_n - \omega_n) \cap (\rho\mathbb{B}) = \emptyset \quad \textit{and} \quad \max_{1 \le i \le n-1} \|a_i\| < \varepsilon\rho, \tag{3.51}$$

and condition (3.44).

The proof below is a modification of that of Proposition 3.3.18.

Proof. Choose vectors $y_i^* \in N^F_{\Omega_i}(\omega_i)$ $(i = 1, \dots, n)$ and a positive number $\varepsilon' < \varepsilon$ such that

$$\sum_{i=1}^{n} \|x_i^* - y_i^*\| < \varepsilon'. \tag{3.52}$$

Then choose numbers $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$ such that $\varepsilon_1 + \varepsilon_2 < \varepsilon - \varepsilon'$ and $\varepsilon_2 < (1-\tau)\varepsilon$. By the definition of the Fréchet normal cone, there is a $\delta > 0$ such that

$$\langle y_i^*, \omega - \omega_i \rangle \le \frac{\varepsilon_1}{n} \|\omega - \omega_i\| \quad \text{for all} \quad \omega \in \Omega_i \cap B_\delta(\omega_i) \ \ (i = 1, \dots, n). \tag{3.53}$$

Let $\rho \in]0, \delta[$. Choose vectors $a_i \in X$ $(i = 1, \dots, n)$ satisfying

$$\|a_i\| < \varepsilon\rho \quad \text{and} \quad \langle x_i^*, a_i \rangle > \varepsilon\rho \|x_i^*\| - \frac{\varepsilon_2 \rho}{n-1} \quad (i = 1, \dots, n-1).$$

By (D6), we have

$$\sum_{i=1}^{n-1} \langle x_i^*, a_i \rangle > \varepsilon\rho - \varepsilon_2\rho. \tag{3.54}$$

The inequality in (3.51) and condition (3.44) follow. Suppose that the equality in (3.51) is not satisfied. Then there exist points $\omega_i' \in \Omega_i \cap B_\rho(\omega_i)$ $(i = 1, \dots, n)$ and an $x \in \rho\mathbb{B}$ such that

$$\omega_1' - \omega_1 - a_1 = \ldots = \omega_{n-1}' - \omega_{n-1} - a_{n-1} = \omega_n' - \omega_n = x. \tag{3.55}$$

Hence, making use of (3.55), (D6), (3.54), (3.53) and (3.52), we have

$$\begin{aligned}
0 &= \sum_{i=1}^{n-1} \langle x_i^*, (\omega_n' - \omega_n) - (\omega_i' - \omega_i - a_i) \rangle = -\sum_{i=1}^{n-1} \langle x_i^*, \omega_i' - \omega_i - a_i \rangle - \langle x_n^*, \omega_n' - \omega_n \rangle \\
&= \sum_{i=1}^{n-1} \langle x_i^*, a_i \rangle - \sum_{i=1}^{n} \langle x_i^*, \omega_i' - \omega_i \rangle = \sum_{i=1}^{n-1} \langle x_i^*, a_i \rangle - \sum_{i=1}^{n} \langle y_i^*, \omega_i' - \omega_i \rangle + \sum_{i=1}^{n} \langle y_i^* - x_i^*, \omega_i' - \omega_i \rangle \\
&> \varepsilon\rho - \varepsilon_2\rho - \frac{\varepsilon_1}{n} \sum_{i=1}^{n} \|\omega_i' - \omega_i\| - \varepsilon' \max_{1 \le i \le n} \|\omega_i' - \omega_i\| \ge (\varepsilon - \varepsilon' - \varepsilon_1 - \varepsilon_2)\rho > 0.
\end{aligned}$$

This contradiction proves the proposition. □

Chapter 4

Nonlinear Characterisations

The conventional extremal principle and all its subsequent extensions, including the recent dual generalised separation characterisations of the elementary non-intersection properties in [29, 30, 97, 170, 172, 174], use only 'linear' estimates. Motivated partially by the very recent developments in [40,42,43], where nonlinear transversality properties are studied, in Chapter 4, we target linear and nonlinear primal and dual characterisations of the elementary non-intersection properties involved in Definition 2.3.1. Note that, unlike the corresponding transversality properties, the properties in Definition 2.3.1 do not contain explicitly any nonlinearity. Nevertheless, nonlinear estimates can be added naturally to their characterisations. As discussed above, when proving dual generalised separation statements, non-intersection properties of sets are first reformulated in terms of functions, which admit application of the Ekeland variational principle. Starting from [101–103], distance-type functions are normally used for that purpose. It is elementary to observe that more general nonlinear functions can be used as well. The functions do not even have to be continuous (the Ekeland variational principle only requires the function to be lower semicontinuous), but we do not go that far in this context.

In this chapter, we establish nonlinear primal (slope) and dual (normal cone) characterisations of the key non-intersection properties (3.35) and (3.12), with fixed vectors a_i's. These ubiquitous properties are present in one form or another in all four parts of Definition 2.3.1 and Proposition 2.3.4, as well as in all known extensions of the extremality/stationarity properties. Thus, any necessary characterisations of these properties translate into the corresponding characterisations of the properties in Definition 2.3.1. In particular, their dual characterisations lie at the heart of the conventional extremal principle. Besides, examining the elementary non-intersection properties independently of the containing them four conventional properties in Definition 2.3.1 opens a way for studying other related properties, e.g., in the important for applications setting when all the sets lie in a subspace of X.

(Linear) metric and dual characterisations of (3.35) and (3.12) have been studied in [30]. Here we aim at establishing more general nonlinear characterisations (also in the slope form). The nonlinearity in our model is determined by a continuous strictly increasing function $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying $\varphi(0) = 0$. The family of all such functions is denoted by $\mathcal{C}$. We denote by $\mathcal{C}^1$ the subfamily of functions $\varphi \in \mathcal{C}$ which are continuously differentiable on $]0, \infty[$ with $\varphi'(t) > 0$ for all $t > 0$. Obviously, if $\varphi \in \mathcal{C}$ ($\varphi \in \mathcal{C}^1$), then $\varphi^{-1} \in \mathcal{C}$ ($\varphi^{-1} \in \mathcal{C}^1$).

Along with the conventional maximum norm, we are going to consider on the product X^n of $n > 1$ copies of the space X the following parametric norm depending on a number $\gamma > 0$:

$$\|(u_1, \ldots, u_n)\|_\gamma := \max\{\|u_1\|, \ldots, \|u_{n-1}\|, \gamma\|u_n\|\}, \quad u_1, \ldots, u_n \in X, \tag{4.1}$$

where the parameter is always associated with the last component.

4.1 Slope Characterisations

4.1.1 Slopes

Let $\psi : X \to \mathbb{R}\cup\{+\infty\}$ be an extended-real-valued function on a metric space. The *slope* [45] (cf. [6,75]) of ψ at $x \in \operatorname{dom}\psi$ is defined by

$$|\nabla\psi|(x) := \limsup_{u\to x,\, u\neq x} \frac{[\psi(x)-\psi(u)]_+}{d(x,u)},$$

where $\alpha_+ := \max\{0,\alpha\}$ for any $\alpha \in \mathbb{R}$. If $\psi(x) \geq 0$, one can define the *nonlocal slope* [94] (cf. [126]) of ψ at x:

$$|\nabla\psi|^{\diamond}(x) := \sup_{u\neq x} \frac{[\psi(x)-\psi(u)_+]_+}{d(x,u)}.$$

When $x \notin \operatorname{dom}\psi$, we set $|\nabla\psi|(x) := |\nabla\psi|^{\diamond}(x) := +\infty$. Obviously, $0 \leq |\nabla\psi|(x) \leq |\nabla\psi|^{\diamond}(x)$ for all $x \in X$ (with $\psi(x) \geq 0$), and both quantities can be infinite.

The next lemma from [42] provides a chain rule for slopes. It slightly improves [7, Lemma 4.1], where ψ and φ were assumed lower semicontinuous and continuously differentiable, respectively. The composition $\varphi\circ\psi$ of a function $\psi : X \to \mathbb{R}\cup\{+\infty\}$ on a metric space and a function $\varphi : \mathbb{R} \to \mathbb{R}\cup\{+\infty\}$ is understood in the usual sense with the natural convention that $(\varphi\circ\psi)(x) = +\infty$ if $\psi(x) = +\infty$.

Lemma 4.1.1 (Slope chain rule)**.** *Let X be a metric space, $\psi : X \to \mathbb{R}\cup\{+\infty\}$, $\varphi : \mathbb{R} \to \mathbb{R}\cup\{+\infty\}$, $x \in \operatorname{dom}\psi$ and $\psi(x) \in \operatorname{dom}\varphi$. Suppose φ is nondecreasing on $\mathbb{R}$ and differentiable at $\psi(x)$, and either $\varphi'(\psi(x)) > 0$ or $|\nabla\psi|(x) < +\infty$. Then*

$$|\nabla(\varphi\circ\psi)|(x) = \varphi'(\psi(x))|\nabla\psi|(x).$$

Remark 4.1.2. The chain rule in Lemma 4.1.1 is a local result. Instead of assuming that φ is defined on the whole real line, one can assume that φ is defined and finite on a closed interval $[\alpha,\beta]$ around the point $\psi(x)$: $\alpha < \psi(x) < \beta$. It is sufficient to redefine the composition $\varphi\circ\psi$ for x with $\psi(x) \notin [\alpha,\beta]$ as follows: $(\varphi\circ\psi)(x) := \varphi(\alpha)$ if $\psi(x) < \alpha$, and $(\varphi\circ\psi)(x) := \varphi(\beta)$ if $\psi(x) > \beta$. This does not affect the conclusion of the lemma.

4.1.2 Slope Characterisations

The next statement gives slope characterisations for the non-intersection property (3.12), where the vectors $a_1,\ldots,a_{n-1} \in X$ are fixed. If $\cap_{i=1}^n \Omega_i \neq \emptyset$, then condition (3.12) implies $\max\limits_{1\leq i\leq n-1}\|a_i\| > 0$. Note that condition (3.12) is not symmetric: the role of the set Ω_n differs from that of the other sets $\Omega_1,\ldots,\Omega_{n-1}$. This difference is exploited in the subsequent statements.

To quantify non-intersection properties, we are going to use the the asymmetric distance-like quantity (2.32) (*nonintersect index* [172]):

$$d_1(\Omega_1,\ldots,\Omega_n) := \inf_{u_i\in\Omega_i\ (i=1,\ldots,n)} \max_{1\leq i\leq n-1} \|u_n - u_i\|.$$

Theorem 4.1.3. *Let $\Omega_1,\ldots,\Omega_n$ be closed subsets of a Banach space X, $\bar{x} \in \cap_{i=1}^n\Omega_i$, $a_i \in X$ ($i = 1,\ldots,n-1$), $\varepsilon > 0$ and $\varphi \in \mathcal{C}$. Suppose that condition* (3.12) *is satisfied, and*

$$\varphi\left(\max_{1\leq i\leq n-1}\|a_i\|\right) < \varphi(d_1(\Omega_1 - a_1,\ldots,\Omega_{n-1}-a_{n-1},\Omega_n) + \varepsilon \tag{4.2}$$

(or simply $\varphi\left(\max_{1\le i\le n-1}\|a_i\|\right)<\varepsilon$*). Then, for any* $\lambda>0$ *and* $\eta>0$*, there exist points* $\omega_i\in\Omega_i\cap B_\lambda(\bar{x})$ $(i=1,\dots,n-1)$ *and* $\omega_n\in\Omega_n\cap B_\eta(\bar{x})$ *such that*

$$\sup_{\substack{u_i\in\Omega_i\ (i=1,\dots,n)\\(u_1,\dots,u_n)\ne(\omega_1,\dots,\omega_n)}}\frac{\varphi\Big(\max_{1\le i\le n-1}\|\omega_n+a_i-\omega_i\|\Big)-\varphi\Big(\max_{1\le i\le n-1}\|u_n+a_i-u_i\|\Big)}{\|(u_1-\omega_1,\dots,u_n-\omega_n)\|_\gamma}<\frac{\varepsilon}{\lambda},\tag{4.3}$$

$$0<\max_{1\le i\le n-1}\|\omega_n+a_i-\omega_i\|\le\max_{1\le i\le n-1}\|a_i\|,\tag{4.4}$$

where $\gamma:=\dfrac{\lambda}{\eta}$*. As a consequence,*

$$\limsup_{\substack{\Omega_i\ni u_i\to\omega_i\ (i=1,\dots,n)\\(u_1,\dots,u_n)\ne(\omega_1,\dots,\omega_n)}}\frac{\varphi\Big(\max_{1\le i\le n-1}\|\omega_n+a_i-\omega_i\|\Big)-\varphi\Big(\max_{1\le i\le n-1}\|u_n+a_i-u_i\|\Big)}{\|(u_1-\omega_1,\dots,u_n-\omega_n)\|_\gamma}<\frac{\varepsilon}{\lambda}.\tag{4.5}$$

Moreover, if φ *is differentiable at* $\max_{1\le i\le n-1}\|\omega_n+a_i-\omega_i\|$*, then*

$$\begin{aligned}&\varphi'\Big(\max_{1\le i\le n-1}\|\omega_n+a_i-\omega_i\|\Big)\\&\qquad\times\limsup_{\substack{\Omega_i\ni u_i\to\omega_i\ (i=1,\dots,n)\\(u_1,\dots,u_n)\ne(\omega_1,\dots,\omega_n)}}\frac{\max_{1\le i\le n-1}\|\omega_n+a_i-\omega_i\|-\max_{1\le i\le n-1}\|u_n+a_i-u_i\|}{\|(u_1-\omega_1,\dots,u_n-\omega_n)\|_\gamma}<\frac{\varepsilon}{\lambda},\end{aligned}\tag{4.6}$$

with the convention $0\cdot(+\infty)=0$*.*

Proof. Let $\lambda>0$, $\eta>0$, $\gamma:=\dfrac{\lambda}{\eta}$, and a number ε' satisfy

$$\varphi\left(\max_{1\le i\le n-1}\|a_i\|\right)-\varphi(d_1(\Omega_1-a_1,\dots,\Omega_{n-1}-a_{n-1},\Omega_n)<\varepsilon'<\varepsilon.\tag{4.7}$$

Consider a continuous function $f:X^n\to\mathbb{R}_+$:

$$f(u_1,\dots,u_n):=\varphi\Big(\max_{1\le i\le n-1}\|u_n+a_i-u_i\|\Big),\quad u_1,\dots,u_n\in X.\tag{4.8}$$

It follows from (3.12) and (4.7) that

$$f(u_1,\dots,u_n)>0\quad\text{for all}\quad u_i\in\Omega_i\ (i=1,\dots,n),$$
$$f(\bar{x},\dots,\bar{x})=\varphi\left(\max_{1\le i\le n-1}\|a_i\|\right)<\varphi(d_1(\Omega_1-a_1,\dots,\Omega_{n-1}-a_{n-1},\Omega_n)+\varepsilon'.$$

Applying the Ekeland Variational Principle (Lemma 3.1.1) to the restriction of the function f to the complete metric space $\Omega_1\times\ldots\times\Omega_n$ with the metric induced by the norm (4.1), we find points $\omega_i\in\Omega_i\cap B_\lambda(\bar{x})$ $(i=1,\dots,n-1)$ and $\omega_n\in\Omega_n\cap B_\eta(\bar{x})$ such that

$$0<f(\omega_1,\dots,\omega_n)\le f(\bar{x},\dots,\bar{x}),$$
$$f(u_1,\dots,u_n)+\frac{\varepsilon'}{\lambda}\|(u_1-\omega_1,\dots,u_n-\omega_n)\|_\gamma\ge f(\omega_1,\dots,\omega_n)$$

for all $u_i\in\Omega_i$ $(i=1,\dots,n)$. In view of the monotonicity of φ, the last two inequalities imply conditions (4.3) and (4.4). Condition (4.3) obviously yields (4.5). If φ is differentiable at $\max_{1\le i\le n-1}\|\omega_n+a_i-\omega_i\|$, then condition (4.5) implies (4.6) thanks to Lemma 4.1.1. □

Remark 4.1.4. (i) The expressions in the left-hand sides of the inequalities (4.3) and (4.5) are the main ingredients of the, respectively, global and local slopes of the restriction of the function f given by (4.8) to the complete metric space $\Omega_1 \times \ldots \times \Omega_n$ with the metric induced by the norm (4.1) (γ-slopes [94]), and can be equivalently replaced in these inequalities by the respective slopes. This observation justifies the name 'slope characterisations' adopted in the thesis for this type of estimates as well as the reference to Lemma 4.1.1 in the proof of Theorem 4.1.3.

(ii) The functions in the left-hand sides of the inequalities (4.3), (4.5) and (4.6) are computed at some points $\omega_i \in \Omega_i$ ($i = 1, \ldots, n$) in a neighbourhood of the reference point $\bar{x}$. The size of the neighbourhoods is controlled by the parameters λ and η, which can be chosen arbitrarily small. Note that both the left and the right-hand sides of (4.3), (4.5) and (4.6) also depend on λ and η, and decreasing their values weakens these conditions. Note also that the neighbourhoods for ω_i ($i = 1, \ldots, n-1$) on one hand, and ω_n on the other hand are controlled by different parameters. This reflects the fact that condition (3.12) is not symmetric.

(iii) It is easy to see that the conditions under sup in (4.3) and under lim sup in (4.5) and (4.6) can be complimented by the inequality

$$\max_{1 \le i \le n-1} \|u_n + a_i - u_i\| < \max_{1 \le i \le n-1} \|\omega_n + a_i - \omega_i\|$$

with the convention that supremum over the empty set equals 0.

(iv) Condition (4.4) relates points $\omega_i, \ldots, \omega_n$ to the given vectors $a_1, \ldots, a_n$ and complements the other conditions in Theorem 4.1.3 on the choice of these points. The smaller the norms of these vectors are, the more binding condition (4.4) is; for instance, it follows from conditions $\omega_i \in B_\lambda(\bar{x})$ ($i = 1, \ldots, n-1$) and $\omega_n \in B_\eta(\bar{x})$ that, for each $i = 1, \ldots, n-1$, $\|\omega_n - \omega_i\| < \lambda + \eta$, while condition (4.4) gives an alternative estimate: $\|\omega_n - \omega_i\| < 2 \max_{1 \le i \le n-1} \|a_i\|$, which obviously 'outperforms' the first one when $\max_{1 \le i \le n-1} \|a_i\| < (\lambda + \eta)/2$.

(v) Condition (4.2) in Theorem 4.1.3 can obviously be replaced by a simpler (though stronger!) condition $\varphi\left(\max_{1 \le i \le n-1} \|a_i\|\right) < \varepsilon$. This weakened version of Theorem 4.1.3 is sufficient for characterising the conventional extremality/stationarity properties in Definition 2.3.1. One can go even further and require simply that $\max_{1 \le i \le n-1} \|a_i\| < \varepsilon$. Of course, in this case ε in the inequalities (4.3), (4.5) and (4.6) must be replaced by $\varphi(\varepsilon)$. This creates an interesting phenomenon: φ disappears completely from the assumptions of Theorem 4.1.3 and remains only in its conclusions (which must hold true for any $\varphi \in \mathcal{C}$!)

The importance of the full version of a condition of the type (4.2) for some applications was demonstrated in [170, 172].

(vi) The conclusions of Theorem 4.1.3 are true with any positive number $\gamma \le \dfrac{\lambda}{\eta}$.

Theorem 4.1.3 gives nonlinear primal (slope) characterisations of the asymmetric non-intersection property (3.12), which is a special case of the key non-intersection property (3.35) in the definition of extremality. Now observe that characterisations of even more general than (3.35) symmetric 'local' non-intersection property (P2) can be straightforwardly deduced from Theorem 4.1.3. It is sufficient to add to the given collection of n sets a closed ball $\overline{B}_\eta(\bar{x}) \subset B_\rho(\bar{x})$ and apply Theorem 4.1.3.

Theorem 4.1.5. *Let $\Omega_1, \ldots, \Omega_n$ be closed subsets of a Banach space X, $\bar{x} \in \cap_{i=1}^n \Omega_i$, $a_i \in X$ ($i = 1, \ldots, n$), $\varepsilon > 0$ and $\varphi \in \mathcal{C}$. Suppose that condition* (P2) *is satisfied with some $\rho \in]0, \infty]$, and*

$$\varphi\left(\max_{1 \le i \le n} \|a_i\|\right) < \varphi(d_1(\Omega_1 - a_1, \ldots, \Omega_n - a_n, B_\eta(\bar{x}))) + \varepsilon \tag{4.9}$$

(or simply $\varphi\left(\max_{1\le i\le n}\|a_i\|\right)<\varepsilon$*). Then, for any* $\lambda>0$ *and* $\eta\in]0,\rho[$*, there exist points* $\omega_i\in\Omega_i\cap B_\lambda(\bar{x})$ $(i=1,\ldots,n)$ *and* $x\in B_\eta(\bar{x})$ *such that*

$$\sup_{\substack{u_i\in\Omega_i\ (i=1,\ldots,n),\, u\in B_\eta(\bar{x})\\ (u_1,\ldots,u_n,u)\neq(\omega_1,\ldots,\omega_n,x)}}\frac{\varphi\Big(\max_{1\le i\le n}\|x+a_i-\omega_i\|\Big)-\varphi\Big(\max_{1\le i\le n}\|u+a_i-u_i\|\Big)}{\|(u_1-\omega_1,\ldots,u_n-\omega_n,u-x)\|_\gamma}<\frac{\varepsilon}{\lambda}, \tag{4.10}$$

$$0<\max_{1\le i\le n}\|x+a_i-\omega_i\|\le\max_{1\le i\le n}\|a_i\|, \tag{4.11}$$

where $\gamma:=\dfrac{\lambda}{\eta}$*. As a consequence,*

$$\limsup_{\substack{\Omega_i\ni u_i\to\omega_i\ (i=1,\ldots,n),\, u\to x\\ (u_1,\ldots,u_n,u)\neq(\omega_1,\ldots,\omega_n,x)}}\frac{\varphi\Big(\max_{1\le i\le n}\|x+a_i-\omega_i\|\Big)-\varphi\Big(\max_{1\le i\le n}\|u+a_i-u_i\|\Big)}{\|(u_1-\omega_1,\ldots,u_n-\omega_n,u-x)\|_\gamma}<\frac{\varepsilon}{\lambda}. \tag{4.12}$$

Moreover, if φ *is differentiable at* $\max_{1\le i\le n}\|x+a_i-\omega_i\|$*, then*

$$\varphi'\Big(\max_{1\le i\le n}\|x+a_i-\omega_i\|\Big)\times\limsup_{\substack{\Omega_i\ni u_i\to\omega_i\ (i=1,\ldots,n),\, u\to x\\ (u_1,\ldots,u_n,u)\neq(\omega_1,\ldots,\omega_n,x)}}\frac{\max_{1\le i\le n}\|x+a_i-\omega_i\|-\max_{1\le i\le n}\|u+a_i-u_i\|}{\|(u_1-\omega_1,\ldots,u_n-\omega_n,u-x)\|_\gamma}<\frac{\varepsilon}{\lambda}, \tag{4.13}$$

with the convention $0\cdot(+\infty)=0$.

Remark 4.1.6. The comments concerning Theorem 4.1.3 made in Remark 4.1.4 are, with obvious modifications, applicable to Theorem 4.1.5 and the subsequent statements in this thesis.

All the properties in Definition 2.3.1 as well as the nonlinear primal characterisations of the non-intersection properties in Theorems 4.1.3 and 4.1.5 presume that the sets have a common point. Fortunately Theorem 4.1.3 is rich enough to characterise a non-intersection property without this assumption. Indeed, if $\cap_{i=1}^n\Omega_i=\emptyset$, then, for any points $\omega_i\in\Omega_i$ $(i=1,\ldots,n)$, one can consider the sets $\Omega_i':=\Omega_i-\omega_i$ $(i=1,\ldots,n)$, which obviously satisfy $0\in\cap_{i=1}^n\Omega_i'$. Moreover, after setting $a_i:=\omega_n-\omega_i$ $(i=1,\ldots,n-1)$, one has

$$\bigcap_{i=1}^{n-1}(\Omega_i'-a_i)\cap\Omega_n'=\bigcap_{i=1}^{n-1}(\Omega_i-\omega_n)\cap(\Omega_n-\omega_n)=\bigcap_{i=1}^{n}\Omega_i-\omega_n=\emptyset. \tag{4.14}$$

Thus, Theorem 4.1.3 is applicable and we immediately arrive at the next statement. Note that $d(\Omega_1'-a_1,\ldots,\Omega_{n-1}'-a_{n-1},\Omega_n')=d(\Omega_1,\ldots,\Omega_n)$.

Proposition 4.1.7. *Let* $\Omega_1,\ldots,\Omega_n$ *be closed subsets of a Banach space* X*,* $\omega_i\in\Omega_i$ $(i=1,\ldots,n)$*,* $\varepsilon>0$ *and* $\varphi\in\mathcal{C}$*. Suppose that* $\cap_{i=1}^n\Omega_i=\emptyset$ *and*

$$\varphi\Big(\max_{1\le i\le n-1}\|\omega_n-\omega_i\|\Big)<\varphi(d_1(\Omega_1,\ldots,\Omega_n))+\varepsilon$$

(or simply $\varphi\left(\max_{1\le i\le n-1}\|\omega_n-\omega_i\|\right)<\varepsilon$*). Then, for any* $\lambda>0$ *and* $\eta>0$*, there exist points* $\omega_i'\in\Omega_i\cap B_\lambda(\omega_i)$ $(i=1,\ldots,n-1)$ *and* $\omega_n'\in\Omega_n\cap B_\rho(\omega_n)$ *such that*

$$\sup_{\substack{u_i\in\Omega_i\ (i=1,\ldots,n)\\ (u_1,\ldots,u_n)\neq(\omega_1',\ldots,\omega_n')}}\frac{\varphi\Big(\max_{1\le i\le n-1}\|\omega_n'-\omega_i'\|\Big)-\varphi\Big(\max_{1\le i\le n-1}\|u_n-u_i\|\Big)}{\|(u_1-\omega_1',\ldots,u_n-\omega_n')\|_\gamma}<\frac{\varepsilon}{\lambda},$$

$$0<\max_{1\le i\le n-1}\|\omega_n'-\omega_i'\|\le\max_{1\le i\le n-1}\|\omega_n-\omega_i\|, \tag{4.15}$$

where $\gamma := \dfrac{\lambda}{\eta}$. *As a consequence,*

$$\limsup_{\substack{\Omega_i \ni u_i \to \omega'_i\ (i=1,\ldots,n)\\ (u_1,\ldots,u_n)\neq(\omega'_1,\ldots,\omega'_n)}} \frac{\varphi\Big(\max\limits_{1\le i\le n-1}\|\omega'_n-\omega'_i\|\Big)-\varphi\Big(\max\limits_{1\le i\le n-1}\|u_n-u_i\|\Big)}{\|(u_1-\omega'_1,\ldots,u_n-\omega'_n)\|_\gamma} < \frac{\varepsilon}{\lambda}.$$

Moreover, if φ *is differentiable at* $\max\limits_{1\le i\le n-1}\|\omega'_n-\omega'_i\|$, *then*

$$\varphi'\Big(\max_{1\le i\le n-1}\|\omega'_n-\omega'_i\|\Big)\limsup_{\substack{\Omega_i \ni u_i \to \omega'_i\ (i=1,\ldots,n)\\ (u_1,\ldots,u_n)\neq(\omega'_1,\ldots,\omega'_n)}} \frac{\max\limits_{1\le i\le n-1}\|\omega'_n-\omega'_i\|-\max\limits_{1\le i\le n-1}\|u_n-u_i\|}{\|(u_1-\omega'_1,\ldots,u_n-\omega'_n)\|_\gamma} < \frac{\varepsilon}{\lambda},$$

with the convention $0\cdot(+\infty)=0$.

4.2 Dual Characterisations

In this subsection, we require the function φ to be continuously differentiable.

The dual norm on $(X^*)^{n+1}$ corresponding to (4.1) has the following form:

$$\|(x_1^*,\ldots,x_n^*)\|_\gamma = \sum_{i=1}^{n-1}\|x_i^*\| + \frac{1}{\gamma}\|x_n^*\|, \quad x_1^*,\ldots,x_n^*\in X^*.$$

The next theorem is a dual counterpart of Theorem 4.1.3, providing dual characterisations of the non-intersection property (3.12). It generalises and improves [30, Theorem 6.1].

Theorem 4.2.1. *Let* $\Omega_1,\ldots,\Omega_n$ *be closed subsets of a Banach space* X, $\bar{x}\in\cap_{i=1}^n\Omega_i$, $a_i\in X$ $(i=1,\ldots,n-1)$, $\varepsilon>0$ *and* $\varphi\in\mathcal{C}^1$. *Suppose that conditions* (3.12) *and* (4.2) *(or simply* $\varphi\left(\max\limits_{1\le i\le n-1}\|a_i\|\right)<\varepsilon$*) are satisfied. Then, for any* $\lambda>0$ *and* $\eta>0$,

(i) *there exist points* $\omega_i\in\Omega_i\cap B_\lambda(\bar{x})$ $(i=1,\ldots,n-1)$ *and* $\omega_n\in\Omega_n\cap B_\eta(\bar{x})$ *satisfying condition* (4.4), *and vectors* $x_i^*\in X^*$ $(i=1,\ldots,n)$ *such that*

$$\sum_{i=1}^n x_i^*=0, \quad \sum_{i=1}^{n-1}\|x_i^*\|=1, \tag{4.16}$$

$$\varphi'\Big(\max_{1\le i\le n-1}\|\omega_n+a_i-\omega_i\|\Big)\left(\lambda\sum_{i=1}^{n-1}d\left(x_i^*,N_{\Omega_i}(\omega_i)\right)+\eta d\left(x_n^*,N_{\Omega_n}(\omega_n)\right)\right)<\varepsilon, \tag{4.17}$$

$$\sum_{i=1}^{n-1}\langle x_i^*,\omega_n+a_i-\omega_i\rangle = \max_{1\le i\le n-1}\|\omega_n+a_i-\omega_i\|, \tag{4.18}$$

where N *stands for the Clarke normal cone* ($N=N^C$);

(ii) *if* X *is Asplund, then, for any* $\tau\in]0,1[$, *there exist points* $\omega_i\in\Omega_i\cap B_\lambda(\bar{x})$ $(i=1,\ldots,n-1)$ *and* $\omega_n\in\Omega_n\cap B_\eta(\bar{x})$ *satisfying*

$$0<\max_{1\le i\le n-1}\|\omega_n+a_i-\omega_i\|<\varepsilon, \tag{4.19}$$

and vectors $x_i^* \in X^*$ $(i=1,\ldots,n)$ satisfying conditions (4.16), (4.17), *with N standing for the Fréchet normal cone ($N=N^F$), and*

$$\sum_{i=1}^{n-1} \langle x_i^*, \omega_n + a_i - \omega_i \rangle > \tau \max_{1\le i\le n-1} \|\omega_n + a_i - \omega_i\|. \tag{4.20}$$

Proof. Choose a number ε' satisfying condition (4.7). By Theorem 4.1.3, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i=1,\ldots,n-1)$ and $\omega_n \in \Omega_n \cap B_\eta(\bar{x})$ satisfying condition (4.4) such that condition (4.6) holds with $\gamma := \dfrac{\eta}{\lambda}$ and ε' in place of ε. The last condition yields

$$0 \in \partial^F (g + g_1 + g_2)(\omega_1,\ldots,\omega_n), \tag{4.21}$$

where, for all $u_1,\ldots,u_n \in X$,

$$\begin{aligned} g(u_1,\ldots,u_n) &:= \max_{1\le i\le n-1} \|u_n + a_i - u_i\|, \\ g_1(u_1,\ldots,u_n) &:= \frac{\varepsilon'}{\lambda\varphi'(g(\omega_1,\ldots,\omega_n))} \|(u_1-\omega_1,\ldots,u_n-\omega_n)\|_\gamma, \\ g_2(u_1,\ldots,u_n) &:= \begin{cases} 0 & \text{if } u_i \in \Omega_i \ (i=1,\ldots,n), \\ \infty & \text{otherwise.} \end{cases} \end{aligned} \tag{4.22}$$

Functions g and g_1 are convex and Lipschitz continuous, and g_2 is lower semicontinuous.

From this point we split the proof into two cases.

(i) X is a general Banach space. Condition (4.21) obviously implies $0 \in \partial^C (g+g_1+g_2)(\omega_1,\ldots,\omega_n)$. By the Clarke–Rockafellar subdifferential sum rule (Lemma 2.2.1(ii)), there exist three subgradients: $(v_1^*,\ldots,v_n^*) \in \partial g(\omega_1,\ldots,\omega_n)$, $(v_{11}^*,\ldots,v_{1n}^*) \in \partial g_1(\omega_1,\ldots,\omega_n)$ and $(v_{21}^*,\ldots,v_{2n}^*) \in \partial^C g_2(\omega_1,\ldots,\omega_n)$ such that

$$v_i^* + v_{1i}^* + v_{2i}^* = 0 \quad (i=1,\ldots,n).$$

Observe that $g(\omega_1,\ldots,\omega_n) > 0$ (thanks to (4.4)), and g_2 is the indicator function of the set $\Omega_1 \times \ldots \times \Omega_n$. Hence, by Lemmas 2.2.3 and 2.2.4,

$$\sum_{i=1}^{n} v_i^* = 0, \quad \sum_{i=1}^{n-1} \|v_i^*\| = 1, \tag{4.23}$$

$$\sum_{i=1}^{n-1} \langle v_i^*, \omega_i - a_i - \omega_n \rangle = \max_{1\le i\le n-1} \|\omega_i - a_i - \omega_n\|, \tag{4.24}$$

$$\varphi'(g(\omega_1,\ldots,\omega_n)) \left(\lambda \sum_{i=1}^{n-1} \|v_{1i}^*\| + \eta \|v_{1n}^*\| \right) \le \varepsilon', \tag{4.25}$$

and $v_{2i}^* \in N^C_{\Omega_i}(\omega_i)$ $(i=1,\ldots,n)$. Setting $x_i^* := -v_i^*$ $(i=1,\ldots,n)$, we immediately get conditions (4.16) and (4.18). Moreover,

$$\begin{aligned} \lambda \sum_{i=1}^{n-1} d\left(x_i^*, N^C_{\Omega_i}(\omega_i)\right) + \eta d\left(x_n^*, N^C_{\Omega_n}(\omega_n)\right) &\le \lambda \sum_{i=1}^{n-1} \|v_i^* + v_{2i}^*\| + \eta \|v_n^* + v_{2n}^*\| \\ &= \lambda \sum_{i=1}^{n-1} \|v_{1i}^*\| + \eta \|v_{1n}^*\|, \end{aligned}$$

and condition (4.17) is a consequence of (4.25).

(ii) Let X be Asplund, and $\tau \in]0,1[$. In view of (4.21), we can apply the fuzzy sum rule (Lemma 2.2.1(iii)) to the sum of $g+g_1$ and g_2 followed by the convex sum rule (Lemma 2.2.1(i)) applied to the sum of g and g_1: for any $\xi>0$, there are points $x_i \in X$ and $\omega'_i \in \Omega_i$ $(i=1,\ldots,n)$ and subgradients $(v^*_1,\ldots,v^*_n) \in \partial g(x_1,\ldots,x_n)$, $(v^*_{11},\ldots,v^*_{1n}) \in \partial g_1(x_1,\ldots,x_n)$ and $(v^*_{21},\ldots,v^*_{2n}) \in \partial^F g_2(\omega'_1,\ldots,\omega'_n)$ such that

$$\max_{1\le i\le n} \|x_i-\omega_i\| < \xi, \quad \max_{1\le i\le n} \|\omega'_i-\omega_i\| < \xi, \quad \sum_{i=1}^n \|v^*_i+v^*_{1i}+v^*_{2i}\| < \xi.$$

The number ξ can be chosen small enough so that $\omega'_i \in B_\lambda(\bar{x})$ $(i=1,\ldots,n-1)$, $\omega'_n \in B_\eta(\bar{x})$, $g(x_1,\ldots,x_n)>0$, $g(\omega'_1,\ldots,\omega'_n)>0$, conditions (4.19) and (4.20) are satisfied with ω'_i in place of ω_i $(i=1,\ldots,n)$, and, taking into account the continuity of g and φ',

$$\frac{\varepsilon'}{\varphi'\left(g(x_1,\ldots,x_n)\right)} + \max\{\lambda,\eta\}\xi < \frac{\varepsilon}{\varphi'\left(g(\omega'_1,\ldots,\omega'_n)\right)}.$$

By Lemmas 2.2.3 and 2.2.4, vectors $(v^*_1,\ldots,v^*_n)$ and $(v^*_{11},\ldots,v^*_{1n})$ satisfy conditions (4.23), (4.24) and (4.25) with x_i in place of ω_i $(i=1,\ldots,n)$, and $v^*_{2i} \in N^F_{\Omega_i}(\omega'_i)$ $(i=1,\ldots,n)$. Set $x^*_i := -v^*_i$ $(i=1,\ldots,n)$. Conditions (4.16) follow immediately. Moreover,

$$\begin{aligned}\lambda\sum_{i=1}^{n-1} d\left(x^*_i, N^F_{\Omega_i}(\omega'_i)\right) + \eta d\left(x^*_n, N^F_{\Omega_n}(\omega'_n)\right) &\le \lambda\sum_{i=1}^{n-1}\|v^*_i+v^*_{2i}\| + \eta\|v^*_n+v^*_{2n}\| \\ &\le \lambda\sum_{i=1}^{n-1}\|v^*_{1i}\| + \eta\|v^*_{1n}\| + \max\{\lambda,\eta\}\sum_{i=1}^{n}\|v^*_i+v^*_{1i}+v^*_{2i}\| \\ &\le \frac{\varepsilon'}{\varphi'\left(g(x_1,\ldots,x_n)\right)} + \max\{\lambda,\eta\}\xi < \frac{\varepsilon}{\varphi'\left(g(\omega'_1,\ldots,\omega'_n)\right)},\end{aligned}$$

i.e. condition (4.17) is satisfied with ω'_i in place of ω_i $(i=1,\ldots,n)$. □

Remark 4.2.2. (i) Inequality (4.17) together with the first equality in (4.16) play the key role in asymmetric dual characterisations of extremality/stationarity properties. The inequality ensures that the dual vectors $x^*_1,\ldots,x^*_n$, whose sum is zero, are close to the corresponding normal cones. The second equality in (4.16) is the normalisation condition for the collection of dual vectors; it ensures that the conditions are nontrivial.

(ii) Note that, when φ is linear, the left-hand side of (4.17) is independent of the vectors $a_1,\ldots,a_{n-1}$.

(iii) Conditions (4.18) and (4.20) first appeared explicitly in [170] and were explored further in [30,172]. Conditions of this type relate dual vectors x^*_i and primal space vectors $\omega_n+a_i-\omega_i$ $(i=1,\ldots,n-1)$, and allow to reduce the number of dual vectors involved in checking dual characterisations of extremality/stationarity properties. Such conditions also play an important role in characterisations of intrinsic transversality [127].

(iv) Primal space conditions (4.4) and (4.19) provide additional characterisations of non-intersection properties (cf. Remark 4.1.4(iv)). They have not been used in this context before.

(v) Condition (4.17) with Fréchet normal cones is obviously stronger than its version with Clarke normal cones. On the other hand, conditions (4.19) and (4.20) in the second part of Theorem 4.2.1 are weaker than the corresponding conditions (4.4) and (4.18) in its first part. This is because of the fuzzy sum rule used in its proof.

(vi) Clarke normal cones in part (i) of Theorem 4.2.1 and the other dual space characterisations in this thesis can be replaced by the *G-normal cones* by Ioffe [75], corresponding to the *approximate G-subdifferentials*, which, similar to the Clarke ones, possess an exact sum rule in general Banach spaces; cf. [75, Theorem 4.69].

As before, characterisations of the more general than (3.12) symmetric local non-intersection property (P2) can be straightforwardly deduced from Theorem 4.2.1 by using the same simple trick: adding to the given collection of n sets a closed ball $\overline{B}_\eta(\bar{x}) \subset B_\rho(\bar{x})$. The next statement generalises and improves [30, Theorem 6.3].

Theorem 4.2.3. *Let $\Omega_1,\ldots,\Omega_n$ be closed subsets of a Banach space X, $\bar{x} \in \cap_{i=1}^n \Omega_i$, $a_i \in X$ $(i = 1,\ldots,n)$, $\rho \in]0,\infty]$, $\varepsilon > 0$ and $\varphi \in \mathcal{C}$. Suppose that conditions* (3.35) *and* (4.9) *(or simply $\varphi\left(\max_{1\le i\le n} \|a_i\|\right) < \varepsilon$) are satisfied. Then, for any $\lambda > 0$ and $\eta \in]0,\rho[$,*

(i) *there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1,\ldots,n)$ and $x \in B_\eta(\bar{x})$ satisfying condition* (4.11), *and vectors $x_i^* \in X^*$ $(i = 1,\ldots,n)$ such that*

$$\sum_{i=1}^n \|x_i^*\| = 1, \tag{4.26}$$

$$\varphi'\left(\max_{1\le i\le n} \|x + a_i - \omega_i\|\right)\left(\lambda \sum_{i=1}^n d\left(x_i^*, N_{\Omega_i}(\omega_i)\right) + \eta \left\|\sum_{i=1}^n x_i^*\right\|\right) < \varepsilon, \tag{4.27}$$

$$\sum_{i=1}^n \langle x_i^*, x + a_i - \omega_i\rangle = \max_{1\le i\le n} \|x + a_i - \omega_i\|, \tag{4.28}$$

where N stands for the Clarke normal cone ($N = N^C$);

(ii) *if X is Asplund, then, for any $\tau \in]0,1[$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1,\ldots,n)$ and $x \in B_\eta(\bar{x})$ satisfying*

$$0 < \max_{1\le i\le n} \|x + a_i - \omega_i\| < \varepsilon, \tag{4.29}$$

and vectors $x_i^* \in X^*$ $(i = 1,\ldots,n)$ satisfying conditions (4.26), (4.27), *with N standing for the Fréchet normal cone ($N = N^F$), and*

$$\sum_{i=1}^n \langle x_i^*, x + a_i - \omega_i\rangle > \tau \max_{1\le i\le n} \|x + a_i - \omega_i\|. \tag{4.30}$$

Proof. The statement is a direct consequence of Theorem 4.2.1 applied to the collection of $n+1$ closed sets $\Omega_1,\ldots,\Omega_n,\overline{B}_\eta(\bar{x})$. It is sufficient to notice that, once $x \in B_\eta(\bar{x})$, we have $N_{B_\eta(\bar{x})}(x) = \{0\}$, and consequently, $d\left(-\sum_{i=1}^n x_i^*, N_{B_\eta(\bar{x})}(x)\right) = \left\|\sum_{i=1}^n x_i^*\right\|$.

□

Remark 4.2.4. The comments concerning Theorem 4.2.1 made in Remark 4.2.2 are with obvious modifications applicable to Theorem 4.2.3 and the subsequent statements in this thesis.

The single common point $\bar{x} \in \cap_{i=1}^n \Omega_i$ in Theorems 4.2.1 and 4.2.3 can be replaced by a collection of individual points $\omega_i \in \Omega_i$ $(i = 1,\ldots,n)$ (which always exist as long as the sets are nonempty). The next statement is a consequence of Theorem 4.2.1 applied to the collection of sets $\Omega_i' := \Omega_i - \omega_i$ $(i = 1,\ldots,n)$, which obviously have a common point $0 \in \cap_{i=1}^n \Omega_i'$. It generalises and improves [30, Corollary 6.1].

Proposition 4.2.5. *Let $\Omega_1,\ldots,\Omega_n$ be closed subsets of a Banach space X, $\omega_i \in \Omega_i$ $(i = 1,\ldots,n)$, $a_i \in X$ $(i = 1,\ldots,n-1)$, $\varepsilon > 0$ and $\varphi \in \mathcal{C}^1$. Suppose that*

$$\bigcap_{i=1}^{n-1} (\Omega_i - \omega_i - a_i) \cap (\Omega_n - \omega_n) = \emptyset,$$

$$\varphi\left(\max_{1\le i\le n-1} \|a_i\|\right) < \varphi(d(\Omega_1 - \omega_i - a_1,\ldots,\Omega_{n-1} - \omega_{n-1} - a_{n-1},\Omega_n - \omega_n) + \varepsilon$$

(or simply $\varphi\left(\max\limits_{1\le i\le n-1}\|a_i\|\right)<\varepsilon$*). Set* $a'_i:=a_i+\omega_i-\omega_n$ $(i=1,\ldots,n-1)$*. Then, for any* $\lambda>0$*,* $\eta>0$*,*

(i) *there exist points* $\omega'_i\in\Omega_i\cap B_\lambda(\omega_i)$ $(i=1,\ldots,n-1)$ *and* $\omega'_n\in\Omega_n\cap B_\eta(\omega_n)$ *satisfying*

$$0<\max_{1\le i\le n-1}\|\omega'_n+a'_i-\omega'_i\|\le\max_{1\le i\le n-1}\|a_i\|, \tag{4.31}$$

and vectors $x_i^*\in X^*$ $(i=1,\ldots,n)$ *satisfying conditions* (4.16) *and*

$$\varphi'\left(\max_{1\le i\le n-1}\|\omega'_n+a'_i-\omega'_i\|\right)\times\left(\lambda\sum_{i=1}^{n-1}d\left(x_i^*,N_{\Omega_i}(\omega'_i)\right)+\eta d\left(x_n^*,N_{\Omega_n}(\omega'_n)\right)\right)<\varepsilon, \tag{4.32}$$

$$\sum_{i=1}^{n-1}\langle x_i^*,\omega'_n+a'_i-\omega'_i\rangle=\max_{1\le i\le n-1}\|\omega'_n+a'_i-\omega'_i\|,$$

where N *stands for the Clarke normal cone* ($N=N^C$)*;*

(ii) *if* X *is Asplund, then, for any* $\tau\in]0,1[$*, there exist points* $\omega'_i\in\Omega_i\cap B_\lambda(\omega_i)$ $(i=1,\ldots,n-1)$ *and* $\omega'_n\in\Omega_n\cap B_\eta(\omega_n)$ *satisfying*

$$0<\max_{1\le i\le n-1}\|\omega'_n+a'_i-\omega'_i\|<\varepsilon,$$

and vectors $x_i^*\in X^*$ $(i=1,\ldots,n)$ *satisfying conditions* (4.16), (4.32)*, with* N *standing for the Fréchet normal cone* ($N=N^F$)*, and*

$$\sum_{i=1}^{n-1}\langle x_i^*,\omega'_n+a'_i-\omega'_i\rangle>\tau\max_{1\le i\le n-1}\|\omega'_n+a'_i-\omega'_i\|.$$

Remark 4.2.6. (i) In the particular case when all the points $\omega_i\in\Omega_i$ $(i=1,\ldots,n)$ coincide, i.e. $\omega_1=\ldots=\omega_n=:\bar{x}\in\cap_{i=1}^n\Omega_i$, Proposition 4.2.5 reduces to Theorem 4.2.1.

(ii) Proposition 4.2.5 does not assume the sets $\Omega_1,\ldots,\Omega_n$ to have a common point and, given some individual points $\omega_i\in\Omega_i$ $(i=1,\ldots,n)$, establishes the existence of another collection of points $\omega'_i\in\Omega_i$ $(i=1,\ldots,n)$ with certain properties, each point in a neighbourhood of the corresponding given one. If the sets do have a common point $\bar{x}\in\cap_{i=1}^n\Omega_i$, then, with $\xi:=\max\limits_{1\le i\le n}\|\omega_i-\bar{x}\|$, the estimates in Proposition 4.2.5 yield $\omega'_i\in B_{\lambda+\xi}(\bar{x})$ $(i=1,\ldots,n-1)$ and $\omega'_n\in B_{\eta+\xi}(\bar{x})$. In this form, Proposition 4.2.5 can be considered as a nonlinear extension of [97, Theorem 3.1], which served as the main tool when extending the extremal principle to infinite collections of sets. Note that in the linear case, the conclusions of Proposition 4.2.5 admit significant simplifications (see Section 5.2.5, especially item 3) which make the reduction of Proposition 4.2.5 to (the improved version of) [97, Theorem 3.1] straightforward; cf. [30, Proposition 6.1].

Another consequence of Theorem 4.2.1 provides dual characterisations for a collection of sets with empty intersection. It generalises and improves [30, Theorem 6.2].

Proposition 4.2.7. *Let* $\Omega_1,\ldots,\Omega_n$ *be closed subsets of a Banach space* X*,* $\omega_i\in\Omega_i$ $(i=1,\ldots,n)$*,* $\varepsilon>0$ *and* $\varphi\in\mathcal{C}$*. Suppose that* $\cap_{i=1}^n\Omega_i=\emptyset$ *and*

$$\varphi\left(\max_{1\le i\le n-1}\|\omega_n-\omega_i\|\right)<\varphi(d(\Omega_1,\ldots,\Omega_n))+\varepsilon. \tag{4.33}$$

Then, for any $\lambda>0$ *and* $\eta>0$*,*

(i) *there exist points* $\omega_i' \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1,\ldots,n-1)$ *and* $\omega_n' \in \Omega_n \cap B_\eta(\omega_n)$ *satisfying condition* (4.15), *and vectors* $x_i^* \in X^*$ $(i = 1,\ldots,n)$ *satisfying conditions* (4.16) *and*

$$\varphi'\left(\max_{1\le i\le n-1}\|\omega_n'-\omega_i'\|\right)\left(\lambda\sum_{i=1}^{n-1} d\left(x_i^*, N_{\Omega_i}(\omega_i)\right) + \eta d\left(x_n^*, N_{\Omega_n}(\omega_n)\right)\right) < \varepsilon, \tag{4.34}$$

$$\sum_{i=1}^{n-1}\langle x_i^*, \omega_n'-\omega_i'\rangle = \max_{1\le i\le n-1}\|\omega_n'-\omega_i'\|,$$

where N *stands for the Clarke normal cone* $(N = N^C)$*;*

(ii) *if* X *is Asplund, then, for any* $\tau \in]0,1[$*, there exist points* $\omega_i' \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1,\ldots,n-1)$ *and* $\omega_n' \in \Omega_n \cap B_\eta(\omega_n)$ *satisfying*

$$0 < \max_{1\le i\le n-1}\|\omega_n'-\omega_i'\| < \varepsilon,$$

and vectors $x_i^* \in X^*$ $(i = 1,\ldots,n)$ *satisfying conditions* (4.16), (2.49)*, with* N *standing for the Fréchet normal cone* $(N = N^F)$*, and*

$$\sum_{i=1}^{n-1}\langle x_i^*, \omega_n'-\omega_i'\rangle > \tau \max_{1\le i\le n-1}\|\omega_n'-\omega_i'\|.$$

Proof. It is sufficient to notice that the sets $\Omega_i' := \Omega_i - \omega_i$ $(i = 1,\ldots,n)$ and vectors $a_i := \omega_n - \omega_i$ $(i = 1,\ldots,n-1)$ satisfy $0 \in \bigcap_{i=1}^{n} \Omega_i'$ and (4.14), and apply Theorem 4.2.1. □

Remark 4.2.8. Proposition 4.2.7 can be considered as a nonlinear extension of the two *unified separation theorems* due to Zheng and Ng [172, Theorems 3.1 and 3.4], which correspond to φ being the identity function, and $\lambda = \eta$; cf. [30, Theorem 6.2].

In [172], instead of the distance-like quantity (2.35) in condition (4.33), a slightly more general p-weighted nonintersect index was used with the corresponding q-weighted sums replacing the usual ones in (4.16) and (4.34). This corresponds to considering ℓ_p norms on product spaces and the corresponding ℓ_q dual norms; cf. Remark 2.2.5(ii). In the current thesis, for simplicity only the maximum norm on product spaces is considered together with the corresponding sum norm in the dual space.

Chapter 5

Nonlinear Characterisations of Extremality/Stationarity Properties

In this Chapter, we illustrate the general necessary slope and dual characterisations of non-intersection properties established in Subsection 2.3.1 by applying them to characterising each of the properties in Definition 2.3.1.

Below we formulate a series of necessary conditions for each of the properties. They all follow straightforwardly from the definitions and corresponding statements in Subsection 2.3.1.

Let $\Omega_1, \ldots, \Omega_n$ be closed subsets of a Banach space X and $\bar{x} \in \cap_{i=1}^n \Omega_i$.

5.1 Nonlinear Characterisations

5.1.1 Extremality

Suppose that the collection $\{\Omega_1, \ldots, \Omega_n\}$ is extremal at $\bar{x}$. Then the conditions below hold true.

We start with primal space (slope) necessary extremality conditions.

EC 5.1.1. *For any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying condition*

$$\max_{1 \le i \le n} \|a_i\| < \varepsilon, \tag{5.1}$$

and such that, for any $\lambda > 0$, $\eta > 0$ and $\varphi \in \mathcal{C}$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ and $x \in B_\eta(\bar{x})$ satisfying condition (4.11) and

$$\sup_{\substack{u_i \in \Omega_i \ (i=1,\ldots,n),\, u \in B_\eta(\bar{x}) \\ (u_1,\ldots,u_n,u) \neq (\omega_1,\ldots,\omega_n,x)}} \frac{\varphi\Big(\max\limits_{1 \le i \le n} \|x + a_i - \omega_i\|\Big) - \varphi\Big(\max\limits_{1 \le i \le n} \|u + a_i - u_i\|\Big)}{\|(u_1 - \omega_1, \ldots, u_n - \omega_n, u - x)\|_\gamma} < \frac{\varphi(\varepsilon)}{\lambda}, \tag{5.2}$$

where $\gamma := \dfrac{\lambda}{\eta}$. As a consequence,

$$\limsup_{\substack{\Omega_i \ni u_i \to \omega_i \ (i=1,\ldots,n),\, u \to x \\ (u_1,\ldots,u_n,u) \neq (\omega_1,\ldots,\omega_n,x)}} \frac{\varphi\Big(\max\limits_{1 \le i \le n} \|x + a_i - \omega_i\|\Big) - \varphi\Big(\max\limits_{1 \le i \le n} \|u + a_i - u_i\|\Big)}{\|(u_1 - \omega_1, \ldots, u_n - \omega_n, u - x)\|_\gamma} < \frac{\varphi(\varepsilon)}{\lambda}. \tag{5.3}$$

Moreover, if φ *is differentiable at* $\max\limits_{1\le i\le n}\|x+a_i-\omega_i\|$*, then*

$$\varphi'\Big(\max_{1\le i\le n}\|x+a_i-\omega_i\|\Big) \times \limsup_{\substack{\Omega_i\ni u_i\to\omega_i\ (i=1,\ldots,n),\ u\to x\\ (u_1,\ldots,u_n,u)\ne(\omega_1,\ldots,\omega_n,x)}} \frac{\max\limits_{1\le i\le n}\|x+a_i-\omega_i\|-\max\limits_{1\le i\le n}\|u+a_i-u_i\|}{\|(u_1-\omega_1,\ldots,u_n-\omega_n,u-x)\|_\gamma} < \frac{\varphi(\varepsilon)}{\lambda}, \tag{5.4}$$

with the convention $0\cdot(+\infty)=0$.

Proof. The necessity follows from Definition 2.3.1(i) and Theorem 4.1.5 (with $\rho=+\infty$).

□

EC 5.1.2. *For any* $\varepsilon>0$*, there exist vectors* $a_i\in X$ $(i=1,\ldots,n-1)$ *satisfying condition* (5.1) *and such that, for any* $\lambda>0$*,* $\eta>0$ *and* $\varphi\in\mathcal{C}$*, there exist points* $\omega_i\in\Omega_i\cap B_\lambda(\bar x)$ $(i=1,\ldots,n-1)$ *and* $\omega_n\in\Omega_n\cap B_\eta(\bar x)$ *satisfying condition* (4.4) *and*

$$\sup_{\substack{u_i\in\Omega_i\ (i=1,\ldots,n)\\ (u_1,\ldots,u_n)\ne(\omega_1,\ldots,\omega_n)}} \frac{\varphi\Big(\max\limits_{1\le i\le n-1}\|\omega_n+a_i-\omega_i\|\Big)-\varphi\Big(\max\limits_{1\le i\le n-1}\|u_n+a_i-u_i\|\Big)}{\|(u_1-\omega_1,\ldots,u_n-\omega_n)\|_\gamma} < \frac{\varphi(\varepsilon)}{\lambda}, \tag{5.5}$$

where $\gamma:=\dfrac{\lambda}{\eta}$*. As a consequence,*

$$\limsup_{\substack{\Omega_i\ni u_i\to\omega_i\ (i=1,\ldots,n)\\ (u_1,\ldots,u_n)\ne(\omega_1,\ldots,\omega_n)}} \frac{\varphi\Big(\max\limits_{1\le i\le n-1}\|\omega_n+a_i-\omega_i\|\Big)-\varphi\Big(\max\limits_{1\le i\le n-1}\|u_n+a_i-u_i\|\Big)}{\|(u_1-\omega_1,\ldots,u_n-\omega_n)\|_\gamma} < \frac{\varphi(\varepsilon)}{\lambda}. \tag{5.6}$$

Moreover, if φ *is differentiable at* $\max\limits_{1\le i\le n-1}\|\omega_n+a_i-\omega_i\|$*, then*

$$\varphi'\Big(\max_{1\le i\le n-1}\|\omega_n+a_i-\omega_i\|\Big) \times \limsup_{\substack{\Omega_i\ni u_i\to\omega_i\ (i=1,\ldots,n)\\ (u_1,\ldots,u_n)\ne(\omega_1,\ldots,\omega_n)}} \frac{\max\limits_{1\le i\le n-1}\|\omega_n+a_i-\omega_i\|-\max\limits_{1\le i\le n-1}\|u_n+a_i-u_i\|}{\|(u_1-\omega_1,\ldots,u_n-\omega_n)\|_\gamma} < \frac{\varphi(\varepsilon)}{\lambda}, \tag{5.7}$$

with the convention $0\cdot(+\infty)=0$.

Proof. The necessity follows from Proposition 2.3.4(i) and Theorem 4.1.3.

□

Remark 5.1.3. (i) Condition 5.1.1 is symmetric. It corresponds to the original Definition 2.3.1(i) of extremality; all the sets $\Omega_1,\ldots,\Omega_n$ play the same role; apart from the points $\omega_i\in\Omega_i$ $(i=1,\ldots,n)$, the condition involves an additional point x not belonging to any set. On the other hand, Condition 5.1.2 corresponds to the amended asymmetric characterisation of extremality in Proposition 2.3.4(i) with the last set (it can be any set) playing a special role; the condition does not involve additional points.

(ii) The necessity of Condition 5.1.1 can also be deduced from Condition 5.1.2 applied to the extremal collection of $n+1$ sets $\{\Omega_1,\ldots,\Omega_n,X\}$.

Now we formulate dual space (normal cone) necessary extremality conditions. The next two conditions are for the case of a general Banach space. They employ Clarke normal cones.

EC 5.1.4. *For any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying condition* (5.1) *and such that, for any $\lambda > 0$, $\eta > 0$ and $\varphi \in \mathcal{C}^1$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ and $x \in B_\eta(\bar{x})$ satisfying condition* (4.11)*, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (4.26), (4.28) *and*

$$\varphi'\left(\max_{1\le i\le n} \|x + a_i - \omega_i\|\right)\left(\lambda \sum_{i=1}^{n} d\left(x_i^*, N_{\Omega_i}(\omega_i)\right) + \eta \left\|\sum_{i=1}^{n} x_i^*\right\|\right) < \varphi(\varepsilon). \tag{5.8}$$

Proof. The necessity follows from Definition 2.3.1(i) and Theorem 4.2.3(i) (with $\rho = +\infty$).

□

EC 5.1.5. *For any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n-1)$ satisfying condition*

$$\max_{1\le i\le n-1} \|a_i\| < \varepsilon, \tag{5.9}$$

and such that, for any $\lambda > 0$, $\eta > 0$ and $\varphi \in \mathcal{C}^1$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n-1)$ and $\omega_n \in \Omega_n \cap B_\eta(\bar{x})$ satisfying condition (4.4)*, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (4.16), (4.18) *and*

$$\varphi'\Big(\max_{1\le i\le n-1} \|\omega_n + a_i - \omega_i\|\Big)\left(\lambda \sum_{i=1}^{n-1} d\left(x_i^*, N_{\Omega_i}(\omega_i)\right) + \eta d\left(x_n^*, N_{\Omega_n}(\omega_n)\right)\right) < \varphi(\varepsilon). \tag{5.10}$$

Proof. The necessity follows from Proposition 2.3.4(i) and Theorem 4.2.1(i).

□

The next two conditions are versions of, respectively, Conditions 5.1.4 and 5.1.5 for the case when X is an Asplund space. They employ Fréchet normal cones.

EA 5.1.6. *For any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying condition* (5.1) *and such that, for any $\lambda > 0$, $\eta > 0$, $\tau \in]0, 1[$ and $\varphi \in \mathcal{C}^1$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ and $x \in B_\eta(\bar{x})$ satisfying condition* (4.29)*, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (4.26), (4.30) *and* (5.8)*.*

Proof. The necessity follows from Definition 2.3.1(i) and Theorem 4.2.3(ii) (with $\rho = +\infty$).

□

EA 5.1.7. *For any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n-1)$ satisfying condition* (5.9) *and such that, for any $\lambda > 0$, $\eta > 0$, $\tau \in]0, 1[$ and $\varphi \in \mathcal{C}^1$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n-1)$ and $\omega_n \in \Omega_n \cap B_\eta(\bar{x})$ satisfying condition* (4.19)*, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (4.16), (4.20) *and* (5.10)*.*

Proof. The necessity follows from Proposition 2.3.4(i) and Theorem 4.2.1(ii).

□

Remark 5.1.8. (i) Conditions 5.1.4 and 5.1.6 are symmetric. They correspond to the original Definition 2.3.1(i) of extremality. On the other hand, Conditions 5.1.5 and 5.1.7 correspond to the amended asymmetric definition of extremality in Proposition 2.3.4(i). Cf. Remark 5.1.3(i).

(ii) The necessity of Conditions 5.1.4 and 5.1.6 can also be deduced from Conditions 5.1.5 and 5.1.7 applied to the extremal collection of $n + 1$ sets $\{\Omega_1, \ldots, \Omega_n, X\}$; cf. Remark 5.1.3(ii).

(iii) Inequalities (4.17) and (4.27) play the key role in dual characterisations of extremality properties.

(iv) The derivative φ' involved in the left-hand sides of inequalities (4.17) and (4.27) plays no role when φ is linear. Otherwise, in view of conditions (4.4), (4.11), (4.19) and (4.29), the behaviour of φ' near 0 becomes important, e.g., when φ is a power function, i.e. $\varphi(t) = t^q$, the cases $0 < q < 1$ and $q > 1$ are strongly different.

(v) Conditions (4.18), (4.20), (4.28) and (4.30) relate the corresponding primal and dual vectors involved in the dual characterisations of extremality; cf. Remark 4.2.2(iii). Conditions of this type have been used in [30, 170, 172].

(vi) Primal space conditions (4.4), (4.11), (4.19) and (4.29), provide additional characterisations of non-intersection properties. They have not been used in this context before.

The comments concerning the necessary extremality conditions made in Remark 5.1.8 are with obvious modifications applicable to the formulated below necessary conditions for the other extremality/stationarity properties.

5.1.2 Local Extremality

Suppose that the collection $\{\Omega_1, \ldots, \Omega_n\}$ is locally extremal at $\bar{x}$ with some $\rho \in]0, +\infty]$. Then the conditions below hold true.

We start with primal space (slope) necessary local extremality conditions.

LEC 5.1.9. *For any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying condition* (5.1) *and such that, for any $\lambda > 0$, $\eta \in]0, \rho[$ and $\varphi \in \mathcal{C}$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ and $x \in B_\eta(\bar{x})$ satisfying conditions* (4.11) *and* (5.2), *where $\gamma := \dfrac{\lambda}{\eta}$. As a consequence, condition* (5.3) *is satisfied. Moreover, if φ is differentiable at $\max\limits_{1 \le i \le n} \|x + a_i - \omega_i\|$, then condition* (5.4) *is satisfied, with the convention $0 \cdot (+\infty) = 0$.*

Proof. The necessity follows from Definition 2.3.1(ii) and Theorem 4.1.5 applied to the collection of $n+1$ closed sets $\{\Omega_1, \ldots, \Omega_n, \overline{B}_\eta(\bar{x})\}$.

□

LEC 5.1.10. *For any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n-1)$ satisfying condition* (5.9) *and such that, for any $\lambda > 0$, $\eta \in]0, \rho[$ and $\varphi \in \mathcal{C}$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n-1)$ and $\omega_n \in \Omega_n \cap B_\eta(\bar{x})$ satisfying conditions* (4.4) *and* (5.5), *where $\gamma := \dfrac{\lambda}{\eta}$. As a consequence, condition* (5.6) *is satisfied. Moreover, if φ is differentiable at $\max\limits_{1 \le i \le n-1} \|\omega_n + a_i - \omega_i\|$, then condition* (5.7) *is satisfied, with the convention $0 \cdot (+\infty) = 0$.*

Proof. The necessity follows from Proposition 2.3.4(ii) and Theorem 4.1.3 applied to the collection of n closed sets $\{\Omega_1, \ldots, \Omega_{n-1}, \Omega_n \cap \overline{B}_\eta(\bar{x})\}$.

□

Now we formulate dual space (normal cone) necessary local extremality conditions. The next two conditions are for the case of a general Banach space. They employ Clarke normal cones.

LEC 5.1.11. *For any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying condition* (5.1) *and such that, for any $\lambda > 0$, $\eta \in]0, \rho[$ and $\varphi \in \mathcal{C}^1$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ and $x \in B_\eta(\bar{x})$ satisfying condition* (4.11)*, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (4.26)*,* (4.28) *and* (5.8)*.*

Proof. The necessity follows from Definition 2.3.1(ii) and Theorem 4.2.3(i).

□

LEC 5.1.12. *For any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n-1)$ satisfying condition* (5.9) *and such that, for any $\lambda > 0$, $\eta \in]0, \rho[$ and $\varphi \in \mathcal{C}^1$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n-1)$ and $\omega_n \in \Omega_n \cap B_\eta(\bar{x})$ satisfying condition* (4.4)*, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (4.16)*,* (4.18) *and* (5.10)*.*

Proof. The necessity follows from Proposition 2.3.4(ii) and Theorem 4.2.1(i).

□

The next two conditions are versions of, respectively, Conditions 5.1.11 and 5.1.12 for the case when X is an Asplund space. They employ Fréchet normal cones.

LEA 5.1.13. *For any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying condition* (5.1) *and such that, for any $\lambda > 0$, $\eta \in]0, \rho[$, $\tau \in]0, 1[$ and $\varphi \in \mathcal{C}^1$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ and $x \in B_\eta(\bar{x})$ satisfying condition* (4.29)*, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (4.26)*,* (4.30) *and* (5.8)*.*

Proof. The necessity follows from Definition 2.3.1(ii) and Theorem 4.2.3(ii).

□

LEA 5.1.14. *For any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n-1)$ satisfying condition* (5.9) *and such that, for any $\lambda > 0$, $\eta \in]0, \rho[$, $\tau \in]0, 1[$ and $\varphi \in \mathcal{C}^1$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n-1)$ and $\omega_n \in \Omega_n \cap B_\eta(\bar{x})$ satisfying condition* (4.19)*, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (4.16)*,* (4.20) *and* (5.10)*.*

Proof. The necessity follows from Proposition 2.3.4(ii) and Theorem 4.2.1(ii).

□

Remark 5.1.15. Conditions 5.1.9–5.1.14 are weaker than the corresponding Conditions 5.1.1–5.1.7 for the extremality because of the additional requirement $\eta < \rho$. Note that η not only controls the choice of x and ω_n, but is also involved in conditions (4.3), (4.10), (4.17) and (4.27).

5.1.3 Stationarity

Suppose that the collection $\{\Omega_1, \ldots, \Omega_n\}$ is stationary at $\bar{x}$. Then the conditions below hold true.

We start with primal space (slope) necessary stationarity conditions.

SC 5.1.16. *For any $\varepsilon > 0$, there exist a number $\rho \in]0, \varepsilon[$ and vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying condition*

$$\max_{1 \le i \le n} \|a_i\| < \rho\varepsilon, \tag{5.11}$$

and such that, for any $\lambda > 0$, $\eta \in]0, \rho]$ *and* $\varphi \in \mathcal{C}$, *there exist points* $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ *and* $x \in B_\eta(\bar{x})$ *satisfying conditions* (4.11) *and* (5.2), *where* $\gamma := \dfrac{\lambda}{\eta}$. *As a consequence, condition* (5.3) *is satisfied. Moreover, if* φ *is differentiable at* $\max\limits_{1 \le i \le n} \|x + a_i - \omega_i\|$, *then condition* (5.4) *is satisfied, with the convention* $0 \cdot (+\infty) = 0$.

Proof. The necessity follows from Definition 2.3.1(iii) and Theorem 4.1.5.

□

SC 5.1.17. *For any* $\varepsilon > 0$, *there exist a number* $\rho \in]0, \varepsilon[$ *and vectors* $a_i \in X$ $(i = 1, \ldots, n-1)$, *satisfying condition*

$$\max_{1 \le i \le n-1} \|a_i\| < \rho\varepsilon, \tag{5.12}$$

and such that, for any $\lambda > 0$, $\eta \in]0, \rho]$ *and* $\varphi \in \mathcal{C}$, *there exist points* $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n-1)$ *and* $\omega_n \in \Omega_n \cap B_\eta(\bar{x})$ *satisfying conditions* (4.4) *and* (5.5), *where* $\gamma := \dfrac{\lambda}{\eta}$. *As a consequence, condition* (5.6) *is satisfied. Moreover, if* φ *is differentiable at* $\max\limits_{1 \le i \le n-1} \|\omega_n + a_i - \omega_i\|$, *then condition* (5.7) *is satisfied, with the convention* $0 \cdot (+\infty) = 0$.

Proof. The necessity follows from Proposition 2.3.4(iii) and Theorem 4.1.3.

□

Now we formulate dual space (normal cone) necessary stationary conditions. The next two conditions are for the case of a general Banach space. They employ Clarke normal cones.

SC 5.1.18. *For any* $\varepsilon > 0$, *there exist a number* $\rho \in]0, \varepsilon[$ *and vectors* $a_i \in X$ $(i = 1, \ldots, n)$ *satisfying condition* (5.11) *and such that, for any* $\lambda > 0$, $\eta \in]0, \rho]$ *and* $\varphi \in \mathcal{C}^1$, *there exist points* $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ *and* $x \in B_\eta(\bar{x})$ *satisfying condition* (4.11), *and vectors* $x_i^* \in X^*$ $(i = 1, \ldots, n)$ *satisfying conditions* (4.26), (4.28) *and* (5.8).

Proof. The necessity follows from Definition 2.3.1(iii) and Theorem 4.2.3(i).

□

SC 5.1.19. *For any* $\varepsilon > 0$, *there exist a number* $\rho \in]0, \varepsilon[$ *and vectors* $a_i \in X$ $(i = 1, \ldots, n-1)$ *satisfying condition* (5.12) *and such that, for any* $\lambda > 0$, $\eta \in]0, \rho]$ *and* $\varphi \in \mathcal{C}^1$, *there exist points* $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n-1)$ *and* $\omega_n \in \Omega_n \cap B_\eta(\bar{x})$ *satisfying condition* (4.4), *and vectors* $x_i^* \in X^*$ $(i = 1, \ldots, n)$ *satisfying conditions* (4.16), (4.18) *and* (5.10).

Proof. The necessity follows from Proposition 2.3.4(iii) and Theorem 4.2.1(i).

□

The next two conditions are versions of, respectively, Conditions 5.1.18 and 5.1.19 for the case when X is an Asplund space. They employ Fréchet normal cones.

SA 5.1.20. *For any* $\varepsilon > 0$, *there exist a number* $\rho \in]0, \varepsilon[$ *and vectors* $a_i \in X$ $(i = 1, \ldots, n)$ *satisfying condition* (5.11) *and such that, for any* $\lambda > 0$, $\eta \in]0, \rho]$, $\tau \in]0, 1[$ *and* $\varphi \in \mathcal{C}^1$, *there exist points* $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ *and* $x \in B_\eta(\bar{x})$ *satisfying*

$$0 < \max_{1 \le i \le n} \|x + a_i - \omega_i\| < \varepsilon\rho, \tag{5.13}$$

and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ *satisfying conditions* (4.26), (4.30) *and* (5.8).

Proof. The necessity follows from Definition 2.3.1(iii) and Theorem 4.2.3(ii).

□

SA 5.1.21. *For any $\varepsilon > 0$, there exist a number $\rho \in]0, \varepsilon[$ and vectors $a_i \in X$ $(i = 1, \ldots, n-1)$ satisfying condition* (5.12) *and such that, for any $\lambda > 0$, $\eta \in]0, \rho]$, $\tau \in]0, 1[$ and $\varphi \in \mathcal{C}^1$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n-1)$ and $\omega_n \in \Omega_n \cap B_\eta(\bar{x})$ satisfying*

$$0 < \max_{1 \le i \le n-1} \|\omega_n + a_i - \omega_i\| < \varepsilon\rho, \tag{5.14}$$

and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions (4.16), (4.20) *and* (5.10).

Proof. The necessity follows from Proposition 2.3.4(iii) and Theorem 4.2.1(ii).

□

Remark 5.1.22. Conditions 5.1.16–5.1.21 are weaker than the corresponding Conditions 5.1.9–5.1.14 for the local extremality because of the additional requirement $\rho < \varepsilon$.

5.1.4 Approximate Stationarity

Suppose that the collection $\{\Omega_1, \ldots, \Omega_n\}$ is approximately stationary at $\bar{x}$. Then the conditions below hold true.

We start with primal space (slope) necessary approximate stationarity conditions.

ASC 5.1.23. *For any $\varepsilon > 0$, there exist a number $\rho \in]0, \varepsilon[$, points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ and vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying condition* (5.11) *and such that, for any $\lambda > 0$, $\eta \in]0, \rho]$ and $\varphi \in \mathcal{C}$, there exist points $\omega_i' \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1, \ldots, n)$ and $x \in \eta\mathbb{B}$ such that*

$$\sup_{\substack{u_i \in \Omega_i\ (i=1,\ldots,n),\ u \in \eta\mathbb{B} \\ (u_1,\ldots,u_n,u) \neq (\omega_1',\ldots,\omega_n',x)}} \frac{\varphi\Big(\max\limits_{1 \le i \le n} \|x + a_i' - \omega_i'\|\Big) - \varphi\Big(\max\limits_{1 \le i \le n} \|u + a_i' - u_i\|\Big)}{\|(u_1 - \omega_1', \ldots, u_n - \omega_n', u - x)\|_\gamma} < \frac{\varphi(\varepsilon)}{\lambda},$$

$$0 < \max_{1 \le i \le n} \|x + a_i' - \omega_i'\| \le \max_{1 \le i \le n} \|a_i\|, \tag{5.15}$$

where $\gamma := \dfrac{\lambda}{\eta}$ and $a_i' := a_i + \omega_i$ $(i = 1, \ldots, n)$. As a consequence

$$\limsup_{\substack{\Omega_i \ni u_i \to \omega_i'\ (i=1,\ldots,n),\ u \to 0 \\ (u_1,\ldots,u_n,u) \neq (\omega_1',\ldots,\omega_n',x)}} \frac{\varphi\Big(\max\limits_{1 \le i \le n} \|x + a_i' - \omega_i'\|\Big) - \varphi\Big(\max\limits_{1 \le i \le n} \|u + a_i' - u_i\|\Big)}{\|(u_1 - \omega_1', \ldots, u_n - \omega_n', u - x)\|_\gamma} < \frac{\varphi(\varepsilon)}{\lambda}.$$

Moreover, if φ is differentiable at $\max\limits_{1 \le i \le n} \|x + a_i' - \omega_i'\|$, then

$$\varphi'\Big(\max_{1 \le i \le n} \|x + a_i' - \omega_i'\|\Big) \times \limsup_{\substack{\Omega_i \ni u_i \to \omega_i'\ (i=1,\ldots,n),\ u \to 0 \\ (u_1,\ldots,u_n,u) \neq (\omega_1',\ldots,\omega_n',x)}} \frac{\max\limits_{1 \le i \le n} \|x + a_i' - \omega_i'\| - \max\limits_{1 \le i \le n} \|u + a_i' - u_i\|}{\|(u_1 - \omega_1', \ldots, u_n - \omega_n', u - x)\|_\gamma} < \frac{\varphi(\varepsilon)}{\lambda}.$$

Proof. The necessity follows from Definition 2.3.1(iv) and Theorem 4.1.5 applied to the collection of closed sets $\{\Omega_1 - \omega_1, \ldots, \Omega_n - \omega_n\}$ and their common point 0.

□

ASC 5.1.24. *For any $\varepsilon > 0$, there exist a number $\rho \in]0, \varepsilon[$, points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ $(i = 1, \ldots, n)$ and vectors $a_i \in X$ $(i = 1, \ldots, n-1)$ satisfying condition* (5.12) *and such that, for any $\lambda > 0$, $\eta \in]0, \rho]$ and $\varphi \in \mathcal{C}$, there exist points $\omega_i' \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1, \ldots, n-1)$ and $\omega_n' \in \Omega_n \cap B_\eta(\omega_n)$ satisfying condition* (4.31) *and*

$$\sup_{\substack{u_i \in \Omega_i\ (i=1,\ldots,n),\, u_n \in B_\eta(\omega_n) \\ (u_1,\ldots,u_n) \neq (\omega_1',\ldots,\omega_n')}} \frac{\varphi\Big(\max\limits_{1 \le i \le n-1} \|\omega_n' + a_i' - \omega_i'\|\Big) - \varphi\Big(\max\limits_{1 \le i \le n-1} \|u_n + a_i' - u_i\|\Big)}{\|(u_1 - \omega_1', \ldots, u_n - \omega_n')\|_\gamma} < \frac{\varphi(\varepsilon)}{\lambda},$$

where $\gamma := \dfrac{\lambda}{\eta}$ and $a_i' := a_i + \omega_i - \omega_n$ $(i = 1, \ldots, n-1)$. As a consequence,

$$\limsup_{\substack{\Omega_i \ni u_i \to \omega_i'\ (i=1,\ldots,n) \\ (u_1,\ldots,u_n) \neq (\omega_1',\ldots,\omega_n')}} \frac{\varphi\Big(\max\limits_{1 \le i \le n-1} \|\omega_n' + a_i' - \omega_i'\|\Big) - \varphi\Big(\max\limits_{1 \le i \le n-1} \|u_n + a_i' - u_i\|\Big)}{\|(u_1 - \omega_1', \ldots, u_n - \omega_n')\|_\gamma} < \frac{\varphi(\varepsilon)}{\lambda},$$

Moreover, if φ is differentiable at $\max\limits_{1 \le i \le n-1} \|\omega_n' + a_i' - \omega_i'\|$, then

$$\varphi'\Big(\max_{1 \le i \le n-1} \|\omega_n' + a_i' - \omega_i'\|\Big) \times \limsup_{\substack{\Omega_i \ni u_i \to \omega_i'\ (i=1,\ldots,n) \\ (u_1,\ldots,u_n) \neq (\omega_1',\ldots,\omega_n')}} \frac{\max\limits_{1 \le i \le n-1} \|\omega_n' + a_i' - \omega_i'\| - \max\limits_{1 \le i \le n-1} \|u_n + a_i' - u_i\|}{\|(u_1 - \omega_1', \ldots, u_n - \omega_n')\|_\gamma} < \frac{\varphi(\varepsilon)}{\lambda}.$$

Proof. The necessity follows from Proposition 2.3.4(iv) and Theorem 4.1.3 applied to the collection of closed sets $\{\Omega_1 - \omega_1, \ldots, \Omega_{n-1} - \omega_{n-1}, (\Omega_n - \omega_n) \cap (\rho\overline{\mathbb{B}})\}$ and their common point 0.

□

Now we formulate dual space (normal cone) necessary approximate stationary conditions. The next two conditions are for the case of a general Banach space. They employ Clarke normal cones.

ASC 5.1.25. *For any $\varepsilon > 0$, there exist a number $\rho \in]0, \varepsilon[$, points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ and vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying condition* (5.11) *and such that, for any $\lambda > 0$, $\eta \in]0, \rho]$ and $\varphi \in \mathcal{C}$, there exist points $\omega_i' \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1, \ldots, n)$ and $x \in \eta\mathbb{B}$ satisfying condition* (5.15), *and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying condition* (4.26) *and*

$$\varphi'\left(\max_{1 \le i \le n} \|x + a_i' - \omega_i'\|\right)\left(\lambda \sum_{i=1}^n d\left(x_i^*, N_{\Omega_i}(\omega_i')\right) + \eta \left\|\sum_{i=1}^n x_i^*\right\|\right) < \varphi(\varepsilon), \tag{5.16}$$

$$\sum_{i=1}^n \langle x_i^*, x + a_i' - \omega_i' \rangle = \max_{1 \le i \le n} \|x + a_i' - \omega_i'\|, \tag{5.17}$$

where $a_i' := a_i + \omega_i$ $(i = 1, \ldots, n)$.

Proof. The necessity follows from Definition 2.3.1(iv) and Theorem 4.2.3(i) applied to the collection of closed sets $\{\Omega_1 - \omega_1, \ldots, \Omega_n - \omega_n\}$ and their common point 0.

□

ASC 5.1.26. *For any $\varepsilon > 0$, there exist a number $\rho \in]0, \varepsilon[$, points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ $(i = 1, \ldots, n)$ and vectors $a_i \in X$ $(i = 1, \ldots, n-1)$ satisfying condition* (5.12) *and such that, for any $\lambda > 0$, $\eta \in]0, \rho]$ and*

$\varphi \in \mathcal{C}^1$, *there exist points* $\omega_i' \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1, \ldots, n-1)$ *and* $\omega_n' \in \Omega_n \cap B_\eta(\omega_n)$ *satisfying condition* (4.31), *and vectors* $x_i^* \in X^*$ $(i = 1, \ldots, n)$ *satisfying condition* (4.16) *and*

$$\varphi'\left(\max_{1\le i\le n-1} \|\omega_n' + a_i' - \omega_i'\|\right)\left(\lambda \sum_{i=1}^{n-1} d\left(x_i^*, N_{\Omega_i}(\omega_i')\right) + \eta d\left(x_n^*, N_{\Omega_n}(\omega_i')\right)\right) < \varphi(\varepsilon), \tag{5.18}$$

$$\sum_{i=1}^{n-1} \langle x_i^*, \omega_n' + a_i' - \omega_i'\rangle = \max_{1\le i\le n-1} \|\omega_n' + a_i' - \omega_i'\|, \tag{5.19}$$

where $a_i' := a_i + \omega_i - \omega_n$ $(i = 1, \ldots, n-1)$.

Proof. The necessity follows from Proposition 2.3.4(iv) and Theorem 4.2.1(i).

□

The next two conditions are versions of, respectively, Conditions 5.1.25 and 5.1.26 for the case when X is an Asplund space. They employ Fréchet normal cones.

ASA 5.1.27. *For any* $\varepsilon > 0$, *there exist a number* $\rho \in]0, \varepsilon[$, *points* $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ *and vectors* $a_i \in X$ $(i = 1, \ldots, n)$ *satisfying condition* (5.11) *and such that, for any* $\lambda > 0$, $\eta \in]0, \rho]$, $\tau \in]0, 1[$ *and* $\varphi \in \mathcal{C}$, *there exist points* $\omega_i' \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1, \ldots, n)$ *and* $x \in \eta\mathbb{B}$ *satisfying*

$$0 < \max_{1\le i\le n} \|x + a_i' - \omega_i'\| < \varepsilon\rho, \tag{5.20}$$

where $a_i' := a_i + \omega_i$ $(i = 1, \ldots, n)$, *and vectors* $x_i^* \in X^*$ $(i = 1, \ldots, n)$ *satisfying conditions* (4.26), (5.16) *and*

$$\sum_{i=1}^{n} \langle x_i^*, x + a_i' - \omega_i'\rangle > \tau \max_{1\le i\le n} \|x + a_i' - \omega_i'\|. \tag{5.21}$$

Proof. The necessity follows from Definition 2.3.1(iv) and Theorem 4.2.3(ii) applied to the collection of closed sets $\{\Omega_1 - \omega_1, \ldots, \Omega_n - \omega_n\}$ and their common point 0.

□

ASA 5.1.28. *For any* $\varepsilon > 0$, *there exist a number* $\rho \in]0, \varepsilon[$, *points* $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ $(i = 1, \ldots, n)$ *and vectors* $a_i \in X$ $(i = 1, \ldots, n-1)$ *satisfying condition* (5.12) *and such that, for any* $\lambda > 0$, $\eta \in]0, \rho]$, $\tau \in]0, 1[$ *and* $\varphi \in \mathcal{C}^1$, *there exist points* $\omega_i' \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1, \ldots, n-1)$ *and* $\omega_n' \in \Omega_n \cap B_\eta(\omega_n)$ *satisfying*

$$0 < \max_{1\le i\le n-1} \|\omega_n' + a_i' - \omega_i'\| < \varepsilon\rho, \tag{5.22}$$

and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ *satisfying conditions* (4.16), (5.18) *and*

$$\sum_{i=1}^{n-1} \langle x_i^*, \omega_n' + a_i' - \omega_i'\rangle > \tau \max_{1\le i\le n-1} \|\omega_n' + a_i' - \omega_i'\|, \tag{5.23}$$

where $a_i' := a_i + \omega_i - \omega_n$ $(i = 1, \ldots, n-1)$.

Proof. The necessity follows from Proposition 2.3.4(iv) and Theorem 4.2.1(ii).

□

Remark 5.1.29. Conditions 5.1.23–5.1.28 are weaker than the corresponding Conditions 5.1.16–5.1.21 for the stationarity because of the additional collection of points $\omega_i \in \Omega_i$ $(i = 1, \ldots, n)$ present in all of them. Conditions 5.1.16–5.1.21 correspond to setting $\omega_i := \bar{x}$ $(i = 1, \ldots, n)$ in Conditions 5.1.23–5.1.28. Note that the other collection $\omega_i' \in \Omega_i$ $(i = 1, \ldots, n)$ in Conditions 5.1.23–5.1.28 corresponds to $\omega_i \in \Omega_i$ $(i = 1, \ldots, n)$ in Conditions 5.1.16–5.1.21.

5.2 Hölder Characterisations

In this section, we formulate realisations of the necessary conditions of the extremality/stationarity properties from Section 5.1 in the Hölder setting.

Let $\Omega_1, \ldots, \Omega_n$ be closed subsets of a Banach space X, $\bar{x} \in \cap_{i=1}^n \Omega_i$ and $q > 0$.

The conditions below correspond to setting $\varphi(t) := \alpha t^q$ $(t > 0)$ with some $\alpha > 0$ in the corresponding conditions from Section 5. 'H' in the labels of the conditions stands for 'Hölder'.

5.2.1 Extremality

Suppose that the collection $\{\Omega_1, \ldots, \Omega_n\}$ is extremal at $\bar{x}$. Then the conditions below hold true.

HE 5.2.1. *For any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying condition* (5.1) *and such that, for any $\lambda > 0$ and $\eta > 0$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ and $x \in B_\eta(\bar{x})$ satisfying condition* (4.11) *and*

$$\sup_{\substack{u_i \in \Omega_i\ (i=1,\ldots,n),\, u \in B_\eta(\bar{x}) \\ (u_1,\ldots,u_n,u) \neq (\omega_1,\ldots,\omega_n,x)}} \frac{\max\limits_{1\le i\le n} \|x + a_i - \omega_i\|^q - \max\limits_{1\le i\le n} \|u + a_i - u_i\|^q}{\|(u_1 - \omega_1, \ldots, u_n - \omega_n, u - x)\|_\gamma} < \frac{\varepsilon^q}{\lambda}, \tag{5.24}$$

where $\gamma := \dfrac{\lambda}{\eta}$. As a consequence,

$$q \max_{1\le i\le n} \|x + a_i - \omega_i\|^{q-1} \times \limsup_{\substack{\Omega_i \ni u_i \to \omega_i\ (i=1,\ldots,n),\, u \to x \\ (u_1,\ldots,u_n,u) \neq (\omega_1,\ldots,\omega_n,x)}} \frac{\max\limits_{1\le i\le n} \|x + a_i - \omega_i\| - \max\limits_{1\le i\le n} \|u + a_i - u_i\|}{\|(u_1 - \omega_1, \ldots, u_n - \omega_n, u - x)\|_\gamma} < \frac{\varepsilon^q}{\lambda}. \tag{5.25}$$

HE 5.2.2. *For any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n-1)$ satisfying condition* (5.9) *and such that, for any $\lambda > 0$ and $\eta > 0$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n-1)$ and $\omega_n \in \Omega_n \cap B_\eta(\bar{x})$ satisfying condition* (4.4) *and*

$$\sup_{\substack{u_i \in \Omega_i\ (i=1,\ldots,n) \\ (u_1,\ldots,u_n) \neq (\omega_1,\ldots,\omega_n)}} \frac{\max\limits_{1\le i\le n-1} \|\omega_n + a_i - \omega_i\|^q - \max\limits_{1\le i\le n-1} \|u_n + a_i - u_i\|^q}{\|(u_1 - \omega_1, \ldots, u_n - \omega_n)\|_\gamma} < \frac{\varepsilon^q}{\lambda}, \tag{5.26}$$

where $\gamma := \dfrac{\lambda}{\eta}$. As a consequence,

$$q \max_{1\le i\le n-1} \|\omega_n + a_i - \omega_i\|^{q-1} \times \limsup_{\substack{\Omega_i \ni u_i \to \omega_i\ (i=1,\ldots,n) \\ (u_1,\ldots,u_n) \neq (\omega_1,\ldots,\omega_n)}} \frac{\max\limits_{1\le i\le n-1} \|\omega_n + a_i - \omega_i\| - \max\limits_{1\le i\le n-1} \|u_n + a_i - u_i\|}{\|(u_1 - \omega_1, \ldots, u_n - \omega_n)\|_\gamma} < \frac{\varepsilon^q}{\lambda}. \tag{5.27}$$

The next two conditions are for the case of a general Banach space. They employ Clarke normal cones.

HE 5.2.3. *For any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying condition* (5.1) *and such that, for any $\lambda > 0$ and $\eta > 0$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ and $x \in B_\eta(\bar{x})$ satisfying condition* (4.11)*, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (4.26), (4.28) *and*

$$q \max_{1\le i\le n} \|x + a_i - \omega_i\|^{q-1} \left(\lambda \sum_{i=1}^n d\left(x_i^*, N_{\Omega_i}(\omega_i)\right) + \eta \left\| \sum_{i=1}^n x_i^* \right\| \right) < \varepsilon^q. \tag{5.28}$$

HE 5.2.4. *For any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n-1)$ satisfying condition* (5.9) *and such that, for any $\lambda > 0$ and $\eta > 0$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n-1)$ and $\omega_n \in \Omega_n \cap B_\eta(\bar{x})$ satisfying condition* (4.4)*, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (4.16), (4.18) *and*

$$q \max_{1 \le i \le n-1} \|\omega_n + a_i - \omega_i\|^{q-1} \left(\lambda \sum_{i=1}^{n-1} d\left(x_i^*, N_{\Omega_i}(\omega_i)\right) + \eta d\left(x_n^*, N_{\Omega_n}(\omega_n)\right) \right) < \varepsilon^q. \tag{5.29}$$

The next two conditions are versions of, respectively, Conditions 5.2.3 and 5.2.4 for the case when X is an Asplund space. They employ Fréchet normal cones.

HEA 5.2.5. *For any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying condition* (5.1) *and such that, for any $\lambda > 0$, $\eta > 0$ and $\tau \in]0, 1[$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ and $x \in B_\eta(\bar{x})$ satisfying condition* (4.29)*, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (4.26), (4.30) *and* (5.28)*.*

HEA 5.2.6. *For any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n-1)$ satisfying condition* (5.9) *and such that, for any $\lambda > 0$, $\eta > 0$ and $\tau \in]0, 1[$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n-1)$ and $\omega_n \in \Omega_n \cap B_\eta(\bar{x})$ satisfying condition* (4.19)*, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (4.16), (4.20) *and* (5.29)*.*

5.2.2 Local Extremality

Suppose that the collection $\{\Omega_1, \ldots, \Omega_n\}$ is locally extremal at $\bar{x}$ with some $\rho \in]0, +\infty]$. Then the conditions below hold true.

HLE 5.2.7. *For any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying condition* (5.1) *and such that, for any $\lambda > 0$ and $\eta \in]0, \rho[$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ and $x \in B_\eta(\bar{x})$ satisfying conditions* (4.11) *and* (5.24)*. As a consequence, condition* (5.25) *is satisfied.*

HLE 5.2.8. *For any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n-1)$ satisfying condition* (5.9) *and such that, for any $\lambda > 0$ and $\eta \in]0, \rho[$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n-1)$ and $\omega_n \in \Omega_n \cap B_\eta(\bar{x})$ satisfying condition* (4.4) *and*

$$\sup_{\substack{u_i \in \Omega_i \ (i=1,\ldots,n),\, u_n \in B_\eta(\bar{x}) \\ (u_1,\ldots,u_n) \ne (\omega_1,\ldots,\omega_n)}} \frac{\max\limits_{1 \le i \le n-1} \|\omega_n + a_i - \omega_i\|^q - \max\limits_{1 \le i \le n-1} \|u_n + a_i - u_i\|^q}{\|(u_1 - \omega_1, \ldots, u_n - \omega_n)\|_\gamma} < \frac{\varepsilon^q}{\lambda}.$$

As a consequence, condition (5.27) *is satisfied.*

The next two conditions are for the case of a general Banach space. They employ Clarke normal cones.

HLE 5.2.9. *For any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying condition* (5.1) *and such that, for any $\lambda > 0$ and $\eta \in]0, \rho[$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ and $x \in B_\eta(\bar{x})$ satisfying condition* (4.11)*, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (4.26), (4.28) *and* (5.28)*.*

HLE 5.2.10. *For any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n-1)$ satisfying condition* (5.9) *and such that, for any $\lambda > 0$ and $\eta \in]0, \rho[$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n-1)$ and $\omega_n \in \Omega_n \cap B_\eta(\bar{x})$ satisfying condition* (4.4)*, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (4.16), (4.18) *and* (5.29)*.*

The next two conditions are versions of, respectively, Conditions 5.2.9 and 5.2.10 for the case when X is an Asplund space. They employ Fréchet normal cones.

HLEA 5.2.11. *For any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying condition* (5.1) *and such that, for any $\lambda > 0$, $\eta \in]0, \rho[$ and $\tau \in]0, 1[$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ and $x \in B_\eta(\bar{x})$ satisfying condition* (4.29)*, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (4.26)*,* (4.30) *and* (5.28)*.*

HLEA 5.2.12. *For any $\varepsilon > 0$, there exist vectors $a_i \in X$ $(i = 1, \ldots, n-1)$ satisfying condition* (5.9) *and such that, for any $\lambda > 0$, $\eta \in]0, \rho[$ and $\tau \in]0, 1[$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n-1)$ and $\omega_n \in \Omega_n \cap B_\eta(\bar{x})$ satisfying condition* (4.19)*, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (4.16)*,* (4.20) *and* (5.29)*.*

5.2.3 Stationarity

Suppose that the collection $\{\Omega_1, \ldots, \Omega_n\}$ is stationary at $\bar{x}$. Then the conditions below hold true.

HS 5.2.13. *For any $\varepsilon > 0$, there exist a number $\rho \in]0, \varepsilon[$ and vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying condition* (5.11) *and such that, for any $\lambda > 0$ and $\eta \in]0, \rho]$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ and $x \in B_\eta(\bar{x})$ satisfying conditions* (4.11) *and* (5.24)*, where $\gamma := \dfrac{\lambda}{\eta}$. As a consequence, condition* (5.25) *is satisfied.*

HS 5.2.14. *For any $\varepsilon > 0$, there exist a number $\rho \in]0, \varepsilon[$ and vectors $a_i \in X$ $(i = 1, \ldots, n-1)$, satisfying condition* (5.12) *and such that, for any $\lambda > 0$ and $\eta \in]0, \rho]$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n-1)$ and $\omega_n \in \Omega_n \cap B_\eta(\bar{x})$ satisfying conditions* (4.4) *and* (5.26)*, where $\gamma := \dfrac{\lambda}{\eta}$. As a consequence, condition* (5.27) *is satisfied.*

The next two conditions are for the case of a general Banach space. They employ Clarke normal cones.

HS 5.2.15. *For any $\varepsilon > 0$, there exist a number $\rho \in]0, \varepsilon[$ and vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying condition* (5.11) *and such that, for any $\lambda > 0$ and $\eta \in]0, \rho]$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ and $x \in B_\eta(\bar{x})$ satisfying condition* (4.11)*, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (4.26)*,* (4.28) *and* (5.28)*.*

HS 5.2.16. *For any $\varepsilon > 0$, there exist a number $\rho \in]0, \varepsilon[$ and vectors $a_i \in X$ $(i = 1, \ldots, n-1)$ satisfying condition* (5.12) *and such that, for any $\lambda > 0$ and $\eta \in]0, \rho]$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n-1)$ and $\omega_n \in \Omega_n \cap B_\eta(\bar{x})$ satisfying condition* (4.4)*, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (4.16)*,* (4.18) *and* (5.29)*.*

The next two conditions are versions of, respectively, Conditions 5.2.15 and 5.2.16 for the case when X is an Asplund space. They employ Fréchet normal cones.

HSA 5.2.17. *For any $\varepsilon > 0$, there exist a number $\rho \in]0, \varepsilon[$ and vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying condition* (5.11) *and such that, for any $\lambda > 0$, $\eta \in]0, \rho]$ and $\tau \in]0, 1[$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ and $x \in B_\eta(\bar{x})$ satisfying condition* (5.13)*, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (4.26)*,* (4.30) *and* (5.28)*.*

HSA 5.2.18. *For any $\varepsilon > 0$, there exist a number $\rho \in]0, \varepsilon[$ and vectors $a_i \in X$ $(i = 1, \ldots, n-1)$ satisfying condition* (5.12) *and such that, for any $\lambda > 0$, $\eta \in]0, \rho]$ and $\tau \in]0, 1[$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n-1)$ and $\omega_n \in \Omega_n \cap B_\eta(\bar{x})$ satisfying condition* (5.14)*, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (4.16)*,* (4.20) *and* (5.29)*.*

5.2.4 Approximate Stationarity

Suppose that the collection $\{\Omega_1, \ldots, \Omega_n\}$ is approximately stationary at $\bar{x}$. Then the conditions below hold true.

HAS 5.2.19. *For any $\varepsilon > 0$, there exist a number $\rho \in]0, \varepsilon[$, points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ and vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying condition* (5.11) *and such that, for any $\lambda > 0$ and $\eta \in]0, \rho]$, there exist points $\omega'_i \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1, \ldots, n)$ and $x \in \eta\mathbb{B}$ satisfying condition* (5.15) *and*

$$\sup_{\substack{u_i \in \Omega_i\ (i=1,\ldots,n),\, u \in B_\eta(\bar{x}) \\ (u_1,\ldots,u_n,u) \neq (\omega_1,\ldots,\omega_n,x)}} \frac{\max\limits_{1 \le i \le n} \|x + a'_i - \omega'_i\|^q - \max\limits_{1 \le i \le n} \|u + a'_i - u_i\|^q}{\|(u_1 - \omega'_1, \ldots, u_n - \omega'_n, u - x)\|_\gamma} < \frac{\varepsilon^q}{\lambda},$$

where $\gamma := \dfrac{\lambda}{\eta}$, and $a'_i := a_i + \omega_i$ $(i = 1, \ldots, n)$. As a consequence,

$$q \max_{1 \le i \le n} \|x + a'_i - \omega'_i\|^{q-1} \times \limsup_{\substack{\Omega_i \ni u_i \to \omega'_i\ (i=1,\ldots,n),\, u \to 0 \\ (u_1,\ldots,u_n,u) \neq (\omega'_1,\ldots,\omega'_n,x)}} \frac{\max\limits_{1 \le i \le n} \|x + a'_i - \omega'_i\| - \max\limits_{1 \le i \le n} \|u + a'_i - u_i\|}{\|(u_1 - \omega'_1, \ldots, u_n - \omega'_n, u - x)\|_\gamma} < \frac{\varepsilon^q}{\lambda}.$$

HAS 5.2.20. *For any $\varepsilon > 0$, there exist a number $\rho \in]0, \varepsilon[$, points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ $(i = 1, \ldots, n)$ and vectors $a_i \in X$ $(i = 1, \ldots, n-1)$ satisfying* (5.12) *and such that, for any $\lambda > 0$ and $\eta \in]0, \rho]$, there exist points $\omega'_i \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1, \ldots, n-1)$ and $\omega'_n \in \Omega_n \cap B_\eta(\omega_n)$ satisfying* (4.31) *and*

$$\sup_{\substack{u_i \in \Omega_i\ (i=1,\ldots,n),\, u_n \in B_\eta(\omega_n) \\ (u_1,\ldots,u_n) \neq (\omega_1,\ldots,\omega_n)}} \frac{\max\limits_{1 \le i \le n-1} \|\omega'_n + a'_i - \omega'_i\|^q - \max\limits_{1 \le i \le n-1} \|u_n + a'_i - u_i\|^q}{\|(u_1 - \omega'_1, \ldots, u_n - \omega'_n)\|_\gamma} < \frac{\varepsilon^q}{\lambda},$$

where $\gamma := \dfrac{\lambda}{\eta}$ and $a'_i := a_i + \omega_i - \omega_n$ $(i = 1, \ldots, n-1)$. As a consequence,

$$q \max_{1 \le i \le n-1} \|\omega'_n + a'_i - \omega'_i\|^{q-1} \times \limsup_{\substack{\Omega_i \ni u_i \to \omega'_i\ (i=1,\ldots,n) \\ (u_1,\ldots,u_n) \neq (\omega_1,\ldots,\omega_n)}} \frac{\max\limits_{1 \le i \le n-1} \|\omega'_n + a'_i - \omega'_i\| - \max\limits_{1 \le i \le n-1} \|u_n + a'_i - u_i\|}{\|(u_1 - \omega'_1, \ldots, u_n - \omega'_n)\|_\gamma} < \frac{\varepsilon^q}{\lambda}.$$

The next two conditions are for the case of a general Banach space. They employ Clarke normal cones.

HAS 5.2.21. *For any $\varepsilon > 0$, there exist a number $\rho \in]0, \varepsilon[$, points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ and vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying condition* (5.11) *and such that, for any $\lambda > 0$ and $\eta \in]0, \rho]$, there exist points $\omega'_i \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1, \ldots, n)$ and $x \in \eta\mathbb{B}$ satisfying condition* (5.15)*, and vectors $x^*_i \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (4.26), (5.17) *and*

$$q \max_{1 \le i \le n} \|x + a'_i - \omega'_i\|^{q-1} \left(\lambda \sum_{i=1}^n d\left(x^*_i, N_{\Omega_i}(\omega'_i)\right) + \eta \left\| \sum_{i=1}^n x^*_i \right\| \right) < \varepsilon^q, \tag{5.30}$$

where $a'_i := a_i + \omega_i$ $(i = 1, \ldots, n)$.

HAS 5.2.22. *For any $\varepsilon > 0$, there exist a number $\rho \in]0, \varepsilon[$, points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ $(i = 1, \ldots, n)$ and vectors $a_i \in X$ $(i = 1, \ldots, n-1)$ satisfying condition* (5.12) *and such that, for any $\lambda > 0$ and $\eta \in]0, \rho]$, there exist points $\omega'_i \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1, \ldots, n-1)$ and $\omega'_n \in \Omega_n \cap B_\eta(\omega_n)$ satisfying condition* (4.31)*, and vectors $x^*_i \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (4.16), (5.19) *and*

$$q \max_{1 \le i \le n-1} \|\omega'_n + a'_i - \omega'_i\|^{q-1} \left(\lambda \sum_{i=1}^{n-1} d\left(x^*_i, N_{\Omega_i}(\omega'_i)\right) + \eta d\left(x^*_n, N_{\Omega_n}(\omega'_i)\right) \right) < \varepsilon^q, \tag{5.31}$$

where $a'_i := a_i + \omega_i - \omega_n$ $(i = 1, \ldots, n-1)$.

The next two conditions are versions of, respectively, Conditions 5.2.21 and 5.2.22 for the case when X is an Asplund space. They employ Fréchet normal cones.

HASA 5.2.23. *For any $\varepsilon > 0$, there exist a number $\rho \in]0, \varepsilon[$, points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ and vectors $a_i \in X$ $(i = 1, \ldots, n)$ satisfying condition (5.11) and such that, for any $\lambda > 0$, $\eta \in]0, \rho]$ and $\tau \in]0, 1[$, there exist points $\omega'_i \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1, \ldots, n)$ and $x \in \eta\mathbb{B}$ satisfying condition (5.20), and vectors $x^*_i \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions (4.26), (5.21) and (5.30), where $a'_i := a_i + \omega_i$ $(i = 1, \ldots, n)$.*

HASA 5.2.24. *For any $\varepsilon > 0$, there exist a number $\rho \in]0, \varepsilon[$, points $\omega_i \in \Omega_i \cap B_\varepsilon(\bar{x})$ $(i = 1, \ldots, n)$ and vectors $a_i \in X$ $(i = 1, \ldots, n-1)$ satisfying condition (5.12) and such that, for any $\lambda > 0$, $\eta \in]0, \rho]$ and $\tau \in]0, 1[$, there exist points $\omega'_i \in \Omega_i \cap B_\lambda(\omega_i)$ $(i = 1, \ldots, n-1)$ and $\omega'_n \in \Omega_n \cap B_\eta(\omega_n)$ satisfying condition (5.22), and vectors $x^*_i \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions (4.16), (5.23) and (5.31), where $a'_i := a_i + \omega_i - \omega_n$ $(i = 1, \ldots, n-1)$.*

5.2.5 Linear Dual Characterisations

In this subsection, we briefly discuss simplifications in the above necessary characterisations of the extremality/stationarity properties in the linear case, i.e. when $q = 1$. We limit ourselves to the dual necessary characterisations, where the most important simplifications appear.

1. In each of the dual necessary characterisations of extremality/stationarity properties, the Hölder parameter q is only involved in a single condition: see the key normal cone conditions (5.28)–(5.31). In the linear case, all these conditions get much simpler as the terms containing expressions to the power $q-1$ in the left-hand sides of the inequalities disappear. As a result, four different expressions reduce to just two forms. Specifically, inequalities (5.28) and (5.30) take the form

$$\lambda \sum_{i=1}^{n} d\left(x^*_i, N_{\Omega_i}(\omega_i)\right) + \eta \left\| \sum_{i=1}^{n} x^*_i \right\| < \varepsilon, \tag{5.32}$$

while inequalities (5.29) and (5.31) take the form

$$\lambda \sum_{i=1}^{n-1} d\left(x^*_i, N_{\Omega_i}(\omega_i)\right) + \eta d\left(x^*_n, N_{\Omega_n}(\omega_n)\right) < \varepsilon, \tag{5.33}$$

with N in both cases standing for either the Clarke or the Fréchet normal cone. Thus, inequalities (5.28)–(5.31) become independent on the vectors a_i's.

The other simplifications in the necessary characterisations of extremality/stationarity properties discussed below come at the expense of weakening the conditions.

2. The vectors a_i's (coming from the original definitions of the respective properties) remain in two conditions in each of the dual necessary characterisations: conditions (4.4), (4.11), (4.19), (4.29), (4.31), (5.13), (5.14) or (5.15) on the choice of points $\omega_i \in \Omega_i$ (or $\omega'_i \in \Omega_i$) $(i = 1, \ldots, n)$, and conditions (4.18), (4.20), (4.28), (4.30), (5.17), (5.19), (5.21) or (5.23) coupling the primal and dual vectors involved in the characterisations. Both groups of conditions are important for detecting the respective properties. At the same time, finding the vectors a_i's is not easy in general, which makes checking these conditions difficult. Removing the conditions listed above (together with the number τ involved in the characterisations employing Fréchet normal cones) from the corresponding necessary characterisations makes them simpler and more practical, though weaker in general. For instance, the dual necessary characterisations of extremality 5.2.3 and 5.2.4, and their Asplund space versions 5.2.5 and 5.2.6 reduce to the following two conditions, where N stands for either the Clarke normal cone ($N = N^C$) in the case of a general Banach space, or the Fréchet normal cone ($N = N^F$) if the space is Asplund.

EL 5.2.25. *For any $\varepsilon > 0$, $\lambda > 0$ and $\eta > 0$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$ and $x \in B_\eta(\bar{x})$, and vectors $x^*_i \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions (4.26) and (5.32).*

EL 5.2.26. *For any $\varepsilon > 0$, $\lambda > 0$ and $\eta > 0$, there exist points $\omega_i \in \Omega_i \cap B_\lambda(\bar{x})$ $(i = 1, \ldots, n-1)$ and $\omega_n \in \Omega_n \cap B_\eta(\bar{x})$, and vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying conditions* (4.16) *and* (5.33).

The corresponding simplified versions of the necessary characterisations of local extremality, stationarity and approximate stationarity contain the same dual conditions (5.32) and (5.33).

3. Unlike condition (5.33), which basically requires the dual vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ to be close to the respective normal cones, condition (5.32) combines two different types of constraints on theses vectors: they must be close to the respective normal cones and their sum must be small. Within Condition 5.2.26, the second constraint is taken care of by condition (4.16), which requires even more: the sum of the vectors must **equal** 0.

Fortunately, the two types of constraints combined in condition (5.32) can be easily separated and, using elementary arguments hidden in multiple proofs of 'generalised separation' statements and formulated explicitly in [29, Lemma 1], the collection of dual vectors $x_i^* \in X^*$ $(i = 1, \ldots, n)$ satisfying (5.32) can be replaced in Condition 5.2.25 and its analogues by another collection of vectors **belonging** to the respective normal cones: $x_i^{*\prime} \in N_{\Omega_i}(\omega_i)$ $(i = 1, \ldots, n)$, satisfying (4.26) and with $\left\| \sum_{i=1}^{n} x_i^{*\prime} \right\|$ arbitrarily small.

4. Finally, it is not difficult to observe that the parameters λ and η in Conditions 5.2.25 and 5.2.26 are redundant. They both can be replaced in conditions $\omega_i \in B_\lambda(\bar{x})$ $(i = 1, \ldots, n)$, $x \in B_\eta(\bar{x})$ in 5.2.25 and $\omega_i \in B_\lambda(\bar{x})$ $(i = 1, \ldots, n-1)$, $\omega_n \in B_\eta(\bar{x})$ in 5.2.26 by ε. Moreover, taking into account the observations in item 3, condition (5.32) can be replaced by $x_i^* \in N_{\Omega_i}(\omega_i)$ $(i = 1, \ldots, n)$ and $\left\| \sum_{i=1}^{n} x_i^* \right\| < \varepsilon$, while condition (5.33) can be replaced by $\sum_{i=1}^{n} d\left(x_i^*, N_{\Omega_i}(\omega_i)\right) < \varepsilon$. Thus, Conditions 5.2.25 and 5.2.26 reduce, respectively, to the well known conditions (ii) and (iii) in Theorem 2.7.3.

Similar observations apply to the necessary characterisations of local extremality, stationarity and approximate stationarity. The auxiliary points $\omega_i \in \Omega_i$ $(i = 1, \ldots, n)$ in the characterisations of approximate stationarity can be disregarded. As a result, the respective analogues of Conditions 5.2.25 and 5.2.26 reduce to the same conventional conditions (ii) and (iii) in Theorem 2.7.3.

Thus, the equivalent conditions (ii) and (iii) in Theorem 2.7.3 represent universal dual necessary characterisations of all four properties: extremality, local extremality, stationarity and approximate stationarity in the linear setting. In fact, as it follows from Theorem 2.7.3, these conditions are indeed good for characterising approximate stationarity. For the other three properties, which are stronger, in some cases these characterisations can be too weak. In such cases, one has an option of employing the original (not simplified) characterisations, involving the vectors a_i's.

Chapter 6

Some New Characterisations of Intrinsic Transversality in Hilbert Spaces

Motivated mainly by the above open questions, this part of the thesis is devoted to investigating further characterisations of intrinsic transversality in connection with other transversality-type properties. Apart from the appeal to address the aforementioned questions in Chapter 1, this work was also motivated by the potential for meaningful applications of these properties, for example, in establishing convergence criteria for more involved projection algorithms (rather than alternating projections) and in formulating calculus rules for *relative limiting normal cones* (see Definition 6.3.1).

New results in terms of elements of normal cones are presented in Section 6.1 with the key quantitative estimate formulated in Lemma 6.1.11. Theorem 6.1.13 establishes the equivalence of intrinsic transversality, weak intrinsic transversality [96] and the sufficient condition for subtransversality formulated in [99, Theorem 2]. This result significantly clarifies the picture of transversality-type properties and unifies several of them. As by-products, we address several important questions concerning these properties in the Hilbert space setting, see Questions 1–3. In Section 6.2, for the first time, intrinsic transversality is characterised by an equivalent property which does not involve normal vectors, see Theorem 6.2.9. This result, which was motivated by a question (see Question 4) raised by Ioffe [74, page 358], opens a new perspective on intrinsic transversality. Intrinsic transversality in Euclidean spaces is studied in Section 6.3. Lemma 6.3.5 establishes a geometric counterpart of the analytic condition under which the complete quantitative results of Theorem 6.1.13(ii) are obtained. Theorem 6.2.9 gives a new characterisation of intrinsic transversality, which refines the corresponding result of [96, Corollary 3]. Theorem 6.3.7 yields further insight on the quantitative results established in Section 6.1 when specialised to the Euclidean space setting. As by-products, we address a couple of interesting questions concerning intrinsic transversality in Euclidean spaces raised by Kruger [96, page 140], see Questions 5 and 6.

6.1 Subtransversality, Transversality and Intrinsic Transversality

The following definition simplifies Definition 2.3.9(ii), (iii) for a pair of sets.

Definition 6.1.1 (subtransversality & transversality)**.** Let $\{A, B\}$ be a pair of sets and $\bar{x} \in A \cap B$.

(i) $\{A, B\}$ is *subtransversal* at $\bar{x}$ if there exist numbers $\alpha \in]0,1[$ and $\delta > 0$ such that

$$\alpha \text{dist}\,(x, A \cap B) \leq \max\{\text{dist}\,(x, A), \text{dist}\,(x, B)\} \quad \forall x \in \mathbb{B}_\delta(\bar{x}). \tag{6.1}$$

(ii) $\{A, B\}$ is *transversal* at $\bar{x}$ if there exist numbers $\alpha \in]0,1[$ and $\delta > 0$ such that

$$\alpha \text{dist}\,(x, (A - x_1) \cap (B - x_2)) \leq \max\{\text{dist}\,(x, A - x_1), \text{dist}\,(x, B - x_2)\} \qquad \forall x \in \mathbb{B}_\delta(\bar{x}),\ \forall x_1, x_2 \in \delta\mathbb{B}. \tag{6.2}$$

The exact upper bound of all $\alpha \in]0,1[$ such that condition (6.1) or condition (6.2) is satisfied for some $\delta > 0$ is denoted by $\text{str}[A, B](\bar{x})$ or $\text{tr}[A, B](\bar{x})$, respectively, with the convention that the supremum over the empty set equals 0.

Remark 6.1.2. (i) (*subtransversality*) The subtransversality property can be traced back to at least the early 80's thanks to Dolecki under the name *decisive separation* [49, 50], where it was known as a sufficient (and also necessary in the convex setting) condition for the tangent cone of the intersection of a pair of sets at a reference point being equal to the intersection of the two tangent cones of the sets at that point [49, Propositions 5.3 and 5.4]. In the surveys [68, 70], Ioffe used the property (without a name) as a qualification condition for establishing calculus rules for normal cones and subdifferentials. Subtransversality was studied by Bauschke and Borwein [12] under the name *linear regularity* as a sufficient condition for linear convergence of the alternating projection algorithm for solving convex feasibility problems in Hilbert spaces, and became widely known thanks to this important application. Their results were extended to the cyclic projection algorithm for solving feasibility problems involving a finite number of convex sets [13]. The term *linear regularity* was widely adapted in the community of variational analysis and optimisation for several decades, for example, Bakan et al. [8], Bauschke et al. [14, 15], Li et al. [111], Ng and Zang [124], Zheng and Ng [171], Kruger and his collaborators [90–92, 104–106, 158]. Ngai and Théra [125] referred to this property as *metric inequality* and used it to establish calculus rules for the limiting Fréchet subdifferential. Penot [129] referred to the property as *linear coherence* and applied it to calculus rules for the *viscosity Fréchet* and *viscosity Hadamard* subdifferentials. The name (sub)transversality was coined by Ioffe in the 2016 survey [72, Definition 6.14]. In his 2017 book [74, page 301] he explained that "Regularity is a property of a single object while transversality relates to the interaction of two or more independent objects". In spite of the relatively long history with many important features of subtransversality, for example, those in connection with *metric subregularity*, *error bounds*, *weak sharp minima*, *growth conditions* and *conditions involving primal and dual slopes*, useful applications of the property keep being discovered. For example, Luke et al. [114, Theorem 8] have very recently proved that subtransversality is not only sufficient but also necessary for linear convergence of convex alternating projections. This complements the aforementioned result by Bauschke and Borwein [12] obtained 25 years earlier. Luke et al. [114, Section 4] also reveal that the property has been imposed either explicitly or implicitly in all existing linear convergence criteria for nonconvex alternating projections, and hence conjecture that subtransversality is a necessary condition for linear convergence of the algorithm also in the nonconvex setting.

(ii) (*transversality*) The origin of the concept of *transversality* can be traced back to at least the 19th century (cf. [63, 66]) in differential geometry which deals of course with smooth manifolds, where transversality of a pair of smooth manifolds $\{A, B\}$ at a common point $\bar{x}$ can also be characterised by condition [1]

$$\overline{N}_A(\bar{x}) \cap \left(-\overline{N}_B(\bar{x})\right) = \{0\}. \tag{6.3}$$

The property is known as a sufficient condition for the intersection $A \cap B$ to be also a smooth manifold around $\bar{x}$. To the best of our awareness, transversality of pairs (collections) of general sets in normed linear spaces was first investigated by Kruger in a systematic picture of mutual arrangement properties of sets. The property has been known under quite a number of other names including *regularity*, *strong regularity*, *property* $(UR)_S$, *uniform regularity*, *strong metric inequality* [90–92, 104]

[1] In this special setting, the normal cones appearing in (6.3) are the *normal spaces* (i.e., orthogonal complements to the *tangent spaces*) to the manifolds at $\bar{x}$. The minus sign in (6.3) can be omitted.

and *linear regular intersection* [109]. Plenty of primal and dual space characterisations of transversality as well as its close connections to important concepts in optimisation and variational analysis such as *metric regularity*, *(extended) extremal principles*, *open mapping theorems* and other types of mutual arrangement properties of collections of sets have been established and extended to more general *nonlinear settings* in a series of papers by Kruger and his collaborators [81,90–95,105,106]. Apart from classical applications of the property, for example, as constraint qualification conditions for establishing calculus rules for the limiting normal cones [120, page 265] and coderivatives (in connection with *metric regularity*, the counterpart of transversality in terms of set-valued mappings) [51,143], important applications have also emerged in the field of numerical analysis. Lewis et al. [109,110] applied the property to establish the first linear convergence criteria for nonconvex alternating and averaged projections. Transversality was also used to prove linear convergence of the Douglas-Rachford algorithm [65,135] and its relaxations [159]. A practical application of these results is the *phase retrieval problem* where transversality is sufficient for linear convergence of alternating projections, the Douglas-Rachford algorithm and actually any convex combinations of the two algorithms [160].

We refer the reader to the recent surveys by Kruger et al. [99, 100] for a more comprehensive discussion about the two properties.

A number of characterisations of transversality in terms of normal vectors, especially in the Euclidean space setting, have been established [90–92,100,104,106,109] and applied, for example, [109,120,135,159]. The situation is very much different for subtransversality. For collections of closed and convex sets, the following characterisation of subtransversality is due to Kruger.

Proposition 6.1.3 (normal-vector-based characterisation of subtransversality with convexity)**.** *[96, Theorem 3]*[2] *A pair of closed and convex sets $\{A, B\}$ is subtransversal at a point $\bar{x} \in A \cap B$ if and only if there exist numbers $\alpha \in]0,1[$ and $\delta > 0$ such that $\|v_1 + v_2\| > \alpha$ for all $a \in (A \setminus B) \cap \mathbb{B}_\delta(\bar{x})$, $b \in (B \setminus A) \cap \mathbb{B}_\delta(\bar{x})$, $x \in \mathbb{B}_\delta(\bar{x})$ with $\|x - a\| = \|x - b\|$ and $v_1, v_2 \in X$ satisfying*

$$\operatorname{dist}(v_1, N_A(a)) < \delta,\ \operatorname{dist}(v_2, N_B(b)) < \delta,$$
$$\|v_1\| + \|v_2\| = 1,\ \langle v_1, x - a\rangle = \|v_1\|\,\|x - a\|,\ \langle v_2, x - b\rangle = \|v_2\|\,\|x - b\|.$$

Let $\operatorname{itr}_c[A, B](\bar{x})$ denote the exact upper bound of all $\alpha \in]0,1[$ such that the conditions in Proposition 6.1.3 (equivalent to subtransversality in the convex setting) are satisfied for some $\delta > 0$, with the convention that the supremum over the empty set equals 0. This quantity is well defined regardless of the convexity of the sets, and the strict inequality $\operatorname{itr}_c[A, B](\bar{x}) > 0$ characterises a certain transversality-type property in not necessarily convex setting. The constant $\operatorname{itr}_c[A, B](\bar{x})$ is going to play a central role in the analysis in this part of thesis.

In the nonconvex setting, the first sufficient condition for subtransversality in terms of normal vectors was formulated in [105, Theorem 4.1] following the routine of deducing metric subregularity characterisations for set-valued mappings in [94]. The result was then refined successively in [100, Theorem 4(ii)], [99, Theorem 2] and finally in [96] in the following form.

Proposition 6.1.4 (sufficient condition for subtransversality)**.** *[96, combination of Definition 2 and Corollary 2]*[3] *A pair of closed sets $\{A, B\}$ is subtransversal at a point $\bar{x} \in A \cap B$ if there exist numbers $\alpha \in]0,1[$ and $\delta > 0$ such that, for all $a \in (A \setminus B) \cap \mathbb{B}_\delta(\bar{x})$, $b \in (B \setminus A) \cap \mathbb{B}_\delta(\bar{x})$ and $x \in \mathbb{B}_\delta(\bar{x})$ with $\|x - a\| = \|x - b\|$, one has $\|v_1 + v_2\| > \alpha$ for some $\varepsilon > 0$ and all $a' \in A \cap \mathbb{B}_\varepsilon(a)$, $b' \in B \cap \mathbb{B}_\varepsilon(b)$, $x_1' \in \mathbb{B}_\varepsilon(a)$, $x_2' \in \mathbb{B}_\varepsilon(b)$ with $\|x - x_1'\| = \|x - x_2'\|$, and $v_1, v_2 \in X$ satisfying*

$$\operatorname{dist}(v_1, N_A(a')) < \delta,\ \operatorname{dist}(v_2, N_B(b')) < \delta,$$
$$\|v_1\| + \|v_2\| = 1,\ \langle v_1, x - x_1'\rangle = \|v_1\|\|x - x_1'\|,\ \langle v_2, x - x_2'\rangle = \|v_2\|\|x - x_2'\|.$$

Let $\operatorname{itr}_w[A, B](\bar{x})$ denote the exact upper bound of all $\alpha \in]0,1[$ such that the above sufficient condition

[2] The result is valid in Banach spaces.

[3] The result is valid in Asplund spaces.

for subtransversality is satisfied for some $\delta > 0$, with the convention that the supremum over the empty set equals 0.

To this end, the following question about the above result is of importance. Our subsequent analysis will give the negative answer to it (see Remark 6.1.17).

Question 1. *Is the sufficient condition formulated in Proposition 6.1.4, i.e.* $\mathrm{itr}_w[A,B](\bar{x}) > 0$*, also necessary for subtransversality?*

We next recall the central concept of this chapter, i.e., the *intrinsic transversality* property of pairs of sets. Compared to its better known siblings: transversality and subtransversality, intrinsic transversality came to life much later.

Definition 6.1.5 (intrinsic transversality in Euclidean spaces)**.** [52, Definition 3.1] A pair of closed sets $\{A,B\}$ in a Euclidean space is *intrinsically transversal* at a point $\bar{x} \in A \cap B$ if there exists an angle $\alpha > 0$ together with a number $\delta > 0$ such that any two points $a \in (A \setminus B) \cap \mathbb{B}_\delta(\bar{x})$ and $b \in (B \setminus A) \cap \mathbb{B}_\delta(\bar{x})$ cannot have difference $a - b$ simultaneously making an angle strictly less than α with the two proximal normal cones $N_B^p(b)$ and $-N_A^p(a)$.

The above property was originally introduced in 2015 by Drusvyatskiy et al. [52] as a sufficient condition for local linear convergence of the alternating projection algorithm for solving nonconvex feasibility problems in Euclidean spaces. As demonstrated by Ioffe [75], Kruger et al. [96, 99] and will also be confirmed in this chapter, intrinsic transversality turns out to be an important qualification property in the framework of variational analysis. Kruger [96] recently extended and investigated intrinsic transversality in more general underlying spaces[4].

Definition 6.1.6 (intrinsic transversality)**.** [99, Definition 4(ii)] & [96, Definition 2(ii)][5] A pair of closed sets $\{A,B\}$ is *intrinsically transversal* at a point $\bar{x} \in A \cap B$ if there exist numbers $\alpha \in]0,1[$ and $\delta > 0$ such that $\|v_1 + v_2\| > \alpha$ for all $a \in (A \setminus B) \cap \mathbb{B}_\delta(\bar{x})$, $b \in (B \setminus A) \cap \mathbb{B}_\delta(\bar{x})$, $x \in \mathbb{B}_\delta(\bar{x})$ with $x \neq a$, $x \neq b$, $1 - \delta < \dfrac{\|x-a\|}{\|x-b\|} < 1 + \delta$, and $v_1 \in N_A(a) \setminus \{0\}$, $v_2 \in N_B(b) \setminus \{0\}$ satisfying

$$\|v_1\| + \|v_2\| = 1, \ \frac{\langle v_1, x-a \rangle}{\|v_1\|\|x-a\|} > 1 - \delta, \ \frac{\langle v_2, x-b \rangle}{\|v_2\|\|x-b\|} > 1 - \delta.$$

The exact upper bound of all $\alpha \in]0,1[$ that together with some $\delta > 0$ satisfies the above description of intrinsic transversality is denoted by $\mathrm{itr}[A,B](\bar{x})$, with the convention that the supremum over the empty set equals 0.

Remark 6.1.7. In Definition 6.1.6 it can be assumed without loss of generality that $\delta \in]0,1[$. In this case, the three conditions $a \in A \setminus B$, $b \in B \setminus A$ and $1 - \delta < \dfrac{\|x-a\|}{\|x-b\|} < 1 + \delta$ imply conditions $x \neq a$ and $x \neq b$. In similar contexts in the sequel, the latter two conditions will be omitted for the sake of brevity, for example, in the proofs of Lemmas 6.1.11 & 6.2.8, representation (6.70) and Definition 6.3.1(i).

Making use of the quantities $\mathrm{str}[A,B](\bar{x})$, $\mathrm{tr}[A,B](\bar{x})$, $\mathrm{itr}_c[A,B](\bar{x})$, $\mathrm{itr}_w[A,B](\bar{x})$ and $\mathrm{itr}[A,B](\bar{x})$, we can concisely summarize the facts recalled so far in this section. The definitions of subtransversality, transversality and intrinsic transversality and Propositions 6.1.3 & 6.1.4 respectively admit more concise descriptions.

Proposition 6.1.8 (summary)**.** *Let $\{A,B\}$ be a pair of closed sets and $\bar{x} \in A \cap B$.*

(i) *$\{A,B\}$ is subtransversal at $\bar{x}$ if and only if* $\mathrm{str}[A,B](\bar{x}) > 0$.

[4] It is worth noting that the extension from Definition 6.1.5 to Definition 6.1.6 of intrinsic transversality is not trivial, and the coincidence of the two definitions in the Euclidean space setting was shown in [96, Proposition 8(iii)].

[5] The property was defined and investigated in general normed linear spaces.

(ii) $\{A,B\}$ *is transversal at* $\bar{x}$ *if and only if* $\mathrm{tr}[A,B](\bar{x})>0$.

(iii) $\{A,B\}$ *is intrinsically transversal at* $\bar{x}$ *if and only if* $\mathrm{itr}[A,B](\bar{x})>0$.

(iv) *If the sets* A,B *are convex, then* $\{A,B\}$ *is subtransversal at* $\bar{x}$ *if and only if* $\mathrm{itr}_c[A,B](\bar{x})>0$.

(v) $\{A,B\}$ *is subtransversal at* $\bar{x}$ *if* $\mathrm{itr}_w[A,B](\bar{x})>0$.

In this section, we are particularly interested in the result established by Kruger [96, Theorem 4] that intrinsic transversality implies the sufficient condition of subtransversality stated in Proposition 6.1.4, which in turn implies the one stated in Proposition 6.1.3. These implications are captured by Proposition 6.1.9(i) via the relationships between the corresponding quantities. For completeness, more comprehensive relationships between the transversality-type properties are also presented here.

Proposition 6.1.9 (relationships between quantitative constants)**.** *[96, Proposition 1] Let* $\{A,B\}$ *be a pair of closed sets and* $\bar{x}\in A\cap B$.

(i) $0\le \mathrm{tr}[A,B](\bar{x})\le \mathrm{itr}[A,B](\bar{x})\le \mathrm{itr}_w[A,B](\bar{x})\le \mathrm{itr}_c[A,B](\bar{x})\le 1.$[6]

(ii) $\mathrm{itr}_w[A,B](\bar{x})\le \mathrm{str}[A,B](\bar{x}).$[7]

(iii) *If* A *and* B *are convex, then* $\mathrm{itr}_c[A,B](\bar{x})=\mathrm{str}[A,B](\bar{x})$.

(iv) *If* $\dim X<\infty$ *and* A,B *are convex, then* $\mathrm{itr}_w[A,B](\bar{x})=\mathrm{itr}_c[A,B](\bar{x})=\mathrm{str}[A,B](\bar{x})$.

Remark 6.1.10 (notation & terminology). In view of Proposition 6.1.9(i)&(ii), the strict inequality $\mathrm{itr}_w[A,B](\bar{x})>0$ corresponds to a property, which is weaker than intrinsic transversality and stronger than subtransversality. That property is called *weak intrinsic transversality* in [96, 99]. This in particular explains why the letter "w" is used in the notation $\mathrm{itr}_w[A,B](\bar{x})$. Similarly, the strict inequality $\mathrm{itr}_c[A,B](\bar{x})>0$ corresponds to a weaker property than weak intrinsic transversality. Such a property has not been named yet, but it has played an important role in the analysis of transversality-type properties mainly in the convex setting [96]. This particularly explains why the letter "c" is used in the notation $\mathrm{itr}_c[A,B](\bar{x})$.

Proposition 6.1.9(iv) in particular claims the equivalence between weak intrinsic transversality and intrinsic transversality in the convex and finite dimensional setting. The following question about their relationship in more general settings is of interest. Subsequent analysis will give the answer to this question in the Hilbert space setting (see Remark 6.1.16).

Question 2. *[96, question 3, page 140] What is the relationship between weak intrinsic transversality and intrinsic transversality in the general nonconvex setting?*

The next result establishes the main quantitative estimate of this section. Though the statement and its proof are rather technical, its meaningful consequences will follow shortly.

Lemma 6.1.11 (quantitative estimate)**.** *Let* $\{A,B\}$ *be a pair of closed sets and* $\bar{x}\in A\cap B$. *It holds*

$$\min\left\{\mathrm{itr}_c[A,B](\bar{x}),1/\sqrt{2}\right\}\le \mathrm{itr}[A,B](\bar{x}). \tag{6.4}$$

Proof. To proceed with the proof, let us suppose that $\mathrm{itr}_c[A,B](\bar{x})>0$ since there is nothing to prove in the case $\mathrm{itr}_c[A,B](\bar{x})=0$. Let us fix an arbitrary number

$$\beta\in\left]0,\min\left\{\mathrm{itr}_c[A,B](\bar{x}),1/\sqrt{2}\right\}\right[\tag{6.5}$$

[6] The statement is valid in Banach spaces.

[7] The statement is valid in Asplund spaces.

and prove that $\operatorname{itr}[A,B](\bar{x}) \geq \beta$. By the definition of $\operatorname{itr}_c[A,B](\bar{x})$, there exist numbers

$$\alpha \in \left]\beta, \min\left\{\operatorname{itr}_c[A,B](\bar{x}), 1/\sqrt{2}\right\}\right[$$

and $\delta > 0$ such that, for all $a \in (A \setminus B) \cap \mathbb{B}_\delta(\bar{x})$, $b \in (B \setminus A) \cap \mathbb{B}_\delta(\bar{x})$ and $x \in \mathbb{B}_\delta(\bar{x})$ with $\|x-a\| = \|x-b\|$, one has

$$\|v_1 + v_2\| > \alpha \tag{6.6}$$

for all $v_1, v_2 \in X$ satisfying

$$\operatorname{dist}(v_1, N_A(a)) < \delta,\ \operatorname{dist}(v_2, N_B(b)) < \delta, \tag{6.7}$$

$$\|v_1\| + \|v_2\| = 1,\ \langle v_1, x-a\rangle = \|v_1\|\|x-a\|,\ \langle v_2, x-b\rangle = \|v_2\|\|x-b\|. \tag{6.8}$$

Choose a number $\delta' \in]0, \delta/3[$ and satisfying

$$2\left(\sqrt{\delta'} + \delta'\right) < 1/2 - \beta^2, \tag{6.9}$$

$$\sqrt{2\delta'} + 2\sqrt{\frac{2\delta' - \delta'^2}{4 - 6\delta' + 3\delta'^2}} < \min\{\delta, \alpha - \beta\}. \tag{6.10}$$

Such a number δ' exists since $1/2 - \beta^2 > 0$, $\min\{\delta, \alpha-\beta\} > 0$ and

$$\lim_{t \downarrow 0} 2\left(\sqrt{t} + t\right) = 0,\ \lim_{t \downarrow 0}\left(\sqrt{2t} + 2\sqrt{\frac{2t - t^2}{4 - 6t + 3t^2}}\right) = 0.$$

We are going to prove $\operatorname{itr}[A,B](\bar{x}) \geq \beta$ with the technical constant $\delta' > 0$. To begin, let us take any $a \in (A \setminus B) \cap \mathbb{B}_{\delta'}(\bar{x})$, $b \in (B \setminus A) \cap \mathbb{B}_{\delta'}(\bar{x})$ and $x \in \mathbb{B}_{\delta'}(\bar{x})$ with $x \neq a$, $x \neq b$,

$$1 - \delta' < \frac{\|x-a\|}{\|x-b\|} < 1 + \delta', \tag{6.11}$$

and $v_1 \in N_A(a) \setminus \{0\}$, $v_2 \in N_B(b) \setminus \{0\}$ satisfying

$$\|v_1\| + \|v_2\| = 1,\ \frac{\langle v_1, x-a\rangle}{\|v_1\|\|x-a\|} > 1 - \delta',\ \frac{\langle v_2, x-b\rangle}{\|v_2\|\|x-b\|} > 1 - \delta'. \tag{6.12}$$

All we need is to show that

$$\|v_1 + v_2\| > \beta. \tag{6.13}$$

We first observe from (6.12) that

$$\begin{aligned} \left\|\frac{v_1}{\|v_1\|} - \frac{x-a}{\|x-a\|}\right\|^2 &= 2 - 2\frac{\langle v_1, x-a\rangle}{\|v_1\|\|x-a\|} < 2 - 2(1-\delta') = 2\delta', \\ \left\|\frac{v_2}{\|v_2\|} - \frac{x-b}{\|x-b\|}\right\|^2 &= 2 - 2\frac{\langle v_2, x-b\rangle}{\|v_2\|\|x-b\|} < 2 - 2(1-\delta') = 2\delta'. \end{aligned} \tag{6.14}$$

We take care of two possibilities concerning the value of $\langle x-a, x-b\rangle$ as follows.

Case 1. $\langle x-a, x-b\rangle > 0$. Then

$$\left\|\frac{x-a}{\|x-a\|} - \frac{x-b}{\|x-b\|}\right\|^2 = 2 - 2\frac{\langle x-a, x-b\rangle}{\|x-a\|\|x-b\|} < 2.$$

Equivalently,

$$\left\|\frac{x-a}{\|x-a\|} - \frac{x-b}{\|x-b\|}\right\| < \sqrt{2}. \tag{6.15}$$

By the triangle inequality and estimates (6.15), (6.14), we get that

$$\begin{aligned} \left\|\frac{v_1}{\|v_1\|} - \frac{v_2}{\|v_2\|}\right\| &\leq \left\|\frac{x-a}{\|x-a\|} - \frac{x-b}{\|x-b\|}\right\| + \left\|\frac{v_1}{\|v_1\|} - \frac{x-a}{\|x-a\|}\right\| + \left\|\frac{v_2}{\|v_2\|} - \frac{x-b}{\|x-b\|}\right\| \\ &< \sqrt{2} + 2\sqrt{2\delta'} = \sqrt{2}\left(1 + 2\sqrt{\delta'}\right). \end{aligned}$$

This implies that

$$\left\|\frac{v_1}{\|v_1\|}-\frac{v_2}{\|v_2\|}\right\|^2 = 2-2\frac{\langle v_1,v_2\rangle}{\|v_1\|\,\|v_2\|} < 2\left(1+2\sqrt{\delta'}\right)^2 \Leftrightarrow \langle v_1,v_2\rangle > -4\left(\sqrt{\delta'}+\delta'\right)\|v_1\|\,\|v_2\|\,. \tag{6.16}$$

Using $\|v_1\|+\|v_2\|=1$ which implies $\|v_1\|\,\|v_2\|\leq 1/4$ and (6.16), respectively, we obtain that

$$\begin{aligned}\|v_1+v_2\|^2 &= \|v_1\|^2+\|v_2\|^2+2\langle v_1,v_2\rangle = 1-2\|v_1\|\,\|v_2\|+2\langle v_1,v_2\rangle \\ &> 1-2\|v_1\|\,\|v_2\|-8\left(\sqrt{\delta'}+\delta'\right)\|v_1\|\,\|v_2\| \geq \frac{1}{2}-2\left(\sqrt{\delta'}+\delta'\right).\end{aligned}$$

This combining with (6.9) yields that

$$\|v_1+v_2\| > \sqrt{\frac{1}{2}-2\left(\sqrt{\delta'}+\delta'\right)} > \beta.$$

Case 2.

$$\langle x-a,x-b\rangle \leq 0. \tag{6.17}$$

Let us define $m=\dfrac{a+b}{2}$ and

$$x' = x-\frac{\langle b-a,x-m\rangle}{\|b-a\|^2}(b-a). \tag{6.18}$$

We first check that

$$\|x'-a\| = \|x'-b\|. \tag{6.19}$$

Indeed,

$$\begin{aligned}\|x'-a\|^2-\|x'-b\|^2 &= \|x-a\|^2-\|x-b\|^2-2\frac{\langle b-a,x-m\rangle}{\|b-a\|^2}\langle x-a,b-a\rangle \\ &\quad+2\frac{\langle b-a,x-m\rangle}{\|b-a\|^2}\langle x-b,b-a\rangle \\ &= \|x-a\|^2-\|x-b\|^2-2\frac{\langle b-a,x-m\rangle}{\|b-a\|^2}\langle b-a,b-a\rangle \\ &= \|x-a\|^2-\|x-b\|^2-2\langle b-a,x-m\rangle \\ &= \|x-a\|^2-\|x-b\|^2-\langle (x-a)-(x-b),(x-a)+(x-b)\rangle = 0.\end{aligned}$$

We next check that

$$\langle x-x',x'-m\rangle = 0. \tag{6.20}$$

Indeed, by (6.18), it holds that

$$\langle x-x',x'-m\rangle = \frac{\langle b-a,x-m\rangle}{\|b-a\|^2}\langle b-a,x'-m\rangle,$$

from which (6.20) follows since

$$\begin{aligned}\langle b-a,x'-m\rangle &= \left\langle b-a,x-m-\frac{\langle b-a,x-m\rangle}{\|b-a\|^2}(b-a)\right\rangle \\ &= \langle b-a,x-m\rangle-\frac{\langle b-a,x-m\rangle}{\|b-a\|^2}\langle b-a,b-a\rangle = 0.\end{aligned}$$

Let us define also

$$v_1' = \frac{\|v_1\|}{\|x'-a\|}(x'-a),\ v_2' = \frac{\|v_2\|}{\|x'-b\|}(x'-b). \tag{6.21}$$

It is clear that

$$\begin{aligned}&\|v_1'\|=\|v_1\|,\ \|v_2'\|=\|v_2\|,\ \|v_1'\|+\|v_2'\|=1,\\ &\langle v_1',x'-a\rangle = \|v_1\|\|x'-a\| = \|v_1'\|\|x'-a\|,\\ &\langle v_2',x'-b\rangle = \|v_2\|\|x'-b\| = \|v_2'\|\|x'-b\|.\end{aligned} \tag{6.22}$$

We next check that

$$\operatorname{dist}(v_1', N_A(a)) < \delta,\ \operatorname{dist}(v_2', N_B(b)) < \delta. \tag{6.23}$$

Let us prove $\operatorname{dist}(v_1', N_A(a)) < \delta$. Indeed, since $v_1 \in N_A(a)$, it holds by (6.21) that

$$\begin{aligned}\operatorname{dist}(v_1', N_A(a)) \leq \|v_1' - v_1\| &= \left\| \frac{\|v_1\|}{\|x'-a\|}(x'-a) - v_1 \right\| \\ &= \|v_1\| \left\| \frac{x'-a}{\|x'-a\|} - \frac{v_1}{\|v_1\|} \right\| \\ &\leq \|v_1\| \left(\left\| \frac{x-a}{\|x-a\|} - \frac{v_1}{\|v_1\|} \right\| + \left\| \frac{x'-a}{\|x'-a\|} - \frac{x-a}{\|x-a\|} \right\| \right).\end{aligned} \tag{6.24}$$

An upper bound of $\left\| \frac{x-a}{\|x-a\|} - \frac{v_1}{\|v_1\|} \right\|$ has been given by (6.14):

$$\left\| \frac{x-a}{\|x-a\|} - \frac{v_1}{\|v_1\|} \right\| < \sqrt{2\delta'}. \tag{6.25}$$

We now establish an upper bound of $\left\| \frac{x'-a}{\|x'-a\|} - \frac{x-a}{\|x-a\|} \right\|$ via three steps as follows.

Step 1. We show that

$$\|x - x'\|^2 \leq \tfrac{2\delta' - \delta'^2}{4(1-\delta')^2} \min\left\{\|x-a\|^2, \|x-b\|^2\right\}. \tag{6.26}$$

If $\|x-a\| \geq \|x-b\|$, then

$$\begin{aligned}&\|x-a\|^2 - \|x-b\|^2 \geq 0 \\ \Leftrightarrow\ &\|x-m\|^2 + \|m-a\|^2 + 2\langle x-m, m-a\rangle - \|x-m\|^2 - \|m-b\|^2 - 2\langle x-m, m-b\rangle \geq 0 \\ \Leftrightarrow\ &\langle b-a, x-m\rangle \geq 0.\end{aligned} \tag{6.27}$$

Note from (6.17) that

$$\langle b-a, x-b\rangle = \langle b-x, x-b\rangle + \langle x-a, x-b\rangle = -\|x-b\|^2 + \langle x-a, x-b\rangle \leq 0. \tag{6.28}$$

Taking (6.18), (6.27) and (6.28) into account, we have that

$$\begin{aligned}\|x'-b\|^2 - \|x-b\|^2 &= \|x'-x\|^2 + 2\langle x'-x, x-b\rangle \\ &= \|x'-x\|^2 - 2\frac{\langle b-a, x-m\rangle}{\|b-a\|^2}\langle b-a, x-b\rangle \\ &\geq \|x'-x\|^2.\end{aligned} \tag{6.29}$$

By (6.19) and (6.20) we get that

$$\begin{aligned}\|x-a\|^2 + \|x-b\|^2 &= 2\|x-x'\|^2 + \|x'-a\|^2 + \|x'-b\|^2 + 2\langle x-x', 2x'-(a+b)\rangle \\ &= 2\|x-x'\|^2 + 2\|x'-b\|^2 + 4\langle x-x', x'-m\rangle \\ &= 2\|x-x'\|^2 + 2\|x'-b\|^2.\end{aligned}$$

This together with (6.11) and (6.29) yields that

$$\begin{aligned}2\|x-x'\|^2 &= \|x-a\|^2 + \|x-b\|^2 - 2\|x'-b\|^2 \\ &\leq (1+\delta')^2\|x-b\|^2 + \|x-b\|^2 - 2\|x'-b\|^2 \\ &\leq (1+\delta')^2\|x-b\|^2 + \|x-b\|^2 - 2\left(\|x-b\|^2 + \|x'-x\|^2\right).\end{aligned}$$

Equivalently,

$$4\|x-x'\|^2 \leq \left(2\delta' + \delta'^2\right) \|x-b\|^2 = \left(2\delta' + \delta'^2\right) \min\{\|x-a\|^2, \|x-b\|^2\} \tag{6.30}$$

since $\|x-a\| \geq \|x-b\|$ in this case.

By a similar argument, if $\|x-a\| \leq \|x-b\|$, then

$$\langle b-a, x-m\rangle \leq 0,\ \langle b-a, x-a\rangle \geq 0.$$

Thus

$$\begin{aligned} \|x'-a\|^2 - \|x-a\|^2 &= \|x'-x\|^2 + 2\left\langle x'-x, x-a\right\rangle \\ &= \|x'-x\|^2 - 2\frac{\langle b-a, x-m\rangle}{\|b-a\|^2}\left\langle b-a, x-a\right\rangle \\ &\geq \|x'-x\|^2. \end{aligned} \tag{6.31}$$

By (6.19) and (6.20) we get that

$$\|x-a\|^2 + \|x-b\|^2 = 2\|x-x'\|^2 + 2\|x'-a\|^2,$$

which together with (6.11) and (6.31) yields that

$$2\|x-x'\|^2 \leq \|x-a\|^2 + \tfrac{1}{(1-\delta')^2}\|x-a\|^2 - 2\left(\|x-a\|^2 + \|x'-x\|^2\right).$$

Equivalently,

$$4\|x-x'\|^2 \leq \tfrac{2\delta'-\delta'^2}{(1-\delta')^2}\|x-a\|^2 = \tfrac{2\delta'-\delta'^2}{(1-\delta')^2}\min\{\|x-a\|^2, \|x-b\|^2\} \tag{6.32}$$

since $\|x-a\| \leq \|x-b\|$ in this case.

Combining (6.30) and (6.32) and noting that $2\delta' + \delta'^2 < \dfrac{2\delta'-\delta'^2}{(1-\delta')^2}$, we obtain (6.26) as claimed.

Step 2. We show that

$$\|x'-a\|^2 \geq \tfrac{4-6\delta'+3\delta'^2}{2\delta'-\delta'^2}\|x-x'\|^2. \tag{6.33}$$

Indeed, if $\|x-a\| \leq \|x-b\|$, then the use of (6.31) and (6.32) yields (6.33):

$$\begin{aligned} \|x'-a\|^2 &\geq \|x-a\|^2 + \|x-x'\|^2 \\ &\geq \tfrac{4(1-\delta')^2)}{2\delta'-\delta'^2}\|x-x'\|^2 + \|x-x'\|^2 = \tfrac{4-6\delta'+3\delta'^2}{2\delta'-\delta'^2}\|x-x'\|^2. \end{aligned}$$

Otherwise, i.e., $\|x-a\| \geq \|x-b\|$, then the use of (6.19), (6.29) and (6.30) successively implies that

$$\begin{aligned} \|x'-a\|^2 = \|x'-b\|^2 &\geq \|x-b\|^2 + \|x-x'\|^2 \\ &\geq \tfrac{4}{2\delta'+\delta'^2}\|x-x'\|^2 + \|x-x'\|^2 = \tfrac{4+2\delta'+\delta'^2}{2\delta'+\delta'^2}\|x-x'\|^2, \end{aligned}$$

which also yields (6.33) since $\dfrac{4+2\delta'+\delta'^2}{2\delta'+\delta'^2} > \dfrac{4-6\delta'+3\delta'^2}{2\delta'-\delta'^2}$. Hence (6.33) has been proved.

Step 3. We show that

$$\left\| \frac{x'-a}{\|x'-a\|} - \frac{x-a}{\|x-a\|} \right\| \leq 2\frac{\|x-x'\|}{\|x'-a\|}. \tag{6.34}$$

Indeed,

$$\begin{aligned} \left\| \frac{x'-a}{\|x'-a\|} - \frac{x-a}{\|x-a\|} \right\| &\leq \left\| \frac{x'-a}{\|x'-a\|} - \frac{x-a}{\|x'-a\|} \right\| + \left\| \frac{x-a}{\|x'-a\|} - \frac{x-a}{\|x-a\|} \right\| \\ &= \frac{\|x-x'\|}{\|x'-a\|} + \left| \frac{\|x-a\|}{\|x'-a\|} - 1 \right|. \end{aligned}$$

If $\|x-a\| \geq \|x'-a\|$, then (6.34) holds true since

$$\left|\frac{\|x-a\|}{\|x'-a\|}-1\right| = \frac{\|x-a\|}{\|x'-a\|}-1 \leq \frac{\|x-x'\|+\|x'-a\|}{\|x'-a\|}-1 = \frac{\|x-x'\|}{\|x'-a\|}.$$

Otherwise, i.e., $\|x-a\| < \|x'-a\|$, then (6.34) also holds true since

$$\left|\frac{\|x-a\|}{\|x'-a\|}-1\right| = 1-\frac{\|x-a\|}{\|x'-a\|} \leq 1-\frac{\|x'-a\|-\|x-x'\|}{\|x'-a\|} = \frac{\|x-x'\|}{\|x'-a\|}.$$

Hence (6.34) has been proved.

A combination of (6.33) and (6.34) yields that

$$\left\|\frac{x'-a}{\|x'-a\|}-\frac{x-a}{\|x-a\|}\right\| \leq 2\sqrt{\tfrac{2\delta'-\delta'^2}{4-6\delta'+3\delta'^2}}. \tag{6.35}$$

Substitute (6.25) and (6.35) into (6.24) and using (6.10), we obtain that

$$\begin{aligned}\operatorname{dist}(v_1', N_A(a)) \leq \|v_1'-v_1\| \leq \|v_1\|\left(\sqrt{2\delta'}+2\sqrt{\frac{2\delta'-\delta'^2}{4-6\delta'+3\delta'^2}}\right)\\ < \sqrt{2\delta'}+2\sqrt{\frac{2\delta'-\delta'^2}{4-6\delta'+3\delta'^2}} < \delta.\end{aligned} \tag{6.36}$$

The proof of $\operatorname{dist}(v_2', N_B(b)) < \delta$ is analogous, and we also obtain that

$$\begin{aligned}\operatorname{dist}(v_2', N_B(b)) \leq \|v_2'-v_2\| \leq \|v_2\|\left(\sqrt{2\delta'}+2\sqrt{\frac{2\delta'-\delta'^2}{4-6\delta'+3\delta'^2}}\right)\\ < \sqrt{2\delta'}+2\sqrt{\frac{2\delta'-\delta'^2}{4-6\delta'+3\delta'^2}} < \delta.\end{aligned} \tag{6.37}$$

Hence (6.23) has been proved.

Conditions (6.23) and (6.22) ensure that the pair of vectors $\{v_1', v_2'\}$ satisfies conditions (6.7) and (6.8), respectively. It is trivial from the choice of δ' in (6.10) that $a \in (A\setminus B)\cap \mathbb{B}_\delta(\bar{x})$, $b \in (B\setminus A)\cap \mathbb{B}_\delta(\bar{x})$. We also have $x' \in \mathbb{B}_\delta(\bar{x})$ since

$$\begin{aligned}\|x'-\bar{x}\| &= \left\|x-\frac{\langle b-a, x-m\rangle}{\|b-a\|^2}(b-a)-\bar{x}\right\| \leq \|x-\bar{x}\|+\|x-m\| \leq \delta'+\max\{\|x-a\|,\|x-b\|\}\\ &\leq \delta'+\|x-\bar{x}\|+\max\{\|a-\bar{x}\|,\|b-\bar{x}\|\} \leq 3\delta' < \delta.\end{aligned}$$

Hence, the estimate (6.6) is applicable to $\{v_1', v_2'\}$. That is,

$$\|v_1'+v_2'\| > \alpha. \tag{6.38}$$

Now using the triangle inequality, (6.36), (6.37), (6.38), (6.10) and (6.5) successively, we obtain the desired estimate:

$$\begin{aligned}\|v_1+v_2\| &= \|v_1'+v_2'+v_1-v_1'+v_2-v_2'\|\\ &\geq \|v_1'+v_2'\|-\|v_1'-v_1\|-\|v_2'-v_2\|\\ &\geq \alpha-(\|v_1\|+\|v_2\|)\left(\sqrt{2\delta'}+2\sqrt{\frac{2\delta'+\delta'^2}{4+2\delta'+\delta'^2}}\right)\\ &= \alpha-\left(\sqrt{2\delta'}+2\sqrt{\frac{2\delta'-\delta'^2}{4-6\delta'+3\delta'^2}}\right)\\ &> \alpha-(\alpha-\beta) = \beta.\end{aligned}$$

This completes *Case 2* and (6.13) has been proved.

Hence, we have proved that $\{A,B\}$ is intrinsically transversal at $\bar{x}$ with $\text{itr}[A,B](\bar{x}) \geq \beta$. Since β can be arbitrarily close to $\min\left\{\text{itr}_c[A,B](\bar{x}), 1/\sqrt{2}\right\}$, we also obtain the estimate (6.4) and the proof is complete. □

Remark 6.1.12. The idea behind Lemma 6.1.11 comes from two observations. First, when pairs of vectors (v_1, v_2) appearing in the definitions of $\text{itr}[A,B](\bar{x})$ and $\text{itr}_c[A,B](\bar{x})$ are further restricted to the constraint $\langle v_1, v_2\rangle < 0$, the two groups of conditions defining the two constants become equivalent. Second, this additional constraint, which also gives rise to the number $1/\sqrt{2}$ in Lemma 6.1.11, does not qualitatively affect the properties characterised by these constants. See also Lemma 6.3.5 for a geometric counterpart of this constraint in the Euclidean space setting.

Combining Lemma 6.1.11 with Proposition 6.1.9, we obtain the main result of this section.

Theorem 6.1.13 (complete quantitative relationships). *Let $\{A,B\}$ be a pair of closed sets and $\bar{x} \in A\cap B$. Then the following statements hold true.*

(i) *If* $\text{itr}_c[A,B](\bar{x}) \geq 1/\sqrt{2}$, *then*

$$\min\{\text{itr}[A,B](\bar{x}), \text{itr}_w[A,B](\bar{x}), \text{itr}_c[A,B](\bar{x})\} \geq 1/\sqrt{2}. \tag{6.39}$$

(ii) *If* $\text{itr}_c[A,B](\bar{x}) < 1/\sqrt{2}$, *then*

$$\text{itr}_c[A,B](\bar{x}) = \text{itr}[A,B](\bar{x}) = \text{itr}_w[A,B](\bar{x}). \tag{6.40}$$

As a consequence, the three transversality-type properties characterised by the above constants are equivalent. In particular, $\text{itr}_c[A,B](\bar{x})$ *is a refined equivalent characterisation of intrinsic transversality.*

Proof. (i) In this case, it holds $\min\left\{\text{itr}_c[A,B](\bar{x}), 1/\sqrt{2}\right\} = 1/\sqrt{2}$. Lemma 6.1.11 then yields $\text{itr}[A,B](\bar{x}) \geq 1/\sqrt{2}$, which in turn implies (6.39) because the left-hand-side of (6.39) equals $\text{itr}[A,B](\bar{x})$ in view of Proposition 6.1.9(i).

(ii) In this case, it holds $\min\left\{\text{itr}_c[A,B](\bar{x}), 1/\sqrt{2}\right\} = \text{itr}_c[A,B](\bar{x})$. Lemma 6.1.11 then implies $\text{itr}_c[A,B](\bar{x}) \leq \text{itr}[A,B](\bar{x})$, which together with Proposition 6.1.9(i) yields the equalities in (6.40).

□

In view of Theorem 6.1.13, the following result covers both Propositions 6.1.3 & 6.1.4 in the Hilbert space setting. More importantly, it refines Proposition 6.1.4 which establishes the weakest sufficient condition in terms of normal vectors for subtransversality in the nonconvex setting.

Corollary 6.1.14 (refined sufficient condition for subtransversality). *A pair of closed sets $\{A,B\}$ is subtransversal at $\bar{x} \in A\cap B$ if* $\text{itr}_c[A,B](\bar{x}) > 0$. *The inverse implication is also true when the sets are convex.*

Proof. Let us suppose $\text{itr}_c[A,B](\bar{x}) > 0$. In both cases of either $\text{itr}_c[A,B](\bar{x}) \geq 1/\sqrt{2}$ or $\text{itr}_c[A,B](\bar{x}) < 1/\sqrt{2}$, Theorem 6.1.13 ensures that $\text{itr}_w[A,B](\bar{x}) > 0$. In view of Proposition 6.1.9(ii), the latter implies subtransversal of $\{A,B\}$ at $\bar{x}$. The statement about the inverse implication in the convex setting follows from Proposition 6.1.3.

□

Corollary 6.1.14 gives rise to the following question, which will be addressed in Remark 6.1.18.

Question 3. *Is the sufficient condition* $\mathrm{itr}_c[A,B](\bar{x}) > 0$ *also necessary for subtransversality in the nonconvex setting?*

The following result is of importance.

Corollary 6.1.15 (intrinsic transversality is subtransversality in the convex setting)**.** *For pairs of closed and convex sets, subtransversality and intrinsic transversality are equivalent.*

Proof. By Corollary 6.1.14, a pair of closed and convex sets $\{A,B\}$ is intrinsically transversal at $\bar{x} \in A \cap B$ if and only if $\mathrm{itr}_c[A,B](\bar{x}) > 0$. In view of Theorem 6.1.13, the latter is equivalent to $\mathrm{itr}_w[A,B](\bar{x}) > 0$, which in turns amounts to $\mathrm{str}[A,B](\bar{x}) > 0$ thanks to Proposition 6.1.9(iii). This is equivalent to subtransversality of $\{A,B\}$ at $\bar{x}$ due to Proposition 6.1.8(i).

□

As by-products, we address Questions 1 and 2 in the remainder of this section.

Remark 6.1.16 (answer to Question 2)*.* Theorem 6.1.13 clearly shows that weak intrinsic transversality (i.e., $\mathrm{itr}_w[A,B](\bar{x}) > 0$) and intrinsic transversality (i.e., $\mathrm{itr}[A,B](\bar{x}) > 0$) are equivalent in the Hilbert space setting. Note that it can happen that $\mathrm{itr}_w[A,B](\bar{x}) \neq \mathrm{itr}[A,B](\bar{x})$ if $\mathrm{itr}_c[A,B](\bar{x}) > 1/\sqrt{2}$.

It is worth keeping in mind that the question remains open in more general underlying spaces.

Remark 6.1.17 (answer to Question 1)*.* In view of Remark 6.1.16, the condition $\mathrm{itr}_w[A,B](\bar{x}) > 0$ (i.e, weak intrinsic transversality) is equivalent to intrinsic transversality of $\{A,B\}$ at $\bar{x}$. But it is widely known that the latter is not necessary for subtransversality in the nonconvex setting [53, 100, 128]. This establishes the negative answer to Question 1.

Remark 6.1.18 (answer to Question 3)*.* Thanks to Theorem 6.1.13, the two conditions $\mathrm{itr}_c[A,B](\bar{x}) > 0$ and $\mathrm{itr}_w[A,B](\bar{x}) > 0$ are equivalent. The latter is not necessary for subtransversality as explained in Remark 6.1.17. As a consequence, the condition $\mathrm{itr}_c[A,B](\bar{x}) > 0$ is not necessary for subtransversality in the nonconvex setting.

In conclusion, Theorem 6.1.13 allows us to unify and refine a number of transversality-type properties in Hilbert spaces including intrinsic transversality [52], weak intrinsic transversality [96], the sufficient conditions for subtransversality [99, 100, 106] and the characterisation of subtransversality with convexity [96]. This result significantly clarifies the picture of transversality-type properties.

6.2 Intrinsic Transversality in Primal Space Terms

In view of Corollary 6.1.15, for pairs of closed and convex sets in Hilbert spaces, intrinsic transversality is equivalent to subtransversality which is defined in primal space[8] terms. The situation for pairs of nonconvex sets has not been known and requires clarification; see, for example, the following question raised by Ioffe [75].

Question 4. *[75, page 358] What are primal space counterparts of intrinsic transversality?*

The analysis in this section is devoted to addressing the above question. The main goal is to characterise intrinsic transversality by conditions which do not involve normal vectors.

[8]In the framework of this section, the term *primal space* is occasionally used to indicate that the mentioned object/condition/property/representation does not explicitly involve normals.

For convenience, let us recall the notation of distance defined in (2.32) for three sets as follows[9]:

$$d_1(A,B,\Omega) := \inf_{x\in\Omega, a\in A, b\in B} \max\{\|x-a\|, \|x-b\|\}, \quad \forall A,B,\Omega \subset X,$$

with the convention that the infimum over the empty set equals infinity.

Definition 6.2.1 (property $(\mathcal{P})$[10])**.** A pair of closed sets $\{A,B\}$ is said to satisfy property $(\mathcal{P})$ at a point $\bar{x} \in A\cap B$ if there are numbers $\alpha \in]0,1[$ and $\varepsilon > 0$ such that for any $a \in (A\setminus B)\cap \mathbb{B}_\varepsilon(\bar{x})$, $b \in (B\setminus A)\cap \mathbb{B}_\varepsilon(\bar{x})$ and $x \in \mathbb{B}_\varepsilon(\bar{x})$ with $\|x-a\| = \|x-b\|$ and number $\delta > 0$, there exists $\rho \in]0,\delta[$ satisfying

$$d_1\left(A\cap \overline{\mathbb{B}}_\lambda(a), B\cap \overline{\mathbb{B}}_\lambda(b), \mathbb{B}_\rho(x)\right) + \alpha\rho \le \|x-a\|, \text{ where } \lambda := \left(\alpha + 1/\sqrt{\varepsilon}\right)\rho. \tag{6.41}$$

The exact upper bound of all $\alpha \in]0,1[$ such that $\{A,B\}$ satisfies property $(\mathcal{P})$ at $\bar{x}$ for α and some $\varepsilon > 0$ is denoted by $\mathrm{itr}_p[A,B](\bar{x})$.

The following statement is straightforward from the definition.

Proposition 6.2.2. *A pair of closed sets $\{A,B\}$ satisfies property $(\mathcal{P})$ at a point $\bar{x} \in A\cap B$ if and only if* $\mathrm{itr}_p[A,B](\bar{x}) > 0$.

We next reformulate several technical results which are essential for proving the equivalence between property $(\mathcal{P})$ and intrinsic transversality. We first recall Corollary 3.3.3.

Proposition 6.2.3. *[30, Corollary 6.3]*[11] *Let $\{A,B\}$ be a pair of closed sets in X, $\bar{x} \in A\cap B$, $u,v \in X$ and numbers $\rho,\varepsilon > 0$. Suppose that*

$$(A-u)\cap(B-v)\cap \mathbb{B}_\rho(\bar{x}) = \emptyset, \tag{6.42}$$

$$\max\{\|u\|, \|v\|\} < d_1(A-u, B-v, \mathbb{B}_\rho(\bar{x})) + \varepsilon. \tag{6.43}$$

Then for any numbers $\lambda \ge \varepsilon + \rho$ and $\tau \in \left[0, \frac{\lambda-\varepsilon}{\lambda+\varepsilon}\right[$, there exist points $\hat{a} \in A\cap \mathbb{B}_\lambda(\bar{x})$, $\hat{b} \in B\cap \mathbb{B}_\lambda(\bar{x})$, $\hat{x} \in \mathbb{B}_\rho(\bar{x})$ and vectors $v_1 \in N_A(\hat{a})$, $v_2 \in N_B(\hat{b})$ such that

$$\|v_1\| + \|v_2\| = 1, \ \|v_1+v_2\| < \varepsilon/\rho, \tag{6.44}$$

$$\langle v_1, \hat{x}-\hat{a}+u\rangle + \left\langle v_2, \hat{x}-\hat{b}+v\right\rangle > \tau \max\left\{\|\hat{x}-\hat{a}+u\|, \|\hat{x}-\hat{b}+v\|\right\}. \tag{6.45}$$

Remark 6.2.4. Condition (6.45) plays an important role in searching for primal space counterparts of intrinsic transversality because it relates normal vectors to the primal space elements.

Our subsequent analysis requires the following modified version of Proposition 6.2.3, where the reference point $\bar{x} \in A\cap B$ is replaced by a triple of points $(a,b,x) \in A\times B\times X$.

Proposition 6.2.5. *Let $\{A,B\}$ be a pair of closed sets in X, $a \in A$, $b \in B$, $x \in X$ and numbers $\rho,\varepsilon > 0$. Suppose that*

$$0 < d_1(A,B,\mathbb{B}_\rho(x)), \tag{6.46}$$

$$\max\{\|x-a\|, \|x-b\|\} < d_1(A,B,\mathbb{B}_\rho(x)) + \varepsilon. \tag{6.47}$$

Then for any numbers $\lambda \ge \varepsilon + \rho$ and $\tau \in \left[0, \frac{\lambda-\varepsilon}{\lambda+\varepsilon}\right[$, there exist points $\hat{a} \in A\cap \mathbb{B}_\lambda(a)$, $\hat{b} \in B\cap \mathbb{B}_\lambda(b)$, $\hat{x} \in \mathbb{B}_\rho(x)$ and vectors $v_1 \in N_A(\hat{a})$, $v_2 \in N_B(\hat{b})$ satisfying (6.44) *and*

$$\langle v_1, \hat{x}-\hat{a}\rangle + \langle v_2, \hat{x}-\hat{b}\rangle > \tau \max\left\{\|\hat{x}-\hat{a}\|, \|\hat{x}-\hat{b}\|\right\}. \tag{6.48}$$

[9]This is the distance between the two sets $A\times B$ and $\{(x,x) \in X\times X \mid x \in \Omega\}$ in $X\times X$ endowed with the maximum norm.

[10]The definition is valid in normed linear spaces.

[11]The result is valid in Asplund spaces.

Proof. The idea is to apply Proposition 6.2.3 to

$$A' := A - a,\ B' := B - b,\ \bar{x} := 0 \in A' \cap B',\ u := x - a \ \text{ and } \ v := x - b \tag{6.49}$$

by verifying (6.42) and (6.43). Indeed, condition (6.46) implies that

$$A \cap B \cap \mathbb{B}_\rho(x) = \emptyset, \text{ or equivalently, } (A - x) \cap (B - x) \cap (\rho\mathbb{B}) = \emptyset.$$

The latter is exactly (6.42) since by (6.49)

$$(A' - u) \cap (B' - v) \cap \mathbb{B}_\rho(\bar{x}) = (A - x) \cap (B - x) \cap (\rho\mathbb{B}).$$

Note also from (6.49) that

$$d_1\left(A' - u, B' - v, \mathbb{B}_\rho(\bar{x})\right) = d_1\left(A - x, B - x, \rho\mathbb{B}\right) = d_1\left(A, B, \mathbb{B}_\rho(x)\right).$$

This together with (6.47) yields condition (6.43).

In view of Proposition 6.2.3, there exist points

$$a' \in A' \cap (\lambda\mathbb{B}),\ b' \in B' \cap (\lambda\mathbb{B}),\ x' \in \rho\mathbb{B} \tag{6.50}$$

and vectors $v_1 \in N_{A'}(a')$, $v_2 \in N_{B'}(b')$ satisfying conditions (6.44) and (6.45). Let us define

$$\hat{a} := a' + a,\ \hat{b} := b' + b \ \text{ and } \ \hat{x} := x' + x.$$

This together with (6.49) and (6.50) ensures that $\hat{a} \in A \cap \mathbb{B}_\lambda(a)$, $\hat{b} \in B \cap \mathbb{B}_\lambda(b)$ and $\hat{x} \in \mathbb{B}_\rho(x)$. Note also that

$$N_A(\hat{a}) = N_{A'}(a'),\ N_B(\hat{b}) = N_{B'}(b'), \tag{6.51}$$

$$\hat{x} - \hat{a} = x' - a' + x - a = x' - a' + u,\ \hat{x} - \hat{b} = x' - b' + x - b = x' - b' + v. \tag{6.52}$$

The combination of (6.45) and (6.52) yields (6.48) while condition (6.51) ensures that $v_1 \in N_A(\hat{a})$, $v_2 \in N_B(\hat{b})$.

Therefore, the points $\hat{a}$, $\hat{b}$, $\hat{x}$ and vectors v_1, v_2 satisfy all the required conditions of the proposition. The proof is complete.

□

Another technical result is needed for our analysis.

Proposition 6.2.6. [12] *Let $\{A, B\}$ be a pair of closed sets in X, $a \in A$, $b \in B$, $x \in X$ with $\|x - a\| = \|x - b\| > 0$, $\alpha, \varepsilon > 0$ and vectors $v_1, v_2 \in X$ with $\|v_1\| + \|v_2\| = 1$. Suppose that the following conditions are satisfied:*

$$\|v_1 + v_2\| + 2\varepsilon\left(\alpha + 1/\sqrt{\varepsilon}\right) < \alpha, \tag{6.53}$$

$$\operatorname{dist}(v_1, N_A(a)) < \varepsilon,\ \operatorname{dist}(v_2, N_B(b)) < \varepsilon, \tag{6.54}$$

$$\langle v_1, x - a\rangle = \|v_1\|\,\|x - a\|,\ \langle v_2, x - b\rangle = \|v_2\|\,\|x - b\|. \tag{6.55}$$

Then there exists a number $\delta > 0$ such that

$$d_1\left(A \cap \overline{\mathbb{B}}_\lambda(a), B \cap \overline{\mathbb{B}}_\lambda(b), \mathbb{B}_\rho(x)\right) + \alpha\rho > \|x - a\|,\ \textit{where } \lambda := \left(\alpha + 1/\sqrt{\varepsilon}\right)\rho, \tag{6.56}$$

holds true for all $\rho \in]0, \delta[$.

[12] The result is valid in normed linear spaces.

Proof. By (6.54), there exist vectors $u_1 \in N_A(a)$ and $u_2 \in N_B(b)$ such that

$$\|v_1 - u_1\| < \varepsilon, \ \|v_2 - u_2\| < \varepsilon. \tag{6.57}$$

Since $\alpha - \|v_1 + v_2\| > 2\varepsilon(\alpha + 1/\sqrt{\varepsilon})$ due to (6.53), there are positive numbers α_1 and α_2 such that

$$\alpha > \alpha_1 > \alpha_2 > \|v_1 + v_2\|, \ \alpha_1 - \alpha_2 > 2\varepsilon(\alpha + 1/\sqrt{\varepsilon}). \tag{6.58}$$

Choose a number $\beta > 0$ such that

$$\beta < \frac{\alpha_1 - \alpha_2}{\alpha + 1/\sqrt{\varepsilon}} - 2\varepsilon, \text{ equivalently, } \alpha_1 - \alpha_2 - (2\varepsilon + \beta)(\alpha + 1/\sqrt{\varepsilon}) > 0. \tag{6.59}$$

By the definition (2.1) of the Fréchet normal cone, there is a number $\delta' > 0$ such that

$$\langle u_1, a' - a\rangle \le \tfrac{\beta}{2}\|a' - a\|, \ \langle u_2, b' - b\rangle \le \tfrac{\beta}{2}\|b' - b\|, \quad \forall a' \in A \cap \mathbb{B}_{\delta'}(a), \, b' \in B \cap \mathbb{B}_{\delta'}(b). \tag{6.60}$$

Let us define

$$\delta := \tfrac{\delta'}{\alpha + 1/\sqrt{\varepsilon}} > 0 \tag{6.61}$$

and show that δ fulfills the requirement of the proposition. Indeed, let us suppose to the contrary that condition (6.56) is not satisfied for some $\rho \in]0, \delta[$. That is,

$$d_1\left(A \cap \overline{\mathbb{B}}_\lambda(a), B \cap \overline{\mathbb{B}}_\lambda(b), \mathbb{B}_\rho(x)\right) + \rho\alpha \le \|x - a\|, \text{ where } \lambda := \left(\alpha + 1/\sqrt{\varepsilon}\right)\rho.$$

Since $\alpha_1 < \alpha$, the above inequality ensures the existence of $\hat{a} \in A \cap \overline{\mathbb{B}}_\lambda(a)$, $\hat{b} \in B \cap \overline{\mathbb{B}}_\lambda(b)$ and $\hat{x} \in \mathbb{B}_\rho(x)$ such that

$$\max\{\|\hat{x} - \hat{a}\|, \|\hat{x} - \hat{b}\|\} < \|x - a\| - \rho\alpha_1. \tag{6.62}$$

Note that $\lambda = \left(\alpha + 1/\sqrt{\varepsilon}\right)\rho < (\alpha + 1/\sqrt{\varepsilon})\delta = \delta'$ thanks to (6.61). Then in view of (6.60), we have

$$\langle u_1, \hat{a} - a\rangle \le \tfrac{\beta}{2}\|\hat{a} - a\|, \ \langle u_2, \hat{b} - b\rangle \le \tfrac{\beta}{2}\|\hat{b} - b\|.$$

This implies that

$$\langle u_1, \hat{a} - a\rangle + \langle u_2, \hat{b} - b\rangle \le \tfrac{\beta}{2}\|\hat{a} - a\| + \tfrac{\beta}{2}\|\hat{b} - b\| \le \tfrac{\beta}{2}\lambda + \tfrac{\beta}{2}\lambda = \lambda\beta. \tag{6.63}$$

By (6.55), $\|v_1\| + \|v_2\| = 1$ and $\|x - a\| = \|x - b\|$, it holds

$$\langle v_1, x - a\rangle + \langle v_2, x - b\rangle = \|v_1\|\|x - a\| + \|v_2\|\|x - b\| = \|x - a\|. \tag{6.64}$$

By (6.62) and $\|v_1\| + \|v_2\| = 1$, it holds

$$\langle v_1, \hat{x} - \hat{a}\rangle + \langle v_2, \hat{x} - \hat{b}\rangle \le \max\{\|\hat{x} - \hat{a}\|, \|\hat{x} - \hat{b}\|\} < \|x - a\| - \rho\alpha_1. \tag{6.65}$$

By the Cauchy-Schwarz inequality and $\alpha_2 > \|v_1 + v_2\|$ in view of (6.58), it holds

$$\langle v_1 + v_2, x - \hat{x}\rangle \le \|v_1 + v_2\|\,\|x - \hat{x}\| < \rho\alpha_2. \tag{6.66}$$

By the Cauchy-Schwarz inequality and (6.57), it holds

$$\langle v_1 - u_1, \hat{a} - a\rangle + \langle v_2 - u_2, \hat{b} - b\rangle \le \varepsilon\|\hat{a} - a\| + \varepsilon\|\hat{b} - b\| \le 2\lambda\varepsilon. \tag{6.67}$$

Adding (6.67) and (6.63) yields that

$$\langle v_1, \hat{a} - a\rangle + \langle v_2, \hat{b} - b\rangle \le 2\lambda\varepsilon + \lambda\beta. \tag{6.68}$$

Making use of (6.64), (6.65), (6.66), (6.68), $\lambda = \left(\alpha + 1/\sqrt{\varepsilon}\right)\rho$ and (6.59) successively, we come up with

$$\begin{aligned}
\|x - a\| &= \langle v_1, x - a\rangle + \langle v_2, x - b\rangle \\
&= \langle v_1, \hat{x} - \hat{a}\rangle + \langle v_2, \hat{x} - \hat{b}\rangle + \langle v_1 + v_2, x - \hat{x}\rangle + \langle v_1, \hat{a} - a\rangle + \langle v_2, \hat{b} - b\rangle \\
&< \|x - a\| - \rho\alpha_1 + \rho\alpha_2 + 2\lambda\varepsilon + \lambda\beta \\
&= \|x - a\| - \rho\left(\alpha_1 - \alpha_2 - (2\varepsilon + \beta)(\alpha + 1/\sqrt{\varepsilon})\right) < \|x - a\|,
\end{aligned}$$

which is a contradiction and hence the proof is complete.

□

Remark 6.2.7. Condition (6.56) holding true for all $\rho \in]0, \delta[$ is the negation of condition (6.41) holding true for some $\rho \in]0, \delta[$.

The next lemma establishes the key quantitative estimates of this section.

Lemma 6.2.8 (quantitative estimates[13]). *Let $\{A, B\}$ be a pair of closed sets and $\bar{x} \in A \cap B$. Then*

$$\text{itr}[A, B](\bar{x}) \leq \text{itr}_p[A, B](\bar{x}) \leq \text{itr}_c[A, B](\bar{x}). \tag{6.69}$$

Proof. We first prove $\text{itr}[A, B](\bar{x}) \leq \text{itr}_p[A, B](\bar{x})$. Since the inequality becomes trivial when $\text{itr}_p[A, B](\bar{x}) = 1$, we only need to prove the inequality for the case $\text{itr}_p[A, B](\bar{x}) < 1$. We take an arbitrary number α satisfying $\text{itr}_p[A, B](\bar{x}) < \alpha \leq 1$ and show that $\text{itr}[A, B](\bar{x}) \leq \alpha$. To do this, let us first recall the following representation of $\text{itr}[A, B](\bar{x})$ [99, Equation (72)]:

$$\text{itr}[A, B](\bar{x}) = \liminf_{\substack{a \to \bar{x},\ b \to \bar{x},\ x \to \bar{x},\ a \in A \setminus B,\ b \in B \setminus A \\ v_1 \in N_A(a) \setminus \{0\},\ v_2 \in N_B(b) \setminus \{0\},\ \|v_1\| + \|v_2\| = 1 \\ \frac{\|x-a\|}{\|x-b\|} \to 1,\ \frac{\langle v_1, x-a \rangle}{\|v_1\| \|x-a\|} \to 1,\ \frac{\langle v_2, x-b \rangle}{\|v_2\| \|x-b\|} \to 1}} \|v_1 + v_2\|, \tag{6.70}$$

with the convention that the infimum over the empty set equals 1. In view of (6.70), all we need is to show that for any (arbitrarily small) number $\varepsilon > 0$, there exists a pair of vectors (v_1, v_2) satisfying the constraints under the $\liminf$ in (6.70) and $\|v_1 + v_2\| < \alpha$.

By the definition of $\text{itr}_p[A, B](\bar{x})$ and the inequality $\text{itr}_p[A, B](\bar{x}) < \alpha$, we have that for any number $\varepsilon > 0$, there exist points $a \in (A \setminus B) \cap \mathbb{B}_\varepsilon(\bar{x})$, $b \in (B \setminus A) \cap \mathbb{B}_\varepsilon(\bar{x})$ and $x \in \mathbb{B}_\varepsilon(\bar{x})$ with $\|x - a\| = \|x - b\|$ and a number $\delta > 0$ such that the inequality (6.56) holds true for all $\rho \in]0, \delta[$.

Since $A \cap B$ is a closed set and $a, b \notin A \cap B$, there is a number $\gamma \in]0, \varepsilon[$ such that

$$\mathbb{B}_\gamma(a) \cap (A \cap B) = \emptyset,\ \mathbb{B}_\gamma(b) \cap (A \cap B) = \emptyset. \tag{6.71}$$

Let us take an arbitrary number $\varepsilon \in]0, 1[$ and choose a number $\rho > 0$ satisfying

$$\rho < \min\left\{\delta,\ \tfrac{\gamma}{\alpha + 1/\sqrt{\varepsilon}},\ \varepsilon,\ \tfrac{\varepsilon\|x-a\|}{1+\alpha+1/\sqrt{\varepsilon}}\right\}. \tag{6.72}$$

Define a number $\varepsilon' := \alpha\rho > 0$ and note from (6.72) that

$$\varepsilon' = \alpha\rho < \left(\alpha + 1 + 1/\sqrt{\varepsilon}\right)\rho < \varepsilon\|x - a\| < \|x - a\| = \max\{\|x - a\|, \|x - b\|\}. \tag{6.73}$$

We are going to apply Proposition 6.2.5 to the sets $A' := A \cap \overline{\mathbb{B}}_\lambda(a)$, $B' := B \cap \overline{\mathbb{B}}_\lambda(b)$, the points $a \in A'$, $b \in B'$, $x \in X$ and numbers ρ, ε', λ and $\tau := \frac{\lambda - 2\alpha\rho}{\lambda + \alpha\rho}$. Let us verify all the conditions of the proposition. First, the inequality (6.56) reduces to (6.47). It also implies (6.46) since $d_1(A', B', \mathbb{B}_\rho(x)) > \|x - a\| - \alpha\rho > 0$ thanks to (6.73). It is clear that $\varepsilon' + \rho = (\alpha + 1)\rho < \left(\alpha + 1/\sqrt{\varepsilon}\right)\rho = \lambda$ as $\varepsilon \in]0, 1[$ and $\tau = \frac{\lambda - 2\alpha\rho}{\lambda + \alpha\rho} < \frac{\lambda - \alpha\rho}{\lambda + \alpha\rho} = \frac{\lambda - \varepsilon'}{\lambda + \varepsilon'}$. We have checked all the conditions of Proposition 6.2.5. Therefore, in view of Proposition 6.2.5, there exist points $\hat{a} \in A' \cap \mathbb{B}_\lambda(a)$, $\hat{b} \in B' \cap \mathbb{B}_\lambda(b)$, $\hat{x} \in \mathbb{B}_\rho(x)$ and vectors $v_1 \in N_{A'}(\hat{a})$ and $v_2 \in N_{B'}(\hat{b})$ satisfying (6.48) and

$$\|v_1\| + \|v_2\| = 1,\ \|v_1 + v_2\| < \tfrac{\varepsilon'}{\rho} = \alpha. \tag{6.74}$$

The following observations verify that the vectors (v_1, v_2) in conjunction with the points $(\hat{a}, \hat{b}, \hat{x})$ fulfill the constraints under the $\liminf$ in (6.70).

[13]The first inequality in (6.69) holds true in Asplund spaces while the second one holds true in normed linear spaces.

- By the triangle inequality and $\rho < \varepsilon$ due to (6.72), it holds that

$$\begin{aligned}\|\bar{x}-\hat{x}\| &\le \|\bar{x}-x\| + \|x-\hat{x}\| < \varepsilon + \rho \le 2\varepsilon,\\ \|\bar{x}-\hat{a}\| &\le \|\bar{x}-a\| + \|a-\hat{a}\| < \varepsilon + \lambda = \varepsilon + \left(\alpha + 1/\sqrt{\varepsilon}\right)\rho \le (1+\alpha)\varepsilon + \sqrt{\varepsilon},\\ \|\bar{x}-\hat{b}\| &\le \|\bar{x}-b\| + \|b-\hat{b}\| < \varepsilon + \lambda = \varepsilon + (\alpha + 1/\sqrt{\varepsilon})\rho \le (1+\alpha)\varepsilon + \sqrt{\varepsilon}.\end{aligned}$$

  This implies that $\hat{x} \to \bar{x}$, $\hat{a} \to \bar{x}$, $\hat{b} \to \bar{x}$ as $\varepsilon \downarrow 0$.

- Since $\lambda < \gamma$ by the choice of ρ in (6.72), we have $\hat{a} \in A \cap \mathbb{B}_\gamma(a)$ which together with (6.71) implies that $\hat{a} \notin B$. That is $\hat{a} \in A \setminus B$. Similarly, we also have $\hat{b} \in B \setminus A$.

- It holds $v_1 \neq 0$ and $v_2 \neq 0$. Indeed, if otherwise, (6.74) implies $\|v_1 + v_2\| = 1 < \alpha$, which is a contradiction to the inequality $\alpha \le 1$.

- Since $\hat{a} \in A' \cap \mathbb{B}_\lambda(a) = A \cap \mathbb{B}_\lambda(a)$ and $\hat{b} \in B' \cap \mathbb{B}_\lambda(b) = B \cap \mathbb{B}_\lambda(b)$ by the definitions of A' and B', we have

$$v_1 \in N_{A'}(\hat{a}) = N_{A\cap\overline{\mathbb{B}}_\lambda(a)}(\hat{a}) = N_A(\hat{a}),\ v_2 \in N_{A'}(\hat{a}) = N_{B\cap\overline{\mathbb{B}}_\lambda(b)}(\hat{b}) = N_B(\hat{b}).$$

- By the triangle inequality and $\left(\alpha + 1 + 1/\sqrt{\varepsilon}\right)\rho < \varepsilon\,\|x-a\|$ in view of (6.72), it holds that

$$\begin{aligned}\|\hat{x}-\hat{a}\| &\le \|x-a\| + \|\hat{x}-x\| + \|a-\hat{a}\| \le \|x-a\| + \varepsilon' + \lambda\\ &= \|x-a\| + \left(\alpha + 1 + 1/\sqrt{\varepsilon}\right)\rho < (1+\varepsilon)\,\|x-a\|,\\ \|\hat{x}-\hat{a}\| &\ge \|x-a\| - \|\hat{x}-x\| - \|a-\hat{a}\| \ge \|x-a\| - \varepsilon' - \lambda\\ &= \|x-a\| - \left(\alpha + 1 + 1/\sqrt{\varepsilon}\right)\rho > (1-\varepsilon)\,\|x-a\|.\end{aligned}$$

  Hence, we have

$$(1-\varepsilon)\,\|x-a\| \le \|\hat{x}-\hat{a}\| \le (1+\varepsilon)\,\|x-a\|.$$

  Using similar estimates, we also have

$$(1-\varepsilon)\,\|x-b\| \le \|\hat{x}-\hat{b}\| \le (1+\varepsilon)\,\|x-b\|.$$

  The above estimates together with $\|x-a\| = \|x-b\| \neq 0$ imply

$$\frac{1-\varepsilon}{1+\varepsilon} \le \frac{\|\hat{x}-\hat{a}\|}{\|\hat{x}-\hat{b}\|} \le \frac{1+\varepsilon}{1-\varepsilon},$$

  which in turn implies that $\dfrac{\|\hat{x}-\hat{a}\|}{\|\hat{x}-\hat{b}\|} \to 1$ as $\varepsilon \downarrow 0$.

- By (6.74), the Cauchy-Schwarz inequality, (6.48) and the definition of τ, we have

$$\begin{aligned}1 = \|v_1\| + \|v_2\| \quad &\ge \quad \frac{\langle v_1, \hat{x}-\hat{a}\rangle}{\|\hat{x}-\hat{a}\|} + \frac{\langle v_2, \hat{x}-\hat{b}\rangle}{\|\hat{x}-\hat{b}\|}\\ &\ge \quad \frac{\langle v_1, \hat{x}-\hat{a}\rangle}{\max\left\{\|\hat{x}-\hat{a}\|, \|\hat{x}-\hat{b}\|\right\}} + \frac{\langle v_2, \hat{x}-\hat{b}\rangle}{\max\left\{\|\hat{x}-\hat{a}\|, \|\hat{x}-\hat{b}\|\right\}}\\ &> \quad \tau = \frac{\lambda - 2\alpha\rho}{\lambda + \alpha\rho} = \frac{(\alpha + 1/\sqrt{\varepsilon})\,\rho - 2\alpha\rho}{(\alpha + 1/\sqrt{\varepsilon})\,\rho + \alpha\rho} = \frac{1/\sqrt{\varepsilon} - \alpha}{1/\sqrt{\varepsilon} + 2\alpha},\end{aligned}$$

  which tends to 1 as $\varepsilon \downarrow 0$. Thus,

$$\frac{\langle v_1, \hat{x}-\hat{a}\rangle}{\|\hat{x}-\hat{a}\|} + \frac{\langle v_2, \hat{x}-\hat{b}\rangle}{\|\hat{x}-\hat{b}\|} \to 1 \quad \text{as} \quad \varepsilon \downarrow 0.$$

  Due to the Cauchy-Schwarz inequality and $\|v_1\| + \|v_2\| = 1$, the above convergence happens if and only if

$$\frac{\langle v_1, \hat{x}-\hat{a}\rangle}{\|v_1\|\,\|\hat{x}-\hat{a}\|} \to 1 \quad \text{and} \quad \frac{\langle v_2, \hat{x}-\hat{b}\rangle}{\|v_2\|\,\|\hat{x}-\hat{b}\|} \to 1 \quad \text{as} \quad \varepsilon \downarrow 0.$$

By (6.74) and the above observations, letting $\varepsilon \downarrow 0$ implies $\text{itr}[A,B](\bar{x}) \le \alpha$ in view of (6.70).

We now prove $\text{itr}_p[A,B](\bar{x}) \le \text{itr}_c[A,B](\bar{x})$. Since the inequality becomes trivial when $\text{itr}_c[A,B](\bar{x}) = 1$, we only need a proof for the case $\text{itr}_c[A,B](\bar{x}) < 1$. We take an arbitrary number α satisfying $\text{itr}_c[A,B](\bar{x}) < \alpha < 1$ and prove that $\text{itr}_p[A,B](\bar{x}) \le \alpha$. From Definition 6.2.1, to obtain this inequality, it suffices to show that $\{A,B\}$ does not satisfy property $(\mathcal{P})$ at $\bar{x}$ with α (and any number $\varepsilon > 0$).

Fix a number α_1 satisfying $\text{itr}_c[A,B](\bar{x}) < \alpha_1 < \alpha$. Since $2\varepsilon(\alpha + 1/\sqrt{\varepsilon}) \downarrow 0$ as $\varepsilon \downarrow 0$ and $\alpha - \alpha_1 > 0$, there exists a number $\varepsilon_0 > 0$ such that $\varepsilon \in]0, \varepsilon_0[$ is equivalent to $\varepsilon > 0$ satisfying

$$2\varepsilon\left(\alpha + 1/\sqrt{\varepsilon}\right) < \alpha - \alpha_1. \tag{6.75}$$

We claim that for any $\varepsilon \in]0, \varepsilon_0[$, there exist $a \in (A \setminus B) \cap \mathbb{B}_\varepsilon(\bar{x})$, $b \in (B \setminus A) \cap \mathbb{B}_\varepsilon(\bar{x})$, $x \in \mathbb{B}_\varepsilon(\bar{x})$ with $\|x - a\| = \|x - b\|$ and a number $\delta > 0$ such that (6.56) holds true for all $\rho \in]0, \delta[$.

To prove the above claim, we first recall the following representation of $\text{itr}_c[A,B](\bar{x})$ [96, Equation (15)]:

$$\text{itr}_c[A,B](\bar{x}) = \liminf_{\substack{a \to \bar{x},\ b \to \bar{x},\ x \to \bar{x} \\ a \in A \setminus B,\ b \in B \setminus A,\ \|x-a\| = \|x-b\| \\ \text{dist}\,(v_1, N_A(a)) \to 0,\ \text{dist}\,(v_2, N_B(b)) \to 0,\ \|v_1\| + \|v_2\| = 1 \\ \langle v_1, x-a \rangle = \|v_1\|\,\|x-a\|,\ \langle v_2, x-b \rangle = \|v_2\|\,\|x-b\|}} \|v_1 + v_2\|, \tag{6.76}$$

with the convention that the infimum over the empty set equals 1.

In view of (6.76) and $\text{itr}_c[A,B](\bar{x}) < \alpha_1$, there exist $a \in (A \setminus B) \cap \mathbb{B}_\varepsilon(\bar{x})$, $b \in (B \setminus A) \cap \mathbb{B}_\varepsilon(\bar{x})$, $x \in \mathbb{B}_\varepsilon(\bar{x})$ with $\|x - a\| = \|x - b\|$ and $v_1, v_2 \in X$ satisfying

$$\begin{gathered} \text{dist}\,(v_1, N_A(a)) < \varepsilon,\ \text{dist}\,(v_2, N_B(b)) < \varepsilon,\ \|v_1\| + \|v_2\| = 1,\ \|v_1 + v_2\| < \alpha_1, \\ \langle v_1, x - a \rangle = \|v_1\|\,\|x - a\|\,,\ \langle v_2, x - b \rangle = \|v_2\|\,\|x - b\|\,. \end{gathered} \tag{6.77}$$

The last inequality in (6.77) and (6.75) imply $\|v_1 + v_2\| + 2\varepsilon\left(\alpha + 1/\sqrt{\varepsilon}\right) < \alpha$, which is (6.53).

Thus, the points a, b, x together with the numbers $\alpha, \varepsilon > 0$ and the vectors v_1, v_2 satisfy all the assumptions of Proposition 6.2.6. In view of this proposition, there exists a number $\delta > 0$ such that (6.56) holds true for all $\rho \in]0, \delta[$.

Hence, we have proven the above claim, which in turn implies that $\{A,B\}$ does not satisfy property $(\mathcal{P})$ at $\bar{x}$ with the number α. This is because in view of Remark 6.2.7, the statement that (6.56) holds true for all $\rho \in]0, \delta[$ is the negation of the statement that there exists $\rho \in]0, \delta[$ such that (6.41) holds true. Then from Definition 6.2.1, we have $\text{itr}_p[A,B](\bar{x}) \le \alpha$.

The proof is complete.

□

The first estimate in (6.69) shows that property $(\mathcal{P})$ is a necessary condition for intrinsic transversality while the second one shows that it is a sufficient condition for the property characterised by $\text{itr}_c[A,B](\bar{x})$. A combination of Lemma 6.2.8 and Theorem 6.1.13 eliminates the gap between these properties in the Hilbert space setting.

Theorem 6.2.9 (equivalences)**.** *Let $\{A,B\}$ be a pair of closed sets and $\bar{x} \in A \cap B$. Then the following statements hold true.*

(i) *If* $\operatorname{itr}_c[A,B](\bar{x}) \geq 1/\sqrt{2}$, *then*

$$\min\{\operatorname{itr}[A,B](\bar{x}), \operatorname{itr}_w[A,B](\bar{x}), \operatorname{itr}_p[A,B](\bar{x}), \operatorname{itr}_c[A,B](\bar{x})\} \geq 1/\sqrt{2}.$$

(ii) *If* $\operatorname{itr}_c[A,B](\bar{x}) < 1/\sqrt{2}$, *then*

$$\operatorname{itr}[A,B](\bar{x}) = \operatorname{itr}_w[A,B](\bar{x}) = \operatorname{itr}_p[A,B](\bar{x}) = \operatorname{itr}_c[A,B](\bar{x}).$$

Proof. Item (i) follows from Theorem 6.1.13(i) and (6.69) while item (ii) follows from Theorem 6.1.13(ii) and (6.69).

□

The following observation is obvious in view of Theorem 6.2.9.

Remark 6.2.10 (answer to Question 4). In Hilbert spaces, intrinsic transversality is equivalent to property $(\mathcal{P})$. Note that the latter does not involve elements of normal cones. Note also that it can happen that $\operatorname{itr}_p[A,B](\bar{x}) \neq \operatorname{itr}[A,B](\bar{x})$ if $\operatorname{itr}_c[A,B](\bar{x}) > 1/\sqrt{2}$.

6.3 Intrinsic Transversality in Euclidean Spaces

We first recall definitions of the relative limiting normals which are motivated by the compactness of the unit sphere in finite dimensional spaces as well as the fact that not all normal vectors are always involved in characterising transversality-type properties. These notions were shown to be useful for analysing intrinsic transversality and its variants, see [96, page 123] for a more thorough discussion.

In this section, X is a Euclidean space.

Definition 6.3.1 (relative limiting normal)**.** [96, Definition 2] Let A, B be closed sets and $\bar{x} \in A \cap B$.

(i) A pair $(v_1, v_2) \in X \times X$ is called a *pair of relative limiting normals* to $\{A, B\}$ at $\bar{x}$ if there exist sequences $(a_k) \subset A \setminus B$, $(b_k) \subset B \setminus A$, $(x_k) \subset X$ and $(v_{1k}), (v_{2k}) \subset X$ such that $a_k \to \bar{x}$, $b_k \to \bar{x}$, $x_k \to \bar{x}$, $v_{1k} \to v_1$, $v_{2k} \to v_2$, and

$$v_{1k} \in N_A(a_k),\ v_{2k} \in N_B(b_k) \quad (k = 1, 2, \ldots),$$
$$\frac{\|x_k - a_k\|}{\|x_k - b_k\|} \to 1,\ \frac{\langle v_{1k}, x_k - a_k\rangle}{\|v_{1k}\|\,\|x_k - a_k\|} \to 1,\ \frac{\langle v_{2k}, x_k - b_k\rangle}{\|v_{2k}\|\,\|x_k - b_k\|} \to 1,$$

with the convention that $\frac{0}{0} = 1$. The set of all pairs of relative limiting normals to $\{A, B\}$ at $\bar{x}$ is denoted by $\overline{N}_{A,B}(\bar{x})$.

(ii) A pair $(v_1, v_2) \in X \times X$ is called a *pair of restricted relative limiting normals* to $\{A, B\}$ at $\bar{x}$ if there exist sequences $(a_k) \subset A \setminus B$, $(b_k) \subset B \setminus A$, $(x_k) \subset X$ and $(v_{1k}), (v_{2k}) \subset X$ such that $\|x_k - a_k\| = \|x_k - b_k\|$ $(k = 1, 2, \ldots)$, $a_k \to \bar{x}$, $b_k \to \bar{x}$, $x_k \to \bar{x}$, $v_{1k} \to v_1$, $v_{2k} \to v_2$, and

$$\operatorname{dist}(v_{1k}, N_A(a_k)) \to 0,\ \operatorname{dist}(v_{2k}, N_B(b_k)) \to 0,$$
$$\langle v_{1k}, x_k - a_k\rangle = \|v_{1k}\|\,\|x_k - a_k\|,\ \langle v_{2k}, x_k - b_k\rangle = \|v_{2k}\|\,\|x_k - b_k\| \quad (k = 1, 2, \ldots).$$

The set of all pairs of restricted relative limiting normals to $\{A, B\}$ at $\bar{x}$ is denoted by $\overline{N}^c_{A,B}(\bar{x})$.

The following statement recalls a basic property of $\overline{N}_{A,B}(\bar{x})$ and $\overline{N}^c_{A,B}(\bar{x})$. In particular, they are cones[14].

[14] Here, the empty set is viewed as a cone by convention.

Proposition 6.3.2. *[96, Proposition 2(i)] Let A, B be closed sets and $\bar{x} \in A \cap B$. The sets $\overline{N}_{A,B}(\bar{x})$ and $\overline{N}^c_{A,B}(\bar{x})$ are closed cones in $X \times X$. Moreover, if $(v_1, v_2) \in \overline{N}_{A,B}(\bar{x})$ (respectively, $\overline{N}^c_{A,B}(\bar{x})$), then $(t_1 v_1, t_2 v_2) \in \overline{N}_{A,B}(\bar{x})$ (respectively, $\overline{N}^c_{A,B}(\bar{x})$) for all $t_1, t_2 > 0$.*

The following equalities show the important role of the cones $\overline{N}_{A,B}(\bar{x})$ and $\overline{N}^c_{A,B}(\bar{x})$ in characterising intrinsic transversality [96, Equations (19)&(20)]:

$$\mathrm{itr}[A, B](\bar{x}) = \min_{\substack{(v_1,v_2)\in\overline{N}_{A,B}(\bar{x}) \\ \|v_1\|+\|v_2\|=1}} \|v_1 + v_2\|, \tag{6.78}$$

$$\mathrm{itr}_c[A, B](\bar{x}) = \min_{\substack{(v_1,v_2)\in\overline{N}^c_{A,B}(\bar{x}) \\ \|v_1\|+\|v_2\|=1}} \|v_1 + v_2\|, \tag{6.79}$$

with the convention that the minimum over the empty set equals 1. Note that the minima in (6.78) and (6.79) are attainable thanks to the compactness of the constraint sets under the minima.

Remark 6.3.3. One can formulate a similar representation for $\mathrm{itr}_w[A, B](\bar{x})$ using another cone $\overline{N}^w_{A,B}(\bar{x})$ of 'weak' relative limiting normals satisfying

$$\overline{N}^c_{A,B}(\bar{x}) \subset \overline{N}^w_{A,B}(\bar{x}) \subset \overline{N}_{A,B}(\bar{x}). \tag{6.80}$$

The objective of this section consists in 1) giving a geometric counterpart of the analytic condition under which Theorem 6.1.13(ii) is valid - see Lemma 6.3.5; 2) deducing new characterisations of intrinsic transversality in terms of relative limiting normals - see Theorem 6.3.6; and 3) giving further insight into the quantitative results established in Section 6.1 when specialized to the Euclidean space setting - see Theorem 6.3.7. As by-products, we address the following questions raised by Kruger [96] regarding the cone objects appearing in (6.80) - see Remarks 6.3.8 and 6.3.9.

Question 5. *[96, question 4, page 140] What is the relationship between the two cones $\overline{N}_{A,B}(\bar{x})$ and $\overline{N}^c_{A,B}(\bar{x})$?*

Question 6. *[96, question 5, page 140] How can $\overline{N}^w_{A,B}(\bar{x})$ be used in characterising intrinsic transversality?*

Recall from Theorem 6.1.13 that if $\mathrm{itr}_c[A, B](\bar{x}) \geq 1/\sqrt{2}$, then the intrinsic transversality property is satisfied with a quantitative constant at least $1/\sqrt{2}$. This indicates that the other case, i.e.,

$$\mathrm{itr}_c[A, B](\bar{x}) < 1/\sqrt{2}, \tag{6.81}$$

is of the main interest in studying the property.

The analysis in this section is tuned to the scenario of (6.81). To begin, let us define the set:

$$C := \{(v_1, v_2) \in X \times X \mid \langle v_1, v_2\rangle < 0\}. \tag{6.82}$$

Proposition 6.3.4. *The set C defined above is a cone in $X \times X$. Moreover, if $(v_1, v_2) \in C$, then $(t_1 v_1, t_2 v_2) \in C$ for all $t_1, t_2 > 0$.*

Proof. The proof is straightforward from the definition of C. □

The following result explains the role of C in the analysis of intrinsic transversality in this chapter.

For convenience, we also define the set

$$S := \{(v_1, v_2) \in X \times X \mid \|v_1\| + \|v_2\| = 1\},$$

and note that $\overline{N}^c_{A,B}(\bar{x}) \cap S$ is the feasible set under the minimization in (6.79).

Lemma 6.3.5. *Let A, B be closed sets and $\bar{x} \in A \cap B$. Then*

$$C \cap \overline{N}^c_{A,B}(\bar{x}) \cap S \neq \emptyset \quad \Leftrightarrow \quad \text{itr}_c[A,B](\bar{x}) < 1/\sqrt{2}.$$

Proof. ($\Rightarrow$) We suppose $C \cap \overline{N}^c_{A,B}(\bar{x}) \cap S \neq \emptyset$ and prove that $\text{itr}_c[A,B](\bar{x}) < 1/\sqrt{2}$. Take a pair $(v_1, v_2) \in C \cap \overline{N}^c_{A,B}(\bar{x}) \cap S$. We note that $v_1 \neq 0$ and $v_2 \neq 0$ as $(v_1, v_2) \in C$, and set $v_1' := \dfrac{v_1}{2\,\|v_1\|}$ and $v_2' := \dfrac{v_2}{2\,\|v_2\|}$. It holds $(v_1', v_2') \in S$ since $\|v_1'\| + \|v_2'\| = 1/2 + 1/2 = 1$. Thanks to Propositions 6.3.2 and 6.3.4, it also holds $(v_1', v_2') \in C \cap \overline{N}^c_{A,B}(\bar{x})$. Then we have in view of (6.79) that

$$\text{itr}_c[A,B](\bar{x})^2 = \min_{(v_1,v_2)\in \overline{N}^c_{A,B}(\bar{x})\cap S} \|v_1 + v_2\|^2 \leq \|v_1' + v_2'\|^2 = 1/2 + 2\,\langle v_1', v_2'\rangle < 1/2,$$

where the last inequality holds true due to (6.82) as $(v_1', v_2') \in C$. Hence $\text{itr}_c[A,B](\bar{x}) < 1/\sqrt{2}$.

($\Leftarrow$) We suppose $C \cap \overline{N}^c_{A,B}(\bar{x}) \cap S = \emptyset$ and prove that $\text{itr}_c[A,B](\bar{x}) \geq 1/\sqrt{2}$. If $\overline{N}^c_{A,B}(\bar{x}) \cap S = \emptyset$, then (6.79) yields $\text{itr}_c[A,B](\bar{x}) = 1 > 1/\sqrt{2}$. We consider the case $\overline{N}^c_{A,B}(\bar{x}) \cap S \neq \emptyset$. Take an arbitrary pair $(v_1, v_2) \in \overline{N}^c_{A,B}(\bar{x}) \cap S$. Then by the assumption, it holds $(v_1, v_2) \notin C$, i.e., $\langle v_1, v_2\rangle \geq 0$. This implies

$$\|v_1 + v_2\|^2 = \|v_1\|^2 + \|v_2\|^2 + 2\,\langle v_1, v_2\rangle \geq \|v_1\|^2 + \|v_2\|^2 \geq \frac{1}{2}\left(\|v_1\| + \|v_2\|\right)^2 = 1/2.$$

We have just proved that

$$\|v_1 + v_2\| \geq 1/\sqrt{2} \quad \text{for all } (v_1, v_2) \in \overline{N}^c_{A,B}(\bar{x}) \cap S.$$

This together with the equality (6.79) implies that $\text{itr}_c[A,B](\bar{x}) \geq 1/\sqrt{2}$.

□

Lemma 6.3.5 provides an intuitive geometric counterpart of the quite mysterious analytic condition (6.81). It also gives rise to the following characterisations of intrinsic transversality.

Theorem 6.3.6 (refined characterisations of intrinsic transversality)**.** *Let A, B be closed sets and $\bar{x} \in A \cap B$. Then the following statements are equivalent.*

(i) *$\{A, B\}$ is intrinsically transversal at $\bar{x}$.*

(ii) *There exists a number $\alpha \in]0, 1[$ such that*

$$\|v_1 + v_2\| > \alpha \quad \text{for all} \quad (v_1, v_2) \in C \cap \overline{N}^c_{A,B}(\bar{x}) \cap S. \tag{6.83}$$

(iii) $\left\{v \in X \mid (v, -v) \in C \cap \overline{N}^c_{A,B}(\bar{x})\right\} \subset \{0\}$.

Proof. (i) $\Rightarrow$ (ii). We suppose that $\{A, B\}$ is intrinsically transversal at $\bar{x}$. This is equivalent to $\text{itr}_c[A,B](\bar{x}) > 0$ in view of Theorem 6.1.13. If $\text{itr}_c[A,B](\bar{x}) \geq 1/\sqrt{2}$, then by Lemma 6.3.5 the intersection $C \cap \overline{N}^c_{A,B}(\bar{x}) \cap S$ is empty, and hence the conclusion is trivial. We analyse the case $\text{itr}_c[A,B](\bar{x}) < 1/\sqrt{2}$. Take any α satisfying $0 < \alpha < \text{itr}_c[A,B](\bar{x})$. Then in view of (6.79) we have that

$$\|v_1 + v_2\| \geq \text{itr}_c[A,B](\bar{x}) > \alpha \quad \text{for all} \quad (v_1, v_2) \in \overline{N}^c_{A,B}(\bar{x}) \cap S,$$

which obviously implies (6.83).

(ii) $\Rightarrow$ (iii). Suppose that (iii) is violated, i.e., there exists $v \neq 0$ such that $(v, -v) \in C \cap \overline{N}^c_{A,B}(\bar{x})$. Then the pair $(v_1, v_2) := \left(\dfrac{v}{2\|v\|}, -\dfrac{v}{2\|v\|}\right) \in C \cap \overline{N}^c_{A,B}(\bar{x}) \cap S$, but $\|v_1 + v_2\| = 0$. That is, (ii) is violated.

(iii) $\Rightarrow$ (i). Suppose that (i) is violated, i.e., $\text{itr}_c[A,B](\bar{x}) = 0$. By [96, Corollary 3, (i)$\Leftrightarrow$(ii)], there exists a pair of vectors $(v_1, v_2) \in \overline{N}^c_{A,B}(\bar{x}) \cap S$ satisfying $\|v_1 + v_2\| = 0$. The latter conditions trivially imply $v_1 = -v_2 \neq 0$ and $\langle v_1, v_2\rangle < 0$. Then the vector $v \neq 0$ satisfies $(v, -v) := (v_1, v_2) \in C \cap \overline{N}^c_{A,B}(\bar{x})$. That is, (iii) is violated.

□

An implication of Theorem 6.3.6 is that only pairs of relative limiting normals in C are relevant for characterising intrinsic transversality. This observation guides us to the following concrete relationship between the two cones $\overline{N}_{A,B}(\bar{x})$ and $\overline{N}^c_{A,B}(\bar{x})$.

Theorem 6.3.7. *Let A, B be closed sets and $\bar{x} \in A \cap B$. Then*

$$\overline{N}^c_{A,B}(\bar{x}) \cap C = \overline{N}_{A,B}(\bar{x}) \cap C. \tag{6.84}$$

Proof. It is known by [96, Proposition 2(ii)] that $\overline{N}^c_{A,B}(\bar{x}) \subset \overline{N}_{A,B}(\bar{x})$. Thus, it is sufficient to show that

$$\overline{N}_{A,B}(\bar{x}) \cap C \subset \overline{N}^c_{A,B}(\bar{x}) \cap C.$$

Let us take any $(v_1, v_2) \in \overline{N}_{A,B}(\bar{x}) \cap C$. Then by the definition of $\overline{N}_{A,B}(\bar{x})$, there exist sequences $(a_k) \subset A \setminus B$, $(b_k) \subset B \setminus A$, $(x_k) \subset X$ and $(v_{1k}), (v_{2k}) \subset X$ such that $a_k \to \bar{x}$, $b_k \to \bar{x}$, $x_k \to \bar{x}$, $v_{1k} \to v_1$, $v_{2k} \to v_2$ and

$$\begin{gathered} v_{1k} \in N_A(a_k),\ v_{2k} \in N_B(b_k) \quad (k = 1, 2, \ldots), \\ \frac{\|x_k - a_k\|}{\|x_k - b_k\|} \to 1,\ \frac{\langle v_{1k}, x_k - a_k\rangle}{\|v_{1k}\|\,\|x_k - a_k\|} \to 1,\ \frac{\langle v_{2k}, x_k - b_k\rangle}{\|v_{1k}\|\,\|x_k - b_k\|} \to 1. \end{gathered} \tag{6.85}$$

Note that $v_1 \neq 0$ and $v_2 \neq 0$ as $(v_1, v_2) \in C$. Then thanks to $v_{1k} \to v_1$ and $v_{2k} \to v_2$, we can assume that $v_{1k} \neq 0$ and $v_{2k} \neq 0$ for all $k \in \mathbb{N}$. In view of Remark 6.1.7, we can also assume that $x_k \neq a_k$ and $x_k \neq b_k$ for all $k \in \mathbb{N}$.

Since $(v_1, v_2) \in C$, it holds that $\langle v_1, v_2\rangle < 0$, equivalently,

$$\left\| \frac{v_1}{\|v_1\|} - \frac{v_2}{\|v_2\|} \right\| > \sqrt{2}. \tag{6.86}$$

To complete the proof, it suffices to prove that $(v_1, v_2) \in \overline{N}^c_{A,B}(\bar{x})$. For each $k = 1, 2, \ldots$, let us define:

$$m_k := \frac{a_k + b_k}{2},$$

$$x'_k := x_k - \frac{\langle b_k - a_k, x_k - m_k\rangle}{\|b_k - a_k\|^2}(b_k - a_k), \tag{6.87}$$

$$v'_{1k} := \frac{\|v_{1k}\|}{\|x'_k - a_k\|}(x'_k - a_k),\ v'_{2k} := \frac{\|v_{2k}\|}{\|x'_k - b_k\|}(x'_k - b_k). \tag{6.88}$$

All we need is to verify the following four conditions:

(i)
$$\|x'_k - a_k\| = \|x'_k - b_k\| \quad (k = 1, 2, \ldots); \tag{6.89}$$

(ii)
$$x'_k \to \bar{x},\ v'_{1k} \to v_1,\ v'_{2k} \to v_2 \quad (k = 1, 2, \ldots);$$

(iii)
$$\text{dist}\,(v'_{1k}, N_A(a_k)) \to 0,\ \text{dist}\,(v'_{2k}, N_B(b_k)) \to 0;$$

(iv)

$$\langle v'_{1k}, x'_k - a_k\rangle = \|v'_{1k}\|\|x'_k - a_k\|,\ \langle v'_{2k}, x'_k - b_k\rangle = \|v'_{2k}\|\|x'_k - b_k\|.$$

Condition (i). This follows from (6.87) since for each $k = 1, 2, \ldots$, we have that

$$\begin{aligned}\|x'_k - a_k\|^2 - \|x'_k - b_k\|^2 &= \|x_k - a_k\|^2 - \|x_k - b_k\|^2 - 2\frac{\langle b_k - a_k, x_k - m_k\rangle}{\|b_k - a_k\|^2}\langle x_k - a_k, b_k - a_k\rangle \\ &\quad + 2\frac{\langle b_k - a_k, x_k - m_k\rangle}{\|b_k - a_k\|^2}\langle x_k - b_k, b_k - a_k\rangle \\ &= \|x_k - a_k\|^2 - \|x_k - b_k\|^2 - 2\frac{\langle b_k - a_k, x_k - m_k\rangle}{\|b_k - a_k\|^2}\langle b_k - a_k, b_k - a_k\rangle \\ &= \|x_k - a_k\|^2 - \|x_k - b_k\|^2 - 2\langle b_k - a_k, x_k - m_k\rangle \\ &= \|x_k - a_k\|^2 - \|x_k - b_k\|^2 - \langle (x_k - a_k) - (x_k - b_k), (x_k - a_k) + (x_k - b_k)\rangle = 0.\end{aligned}$$

Condition (iv). We first infer from (i) and $a_k \neq b_k$ that $x'_k \neq a_k$ and $x'_k \neq b_k$. This ensures that the definitions of (v'_{1k}) and (v'_{2k}) by (6.88) are well defined. Then thanks to (6.88), we have that

$$\begin{aligned}\|v'_{1k}\| = \|v_{1k}\|,\ \|v'_{2k}\| = \|v_{2k}\|,\\ \langle v'_{1k}, x'_k - a_k\rangle = \|v_{1k}\|\|x'_k - a_k\| = \|v'_{1k}\|\|x'_k - a_k\|,\\ \langle v'_{2k}, x'_k - b_k\rangle = \|v_{2k}\|\|x'_k - b_k\| = \|v'_{2k}\|\|x'_k - b_k\|.\end{aligned}$$

Condition (iii). Since $v_{1k} \in N_A(a_k)$ and $v_{2k} \in N_A(b_k)$ $(k = 1, 2, \ldots)$, whenever condition (ii) has been verified, we have that

$$\begin{aligned}&\text{dist }(v'_{1k}, N_A(a_k)) \le \|v'_{1k} - v_{1k}\| \le \|v'_{1k} - v_1\| + \|v_{1k} - v_1\| \to 0,\\ &\text{dist }(v'_{2k}, N_B(b_k)) \le \|v'_{2k} - v_{2k}\| \le \|v'_{2k} - v_2\| + \|v_{2k} - v_2\| \to 0.\end{aligned}$$

Condition (ii). Since $x_k \to \bar{x}$, $a_k \to \bar{x}$ and $b_k \to \bar{x}$, it holds by (6.87) that

$$\|x'_k - x_k\| = \left\|\frac{\langle b_k - a_k, x - m_k\rangle}{\|b_k - a_k\|^2}(b_k - a_k)\right\| \le \|x_k - m_k\| = \left\|x_k - \frac{a_k + b_k}{2}\right\| \to 0.$$

Then

$$\|x'_k - \bar{x}\| \le \|x'_k - x_k\| + \|x_k - \bar{x}\| \to 0.$$

In the remainder of the proof, we show that $v'_{1k} \to v_1$ while the condition $v'_{2k} \to v_2$ is obtained in a similar manner. Since $v_{1k} \to v_1$, all we need is to show that $\|v'_{1k} - v_{1k}\| \to 0$. Note that by (6.88) it holds

$$\begin{aligned}\|v'_{1k} - v_{1k}\| &= \left\|\frac{\|v_{1k}\|}{\|x'_k - a_k\|}(x'_k - a_k) - v_{1k}\right\| = \|v_{1k}\|\left\|\frac{x'_k - a_k}{\|x'_k - a_k\|} - \frac{v_{1k}}{\|v_{1k}\|}\right\| \\ &\le \|v_{1k}\|\left(\left\|\frac{x_k - a_k}{\|x_k - a_k\|} - \frac{v_{1k}}{\|v_{1k}\|}\right\| + \left\|\frac{x'_k - a_k}{\|x'_k - a_k\|} - \frac{x_k - a_k}{\|x_k - a_k\|}\right\|\right).\end{aligned} \tag{6.90}$$

Note also that due to (6.85),

$$\left\|\frac{x_k - a_k}{\|x_k - a_k\|} - \frac{v_{1k}}{\|v_{1k}\|}\right\| = \sqrt{2 - 2\frac{\langle v_{1k}, x_k - a_k\rangle}{\|v_{1k}\|\,\|x_k - a_k\|}} \to 0. \tag{6.91}$$

In view of (6.90) and (6.91), in order to obtain $\|v'_{1k} - v_{1k}\| \to 0$, it suffices to prove that

$$\left\|\frac{x'_k - a_k}{\|x'_k - a_k\|} - \frac{x_k - a_k}{\|x_k - a_k\|}\right\| \to 0.$$

To proceed, let us take any number $\varepsilon > 0$ which can be arbitrarily small and show the existence of a natural $N \in \mathbb{N}$ such that

$$\left\| \frac{x_k' - a_k}{\|x_k' - a_k\|} - \frac{x_k - a_k}{\|x_k - a_k\|} \right\| < \varepsilon, \quad \forall k \geq N. \tag{6.92}$$

Choose a number $\varepsilon' > 0$ to satisfy

$$2\sqrt{\frac{2\varepsilon' - \varepsilon'^2}{4 - 6\varepsilon' + 3\varepsilon'^2}} < \varepsilon. \tag{6.93}$$

Such a number ε' exists since $\varepsilon > 0$ and $\lim_{t \downarrow 0} 2\sqrt{\frac{2t - t^2}{4 - 6t + 3t^2}} = 0$.
By the convergence conditions in (6.85), there exists a natural number $N \in \mathbb{N}$ such that $\forall k \geq N$,

$$1 - \varepsilon' < \frac{\|x_k - a_k\|}{\|x_k - b_k\|} < 1 + \varepsilon', \tag{6.94}$$

$$\frac{\langle v_{1k}, x_k - a_k \rangle}{\|v_{1k}\| \, \|x_k - a_k\|} > 1 - \varepsilon', \; \frac{\langle v_{2k}, x_k - b_k \rangle}{\|v_{1k}\| \, \|x_k - b_k\|} > 1 - \varepsilon'. \tag{6.95}$$

The estimates in (6.95) amount to

$$\left\| \frac{v_{1k}}{\|v_{1k}\|} - \frac{x_k - a_k}{\|x_k - a_k\|} \right\| < \sqrt{2\varepsilon'}, \; \left\| \frac{v_{2k}}{\|v_{2k}\|} - \frac{x_k - b_k}{\|x_k - b_k\|} \right\| < \sqrt{2\varepsilon'}. \tag{6.96}$$

In order to prove (6.92), we first note that

$$\langle x_k - x_k', x_k' - m_k \rangle = 0. \tag{6.97}$$

Indeed, by (6.87), it holds that

$$\langle x_k - x_k', x_k' - m_k \rangle = \frac{\langle b_k - a_k, x_k - m_k \rangle}{\|b_k - a_k\|^2} \langle b_k - a_k, x_k' - m_k \rangle,$$

from which (6.97) follows since

$$\begin{aligned} \langle b_k - a_k, x_k' - m_k \rangle &= \left\langle b_k - a_k, x_k - m_k - \frac{\langle b_k - a_k, x_k - m_k \rangle}{\|b_k - a_k\|^2}(b_k - a_k) \right\rangle \\ &= \langle b_k - a_k, x_k - m_k \rangle - \frac{\langle b_k - a_k, x_k - m_k \rangle}{\|b_k - a_k\|^2} \langle b_k - a_k, b_k - a_k \rangle = 0. \end{aligned}$$

Second, we show that $\forall k \geq N$,

$$\|x_k - x_k'\|^2 \leq \tfrac{2\varepsilon' - \varepsilon'^2}{4(1-\varepsilon')^2} \min\left\{ \|x_k - a_k\|^2, \|x_k - b_k\|^2 \right\}. \tag{6.98}$$

If $\|x_k - a_k\| \geq \|x_k - b_k\|$, then

$$\langle b_k - a_k, x_k - m_k \rangle \geq 0. \tag{6.99}$$

Note from (6.86) and (6.96) that

$$\begin{aligned} \left\| \frac{x_k - a_k}{\|x_k - a_k\|} - \frac{x_k - b_k}{\|x_k - b_k\|} \right\| &\geq \left\| \frac{v_{1k}}{\|v_{1k}\|} - \frac{v_{2k}}{\|v_{2k}\|} \right\| - \left\| \frac{v_{1k}}{\|v_{1k}\|} - \frac{x_k - a_k}{\|x_k - a_k\|} \right\| - \left\| \frac{v_{2k}}{\|v_{2k}\|} - \frac{x_k - b_k}{\|x_k - b_k\|} \right\| \\ &> \sqrt{2} - 2\sqrt{2\varepsilon'}. \end{aligned}$$

This combining with (6.94) yields that

$$\langle x_k - a_k, x_k - b_k \rangle < 4\left(\sqrt{\varepsilon'} - \varepsilon'\right) \|x_k - a_k\| \|x_k - b_k\| < 4\left(\sqrt{\varepsilon'} - \varepsilon'\right)(1 + \varepsilon') \|x_k - b_k\|^2.$$

Then

$$\begin{aligned}\langle b_k - a_k, x_k - b_k \rangle &= \langle b_k - x_k, x_k - b_k \rangle + \langle x_k - a_k, x_k - b_k \rangle \\ &= -\|x_k - b_k\|^2 + \langle x_k - a_k, x_k - b_k \rangle \\ &< -\|x_k - b_k\|^2 + 4\left(\sqrt{\varepsilon'} - \varepsilon'\right)(1+\varepsilon')\|x_k - b_k\|^2 \\ &= -\left(1 - 4\left(\sqrt{\varepsilon'} - \varepsilon'\right)(1+\varepsilon')\right)\|x_k - b_k\|^2 < 0.\end{aligned} \tag{6.100}$$

From (6.87), (6.99) and (6.100) we have that

$$\begin{aligned}\|x'_k - b_k\|^2 - \|x_k - b_k\|^2 &= \|x'_k - x_k\|^2 + 2\langle x'_k - x_k, x_k - b_k \rangle \\ &= \|x'_k - x_k\|^2 - 2\frac{\langle b_k - a_k, x_k - m_k \rangle}{\|b_k - a_k\|^2}\langle b_k - a_k, x_k - b_k \rangle \\ &\geq \|x'_k - x_k\|^2.\end{aligned} \tag{6.101}$$

By (6.89) and (6.97) we get that

$$\begin{aligned}\|x_k - a_k\|^2 + \|x_k - b_k\|^2 &= 2\|x_k - x'_k\|^2 + \|x'_k - a_k\|^2 + \|x'_k - b_k\|^2 + 2\langle x_k - x'_k, 2x'_k - (a_k + b_k)\rangle \\ &= 2\|x_k - x'_k\|^2 + 2\|x'_k - b_k\|^2 + 4\langle x_k - x'_k, x'_k - m_k \rangle \\ &= 2\|x_k - x'_k\|^2 + 2\|x'_k - b_k\|^2.\end{aligned}$$

This together with (6.94) and (6.101) yields

$$\begin{aligned}2\|x_k - x'_k\|^2 &= \|x_k - a_k\|^2 + \|x_k - b_k\|^2 - 2\|x'_k - b_k\|^2 \\ &\leq (1+\varepsilon')^2\|x_k - b_k\|^2 + \|x_k - b_k\|^2 - 2\|x'_k - b_k\|^2 \\ &\leq (1+\varepsilon')^2\|x_k - b_k\|^2 + \|x_k - b_k\|^2 - 2\left(\|x_k - b_k\|^2 + \|x'_k - x_k\|^2\right).\end{aligned}$$

Hence

$$4\|x_k - x'_k\|^2 \leq \left(2\varepsilon' + \varepsilon'^2\right)\|x_k - b_k\|^2 = \left(2\varepsilon' + \varepsilon'^2\right)\min\{\|x_k - a_k\|^2, \|x_k - b_k\|^2\} \tag{6.102}$$

since $\|x_k - a_k\| \geq \|x_k - b_k\|$ in this case.

By a similar argument, if $\|x_k - a_k\| \leq \|x_k - b_k\|$, then

$$\langle b_k - a_k, x_k - m_k \rangle \leq 0, \ \langle b_k - a_k, x_k - a_k \rangle \geq 0.$$

Thus

$$\begin{aligned}\|x'_k - a_k\|^2 - \|x_k - a_k\|^2 &= \|x'_k - x_k\|^2 + 2\langle x'_k - x_k, x_k - a_k \rangle \\ &= \|x'_k - x_k\|^2 - 2\frac{\langle b_k - a_k, x_k - m_k \rangle}{\|b_k - a_k\|^2}\langle b_k - a_k, x_k - a_k \rangle \\ &\geq \|x'_k - x_k\|^2.\end{aligned} \tag{6.103}$$

By (6.89) and (6.97) we get that

$$\|x_k - a_k\|^2 + \|x_k - b_k\|^2 = 2\|x_k - x'_k\|^2 + 2\|x'_k - a_k\|^2,$$

which together with (6.94) and (6.103) yields that

$$2\|x_k - x'_k\|^2 \leq \|x_k - a_k\|^2 + \tfrac{1}{(1-\varepsilon')^2}\|x_k - a_k\|^2 - 2\left(\|x_k - a_k\|^2 + \|x'_k - x_k\|^2\right).$$

Equivalently,

$$4\|x_k - x'_k\|^2 \leq \tfrac{2\varepsilon' - \varepsilon'^2}{(1-\varepsilon')^2}\|x_k - a_k\|^2 = \tfrac{2\varepsilon' - \varepsilon'^2}{(1-\varepsilon')^2}\min\{\|x_k - a_k\|^2, \|x_k - b_k\|^2\} \tag{6.104}$$

since $\|x_k - a_k\| \leq \|x_k - b_k\|$ in this case.

Combining (6.102) and (6.104) and noting that $2\varepsilon' + \varepsilon'^2 < \frac{2\varepsilon' - \varepsilon'^2}{(1-\varepsilon')^2}$, we obtain (6.98) as claimed.

Third, we show that $\forall k \geq N$

$$\|x'_k - a_k\|^2 \geq \tfrac{4-6\varepsilon'+3\varepsilon'^2}{2\varepsilon'-\varepsilon'^2}\|x_k - x'_k\|^2. \tag{6.105}$$

Indeed, if $\|x_k - a_k\| \leq \|x_k - b_k\|$, then the use of (6.103) and (6.104) yields (6.105):

$$\begin{aligned}\|x'_k - a_k\|^2 &\geq \|x_k - a_k\|^2 + \|x_k - x'_k\|^2 \\ &\geq \tfrac{4(1-\varepsilon')^2)}{2\varepsilon'-\varepsilon'^2}\|x_k - x'_k\|^2 + \|x_k - x'_k\|^2 = \tfrac{4-6\varepsilon'+3\varepsilon'^2}{2\varepsilon'-\varepsilon'^2}\|x_k - x'_k\|^2.\end{aligned}$$

Otherwise, i.e., if $\|x_k - a_k\| \geq \|x_k - b_k\|$, then the use of (6.89), (6.101) and (6.98) successively implies that

$$\begin{aligned}\|x'_k - a_k\|^2 = \|x'_k - b_k\|^2 \geq \|x_k - b_k\|^2 + \|x_k - x'_k\|^2 \\ \geq \tfrac{4}{2\varepsilon'+\varepsilon'^2}\|x_k - x'_k\|^2 + \|x'_k - x_k\|^2 = \tfrac{4+2\varepsilon'+\varepsilon'^2}{2\varepsilon'+\varepsilon'^2}\|x_k - x'_k\|^2,\end{aligned}$$

which also yields (6.105) since $\dfrac{4+2\varepsilon'+\varepsilon'^2}{2\varepsilon'+\varepsilon'^2} > \dfrac{4-6\varepsilon'+3\varepsilon'^2}{2\varepsilon'-\varepsilon'^2}$. Hence (6.105) has been proved.

Fourth, we show that $\forall k \geq N$

$$\left\|\frac{x'_k - a_k}{\|x'_k - a_k\|} - \frac{x_k - a_k}{\|x_k - a_k\|}\right\| \leq 2\frac{\|x_k - x'_k\|}{\|x'_k - a_k\|}. \tag{6.106}$$

Indeed,

$$\begin{aligned}\left\|\frac{x'_k - a_k}{\|x'_k - a_k\|} - \frac{x_k - a_k}{\|x_k - a_k\|}\right\| &\leq \left\|\frac{x'_k - a_k}{\|x'_k - a_k\|} - \frac{x_k - a_k}{\|x'_k - a_k\|}\right\| + \left\|\frac{x_k - a_k}{\|x'_k - a_k\|} - \frac{x_k - a_k}{\|x_k - a_k\|}\right\| \\ &\leq \frac{\|x_k - x'_k\|}{\|x'_k - a_k\|} + \left|\frac{\|x_k - a_k\|}{\|x'_k - a_k\|} - 1\right|.\end{aligned}$$

If $\|x_k - a_k\| \geq \|x'_k - a_k\|$, then (6.106) holds true since

$$\left|\frac{\|x_k - a_k\|}{\|x'_k - a_k\|} - 1\right| = \frac{\|x_k - a_k\|}{\|x'_k - a_k\|} - 1 \leq \frac{\|x_k - x'_k\| + \|x'_k - a_k\|}{\|x'_k - a_k\|} - 1 = \frac{\|x_k - x'_k\|}{\|x'_k - a_k\|}.$$

Otherwise, i.e., if $\|x_k - a_k\| < \|x'_k - a_k\|$, (6.106) also holds true since

$$\left|\frac{\|x_k - a_k\|}{\|x'_k - a_k\|} - 1\right| = 1 - \frac{\|x_k - a_k\|}{\|x'_k - a_k\|} \leq 1 - \frac{\|x'_k - a_k\| - \|x_k - x'_k\|}{\|x'_k - a_k\|} = \frac{\|x_k - x'_k\|}{\|x'_k - a_k\|}.$$

Hence (6.106) has been proved.

Finally, a combination of (6.105), (6.106) and (6.93) yields that

$$\left\|\frac{x'_k - a_k}{\|x'_k - a_k\|} - \frac{x_k - a_k}{\|x_k - a_k\|}\right\| \leq 2\frac{\|x_k - x'_k\|}{\|x'_k - a_k\|} \leq 2\sqrt{\frac{2\varepsilon' - \varepsilon'^2}{4 - 6\varepsilon' + 3\varepsilon'^2}} < \varepsilon, \quad \forall k \geq N,$$

which is (6.92) and hence the proof is complete.

□

To this end, we have sufficient information for addressing Questions 5 and 6.

Remark 6.3.8 (answer to Question 5). In view of Theorem 6.3.7, the two cones $\overline{N}^c_{A,B}(\bar{x})$ and $\overline{N}_{A,B}(\bar{x})$ are equal when restricted to the cone C given by (6.82). Recall that elements of C are sufficient for characterising intrinsic transversality in view of Theorem 6.3.6. Their equality outside of C remains an open question.

Remark 6.3.9 (answer to Question 6). The combination of (6.80) and (6.84) yields that

$$\overline{N}^{c}_{A,B}(\bar{x}) \cap C = \overline{N}^{w}_{A,B}(\bar{x}) \cap C = \overline{N}_{A,B}(\bar{x}) \cap C. \tag{6.107}$$

In view of Theorem 6.3.6, this particularly implies that the three cones have equivalent roles in characterising intrinsic transversality.

Remark 6.3.10. Corresponding to $\mathrm{itr}_w[A,B](\bar{x})$, the cone $\overline{N}^{w}_{A,B}(\bar{x})$ would not be essential for the analysis in this section due to (6.80) and (6.107). This explains why we chose not to introduce its definition for the sake of brevity in terms of terminology. It occasionally comes up in the discussion only for the purpose of addressing Question 6.

Remark 6.3.11. Thanks to (6.78) and (6.79), Theorem 6.1.13(ii) in the Euclidean space setting can be deduced from Theorem 6.3.7. But the inverse implication is not trivial since the minimal objective values in (6.78) and (6.79) being equal does not tell much about the relationship between the corresponding feasible sets. Instead, Theorem 6.3.7 complements Theorem 6.1.13(ii) and further clarifies the characterisations of intrinsic transversality in terms of relative limiting normals. It is worth noting that the result of Theorem 6.3.7 is also inspired by the importance of the cones themselves, see [96, page 123].

We conclude this chapter with a list of characterisations of intrinsic transversality in the Euclidean space setting.

Proposition 6.3.12 (characterisations of intrinsic transversality in Euclidean spaces)**.** *Let A, B be closed sets and $\bar{x} \in A \cap B$. The following conditions are equivalent:*

(i) *$\{A, B\}$ is intrinsically transversal at $\bar{x}$;*

(ii) $\mathrm{itr}_c[A,B](\bar{x}) > 0$;

(iii) *there exists a number $\alpha \in]0,1[$ such that $\|v_1 + v_2\| > \alpha$ for all $(v_1, v_2) \in C \cap \overline{N}^{c}_{A,B}(\bar{x}) \cap S$;*

(iv) $\left\{v \in X \mid (v, -v) \in C \cap \overline{N}^{c}_{A,B}(\bar{x})\right\} \subset \{0\}$;

(v) *there exists a number $\alpha \in]0,1[$ such that $\|v_1 + v_2\| > \alpha$ for all $(v_1, v_2) \in \overline{N}^{c}_{A,B}(\bar{x}) \cap S$;*

(vi) $\left\{v \in X \mid (v, -v) \in \overline{N}^{c}_{A,B}(\bar{x})\right\} \subset \{0\}$;

(vii) $\mathrm{itr}[A,B](\bar{x}) > 0$;

(viii) *there exists a number $\alpha \in]0,1[$ such that $\|v_1 + v_2\| > \alpha$ for all $(v_1, v_2) \in C \cap \overline{N}_{A,B}(\bar{x}) \cap S$;*

(ix) $\left\{v \in X \mid (v, -v) \in C \cap \overline{N}_{A,B}(\bar{x})\right\} \subset \{0\}$;

(x) *there exists a number $\alpha \in]0,1[$ such that $\|v_1 + v_2\| > \alpha$ for all $(v_1, v_2) \in \overline{N}_{A,B}(\bar{x}) \cap S$;*

(xi) $\left\{v \in X \mid (v, -v) \in \overline{N}_{A,B}(\bar{x})\right\} \subset \{0\}$;

(xii) $\mathrm{itr}_w[A,B](\bar{x}) > 0$;

(xiii) *there exists a number $\alpha \in]0,1[$ such that $\|v_1 + v_2\| > \alpha$ for all $(v_1, v_2) \in C \cap \overline{N}^{w}_{A,B}(\bar{x}) \cap S$;*

(xiv) $\left\{v \in X \mid (v, -v) \in C \cap \overline{N}^{w}_{A,B}(\bar{x})\right\} \subset \{0\}$;

(xv) *there exists a number $\alpha \in]0,1[$ such that $\|v_1 + v_2\| > \alpha$ for all $(v_1, v_2) \in \overline{N}^{w}_{A,B}(\bar{x}) \cap S$;*

(xvi) $\left\{v \in X \mid (v, -v) \in \overline{N}^{w}_{A,B}(\bar{x})\right\} \subset \{0\}$;

(xvii) $\mathrm{itr}_p[A,B](\bar{x}) > 0$;

(xviii) *$\{A, B\}$ satisfies property $(\mathcal{P})$ at $\bar{x}$.*

If, in addition, the sets A, B are convex, then the following item can be added to the above list:

(xix) $\{A, B\}$ *is subtransversal at* $\bar{x}$*.*

Proof. (i) $\Leftrightarrow$ (ii) follows from Theorem 6.1.13. (i) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (iv) follow from Theorem 6.3.6. (i) $\Leftrightarrow$ (v) $\Leftrightarrow$ (vi) follow from Corollary 3 of [96]. (i) $\Leftrightarrow$ (vii) follows from Proposition 6.1.8(iii). (i) $\Leftrightarrow$ (viii) $\Leftrightarrow$ (ix) follow from Theorem 6.3.6 and Theorem 6.3.7. (i) $\Leftrightarrow$ (x) $\Leftrightarrow$ (xi) follow from Theorem 5 of [96]. (i) $\Leftrightarrow$ (xii) $\Leftrightarrow$ (xiii) $\Leftrightarrow$ (xiv) $\Leftrightarrow$ (xv) $\Leftrightarrow$ (xvi) are consequences of the previous equivalences in view of Proposition 6.1.9(i) and the inclusions in (6.80). (i) $\Leftrightarrow$ (xvii) $\Leftrightarrow$ (xviii) follow from Theorem 6.2.9 and Proposition 6.2.2, respectively. (i) $\Leftrightarrow$ (xix) in the convex setting follows from Proposition 6.1.8(iv).

□

Part II

Generalised Convexity

Chapter 7

Robustly Quasiconvex Functions

Our aim in the chapter is to establish the first–order characterisations for the robust quasiconvexity of lower semicontinuous functions using means of Fréchet subdifferential in Asplund spaces. First, some existing results regarding to the properties of subdifferential operators of convex, quasiconvex functions are recalled in Section 7.1, where the definitions and some basic results are given as well. Besides, necessary and sufficient first-order conditions for a lower semicontinuous function to be quasiconvex are reconsidered. Those characterisations moreover could be used to characterise the Asplund property of the given space. Second, two criteria for the robust quasiconvexity of lower semicontinuous functions in Asplund spaces are obtained by using Fréchet subdifferentials in Section 3. Each criterion corresponds to each type of analogous conditions for quasiconvexity. The first one is based on the zero and first order condition for quasiconvexity (see Theorem 7.1.5(b) in Section 2). It extends [11, Proposition 5.3] from finite dimensional spaces to Asplund spaces. Moreover, its proof also overcomes a glitch in the proof of the sufficient condition of [11, Proposition 5.3]. The second criterion is totally new. It is settled from the equivalence of the quasiconvexity of lower semicontinuous functions and the quasimonotonicity of their subdifferential operators (see Theorem 7.1.5(c) in Section 2).

7.1 Basic results

An operator A is monotone if for all $x, y \in \text{dom}A$, one has $\langle x^* - y^*, x - y\rangle \geq 0$ with $x^* \in A(x), y^* \in A(y)$. It is well-known that when φ is convex, the operator $\partial^F \varphi$ is monotone [140]. The inverse implication also holds in Asplund space [120, Theorem 3.56]; but it is not true in general Banach spaces. The reader is referred to the proof of the reverse implication in [163, Theorem 2.4] for a counter-example.

Let us recall some notions of generalised convex functions.

Definition 7.1.1. A function $\varphi : X \to \overline{\mathbb{R}}$ is

(i) **quasiconvex** if

$$\forall x, y \in X, \lambda \in]0,1[, \quad f(\lambda x + (1-\lambda)y) \leq \max\{f(x), f(y)\}. \tag{7.1}$$

(ii) **α-robustly quasiconvex** with $\alpha > 0$ if, for every $v^* \in \alpha\mathbb{B}^*$, the function $\varphi_{v^*} : x \mapsto \varphi(x) + \langle v^*, x\rangle$ is quasiconvex.

Clearly, φ is α-robustly quasiconvex if and only if the function φ_{v^*} is quasiconvex for all $v^* \in X^*$ such that $\|v^*\| < \alpha$.

Definition 7.1.2. An operator $A : X \rightrightarrows X^*$ is quasimonotone if for all $x, y \in X$ and $x^* \in A(x), y^* \in A(y)$ we have $\min\{\langle x^*, y - x\rangle, \langle y^*, x - y\rangle\} \leq 0$.

Below, we give a short proof to clarify the dual characterisation in terms of Fréchet subdifferentials in Asplund spaces. Our proof relies on the proof scheme of [4] and the following approximate mean value theorem [120, Theorem 3.49].

Theorem 7.1.3. *Let X be an Asplund space and $\varphi : X \to \overline{\mathbb{R}}$ be a proper lower semicontinuous function finite at two given points $a \neq b$. Consider any point $c \in [a, b)$ at which the function*

$$\psi(x) := \varphi(x) - \frac{\varphi(b) - \varphi(a)}{\|a - b\|}\|x - a\|$$

attains its minimum on $[a, b]$; such a point always exists. Then, there are sequences $x_k \xrightarrow{\varphi} c$ and $x_k^* \in \partial^F \varphi(x_k)$ satisfying

$$\liminf_{k\to\infty}\langle x_k^*, b - x_k\rangle \geq \frac{\varphi(b) - \varphi(a)}{\|a - b\|}\|b - c\|, \tag{7.2}$$

$$\liminf_{k\to\infty}\langle x_k^*, b - a\rangle \geq \varphi(b) - \varphi(a). \tag{7.3}$$

Moreover, when $c \neq a$ one has

$$\lim_{k\to\infty}\langle x_k^*, b - a\rangle = \varphi(b) - \varphi(a). \tag{7.4}$$

Theorem 7.1.3 allows us to deduce the following three-points lemma which is similar to [5, Lemma 3.1]. The proof for Lemma 7.1.4 is identical to the proof for [5, Lemma 3.1].

Lemma 7.1.4. *Let $\varphi : X \to \overline{\mathbb{R}}$ be a proper, lower semicontinuous function on an Asplund space X. Let $u, v, w \in X$ such that $v \in [u, w]$, $\varphi(v) > \varphi(u)$ and $\lambda > 0$. Then, there are $\bar{x} \in \operatorname{dom}\varphi$ and $\bar{x}^* \in \partial^F \varphi(\bar{x})$ such that*

$$\bar{x} \in \mathbb{B}_\lambda([u, v]) \quad \textit{and} \quad \langle \bar{x}^*, w - \bar{x}\rangle > 0,$$

where

$$\mathbb{B}_\lambda([u, v]) := \{x \in X : \exists y \in [u, v] \textit{ such that } \|x - y\| < \lambda\}.$$

We are in position to establish characterisations of quasiconvexity in terms of Fréchet subdifferentials in Asplund spaces.

Theorem 7.1.5. *Let $\varphi : X \to \overline{\mathbb{R}}$ be a proper lower semicontinuous function on an Asplund space X. The following statements are equivalent*

(a) *φ is quasiconvex;*

(b) *If there are $x, y \in X$ such that $\varphi(y) \leq \varphi(x)$, then $\langle x^*, y - x\rangle \leq 0$ for all $x^* \in \widehat{\partial}\varphi(x)$.*

(c) *$\partial^F \varphi$ is quasimonotone.*

Proof. (a)⇒(b) Assume that $x, y \in X$, $\varphi(x) \geq \varphi(y)$, and $x^* \in \partial^F \varphi(x)$. Consider $S_x := \{u \in X : \varphi(u) \leq \varphi(x)\}$. Since φ is quasiconvex, then S_x is a convex set. Thus, we have the function $f := \delta_{S_x} + \varphi(x)$ is convex, where δ_{S_x} is equal to 0 for $u \in S_x$ and to ∞ otherwise. On the other hand, $f(x) = \varphi(x)$ and $f(u) \geq \varphi(u)$ for all $u \in X$, thus $\partial^F \varphi(x) \subset \partial^F f(x)$. By the definition of convex subdifferential, since $x^* \in \widehat{\partial}\varphi(x) \subset \partial^F f(x)$, we have $\langle x^*, y - x\rangle \leq 0$.

(b)⇒(c) Assume that there are $x, y \in X$ and $x^* \in \partial^F \varphi(x)$, $y^* \in \partial^F \varphi(y)$ such that $\langle x^*, x - y\rangle < 0$ and $\langle y^*, x - y\rangle > 0$. Then, by (b), $\varphi(x) < \varphi(y)$ and $\varphi(y) < \varphi(x)$, which is a contradiction.

(c)⇒(a) By using Lemma 7.1.4, the proof of this assertion is similar to one in [4, Theorem 4.1]. □

Remark 7.1.6. Observe that the implications $(a) \Rightarrow (b)$ and $(b) \Rightarrow (c)$ hold in Banach spaces while $(c) \Rightarrow (a)$ only holds in Asplund spaces. In fact, the equivalence of these statements actually can characterise the Asplund property in the sense that if X is not an Asplund space, then there is a function φ whose Fréchet subdiferential satisfies (b) and (c) but is not quasiconvex. Such a function φ can be found in [163, Theorem 2.4].

7.2 Characterisations of Robustly Quasiconvex Functions

Here we show that this conclusion in [11, Proposition 5.3] is still correct not only when f is assumed just to be lower semicontinuous, but also when X is only assumed to be an Asplund space. To derive this generalisation, we need the following lemmas, revealing that quasiconvex functions have certain nice properties which resemble those of convex functions.

Lemma 7.2.1. *If $\varphi : X \to \overline{\mathbb{R}}$ is a quasiconvex and lowersemicontinuous function, and $u, v \in X$ are such that $\varphi(v) \geq \varphi(u)$ then*

$$\lim_{t \downarrow 0} \varphi(v + t(u - v)) = \varphi(v). \tag{7.5}$$

Proof. Suppose that $u, v \in X$ and that $\varphi(v) \geq \varphi(u)$. Since φ is quasiconvex, for all $t \in]0, 1[$, we have $\varphi(v + t(u - v)) \leq \max\{\varphi(v), \varphi(u)\} = \varphi(v)$. It follows that $\limsup_{t \downarrow 0} \varphi(v + t(u - v)) \leq \varphi(v)$. Combining the latter with the lower semicontinuity of φ we get (7.5). □

Lemma 7.2.2. *Let $\varphi : X \to \overline{\mathbb{R}}$ be a quasiconvex function and $u, v, w \in X$ such that $v \in]u, w[, \varphi(u) \leq \varphi(w)$. Suppose that there exist $v^* \in X^*$ and $z \in]u, v[$ such that $\varphi_{v^*}(z) > \max\{\varphi_{v^*}(u), \varphi_{v^*}(w)\}$. Then*

$$\varphi(u) < \varphi(z) \leq \varphi(v) \leq \varphi(w). \tag{7.6}$$

Proof. Since $z \in]u, v[\subset]u, w[, \varphi(u) \leq \varphi(w)$ and φ is quasiconvex we have $\varphi(z) \leq \max\{\varphi(u), \varphi(w)\} = \varphi(w)$. Hence, the latter and the inequality $\varphi_{v^*}(z) > \varphi_{v^*}(w)$ implies that $\langle v^*, z\rangle > \langle v^*, w\rangle$. As $\langle v^*, z - w\rangle > 0$ and $z = \lambda u + (1 - \lambda)w$, we have $\langle v^*, z - w\rangle = \lambda \langle v^*, u - w\rangle > 0$ with $\lambda \in]0, 1[$, hence $\langle v^*, u - w\rangle > 0$. Then, $\langle v^*, z - u\rangle = (1-\lambda) \langle v^*, w - u\rangle < 0$ so $\langle v^*, z\rangle < \langle v^*, u\rangle$. Therefore, the inequality $\varphi_{v^*}(u) < \varphi_{v^*}(z)$ yields $\varphi(u) < \varphi(z)$. Since $z \in]u, v[$ and $v \in]z, w[$, we deduce $\varphi(z) \leq \varphi(v) \leq \varphi(w)$ from the latter inequality and the quasiconvexity of φ. Hence, (7.6) holds. □

Lemma 7.2.3. *Let $\varphi : X \to \overline{\mathbb{R}}$ be a quasiconvex, proper, and lower semicontinuous function, and $v^* \in X^*$. If φ_{v^*} is not quasiconvex then there exist $u, v, w \in X$ such that $v \in]u, w[$ and*

$$\varphi(w) \geq \varphi(v) > \varphi(u), \tag{7.7}$$

$$\varphi_{v^*}(v) > \max\{\varphi_{v^*}(u), \varphi_{v^*}(w)\}, \tag{7.8}$$

$$\forall \gamma > 0, \exists v_\gamma \in \mathbb{B}_\gamma(v) \cap]v, w[: \varphi_{v^*}(v) > \varphi_{v^*}(v_\gamma). \tag{7.9}$$

Proof. Since φ_{v^*} is not quasiconvex, there exist $u, w \in X$ such that $u \neq w, \varphi(u) \leq \varphi(w)$ and $v_0 \in]u, w[$ such that $\varphi_{v^*}(v_0) > \max\{\varphi_{v^*}(u), \varphi_{v^*}(w)\}$. Applying Lemma 7.2.1, we get $\lim_{t \downarrow 0} \varphi(w + t(u - w)) = \varphi(w)$, and so $\lim_{t \downarrow 0} \varphi_{v^*}(w + t(u - w)) = \varphi_{v^*}(w)$. Since $\varphi_{v^*}(w) < \varphi_{v^*}(v_0)$, there exists $t_0 \in]0, 1[$ such that

$$\varphi_{v^*}(w + t(u - w)) < \varphi_{v^*}(v_0), \quad \forall t \in]0, t_0[. \tag{7.10}$$

Consider the set

$$\mathscr{L} := \{z \in]u, w[: \varphi_{v^*}(z) \geq \varphi_{v^*}(v_0)\}.$$

Since we already have $v_0 \in \mathscr{L}$, then $\mathscr{L} \neq \emptyset$. For each $z \in \mathscr{L}$ we have $\|z-w\| \geq t_0\|u-w\|$ by (7.10). It follows that

$$r := \inf\{\|z-w\| : z \in \mathscr{L}\} \in [t_0\|u-w\|, \|u-w\|[\, \subset\,]0, \|u-w\|[,$$

$$v := w + r\frac{u-w}{\|u-w\|} \in\,]u,w[.$$

We will show that $v \in \mathscr{L}$ and so (7.8) holds. Suppose on the contrary that $v \notin \mathscr{L}$. Then $v_0 \in]u,v[$ and we get $\varphi(u) < \varphi(v_0) \leq \varphi(v) \leq \varphi(w)$ by Lemma 7.2.2. Applying Lemma 7.2.1, we get $\lim_{t\downarrow 0} \varphi(v+t(u-v)) = \varphi(v)$, and so $\lim_{t\downarrow 0} \varphi_{v^*}(v+t(u-v)) = \varphi_{v^*}(v)$. By the definition of r, there exists a sequence $(z_n) \subset \mathscr{L}$ such that $\|z_n - w\| \to r$ and $\|z_n - w\| > r$ for all $n \in \mathbb{N}$. Therefore,

$$\begin{aligned}
\varphi_{v^*}(v) &= \lim_{t\downarrow 0} \varphi_{v^*}(v+t(u-v)) \\
&= \lim \varphi_{v^*}\left(v + \frac{\|z_n - w\| - r}{\|u-v\|}(u-v)\right) \\
&= \lim \varphi_{v^*}\left(v - \frac{r}{\|u-v\|}(u-v) + \frac{\|z_n - w\|}{\|u-v\|}(u-v)\right) \\
&= \lim \varphi_{v^*}\left(v - \frac{r}{\|u-w\|}(u-w) + \frac{\|z_n - w\|}{\|u-w\|}(u-w)\right) \\
&= \lim \varphi_{v^*}\left(w + \frac{\|z_n - w\|}{\|u-w\|}(u-w)\right) \\
&= \lim \varphi_{v^*}(z_n) \\
&\geq \varphi_{v^*}(v_0),
\end{aligned}$$

which is a contradiction. Now we show that v satisfies (7.9). Let γ be any positive real number and

$$v_\gamma := w + \frac{r - r_\gamma}{\|u-w\|}(u-w) \text{ with } r_\gamma := \min\{r/2, \gamma/2\} > 0.$$

Since $0 < r - r_\gamma < r < \|u-w|$, it implies that $v_\gamma \in]v,w[\setminus \mathscr{L}$. Therefore, $\varphi_{v^*}(v_\gamma) < \varphi_{v^*}(v_0) \leq \varphi_{v^*}(v)$. Furthermore,

$$\|v_\gamma - v\| = \left\| w + \frac{r - r_\gamma}{\|u-w\|}(u-w) - w - r\frac{u-w}{\|u-w\|} \right\| = r_\gamma < \gamma.$$

Hence, v satisfies (7.9). ☐

Theorem 7.2.4. *Let $\varphi : X \to \overline{\mathbb{R}}$ be a proper lower semicontinuous function on a Banach space X, and $\alpha > 0$. Consider the following statements*

(a) *φ is $\alpha-$robustly quasiconvex;*

(b) *For every $x, y \in X$*

$$\varphi(y) \leq \varphi(x) \implies \langle x^*, y-x \rangle \leq -\min\{\alpha\|y-x\|, \varphi(x) - \varphi(y)\}, \ \forall x^* \in \partial^F \varphi(x). \tag{7.11}$$

Then (a)⇒(b). *Additionally, if X is an Asplund space, then* (b)⇒(a).

Proof. Suppose that φ is $\alpha-$robustly quasiconvex, and $x, y \in X$ satisfy $\varphi(y) \leq \varphi(x)$. Assume that $x^* \in \partial^F \varphi(x)$. We will prove

$$\langle x^*, y-x \rangle \leq -\min\{\alpha\|y-x\|, \varphi(x) - \varphi(y)\}.$$

If $x = y$, the above inequality is trivial. Otherwise, we consider two cases:

Case 1. $\alpha\|y-x\| \le \varphi(x)-\varphi(y)$

We then need to prove that

$$\langle x^*, y-x\rangle \le -\alpha\|y-x\|. \tag{7.12}$$

By the Hahn-Banach theorem, there exists $v^* \in X^*$, $\|v^*\|=1$ such that $\langle v^*, y-x\rangle = \|y-x\|$. Consider the function $f: X \to \overline{\mathbb{R}}$ given by

$$f(z) = \varphi(z) + \alpha\langle v^*, z-x\rangle \quad \forall z \in X.$$

Then $f(x)=\varphi(x)$, and

$$f(y) = \varphi(y) + \alpha\langle v^*, y-x\rangle = \varphi(y) + \alpha\|y-x\| \le \varphi(x) = f(x),$$

i.e., $\max\{f(x), f(y)\} = f(x)$. Since φ is $\alpha-$robustly quasiconvex, f is quasiconvex. Therefore for each $t \in [0,1]$, we always have

$$\begin{aligned}\varphi(x) = f(x) = \max\{f(x), f(y)\} &\ge f(x+t(y-x)) \\ &= \varphi(x+t(y-x)) + t\alpha\langle v^*, y-x\rangle \\ &= \varphi(x+t(y-x)) + t\alpha\|y-x\|,\end{aligned}$$

which implies that

$$\varphi(x) - t\alpha\|y-x\| \ge \varphi(x+t(y-x)). \tag{7.13}$$

Since $x^* \in \partial^F\varphi(x)$, for any $\gamma > 0$, there exists a number $r > 0$ such that

$$\varphi(z) \ge \varphi(x) + \langle x^*, z-x\rangle - \gamma\|z-x\| \quad \forall z \in \mathbb{B}_r(x). \tag{7.14}$$

Let $t \in]0,1[$ such that $x+t(y-x) \in \mathbb{B}_r(x)$. It follows from (7.13) and (7.14) that

$$\varphi(x) - t\alpha\|y-x\| \ge \varphi(x) + t\langle x^*, y-x\rangle - t\gamma\|y-x\|,$$

and so

$$\langle x^*, y-x\rangle \le -\alpha\|y-x\| + \gamma\|y-x\|. \tag{7.15}$$

On taking limit on both sides of the above inequality as $\gamma \to 0^+$, we get (7.12).

Case 2. $\alpha\|y-x\| > \varphi(x)-\varphi(y)$

We have $\bar\alpha\|y-x\| = \varphi(x)-\varphi(y)$, where

$$\bar\alpha := \frac{\varphi(x)-\varphi(y)}{\|y-x\|} \in]0,\alpha[.$$

Since φ is $\bar\alpha-$robustly quasiconvex, we derive from Case 1 that

$$\begin{aligned}\langle x^*, y-x\rangle &\le -\bar\alpha\|y-x\| = \varphi(y)-\varphi(x) \\ &= -\min\{\alpha\|y-x\|, \varphi(x)-\varphi(y)\}.\end{aligned}$$

Conversely, assume that X is Asplund, and (b) holds. It follows from Theorem 7.1.5 that φ is quasiconvex. Suppose that φ is not $\alpha-$robustly quasiconvex, i.e., there exists $v^* \in X^* \setminus \{0\}, \|v^*\| < \alpha$ such that φ_{v^*} is not quasiconvex. By Lemma 7.2.3, there are $u, w \in X$ and $v \in]u,w[$ satisfying (7.7),(7.8), and (7.9). Since $\varphi_{v^*}(v) > \varphi_{v^*}(u)$, there exists $\delta > 0$ such that $\bar v^* := (1+\delta)v^*$ satisfies $\|\bar v^*\| < \alpha$ and $\varphi_{\bar v^*}(v) > \varphi_{\bar v^*}(u)$. Thus, we have $\varphi(v) > \varphi(u)$, $\varphi_{v^*}(v) > \varphi_{v^*}(u)$, $\varphi_{\bar v^*}(v) > \varphi_{\bar v^*}(u)$ and the lower semicontinuity of φ, φ_{v^*}, and $\varphi_{\bar v^*}$. This implies the existence of $\gamma > 0$ satisfying

$$\varphi(z) > \varphi(u), \quad \varphi_{v^*}(z) > \varphi_{v^*}(u), \quad \varphi_{\bar v^*}(z) > \varphi_{\bar v^*}(u) \quad \forall z \in \mathbb{B}_\gamma(v). \tag{7.16}$$

By the assertion (7.9), there is $v_\gamma \in \mathbb{B}_\gamma(v) \cap]v, w[$ such that $\varphi_{v^*}(v) > \varphi_{v^*}(v_\gamma)$. Then, v_γ can be written as

$$v_\gamma := v + \lambda(w - v) \text{ with } \lambda \in \left]0, \min\left\{1, \frac{\gamma}{\|w - v\|}\right\}\right[.$$

Since $\varphi_{v^*}(v) > \varphi_{v^*}(w)$ and $\varphi(v) \leq \varphi(w)$, we have $\langle v^*, w - v\rangle < 0$ and so

$$\begin{aligned}\varphi_{\bar{v}^*}(v_\gamma) - \varphi_{\bar{v}^*}(v) &= \varphi_{v^*}(v_\gamma) - \varphi_{v^*}(v) + \delta\langle v^*, v_\gamma - v\rangle \\ &= \varphi_{v^*}(v_\gamma) - \varphi_{v^*}(v) + \delta\lambda\langle v^*, w - v\rangle < 0.\end{aligned}$$

Observe that $\|v_\gamma - v\| = \lambda \|w - v\| < \gamma$. We can take $r \in]\, \|v_\gamma - v\|, \gamma[$. Applying Lemma 7.1.4 for $\varphi_{\bar{v}^*}$, $v \in [v_\gamma, u]$ with $\varphi_{\bar{v}^*}(v) > \varphi_{\bar{v}^*}(v_\gamma)$, there exist $x \in \text{dom}\varphi_{\bar{v}^*}$ and $x^* \in \partial^F \varphi_{\bar{v}^*}(x)$ such that

$$x \in [v_\gamma, v] + (r - \|v_\gamma - v\|)\mathbb{B} \quad \text{and} \quad \langle x^*, u - x\rangle > 0. \tag{7.17}$$

Then $x \in \mathbb{B}_\gamma(v)$ and so $\varphi(x) > \varphi(u)$ by (7.16). By the assumption (b) and the second inequality of (7.17),

$$-\langle \bar{v}^*, u - x\rangle < \langle x^* - \bar{v}^*, u - x\rangle \leq -\min\{\alpha\|u - x\|, \varphi(x) - \varphi(u)\}.$$

Since $\langle \bar{v}^*, u - x\rangle \leq \|\bar{v}^*\|\|u - x\| < \alpha\|u - x\|$, the above inequality implies that $\langle \bar{v}^*, u - x\rangle > \varphi(x) - \varphi(u)$, i.e., $\varphi_{\bar{v}^*}(x) < \varphi_{\bar{v}^*}(u)$ and this contradicts (7.16). □

We next construct a completely new characterisation for the robust quasiconvexity. It is based on the equivalence of the quasiconvexity of a lower semicontinuous function and the quasimonotonicity of its subdifferential operator.

Theorem 7.2.5. *Let $\varphi : X \to \overline{\mathbb{R}}$ be proper, lower semicontinuous on an Asplund space X and $\alpha > 0$. Then, φ is $\alpha-$robustly quasiconvex if and only if for any $(x, x^*), (y, y^*) \in$ graph $\partial^F \varphi$, we have*

$$\min\{\langle x^*, y - x\rangle, \langle y^*, x - y\rangle\} > -\alpha\|y - x\| \Longrightarrow \langle x^* - y^*, x - y\rangle \geq 0. \tag{7.18}$$

Proof. Suppose that φ is $\alpha-$robustly quasiconvex and that there exist $(x, x^*), (y, y^*) \in$ graph $\partial^F \varphi$ such that

$$\min\{\langle x^*, y - x\rangle, \langle y^*, x - y\rangle\} > -\alpha\|y - x\|. \tag{7.19}$$

Since φ is quasiconvex, $\partial^F \varphi$ is quasimonotone by Theorem 7.1.5. It follows that

$$\min\{\langle x^*, y - x\rangle, \langle y^*, x - y\rangle\} \leq 0. \tag{7.20}$$

Combining (7.19) and (7.20), we have

$$0 \leq -\min\left\{\left\langle x^*, \frac{y - x}{\|y - x\|}\right\rangle, \left\langle y^*, \frac{x - y}{\|x - y\|}\right\rangle\right\} < \alpha.$$

Without loss of generality, we may assume

$$\left\langle x^*, \frac{y - x}{\|y - x\|}\right\rangle = \min\left\{\left\langle x^*, \frac{y - x}{\|y - x\|}\right\rangle, \left\langle y^*, \frac{x - y}{\|x - y\|}\right\rangle\right\}.$$

Let $r > 0$ be such that

$$-\left\langle x^*, \frac{y - x}{\|y - x\|}\right\rangle < r \leq \alpha. \tag{7.21}$$

By the Hahn-Banach theorem, there exists $v^* \in X^*$ satisfying $\langle v^*, y - x\rangle = r\|y - x\|$ and $\|v^*\| = r \leq \alpha$. It follows that

$$\langle x^*, y - x\rangle + \langle v^*, y - x\rangle > -r\|y - x\| + r\|y - x\| = 0. \tag{7.22}$$

Consider $\varphi_{v^*} : X \to \overline{\mathbb{R}}$ given by $\varphi_{v^*}(u) = \varphi(u) + \langle v^*, u\rangle$ for any $u \in X$. Then, we have $\partial^F \varphi_{v^*}(u) = \partial^F \varphi(u) + v^*$ for $u \in \text{dom}\varphi$. Hence, by the quasiconvexity of φ_{v^*} and by Theorem 7.1.5, we have

$$\min\{\langle x^*, y - x\rangle + \langle v^*, y - x\rangle, \langle y^*, x - y\rangle + \langle v^*, x - y\rangle\} \leq 0.$$

Combining with (7.22), it implies

$$\langle y^*, x - y\rangle + \langle v^*, x - y\rangle \le 0, \text{ i.e., } \langle y^*, x - y\rangle \le \langle v^*, y - x\rangle = r\|x - y\|.$$

Letting $r \to -\left\langle x^*, \dfrac{y-x}{\|y-x\|}\right\rangle$, we obtain $\langle y^*, x - y\rangle \le \langle x^*, x - y\rangle$ and thus (7.18) holds.

Conversely, assume that (7.18) holds for all $x, y \in X$ and $x^* \in \partial^F\varphi(x), y^* \in \partial^F\varphi(y)$. Taking any v^* in $\alpha\mathbb{B}^*$, we next prove that $\varphi_{v^*} : X \to \mathbb{R}$, defined by $\varphi_{v^*}(u) = \varphi(u) + \langle v^*, u\rangle$ for any $u \in X$, is quasiconvex by showing the quasimonotonicity of $\widehat{\partial}\varphi_{v^*}$. Taking any $x, y \in X$ and $x^* \in \partial^F\varphi_{v^*}(x)$, $y^* \in \partial^F\varphi_{v^*}(y)$, then $x^* - v^* \in \partial^F\varphi(x)$, $y^* - v^* \in \partial^F\varphi(y)$. We then consider two cases.

Case 1. $\min\{\langle x^* - v^*, y - x\rangle, \langle y^* - v^*, x - y\rangle\} \le -\alpha\|y - x\|$

Without loss of generality, assume that

$$\langle x^* - v^*, y - x\rangle = \min\{\langle x^* - v^*, y - x\rangle, \langle y^* - v^*, x - y\rangle\}.$$

Since $\|v^*\| \le \alpha$, we have

$$\begin{aligned}\min\{\langle x^*, y - x\rangle, \langle y^*, x - y\rangle\} &\le \langle x^*, y - x\rangle = \langle x^* - v^*, y - x\rangle + \langle v^*, y - x\rangle \\ &\le -\alpha\|y - x\| + \|v^*\|\|y - x\| \le 0.\end{aligned}$$

Case 2. $\min\{\langle x^* - v^*, y - x\rangle, \langle y^* - v^*, x - y\rangle\} > -\alpha\|y - x\|$

Since (7.18) is satisfied, we have

$$\langle (x^* - v^*) - (y^* - v^*), x - y\rangle \ge 0,$$

i.e., $\langle x^* - y^*, x - y\rangle \ge 0$. It implies that

$$2\min\{\langle x^*, y - x\rangle, \langle y^*, x - y\rangle\} \le \langle x^*, y - x\rangle + \langle y^*, x - y\rangle \le 0.$$

Hence, $\partial^F\varphi_{v^*}$ is quasimonotone and thus φ_{v^*} is quasiconvex for any $v^* \in \alpha\mathbb{B}^*$ by Theorem 7.1.5. This yields the α-robust quasiconvexity of φ. □

Chapter 8

Abstract Convexity

Section 8.1 recalls some preliminary definitions and facts used throughout the chapter. We briefly introduce and study the space of abstract linear functions, and abstract convexity notions; some results are new in the context of abstract convexity. In Section 8.2, we provide some properties that are tantamount to the equality between the conjugate of the sum of some functions and the infimal convolution of the conjugates of those functions, which ensures the zero duality gap. We impose no topological assumptions on the primal space nor on the space of linear functions. A comparison with its forerunner, [23, Theorem 3.2] for convex programming, is established. The necessary and sufficient characterisation of the zero duality gap is provided, which is new even in the standard convex analysis. In Section 8.3, we equip the space of abstract linear functions with the weak* topology to extend some classical convex subdifferential calculus in the framework of abstract convexity. Some of the facts in convex analysis cannot be extended to abstract convexity without imposing additional assumptions. Here, we assume that the epigraphs of the conjugate functions admit the weak* additive property (see (8.38)). This condition holds for lower semicontinuous convex functions in classical convex analysis. Then, the zero duality conditions are exposed fully. In the last section of this chapter, we construct a nontrivial example for which our analysis applies.

In this chapter, when talking about a convex function, or convex set we normally mean conventional convex analysis. Throughout, X is a nonempty set, which can be thought of as a primal space. Note that we do not assume any algebraic or topological structure on X. Let $\mathcal{F} := \{f : X \to \mathbb{R} : f \text{ is a function}\} = X^{\mathbb{R}}$, i.e., $\mathcal{F}$ is the set of all functions acting from X to $\mathbb{R}$. The addition operator in $\mathcal{F}$ is the conventional one, i.e., $(f_1 + f_2)(x) := f_1(x) + f_2(x)$ for all $x \in X$. As mentioned above, the linear functions and their vertical shifts (which are the affine functions) are at the core of convex analysis. They have a crucial role in the definitions of conjugate functions and $\varepsilon-$subdifferentials. In the next section, we define the set $\mathcal{L}$ of abstract linear functions. We make minimal assumptions on $\mathcal{L}$ which trivially hold in the classical convex case. In Proposition 8.1.3 below, we show that some classical features of the conjugate functions and $\varepsilon-$subdifferentials, are still true for this general set $\mathcal{L}$ of linear functions.

8.1 Preliminaries Results on Abstract Convexity

8.1.1 Abstract Linear Space

Definition 8.1.1. A space of *abstract linear functions*, denoted by $\mathcal{L}$, is a subset of $\mathcal{F}$ that satisfies the following properties:

(a) $\mathcal{L}$ is closed with respect to the addition operator i.e. $f_1, f_2 \in \mathcal{L} \Longrightarrow f_1 + f_2 \in \mathcal{L}$;

(b) For every $l \in \mathcal{L}$ and $m \in \mathbb{N}$, there exist $l_1, \ldots, l_m \in \mathcal{L}$ such that

$$l = l_1 + \ldots + l_m. \tag{8.1}$$

Remark 8.1.2. If $0 \in \mathcal{L}$ and $\mathcal{L}$ verifies Definition 8.1.1(a), then $\mathcal{L}$ automatically verifies property (b) in Definition 8.1.1.

Throughout, we assume that $\mathcal{L}$ possesses properties (a) and (b).

Using the set $\mathcal{L}$, we state next the abstract counterparts of infimal convolution, Fenchel conjugate function, and $\varepsilon-$subdifferential.

Definition 8.1.3. (i) Given m functions $\psi_1, \ldots, \psi_m : X \to \mathbb{R}_{+\infty}$, the *infimal convolution* of the functions $\psi_1, \ldots, \psi_m$ is the function $\psi_1 \square \ldots \square \psi_m : X \to \mathbb{R}_{\pm\infty}$ defined by

$$\psi_1 \square \ldots \square \psi_m(x) := \inf_{x_1 + \ldots + x_m = x} \{\psi_1(x_1) + \ldots + \psi_m(x_m)\}, \quad \forall x \in X \tag{8.2}$$

with the convention that infimum over an empty set is $+\infty$.

(ii) Given $\mathcal{L}$ as in Definition 8.1.1, and a function $f : X \to \mathbb{R}_{+\infty}$, the *Fenchel conjugate* of f is the function $f^* : \mathcal{L} \to \mathbb{R}_{\pm\infty}$, defined as

$$f^*(l) := \sup_{x \in X} \{l(x) - f(x)\}. \tag{8.3}$$

(iii) Given $\mathcal{L}$ as in Definition 8.1.1, a number $\varepsilon \geq 0$, and a function $f : X \to \mathbb{R}_{\pm\infty}$, we define the *$\varepsilon-$subdifferential* point–to–set mapping $\partial_\varepsilon f : X \rightrightarrows \mathcal{L}$ at a point $x \in \operatorname{dom} f$ as

$$\partial_\varepsilon f(x) := \{l \in \mathcal{L} : f(y) - f(x) - (l(y) - l(x)) + \varepsilon \geq 0 \text{ for all } y \in X\}. \tag{8.4}$$

If $x \notin \operatorname{dom} f$, then $\partial_\varepsilon f(x) = \emptyset$.

We prove next some properties of the concepts defined in (i)–(iii).

Proposition 8.1.4. *Let $\mathcal{L}$ be a space of abstract linear functions. Then, the following statements hold:*

(i) *For all $x \in X$ and $\varepsilon \geq 0$, we have*

$$l \in \partial_\varepsilon f(x) \Longleftrightarrow f^*(l) + f(x) \leq l(x) + \varepsilon; \tag{8.5}$$

(ii) *Let $f : X \to \mathbb{R}_{+\infty}$. Then, $l \in \operatorname{dom} f^*$ if and only if for any $\varepsilon > 0$, there is an $x \in X$ such that $l \in \partial_\varepsilon f(x)$. In other words,*

$$\operatorname{dom} f^* = \bigcap_{\varepsilon > 0} \partial_\varepsilon f(X);$$

(iii) *For any $\varepsilon \geq 0$, we always have*

$$\bigcap_{\eta > 0} \partial_{\varepsilon + \eta} f(x) = \partial_\varepsilon f(x).$$

Suppose now that $f_1, \ldots, f_m : X \to \mathbb{R}_{+\infty}$ $(m \geq 2)$ are any functions such that $\bigcap_{i=1}^{m} \operatorname{dom} f_i \neq \emptyset$.

(iv) *The following inequality holds*

$$\left(\sum_{i=1}^{m} f_i\right)^* \leq f_1^* \square \ldots \square f_m^* \quad \text{in } \mathcal{L}; \tag{8.6}$$

(v) *For any $x \in X$, and any $\varepsilon \geq 0$, we have the following inclusion*

$$\bigcap_{\eta>0} \bigcup_{\substack{\varepsilon_1+\ldots+\varepsilon_m=\varepsilon+\eta \\ \varepsilon_i \geq 0}} \sum_{i=1}^{m} \partial_{\varepsilon_i} f_i(x) \subset \partial_{\varepsilon}\left(\sum_{i=1}^{m} f_i\right)(x). \tag{8.7}$$

Proof. (i) See [145, Proposition 7.10].

(ii) Let $l \in \operatorname{dom} f^*$. Due to the definition of f^*, $f^*(l) = \sup_{x \in X} \{l(x) - f(x)\}$, then for all $\varepsilon > 0$ we can find an $x \in X$ such that $l(x) - f(x) \geq f^*(l) - \varepsilon$, or $l(x) + \varepsilon \geq f^*(l) + f(x)$. By (i), the converse implication is equivalent to having $l \in \partial_\varepsilon f(x)$.

(iii) Fix $\varepsilon > 0$. It is clear from the definition that $\partial_\varepsilon f(x) \subset \partial_{\varepsilon+\eta} f(x)$ for every $\eta > 0$. Hence we deduce that $\partial_\varepsilon f(x) \subset \bigcap_{\eta>0} \partial_{\varepsilon+\eta} f(x)$. For the opposite inclusion, fix $\eta > 0$ and take $l \in \bigcap_{\eta>0} \partial_{\varepsilon+\eta} f(x)$. By (i), the latter is equivalent to having $-l(x) + f(x) + f^*(l) \leq \eta + \varepsilon$ for all $\eta > 0$. Thus, $-l(x) + f(x) + f^*(l) \leq \varepsilon$. Using (i) again, we deduce that $l \in \partial_\varepsilon f(x)$.

(iv) Take $l \in \mathcal{L}$. We consider two cases.

Case 1. Assume that $l \notin \operatorname{dom} \left(\sum_{i=1}^{m} f_i\right)^*$. We have $\sup_{x \in X} \left\{ l(x) - \sum_{i=1}^{m} f_i(x) \right\} = +\infty$. Take any additive decomposition of l, i.e., take any finite collection $l_1, \ldots, l_m$ such that $l_1 + \ldots + l_m = l$. We have

$$\sum_{i=1}^{m} f_i^*(l_i) \geq \sup_{x \in X} \left\{ \sum_{i=1}^{m} (l_i(x) - f_i(x)) \right\} = +\infty.$$

Taking infimum over all possible additive decompositions of l, we deduce that $f_1^* \square \ldots \square f_m^*(l) = +\infty$.

Case 2. Assume that $l \in \operatorname{dom} \left(\sum_{i=1}^{m} f_i\right)^*$. Take an arbitrary additive decomposition of l, i.e., take any finite collection $l_1, \ldots, l_m$ such that $l_1 + \ldots + l_m = l$. By definition of conjugate function, we have that for every $\varepsilon > 0$, there is an $x \in X$ such that

$$\begin{aligned} \left(\sum_{i=1}^{m} f_i\right)^*(l) &\leq l(x) - \sum_{i=1}^{m} f_i(x) + \varepsilon \\ &= \sum_{i=1}^{m} (l_i(x) - f_i(x)) + \varepsilon \\ &\leq \sum_{i=1}^{m} f_i^*(l_i) + \varepsilon, \end{aligned}$$

where we used the definition of $l_1, \ldots, l_m$ as additive decomposition of l in the equality and the definition of conjugate function in the last inequality. Since the additive decomposition is arbitrary, the expression above yields

$$\left(\sum_{i=1}^{m} f_i\right)^*(l) \leq \inf_{l_1+\ldots+l_m=l} \{f_1^*(l_1) + \ldots + f_m^*(l_m)\} + \varepsilon = f_1^* \square \ldots \square f_m^*(l) + \varepsilon.$$

Since the inequality holds for all $l \in \mathcal{L}$ and $\varepsilon > 0$, we obtain (8.6).

(v) Take $\eta > 0$. We claim that (v) is true if the inclusion

$$\sum_{i=1}^{m} \partial_{\varepsilon_i} f_i(x) \subset \partial_{\varepsilon+\eta}\left(\sum_{i=1}^{m} f_i\right)(x), \tag{8.8}$$

is true for every $\eta > 0$ and every additive decomposition $\varepsilon_1, \ldots, \varepsilon_m$ of $\varepsilon + \eta$. Indeed, consider m non-negative numbers $\varepsilon_1, \ldots, \varepsilon_m$ such that $\varepsilon_1 + \ldots + \varepsilon_m = \varepsilon + \eta$ and assume that (8.8) holds. Since the right hand side of (8.8) does not depend on the choice of $\varepsilon_1, \ldots, \varepsilon_m$, we have that

$$\bigcup_{\substack{\varepsilon_1+\ldots+\varepsilon_m=\varepsilon+\eta \\ \varepsilon_i \geq 0}} \sum_{i=1}^{m} \partial_{\varepsilon_i} f_i(x) \subset \partial_{\varepsilon+\eta}\left(\sum_{i=1}^{m} f_i\right)(x).$$

Now (v) will follow by taking intersection for all $\eta > 0$ in both sides of the expression above and then using (iii). Therefore, our claim is true and we proceed to establish (8.8) for every $\eta > 0$ and every additive decomposition $\varepsilon_1, \ldots, \varepsilon_m$ of $\varepsilon + \eta$.

If $\sum_{i=1}^{m} \partial_{\varepsilon_i} f_i(x) = \emptyset$, then inclusion (8.8) trivially holds. Take $l \in \sum_{i=1}^{m} \partial_{\varepsilon_i} f_i(x)$. Then there are $l_i \in \partial_{\varepsilon_i} f_i(x)$ $(i = 1, \ldots, m)$ such that $l = l_1 + \ldots + l_m$. Using (ii) we can write $f_i^*(l_i) + f_i(x) \leq l_i(x) + \varepsilon_i$ for all $i = 1, \ldots, m$. Add up the inequalities above, and use the fact that $\varepsilon_1 + \ldots + \varepsilon_m = \varepsilon + \eta$, to obtain

$$\sum_{i=1}^{m} f_i^*(l_i) + \sum_{i=1}^{m} f_i(x) \leq \sum_{i=1}^{m} l_i(x) + \varepsilon + \eta = l(x) + \varepsilon + \eta, \tag{8.9}$$

where we used the definition of $l_1, \ldots, l_m$ in the rightmost equality. Using (iv), we have

$$\sum_{i=1}^{m} f_i^*(l_i) \geq \left(\sum_{i=1}^{m} f_i\right)^* (l),$$

which combined with the previous inequality yields

$$(8.9) \Longleftrightarrow l \in \partial_{\varepsilon+\eta}\left(\sum_{i=1}^{m} f_i\right)(x),$$

where we used (i) in the equivalence. This establishes (8.8), and the proof of (v) is complete.

□

Remark 8.1.5. In the conventional convex setting, the assertions in Proposition 8.1.4 are well–known.

8.1.2 Abstract Convex Functions and Abstract Convex Sets

We start this subsection by defining the abstract affine functions, which are, as in the standard convex analysis, the vertical shifts of the abstract linear functions.

Definition 8.1.6. Let X and $\mathcal{L}$ be as in Definition 8.1.1. The space of *abstract affine functions* is defined as $\mathcal{H} := \{l + c : l \in \mathcal{L},\ c \in \mathbb{R}\}$.

Remark 8.1.7. The space of affine function $\mathcal{H}$ can be defined independently of the space of abstract linear functions $\mathcal{L}$ as an arbitrary set of functions which is closed with respect to the vertical shifts.

Equipped with the set $\mathcal{H}$, we can now extend the classical notions of convex function and convex set to our abstract framework.

Definition 8.1.8. [145, Definition 7.2] Let $X, \mathcal{L}$ and $\mathcal{H}$ be as in Definitions 8.1.6 and 8.1.1.

(i) Given any function $f : X \to \mathbb{R}_{+\infty}$, the set

$$\text{supp} f := \{h \in \mathcal{H} : h(x) \leq f(x), \forall x \in X\} \subset \mathcal{H}, \tag{8.10}$$

is called the *support set of* f.

(ii) A function $f : X \to \mathbb{R}_{+\infty}$ is said to be $\mathcal{L}-$convex if there is a subset $L \subset \mathcal{L}$ such that

$$f(x) = \sup_{l \in L} l(x), \quad \forall x \in X.$$

(iii) A set $C \subset \mathcal{L}$ is called $\mathcal{L}-$convex if for every element $l_0 \notin C$, there is an $x \in X$ such that

$$l_0(x) > \sup_{l \in C} l(x).$$

The $\mathcal{L}-$convex hull of a set $A \subset \mathcal{L}$ is the smallest $\mathcal{L}-$convex set that contains A.

Remark 8.1.9. (a) The definition of $\mathcal{L}-$convex set in Definition 8.1.8 (iii) is different from that in [145, Definition 1.4] and in [78] (we formulate [145, Definition 1.4] in Proposition 8.1.12 (i) below). Each definition reflects different ideas of convexity. Definition 8.1.8 (iii) is based on the *separation* property, whereas the rationale in Proposition 8.1.12 (i) below is based on the notion of convex combination.

(b) From Definition 8.1.8 (ii), we also have the definition for $\mathcal{H}-$convex functions, that is, a function $f : X \to \mathbb{R}_{+\infty}$ is said to be $\mathcal{H}-$*convex* if there is a non-empty subset $H \subset \mathcal{H}$ such that

$$f(x) = \sup_{h \in H} h(x), \quad \forall x \in X. \tag{8.11}$$

(c) From Definition 8.1.8 (iii), we also can obtain the definition for $\mathcal{H}-$convex sets, that is, a set $C \subset \mathcal{H}$ is $\mathcal{H}-$convex if for any $h_0 \notin C$ there is an $x \in X$ such that

$$h_0(x) > \sup_{h \in C} h(x).$$

(d) If f is $\mathcal{H}-$convex, and $C := \mathrm{supp} f$, then C is $\mathcal{H}-$convex. Similarly, if C is $\mathcal{H}-$convex and $f = \sup_{l \in C} l$, then $C = \mathrm{supp} f$.

Our next proposition collects some properties of $\mathcal{H}-$convex functions.

Proposition 8.1.10. *Let $X, \mathcal{L}$ and $\mathcal{H}$ be as in Definitions 8.1.8 and 8.1.1, and $f : X \to \mathbb{R}_{+\infty}$. The following assertions hold.*

(i) *For all $x \in X$, we have*

$$\sup_{h \in \mathrm{supp} f} h(x) \le f(x). \tag{8.12}$$

Equality holds in (8.12) for all $x \in X$ if and only if f is $\mathcal{H}-$convex.

(ii) *The following equality holds*

$$\mathrm{epi}\, f^* = \{(l, r) \in \mathcal{L} \times \mathbb{R} : l - r \in \mathrm{supp} f\}. \tag{8.13}$$

(iii) *If the function f is $\mathcal{H}-$convex, then for all $x \in \mathrm{dom}\, f$ and $\varepsilon > 0$, we have $\partial_\varepsilon f(x) \neq \emptyset$. Conversely, if for all $x \in X$, and $\varepsilon > 0$, we always have $\partial_\varepsilon f(x) \neq \emptyset$, then the function f is $\mathcal{H}-$convex.*

(iv) *If $(l, r) \in \mathrm{epi}\, f^*$, then $l \in \partial_{r + f(x) - l(x)} f(x)$ for all $x \in \mathrm{dom}\, f$.*

(v) *(Fenchel–Moreau) For all $x \in X$, we have*

$$f^{**}(x) \le f(x). \tag{8.14}$$

Equality holds in (8.14) for all $x \in X$ if and only if f is $\mathcal{H}-$convex.

Proof. (i) Inequality (8.12) holds trivially by the definition of supp$f$. When the equality holds in (8.12) for all $x \in X$, then f is $\mathcal{H}-$convex by Definition 8.1.8 (ii). Conversely, assume that f is $\mathcal{H}-$convex. By definition, there exists a set $H \subset \mathcal{H}$ such that $f(x) = \sup_{h \in H} h(x)$. It is easy to see from the definitions that $H \subset \text{supp} f$. Since

$$f(x) = \sup_{h \in H} h(x) \leq \sup_{h \in \text{supp} f} h(x) \leq f(x),$$

then $\sup_{h \in \text{supp} f} h(x) = f(x)$, for all $x \in X$. This proves that equality holds in (8.12).

(ii) See [78, Equation (1), page 444].

(iii) Assuming the function f is $\mathcal{H}-$convex, by (i), we have $\sup_{h \in \text{supp} f} h(x) = f(x)$. Then for all $x \in \text{dom} f$ and $\varepsilon > 0$, there is an affine function $l + c \in \text{supp} f$ with $l \in \mathcal{L}$, $c \in \mathbb{R}$ such that

$$l(x) + c + \varepsilon \geq f(x), \quad \text{and} \quad f(y) \geq l(y) + c, \ \forall y \in X.$$

Consequently, $f(y) - f(x) \geq l(y) - l(x) - \varepsilon$, $\forall y \in X$, which implies $l \in \partial_\varepsilon f(x)$.

Conversely, assume that for all $x \in X$ and $\varepsilon > 0$, we always have $\partial_\varepsilon f(x) \neq \emptyset$. We will prove that $f(x) = \sup_{h \in \text{supp} f} h(x)$. By (i), we always have

$$f(x) \geq \sup_{h \in \text{supp} f} h(x). \tag{8.15}$$

Take $\varepsilon > 0$. By our assumption, there is a linear function $l \in \partial_\varepsilon f(x)$, or, equivalently,

$$f(y) \geq l(y) - l(x) - \varepsilon + f(x), \quad \forall y \in X.$$

Then the affine function $\hat{l}(.) := l(\cdot) - l(x) - \varepsilon + f(x)$ belongs to the set supp$f$ and $\hat{l}(x) = l(x) - \varepsilon + f(x) - l(x) = f(x) - \varepsilon$. Hence,

$$f(x) - \varepsilon = \hat{l}(x) \leq \sup_{h \in \text{supp} f} h(x),$$

where the inequality holds because $\hat{l} \in \text{supp} f$. Since the above inequality holds for all $\varepsilon > 0$, inequality (8.15) now implies that $f(x) = \sup_{h \in \text{supp} f} h(x)$. Using now the last statement in (i), we conclude f is a $\mathcal{H}-$convex function.

(iv) Assume $(l, r) \in \text{epi} f^*$, then $f^*(l) \leq r$. Take $x \in \text{dom} f$. By definition of f^* (see Definition 8.1.3 (ii)), $r + f(x) - l(x) \geq 0$. We have $f^*(l) + f(x) \leq r + f(x) = l(x) + (r + f(x) - l(x))$. By Proposition 8.1.4 (i), $l \in \partial_{r+f(x)-l(x)} f(x)$.

(v) See [145, Theorem 7.1].

□

Remark 8.1.11. (a) Inequality (8.12) holds for $f : X \to \mathbb{R}_{+\infty}$ even when supp$f = \emptyset$. Indeed, in this case we have $\sup_{h \in \text{supp} f} h(x) = -\infty < f(x)$. In this situation, however, f is not an $\mathcal{H}$-convex function. Consequently, if f is $\mathcal{H}$-convex, we must have supp$f \neq \emptyset$.

(b) Proposition 8.1.10 (iii) is not an "if and only if " statement. The first implication holds for all $\mathcal{H}-$convex function, whereas the converse implication needs $\text{dom} f = X$.

The next proposition provides some properties of $\mathcal{L}-$convex sets and $\mathcal{H}-$convex sets used in the subsequent sections.

Proposition 8.1.12. *Let $X, \mathcal{L}$ and $\mathcal{H}$ be as in Definition 8.1.8, and $C \subset \mathcal{L}$. The following assertions hold.*

(i) *The set C is $\mathcal{L}-$convex if and only if there is an $\mathcal{L}-$convex function $f : X \to \mathbb{R}$ such that*

$$C = \text{supp} f.$$

In this case, $\text{supp} f \subset \mathcal{L}$*, i.e.* $\text{supp} f := \{l \in \mathcal{L} : l(x) \leq f(x), \forall x \in X\}$.

(ii) *The nonempty set $C \subset \mathcal{L}$ is $\mathcal{L}-$convex if and only if there is a $\mathcal{L}-$convex function $f : X \to \mathbb{R}_{+\infty}$ such that*

$$C = S^{\leq}_{f^*}(0) := \{l \in \mathcal{L} : f^*(l) \leq 0\}. \tag{8.16}$$

(iii) *Suppose in this part that $\mathcal{L}$ has linear structure (i.e. closed with respect to addition and multiplication by a scalar). If the set $C \subset \mathcal{L}$ is $\mathcal{L}-$convex, then C is closed for convex combinations of its elements. Namely, for all $l_1, l_2 \in C$ and $\alpha \in [0, 1]$, we have $\alpha l_1 + (1 - \alpha)l_2 \in C$.*

Proof. (i) See [145, Lemma 1.1, Page 6].

(ii) Assume that $C \subset \mathcal{L}$ is a nonempty $\mathcal{L}-$convex set. Let $f := \sup_{l \in C} l$. Then, f is $\mathcal{L}$-convex and $C = \text{supp} f$. By Proposition 8.1.10 (i), we obtain (8.16). For the opposite inclusion, Proposition 8.1.10 (i) implies $C = S^{\leq}_{f^*}(0) = \text{supp} f$ is an $\mathcal{L}-$convex set.

(iii) Suppose C is a $\mathcal{L}-$convex set. Take $l_1, l_2 \in C$ and $\alpha \in [0, 1]$. We will show that $\alpha l_1 + (1 - \alpha)l_2 \in C$ Let $f := \sup_{l \in C} l$ which is $\mathcal{L}-$convex by construction. By (i), we have $C \subset \text{supp} f$. Hence, for all $x \in X$ we have $l_1(x) \leq f(x)$, and $l_2(x) \leq f(x)$. Then, it is clear that $\alpha l_1(x) + (1 - \alpha)l_2(x) \leq f(x)$. This implies $\alpha l_1 + (1 - \alpha)l_2 \in \text{supp} f = C$.

□

Remark 8.1.13. The converse of Proposition 8.1.12 (iii) is not true. In the example provided in Section 8.4, the characterisations of $\mathcal{L}-$convex sets and $\mathcal{H}-$convex sets respectively in Propositions 8.4.3 and 8.4.5 show that not all sets which are closed for convex combinations are $\mathcal{L}-$convex or $\mathcal{H}-$convex.

8.2 Conditions for Zero Duality Gap

Let X and $\mathcal{L}$ be as in Definition 8.1.1. Given m functions $f_1, \ldots, f_m : X \to \mathbb{R}_{+\infty}$ ($m \geq 2$), consider the minimization problem

$$\begin{aligned} p := \inf & \left(\sum_{i=1}^{m} f_i(x) \right), \\ \text{s.t.} \quad & x \in X. \end{aligned} \tag{P}$$

The dual problem of (P) is given as follows:

$$\begin{aligned} d := \sup & \left(\sum_{i=1}^{m} -f_i^*(l_i) \right), \\ \text{s.t.} \quad & l_1, \ldots, l_m \in \mathcal{L}, \\ & l_1 + \ldots + l_m = 0. \end{aligned} \tag{D}$$

We refer the readers to [23] for comments and further explanation on the zero duality gap. Problem (P) is a very general minimization problem in which X is a general nonempty set, and there is no assumption on the convexity of the functions $f_1, \ldots, f_m$.

Denote by $v(P)$, $v(D)$, the optimal values of (P) and (D), respectively. We say that a *zero duality gap holds* for problems (P) and (D) if $v(P) = v(D)$. In general, we have the inequality $v(P) \geq v(D)$.

The following characterises the zero duality gap property for (P) and (D), using the infimal convolution of the conjugate functions f_i^*.

$$p = \inf_{x \in X} \left(\sum_{i=1}^{m} f_i(x) \right) = -\left(\sum_{i=1}^{m} f_i \right)^* (0); \tag{P1}$$

$$d = \sup_{l_1 + \ldots + l_m = 0} \left(\sum_{i=1}^{m} -f_i^*(l_i) \right) = -(f_1^* \square \ldots \square f_m^*)(0). \tag{D1}$$

Thus, the zero duality gap is equivalent to

$$\left(\sum_{i=1} f_i \right)^* (0) = (f_1^* \square \ldots \square f_m^*)(0). \tag{8.17}$$

Theorem 8.2.1 below extends [23, Theorem 3.2] to our general framework. It characterises the condition

$$\left(\sum_{i=1}^{m} f_i \right)^* = f_1^* \square \ldots \square f_m^*, \quad \text{in } \mathcal{L},$$

which clearly guarantees (8.17).

Theorem 8.2.1. *Let X and $\mathcal{L}$ be as in Definition 8.1.1. Fix $m \in \mathbb{N}$ such that $m \geq 2$. Let $f_1, \ldots, f_m, : X \to \mathbb{R}_{+\infty}$ be such that $\bigcap_{i=1}^{m} \operatorname{dom} f_i \neq \emptyset$. The following statements are equivalent.*

(i) *There is a $K > 0$ such that, for any $x \in X$ and any $\varepsilon > 0$,*

$$\partial_\varepsilon \left(\sum_{i=1}^{m} f_i \right)(x) \subset \sum_{i=1}^{m} \partial_{K\varepsilon} f_i(x). \tag{8.18}$$

(ii) $\left(\sum_{i=1}^{m} f_i \right)^* = f_1^* \square \ldots \square f_m^*$ *in $\mathcal{L}$.*

(iii) *For every $x \in \bigcap_{i=1}^{m} \operatorname{dom} f_i$ and any $\varepsilon \geq 0$,*

$$\partial_\varepsilon \left(\sum_{i=1}^{m} f_i \right)(x) = \bigcap_{\eta > 0} \left[\bigcup_{\substack{\varepsilon_1 + \ldots + \varepsilon_m = \varepsilon + \eta \\ \varepsilon_i \geq 0}} \sum_{i=1}^{m} \partial_{\varepsilon_i + \eta} f_i(x) \right]. \tag{8.19}$$

Proof. (i) $\Rightarrow$ (ii) Let $l \in \mathcal{L}$. By Proposition 8.1.4 (iv), we have

$$\left(\sum_{i=1}^{m} f_i \right)^* (l) \leq f_1^* \square \ldots \square f_m^*(l).$$

Let us show the opposite inequality. It is enough to consider the case in which $l \in \operatorname{dom} \left(\sum_{i=1}^m f_i \right)^*$.

By (ii) in Proposition 8.1.4, for any $\varepsilon > 0$, there is a $x \in X$ that $l \in \partial_\varepsilon \left(\sum_{i=1}^m f_i \right)(x)$. Using now (8.18), $l \in \sum_{i=1}^m \partial_{K\varepsilon} f_i(x)$, and there are $l_i \in \partial_{K\varepsilon} f_i(x)$ $(i = 1, \ldots, m)$ such that $l = l_1 + \ldots + l_m$. The following inequalities hold by Proposition 8.1.4 (i)

$$f_i^*(l_i) + f_i(x) \leq l_i(x) + K\varepsilon, \quad \forall i = 1, \ldots, m.$$

Adding up the inequalities above and using the definition of infimal convolution, we can write

$$\begin{aligned}(f_i^* \Box \ldots \Box f_m^*)(l) &\leq \sum_{i=1}^m f_i^*(l_i) \leq -\sum_{i=1}^m f_i(x) + \sum_{i=1}^m l_i(x) + mK\varepsilon \\ &= -\sum_{i=1}^m f_i(x) + l(x) + mK\varepsilon \leq \left(\sum_{i=1}^m f_i \right)^* (l) + mK\varepsilon.\end{aligned}$$

Since the inequality above holds for all $\varepsilon > 0$, we deduce that $(f_i^* \Box \ldots \Box f_m^*)(l) \leq \left(\sum_{i=1}^m f_i \right)^* (l)$. Using (8.18), we obtain (ii). The proof of (i) $\Rightarrow$ (ii) is complete.

(ii) $\Rightarrow$ (iii). Let ε be a non-negative number, $x \in \bigcap_{i=1}^m \operatorname{dom} f_i$. The inclusion

$$\partial_\varepsilon \left(\sum_{i=1}^m f_i \right)(x) \supset \bigcap_{\eta > 0} \left[\bigcup_{\varepsilon_i \geq 0, \sum_{i=1}^m \varepsilon_i = \varepsilon + \eta} \sum_{i=1}^m \partial_{\varepsilon_i} f_i(x) \right]$$

is shown in Proposition 8.1.4 (v). Let us show the opposite inclusion.

Take $l \in \partial_\varepsilon \left(\sum_{i=1}^m f_i \right)(x)$. By Proposition 8.1.4 (i), we have that

$$\left(\sum_{i=1}^m f_i \right)(x) + \left(\sum_{i=1}^m f_i \right)^* (l) \leq l(x) + \varepsilon. \tag{8.20}$$

Combine assumption (ii) with (8.20) to obtain, for all $\eta > 0$,

$$(f_1^* \Box \ldots \Box f_m^*)(l) + \eta \leq l(x) - \sum_{i=1}^m f_i(x) + \varepsilon + \eta. \tag{8.21}$$

Inequality (8.21) shows that $(f_1^* \Box \ldots \Box f_m^*)(l) < +\infty$. Using the definition of infimal convolution (Definition 8.1.3 (i)), there exist $l_1, \ldots, l_m \in \mathcal{L}$ such that $l = l_1 + \ldots + l_m$ and

$$f_1^*(l_1) + \ldots + f_m^*(l_m) \leq (f_1^* \Box \ldots \Box f_m^*)(l) + \eta.$$

Combine the above inequality with (8.21) to deduce $f_1^*(l_1) + \ldots + f_m^*(l_m) \leq l(x) - \sum_{i=1}^m f_i(x) + \varepsilon + \eta$.

Equivalently,

$$\sum_{i=1}^m (f_i(x) + f_i^*(l_i) - l_i(x)) \leq \varepsilon + \eta. \tag{8.22}$$

Set $\gamma_i := f_i(x) + f_i^*(l_i) - l_i(x)$ $(i = 1, \ldots, m)$. We have that $\gamma_i \geq 0$ by definition of conjugate function. Moreover, from Proposition 8.1.4 (i) we have that $l_i \in \partial_{\gamma_i} f_i(x)$ for all $i = 1, \ldots, m$. Due to (8.22), we obtain $\sum_{i=1}^{n} \gamma_i = \sum_{i=1}^{m}(f_i(x)+f_i^*(l_i)-l_i(x)) \le \varepsilon+\eta$. Choose $\varepsilon_i := \gamma_i+1/m(\varepsilon+\eta-\sum_{j=1}^{m}\gamma_j)$ $(i = 1, \dots, m)$. From the inequality above, we have that $\varepsilon_i \ge \gamma_i$. Hence, $l_i \in \partial_{\gamma_i} f_i(x) \subset \partial_{\varepsilon_i} f_i(x)$ for all $i = 1, \dots, m$, and $\sum_{i=1}^{m} \varepsilon_i = \varepsilon + \eta$. Altogether,

$$l \in \bigcup_{\substack{\varepsilon_1+\ldots+\varepsilon_m=\varepsilon+\eta \\ \varepsilon_i \ge 0}} \sum_{i=1}^{m} \partial_{\varepsilon_i} f_i(x), \quad \forall \eta > 0,$$

which yields (iii). The proof of (ii) $\Rightarrow$ (iii) is complete.

(iii) $\Rightarrow$ (i) We will show that (i) holds for $K = 2$. By (8.19), for all $x \in \bigcap_{i=1}^{m} \operatorname{dom} f_i(x)$ and $\varepsilon > 0$ we have

$$\begin{aligned} \partial_\varepsilon \left(\sum_{i=1}^{m} f_i\right)(x) &= \bigcap_{\eta>0} \left[\bigcup_{\substack{\varepsilon_1+\ldots+\varepsilon_m=\varepsilon+\eta \\ \varepsilon_i \ge 0}} \sum_{i=1}^{m} \partial_{\varepsilon_i} f_i(x) \right] \\ &\subset \bigcap_{\eta>0} \sum_{i=1}^{m} \partial_{\varepsilon+\eta} f_i(x) \subset \sum_{i=1}^{m} \partial_{2\varepsilon} f_i(x), \end{aligned}$$

where we used the fact that $\partial_{\varepsilon_i} f_i(x) \subset \partial_{\varepsilon+\eta} f_i(x)$ for all $i = 1, \dots, m$ in the first inclusion. The last inclusion is obtained by choosing $\eta = \varepsilon$. In the case $x \notin \bigcap_{i=1}^{m} \operatorname{dom} f_i(x)$, we have $\partial_\varepsilon \left(\sum_{i=1}^{m} f_i\right)(x) = \emptyset$. Therefore, we have shown that (8.19) holds for all $x \in X$. The proof of (iii) $\Rightarrow$ (i) is complete.

□

As mentioned above, Theorem 8.2.1 is an extension from the classical convex case to the framework of abstract convexity of the main result in [23, Theorem 3.2]. More precisely, the authors of [23] consider the case $\mathcal{L} = X^*$, the classical dual space of all continuous linear functions, and derive constraint qualifications for zero duality gap of a convex optimization problem. We quote their main result for the convenience of comparison.

Theorem 8.2.2. *[23, Theorem 3.2] Let X be a normed vector space, X^* its conjugate space with weak* topology, $m \in \mathbb{N}$, and $f_i : X \to \mathbb{R}_{+\infty}$ be proper lower semicontinuous convex functions where $i \in \{1, \dots, m\}$. Then the following four conditions are equivalent:*

(i) *There exists $K > 0$ such that for every $x \in \bigcap_{i=1}^{m} \operatorname{dom} f_i$, and every $\varepsilon > 0$,*

$$\operatorname{cl}\left[\sum_{i=1}^{m} \partial_\varepsilon f_i(x)\right] \subset \sum_{i=1}^{m} \partial_{K\varepsilon} f_i(x). \tag{8.23}$$

(ii) $\left(\sum_{i=1}^{m} f_i\right)^* = f_1^* \Box \ldots \Box f_m^*$ *in X^*.*

(iii) $f_1^* \Box \ldots \Box f_m^*$ *is weak* lower semicontinuous.*

(iv) *For every* $x \in X$ *and* $\varepsilon \geq 0$,

$$\partial_\varepsilon(f_1 + \ldots + f_m)(x) = \bigcap_{\eta>0} \left[\bigcup_{\substack{\varepsilon_1+\ldots+\varepsilon_m=\varepsilon+\eta \\ \varepsilon_i \geq 0}} (\partial_{\varepsilon_1} f_1(x) + \ldots + \partial_{\varepsilon_m} f_m(x)) \right]. \tag{8.24}$$

Remark 8.2.3. Note that (8.23) has the same right hand side as (8.18), but in the left hand side of (8.23) a topological closure expression is involved. Therefore in general (8.23) is more restrictive than (8.18). We will show in the next proposition that (8.23) $\iff$ (8.18) holds in any space of abstract linear functions $\mathcal{L}$ as long as the topology in $\mathcal{L}$ possesses property (8.25) stated below.

Recall the following known result for the conventional linear function space X^* equipped with the weak* topology (see [168, Corollary 2.6.7]): for all lower semiconinuous convex functions $f_1, \ldots, f_m$ ($m \geq 2$), $x \in X$ and $\varepsilon \geq 0$, we have

$$\partial_\varepsilon \left(\sum_{i=1}^m f_i \right)(x) = \bigcap_{\eta>0} \operatorname{cl} \left(\bigcup_{\substack{\varepsilon_1+\ldots+\varepsilon_m=\varepsilon+\eta \\ \varepsilon_i \geq 0}} \sum_{i=1}^m \partial_{\varepsilon_i} f_i(x) \right). \tag{8.25}$$

This calculus rule is the key ingredient in the proof of the implication (i) $\Rightarrow$ (ii) in [23, Theorem 3.2].

Proposition 8.2.4. *Given a set* X*, a space of abstract linear functions* $\mathcal{L}$*,* m *functions* $f_1, \ldots, f_m : X \to \mathbb{R}_{+\infty}$ *(*$m \geq 2$*), and* $x \in X$*, for any topology defined on* $\mathcal{L}$*, if the equality* (8.25) *holds, then the conditions* (8.18) *and* (8.23) *are equivalent.*

Proof. Observe that for any positive numbers η, ε with $0 < \eta \leq \varepsilon$, we have

$$\bigcup_{\substack{\varepsilon_1+\ldots+\varepsilon_m=\varepsilon+\eta, \\ \varepsilon_i \geq 0}} \sum_{i=1}^m \partial_{\varepsilon_i} f_i(x) \subset \sum_{i=1}^m \partial_{\varepsilon+\eta} f_i(x) \subset \sum_{i=1}^m \partial_{2\varepsilon} f_i(x). \tag{8.26}$$

Assume (8.23) holds with $K > 0$. Using (8.25), (8.26), and (8.23), we obtain (8.18) as follows:

$$\begin{aligned} \partial_\varepsilon \left(\sum_{i=1}^m f_i \right)(x) &\overset{(8.25)}{=} \bigcap_{\eta>0} \operatorname{cl} \left(\bigcup_{\substack{\varepsilon_1+\ldots+\varepsilon_m=\varepsilon+\eta \\ \varepsilon_i \geq 0}} \sum_{i=1}^m \partial_{\varepsilon_i} f_i(x) \right) \\ &\overset{(8.26)}{\subset} \bigcap_{\eta>0} \operatorname{cl} \left(\sum_{i=1}^m \partial_{2\varepsilon} f_i(x) \right) \overset{(8.23)}{\subset} \sum_{i=1}^m \partial_{2K\varepsilon} f_i(x). \end{aligned}$$

Conversely, assume that condition (8.18) holds with $K > 0$. Take in (8.26) the intersection of all $\eta > 0$ in the left hand side. Use also (8.25) and (8.18), to deduce the following inclusions

$$\begin{aligned} \operatorname{cl} \left(\sum_{i=1}^m \partial_\varepsilon f_i(x) \right) &\subset \bigcap_{\eta>0} \operatorname{cl} \left(\bigcup_{\substack{\varepsilon_1+\ldots+\varepsilon_m=m\varepsilon+\eta, \\ \varepsilon_i \geq 0}} \sum_{i=1}^m \partial_{\varepsilon_i} f_i(x) \right) \\ &\overset{(8.25)}{=} \partial_{m\varepsilon} \left(\sum_{i=1}^m f_i \right)(x) \overset{(8.18)}{\subset} \sum_{i=1}^m \partial_{mK\varepsilon} f_i(x), \end{aligned}$$

which gives (8.23) for $\tilde{K} := mK$ in place of K. □

Remark 8.2.5. (a) From the proof above we see that, if (8.23) holds with $K > 0$, then (8.18) holds with $2K$. On the other hand, if (8.18) holds with $K > 0$, then (8.23) holds with mK.

(b) When $0 \in \mathcal{L}$, part (ii) in Theorem 8.2.1 (or in Theorem 8.2.2 above) ensures the zero duality gap. In our Theorem 8.2.1, we drop the assumption that all the functions $f_1, \ldots, f_m$ are lower semicontinuous convex as in Theorem 8.2.2 and refine the core argument in the proof in [23]. However, without convexity, condition (8.18) might not be easily satisfied. When the functions f_i $(i = 1, \ldots, m)$ are not convex, there exist an $x \in X$ and $\varepsilon > 0$ such that $\partial_\varepsilon f_i(x) = \emptyset$ (see Proposition 8.1.10(iii)). In this situation, condition (8.18) fails.

The zero duality gap property is equivalent to equality (8.17), which is clearly less restrictive than condition (ii) in Theorem 8.2.1. In the next theorem, we relax condition (8.18) as well as item (ii) in Theorem 8.2.1 to obtain a necessary and sufficient condition for the zero duality gap property. This characterisation is new even in the classical convex case.

Theorem 8.2.6. *Suppose X is a set, $\mathcal{L}$ is a space of abstract linear functions with $0 \in \mathcal{L}$, and $f_i : X \to \mathbb{R}_{+\infty}$ $(i = 1, \ldots, m)$ $(m \geq 2)$ are functions with $\bigcap_{i=1}^{m} \operatorname{dom} f_i \neq \emptyset$. Then, the following conditions are equivalent.*

(i) *For all $\varepsilon > 0$, there exists an $x \in X$ such that*

$$\partial_\varepsilon f_1(x) + \ldots + \partial_\varepsilon f_m(x) \ni 0. \tag{8.27}$$

(ii) $\left(\sum_{i=1}^{m} f_i\right)^* (0) = f_1^* \square \ldots \square f_m^*(0) < +\infty.$

Proof. (i) $\Rightarrow$ (ii). Assume that assertion (i) holds. Then, for all $\varepsilon > 0$ there exists $x \in X$ such that

$$0 \in \partial_{\varepsilon/m} f_1(x) + \ldots + \partial_{\varepsilon/m} f_m(x). \tag{8.28}$$

By Proposition 8.1.4 (v), we have

$$0 \in \partial_{\varepsilon/m} f_1(x) + \ldots + \partial_{\varepsilon/m} f_m(x) \subset \bigcap_{\eta>0} \bigcup_{\substack{\varepsilon_1+\ldots+\varepsilon_m=\varepsilon+\eta \\ \varepsilon_i \geq 0}} \sum_{i=1}^{m} \partial_{\varepsilon_i} f_i(x) \overset{(8.7)}{\subset} \partial_\varepsilon \left(\sum_{i=1}^{m} f_i\right)(x).$$

This allows us to use Proposition 8.1.4 (ii) and deduce that $0 \in \operatorname{dom} \left(\sum_{i=1}^{m} f_i\right)^*$, or, equivalently, that $\left(\sum_{i=1}^{m} f_i\right)^* (0) < +\infty$. Due to Proposition 8.1.4(iv), we only need to show that $\left(\sum_{i=1}^{m} f_i\right)^* (0) \geq f_1^* \square \ldots \square f_m^*(0)$. By (8.28) and Proposition 8.1.4 (i), there are $l_i \in \partial_{\varepsilon/m} f_i(x)$ $(i = 1, \ldots, m)$ such that $0 = l_1 + \ldots + l_m$ and $f_i^*(l_i) + f_i(x) \leq l_i(x) + \varepsilon/m$ for all $i = 1, \ldots, m$. This implies $\sum_{i=1}^{m} f_i^*(l_i) + \sum_{i=1}^{m} f_i(x) \leq \varepsilon$. We then have the following inequalities

$$f_1^* \square \ldots \square f_m^*(0) \leq \sum_{i=1}^{m} f_i^*(l_i) \leq \varepsilon - \sum_{i=1}^{m} f_i(x) \leq \varepsilon + \left(\sum_{i=1}^{m} f_i\right)^* (0),$$

where we used the definition of conjugate function in the rightmost inequality. Letting $\varepsilon \downarrow 0$, we obtain (ii).

(ii) $\Rightarrow$ (i). Using the inequality in (ii), we have that $0 \in \operatorname{dom} \left(\sum_{i=1}^{m} f_i\right)^*$. Take $\varepsilon > 0$. By Proposition 8.1.4 (ii) we can find an $x \in X$ such that $0 \in \partial_\varepsilon \left(\sum_{i=1}^{m} f_i\right)(x)$. Using the equality in

(ii) and Proposition 8.1.4 (i), we have

$$f_1^*\square\ldots\square f_m^*(0) = \left(\sum_{i=1}^m f_i\right)^*(0) \overset{(8.5)}{\leq} \varepsilon - \left(\sum_{i=1}^m f_i\right)(x). \tag{8.29}$$

By Definition 8.1.3 (i), there are $l_1, \ldots, l_m \in \mathcal{L}$ such that $l_1 + \ldots + l_m = 0$ and

$$f_1^*(l_1) + \ldots + f_m^*(l_m) \leq f_1^*\square\ldots\square f_m^*(0) - \varepsilon/2. \tag{8.30}$$

Thus, we have the following estimation

$$\begin{aligned} f_1^*(l_1) + \ldots + f_m^*(l_m) &\overset{(8.30)}{\leq} f_1^*\square\ldots\square f_m^*(0) - \varepsilon/2 \overset{(8.29)}{\leq} \varepsilon/2 - \left(\sum_{i=1}^m f_i\right)(x) \\ &\overset{l_1+\ldots+l_m=0}{=} \varepsilon/2 + \sum_{i=1}^m l_i(x) - \left(\sum_{i=1}^m f_i\right)(x). \end{aligned}$$

We derive

$$\varepsilon/2 \geq \sum_{i=1}^m (f_i^*(l_i) + f_i(x) - l_i(x)). \tag{8.31}$$

And we also have $f_i(x) + f_i^*(l_i) - l_i(x) \geq 0$ for all $i = 1, \ldots, m$, then

$$-(f_i^*(l_i) + f_i(x) - l_i(x)) + \varepsilon/2 \overset{(8.31)}{\geq} \sum_{j=1, j\neq i}^m (f_j^*(l_j) + f_j(x) - l_j(x)) \geq 0, \quad \forall i = 1, \ldots, m.$$

Using the inequality above, we can write $f_i^*(l_i) + f_i(x) \leq l_i(x) + \varepsilon/2$ for all $i = 1, \ldots, m$, which by Proposition 8.1.4 (i) yields $l_i \in \partial_{\varepsilon/2} f_i(x)$ for all $i = 1, \ldots, m$. Since $\partial_{\varepsilon/2} f_i(x) \subset \partial_\varepsilon f_i(x)$ we deduce that $0 = l_1 + \ldots + l_m \in \sum_{i=1}^m \partial_\varepsilon f_i(x)$.

□

We will show in the next theorem that if the inclusion (8.27) holds for a fixed $x \in X$ for all $\varepsilon > 0$, or equivalently

$$\bigcap_{\varepsilon>0} (\partial_\varepsilon f_1(x) + \ldots + \partial_\varepsilon f_m(x)) \ni 0, \tag{8.32}$$

then it characterises a stronger property.

Theorem 8.2.7. *Suppose X is a set, $\mathcal{L}$ is a space of abstract linear functions with $0 \in \mathcal{L}$, and $f_i : X \to \mathbb{R}_{+\infty}$ $(i = 1, \ldots, m)$ $(m \geq 2)$ are functions with $\bigcap_{i=1}^m \operatorname{dom} f_i \neq \emptyset$. Then the following conditions are equivalent.*

(i) *There exists an $x \in X$ such that* (8.32) *holds.*

(ii) $\left(\sum_{i=1}^m f_i\right)^*(0) = f_1^*\square\ldots\square f_m^*(0) < +\infty$*, and the x for which condition* (8.32) *holds is a solution of problem* (P)*.*

Proof. (i) $\Rightarrow$ (ii). Suppose there is an $x \in X$ such that (8.32) holds. By Theorem 8.2.6, we deduce the first statement in part (ii). Let us show the second statement in (ii). By (8.32), for all $\varepsilon > 0$, there are $l_1, \ldots, l_m \in \mathcal{L}$ such that $l_1 + \ldots + l_m = 0$, and $l_i \in \partial_\varepsilon f_i(x)$ for all $i = 1, \ldots, m$. We have

$$f_i^*(l_i) + f_i(x) \leq l_i(x) + \varepsilon, \quad \text{for all } i = 1, \ldots, m,$$

which implies $\sum_{i=1}^{m} f_i^*(l_i) + \sum_{i=1}^{m} f_i(x) \leq m\varepsilon$. Then, we have

$$f_1^* \square \ldots \square f_m^*(0) \overset{l_1+\ldots+l_m=0}{\leq} \sum_{i=1}^{m} f_i^*(l_i) \leq -\left(\sum_{i=1}^{m} f_i\right)(x) + m\varepsilon \leq \left(\sum_{i=1}^{m} f_i\right)^*(0) + m\varepsilon.$$

Let $\varepsilon \downarrow 0$, we have equalities in the expression above. Hence,

$$f_1^* \square \ldots \square f_m^*(0) = \left(\sum_{i=1}^{m} f_i\right)^*(0), \quad \text{and} \quad -\left(\sum_{i=1}^{m} f_i\right)(x) = \left(\sum_{i=1}^{m} f_i\right)^*(0),$$

this implies that the supremum in the expression of the conjugate function is attained at x. This establishes the second statement in (ii).

(ii) $\Rightarrow$ (i). Suppose there is an $x \in X$ such that

$$f_1^* \square \ldots \square f_m^*(0) = \left(\sum_{i=1}^{m} f_i\right)^*(0) = -\left(\sum_{i=1}^{m} f_i\right)(x) < +\infty. \tag{8.33}$$

For every $\varepsilon > 0$, there are $l_1, \ldots, l_m \in \mathcal{L}$ such that $l_1 + \ldots + l_m = 0$ and

$$f_1^*(l_1) + \ldots + f_m^*(l_m) \leq f_1^* \square \ldots \square f_m^*(0) + \varepsilon. \tag{8.34}$$

Combining (8.33) and (8.34), we obtain

$$f_1^*(l_1) + \ldots + f_m^*(l_m) \leq -\left(\sum_{i=1}^{m} f_i\right)(x) + \varepsilon. \tag{8.35}$$

Since $f_j^*(l_j) + f_j(x) - l_j(x) \geq 0$ for all $j = 1, \ldots, m$, we have

$$\begin{aligned}
f_i^*(l_i) + f_i(x) - l_i(x) &\leq f_i^*(l_i) + f_i(x) - l_i(x) + \sum_{j=1, j\neq i}^{m} [f_j^*(l_j) + f_j(x) - l_j(x)] \\
&\leq \sum_{j=1}^{m} (f_j^*(l_j)) + \sum_{j=1}^{m} f_j(x) - \sum_{j=1}^{m} l_j(x) \\
&\overset{l_1+\ldots+l_m=0, \text{ and } (8.35)}{\leq} \left(-\sum_{j=1}^{m} f_j(x) + \varepsilon\right) + \sum_{j=1}^{m} f_j(x) = \varepsilon, \quad \forall i = 1, \ldots, m.
\end{aligned}$$

Thus, $l_i \in \partial_\varepsilon f_i(x)$ for all $i = 1, \ldots, m$. We deduce that $0 = l_1 + \ldots + l_m \in \sum_{i=1}^{m} \partial_\varepsilon f_i(x)$. Since $\varepsilon > 0$ is arbitrary, we conclude that $0 \in \bigcap_{\varepsilon>0} \sum_{i=1}^{m} \partial_\varepsilon f_i(x)$, which is (i).

□

8.3 Abstract Convexity with Weak* Topology

In this section, we consider specifically the weak* topology $\sigma(\mathcal{L}, X)$ (see [26, Section 3.3]) on the abstract linear function space $\mathcal{L}$. We expand some fundamental results of standard convex analysis to the framework of abstract convexity. Condition (8.23) is fully extended into the abstract convexity framework using weak* topology.

8.3.1 Weak* Topology

Recall from section 8.1 that $\mathcal{F}$ is the set of all real functions defined on X. On $\mathcal{F}$, the weak* topology $\sigma(\mathcal{F}, X)$ is the weakest topology that makes all the functions $x : \mathcal{F} \to \mathbb{R}, f \mapsto f(x)$ continuous. It follows that, for any sequence $(f_n) \subset \mathcal{F}, f_n \longrightarrow f \in \mathcal{F}$ if and only if $f_n(x) \longrightarrow f(x)$ for all $x \in X$ (pointwise convergence topology) (see [26, Proposition 3.1]). Recall that $\tau(\mathcal{F}, \sigma(\mathcal{F}, X))$ is a Hausdorff topological space (the proof of this fact is similar to the one in [26, Proposition 3.11]).

In the context of [26], the weak* topology is studied for the dual space X^* of a normed linear vector space X. In order to study general abstract linear function spaces $\mathcal{L}$, here, we generalize several fundamental results in [26] to the set of function $\mathcal{L}$ which satisfy certain conditions below. Note that in our analysis, no topology is assumed on X.

We define the multiplication by a scalar on $\mathcal{L}$ as usual: $(\alpha l)(x) = \alpha(l(x))$ for all $x \in X$ and $\alpha \in \mathbb{R}$. We will assume that $\mathcal{L}$ satisfies the following conditions:

A- $\mathcal{L}$ is an $\mathbb{R}-$vector space.

B- $\mathcal{L}$ is weak* closed in $\mathcal{F}$.

Observe that the evaluation functions $x : \mathcal{L} \to \mathbb{R}, l \mapsto l(x)$, $x \in X$ are linear in the conventional sense. Thus, the functions $|x| : \mathcal{L} \to \mathbb{R}, |x|(l) = |l(x)|$ are seminorms (see [168, Page 4]). The next two propositions show that the space $\mathcal{L}$ is a locally convex topological vector space.

Proposition 8.3.1. *Let X be a set, $\mathcal{L}$ a space of abstract linear functions equipped with the weak* topology, $l_0 \in \mathcal{L}$, $\varepsilon > 0$, and a set of vectors $\{x_1, \dots, x_n\}$ in X. Then the set*

$$V(x_1, \dots, x_n|\varepsilon) = \{l \in \mathcal{L} : |l(x_i) - l_0(x_i)| < \varepsilon, \quad \forall i = 1, \dots, n\}$$

is a neighbourhood of l_0 in $\mathcal{L}$. Moreover, the collection of all $V(x_1, \dots, x_n|\varepsilon)$, $n \in \mathbb{N}$, $x_1, \dots, x_n \in X$ and $\varepsilon > 0$ forms a basis of neighbourhoods of l_0 in $\mathcal{L}$.

Proposition 8.3.2. *Suppose X is a set, and $\mathcal{L}$ is a space of abstract linear functions equipped with the weak* topology. Assume further that conditions (A), (B) hold, then $\mathcal{L}$ is a locally convex topological vector space.*

Remark 8.3.3. Proofs of propositions 8.3.1 and 8.3.2 are natural modifications of ones in [26].

Example 8.3.4. [78, Example 2.1] Let $\mathcal{L}$ be a set of functions defined on the Euclidean space $\mathbb{R}^n$, comprising all the functions $l := \sum_{i=0}^{n} a_i h_i \in \mathcal{L}$ where $a_i \in \mathbb{R}$, $i = 1, \dots, n$ and $x = (x_1, \dots, x_n) \in \mathbb{R}^n$

$$h_0(x) = \|x\|^2, \quad h_1(x) = x_1, \dots, h_n(x) = x_n.$$

The set $\mathcal{L}$ possesses the following properties.

Proposition 8.3.5. (i) *$\mathcal{L}$ is closed in $\mathcal{F}$ with respect to weak* topology.*

(ii) *The sequence (or net) $(l_i)_{i \in I}$, $l_i := a_i^0 h_0 + \ldots + a_i^n h_n \to l = a^0 h_0 + \ldots + a^n h_n$ if and only if $a_i^j \to a^j$ for all $j = 0, 1, \dots, n$.*

(iii) *The space $\mathcal{L}$ equipped with weak* topology is homeomorphic to $\mathbb{R}^{n+1}$ with the standard Euclidean norm.*

Proof. (i) We show that $\mathcal{L}$ is closed in $\mathcal{F}$. Let $(l_i)_{i \in I}$ be a net in $\mathcal{L}$ such that $l_i \to l \in \mathcal{F}$ with $l_i := a_i^0 h_0 + \ldots + a_i^n h_n$. We have $l_i(1, 0, \dots, 0) + l_i(-1, 0, \dots, 0) = 2a_i^0 \to l(1, 0, \dots, 0) + l(-1, 0, \dots, 0) =: 2a^0$, or $a_i^0 \to a^0$. At the same time, $l_i(x) - a_i^0 \|x\|^2 \to l(x) - a^0 \|x\|^2$ for all $x \in X$ and $l_i - a_i^0 \|\cdot\|^2$ is an usual linear function in $\mathbb{R}^n$ (for all $i \in I$). Thus, $l - a^0 \|\cdot\|^2$ is also an linear function in $\mathbb{R}^n$ (because it is a pointwise limit of standard linear functions). Hence, $l - a^0 \|\cdot\|^2$ can be represented as a linear combination of $h_1, \dots, h_n$. Therefore, $l \in \mathcal{L}$.

(ii) It is shown in (i) that if $l_i \to l$, then $a_i^0 \to a^0$ and $(l_i - a_i^0 \|\cdot\|^2) \to (l - a^0 \|\cdot\|^2)$. This means $a_j^i \to a_j$ for all $j = 0, \dots, n$.

(iii) Consider the bijective mapping $\varphi : \mathcal{L} \to \mathbb{R}^{n+1}$ with $\varphi(l) := (a_0, a_1, \dots, a_n)$, where $l = \sum_{i=0}^{n} a_i h_i$. In the view of part (ii), φ is continuous.

If $(a_i^0, a_i^1, \dots, a_i^n)_{i \in I} \to (a^0, \dots, a^n)$, then $(\varphi^{-1}(a_i^0, a_i^1, \dots, a_i^n))_{i \in I} = \sum_{j=0}^{n} a_i^j h_j \to \sum_{j=0}^{n} a^j h_j = \varphi^{-1}(a^0, a^1, \dots, a^n)$. Hence, φ^{-1} is continuous.

□

In view of Proposition 8.3.5, we can treat the space of abstract linear functions $\mathcal{L}$ (in Example 8.3.4) as the vector space $\mathbb{R}^{n+1}$ with the basis $h_0, h_1, \dots, h_n$. Moreover, Proposition 8.3.5 (iii) implies that the weak* topology is equivalent to the Euclidean norm topology in $\mathbb{R}^{n+1}$.

When X is a normed vector space, and $\mathcal{L} = X^*$ its conventional dual space, the compactness of the unit ball $\mathbb{B}^*$ is of utmost importance. Here we establish a generalization of the Banach–Alaoglu–Bourbaki theorem (see [26, Theorem 3.16]) to real functions space $\mathcal{F}$.

Theorem 8.3.6. *Suppose that X is a nonempty set (possibly linear vector space), $\mathcal{F}$ is the set of all real functions acting from X to $\mathbb{R}$, and $F, G : X \to \mathbb{R}$ are any real functions defined on X with $G \le F$ over X. Let $K := \{f \in \mathcal{F} : G(x) \le f(x) \le F(x),\ \forall\, x \in X\}$. Then, K is a weak* compact set in $\mathcal{F}$.*

Proof. Consider the product space $Y := \mathbb{R}^X$, and Y is equipped with the product topology. Let $\phi : \mathcal{F} \to Y$ be defined as $\phi(f) := (f(x))_{x \in X}$. It is well known that ϕ is *a homeomorphism* from $\mathcal{F}$ to Y (w.r.t. the product topology). We will show that the set $H := \phi(K) = \{y \in Y : G(x) \le y(x) \le F(x),\ \forall x \in X\}$ is a compact set in Y. Indeed, we have

$$H = \prod_{x \in X} [G(x), F(x)], \tag{8.36}$$

a product of compact sets. By Tychonoff's Theorem, H is compact, and so K is weak* compact.

□

Corollary 8.3.7. *Let X be a linear vector space, $\mathcal{F}$ be equipped with the weak* topology. Assume that A is a weak* closed subset of $\mathcal{F}$. Then A is weak* compact if the functions $F := \sup_{f \in A} f$ and $G := \sup_{f \in A} (-f)$ are finite everywhere i.e. $F(x), G(x) \in \mathbb{R}$ for all $x \in X$.*

Proof. Observe that $G = -\inf_{g \in A} f$. The inclusion $A \subset \{f \in \mathcal{F} : -G(x) \le f(x) \le F(x), \forall x \in X\}$ is trivial. Then we have $A = \{f \in \mathcal{F} : -G(x) \le f(x) \le F(x), \forall x \in X\} \cap A$. Theorem 8.3.6 together with the weak* closedness of A yields the weak* compactness of A. □

Remark 8.3.8. By taking $\mathcal{L} = X^*$ and $G(x) = -1, F(x) = 1$, and $A = \{x^* : \|x^*\| \le 1\}$, we recover the Banach–Alaoglu–Bourbaki Theorem.

8.3.2 Sum Rule for Subdifferentials

In this subsection, we improve one of the main results of [78, Theorem 3.2]. In [78], the authors studied the extended sum rule for abstract convex functions using the additivity property of the epigraph of the conjugate functions. The additivity condition on the sets epi f^* and epi g^* is stated as follows

$$\operatorname{epi} f^* + \operatorname{epi} g^* = \operatorname{epi} (f+g)^*. \tag{8.37}$$

However, this additivity condition may not be valid even for convex functions in the classical sense, because it entails weak* closedness of the set epi f^*+epi g^*. It is well known that the sum of closed sets is not closed in general. Condition (8.37) is even stronger than condition $(f+g)^* = f^* \Box g^*$ (cf. [78, Corollary 5.1]).

Our aim is to sharpen condition (8.37) by the following condition:

$$\operatorname{cl}^*(\operatorname{epi} f^* + \operatorname{epi} g^*) = \operatorname{epi} (f+g)^*, \tag{8.38}$$

where cl $^*A$ denotes the weak* closure of the set A. The weak* topology on $\mathcal{L} \times \mathbb{R}$ is the product topology of the topology $\tau(\mathcal{L}, \sigma(\mathcal{L}, X))$ on $\mathcal{L}$ and the standard topology on $\mathbb{R}$. Since the set on the right-hand side of (8.37) is always weak* closed, the additivity condition (8.37) readily implies (8.38). Moreover, condition (8.38) is more convenient because (as shown in [155] and [78, Corollary 3.1]) it holds for standard convex functions. We will use the less restrictive condition (8.38) in the next section. Namely, it will allow us to apply Theorem 8.2.2 for establishing new zero duality gap characterisations.

We first recall the sum rule in [78].

Theorem 8.3.9. *[78, Theorem 3.2] Let X be a set and let $\mathcal{L}$ be a set of abstract linear functions on X. Let $f, g : X \to \mathbb{R}_{+\infty}$ be $\mathcal{H}-$convex functions such that* $\operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$*. Then, equality* (8.37) *holds if and only if for any $\varepsilon \geq 0$,*

$$\partial_\varepsilon (f+g)(x) = \bigcup_{\varepsilon_1+\varepsilon_2=\varepsilon, \varepsilon_1, \varepsilon_2 \geq 0} \partial_{\varepsilon_1} f(x) + \partial_{\varepsilon_2} g(x), \quad x \in \operatorname{dom} f \cap \operatorname{dom} g. \tag{8.39}$$

We next recover some classic properties.

Proposition 8.3.10. *Suppose that X is a vector space and that $\mathcal{L}$ is a space of abstract functions. Then,*

(i) *for any $x \in X$, the function $x(\cdot) : \mathcal{L} \to \mathbb{R}_{+\infty}$, $l \mapsto l(x)$ is weak* continuous;*

(ii) *for every function $f : X \to \mathbb{R}_{+\infty}$, its conjugate function f^* is weak* lower semicontinuous.*

Proof. (i) Take $l \in \mathcal{L}$. For all $(l_i)_{i \in I} \subset \mathcal{L}$ (I is a directed set) such that $l_i \to l$ w.r.t weak* topology, we have

$$\lim x(l_i) = \lim l_i(x) = l(x) = x(l).$$

Thus, $x(\cdot)$ is weak* continuous.

(ii) Take $f : X \to \mathbb{R}_{+\infty}$, and consider $f^*(l) = \sup_{x \in X}\{l(x) - f(x)\}$ for all $l \in \mathcal{L}$. Then, the epigraph of f^* is

$$\operatorname{epi} f^* := \{(l, \lambda) : f^*(l) \leq \lambda\} = \bigcap_{x \in X} \{(l, \lambda) : l(x) - f(x)) \leq \lambda\} = \bigcap_{x \in X} \operatorname{epi} [x(\cdot) - f(x)].$$

Since the function $[x(\cdot) - f(x)] : \mathcal{L} \to \mathbb{R}_{+\infty}$ is weak* continuous, the set epi $[x(\cdot) - f(x)]$ is weak* closed in $\mathcal{L} \times \mathbb{R}$, and so is epi f^*. This implies the weak* lower semicontinuity of the function f^*.

□

Proposition 8.3.11. *Let X be a set, $\mathcal{L}$ be a space of abstract linear functions equipped with the weak* topology. Assume further that condition (B) holds. Then, the following statements hold.*

(i) *The $\mathcal{L}-$convex sets are weak* closed in $\mathcal{L}$.*

(ii) *The subdifferential set* $\partial_\varepsilon f(x)$ *is weak** *closed for any function* $f : X \to \mathbb{R}_{\pm\infty}$*,* $x \in \operatorname{dom} f$ *and* $\varepsilon \geq 0$*.*

(iii) *For any functions* $f_1, \dots, f_m$ $(m \geq 2)$ $x \in \bigcap_{i=1}^{m} \operatorname{dom} f_i$ *and* $\varepsilon \geq 0$*, we have the following inclusion*

$$\bigcap_{\eta>0} \operatorname{cl}^* \left(\bigcup_{\substack{\varepsilon_1+\ldots+\varepsilon_m=\varepsilon+\eta \\ \varepsilon_i \geq 0}} \sum_{i=1}^{m} \partial_{\varepsilon_i} f_i(x) \right) \subset \partial_\varepsilon \left(\sum_{i=1}^{m} f_i \right)(x). \tag{8.40}$$

Proof. (i) Let $f : X \to \mathbb{R}_{+\infty}$ be a $\mathcal{L}-$convex function such that $\operatorname{supp} f = A$. Consider the set $B := \operatorname{cl}^* A \subset \mathcal{L}$ and the function $g(x) := \sup_{h \in B} h(x)$. We claim that it is sufficient to show that $g \equiv f$. Indeed, if this holds, then for all $l \in B$ we have $l \leq g = f$, and thus $l \in A$, or $A = \operatorname{cl}^* A$. Observe that g is also an $\mathcal{L}-$convex function and by construction $g \geq f$. Thus, we proceed to prove the claim that $g \equiv f$. Take $x \in X$, we need to show that $f(x) \geq g(x)$. Indeed, for all $l \in B$, there exists a net $(l_i)_{i \in I} \subset A$, with I being a directed set, such that $l_i \to l$ w.r.t weak* convergence and $f(x) \geq l_i(x) \to l(x)$. Taking the supremum in the right hand side for all $l \in B$, we conclude that $f(x) \geq g(x)$. Since this holds for any x, the proof is complete.

(ii) By Proposition 8.1.4 (i), for all $x \in \operatorname{dom} f$ and $\varepsilon \geq 0$, we have

$$\partial_\varepsilon f(x) = \{l : f^*(l) - l(x) \leq -f(x) + \varepsilon\} = S^{\leq}_{f^* - x(\cdot)}(-f(x) + \varepsilon).$$

From Proposition 8.3.10, the function $f^* - x(\cdot)$ is weak* lower semicontinuous. Thus, its lower level set $S^{\leq}_{f^* - x(\cdot)}(-f(x) + \varepsilon)$ is weak* closed. Hence, $\partial_\varepsilon f(x)$ is weak* closed.

(iii) Proposition 8.1.4 (v) provides inclusion (8.7). This inclusion and (ii) yield

$$\begin{aligned}\operatorname{cl}^* \left(\bigcap_{\eta>0} \left(\bigcup_{\substack{\varepsilon_1+\ldots+\varepsilon_m=\varepsilon+\eta \\ \varepsilon_i \geq 0}} \sum_{i=1}^{m} \partial_{\varepsilon_i} f_i(x) \right) \right) &\subset \bigcap_{\eta>0} \operatorname{cl}^* \left(\bigcup_{\substack{\varepsilon_1+\ldots+\varepsilon_m=\varepsilon+\eta \\ \varepsilon_i \geq 0}} \sum_{i=1}^{m} \partial_{\varepsilon_i} f_i(x) \right) \\ &\overset{(8.7)}{\subset} \operatorname{cl}^* \left(\partial_\varepsilon \left(\sum_{i=1}^{m} f_i \right)(x) \right) \overset{(ii)}{=} \partial_\varepsilon \left(\sum_{i=1}^{m} f_i \right)(x)\end{aligned}$$

This proves (8.40).

□

We present in the next theorem our main result of this subsection.

Theorem 8.3.12. *Let* X *be a nonempty set, and* $\mathcal{L}$ *an abstract linear function space which is weak** *closed in* $\mathcal{F}$ *and* $f_1, \dots, f_m : X \to \mathbb{R}_\infty$ $(m \geq 2)$ *be functions defined on* X *with* $\cap_{i=1}^m \operatorname{dom} f_i \neq \emptyset$*. Assume that*

$$\operatorname{cl}^* \left(\sum_{i=1}^{m} \operatorname{epi} f_i^* \right) = \operatorname{epi} \left(\sum_{i=1}^{m} f_i \right)^*. \tag{8.41}$$

Then, for any number $\varepsilon \geq 0$*, for all* $x \in \bigcap_{i=1}^{m} \operatorname{dom} f_i$*,* (8.25) *holds for all* $x \in \cap_{i=1}^m \operatorname{dom} f_i$*.*

Proof. Assume that (8.41) holds. Take $\varepsilon \geq 0$. For all $x \in \bigcap_{i=1}^{m} \operatorname{dom} f_i$, due to Proposition 8.3.11(iii), we have

$$\partial_\varepsilon \left(\sum_{j=1}^{m} f_j \right)(x) \supset \bigcap_{\eta>0} \operatorname{cl}^* \left(\bigcup_{\substack{\varepsilon_1+\ldots+\varepsilon_m=\varepsilon+\eta \\ \varepsilon_j \geq 0}} \sum_{j=1}^{m} \partial_{\varepsilon_j} f_j(x) \right).$$

Now we prove the converse inclusion. Let $l \in \partial_\varepsilon \left(\sum_{j=1}^{m} f_j \right)(x)$. Then, by Proposition 8.1.4 (i), we have $\left(\sum_{j=1}^{m} f_j \right)^* (l) \leq l(x) + \varepsilon - \sum_{j=1}^{m} f_j(x)$, thus

$$\left(l, l(x) + \varepsilon - \sum_{j=1}^{m} f_j(x) \right) \in \operatorname{epi} \left(\sum_{j=1}^{m} f_j \right)^* \overset{(8.41)}{=} \operatorname{cl}^* \left(\sum_{j=1}^{m} \operatorname{epi} f_j^* \right).$$

There are nets $(l_{1,i}, \lambda_{1,i})_{i \in I}, \ldots, (l_{m,i}, \lambda_{m,i})_{i \in I}$ with I is a directed set such that

$$(l_{j,i}, \lambda_{j,i}) \in \operatorname{epi} f_j^*, \quad j = 1, \ldots, m,\ i \in I, \tag{8.42}$$

$$\sum_{j=1}^{m} l_{j,i} \overset{w^*}{\to} l, \quad \sum_{j=1}^{m} \lambda_{j,i} \to l(x) + \varepsilon - \sum_{j=1}^{m} f_j(x). \tag{8.43}$$

Set $\gamma_{j,i} := \lambda_{j,i} + f_j(x) - l_{j,i}(x)$ for all $j = 1, \ldots, m$, $i \in I$. By Proposition 8.1.10 (iv), (8.42) implies $\gamma_{j,i} \geq 0$ and $l_{j,i} \in \partial_{\gamma_{j,i}} f_j(x)$ for all $j = 1, \ldots, m$ and $i \in I$. Take $\eta > 0$. By (8.43), there is $i_0 \in I$ such that for all $i \geq i_0$ we have

$$\sum_{j=1}^{m} \lambda_{j,i} \leq \sum_{j=1}^{m} l_{j,i}(x) + (\varepsilon + \eta) - \sum_{j=1}^{m} f_j(x)$$

The above inequality yields $\sum_{j=1}^{m} \gamma_{j,i} \leq \varepsilon + \eta$. Thus, $\sum_{j=1}^{m} l_{j,i} \in \bigcup_{\substack{\varepsilon_1+\ldots+\varepsilon_m=\varepsilon+\eta \\ \varepsilon_i \geq 0}} \sum_{i=1}^{m} \partial_{\varepsilon_i} f_i(x)$.

Thus, $l \in \operatorname{cl}^* \left(\bigcup_{\substack{\varepsilon_1+\ldots+\varepsilon_m=\varepsilon+\eta \\ \varepsilon_i \geq 0}} \sum_{i=1}^{m} \partial_{\varepsilon_i} f_i(x) \right)$ for all $\eta > 0$. Hence, we proved (8.25). □

Remark 8.3.13. (a) If condition (8.38) holds, and additionally $\sum_{i=1}^{m} \operatorname{epi} f_i^*$ is weak* closed, then condition (8.37) holds.

(b) Thank to (8.40) and Proposition 8.3.11 (ii), we always have

$$\bigcup_{\substack{\varepsilon_1+\ldots+\varepsilon_m=\varepsilon \\ \varepsilon_i \geq 0}} \sum_{i=1}^{m} \partial_{\varepsilon_i} f_i(x) \subset \bigcap_{\eta>0} \operatorname{cl}^* \left(\bigcup_{\substack{\varepsilon_1+\ldots+\varepsilon_m=\varepsilon+\eta \\ \varepsilon_i \geq 0}} \sum_{i=1}^{m} \partial_{\varepsilon_i} f_i(x) \right) \subset \partial_\varepsilon \left(\sum_{i=1}^{m} f_i \right)(x).$$

Therefore, if the exact sum rule (8.39) holds, the sum rule (8.25) also holds.

8.3.3 Zero Duality Gap with Weak* Topology

We next present our main theorem of this section.

Theorem 8.3.14. *Let X be a normed vector space, $\mathcal{L}$ a space of abstract linear functions with weak* topology, and $f_i : X \to \mathbb{R}_{+\infty}$ $(i = 1, \dots, m)$ $(m \geq 2)$. Consider the following five conditions:*

(i) *There exists $K > 0$ such that for every $x \in \bigcap_{i=1}^{m} \operatorname{dom} f_i$, and every $\varepsilon > 0$,* (8.18) *holds;*

(ii) *There exists $K > 0$ such that for every $x \in \bigcap_{i=1}^{m} \operatorname{dom} f_i$, and every $\varepsilon > 0$,* (8.23) *holds with respect to weak* topology;*

(iii) $\left(\sum_{i=1}^{m} f_i\right)^* = f_1^* \square \dots \square f_m^*$ *in $\mathcal{L}$;*

(iv) $f_1^* \square \dots \square f_m^*$ *is weak* lower semicontinuous;*

(v) *For every $x \in X$ and $\varepsilon \geq 0$,* (8.24) *holds.*

We have (i) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (v) $\Rightarrow$ (iv) $\Rightarrow$ (ii) .

Additionally, if the sum rule (8.25) *(or condition* (8.41)*) holds, then all five statements are equivalent.*

Proof. All the implications (except (ii) $\Leftrightarrow$ (iii)) are clear due to Theorems 8.2.1 and 8.2.2, Propositions 8.2.4 and 8.3.10.

(i) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (v) is proved in Theorem 8.2.1.

(iii) $\Rightarrow$ (iv) is trivial due to the fact that $\left(\sum_{i=1}^{n} f_i\right)^*$ is weak* lower semicontinuous (see Proposition 8.3.10).

(iv) $\Rightarrow$ (ii). The proof proceeds exactly the same way as the proof of (iii) $\Rightarrow$ (i) in [23, Theorem 3.2].

If the sum rule (8.25) holds, then by Proposition 8.2.4, (i) $\Leftrightarrow$ (ii).

$\square$

Remark 8.3.15. A similar series of intriguing corollaries as in [23] can be deduced with more or less identical proofs to those in [23].

8.4 Zero Duality Gap for a Family of Noncovex Problems

In this section, we consider a nontrivial example in which the weak* closed additivity condition (8.38) holds for any $\mathcal{H}-$convex function. Thus, we show that the important sum rule (8.25) holds in a more general framework. To improve readability, we use here notation different to the one used in previous sections.

Definition 8.4.1. Let $X := \mathbb{R}$, and set $\phi_a(t) := at^2$ for some $a \in \mathbb{R}$, $\Psi_{a,b}(t) := at^2 + b$ for some $a, b \in \mathbb{R}$. We define

(i) the set $\mathcal{L} := \{\phi_a : a \in \mathbb{R}\}$ the space of abstract linear functions;

(ii) the set$\mathcal{H} := \{\Psi_{a,b} : a, b \in \mathbb{R}\}$ the space of abstract affine functions.

Remark 8.4.2. A proof similar to that of Proposition 8.3.5 shows that the weak* topologies on $\mathcal{L}$ and $\mathcal{H}$ are isomorphic to the usual topologies of $\mathbb{R}$ and $\mathbb{R}^2$, respectively. Indeed, the mappings $\Gamma_1 : \mathcal{L} \to \mathbb{R}$, $\Gamma_1(\phi_a) := a$ and $\Gamma_2 : \mathcal{H} \to \mathbb{R}^2$, $\Gamma_2(\Psi_{a,b}) = (a, b)$ are homeomorphism with $\mathbb{R}, \mathbb{R}^2$ being equipped with the usual topologies.

We next characterise $\mathcal{L}-$convex sets, and $\mathcal{H}-$convex sets for this example.

Proposition 8.4.3 (Characterisation of $\mathcal{L}-$convex sets)**.** *A set $C \subset \mathcal{L}$ is $\mathcal{L}-$convex if and only if there exists $A \in \mathbb{R}$ such that*

$$C = \{\phi_a : a \leq A\}.$$

Proof. The set of the form $C := \{\phi_a : a \leq A\}$ ($A \in \mathbb{R}$) is the support set of the function ϕ_A in $\mathcal{L}$. Hence, C is $\mathcal{L}-$convex.

Conversely, suppose C is an $\mathcal{L}-$convex set and let $A := \sup\{a : \phi_a \in C\}$. We have $C \subset \{\phi_a : a \leq A\}$, by the definition of A. On the other hand, for every $a \leq A$, $\phi_a(t) \leq \phi_A(t)$ for all $t \in \mathbb{R}$, then $C \supset \{\phi_a : a \leq A\}$. □

To characterise $\mathcal{H}-$convex sets, we need to use a standard strict convex separation theorem, which we recall next.

Theorem 8.4.4. *[26, Theorem 1.7] Let X be a normed vector space, $A \subset X$ and $B \subset X$ be two nonempty convex subsets such that $A \cap B = \emptyset$. Assume that A is closed and B is compact. Then there exists a closed hyperplane that strictly separates A and B i.e. there is $x^* \in X^* \setminus \{0\}$ such that*

$$\sup_{x \in A} \langle x^*, x\rangle < \inf_{y \in B} \langle x^*, y\rangle.$$

We are now in conditions of characterising $\mathcal{H}-$convex sets.

Proposition 8.4.5 (Characterisation of $\mathcal{H}-$convex sets)**.** *A set C is $\mathcal{H}-$convex, $C \not\equiv \mathcal{H}$ if and only if the following properties hold*

(i) *C is weak* closed, convex, and upper bounded (i.e. there exist $A, B \in \mathbb{R}$ such that for all $\psi_{a,b} \in C$ we have $a \leq A$ and $b \leq B$);*

(ii) *if $\Psi_{a,b} \in C$, then $\Psi_{a',b'} \in C$ for all $a' \leq a, b' \leq b$.*

Proof. Suppose C is a $\mathcal{H}-$convex set and $f := \sup_{h \in C} h$. In view of Proposition 8.1.12(iii) and Proposition 8.3.11, C is weak* closed and convex. Let

$$A := \sup\{a : \Psi_{a,b} \in C, \text{ for some } b \in \mathbb{R}\}, \tag{8.44}$$
$$B := \sup\{b : \Psi_{a,b} \in C, \text{ for some } a \in \mathbb{R}\}. \tag{8.45}$$

We must have $A, B \in \mathbb{R}$, otherwise if $A = +\infty$ or $B = +\infty$, then $f \equiv +\infty$, which implies $C \equiv \mathcal{H}$ by Remark 8.1.9 (c), a contradiction to our assumption. Hence, (i) holds.

For all $\Psi_{a,b} \in C$, and any numbers $a' \leq a$, $b' \leq b$, we have $\psi_{a',b'}(t) \leq \Psi_{a,b}(t) \leq f(t)$ for all $t \in \mathbb{R}$. Thus, $\Psi_{a',b'} \in \operatorname{supp} f = C$. Altogether, (i) and (ii) hold when C is $\mathcal{H}-$convex.

Conversely, suppose the set C satisfies conditions (i) and (ii). We will use Remark 8.1.9 (c) to show that C is $\mathcal{H}-$convex. Suppose $\Psi_{\bar{a},\bar{b}} \notin C$ and let A, B be defined as in (8.44), (8.45). Due to (i), $A, B \in \mathbb{R}$.

Case 1. $\bar{a} > A$.

Observe that $\Psi_{A,B}(t) \geq \sup_{h \in C} h(t)$ for all $t \in \mathbb{R}$. Since $\bar{a} > A$, then for some t with the absolute value $|t|$ sufficiently large, we always have $(\bar{a} - A)t^2 > B - \bar{b}$. Thus, $\Psi_{\bar{a},\bar{b}}(t) > \sup_{h \in C} h(t)$. By Definition 8.1.8 (iii), C is $\mathcal{H}-$convex.

Case 2. $\bar{a} \leq A$. Then $\bar{b} > B$ (otherwise by (ii) $\Psi_{\bar{a},\bar{b}} \in C$).

Consider the set $C^2 := \{(a,b) : \Psi_{a,b} \in C\}$ in $\mathbb{R}^2$. By (i), C is weak* closed and convex and $\Psi_{\bar{a},\bar{b}} \notin C$. By Remark 8.4.2, C^2 is a closed convex set in $\mathbb{R}^2$. Since $\Psi_{\bar{a},\bar{b}} \notin C$, we have $(\bar{a}, \bar{b}) \notin C^2$. By the strict convex separation theorem (Theorem 8.4.4), there is $(x^*, y^*) \in \mathbb{R}^2 \setminus \{(0,0)\}$ such that

$$\bar{a}x^* + \bar{b}y^* > \sup_{(a,b) \in C^2} ax^* + by^*. \tag{8.46}$$

We consider three subcases.

1. If $y^* < 0$, then due to property (ii), for every $\Psi_{a,b} \in C$ we can fix parameter a and let $b \to -\infty$. Then, $\sup_{(a,b) \in C^2} ax^* + by^* \overset{b \to -\infty}{\to} +\infty$, which is a contradiction to (8.46).
2. If $y^* = 0$, then $x^* \neq 0$.

   If $x^* > 0$, then due to our assumption we have $Ax^* \geq \bar{a}x^*$, which contradicts to (8.46) ($\bar{a}x^* > \sup_{(a,b) \in C^2} (ax^*) = Ax^*$).

   If $x^* < 0$, then due to property (ii) and with an argument similar to case (i), we have

   $$\sup_{(a,b) \in C^2} ax^* \overset{a \to -\infty}{\to} +\infty.$$

   Altogether, it yields a contradiction.
3. If $y^* > 0$, then we can divide both sides of (8.46) by y^* to obtain

   $$\bar{a}\frac{x^*}{y^*} + \bar{b} > \sup_{(a,b) \in C^2} a\frac{x^*}{y^*} + b.$$

   Arguing as in (ii), we claim that $\frac{x^*}{y^*} \geq 0$. Let $t := \sqrt{\frac{x^*}{y^*}}$ we have

   $$\bar{a}t^2 + \bar{b} > \sup_{(a,b) \in C^2} at^2 + b, \quad or \quad \Psi_{\bar{a},\bar{b}}(t) > \sup_{\Psi \in C} \Psi(t).$$

Combining **case 1** and **case 2**, we conclude that C is an $\mathcal{H}-$convex set by Remark 8.1.9 (c). □

Remark 8.4.6. (i) Any set of the form $\{\Psi_{a,b} : a \leq A, b \leq B\}$ with $A, B \in \mathbb{R}$ is a $\mathcal{H}-$convex set. However, not every $\mathcal{H}-$convex set is of this form. Indeed, the set $C := \{\Psi_{a,b} : b \leq \frac{1}{a},\ a < 0\}$ (the support set of the function $-2|x|$) satisfies the conditions (i), (ii) in Proposition 8.4.5 and cannot be written under the form $\{\Psi_{a,b} : a \leq A, b \leq B\}$ for some $A, B \in \mathbb{R}$.

(ii) All the $\mathcal{H}-$convex functions are even functions and minored by $\mathcal{H}$, but the converse is not true. Take the function $\sin(t + \pi/2)$ as an example, it is even and minored by $\mathcal{H}$ but not $\mathcal{H}-$convex.

We are now in a position to characterise the set $\mathcal{H}$ as in Definition 8.4.1 and verify condition (8.41).

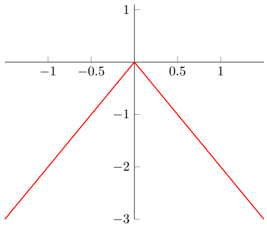


Figure 8.1: Function $-2\,|x|$.

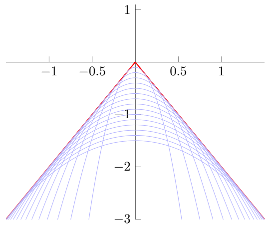


Figure 8.2: Support set of $-2\,|x|$.

Theorem 8.4.7. *Let $\mathcal{H}$ be as in Definition 8.4.1 and f, g be $\mathcal{H}-$convex functions, we have* $\mathrm{cl}^*(\mathrm{supp} f + \mathrm{supp} g)$ *is $\mathcal{H}-$convex. Consequently,*

$$\mathrm{cl}^*(\mathrm{epi}\, f^* + \mathrm{epi}\, g^*) = \mathrm{epi}\,(f+g)^*.$$

Proof. By Remark 8.1.9 (c), the set $\mathrm{supp} f$ and $\mathrm{supp} g$ are $\mathcal{H}-$convex. By Proposition 8.4.5 (i) and Remark 8.4.2, these support sets can be uniquely identified with two closed and convex sets C_f and C_g in $\mathbb{R}^2$. The proposition also yields the existence of upper bounds $A_f, B_f \in \mathbb{R}$ and $A_g, B_g \in \mathbb{R}$ such that for every $(a,b) \in C_f$ we have $a \le A_f$ and $b \in B_f$ and similarly for every $(a,b) \in C_g$ we have $a \le A_g$, $b \le B_g$.

By Proposition 8.4.5 (i) and the convexity of each set C_f, C_g, we conclude that the set $S := C_f + C_g$ is convex. Moreover, by Remark 8.4.2, the set $\mathrm{cl}^*(\mathrm{epi}\, f^* + \mathrm{epi}\, g^*)$ is homeomorphic to the set $\mathrm{cl}\, S$, where the closure in the latter is taken in the usual sense. Since $\mathrm{cl}\,(S)$ is closed and convex in $\mathbb{R}^2$, the set $\mathrm{cl}^*(\mathrm{epi}\, f^* + \mathrm{epi}\, g^*)$ is weak* closed and convex in $\mathcal{H}$. By Remark 8.4.2, the boundedness property holds for $\mathrm{cl}\,(S)$ as well. This implies that condition (i) in Proposition 8.4.5 holds for the set $\mathrm{cl}^*(\mathrm{supp} f + \mathrm{supp} g)$. We proceed now to show that condition (ii) holds for this set.

Take any $\Psi_{a,b} \in \mathrm{cl}^*(\mathrm{supp} f + \mathrm{supp} g)$ with $a, b \in \mathbb{R}$, and any numbers $a' \le a$ and $b' \le b$. We must show that $\Psi_{a',b'} \in \mathrm{cl}^*(\mathrm{supp} f + \mathrm{supp} g)$, equivalently, by Remark 8.4.2, $(a', b') \in \mathrm{cl}\,(S)$. Indeed, there are sequences $(a_{f,n}, b_{f,n}) \subset C_f$, $(a_{g,n}, b_{g,n}) \subset C_g$ such that $(a_{f,n}, b_{f,n}) + (a_{g,n}, b_{g,n}) \to (a,b)$, equivalently $(a_{f,n} + a_{g,n}, b_{f,n} + b_{g,n}) \to (a,b)$. By Proposition 8.4.5 (ii), we have

$$(a_{f,n} - (a-a')/2, b_{f,n} - (b-b')/2) \in C_f, \quad \text{and}$$
$$(a_{g,n} - (a-a')/2, b_{g,n} - (b-b')/2) \in C_g,$$

for all $n \in \mathbb{N}$. Then, $(a_{f,n} + a_{g,n} - (a-a'), b_{f,n} + b_{g,n} - (b-b')) \to (a', b') \in \mathrm{cl}\,(S)$. Hence, if the set $\mathrm{cl}^*(\mathrm{supp} f + \mathrm{supp} g)$ satisfies the two conditions (i), (ii) as in Proposition 8.4.5, then it is an $\mathcal{H}-$convex set.

Since $f = \sup_{\Psi \in \mathrm{supp} f} \Psi$ and $g = \sup_{\Psi \in \mathrm{supp} g} \Psi$ (by Proposition 8.1.10 (i)), and taking into account that the set $\mathrm{cl}^*(\mathrm{supp} f + \mathrm{supp} g)$ is $\mathcal{H}-$convex, then $f + g = \sup_{\Psi \in \mathrm{cl}^*(\mathrm{supp} f + \mathrm{supp} g)} \Psi$ is $\mathcal{H}-$convex with $\mathrm{supp}(f+g) = \mathrm{cl}^*(\mathrm{supp} f + \mathrm{supp} g)$.

By Proposition 8.1.10 (ii), $\mathrm{epi}\,(f+g)^* = \{(\phi_a, b) : \Psi_{a,-b} \in \mathrm{supp}(f+g)\} = \{(\phi_a, b) : \Psi_{a,-b} \in \mathrm{cl}^*(\mathrm{supp} f + \mathrm{supp} g)\}$, then for all $(\phi_a, b) \in \mathrm{epi}\,(f+g)^*$, there are two nets $(\Psi_{a_i,b_i})_{i \in I}$ and $(\Psi_{a'_i,b'_i})_{i \in I}$ (with I being a directed set) such that

$$\Psi_{a_i,b_i} \in \mathrm{supp} f, \quad \Psi_{a'_i,b'_i} \in \mathrm{supp} g, \quad \forall i \in I,$$
$$\Psi_{a_i,b_i} + \Psi_{a'_i,b'_i} \xrightarrow{w^*} \Psi_{a,-b}. \tag{8.47}$$

By Remark 8.4.2, condition (8.47) is equivalent to $(a_i + a'_i, b_i + b'_i) \to (a, -b)$ in the standard Euclidean space $\mathbb{R}^2$. Also by Proposition 8.1.10 (ii), $(\phi_{a_i}, -b_i) \in \mathrm{epi}\, f^*$ and $(\phi_{a'_i}, -b'_i) \in \mathrm{epi}\, g^*$ for all $i \in I$. Then,

$\phi_{a_i} + \phi_{a_i'} \to \phi_a$ and $-b_i - b_i' \to -b$. This implies $(\phi_a, b) \in \mathrm{cl}^*(\mathrm{epi}\, f^* + \mathrm{epi}\, g^*)$. Thus,

$$\mathrm{cl}^*(\mathrm{epi}\, f^* + \mathrm{epi}\, g^*) = \mathrm{epi}\,(f+g)^*.$$

□

Example 8.4.8. Consider the following functions defined in $\mathbb{R}$.

$$f_1(x) := x^4 - x^2; \tag{8.48}$$

$$f_2(x) := 1 - 2\,|x|\,; \tag{8.49}$$

$$f_3(x) := \begin{cases} 1 - 2\,|x| & -0.5 \le x \le 0.5, \\ 0 & \text{otherwise.} \end{cases} \tag{8.50}$$

The functions f_1, f_2, f_3 have their Fenchel conjugate functions respectively

$$f_1^*(\phi_a) = \sup_{x\in\mathbb{R}} \{\phi_a(x) - f_1(x)\} = \begin{cases} \dfrac{(a+1)^2}{4} & a \ge -1, \\ 0 & a < -1; \end{cases} \tag{8.51}$$

$$f_2^*(\phi_a) = \sup_{x\in\mathbb{R}} \{\phi_a(x) - f_2(x)\} = \begin{cases} +\infty & a \ge 0, \\ -1 - \dfrac{1}{a} & a < 0; \end{cases} \tag{8.52}$$

$$f_3^*(\phi_a) = \sup_{x\in\mathbb{R}} \{\phi_a(x) - f_3(x)\} = \begin{cases} +\infty & a > 0, \\ \dfrac{a}{4} & a \in [-2, 0], \\ -1 - \dfrac{1}{a} & a < -2. \end{cases} \tag{8.53}$$

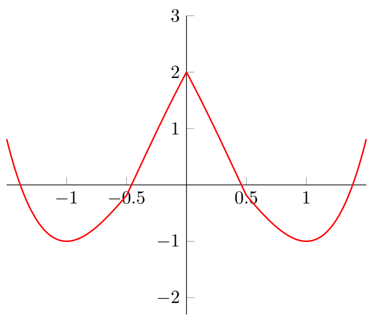


Figure 8.3: Sum function $f_1 + f_2 + f_3$.

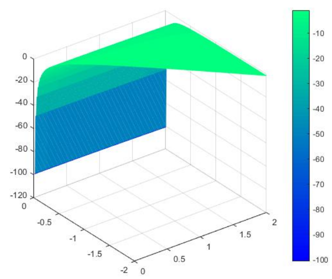


Figure 8.4: The sum of infimal functions $-(f_1^* + f_2^* + f_3^*)$ in the subspace $\phi_{a_1} + \phi_{a_2} + \phi_{a_3} = 0$.

Following Section 3, the minimization problem

$$p := \min\,(f_1 + f_2 + f_3)\,, \tag{(p)}$$

has the dual problem

$$d := \sup_{\phi_{a_1}+\phi_{a_2}+\phi_{a_3}=0} \left(-f_1^*(\phi_{a_1}) - f_2^*(\phi_{a_2}) - f_3^*(\phi_{a_3})\right). \tag{(d)}$$

Note that (d) is a convex problem. Indeed, it is the maximization of a concave function over a subspace. This is always the case because the conjugate functions are suprema of functions which are linear in $\mathcal{L}$. From Figure 8.4.8, we see that $x = \pm 1$ are the global solutions of (p), with optimal value -1. We will obtain these facts as a consequence of Theorem 8.4.11 below.

We will show next that f_1, f_2, f_3 are $\mathcal{H}-$convex and that the zero duality gap holds for problems (p) and (d).

Proposition 8.4.9. *The functions f_1, f_2, f_3 given in Example 8.4.8 are $\mathcal{H}-$convex. The support sets of f_1, f_2, f_3 are*

$$\mathrm{supp} f_1 = \left\{\Psi_{a,b} : b \leq -\frac{(1+a)^2}{4}\right\} \bigcup \left\{\Psi_{a,b} : a \leq -1, b \leq 0\right\}, \tag{8.54}$$

$$\mathrm{supp} f_2 = \left\{\Psi_{a,b} : a \leq 0, b \leq \min_{t\in\mathbb{R}}\{1 - 2|t| - at^2\}\right\}, \tag{8.55}$$

$$\mathrm{supp} f_3 = \left\{\Psi_{a,b} : a \leq 0, b \leq \min_{t\in[-0.5,0.5]}\{1 - 2|t| - at^2\}\right\}. \tag{8.56}$$

Proof. See Section 8.5. □

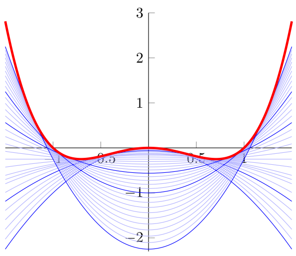


Figure 8.5: Support set of f_1.

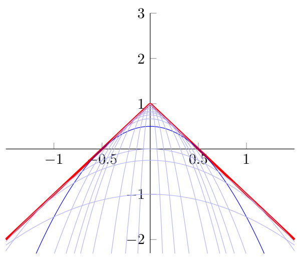


Figure 8.6: Support set of f_2.

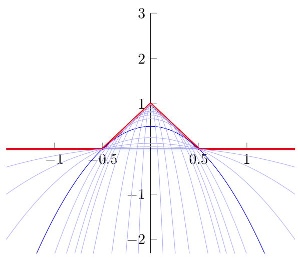


Figure 8.7: Support set of f_3.

Proposition 8.4.10. *The subdifferential operators of the functions f_1, f_2, f_3 as in Example 8.4.8 are as follows*

(i)

$$\partial f_1(x) = \begin{cases} \{\phi_a : a \leq -1\} & x = 0, \\ \{\phi_a : a = 2x^2 - 1\} & \textit{otherwise}; \end{cases} \tag{8.57}$$

(ii)

$$\partial f_2(x) = \begin{cases} \emptyset & x = 0, \\ \left\{\phi_a : a = -\dfrac{1}{|x|}\right\} & \textit{otherwise}; \end{cases} \tag{8.58}$$

(iii)

$$\partial f_3(x) = \begin{cases} \emptyset & x = 0, \\ \left\{\phi_{-\frac{1}{|x|}}\right\} & x \in (-0.5, 0.5) \setminus \{0\}, \\ \{\phi_a : a \in [-2, 0]\} & x = \pm 0.5, \\ \{0\} & \textit{otherwise}. \end{cases} \tag{8.59}$$

Proof. See Section 8.5.

□

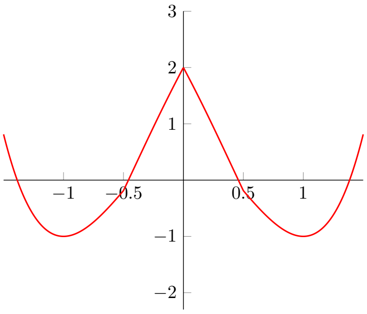


Figure 8.8: Sum function $f_1 + f_2 + f_3$.

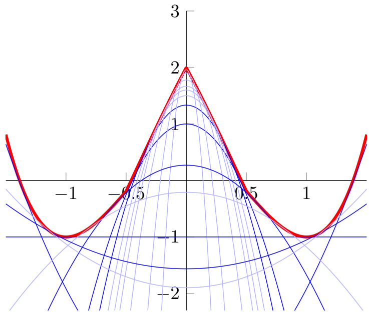


Figure 8.9: Support set of $f_1 + f_2 + f_3$.

Theorem 8.4.11. *Consider problems (p) and (q) as in Example 8.4.8. The zero duality gap of (p) : $f_1 + f_2 + f_3$ and (d) : $-f_1^* - f_2^* - f_3^*$ holds. Moreover, $x = \pm 1$ are the solutions of (p).*

Proof. To prove the claim, we will use Theorem 8.2.7. Namely, we will show that there exists an $x \in \mathbb{R}$ such that (8.32) holds. From Proposition 8.4.10, we observe that

$$0 \in \partial f_1(1) + \partial f_2(1) + \partial f_3(1).$$

So (8.32) holds for $x = 1$. The same expression can analogously be obtained for $x = -1$. By Theorem 8.2.7, we have zero duality gap for (p) and (d). By the last statement in Theorem 8.2.7, $x = \pm 1$ are the solutions of (p). □

8.5 Apendices

In this section, we provide the proofs for Propositions 8.4.9 and 8.4.10.

Proof of Proposition 8.4.9. (i) Take $\Psi_{a,b} \in \mathcal{H}$ with $b \le -\frac{(1+a)^2}{4}$. Then, for all $t \in \mathbb{R}$

$$f_1(t) = t^4 - t^2 = \left(t^2 - \frac{1+a}{2}\right)^2 - \left(\frac{(1+a)^2}{4} + b\right) + (at^2 + b) \ge at^2 + b = \Psi_{a,b}(t).$$

This implies $\Psi_{a,b} \in \operatorname{supp} f_1$. Take $\Psi_{a,b} \in \mathcal{H}$ with $a \le -1, b \le 0$, we also have

$$f_1(t) \ge t^4 + (-a-1)t^2 \ge b, \quad \forall t \in \mathbb{R},$$

which implies $\{\Psi_{a,b} : a \le -1, b \le 0\} \subset \operatorname{supp} f_1$.
Altogether, it yields the inclusion

$$\left\{\Psi_{a,b} : b \le -\frac{(1+a)^2}{4}\right\} \bigcup \{\Psi_{a,b} : a \le -1, b \le 0\} \subset \operatorname{supp} f_1.$$

Take $\Psi_{a',b'} \in \operatorname{supp} f_1$ $(a', b' \in \mathbb{R})$ and assume that

$$\Psi_{a',b'} \notin \left\{\Psi_{a,b} : b \le -\frac{(1+a)^2}{4}\right\} \bigcup \{\Psi_{a,b} : a \le -1, b \le 0\}.$$

This implies $b' + \frac{(1+a')^2}{4} > 0$ and $(a' + 1 > 0$ or $b' > 0)$.
If $b' > 0$, then set $t := 0$. We have $t^4 - (a'+1)t^2 - b' < 0$. So, $\Psi_{a',b'} \notin \operatorname{supp} f_1$.

If $b' \leq 0$, then $a'+1>0$. Set $t := \sqrt{(1+a')/2}$. We have $t^4-(a'+1)t^2-b' = -\dfrac{(a'+1)^2}{4} - b' < 0$. So, $\Psi_{a',b'} \notin \text{supp}f_1$.

In either case, there is a $t \in \mathbb{R}$ such that $t^4 - t^2 < a't^2 + b'$, contradicting the assumption $\Psi_{a',b'} \in \text{supp}f_1$. Hence,

$$\text{supp}f_1 = \left\{\Psi_{a,b} : b \leq -\frac{(1+a)^2}{4}\right\} \bigcup \{\Psi_{a,b} : a \leq -1, b \leq 0\}.$$

It remains to show that $f_1 = \sup\limits_{\Psi \in \text{supp}f_1} \Psi$. For any fixed $t \in \mathbb{R}$, define $a := 2t^2 - 1$, $b := -t^4$. Then, $b = -\dfrac{(a+1)^2}{4}$. By (8.54), $\Psi_{a,b} \in \text{supp}f_1$, and $f_1(t) = t^4 - t^2 = at^2 + b = \Psi_{a,b}(t)$.

(ii) The proof of (8.55) for f_2 can proceed in a similar way as the one in part (iii) below.

(iii) We prove here that f_3 is an $\mathcal{H}-$convex function, and that (8.56) holds. For all $\Psi_{a,b} \in \mathcal{H}$ with $a \leq 0$, and $b \leq \min\limits_{t\in[-0.5,0.5]}\{1-2|t|-at^2\}$, we have

$$b \leq 1-2|t|-at^2, \quad \forall t \in [-0.5, 0.5].$$

This implies $\Psi_{a,b}(t) = at^2 + b \leq 1 - 2|t| = f_3(t)$ for all $t \in [-0.5, 0.5]$.

On the other hand, we also have

$$\Psi_{a,b}(0.5) = \Psi_{a,b}(-0,5) \leq f_3(0.5) = f_3(-0,5) = 0.$$

Since $a \leq 0$, the function $\Psi_{a,b}$ increases on $(-\infty, 0)$, and decreases on $(0, +\infty)$. Hence,

$$\Psi_{a,b}(t) \leq \Psi_{a,b}(-0.5) \leq 0 = f_3(t), \ \forall t \in (-\infty, -0.5),$$
$$\Psi_{a,b}(t) \leq \Psi_{a,b}(0.5) \leq 0 = f_3(t), \ \forall t \in (0.5, +\infty).$$

Altogether, it yields the inclusion

$$\text{supp}f_3 \supset \left\{\Psi_{a,b} : a \leq 0, b \leq \min_{t\in[-0.5,0.5]}\{1-2|t|-at^2\}\right\}.$$

Conversely, take $\Psi_{a,b} \in \text{supp}f_3$.

We must have $a \leq 0$, since otherwise $\lim\limits_{t\to\infty} \Psi_{a,b}(t) = +\infty$, which contradicts $\Psi_{a,b}(t) \leq f_3(t)$ for all $t \in \mathbb{R}$.

If $b > \min\limits_{t\in[-0.5,0.5]}\{1-2|t|-at^2\}$, then there is a $t_0 \in [-0.5, 0.5]$ such that

$$b > 1-2|t_0|-at_0^2 \geq \min_{t\in[-0.5,0.5]}\{1-2|t|-at^2\}.$$

Thus, $\Psi_{a,b}(t_0) = at_0^2 + b > 1 - 2|t_0| = f_3(t_0)$, a contradiction.

Hence, $b \leq \min\limits_{t\in[-0.5,0.5]}\{1-2|t|-at^2\}$. Thus,

$$\text{supp}f_3 \subset \left\{\Psi_{a,b} : a \leq 0, b \leq \min_{t\in[-0.5,0.5]}\{1-2|t|-at^2\}\right\}.$$

Thus, equality (8.56) holds.

Now, we prove that f_3 is $\mathcal{H}-$convex by showing that

$$f_3(x) = \sup_{\Psi\in\text{supp}f_3} \Psi(x), \quad x \in \mathbb{R}. \tag{8.60}$$

Obviously, $f_3(x) \geq \sup\limits_{\Psi\in\text{supp}f_3} \Psi(x), \ \forall x \in \mathbb{R}$.

We aim to show now that $f_3(x) \leq \sup\limits_{\Psi\in\text{supp}f_3} \Psi(x), \ \forall x \in \mathbb{R}$.

1.1 For $x \notin (-0.5, 0.5)$, consider $\Psi \equiv 0 \in \text{supp} f_3$. We have $f_3(x) = \Psi(x) = 0$.

1.2 For $x \in (-0.5, 0.5) \setminus \{0\}$, consider $\Psi_{a,b} \in \text{supp} f_3$ with $a := \dfrac{-1}{|x|} \leq 0$ and

$$b := \min_{t \in [-0.5, 0.5]} \{1 - 2|t| - at^2\} = \min_{t \in [-0.5, 0.5]} \left\{1 - 2|t| + \frac{1}{|x|} t^2\right\} = 1 - |x|.$$

Then, $\Psi_{a,b}(x) = ax^2 + b = 1 - 2|x| = f_3(x)$.

1.3 For $x = 0$, $f_3(0) = 1$. Take the sequence of affine functions $(\Psi_{a_n, b_n}) \subset \text{supp} f_3$, with $a_n := -\dfrac{1}{n}$ and $b_n := 1 - \dfrac{1}{n}$, $n \geq 2$. Observe that $a_n < 0$ and

$$b_n = \min_{t \in [-0.5, 0.5]} \left\{1 - 2|t| - \frac{1}{n} t^2\right\}, \ \forall n \geq 2.$$

Furthermore, we also have

$$f_3(0) = 1 \geq \sup_{\Psi \in \text{supp} f_3} \Psi(0) \geq \lim_{n \to \infty} \Psi_{a_n, b_n}(0) = \lim_{n \to \infty} \left(1 - \frac{1}{n}\right) = 1 = f_3(0).$$

Altogether, (8.60) holds, hence, f_3 is $\mathcal{H}-$convex.

□

proof of Proposition 8.4.10. (i) Take $x \in \mathbb{R}$. By Proposition 8.1.4 and the definition of $\text{supp} f_1$, the linear function $\phi_a \in \partial f_1(x)$ if and only if

$$\phi_a + (f_1(x) - ax^2) \in \text{supp} f_1. \tag{8.61}$$

By (8.54), $f_1(x) - ax^2 \leq -\dfrac{(1+a)^2}{4}$ or $a \leq -1$ and $f_1(x) - ax^2 \leq 0$. If $x = 0$, then $a \leq -1$. If $x \neq 0$, then $x^4 - (a+1)x^2 \leq -\dfrac{(1+a)^2}{4}$, equivalently $\left(x^2 - \dfrac{1+a}{2}\right)^2 \leq 0$, hence $a = 2x^2 - 1$.

(ii) The argument for (8.56) is similar to the argument for (8.59) as below.

(iii) Consider the Fenchel conjugate function

$$f_3^*(\phi_a) = \sup_{x \in \mathbb{R}} \{\phi_a(x) - f_3(x)\} = \begin{cases} +\infty & a > 0, \\ \dfrac{a}{4} & a \in [-2, 0], \\ -1 - \dfrac{1}{a} & a < -2. \end{cases}$$

Then,

$$\phi_a \in \partial f_3(x) \Longleftrightarrow f_3^*(\phi_a) + f_3(x) = \phi_a(x).$$

We consider three cases.

1. $x = 0$, $f_3(x) = 1$. Then, $\phi_a \in \partial f_3(0) \Longleftrightarrow f_3^*(\phi_a) + 1 = 0$. There does not exist $a \in \mathbb{R}$ such that $f_3^*(\phi_a) + 1 = 0$, thus $\partial f_3(0) = \emptyset$.
2. $x \in (-0.5, 0.5) \setminus \{0\}$, $f_3(x) = 1 - 2|x|$. Then, $\phi_a \in \partial f_3(x) \Longleftrightarrow f_3^*(\phi_a) + 1 - 2|x| = \phi_a(x)$. Observe that we must have $a \leq 0$.
   If $a \in [-2, 0]$, then $\dfrac{a}{4} + 1 - 2|x| = \phi_a(x) = ax^2$, or $a(1/4 - x^2) = 2|x| - 1$. This implies $a = \dfrac{2|x| - 1}{1/4 - x^2} = \dfrac{-4}{1 + 2|x|}$. Since $|x| < 0.5$, then $a < -2$, a contradiction.
   If $a < -2$, then

$$-1 - \frac{1}{a} + 1 - 2|x| = ax^2, \text{ equivalently } \quad (ax)^2 + 2a|x| + 1 = 0.$$

   Hence, $a = \dfrac{-1}{|x|}$, and $\phi_a(t) = \dfrac{-1}{|x|} t^2$.

3. $x \notin (-0.5, 0.5)$, $f_3(x) = 0$. Then, $\phi_a \in \partial f_3(x) \iff f_3^*(\phi_a) = ax^2$.

   If $a < -2$, then $-1 - \dfrac{1}{a} = ax^2$, or $-a - 1 = (ax)^2 \geq (0.5a)^2$, which yields $(0.5a + 1)^2 \leq 0$, a contradiction.

   If $a \in [-2, 0]$, then $\dfrac{a}{4} = ax^2$, which only holds true when $a = 0$ or $x = \pm 0.5$.

   Hence, $a \in [-2, 0]$ if $x = \pm 0.5$, and $a = 0$ otherwise.

This establishes (8.59).

□

Part III

Convex Polytopes

Chapter 9

Preliminary Results on Cubical Polytopes

This Preliminary groups a number of results that will be used in later chapters of the thesis.

Unless otherwise stated, the graph theoretical notation and terminology follow from [46] and the polytope theoretical notation and terminology from [175]. Moreover, when referring to graph-theoretical properties of a polytope such as minimum degree, linkedness and connectivity, we mean properties of its graph.

A *cubical* d-polytope is a polytope with all its facets being cubes. By a cube we mean any polytope that is combinatorially equivalent to a cube; that is, one whose face lattice is isomorphic to the face lattice of a cube.

The definitions of polytopal complex and strongly connected complex play an important role in the context. A *polytopal complex* $\mathcal{C}$ is a finite nonempty collection of polytopes in $\mathbb{R}^d$ where the faces of each polytope in $\mathcal{C}$ all belong to $\mathcal{C}$ and where polytopes intersect only at faces (if $P_1 \in \mathcal{C}$ and $P_2 \in \mathcal{C}$ then $P_1 \cap P_2$ is a face of both P_1 and P_2). The empty polytope is always in $\mathcal{C}$. The *dimension* of a complex $\mathcal{C}$ is the largest dimension of a polytope in $\mathcal{C}$; if $\mathcal{C}$ has dimension d we say that C is a *d-complex*. Faces of a complex of largest and second largest dimension are called *facets* and *ridges*, respectively. If each of the faces of a complex $\mathcal{C}$ is contained in some facet we say that $\mathcal{C}$ is *pure*.

Given a polytopal complex $\mathcal{C}$ with vertex set V and a subset X of V, the subcomplex of $\mathcal{C}$ formed by all the faces of $\mathcal{C}$ containing only vertices from X is called *induced* and is denoted by $\mathcal{C}[X]$. Removing from $\mathcal{C}$ all the vertices in a subset $X \subset V(\mathcal{C})$ results in the subcomplex $\mathcal{C}[V(\mathcal{C}) \setminus X]$, which we write as $\mathcal{C} - X$. We say that a subcomplex $\mathcal{C}'$ is a *spanning* subcomplex of $\mathcal{C}$ if $V(\mathcal{C}') = V(\mathcal{C})$. The *graph* of a complex is the undirected graph formed by the vertices and edges of the complex. As in the case of polytopes, we denote the graph of a complex $\mathcal{C}$ by $G(\mathcal{C})$. A pure polytopal complex $\mathcal{C}$ is *strongly connected* if every pair of facets F and F' is connected by a path $F_1 \dots F_n$ of facets in $\mathcal{C}$ such that $F_i \cap F_{i+1}$ is a ridge of $\mathcal{C}$, $F_1 = F$, and $F_n = F'$; we say that such a path is a *$(d-1, d-2)$-path* or a *facet-ridge path* if the dimensions of the faces can be deduced from the context. From the definition, it follows that every 0-complex is trivially strongly connected and that every complex contains a spanning 0-subcomplex.

The relevance of strongly connected complexes stems from the ensuing result of Sallee.

Proposition 9.0.1 ([150, Section 2]). *The graph of a strongly connected d-complex is d-connected.*

Strongly connected complexes can be defined from a d-polytope P. Two basic examples are given by the complex of all faces of P, called the *complex* of P and denoted by $\mathcal{C}(P)$, and the complex of all proper faces of P, called the *boundary complex* of P and denoted by $\mathbb{B}(P)$. For a polytopal complex $\mathcal{C}$, the *star* of a face F of $\mathcal{C}$, denoted $\operatorname{star}(F, \mathcal{C})$, is the subcomplex of $\mathcal{C}$ formed by all the faces containing F, and their faces; the *antistar* of a face F of $\mathcal{C}$, denoted $\operatorname{astar}(F, \mathcal{C})$, is the subcomplex of $\mathcal{C}$ formed by all the faces disjoint from F. That is, $\operatorname{astar}(F, \mathcal{C}) = \mathcal{C} - V(F)$. Unless otherwise stated, when defining stars and antistars in a polytope, we always assume the underlying complex is the boundary complex of the polytope.

Some complexes defined from a d-polytope are strongly connected $(d-1)$-complexes, as the next proposition attests; the parts about the boundary complex and the antistar of a vertex already appeared in [150].

Proposition 9.0.2 ([150, Cor. 2.11, Theorem 3.5])**.** *Let P be a d-polytope. Then, the boundary complex $\mathbb{B}(P)$ of P, and the star and antistar of a vertex in $\mathbb{B}(P)$, are all strongly connected $(d-1)$-complexes of P.*

Proof. Let ψ define the natural anti-isomorphism from the face lattice of P to the face lattice of its dual P^*.

The three complexes are pure. The complex $\mathbb{B}(P)$ is clearly pure, and so is the star of a vertex. Perhaps a sentence may be appropriate for the antistar of a vertex: a face of P that does not contain the vertex must lie in a facet that does not contain the vertex. We proceed to prove the strong connectivity of the complexes.

The statement about $\mathbb{B}(P)$ was already proved in [150, Cor. 2.11]. The facets in $\mathbb{B}(P)$ correspond to vertices in P^*. The existence of a facet-ridge path in $\mathbb{B}(P)$ between any two facets F_1 and F_2 of $\mathbb{B}(P)$ amounts to the existence of a vertex-edge path in P^* between the vertices $\psi(F_1)$ and $\psi(F_2)$ of P^*. That $\mathbb{B}(P)$ is a strongly connected $(d-1)$-complex now follows from the connectivity of the graph of P^* (Balinski's theorem).

The assertion about the star of a vertex does not seem to explicitly appear in [150]. The facets in the star $\mathcal{S}$ of a vertex v in $\mathbb{B}(P)$ correspond to the vertices in the facet $\psi(v)$ in P^*. The existence of a facet-ridge path in $\mathcal{S}$ between any two facets F_1 and F_2 of $\mathcal{S}$ amounts to the existence of a vertex-edge path in $\psi(v)$ between the vertices $\psi(F_1)$ and $\psi(F_2)$ of $\psi(v)$. That $\mathcal{S}$ is a strongly connected $(d-1)$-complex follows from the connectivity of the graph of $\psi(v)$ (Balinski's theorem).

The assertion about the antistar of a vertex v was first shown in [150, Theorem 3.5]. The facets in $\operatorname{astar}(v)$ correspond to the vertices of P^* that are not in $\psi(v)$. That is, if F_1 and F_2 are any two facets of $\operatorname{astar}(v)$, then $\psi(F_1), \psi(F_2) \in V(P^*) \setminus V(\psi(v))$. The existence of a facet-ridge path between F_1 and F_2 in $\operatorname{astar}(v)$ amounts to the existence of a vertex-edge path between $\psi(F_1)$ and $\psi(F_2)$ in the subgraph $G(P^*) - V(\psi(v))$ of $G(P^*)$. The removal of the vertices of a facet does not disconnect the graph of a polytope [150, Theorem 3.1], wherefrom it follows that $G(P^*) - V(\psi(v))$ is connected, as desired. □

The next two propositions follow from the characterisation of 2-linked graphs carried out in [153,161]. Both propositions also have proofs stemming from arguments in the form of Lemma 9.0.3, a lemma used implicitly in the original proof of Balinski's theorem (Theorem 9.0.4) and made explicit in [150, Theorem 3.1]; for the sake of completeness we give such proofs.

Lemma 9.0.3 ([150, Theorem 3.1])**.** *Let P be a d-polytope, and let f be a linear function on $\mathbb{R}^d$ satisfying $f(x) > 0$ for some $x \in P$. If u and v are vertices of P with $f(u) \geq 0$ and $f(v) \geq 0$, then there exists a $u-v$ path $x_0x_1 \ldots x_n$ with $x_0 = u$ and $x_n = v$ such that $f(x_i) > 0$ for $i \in [1, n-1]$.*

Theorem 9.0.4 (Balinski [9])**.** *For every $d \geq 1$, the graph of a d-polytope is d-connected.*

Let X be a set of vertices in a graph G. A path in the graph is called *X-valid* if no inner vertex of the path is in X. A sequence $a_1, \ldots, a_n$ of vertices in a cycle is in *cyclic order* if, while traversing the cycle, the sequence appears in clockwise or counterclockwise order.

Proposition 9.0.5. *Let G be the graph of a 3-polytope and let X be a set of four vertices of G. The set X is linked in G if and only if there is no facet of the polytope containing all the vertices of X.*

Proof. Let P be a 3-polytope embedded in $\mathbb{R}^3$ and let X be an arbitrary set of four vertices in G. We first establish the necessary condition by proving the contrapositive. Let F be a 2-face containing the vertices of X and consider a planar embedding of G in which F is the outer face. Label the vertices of X so that they appear in the cyclic order $s_1s_2t_1t_2$. Then $s_1 - t_1$ and $s_2 - t_2$ paths in G must inevitably intersect, implying that X is not linked.

Assume there is no 2-face of P containing all the vertices of X. Let H be a (linear) hyperplane that contains s_1, s_2 and t_1, and let f be a linear function that vanishes on H (this may require a translation of the polytope). Without loss of generality, assume that $f(x) > 0$ for some $x \in P$ and that $f(t_2) \geq 0$.

First consider the case that H is a supporting hyperplane of a 2-face F. The subgraph $G(F) - \{s_2\}$ is connected, and so there is an X-valid $L_1 := s_1 - t_1$ path on $G(F)$. Then, use Lemma 9.0.3 to find an $L_2 := s_2 - t_2$ path in which each inner vertex has positive f-value. The paths L_1 and L_2 are clearly disjoint.

Now consider the case that H intersects the interior of P. Then there is a vertex in P with f-value greater than zero and a vertex with f-value less than zero. Use Lemma 9.0.3 to find an $s_1 - t_1$ path in which each inner vertex has negative f-value and an $s_2 - t_2$ path in which each inner vertex has positive f-value. □

The subsequent corollary follows at once from Proposition 9.0.5.

Corollary 9.0.6. *No nonsimplicial 3-polytope is 2-linked.*

Proposition 9.0.5 motivates the ensuing definition.

Definition 9.0.7 (Configuration F)**.** Let $Y := \{\{s_1, t_1\}, \{s_2, t_2\}\}$ be a labelling and pairing of four vertices in a 3-cube. *Configuration F* is a configuration in the cube where the vertices of Y appear in cyclic order $s_1s_2t_1t_2$ in a 2-face.

Configuration F is the only configuration in a 3-cube that prevents the linkedness of a pairing Y of four vertices.

The same reasoning employed in the proof of the sufficient condition of Proposition 9.0.5 settles Proposition 9.0.8.

Proposition 9.0.8 (2-linkedness of 4-polytopes)**.** *Every 4-polytope is 2-linked.*

Proof. Let G be the graph of a 4-polytope embedded in $\mathbb{R}^4$. Let X be a given set of four vertices in G and let $Y := \{\{s_1, s_2\}, \{t_1, t_2\}\}$ a labelling and pairing of the vertices in X.

Consider a linear function f that vanishes on a linear hyperplane H passing through X. Consider the two cases in which either H is a supporting hyperplane of a facet F of P or H intersects the interior of P.

Suppose H is a supporting hyperplane of a facet F. First, find an $s_1 - t_1$ path in the subgraph $G(F) - \{s_2, t_2\}$, which is connected by Balinski's theorem. Second, use Lemma 9.0.3 to find an $s_2 - t_2$ path that touches F only at $\{s_2, t_2\}$.

If instead H intersects the interior of P then there is a vertex in P with f-value greater than zero and a vertex with f-value less than zero. Use Lemma 9.0.3 to find an $s_1 - t_1$ path in which each inner vertex has negative f-value and an $s_2 - t_2$ path in which each inner vertex has positive f-value. □

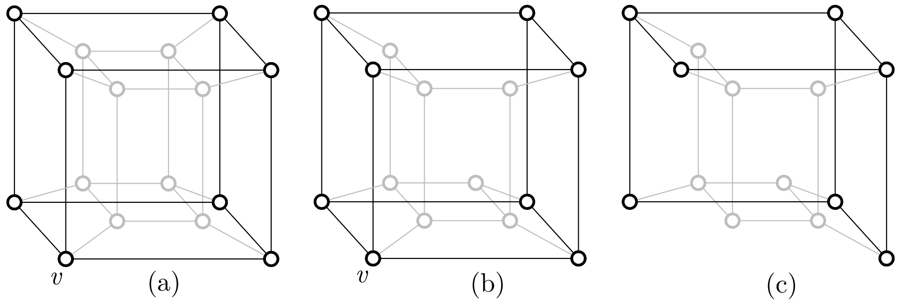


Figure 9.1: Complexes in the 4-cube. (a) The 4-cube with a vertex v highlighted. (b) The star of the vertex v. (c) The link of the vertex v.

The *distance* between two vertices s and t in a graph G, denoted $\operatorname{dist}_G(s,t)$, is the length of a shortest path between the vertices. Let v be a vertex in a d-cube Q_d and let v^o denote the vertex at distance d from v, called the vertex *opposite* to v. The star of a vertex v in the boundary complex of a d-cube Q_d is the subcomplex $Q_d - v^o$, the subcomplex induced by $V(Q_d) \setminus \{v^o\}$.

Remark 9.0.9. The antistar of v coincides with the star of v^o. Consequently, the link of v in a d-cube Q_d is the subcomplex $Q_d - \{v, v^o\}$.

Figure 9.1 depicts the star and link of a vertex in the 4-cube.

By considering a point v' in $\mathbb{R}^d$ beyond a vertex v of a d-polytope P and using [62, Theorem 5.2.1], we get a statement similar to Proposition 9.0.2 for the link of a vertex in $\mathbb{B}(P)$: Proposition 9.0.11. We provide all the details, for the sake of completeness.

Following [175, pp. 78, 241], we say that a facet F is *visible* from a point v' in $\mathbb{R}^d \setminus P$ if v' belongs to the open halfspace that is determined by aff F, the affine hull of the facet, and is disjoint from P; if instead v' belongs to the open halfspace that contains the interior of P, we say that the facet is *nonvisible* from v'. Further we say that a point v' in $\mathbb{R}^d$ is *beyond* a face K of P if the facets containing K are precisely those visible from v'.

Theorem 9.0.10 ([62, Theorem 5.2.1]). *Let P and P' be two d-polytopes in $\mathbb{R}^d$, and let v' be a vertex of P' such that $v' \notin P$ and $P' = \operatorname{conv}(P \cup \{v'\})$. Then*

(i) a face F of P is a face of P' if and only if there exists a facet of P containing F that is nonvisible from v;

(ii) if F is a face of P then $F' := \operatorname{conv}(F \cup \{v'\})$ is a face of P' if

(a) either $v' \in \operatorname{aff} F$;

(b) or among the facets of P containing F there is at least one that is visible from v' and at least one that is nonvisible.

Moreover, each face of P' is of exactly one of the above three types.

Proposition 9.0.11 ([175, Ex. 8.6]). *Let P be a d-polytope. Then the link of a vertex in $\mathbb{B}(P)$ is combinatorially equivalent to the boundary complex of a $(d-1)$-polytope.*

Proof. Let v be a vertex of P and let v' be a point in $\mathbb{R}^d \setminus P$ beyond v so that v' is not on the affine hull of any face of P. Suppose $P' := \operatorname{conv}(P \cup \{v'\})$.

The facets in the star of v in $\mathbb{B}(P)$ are precisely those that are visible from v', and every other facet of P, including the facets in the antistar of v in $\mathbb{B}(P)$, is nonvisible from v'. The link of v is, by

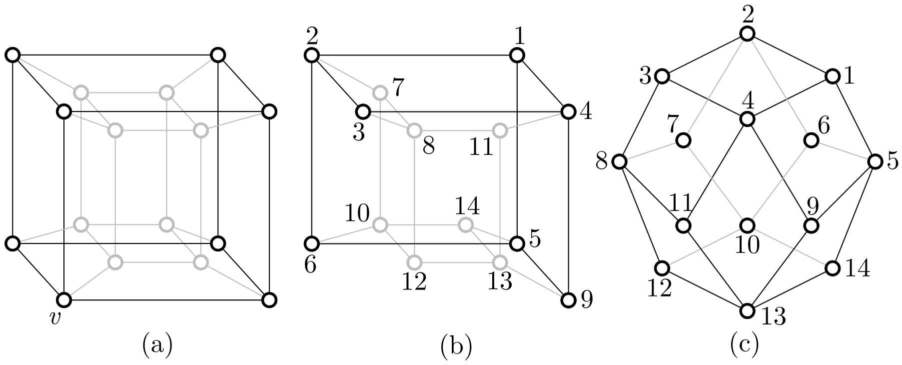


Figure 9.2: The link of a vertex in the 4-cube. (a) The 4-cube with a vertex v highlighted. (b) The link of the vertex v in the 4-cube. (c) The link of the vertex v as the boundary complex of the rhombic dodecahedron (Proposition 9.0.11).

definition, the subcomplex of $\mathbb{B}(P)$ induced by the ridges of P that are contained in a facet of the star of v, a facet visible from v', and a facet of the antistar of v, a facet nonvisible from v'. Consequently, according to Theorem 9.0.10(i), the ridges in $\text{link}(v, \mathbb{B}(P))$ are faces of P'. Furthermore, for every ridge $R \in \text{link}(v, \mathbb{B}(P))$, $R' := \text{conv}\,(R \cup \{v'\})$ is a facet of P' (Theorem 9.0.10(ii-b)), a pyramid over R with apex v'; and every facet in the star of v' in $\mathbb{B}(P')$ is one of these pyramids. Hence, the vertex figure of P' at v', which is a $(d-1)$-polytope [175, Section 2.1], is combinatorially equivalent to the link of v in P, as desired. □

Proposition 9.0.11 is exemplified in Fig. 9.2.

Chapter 10

Connectivity of Cubical Polytopes

Define a *separator* of a polytope as a set of vertices disconnecting the graph of the polytope. Let X be a set of vertices in a graph G. Denote by $G[X]$ the subgraph of G induced by X, the subgraph of G that contains all the edges of G with vertices in X. Write $G - X$ for $G[V(G) \setminus X]$; that is, the subgraph $G - X$ is obtained by removing the vertices in X and their incident edges. The graph G is $k-$connected if and only if the minimum separator is of cardinality k.

10.1 Connectivity of the d-Cube

We unveil some further properties of the cube, whose proofs exploit the realisation of a d-cube as a $0-1$ d-polytope [176]. A $0-1$ *d-polytope* is a d-polytope whose vertices have coordinates in $\{0,1\}^d$. Here $\{0,1\}^d$ denotes the set of all d-element sequences from $\{0,1\}$.

We next give some basic properties of the d-cube, including some specific to its realisation as a $0-1$ polytope.

Remark 10.1.1 (Basic properties of the d-cube). Let $\vec{x} = (x_1, \ldots, x_d)$ with $x_i \in \{0,1\}$ be a vertex of the $0-1$ d-cube Q_d.

(i) Every two facets of Q_d either intersect at a ridge or are disjoint.

(ii) Each of the $2d$ facets of Q_d is the convex hull of a set of the form

$$F_i^0 := \operatorname{conv}\{\vec{x} \in V(Q_d) : x_i = 0\} \text{ or } F_i^1 := \operatorname{conv}\{\vec{x} \in V(Q_d) : x_i = 1\},$$

for i in $[1,d]$, the interval $1, \ldots, d$.

(iii) A $(d-k)$-face is the intersection of exactly k facets, and thus, its vertices have the form

$$\{\vec{x} \in V(Q_d) : x_{i_1} = 0, \ldots, x_{i_r} = 0, x_{i_{r+1}} = 1, \ldots, x_{i_k} = 1\}$$

for $k \in [1,d]$ and $r \in [0,k]$.

While it is true that the antistar of a vertex in a d-polytope is always a strongly connected $(d-1)$-complex (Proposition 9.0.2), it is far from true that this extends to higher dimensional faces. Consider any d-polytope P with a simplex facet J that contains at least one vertex v of degree d in P. Let F be any face in J that does not contain v. Then the vertex v has degree $d - |V(F)|$ in the subcomplex $P - V(F)$. Since every vertex in a pure $(d-1)$-complex has degree at least $d-1$, the antistar of F in $\mathbb{B}(P)$, which contains v, cannot be a pure $(d-1)$-complex for $\dim F \geq 1$. This extension is however possible for the d-cube.

Lemma 10.1.2. *Let F be a proper face in the d-cube Q_d. Then the antistar of F is a strongly connected $(d-1)$-complex.*

Proof. Without loss of generality, assume that Q_d is given as a $0-1$ polytope, and for the sake of concreteness, that our proper face F is defined as $\operatorname{conv}\{\vec{x} \in V(Q_d) : x_1 = 0, \ldots, x_k = 0\}$ (Remark 10.1.1(iii)). That is, $F = F_1^0 \cap \cdots \cap F_k^0$; refer to Remark 10.1.1(ii).

We claim that the antistar of F is the pure $(d-1)$-complex

$$\mathcal{C} := \mathcal{C}(F_1^1) \cup \cdots \cup \mathcal{C}(F_k^1);$$

refer to Remark 10.1.1(ii)-(iii).

We proceed by proving that $\operatorname{astar}(F, Q_d) \subseteq \mathcal{C}$. Take any $(d-l)$-face $K \notin \mathcal{C}$. Then $K = J_1 \cap \cdots \cap J_l$ for some facets J_i of Q_d. A facet J_i is defined by either $\operatorname{conv}\{\vec{x} : V(Q_d) : x_j = 1 \text{ for some } j \in [k+1, d]\}$ or $\operatorname{conv}\{\vec{x} : V(Q_d) : x_j = 0 \text{ for some } j \in [1, d]\}$. According to Remark 10.1.1(iii), for $l \in [1, d]$ and $r \in [0, l]$, we get that

$$K = \operatorname{conv}\{\vec{x} \in V(Q_d) : x_{i_1} = 0, \ldots, x_{i_r} = 0, x_{i_{r+1}} = 1, \ldots, x_{i_l} = 1\}.$$

From the form of the facets J_i it follows that $i_j \geq k+1$ for all $j \in [r+1, l]$. Hence there is a vertex $\vec{x} = (x_1, \ldots, x_d)$ in K satisfying $x_1 = \cdots = x_k = 0$, which implies that $\{\vec{x}\} \subset K \cap F$. That is, $K \notin \operatorname{astar}(F, Q_d)$.

To prove that $\mathcal{C} \subseteq \operatorname{astar}(F, Q_d)$ holds, observe that, if $K \in \mathcal{C}$ then it is in a facet F_i^1 for some $i \in [1, k]$, and therefore, it belongs to $\operatorname{astar}(F, Q_d)$. Hence $\mathcal{C} = \operatorname{astar}(F, Q_d)$.

That $\mathcal{C}$ is strongly connected follows from noting that the facets $F_1^1, \ldots, F_k^1$ are pairwise nondisjoint, and therefore, pairwise intersect at $(d-2)$-faces (Remark 10.1.1(i)). ☐

Proposition 10.1.3 is well known [138, Proposition 1], but we are not aware of a reference for Proposition 10.1.4.

Proposition 10.1.3 ([138, Proposition 1]). *Any separator X of cardinality d in Q_d consists of the d neighbours of some vertex in the cube, and the subgraph $G(Q_d) - X$ has exactly two components, with one of them being the vertex itself.*

Proof. A proof can be found in [138, Proposition 1]: essentially, one proceeds by induction on d, considering the effect of the separator on a pair of disjoint facets. ☐

Proposition 10.1.4. *Let y be a vertex of the d-cube Q_d and let Y be a subset of the neighbours of y in Q_d. Then the subcomplex of Q_d induced by $V(Q_d) \setminus (\{y\} \cup Y)$ contains a spanning strongly connected $(d-2)$-subcomplex.*

Proof. Without loss of generality, assume that Q_d is given as a $0-1$ polytope, and for the sake of concreteness, that $y = (0, \ldots, 0)$ and $Y = \{\vec{e}_1, \ldots, \vec{e}_k\}$ where $\vec{e}_i$ denotes the standard unit vector with the i-entry equal to one.

Let $\mathcal{C} := Q_d - (\{y\} \cup Y)$, the subcomplex of Q_d induced by $V(Q_d) \setminus (\{y\} \cup Y)$. Consider the $d-k$ ridges

$$R_i := \operatorname{conv}\{\vec{x} \in V(Q_d) : x_1 = 0, x_i = 1\} \text{ for } i \in [k+1, d]$$

and the $\binom{d}{2}$ ridges

$$R_{i,j} := \operatorname{conv}\{\vec{x} \in V(Q_d) : x_i = 1, x_j = 1\} \text{for some } i, j \in [1, d] \text{ with } i \neq j.$$

Let $\mathcal{C}' := \mathcal{C}(R_{k+1}) \cup \cdots \cup \mathcal{C}(R_d) \cup \mathcal{C}(R_{1,2}) \cup \cdots \cup \mathcal{C}(R_{d-1,d})$. Then $\mathcal{C}'$ is a pure $(d-2)$-subcomplex of $\mathcal{C}$.

We show that $\mathcal{C}'$ is a spanning subcomplex of $\mathcal{C}$. Let $\vec{x} = (x_1, \ldots, x_d)$ be a vertex in $V(Q_d) \setminus (\{y\} \cup Y)$. Then either $\vec{x} = \vec{e}_i$ for some $i = k+1, \ldots, d$ or $x_i = x_j = 1$ for some $i, j \in [1, d]$ with $i \neq j$; see Remark 10.1.1(iii). In the former case, the vertex $\vec{x}$ lies in the $(d-2)$-face R_i, and in the latter case, the vertex $\vec{x}$ lies in the $(d-2)$-face $R_{i,j}$. Therefore $\vec{x} \in \mathcal{C}'$. We next show that $\mathcal{C}'$ is strongly connected.

Take any two distinct ridges R and R' from $\mathcal{C}'$. We consider three cases based on the form of R and R'.

Suppose that $R = R_i$ and $R' = R_j$ for $i, j \in [k+1, d]$ and $i \neq j$. Then there is a $(d-2, d-3)$-path L of length one from R to R' through their common $(d-3)$-face $\operatorname{conv}\{\vec{x} \in V(Q_d) : x_1 = 0, x_i = 1, x_j = 1\}$. That is, $L := RR'$.

Next suppose that $R = R_i$ and $R' = R_{j,l}$ for $j, l \in [1, d]$ with $j \neq l$. If $i = j$ there is a $(d-2, d-3)$-path of length one from R to R' through the common $(d-3)$-face $\operatorname{conv}\{\vec{x} \in V(Q_d) : x_1 = 0, x_i = 1, x_l = 1\}$. If $i \neq j$, and consequently $i \neq l$, then there is a $(d-2, d-3)$-path L of length two from R to R' through the $(d-2)$ face $R_{i,l}$, which shares the $(d-3)$-face $\operatorname{conv}\{\vec{x} \in V(Q_d) : x_1 = 0, x_i = 1, x_l = 1\}$ with R and the $(d-3)$-face $\operatorname{conv}\{\vec{x} \in V(Q_d) : x_i = 1, x_j = 1, x_l = 1\}$ with R'. That is, $L := RR_{i,l}R'$.

Finally suppose that $R = R_{i,j}$ with $i \neq j$ and $R' = R_{l,m}$ with $l \neq m$. If $i = l$ there is a $(d-2, d-3)$-path from R to R' through the common $(d-3)$-face $\operatorname{conv}\{\vec{x} \in V(Q_d) : x_i = 1, x_j = 1, x_m = 1\}$. If $\{i, j\} \cap \{l, m\} = \emptyset$ then there is a $(d-2, d-3)$-path L of length two from R to R' through the $(d-2)$-face $R_{i,l}$, which shares the $(d-3)$-face $\operatorname{conv}\{\vec{x} \in V(Q_d) : x_i = 1, x_j = 1, x_l = 1\}$ with R and the $(d-3)$-face $\operatorname{conv}\{\vec{x} \in V(Q_d) : x_i = 1, x_l = 1, x_m = 1\}$ with R'. That is, $L := RR_{i,l}R'$. □

Remark 10.1.5. In Proposition 10.1.4, the subcomplex of Q_d induced by $V(Q_d) \setminus (\{y\} \cup Y)$, in the proof of Proposition 10.1.4 denoted by $\mathcal{C}$, is pure if and only if Y is the set of all neighbours of y. Let $Y = \{\vec{e}_1, \ldots, \vec{e}_k\}$ and $y = (0, \ldots, 0)$. If $k < d$ then the facets $\operatorname{conv}\{\vec{x} \in V(Q_d) : x_\ell = 1\}$ for $\ell \in [k+1, d]$ are in $\mathcal{C}$ and the ridge $\operatorname{conv}\{\vec{x} \in V(Q_d) : x_i = 1, x_j = 1\}$ for $i, j \in [1, k]$ and $i \neq j$ is in $\mathcal{C}$ but the facets $\operatorname{conv}\{\vec{x} \in V(Q_d) : x_i = 1\}$ and $\operatorname{conv}\{\vec{x} \in V(Q_d) : x_j = 1\}$ are not in $\mathcal{C}$. Thus $\mathcal{C}$ is nonpure. If instead $k = d$ then no facet is in $\mathcal{C}$, and the vector coordinates of every vertex in $\mathcal{C}$ has at least two entries with ones, and thus, it is contained in some ridge $\operatorname{conv}\{\vec{x} \in V(Q_d) : x_i = 1, x_j = 1\}$ for $i, j \in [1, d]$ and $i \neq j$, which is in $\mathcal{C}$. Thus $\mathcal{C}$ is a pure $(d-2)$-subcomplex of Q_d, and it coincides with the complex $\mathcal{C}'$. Figure 10.1 illustrates Proposition 10.1.4.

10.2 Connectivity of Cubical Polytopes

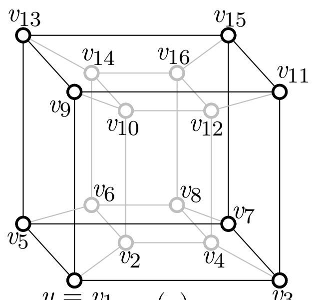
The aim of this section is to prove Theorem 10.2.8, a result that relates the connectivity of a cubical polytope to its minimum degree.

Two vertex-edge paths are *independent* if they share no inner vertex. Similarly, two facet-ridge paths are *independent* if they do not share an inner facet.

Given sets A, B of vertices in a graph, a path from A to B, called an $A-B$ *path*, is a (vertex-edge) path $L := u_0 \ldots u_n$ in the graph such that $V(L) \cap A = \{u_0\}$ and $V(L) \cap B = \{u_n\}$. We write $a-B$ path instead of $\{a\}-B$ path, and likewise, write $A-b$ path instead of $A-\{b\}$.

Our exploration of the connectivity of cubical polytopes starts with a statement about the connectivity of the star of a vertex. But first we need a lemma that holds for all d-polytopes.

Lemma 10.2.1. *Let P be a d-polytope with $d \geq 2$. Then, for any two distinct facets F_1 and F_2 of P, the following hold.*



$v_1 = (0,0,0,0)$	$v_9 = (0,0,0,1)$
$v_2 = (1,0,0,0)$	$v_{10} = (1,0,0,1)$
$v_3 = (0,1,0,0)$	$v_{11} = (0,1,0,1)$
$v_4 = (1,1,0,0)$	$v_{12} = (1,1,0,1)$
$v_5 = (0,0,1,0)$	$v_{13} = (0,0,1,1)$
$v_6 = (1,0,1,0)$	$v_{14} = (1,0,1,1)$
$v_7 = (0,1,1,0)$	$v_{15} = (0,1,1,1)$
$v_8 = (1,1,1,0)$	$v_{16} = (1,1,1,1)$

(b)

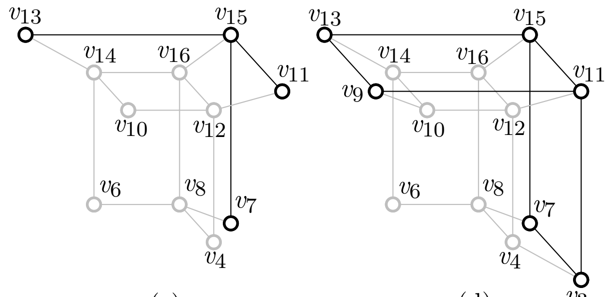


Figure 10.1: Complexes in the 4-cube. (a) The 4-cube with the vertex $y = (0,0,0,0)$ singled out. The vertex labelling corresponds to a realisation of the 4-cube as a $0-1$ polytope. (b) Vertex coordinates as elements of $\{0,1\}^4$. (c) The strongly connected 2-complex $\mathcal{C}$ induced by $V(Q_4) \setminus (\{y\} \cup Y)$ where $Y = \{v_2, v_3, v_5, v_9\}$. Every face of $\mathcal{C}$ is contained in a 2-face of the cube. (d) The non-pure complex $\mathcal{C}$ induced by $V(Q_4) \setminus (\{y\} \cup Y)$ where $Y = \{v_2 = \vec{e}_1, v_5 = \vec{e}_3\}$. The 2-face $\operatorname{conv}\{v_6, v_8, v_{14}, v_{16}\} = \operatorname{conv}\{\vec{x} \in V(Q_4) : x_1 = 1, x_3 = 1\}$ of $\mathcal{C}$ is not contained in any 3-face, and there are two 3-faces in $\mathcal{C}$, namely $\operatorname{conv}\{v_3, v_4, v_7, v_8, v_{11}, v_{12}, v_{15}, v_{16}\} = \operatorname{conv}\{\vec{x} \in V(Q_4) : x_2 = 1\}$ and $\operatorname{conv}\{v_9, v_{10}, v_{11}, v_{12}, v_{13}, v_{14}, v_{15}, v_{16}\} = \operatorname{conv}\{\vec{x} \in V(Q_4) : x_4 = 1\}$.

(i) There are d independent facet-ridge paths between F_1 and F_2 in P.

(ii) Let $\mathcal{S}$ be the star of a vertex and let F be a facet of $\mathcal{S}$. If F_1 and F_2 are in $\mathcal{S}$ and are both different from F, then there exists a $(d-1, d-2)$-path between F_1 and F_2 in $\mathcal{S}$ that does not contain F.

(iii) Let F be a facet of P other than F_1 and F_2. Then there exists a $(d-1, d-2)$-path between F_1 and F_2 in P that does not contain F.

(iv) Let R be an arbitrary ridge of P. Then there exists a facet-ridge path $J_1 \dots J_m$ with $J_1 = F_1$ and $J_m = F_2$ in P such that $J_\ell \cap J_{\ell+1} \neq R$ for each $\ell \in [1, m-1]$.

Proof. The proof of the lemma essentially follows from dualising Balinski's theorem.

Let ψ define the natural anti-isomorphism from the face lattice of P to the face lattice of its dual P^*.

(i). Any two independent vertex-edge paths in P^* between the vertices $\psi(F_1)$ and $\psi(F_2)$ correspond to two independent facet-ridge paths in P between the facets F_1 and F_2. By Balinski's theorem there are d independent $\psi(F_1) - \psi(F_2)$ paths in P^*, and so the assertion follows.

(ii). The facets in the star $\mathcal{S}$ of a vertex s in $\mathbb{B}(P)$ correspond to the vertices in the facet $\psi(s)$ in P^* corresponding to s. The existence of a facet-ridge path in $\mathcal{S}$ between any two facets F_1 and F_2 of $\mathcal{S}$ amounts to the existence of a vertex-edge path in $\psi(s)$ between the vertices $\psi(F_1)$ and $\psi(F_2)$ of $\psi(s)$. Since the graph of the facet $\psi(s)$ is $(d-1)$-connected (Balinski's theorem), by Menger's theorem ([117]; see also [46, Section 3.3]) there are $d-1$ independent paths between $\psi(F_1)$ and $\psi(F_2)$. Hence we can pick one such path L^* that avoids the vertex $\psi(F)$ of $\psi(s)$. Dualising this path L^* gives a $(d-1, d-2)$-path L between F_1 and F_2 in the star $\mathcal{S}$ that does not contain the facet F of P.

(iii). By (i) there are d independent facet-ridge paths between F_1 and F_2 in P, and since $d \geq 2$, we can pick one such path that does not contain F.

(iv). Again by (i), there are d independent facet-ridge paths between F_1 and F_2 in P, and since $d \geq 2$ and the ridge R can be present in at most one such path, there must exist a facet-ridge path that does not contain R. The assertion now follows. ☐

For a path $L := u_0 \dots u_n$ we write $u_i L u_j$ for $0 \leq i \leq j \leq n$ to denote the subpath $u_i \dots u_j$.

Proposition 10.2.2. *Let F be a facet in the star $\mathcal{S}$ of a vertex in a cubical d-polytope. Then the antistar of F in $\mathcal{S}$ is a strongly connected $(d-2)$-complex.*

Proof. Let s be a vertex of a facet F in a cubical d-polytope P and let $F_1, \dots, F_n$ be the facets in the star $\mathcal{S}$ of the vertex s. Let $F_1 = F$. The result is true for $d = 2$: the antistar of F is just a vertex, a strongly connected 0-complex. So assume $d \geq 3$.

According to Lemma 10.1.2, the antistar of $F_i \cap F_1$ in F_i, the subcomplex of F_i induced by $V(F_i) \setminus V(F_i \cap F_1)$, is a strongly connected $(d-2)$-complex for each $i \in [2, n]$. Since

$$\operatorname{astar}(F_1, \mathcal{S}) = \bigcup_{i=2}^{n} \operatorname{astar}(F_i \cap F_1, F_i),$$

it follows that $\operatorname{astar}(F_1, \mathcal{S})$ is a pure $(d-2)$-complex. It remains to prove that there exists a $(d-2, d-3)$-path L between any two ridges R_i and R_j in $\operatorname{astar}(F_1, \mathcal{S})$.

By virtue of Lemma 10.1.2, we can assume that $R_i \in \operatorname{astar}(F_i \cap F_1, F_i)$ and $R_j \in \operatorname{astar}(F_j \cap F_1, F_j)$ for $i \neq j$ and $i, j \in [2, n]$. Since $\mathcal{S}$ is a strongly connected $(d-1)$-complex (Proposition 9.0.2), there exists a $(d-1, d-2)$-path $M := J_1 \dots J_m$ in $\mathcal{S}$, where $J_\ell \cap J_{\ell+1}$ is a ridge for $\ell \in [1, m-1]$, $J_1 = F_i$ and $J_m = F_j$. Let $E_0 := R_i$ and $E_m := R_j$.

We can assume the path M doesn't contain F_1 (Lemma 10.2.1(ii)). Let us show that the path L exists, by proving the following statement by induction.

Claim 1. *If $\ell \leq m$, there exists a $(d-2, d-3)$ path in $\bigcup_{i=1}^{\ell} \operatorname{astar}(J_i \cap F_1, J_i)$ between $E_0 \in \operatorname{astar}(J_1 \cap F_1, J_1)$ and any ridge $E_\ell \in \operatorname{astar}(J_\ell \cap F_1, J_\ell)$.*

Proof. The statement is true for $\ell = 1$. The complex $\operatorname{astar}(J_1 \cap F_1, J_1)$ is a strongly connected $(d-2)$-complex and contains E_0. So an induction on ℓ can start.

Suppose that the statement is true for some $\ell < m$. We show the existence of a $(d-2, d-3)$-path between E_0 and any ridge of $\operatorname{astar}(J_{\ell+1} \cap F_1, J_{\ell+1})$.

Let E_ℓ be a ridge in $\operatorname{astar}(J_\ell \cap F_1, J_\ell)$ such that E_ℓ contains a $(d-3)$-face I_ℓ of $\operatorname{astar}(J_\ell \cap F_1, J_\ell) \cap \operatorname{astar}(J_{\ell+1} \cap F_1, J_{\ell+1})$. By the induction hypothesis there exists a $(d-2, d-3)$ path L_ℓ in $\bigcup_{i=1}^{\ell} \operatorname{astar}(J_i \cap F_1, J_i)$ between E_0 and E_ℓ.

Consider a ridge $E'_{\ell+1} \in \operatorname{astar}(J_{\ell+1} \cap F_1, J_{\ell+1})$ such that $E'_{\ell+1}$ contains the aforementioned $(d-3)$-face I_ℓ. There is a $(d-2, d-3)$-path $L'_{\ell+1}$ in $\operatorname{astar}(J_{\ell+1} \cap F_1, J_{\ell+1})$ from $E'_{\ell+1}$ to any ridge $E_{\ell+1}$ in $\operatorname{astar}(J_{\ell+1} \cap F_1, J_{\ell+1})$, thanks to $\operatorname{astar}(J_{\ell+1} \cap F_1, J_{\ell+1})$ being a strongly connected $(d-2)$-complex.

Since E_ℓ and $E'_{\ell+1}$ share I_ℓ, a path $L_{\ell+1}$ from E_0 to the arbitrary ridge $E_{\ell+1}$ is obtained as $L_{\ell+1} = E_0 L_\ell E_\ell E'_{\ell+1} L'_{\ell+1} E_{\ell+1}$.

For this concatenation to work it remains to prove that the complex $\operatorname{astar}(J_\ell \cap F_1, J_\ell) \cap \operatorname{astar}(J_{\ell+1} \cap F_1, J_{\ell+1})$ contains the aforementioned $(d-3)$-face I_ℓ. Let $K_\ell := J_\ell \cap J_{\ell+1} \cap F_1$. Because $J_\ell \cap J_{\ell+1}$ is a ridge but not of F_1 and because $\{s\} \subseteq V(J_\ell) \cap V(J_{\ell+1}) \cap V(F_1)$, we find that $0 \leq \dim K_\ell \leq d-3$. From $J_\ell \cap J_{\ell+1}$ being a $(d-2)$-cube and $\dim K_\ell \leq d-3$ follows the existence of a $(d-3)$-face in $J_\ell \cap J_{\ell+1}$ that is disjoint from F_1, our I_ℓ. As a consequence, this face $I_\ell \in \operatorname{astar}(J_\ell \cap F_1, J_\ell) \cap \operatorname{astar}(J_{\ell+1} \cap F_1, J_{\ell+1})$, as desired. ☐

Applying the claim to $\ell = m$ gives the existence of a path in $\cup_{i=1}^{m} \operatorname{astar}(J_i \cap F_1, J_i)$ between $E_0 = R_i$ and $E_m = R_j$; this is the desired path L. ☐

The proof method used in Proposition 10.2.2 also proves the following.

Theorem 10.2.3. *Let F be a proper face of a cubical d-polytope P. Then the antistar of F in P contains a spanning strongly connected $(d-2)$-subcomplex.*

Proof. Let $F_1, \ldots, F_n$ be the facets of P and let F be a proper face of P. The result is true for $d = 2$: the antistar of F is a strongly connected 1-complex, and thus, contains a spanning 0-complex. So assume $d \geq 3$.

Let

$$\mathcal{C}_r := \mathbb{B}(F_r) - V(F).$$

If $F_r = F$ then $\mathcal{C}_r = \emptyset$, and if $F_r \cap F = \emptyset$ then $\mathcal{C}_r$ is the boundary complex of F_r, a strongly connected $(d-2)$-subcomplex of F_r (Proposition 9.0.2). Otherwise, $\mathcal{C}_r$ is the antistar of $F_r \cap F$ in F_r, also a strongly connected $(d-2)$-subcomplex of F_r (Lemma 10.1.2).

Let

$$\mathcal{C} := \bigcup_{r=1}^{n} \mathcal{C}_r.$$

We show that $\mathcal{C}$ is the required spanning strongly connected $(d-2)$-subcomplex of $P-V(F)$, the antistar of F in P. It follows that $\mathcal{C}$ is a spanning pure $(d-2)$-subcomplex of $P-V(F)$. It remains to prove that there exists a $(d-2, d-3)$-path L in $\mathcal{C}$ between any two ridges R_i and R_j of $\mathcal{C}$ with $i \neq j$.

If $R_i, R_j \in \mathcal{C}_r$ for some $r \in [1, n]$, then, since $\mathcal{C}_r$ is a strongly connected $(d-2)$-complex (Lemma 10.1.2), there exists a $(d-2, d-3)$-path in $\mathcal{C}_r$ between the two ridges R_i and R_j. Therefore, we can assume that R_i is in $\mathcal{C}_i$ and R_j is in $\mathcal{C}_j$ for $i \neq j$. Observe that $F_i \neq F$ and $F_j \neq F$. Hereafter we let $E_0 := R_i$ and $E_m := R_j$.

Since $\mathbb{B}(P)$ is a strongly connected $(d-1)$-subcomplex of P, there exists a $(d-1, d-2)$-path $M := J_1 \ldots J_m$ in P where $J_\ell \cap J_{\ell+1}$ is a ridge for $\ell \in [1, m-1]$, $J_1 = F_i$ and $J_m = F_j$. Each facet J_r coincides with a facet F_{i_r} for some $i_r \in [1, n]$; we henceforth let $\mathcal{D}_r := \mathcal{C}_{i_r}$.

By Lemma 10.2.1(iii)-(iv) we can assume that $J_r \neq F$ for $r \in [1, m]$ in the case of F being a facet and that $J_\ell \cap J_{\ell+1} \neq F$ for $\ell \in [1, m-1]$ in the case of F being a ridge. As a consequence, $\dim(J_\ell \cap J_{\ell+1} \cap F) \leq d-3$; this in turn implies that, for each $\ell \in [1, m-1]$, $J_\ell \cap J_{\ell+1}$ contains a $(d-3)$-face I_ℓ that is disjoint from F. Hence $I_\ell \in \mathcal{D}_\ell \cap \mathcal{D}_{\ell+1}$ for each $\ell \in [1, m-1]$.

As in the proof or Proposition 10.2.2, we show that the path L exists by proving the following claim by induction.

Claim 2. *If $\ell \leq m$, there exists a $(d-2, d-3)$ path in $\bigcup_{i=1}^{\ell} \mathcal{D}_i$ between $E_0 \in \mathcal{D}_1$ and any ridge $E_\ell \in \mathcal{D}_\ell$.*

Proof. The statement is true for $\ell = 1$. The complex $\mathcal{D}_1$ is a strongly connected $(d-2)$-complex and $E_0 \in \mathcal{D}_1$.

Suppose that the statement is true for some $\ell < m$. We show the existence of a $(d-2, d-3)$-path between $E_0 \in \mathcal{D}_1$ and any ridge of $\mathcal{D}_{\ell+1}$.

Let E_ℓ be a ridge in $\mathcal{D}_\ell$ containing a $(d-3)$ face I_ℓ of $\mathcal{D}_\ell \cap \mathcal{D}_{\ell+1}$; this $(d-3)$-face I_ℓ exists by our previous discussion. By the induction hypothesis, there exists a $(d-2, d-3)$ path L_ℓ in $\bigcup_{i=1}^{\ell} \mathcal{D}_i$ between E_0 and the ridge E_ℓ.

Consider a ridge $E'_{\ell+1} \in \mathcal{D}_{\ell+1}$ containing the face I_ℓ. There is a $(d-2, d-3)$-path $L'_{\ell+1}$ in $\mathcal{D}_{\ell+1}$ from $E'_{\ell+1}$ to any ridge $E_{\ell+1} \in \mathcal{D}_{\ell+1}$, thanks to $\mathcal{D}_{\ell+1}$ being a strongly connected $(d-2)$-complex.

The desired path $L_{\ell+1}$ between E_0 and the arbitrary ridge $E_{\ell+1}$ is obtained as $L_{\ell+1} = E_0 L_\ell E_\ell E'_{\ell+1} L_{\ell+1} E_{\ell+1}$. □

The claim for $\ell = m$ gives the desired $(d-2, d-3)$-path L in $\cup_{i=1}^{m} \mathcal{D}_i \subset \mathcal{C}$ between $E_0 = R_i$ and $E_m = R_j$, which concludes the proof. □

Remark 10.2.4. Theorem 10.2.3 is best possible in the sense that the antistar of a face does not always contain a spanning strongly connected $(d-1)$-subcomplex. The removal of the vertices of the face F in Fig. 1.1 leaves a pure $(d-1)$-subcomplex that is not strongly connected.

The ideas presented in Proposition 10.2.2 and Theorem 10.2.3 play a key role in the proof of the main result of [33].

Before proving the main result of the section, we state a useful corollary that follows from Proposition 9.0.1 and Theorem 10.2.3.

Corollary 10.2.5. *Let P be a cubical d-polytope and let F be a proper face of P. Then the subgraph $G(P) - V(F)$ is $(d-2)$-connected.*

For $d \geq 4$ we define the two functions $f(d)$ and $g(d)$ that we mentioned in the introduction.

(i) The function $f(d)$ gives the maximum number such that every cubical d-polytope with minimum degree $\delta \leq f(d)$ is δ-connected.

(ii) the function $g(d)$ gives the maximum number such that every minimum separator with cardinality at most $g(d)$ of every cubical d-polytope consists of the neighbourhood of some vertex.

The functions $f(3)$ and $g(3)$ are not defined. No cubical 3-polytope has minimum degree $\delta \geq 4$, and so for every positive integer $\delta_0 \geq 3$ it follows that every cubical 3-polytope with minimum degree $\delta \leq \delta_0$ is δ-connected. Figure 1.1 shows cubical 3-polytopes with minimum separators that are not the neighbourhood of a vertex.

The function $f(d)$ is well defined for $d \geq 4$. There is a cubical d-polytope with minimum degree δ for every $\delta \geq d \geq 4$, for instance, a neighbourly cubical d-polytope [79]. Every d-polytope is d-connected by Balinski's theorem. Furthermore, there exists a cubical d-polytope with minimum degree $\delta > 2^{d-1}$ that is not δ-connected: the connected sum of two copies of a neighbourly cubical d-polytope with minimum degree δ. Thus $d \leq f(d) \leq 2^{d-1}$.

At this moment, we don't claim that $g(d)$ exists; this will become evident in the proof of Theorem 10.2.8.

Proposition 10.2.6. *Let P be a cubical d-polytope with $d \geq 4$. If the function $g(d)$ exists and P has minimum degree at least $g(d)+1$, then $G(P)$ is $(g(d)+1)$-connected.*

Proof. Suppose that $G(P)$ is not $(g(d)+1)$-connected. Then there is a minimum separator X with cardinality at most $g(d)$. By the definition of $g(d)$, X consists of all the neighbours of some vertex u. This contradicts the degree of u, which is at least $g(d)+1 > |X|$. □

Corollary 10.2.7. *If the function $g(d)$ exists for $d \geq 4$, then $f(d) > g(d)$.*

Theorem 10.2.8 (Connectivity Theorem)**.** *A cubical d-polytope P with minimum degree δ is $\min\{\delta, 2d-2\}$-connected for every $d \geq 3$.*

Furthermore, for any $d \geq 4$, every minimum separator X of cardinality at most $2d-3$ consists of all the neighbours of some vertex, and the subgraph $G(P)-X$ contains exactly two components, with one of them being the vertex itself.

Proof. Let $0 \leq \alpha \leq d-3$ and let P be a cubical d-polytope with minimum degree at least $d+\alpha$. Let $G := G(P)$.

We first prove that P is $(d+\alpha)$-connected. The case of $d = 3$ follows from Balinski's theorem. So assume $d \geq 4$. Let X be a minimum separator of P. Throughout the proof, let u and v be two distinct vertices that belong to $G-X$ and are disconnected by X. The theorem follows from a number of claims that we prove next.

Claim 3. *If $|X| \leq d+\alpha$ then, for any facet F, the cardinality of $X \cap V(F)$ is at most $d-1$.*

Proof. Suppose otherwise and let F be a facet with $|X \cap V(F)| \geq d$. Let

$$G' := G - V(F).$$

According to Corollary 10.2.5, the subgraph G' is $(d-2)$-connected. Since there are at most $\alpha \leq d-3$ vertices in $V(G') \cap X$, removing from G' the vertices in $V(G') \cap X$ doesn't disconnect G'.

We show there is a $u-v$ path in $G-X$, which would be a contradiction and prove the claim. If $u,v \in V(G') \setminus X$ then there is a $u-v$ path in $G'-X$, as $G'-X$ is connected. So assume $u \in V(F) \setminus X$. Since u has degree at least $d+\alpha$ and since every vertex in F has at least $d+\alpha-(d-1) = \alpha+1$ neighbours outside F (in G'), at least one of them, say $u_{G'}$, is in $V(G') \setminus X$. Likewise either $v \in V(F) \setminus X$ and there is a neighbour $v_{G'}$ of v in $V(G') \setminus X$ or $v \in G'-X$. Therefore, if $v \in V(F) \setminus X$ then there is a $u-v$ path L in $G-X$ that contains a subpath L' in G' between the vertices $u_{G'}$ and $v_{G'}$ in $V(G') \setminus X$; that is, $L = uu_{G'}L'v_{G'}v$. If instead $v \in G'-X$ then there is a $u-v$ path L in $G-X$ passing through the vertex $u_{G'}$ and containing a subpath $L' := u_{G'} - v$ in $G'-X$; that is, $L = uu_{G'}L'v$. Hence there is always a $u-v$ path in $G-X$, and thus, G is not disconnected by X, a contradiction. □

Claim 4. *If $|X| \le d+\alpha$, then there exist facets $F_1, \ldots, F_d$ of P such that $G(F_i)$ is disconnected by X for each $i \in [1,d]$.*

Proof. Suppose by way of contradiction that X disconnects the graphs of at most k facets $F_1, \ldots, F_k$ of P with $k \le d-1$. We find a $u-v$ path in $G-X$, which would contradict X being a separator of G.

There are at least d facets containing u and there are at least d facets containing v. As a result, we can pick facets K_u and K_v with $u \in K_u$ and $v \in K_v$ whose graphs are not disconnected by X; that is $K_u, K_v \notin \{F_1, \ldots, F_k\}$. If $K_u = K_v$ then we can find a $u-v$ path in $G(K_u) - X$. So assume $K_u \ne K_v$. Since $\mathbb{B}(P)$ is a strongly connected $(d-1)$-complex and since there are at least d independent $(d-1, d-2)$-paths from K_u to K_v in $\mathbb{B}(P)$ (Lemma 10.2.1(i)), there exists a $(d-1, d-2)$-path $J_1 \ldots J_n$ in $\mathbb{B}(P)$ with $J_1 = K_u$ and $J_n = K_v$ such that $\{J_1, \ldots, J_n\} \cap \{F_1, \ldots, F_k\} = \emptyset$. As a consequence, the subgraphs $G(J_i)$ are not disconnected by X.

Construct a $u-v$ path L by traversing the facets $J_1, \ldots, J_n$ as follows: find a path L_1 in J_1 from u to a vertex in $J_1 \cap J_2$, then a path L_2 in J_2 from $J_1 \cap J_2$ to $J_2 \cap J_3$ and so on up to a path L_{n-1} in J_{n-1} from $J_{n-2} \cap J_{n-1}$ to $J_{n-1} \cap J_n$; here use the connectivity of the subgraphs $G(J_1) - X, \ldots, G(J_{n-1}) - X$. Finally, find a path L_n in $J_n = K_v$ from $J_{n-1} \cap J_n$ to the vertex v using the connectivity of $G(J_n) - X$. The path L is the concatenation of the paths $L_1, \ldots, L_n$.

The aforementioned concatenation works as long as there is at least one vertex in $V(J_\ell \cap J_{\ell+1}) \setminus X$ for each $\ell \in [1, n-1]$. For $d \ge 4$, it follows that $|V(J_\ell \cap J_{\ell+1})| = 2^{d-2} \ge d$, which is greater that $|V(J_\ell) \cap X| \le d-1$ by Claim 3. Hence $V(J_\ell \cap J_{\ell+1}) \setminus X$ is nonempty, and consequently, the $u-v$ path L always exists and completes the proof of the claim. □

Claim 5. *If $|X| \le d+\alpha$ then $|X| = d+\alpha$.*

Proof. Let F be a facet of P whose graph is disconnected by X, which by Claim 4 exists. Claim 3 together with Balinski's theorem ensures that $|V(F) \cap X| = d-1$. Let $G' := G - V(F)$. By Corollary 10.2.5, G' is a $(d-2)$-connected subgraph of G.

Suppose that a minimum separator X has size at most $d-1+\alpha$; we show that X does not disconnect G by finding a $u-v$ path L between the vertices u and v of $G-X$, which would be a contradiction.

There are at most $\alpha \le d-3$ vertices in $V(G') \cap X$, and so removing $V(G') \cap X$ from G' doesn't disconnect G'.

If u and v are both in G' then there is a $u-v$ path in G' that is disjoint from X. So assume that $u \in V(F) \setminus X$. Let X_1 denote the set of neighbours of u in G'; then $|X_1| \ge \alpha + 1$, since u has at least $d+\alpha$ neighbours in P, with exactly $d-1$ of them in F. As a consequence, there is a neighbour $u_{G'}$ of u in $V(G') \setminus X$. Likewise either $v \in V(F) \setminus X$ and there is a neighbour $v_{G'}$ of v in $V(G') \setminus X$ or $v \in G'-X$. If $v \in V(F) \setminus X$, there is a $u-v$ path in $G-X$ that passes through the vertices $u_{G'}$ and $v_{G'}$ of $V(G') \setminus X$. If instead $v \in G'-X$, there is a $u-v$ path L in $G-X$ that includes a subpath L' in $G'-X$ between $u_{G'}$ and v so that $L = uu_{G'}L'v$. Hence we always have a $u-v$ path in $G-X$. This contradiction shows that a minimum separator has size *exactly* $d+\alpha$. □

From Claim 5 it follows that P is $(d+\alpha)$-connected. The structure of a minimum separator is settled in (6), (7). For every $d \geq 4$, Claim 6 settles the case $\alpha \leq d-4$ and Claim 7 the case $\alpha = d-3$.

Claim 6. *If $\alpha \leq d-4$, then the set X consists of the neighbours of some vertex and the minimum degree of P is exactly $d+\alpha$.*

Proof. As in Claim 5, let F be a facet of P whose graph is disconnected by X and let $G' := G - V(F)$. Then G' is a $(d-2)$-connected subgraph of G (Corollary 10.2.5). Besides, $|V(F) \cap X| = d-1$ by a combination of Claim 3 and Balinski's theorem.

Since there are *exactly* $\alpha + 1 \leq d-3$ vertices in $V(G') \cap X$, removing $V(G') \cap X$ from G' doesn't disconnect G'. We may therefore assume that $u \in V(F) \setminus X$.

If there is a path L_u in $G - X$ from u to a vertex $u_{G'} \in G' - X$ and a path L_v in $G - X$ from v to a vertex $v_{G'}$ in $G' - X$ so that L_u and L_v are both disjoint from X, then we get a $u-v$ path L in $G - X$ defined as $L = uL_u u_{G'} L' v_{G'} L_v v$ where L' is a path in $G' - X$ between $u_{G'}$ and $v_{G'}$. Recall the minimum degree of u is at least $d + \alpha$.

We may therefore assume that u is in $V(F) \setminus X$ and that there is no such path L_u in $G - X$ from u to $G' - X$. The set X_1 of neighbours of u in G' must then be a subset of X, and since $|X_1| \geq \alpha + 1$, it follows that $X_1 = V(G') \cap X$, and thus, that $|X_1| = \alpha + 1$. In addition, every path of length two from u to G' passing through a neighbour of u in F contains some vertex from X; otherwise the aforementioned path L_u would exist. Let X_2 denote the vertices in X that are present in a $u - V(G')$ path of length two passing through a neighbour of u in F. Every vertex of F has a neighbour in G', and so there is a $u - V(G')$ path through each neighbour of u in F, $d-1$ such neighbours in total. Since there are no triangles in P, we get $X_1 \cap X_2 = \emptyset$, which in turn implies that $X_2 \subset V(F)$. Hence $|X_2| = d-1$, and every neighbour of u in F is in X. Consequently, the degree of u, $|X_1| + |X_2|$, is precisely $d + \alpha$, and the set X consists of the $d + \alpha$ neighbours of u in P, as desired. □

Claim 7. *If $\alpha = d-3$, then the set X consists of the neighbours of some vertex and the minimum degree of P is exactly $d+\alpha$.*

Proof. Proceed by contradiction: every vertex in P has at least one neighbour outside X.

By Claim 3 there are at most $d-1$ vertices from X in any facet F of P. If the removal of X disconnects the graph of a facet F, then there would be exactly $d-1$ vertices in $V(F) \cap X$, which constitute the neighbours in F of some vertex of F (Proposition 10.1.3). Consequently, the subgraph $G(F) - X$ would have exactly two components: one being a singleton $z(F)$ and another $Z(F)$ being $(d-3)$-connected by Proposition 10.1.4; if X doesn't disconnect F, we let $z(F) = \emptyset$ and let $Z(F) := G(F) - X$. Hence, for every facet F of P, the subgraph $Z(F)$ is connected, and $V(F) = z(F) \cup V(Z(F)) \cup (V(F) \cap X)$. Abusing terminology, if $z(F) \neq \emptyset$ we make no distinction between the set and its unique element.

Since u and v are separated by X, every $u-v$ path in G contains a vertex from X. Because the vertex u has a neighbour w not in X, there must exist a facet F_u in which $u \in Z(F_u)$: a facet containing the edge uw. Similarly, there exists a facet F_v containing v in which $v \in Z(F_v)$.

Consider an arbitrary $(d-1, d-2)$-path $J_1 \ldots J_n$ in P with $J_1 = F_u$ and $J_n = F_v$. If, for each $i \in [1, n-1]$, there is a vertex $y_i \in V(J_i \cap J_{i+1})$ with $y_i \in V(Z(J_i)) \cap V(Z(J_{i+1}))$, then there would be a $u-v$ path L in $G - X$ and the claim would hold. Indeed, let $y_0 := u$ and $y_n := v$. For all $i \in [0, n-1]$, there would be a path L_{i+1} in $Z(J_{i+1})$ from y_i to y_{i+1}. Concatenating all these paths $L_1, \ldots, L_n$, we would then have a $u-v$ path L in $G - X$, giving a contradiction and settling the claim. We would say that a facet-ridge path $J_1 \ldots J_n$ from F_u to F_v is *valid* if the aforementioned vertex y_i exists for each $i \in [1, n-1]$; otherwise it is *invalid*.

Hence **it remains to show that, for some facet-ridge path $J_1 \ldots J_n$ from $J_1 = F_u$ to $J_n = F_v$, there exists a vertex in $V(Z(J_i)) \cap V(Z(J_{i+1}))$ for each $i \in [1, n-1]$ when $d \geq 4$.** In other words,

it remains to show there exists a valid fact-ridge path from F_u to F_v.

Take a facet-ridge path $J_1 \dots J_n$ from F_u to F_v and suppose it is invalid; that is, $V(Z(J_i)) \cap V(Z(J_{i+1})) = \emptyset$ for some $i \in [1, n-1]$. Then $V(Z(J_i)) \cap V(J_{i+1}) \subset z(J_{i+1})$. Therefore,

$$\begin{aligned} V(J_i \cap J_{i+1}) = V(J_i) \cap V(J_{i+1}) &= [z(J_i) \cup V(Z(J_i)) \cup (V(J_i) \cap X)] \cap V(J_{i+1}) \\ &\subset z(J_i) \cup z(J_{i+1}) \cup (X \cap V(J_i \cap J_{i+1})). \end{aligned} \tag{10.1}$$

If neither $G(J_i)$ nor $G(J_{i+1})$ is disconnected by X, then $z(J_i) = z(J_{i+1}) = \emptyset$, and by Eq. (10.1) and Claim 3,

$$2^{d-2} = |V(J_i \cap J_{i+1})| \leq |X \cap V(J_i)| \leq d-1. \tag{10.2}$$

If instead $G(J_i)$ is disconnected by X, then $X \cap V(J_i)$ consists of all the $d-1$ neighbours of $z(J_i)$ in J_i (Proposition 10.1.3), and thus, $|X \cap V(J_i \cap J_{i+1})| \leq d-2$. In this case, by Eq. (10.1),

$$2^{d-2} = |V(J_i \cap J_{i+1})| \leq 2 + d - 2 = d. \tag{10.3}$$

Equation (10.2) does not hold for $d \geq 4$, while Eq. (10.3) only holds for $d = 4$, in which case it holds with equality. As a consequence, if $d \geq 5$, every facet-ridge path from F_u to F_v is valid. As a result, the aforementioned $u-v$ path L in $G - X$ always exists for $d \geq 5$, a contradiction. **This completes the case $d \geq 5$.**

The case $d = 4$ requires more work. Let

$$X := \{x_1, \dots, x_5\}.$$

Suppose by way of contradiction that every facet-ridge path from F_u to F_v is invalid. Consider a particular such path $M := J_1 \dots J_n$. Then $V(Z(J_i)) \cap V(Z(J_{i+1})) = \emptyset$ for some $i \in [1, n-1]$, and for that index i, Eq. (10.3) must hold with equality, which implies that Eq. (10.1) must also hold with equality. Consequently, the following setting ensues

(i) $|z(J_i) \cup z(J_{i+1})| = 2$; that is, $z(J_i) \neq z(J_{i+1})$;

(ii) the graphs of the facets J_i and J_{i+1} are both disconnected by X;

(iii) the neighbours of $z(J_i)$ in J_i and of $z(J_{i+1})$ in J_{i+1} are all from X;

(iv) the ridge $R_i := J_i \cap J_{i+1}$ consists of four vertices—namely, $z(J_i)$, $z(J_{i+1})$ and two vertices from X, say x_1 and x_2;

(v) each vertex $z(J_i)$ and $z(J_{i+1})$ has a neighbour in $J_i \setminus J_{i+1}$ and $J_{i+1} \setminus J_i$, respectively; and

(vi) there is a vertex from X, say x_5, lying outside $J_i \cup J_{i+1}$.

Any pair of facets in this setting are said to be in *Configuration A* and the ridge in which they intersect is said to be *problematic*. For instance, the pair (J_i, J_{i+1}) is in Configuration A and the ridge R_i is problematic; see Fig. 10.2(a).

For a facet-ridge path from F_u to F_v to be invalid, it must have a pair of facets in Configuration A.

We want to be more careful when selecting the facets F_u and F_v and when selecting the facet-ridge path M from F_u to F_v. We require the following.

> The facets F_u and F_v can be picked so that their graphs are not disconnected by X; that is, $G(F_u) - X$ and $G(F_v) - X$ are both connected subgraphs of $G(F_u)$ and $G(F_v)$, respectively. (*)

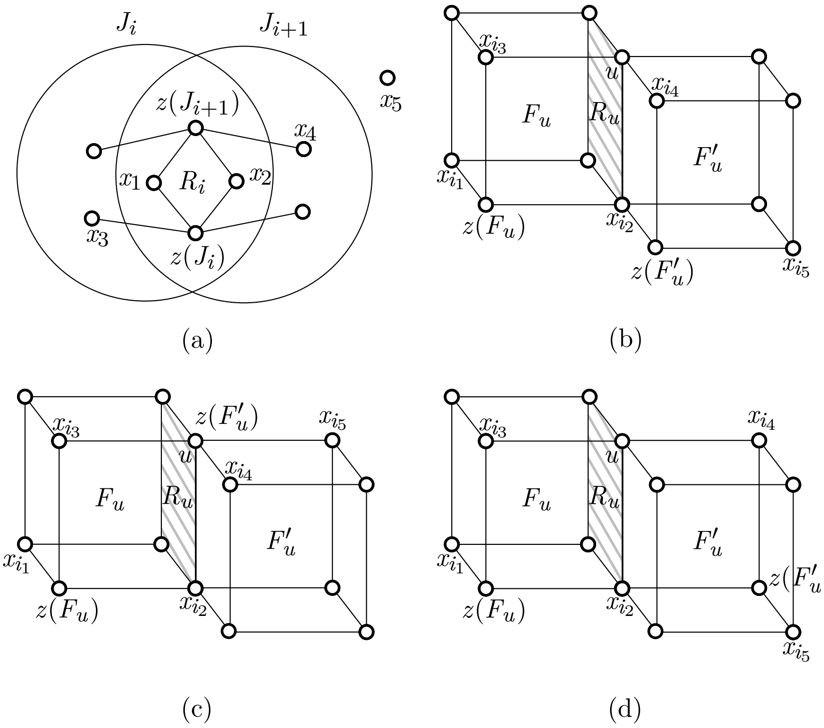


Figure 10.2: Auxiliary figure for Claim 7 of Theorem 10.2.8. (a) Configuration A: two facets J_i and J_{i+1} whose graphs are disconnected by $X = \{x_1, x_2, x_3, x_4, x_5\}$ and a problematic ridge $R_i := J_i \cap J_{i+1}$. (b)–(d) The facets F_u and F_u are both disconnected by X and intersects at an edge. The ridge $R_u := F_u \cap F''_u$ that defines a new facet F''_u is highlighted.

Proof of (*). Suppose that the facet F_u cannot be picked as desired. Then the graph of F_u is disconnected by X, and by Proposition 10.1.3 there is a vertex $z(F_u) \in G(F_u)$ whose neighbours in F_u are all from X. Say $X \cap V(F_u) = \{x_{i_1}, x_{i_2}, x_{i_3}\}$. Recall that $u \in Z(F_u)$.

Since u has degree at least five (it is nonsimple), it follows that there is a facet F'_u in P containing u and intersecting F_u at a vertex or an edge. Since F'_u contains u, its graph must be disconnected by X (otherwise it is the desired facet). Therefore $|X \cap V(F'_u)| = 3$, and thus, $X \cap V(F_u) \cap V(F'_u) \neq \emptyset$. As a consequence, we find that $F_u \cap F'_u$ is an edge between a vertex of X, say x_{i_2}, and u. It follows that $X \cap V(F'_u) = \{x_{i_2}, x_{i_4}, x_{i_5}\}$. Three configurations are possible: Fig. 10.2(b)–(d).

The argument remains unchanged in all the three configurations. Refer to Fig. 10.2(b) for concreteness. Consider the ridge R_u of F_u that contains the edge ux_{i_2} but does not contain the vertex x_{i_3}; the ridge R_u is highlighted in Fig. 10.2(b). Let F''_u be the facet of P that intersects F_u at R_u. Then $X \cap V(F''_u) \subseteq \{x_{i_2}, x_{i_4}\}$, since $F_u \cap F''_u$ and $F'_u \cap F''_u$ are faces that contain u. Therefore, the graph of F''_u is not disconnected by X and F''_u could have been chosen as F_u. As a consequence of this contradiction, the facet F_u can be picked as desired.

Similar analysis shows that the facet F_v can also be picked so that $G(F_v)$ is not disconnected by X. This completes the proof of (*). □

We are now ready to complete the proof of the claim by showing that we can always find a valid facet-ridge $F_u - F_v$ path. The existence of such a path would complete the proof of the claim.

There are at least four independent facet-ridge paths from F_u to F_v (Lemma 10.2.1(i))—say M_a, M_b, M_c and M_d—and at least four pairs of facets exhibiting Configuration A—one per path. Each pair of facets in Configuration A gives rise to a problematic ridge. We may assume that $M = M_a$. The ensuing four points are key.

(i) The facet F_u or F_v does not appear in any Configuration A (by Statement (*)).

(ii) Any facet of P other than F_u and F_v may appear in at most one facet-ridge $F_u - F_v$ path; in particular, it appears in at most one pair exhibiting Configuration A.

(iii) The problematic ridges are pairwise distinct, as the paths M_a, M_b, M_c, M_d as independent.

(iv) Each problematic ridge appears in precisely one of the paths M_a, M_b, M_c, M_d.

(v) Each nonproblematic ridge R of P present in a Configuration A appears in at most two paths in $\{M_a, M_b, M_c, M_d\}$. This is so because R is the intersection of two facets F and F', and the facet F or F' appears in at most one such path.

With a counting argument we show that Configuration A cannot occur in all the four paths. We count the ridges that contain two vertices from X and are present in a Configuration A.

For every pair of facets (J, J') exhibiting Configuration A, there are five ridges in $J \cup J'$ containing two vertices from X. For instance, for the pair (J_i, J_{i+1}) of Fig. 10.2(a), the pairs (x_1, x_2), (x_1, x_3), (x_2, x_3), (x_1, x_4) and (x_2, x_4) induce the five ridges. So, considering the four aforementioned $F_u - F_v$ paths, we have a total of twenty ridges that are present in a Configuration A and contain two vertices from X. Besides, there are ten ways of pairing two vertices from X, and thereby there are at most ten distinct ridges containing two vertices from X.

Each problematic ridge appears in precisely one Configuration A; there are at least four problematic ridges, and therefore, there are at most six nonproblematic ridges. Each nonproblematic ridge that contains two vertices from X appears in two facets, and consequently in at most two pairs of facets exhibiting Configuration A (that is, in at most two Configurations A). We account for the ten ridges containing two vertices from X in the four Configurations A: at least four problematic ridges and at most six nonproblematic ones. This means that we can have at most $4 + 6 \times 2 = 16$ ridges that contain two vertices from X and appear in the four Configurations A. Since the four Configurations A require twenty ridges containing two vertices from X, we can choose a $(d-1, d-2)$-path $F_u - F_v$ in which Configuration A doesn't occur. **This completes the case** $d = 4$, and with it the proof of the claim. $\square$

We now complete the proof of the theorem. (6), (7) ensure that, for $d \geq 4$, a minimum separator X with cardinality at most $2d - 3$ in a cubical d-polytope consists of the neighbours of a vertex. Thus **the function** $g(d)$ **exists and satisfies** $2d - 3 \leq g(d)$**.**

From Corollary 10.2.7 it then follows that $f(d) \geq 2d - 2$; in other words, a cubical d-polytope with minimum degree δ is $\min\{\delta, 2d-2\}$-connected. This completes the proof of the theorem. $\square$

A simple corollary of Theorem 10.2.8 is the following.

Corollary 10.2.9. *A cubical d-polytope with no simple vertices is* $(d+1)$*-connected.*

As we mentioned in the introduction an open problem that nicely arises from Theorem 10.2.8 is Problem 1.2.4.

Chapter 11

Linkedness

A graph with at least $2k$ vertices is k-linked if, for every set of $2k$ distinct vertices organised in arbitrary k unordered pairs of vertices, there are k vertex-disjoint paths joining the vertices in the pairs.

Denote by $V(X)$ the vertex set of a graph or a polytope X. Given sets A, B of vertices in a graph, a path from A to B, called an $A - B$ *path*, is a (vertex-edge) path $L := u_0 \dots u_n$ in the graph such that $V(L) \cap A = \{u_0\}$ and $V(L) \cap B = \{u_n\}$. We write $a - B$ path instead of $\{a\} - B$ path, and likewise, write $A - b$ path instead of $A - \{b\}$.

Let G be a graph and X a subset of $2k$ distinct vertices of G. The elements of X are called *terminals*. Let $Y := \{\{s_1, t_1\}, \dots, \{s_k, t_k\}\}$ be an arbitrary labelling and (unordered) pairing of all the vertices in X. We say that Y is *linked* in G if we can find disjoint $s_i - t_i$ paths for $i \in [1, k]$, the interval $1, \dots, k$. The set X is *linked* in G if every such pairing of its vertices is linked in G. Throughout this part of the thesis, by a set of disjoint paths, we mean a set of vertex-disjoint paths. If G has at least $2k$ vertices and every set of exactly $2k$ vertices is linked in G, we say that G is *k-linked*. If the graph of a polytope is k-linked we say that the polytope is also *k-linked*.

This part of the thesis studies the linkedness of *cubical d-polytopes* (see [32]).

11.1 d-Cube

In the d-cube Q_d, the facet disjoint from a facet F is denoted by F^o, and we say that F and F^o is a pair of *opposite* facets.

Definition 11.1.1 (Projection π)**.** For a pair of opposite facets $\{F, F^o\}$ of Q_d, define a projection $\pi_{F^o}^{Q_d}$ from Q_d to F^o by sending a vertex $x \in F$ to the unique neighbour $x_{F^o}^p$ of x in F^o, and a vertex $x \in F^o$ to itself (that is, $\pi_{F^o}^{Q_d}(x) = x$); write $\pi_{F^o}^{Q_d}(x) = x_{F^o}^p$ to be precise, or write $\pi(x)$ or x^p if the cube Q_d and the facet F^o are understood from the context.

We extend this projection to sets of vertices: given a pair $\{F, F^o\}$ of opposite facets and a set $X \subseteq V(F)$, the projection $X_{F^o}^p$ or $\pi_{F^o}^{Q_d}(X)$ of X onto F^o is the set of the projections of the vertices in X onto F^o. For an i-face $J \subseteq F$, the projection $J_{F^o}^p$ or $\pi_{F^o}^{Q_d}(J)$ of J onto F^o is the i-face consisting of the projections of all the vertices of J onto F^o. For a pair $\{F, F^o\}$ of opposite facets in Q^d, the restrictions of the projection π_{F^o} to F and the projection π_F to F^o are bijections.

Let Z be a set of vertices in the graph of a d-cube Q_d. If, for some pair of opposite facets $\{F, F^o\}$, the set Z contains both a vertex $z \in V(F) \cap Z$ and its projection $z_{F^o}^p \in V(F^o) \cap Z$, we say that the pair $\{F, F^o\}$ is *associated* with the set Z in Q_d and that $\{z, z^p\}$ is an *associating pair*. Note that an associating pair can associate only one pair of opposite facets.

In conjunction with connectivity results around strongly connected complexes in cubical polytopes, the next lemma lies at the core of our methodology.

Lemma 11.1.2. *Let Z be a nonempty subset of $V(Q_d)$. Then the number of pairs $\{F, F^o\}$ of opposite facets associated with Z is at most $|Z| - 1$.*

Proof. Let $G := G(Q_d)$ and let $Z \subset V(Q_d)$ with $|Z| \geq 1$ be given. Consider a pair $\{F, F^o\}$ of opposite facets. Define a *direction* in the cube as the set of the 2^{d-1} edges between F and F^o; each direction corresponds to a pair of opposite facets. The d directions partition the edges of the cube into sets of cardinality 2^{d-1}. (The notion of direction stems from thinking of the cube as a zonotope [175, Section 7.3])

A pair of facets is associated with the set Z if and only if the subgraph $G[Z]$ of G induced by Z contains an edge from the corresponding direction.

If a direction is present in a cycle C of Q_d, then the cycle contains at least two edges from this direction. Indeed, take an edge $e = uv$ on C that belongs to a direction between a pair $\{F, F^o\}$ of opposite facets. After traversing the edge e from $u \in V(F)$ to $v \in V(F^o)$, for the cycle to come back to the facet F, it must contain another edge from the same direction. Hence, by repeatedly removing edges from cycles in $G[Z]$ we obtain a spanning forest of $G[Z]$ that contains an edge for every direction present in $G[Z]$. As a consequence, the number of such directions is at most the number of edges in the forest, which is upper bounded by $|Z| - 1$. (A *forest* is a graph with no cycles.) $\square$

The relevance of the lemma stems from the fact that a pair of opposite facets $\{F, F^o\}$ not associated with a given set of vertices Z allows each vertex z in Z to have "free projection"; that is, for every $z \in Z \cap V(F)$ the projection $\pi_{F^o}(z)$ is not in Z, and for $z \in Z \cap V(F^o)$ the projection $\pi_F(z)$ is not in Z.

11.1.1 Strong Connectivity of the d-Cube

We next unveil some further properties of the cube that will be used in subsequent sections.

While it is true that the antistar of a vertex in a d-polytope is always a strongly connected $(d-1)$-complex (Proposition 9.0.2), it is far from true that this extends to higher dimensional faces. Refer to [33, Section 3] for examples of d-polytopes in which this extension is not possible. This extension is however possible for the d-cube, as shown in Lemma 10.1.2 ([33, Lemma 8]).

Given sets A, B, X of vertices in a graph G, the set X *separates* A from B if every $A - B$ path in the graph contains a vertex from X. A set X separates two vertices a, b not in X if it separates $\{a\}$ from $\{b\}$. We call the set X a *separator* of the graph.

We will also require the following three assertions.

Proposition 11.1.3 ([138, Proposition 1]). *Any separator X of cardinality d in Q_d consists of the d neighbours of some vertex in the cube and the subgraph $G(Q_d) - X$ has exactly two components, with one of them being the vertex itself.*

A set of vertices in a graph is *independent* if no two of its elements are adjacent. Since there are no triangles in a d-cube, Proposition 11.1.3 gives at once the following corollary.

Corollary 11.1.4. *A separator of cardinality d in a d-cube is an independent set.*

Remark 11.1.5. If x and y are vertices of a cube, then they share at most two neighbours. In other words, the complete bipartite graph $K_{2,3}$ is not a subgraph of the cube; in fact, it is not an induced subgraph of any simple polytope [133, Cor. 1.12(iii)].

11.1.2 Linkedness of the d-Cube

The linkedness of a d-cube was first established in [118, Proposition 4.4] as part of a study of linkedness in Cartesian products of graphs. We give an alternative proof of the result. As discussed before (Proposition 9.0.8), the linkedness of Q_4 is easily shown to be two.

Since we make heavy use of Menger's theorem [46, Theorem 3.3.1] henceforth, we next remind the reader of one of its consequences.

Theorem 11.1.6 (Menger [46, Section 3.3])**.** *Let G be a k-connected graph, and let A and B be two subsets of its vertices, each of cardinality at least k. Then there are k disjoint $A-B$ paths in G.*

Two vertex-edge paths are *independent* if they share no inner vertex.

Lemma 11.1.7. *Let P be a cubical d-polytope with $d \geq 4$. Let X be a set of $d+1$ vertices in P, all contained in a facet F. Let $k := \lfloor (d+1)/2 \rfloor$. Arbitrarily label and pair $2k$ vertices in X to obtain $Y := \{\{s_1, t_1\}, \ldots, \{s_k, t_k\}\}$. Then, for at least $k-1$ of these pairs $\{s_i, t_i\}$, there is an X-valid $s_i - t_i$ path in F.*

Proof. If, for each pair in Y there is an X-valid path in F connecting the pair, we are done. So assume there is a pair in Y, say $\{s_1, t_1\}$, for which an X-valid $s_1 - t_1$ path does not exist in F. Since F is $(d-1)$-connected, there are $d-1$ independent $s_1 - t_1$ paths (by Menger's theorem), each containing a vertex from $X \setminus \{s_1, t_1\}$; that is, the set $X \setminus \{s_1, t_1\}$, with cardinality $d-1$, separates s_1 from t_1 in F. By Proposition 11.1.3, the vertices in $X \setminus \{s_1, t_1\}$ are the neighbours of s_1 or t_1 in F, say of s_1.

Take any pair in $Y \setminus \{\{s_1, t_1\}\}$, say $\{s_2, t_2\}$. If there was no X-valid $s_2 - t_2$ path in F, then, by Proposition 11.1.3, the set $X \setminus \{s_2, t_2\}$ would separate s_2 from t_2 and would consist of the neighbours of s_2 or t_2 in F, say of s_2. But in this case, a vertex x in $X \setminus \{s_1, s_2, t_1, t_2\}$, which exists since $|X| \geq 5$, would form a triangle with s_1 and s_2, a contradiction. See also Corollary 11.1.4. Since our choice of $\{s_2, t_2\}$ was arbitrary, we must have an X-valid path in F between any pair $\{s_i, t_i\}$ for $i \in [2, k]$. □

For a set $Y := \{\{s_1, t_1\}, \ldots, \{s_k, t_k\}\}$ of pairs of vertices in a graph, a *Y-linkage* $\{L_1, \ldots, L_k\}$ is a set of disjoint paths with the path L_i joining the pair $\{s_i, t_i\}$ for $i \in [1, k]$. For a path $L := u_0 \ldots u_n$ we often write $u_i L u_j$ for $0 \leq i \leq j \leq n$ to denote the subpath $u_i \ldots u_j$. We are now ready to prove Theorem 11.1.8.

Theorem 11.1.8 (Linkedness of the cube)**.** *For every $d \neq 3$, a d-cube is $\lfloor (d+1)/2 \rfloor$-linked.*

Proof. The cases of $d = 1, 2$ are trivially true. For the remaining values of d, we proceed by induction, with $d = 4$ given by Proposition 9.0.8.

Let $k := \lfloor (d+1)/2 \rfloor$, then $2k-1 \leq d$. Let $X := \{s_1, \ldots, s_k, t_1, \ldots, t_k\}$ be any set of $2k$ vertices, called terminals, in the graph of the d-cube. Also let $Y := \{\{s_1, t_1\}, \ldots, \{s_k, t_k\}\}$. We aim to find a Y-linkage $\{L_1, \ldots, L_k\}$ with L_i joining the pair $\{s_i, t_i\}$ for $i = 1, \ldots, k$.

We consider three scenarios: (1) all the pairs in Y lie in some facet of Q_d, (2) a pair of Y lies in some facet F of Q_d but not every vertex of X is in F, and (3) no pair of Y lies in a facet of Q_d, which amounts to saying that every pair in Y is at distance d in Q_d. For the sake of readability, each scenario is highlighted in bold.

In the first scenario every vertex in X lies in some facet F of Q_d. Hence Lemma 11.1.7 gives an X-valid path L_1 in F joining a pair in Y, say $\{s_1, t_1\}$. The projection in Q_d of every vertex in $(X \setminus \{s_1, t_1\}) \cap V(F)$ onto F^o is not in X. Recall F^o denotes the facet opposite to F. Define $Y^p := \{\{s_2^p, t_2^p\}, \ldots, \{s_k^p, t_k^p\}\}$ as the set of $k-1$ pairs of projections of the corresponding vertices in $X \setminus \{s_1, t_1\}$ onto F^o. By the induction hypothesis on F^o, there is a Y^p-linkage $\{L_2^p, \ldots, L_k^p\}$ with $L_i^p := s_i^p - t_i^p$ for $i \in [2, k]$. Since $V(F^o)$ is disjoint from $V(L_1) \cup X$, each path L_i^p can be extended with s_i and t_i to obtain a path $L_i := s_i - t_i$ for $i \in [2, k]$. And together, all the paths $\{L_1, \ldots, L_k\}$ give the desired Y-linkage in the cube.

In the second scenario a pair of Y, say $\{s_1, t_1\}$, lies in some facet F of Q_d but not every vertex in X is in F. Let $N_K(x)$ denote the set of neighbours of a vertex x in a face K of the cube and let $N(x)$ denote the set of all the neighbours of x in the cube.

In what follows, whenever $x \in X$ we let $\{x, y\} \in Y$. Let $X_F := (X \setminus \{s_1, t_1\}) \cap V(F)$, and partition X_F as follows.

$$
\begin{aligned}
X_0 &:= \{x \in X_F : \{x, y\} \in Y \text{ and } y \in N_F(x)\} \\
X_1 &:= \{x \in X_F \setminus X_0 : \{x, x_{F^o}^p\} \in Y\} \\
X_2 &:= \{x \in X_F \setminus (X_0 \cup X_1) : x_{F^o}^p \notin X\} \\
X_3 &:= \{x \in X_F \setminus (X_0 \cup X_1 \cup X_2) : x_{F^o}^p \in X \text{ and there is a unique } X\text{-valid path } x y_F^p y \\
&\qquad \text{with } \{x, y\} \in Y \text{ and } y \in V(F^o)\} \\
X_4 &:= X_F \setminus (X_0 \cup X_1 \cup X_2 \cup X_3)
\end{aligned}
$$

Claim 8. *For every vertex x in $X_1 \cup X_2 \cup X_3 \cup X_4$, there is an X-valid path M_x of length at most two from x to F^o such that $(V(M_x) \cap X) \subseteq \{x, y\}$ and the $|X_1 \cup X_2 \cup X_3 \cup X_4|$ paths M_x are pairwise disjoint.*

Proof. For $x \in X_1 \cup X_2$, let $M_x = x x_{F^o}^p$. For $x \in X_3$, let M_x be the unique X-valid path $x y_F^p y$ with $\{x, y\} \in Y$ and $y \in V(F^o)$. To find the paths M_x for $x \in X_4$ we run the following algoriTheorem

1: Arbitrarily order the vertices in X_4. Let $i := 1$ and let $O_i := \pi_F(X)$.
2: **while** $i \le |X_4|$ **do**
3: Take the ith element x in the ordering of X_4.
4: Choose $w_x \in N_F(x) \setminus O_i$.
5: Let $M_x := x w_x \pi_{F^o}(w_x)$ and let $O_{i+1} := O_i \cup \{w_x\}$.
6: $i := i + 1$.
7: **end while**

We now show that the algorithm is correct: it produces the $|X_4|$ pairwise disjoint paths M_x for $x \in X_4$. The algorithm correctly terminates provided we can always find a suitable vertex w_x in Step 4; we next show this is the case.

Let $x \in X_4$ be the vertex in the ith position of the ordering. Let

$$O_x := O_i \cap N_F(x).$$

Since the path M_x is of length at most two, the set O_x represents the set of vertices in $N_F(x)$ that cannot be chosen as $w_x \in N_F(x)$ in the particular path M_x.

Excluding the path $x x_{F^o}^p$, there are exactly $d-1$ disjoint paths of length two in Q_d between x and F^o, each going through an element of $N_F(x)$. Thus, to show that there is a suitable vertex $w_x \in N_F(x)$ in Step 4, it suffices to show an injection between O_x and $X \setminus \{x, x_{F^o}^p, y\}$, which would imply $|O_x| \le d-2$. Observe that $y, x_{F^o}^p \in X \setminus O_x$ and $y \neq x_{F^o}^p$.

For every vertex $z \in O_x \cap X$, map z to z. For every $v \in O_x \setminus X$ with $v_{F^o}^p \in X$, map v to $v_{F^o}^p$; note that $v_{F^o}^p \neq y$, since $x \notin X_3$. For a vertex $w_u \in O_x \setminus X$ with $\pi_{F_o}(w_u) \notin X$ there exists a unique vertex $u \in X_4 \setminus O_x$ such that w_u is the *unique* vertex in $N_F(x)$ on the path M_u. Since $u \in X_4$, it

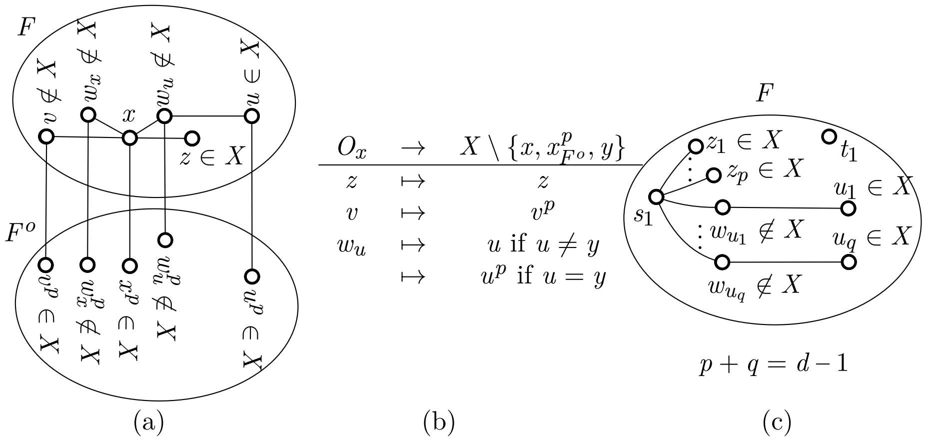


Figure 11.1: Auxiliary figure for the second scenario Theorem 11.1.8. (a) Types of neighbours of a vertex $x \in X_4$ for finding the path M_x. (b) An injective function from O_x to $X \setminus \{x, x^p_{F^o} y\}$. (c) A forbidden configuration where no path $L_1 := s_1 - t_1$ exists in F.

follows that $u^p_{F^o} \in X$. In this case, map w_u to u if $u \neq y$, otherwise map w_u to $u^p_{F^o}$. Note that $u \notin O_x$; otherwise the vertices u, x and w_u would all be pairwise neighbours but there are no triangles in Q_d. See Fig. 11.1(a)-(b) for a depiction of the different types of neighbours of the vertex $x \in X_4$ and the injection from O_x to $X \setminus \{x, x^p_{F^o}, y\}$.

The existence of an injection from O_x to $X \setminus \{x, x^p_{F^o}, y\}$ for $x \in X_4$ shows the existence of the vertex w_x in Step 4 of the algorithm and the termination of the algorithm. Thus the algorithm terminates with the required $|X_4|$ pairwise disjoint paths for vertices in X_4.

It remains to show that the $|X_1 \cup X_2 \cup X_3 \cup X_4|$ paths M_x are pairwise disjoint. Let $x_i \in X_i$ and $x_j \in X_j$ with $i, j \in [1, 4]$. For $i = 1, 2$ a path M_{x_i} is of length one, and necessarily disjoint from any other path. If $i = j$ and $i, j \geq 3$ then the corresponding M_{x_i} and M_{x_j} paths are disjoint by construction; if instead $i \neq j$, this is also clear. □

We now finalise this second scenario. So far the paths M_x have only been defined for terminals in $(X \cap V(F)) \setminus (X_0 \cup \{s_1, t_1\})$. For every vertex $x \in X \cap V(F^o)$, let $M_x := x$; in this way, the paths M_x have been defined for every vertex x in $X \setminus (X_0 \cup \{s_1, t_1\})$. Denote by X' the set of vertices in $M_x \cap V(F^o)$ over all paths M_x with $x \in X \setminus (X_0 \cup \{s_1, t_1\})$. Hence $|X'| \leq d - 1$. Let Y' be the corresponding pairing of the vertices in X': if $\{x, y\} \in Y$ with $x, y \in X \setminus (X_0 \cup \{s_1, t_1\})$, then the corresponding pair in Y' is $\{M_x \cap V(F^o), M_y \cap V(F^o)\}$. Note that, for a vertex $x \in X_1$, since $y = x^p_{F^o}$, the corresponding pair in Y' is $\{x^p_{F^o}, x^p_{F^o}\}$, as $M_x \cap V(F^o) = \{x^p_{F^o}\}$ and $M_{x^p_{F^o}} = x^p_{F^o}$.

The induction hypothesis on F^o gives that Y' is linked in F^o, and therefore, the existence of paths L^p_i in F^o between $M_{s_i} \cap V(F^o)$ and $M_{t_i} \cap V(F^o)$ for $s_i, t_i \in X \setminus (X_0 \cup \{s_1, t_1\})$. Each path L^p_i is then extended with the paths M_{s_i} and M_{t_i} to obtain a path $L_i := s_i - t_i$ for $s_i, t_i \in X \setminus (X_0 \cup \{s_1, t_1\})$. For a vertex $s_i \in X_0$, let $L_i := s_i t_i$.

It only remains to show the existence of a path $L_1 := s_1 - t_1$ in F pairwise disjoint from the paths L_i for $i \in [2, k]$. Assume that we cannot find a path L_1 pairwise disjoint from the other paths L_i with $i \in [2, k]$. Then there would be a set S in $V(F)$ separating s_1 from t_1. The set S would consist of terminal vertices in X_F and nonterminal vertices on some path M_x for $x \in X_3 \cup X_4$. The existence of each nonterminal vertex in S amounts to the existence of a terminal vertex in F^o, namely $x^p_{F^o}$, since $\pi_{F^o}(X_3 \cup X_4) \subset X$. Hence, the cardinality of S would equal $|X_F| + |X_3 \cup X_4|$, and thus, it would be at most $|X_F| + |X \cap V(F^o)| = X \setminus \{s_1, t_1\}| = d - 1$. By the $(d-1)$-connectivity of F, the set S would have cardinality $d - 1$, which would imply that every terminal in X_F and every nonterminal in F that lies on a path M_x for $x \in X_3 \cup X_4$ are in S. Again, by Proposition 11.1.3, the set S would consist of the neighbours of s_1 or t_1, say of s_1. In this configuration all the vertices of X would be in F, which is a contradiction. Indeed, since there is no edge between any two vertices in S (Corollary 11.1.4), no nonterminal on a path M_x is in S, and therefore, $S = X_F$, or equivalently, $X \subset V(F)$, as desired. The existence of the path L_1 finally settles the second scenario. See Fig. 11.1(c).

It is instructive for the reader to convince himself/herself that the proof of the second scenario works well by verifying the existence of the paths M_x and the existence of the path L_1 for the cubes Q_4 and Q_5.

Finally, let us move onto the third and final scenario: every pair in Y is at distance d. From Lemma 11.1.2 it follows that there exists a pair $\{F, F^o\}$ of opposite facets of Q_d that is not associated with $X_{s_1} := X \backslash \{s_1\}$, since $|X \backslash \{s_1\}| \le d$ and there are d pairs of the form $\{F, F^o\}$. This means that for every $x \in X_{s_1} \cap V(F)$, $\pi_{F^o}^{Q_d}(x) \notin X_{s_1}$ and that for every $x \in X_{s_1} \cap V(F^o)$, $\pi_F^{Q_d}(x) \notin X_{s_1}$. Without loss of generality, assume that $s_1, \dots, s_k \in F^o$ and $t_1, \dots, t_k \in F$. We again apply Lemma 11.1.2, now to the set $W := \{\pi_F^{Q_d}(s_2), \dots, \pi_F^{Q_d}(s_k), t_2, \dots, t_k\}$ on the facet F. So there are at most $|W| - 1 \le d - 2$ pairs of opposite ridges of F associated with W in F. That is, there exists a pair $\{R, R^o\}$ that is not associated with W in F. We now consider the projection π^F in F of W onto the ridge opposite to the one containing t_1, say $t_1 \in R$. Let $W^p := \pi_{R^o}^F(W)$. Let $s_i' := \pi_{R^o}^F(\pi_F^{Q_d}(s_i))$ and $t_i' := \pi_{R^o}^F(t_i)$ for $i \ne 1$. Then $W^p = \{s_2', \dots, s_k', t_2', \dots, t_k'\}$ for $i \in [2, k]$; notice that $\operatorname{dist}_F(\pi_F^{Q_d}(s_i), t_i) = d - 1$ for $i \in [2, k]$, and therefore, $\operatorname{dist}_{R^o}(s_i', t_i') = d - 2$ for $i \in [2, k]$.

We find $k - 1$ disjoint paths L_i' in R^o between s_i' and t_i' using the induction hypothesis on R^o for $d \ge 6$. For the case of $d = 5$, use Proposition 9.0.5, since $\operatorname{dist}_{R^o}(s_i', t_i') = 3$ for $i = 2, 3$ in the 3-cube R^o; in other words, the vertices s_2',s_3', t_2' and t_3' are not in a 2-face of R^o. Notice we must use the projections π^F of W onto R^o, since $t_1 \notin W$, which may cause that $\pi_R^F(w) = t_1$ for some $w \in W$. Thus we have disjoint $s_i - t_i$ paths L_i equal to $s_i \pi_F^{Q_d}(s_i) s_i' L_i' t_i' t_i$ for $i \in [2, k]$. Observe that, for $i \in [2, k]$, $t_1 \notin L_i$ and that the subpath $\pi_F^{Q_d}(s_i) L_i t_i$ of L_i is contained in F. By Menger's theorem we can find a path L_1' from $\pi_{F^o}^{Q_d}(t_1)$ to s_1 in F^o that is disjoint from $\{s_2, \dots, s_k\}$, since F^o is $(2k-2)$-connected. Also note that $\pi_{F^o}^{Q_d}(t_1)$ is not in X as a result of $\{F, F^o\}$ being not associated with X_{s_1} and $\operatorname{dist}_{Q_d}(s_1, t_1) = d$. The path $L_1 := s_1 L_1' \pi_{F^o}^{Q_d}(t_1) t_1$ is our final required path. □

11.1.3 Linkedness Inside the Cube

We verify that the link of a vertex in a $(d+1)$-cube, which by Proposition 9.0.11 is combinatorially equivalent to a (cubical) d-polytope, is $\lfloor (d+1)/2 \rfloor$-linked for $d \ne 3$ (Proposition 11.1.9). In another abuse of terminology, we often think of the link as the corresponding (cubical) d-polytope.

Proposition 11.1.9. *For every $d \ne 3$, the link of a vertex in a $(d+1)$-cube Q_{d+1} is $\lfloor (d+1)/2 \rfloor$-linked.*

Proof. The proposition trivially holds for the cases of $d = 1, 2$, so assume $d \ge 4$.

Let $k := \lfloor (d+1)/2 \rfloor$. Let v and v^o be opposite vertices of $G(Q_{d+1})$; that is, $\operatorname{dist}_{Q_{d+1}}(v, v^o) = d+1$. Let X be a given set of $2k$ vertices in $\operatorname{link}(v, Q_{d+1})$ and let $Y := \{\{s_1, t_1\}, \dots, \{s_k, t_k\}\}$ be an arbitrary pairing of the vertices in X. From Remark 9.0.9 it follows that $\operatorname{link}(v, Q_{d+1})$ is the subcomplex $Q_{d+1} - \{v, v^o\}$ of Q_{d+1}. We show that Y is linked in $\operatorname{link}(v, Q_{d+1})$.

Since $|X| - 1 \le d$ and there are $d + 1$ pairs of opposite facets in Q_{d+1}, from Lemma 11.1.2 there exists a pair $\{F, F^o\}$ of opposite facets of Q_{d+1} that is not associated with X. This means that, for every $x \in X \cap V(F)$, its projection $\pi_{F^o}^{Q_{d+1}}(x) \notin X$, and that, for every $x \in X \cap V(F^o)$, its projection $\pi_F^{Q_{d+1}}(x) \notin X$. Henceforth we write π_F rather than $\pi_F^{Q_{d+1}}$. Assume that $v \in F$ and $v^o \in F^o$. We consider two cases based on the number of terminals in the facet F; for the sake of readability, the second case is in turn decomposed into two subcases highlighted in bold.

In what follows we implicitly use the d-connectivity of F or F^o. Without loss of generality assume $|X \cap V(F)| \geq |X \cap V(F^o)|$.

Case 1. $|X \cap V(F)| = d + 1$.

Since F is a d-cube, it is $\lfloor (d+1)/2 \rfloor$-linked by Theorem 11.1.8, and hence, we can find k pairwise disjoint paths $L_1, \ldots, L_k$ in F between s_i and t_i for $i \in [1, k]$. If no path L_i passes through v, we are done. So suppose one of those paths, say L_1, passes through v; there can be only one such path. If neither the projection of s_1 onto F^o nor the projection of t_1 onto F^o is v^o, then find a $\pi_{F^o}(s_1) - \pi_{F^o}(t_1)$ path $\bar{L}_1$ in F^o that avoids v^o. So L_1 would then become $s_1 \pi_{F^o}(s_1) \bar{L}_1 \pi_{F^o}(t_1) t_1$. If the projection of either s_1 or t_1 onto F^o is v^o, say that of s_1, then, since $\operatorname{dist}_{Q^{d+1}}(v, v^o) = d + 1 \geq 5$ and $\operatorname{dist}(s_1, v) = d \geq 4$, there must be a neighbour w of s_1 on L_1 that is different from v. Find a $\pi_{F^o}(w) - \pi_{F^o}(t_1)$ path $\bar{L}_1$ in F^o that avoids v^o. So L_1 would then become $s_1 w \pi_{F^o}(w) \bar{L}_1 \pi_{F^o}(t_1) t_1$.

Case 2. $|X \cap V(F)| \leq d$.

Let $X^p := \pi_F(X)$. The set X^p comprises the terminals in $X \cap V(F)$ together with the projections onto F of vertices in $X \cap V(F^o)$. Then $|X^p| \leq d + 1$.

First suppose that some vertex in $X \cap V(F^o)$, say t_1, is adjacent to v: $\pi_F(t_1) = v$. We must have that either $s_1 \in F^o$ or $s_1 \in F$.

Suppose $s_1 \in F^o$. Find an X-valid path $L_1 := s_1 - t_1$ in F^o using the d-connectivity of F^o as there are at most d terminals in F^o. Thanks to Theorem 11.1.8, F is k-linked, and thus we can find $k - 1$ disjoint paths $\bar{L}_2, \ldots, \bar{L}_k$ between $\pi_F(s_i)$ and $\pi_F(t_i)$ for $i \in [2, k]$, all avoiding v. Each such path $\bar{L}_i$ extends to a path $L_i := s_i - t_i$, if necessary. So we are done in this scenario and ready to assume $s_1 \in F$.

Assume $s_1 \in F$. Observe that now $s_1 \in X^p$. The k-linkedness of F ensures that in F there are k disjoint paths $M_1 := s_1 - v$ and $\bar{L}_i := \pi_F(s_i) - \pi_F(t_i)$ for $i \in [2, k]$. As before, each path $\bar{L}_i$ ($i \in [2, k]$) extends to a path $L_i := s_i - t_i$, if necessary. If v^o is not the projection of s_1 onto F^o, then find an X-valid $\pi_{F^o}(s_1) - t_1$ path $\bar{L}_1$ in F^o using the d-connectivity of F^o. Then L_1 would become $s_1 \pi_{F^o}(s_1) \bar{L}_1 t_1$, and so we are also home in this scenario. Otherwise v^o is the projection of s_1 onto F^o, in which case $\operatorname{dist}_F(s_1, v) = d \geq 4$. There is a neighbour $w \in V(F)$ of s_1 on the path M_1, which is different from v; observe that $\pi_{F^o}(w) \notin X$ since $w \notin X^p$. Find an X-valid path $\pi_{F^o}(w) - t_1$ path $\bar{L}_1$ in F^o (here use again the d-connectivity of F^o). So L_1 would then become $s_1 w \pi_{F^o}(w) \bar{L}_1 t_1$. This settles the subcase of some vertex in $X \cap V(F^o)$ being adjacent to v.

Finally, assume no vertex in $X \cap V(F^o)$ is adjacent to v. This subcase then reduces to Case 1 with the set $X^p \subset V(F)$ playing the role of X, and so we obtain paths $\bar{L}_i := \pi_F(s_i) - \pi_F(t_i)$ in F for $i \in [1, k]$, thanks to the k-linkedness of F. Observe that, following the reasoning in Case 1, if one of the paths $\bar{L}_i$, say $\bar{L}_1$, passes through v, then the projections $\pi_{F^o}(s_1)$ or $\pi_{F^o}(t_1)$ onto F^o may need to be considered. Each path $\bar{L}_i$ ($i \in 1, k$) extends to a path $L_i := s_i - t_i$, if necessary. This completes the proof of the case and of the proposition. □

Proposition 11.1.9 fails for $d = 3$ because of the possible presence of Configuration F (Definition 9.0.7) in the link of a vertex of the 4-cube. For specific examples of Configuration F, consider Fig. 9.2 (b)-(c) and let s_1, s_2, t_1, t_2 be the vertices labelled as $1, 2, 3, 4$, respectively.

11.1.4 Strong Linkedness of the Cube

With Proposition 9.0.8, Lemma 11.1.7, and Theorem 11.1.8 at hand, it can be verified that 4-polytopes and d-cubes for $d \neq 3$ enjoy a property marginally stronger than linkedness: *strong linkedness*. A d-polytope P is *strongly $\lfloor (d+1)/2 \rfloor$-linked* if its graph has at least $d + 1$ vertices and, for every set X of exactly $d+1$ vertices and every pairing Y with $\lfloor (d+1)/2 \rfloor$ pairs from X, the set Y is linked in $G(P)$ and each path joining a pair in Y avoids the vertices in X not being paired in Y. For odd $d = 2k-1$ the properties of strongly k-linkedness and k-linkedness coincide, since every vertex in X is paired in Y; but they differ for even $d = 2k$. Theorem 11.1.10 shows that 4-polytopes are strongly 2-linked while Theorem 11.1.11 shows that d-cubes for $d \neq 3$ are strongly $\lfloor (d+1)/2 \rfloor$-linked.

Theorem 11.1.10 (Strong 2-linkedness of 4-polytopes)**.** *Every cubical 4-polytope is strongly 2-linked.*

Proof. Let G denote the graph of a 4-polytope P embedded in $\mathbb{R}^4$. Let X be a set of five vertices in G. Arbitrarily pair four vertices of X to obtain $Y := \{\{s_1, t_1\}, \{s_2, t_2\}\}$. Let x be the vertex of X not being paired in Y. We aim to find two disjoint paths $L_1 := s_1 - t_1$ and $L_2 := s_2 - t_2$ such that each path L_i avoids the vertex x. The proof is very similar to that of Propositions 9.0.5 and 9.0.8.

Consider a linear function f that vanishes on a linear hyperplane H passing through $\{s_1, s_2, t_1, x\}$. Assume that $f(y) > 0$ for some $y \in P$ and that $f(t_2) \geq 0$.

Suppose first that H is a supporting hyperplane of a facet F of P. If $t_2 \notin V(F)$, then find an X-valid $L_1 := s_1 - t_1$ path in F using the 3-connectivity of F. Then use Lemma 9.0.3 to find an X-valid $s_2 - t_2$ path in which each inner vertex has positive f-value. If instead $t_2 \in F$, then $X \subset V(F)$ and Lemma 11.1.7 ensures the existence of an X-valid $s_i - t_i$ path in F for some $i = 1, 2$, say for $i = 1$. Then use Lemma 9.0.3 to find an X-valid $s_2 - t_2$ path in which each inner vertex has positive f-value. So assume H intersects the interior of P. Then there is a vertex in P with f-value greater than zero and a vertex with f-value less than zero. In this case, use Lemma 9.0.3 to find an X-valid $s_1 - t_1$ path in which each inner vertex has negative f-value and an X-valid $s_2 - t_2$ path in which each inner vertex has positive f-value. □

Not every 4-polytope is strongly 2-linked. Take a two-fold pyramid P over a quadrangle Q. Then P is a 4-polytope on six vertices. Let $V(P) := \{s_1, s_2, t_1, t_2, x\}$ so that the sequence s_1, s_2, t_1, t_2 appears in cyclic order in Q and the vertex x is in $V(P) \setminus V(Q)$. To see that P is not strongly 2-linked, observe that, for every two paths $s_1 - t_1$ and $s_2 - t_2$ in P, either they intersect or one of them contains x.

Theorem 11.1.11 (Strong linkedness of the cube)**.** *For every $d \neq 3$, a d-cube is strongly $\lfloor (d+1)/2 \rfloor$-linked.*

Proof. It suffices to prove the result for $d = 2k$. Let X be a set of $d+1$ vertices in the d-cube for $d \neq 3$. Arbitrarily pair $2k$ vertices in X to obtain $Y := \{\{s_1, t_1\}, \ldots, \{s_k, t_k\}\}$. Let x be the vertex of X not being paired in Y. We aim to find a Y-linkage $\{L_1, \ldots, L_k\}$ where each path L_i joins the pair $\{s_i, t_i\}$ and avoids the vertex x.

The result for $d = 4$ is given by Theorem 11.1.10. So assume $d \geq 6$.

From Lemma 11.1.2 it follows that there exists a pair $\{F, F^o\}$ of opposite facets of Q_d that is not associated with $X_x := X \setminus \{x\}$, since $|X \setminus \{x\}| = d$ and there are d pairs $\{F, F^o\}$ of opposite facets in Q_d. Assume $x \in V(F^o)$. Let $X^p := \pi_F(X_x)$; that is, the set X^p comprises the vertices in $X_x \cap V(F)$ plus the projections of $X_x \cap V(F^o)$ onto F. Denote by Y^p the corresponding pairing of the vertices in X^p; that is, $Y^p := \{\{\pi_F(s_1), \pi_F(t_1)\}, \ldots, \{\pi_F(s_k), \pi_F(t_k)\}\}$. Then $|X^p| = d$ and $|Y^p| = k$. Find a Y^p-linkage $\{L_1^p, \ldots, L_k^p\}$ in F with $L_i^p := \pi_F(s_i) - \pi_F(t_i)$ by resorting to the k-linkedness of F (Theorem 11.1.8). Adding $s_i \in V(F^o)$ or $t_i \in V(F^o)$ to the path L_i^p, if necessary, we extend the linkage $\{L_1^p, \ldots, L_k^p\}$ to the required Y-linkage. □

11.2 Connectivity of Cubical Polytopes

The aim of this section is to present a couple of results related to the connectivity of strongly connected complexes in cubical polytopes, whose proof ideas originated in Proposition 10.2.2.

The proof idea in Proposition 10.2.2 can be pushed a bit further to obtain a rather technical result that we prove next. But first we state two simple but useful remarks.

Remark 11.2.1. Let P be a cubical d-polytope. Let v and F be a vertex and a face containing the vertex in P, respectively. Let v^o be the vertex of F opposite to v. The smallest face in the polytope containing both v and v^o is precisely F.

Remark 11.2.2. For any two faces F, J of a polytope, with F not contained in J, there is a facet containing J but not F. In particular, for any two distinct vertices of a polytope, there is a facet containing one but not the other.

Two facet-ridge paths are *independent* if they do not share an inner facet.

Lemma 11.2.3. *Let P be a cubical d-polytope with $d \geq 4$. Let s_1 be any vertex in P and let $\mathcal{S}_1$ be the star of s_1 in the boundary complex of P. Let s_2 be any vertex in $\mathcal{S}_1$, other than s_1. Define the following sets.*

- *F_1 in $\mathcal{S}_1$, a facet containing s_1 but not s_2.*
- *F_{12} in $\mathcal{S}_1$, a facet containing s_1 and s_2.*
- *$\mathcal{S}_{12}$, the star of s_2 in $\mathcal{S}_1$; that is, the subcomplex of $\mathcal{S}_1$ formed by the facets of P in $\mathcal{S}_1$ containing s_2.*
- *$\mathcal{A}_1$, the antistar of F_1 in $\mathcal{S}_1$.*
- *$\mathcal{A}_{12}$, the subcomplex of $\mathcal{S}_{12}$ induced by $V(\mathcal{S}_{12}) \setminus (V(F_1) \cup V(F_{12}))$.*

Then the following assertions hold.

(i) The complex $\mathcal{S}_{12}$ is a strongly connected $(d-1)$-subcomplex of $\mathcal{S}_1$.

(ii) If there are more than two facets in $\mathcal{S}_{12}$, then, between any two facets of $\mathcal{S}_{12}$ that are different from F_{12}, there exists a $(d-1, d-2)$-path in $\mathcal{S}_{12}$ that does not contain the facet F_{12}.

(iii) If $\mathcal{S}_{12}$ contains more than one facet, then the subcomplex $\mathcal{A}_{12}$ of $\mathcal{S}_{12}$ contains a spanning strongly connected $(d-3)$-subcomplex.

Proof. Let us prove (i). Let ψ define the natural anti-isomorphism from the face lattice of P to the face lattice of its dual P^*. The facets in $\mathcal{S}_1$ correspond to the vertices in the facet $\psi(s_1)$ in P^* corresponding to s_1; likewise for the facets in $\operatorname{star}(s_2, \mathbb{B}(P))$ and the vertices in $\psi(s_2)$. The facets in $\mathcal{S}_{12}$ correspond to the vertices in the nonempty face $\psi(s_1) \cap \psi(s_2)$ of P^*. The existence a facet-ridge path in $\mathcal{S}_{12}$ between any two facets J_1 and J_2 of $\mathcal{S}_{12}$ amounts to the existence of a vertex-edge path in $\psi(s_1) \cap \psi(s_2)$ between $\psi(J_1)$ and $\psi(J_2)$. That $\mathcal{S}_{12}$ is a strongly connected $(d-1)$-complex now follows from the connectivity of the graph of $\psi(s_1) \cap \psi(s_2)$ (Balinski's theorem), as desired.

We proceed with the proof of (ii). If there are more than two facets in $\mathcal{S}_{12}$, then the face $\psi(s_1) \cap \psi(s_2)$ is at least bidimensional. As a result, the graph of $\psi(s_1) \cap \psi(s_2)$ is at least 2-connected by Balinski's theorem. By Menger's theorem, there are at least two independent vertex-edge paths in $\psi(s_1) \cap \psi(s_2)$ between $\psi(J_1)$ and $\psi(J_2)$. Pick one such path L^* that avoids the vertex $\psi(F_{12})$ of $\psi(s_1) \cap \psi(s_2)$. Dualising this path L^* gives a $(d-1, d-2)$-path between J_1 and J_2 in $\mathcal{S}_{12}$ that does not contain the facet F_{12}.

We now prove (iii). Assume that $\mathcal{S}_{12}$ contains more than one facet. Consider now a facet $F \in \mathcal{S}_{12}$ other than F_{12}; it exists by our assumption on $\mathcal{S}_{12}$. We need some additional notation.

- Let $\mathcal{A}_1^F$ denote the subcomplex $F - V(F_1)$; that is, $\mathcal{A}_1^F$ is the antistar of $F \cap F_1$ in F.
- Let $\mathcal{A}_{12}^F$ denote the subcomplex $F - (V(F_1) \cup V(F_{12}))$, the subcomplex of F induced by $V(F) \setminus (V(F_1) \cup V(F_{12}))$.

We require the following claim.

Claim 9. *$\mathcal{A}_{12}^F$ contains a spanning strongly connected $(d-3)$-subcomplex $\mathcal{C}^F$.*

Proof. We first show that $\mathcal{A}_{12}^F \neq \emptyset$. Denoting by s_1^o the vertex in F opposite to s_1, we have that s_1^o is not in F_1 or in F_{12} by Remark 11.2.1. So s_1^o is in $\mathcal{A}_{12}^F$.

Notice that $s_1 \notin \mathcal{A}_1^F$. From Lemma 10.1.2 it follows that $\mathcal{A}_1^F$ is a strongly connected $(d-2)$-subcomplex of F. Write

$$\mathcal{A}_1^F = \mathcal{C}(R_1) \cup \cdots \cup \mathcal{C}(R_m),$$

where R_i is a $(d-2)$-face of F for $i \in [1, m]$. No ridge R_i is contained in F_{12}; otherwise $R_i = F \cap F_{12}$, which implies that $s_1 \in R_i$, and therefore that $s_1 \in \mathcal{A}_1^F$, a contradiction. Moreover, $s_1^o \in R_i$ for every $i \in [1, m]$, since every ridge of F contains either s_1 or s_1^o and $s_1 \notin R_i$.

Let $\mathcal{C}_i := \mathbb{B}(R_i) - V(F_{12})$. As $R_i \not\subset F_{12}$, we have $\dim R_i \cap F_{12} \leq d-3$. Hence $\mathcal{C}_i$ is nonempty. If $R_i \cap F_{12} \neq \emptyset$, then $\mathcal{C}_i$ denotes the antistar of $R_i \cap F_{12}$ in R_i, a spanning strongly connected $(d-3)$-subcomplex of R_i by Lemma 10.1.2. If $R_i \cap F_{12} = \emptyset$, then $\mathcal{C}_i$ denotes the boundary complex of R_i, again a spanning strongly connected $(d-3)$-subcomplex of R_i.

Let $\mathcal{C}^F := \bigcup \mathcal{C}_i$. Then the complex $\mathcal{C}^F$ is a spanning $(d-3)$-subcomplex of $\mathcal{A}_{12}^F$; we show it is strongly connected.

Take any two $(d-3)$-faces W and W' in $\mathcal{C}^F$. We find a $(d-3, d-4)$-path L in $\mathcal{C}^F$ between W and W'. There exist ridges R and R' in $\mathcal{A}_1^F$ with $W \subset R$ and $W' \subset R'$. Since $\mathcal{A}_1^F$ is a strongly connected $(d-2)$-complex, there is a $(d-2, d-3)$-path $R_{i_1} \ldots R_{i_p}$ in $\mathcal{A}_1^F$ between $R_{i_1} = R$ and $R_{i_p} = R'$, with $R_{i_j} \in \mathcal{A}_1^F$ for $j \in [1, p]$.

We already established that $s_1^o \in R_{i_j}$ for every $j \in [1, p]$ and that $s_1^o \notin F_{12}$. Consequently, we get that $R_{i_\ell} \cap R_{i_{\ell+1}} \not\subset F_{12}$, and therefore, that $\dim R_{i_\ell} \cap R_{i_{\ell+1}} \cap F_{12} \leq d-4$ for every $\ell \in [1, p-1]$. Hence the subcomplex $\mathbb{B}_{i_\ell} := \mathbb{B}(R_{i_\ell} \cap R_{i_{\ell+1}}) - V(F_{12})$ of $\mathbb{B}(R_{i_\ell} \cap R_{i_{\ell+1}})$ is a nonempty, strongly connected $(d-4)$-complex by Lemma 10.1.2; in particular, it contains a $(d-4)$-face U_{i_ℓ}. Furthermore, $\mathbb{B}_{i_\ell} \subset \mathcal{C}_{i_\ell} \cap \mathcal{C}_{i_{\ell+1}}$.

Resorting to the strong $(d-3)$-connectivity of $\mathcal{C}_{i_j}$ for $j \in [1, p]$, we build the path L by concatenating $(d-3, d-4)$-paths $L_{i_1}, \ldots, L_{i_p}$ where L_{i_j} lies in $\mathcal{C}_{i_j}$.

- The path L_{i_1} is a $(d-3, d-4)$-path in $\mathcal{C}_{i_1}$ between W and some $(d-3)$-face W'_{i_1} in $\mathcal{C}_{i_1}$ that contains the $(d-4)$-face $U_{i_1} \in \mathcal{C}_{i_1} \cap \mathcal{C}_{i_2}$.
- The path L_{i_2} is a $(d-3, d-4)$-path in $\mathcal{C}_{i_2}$ between a $(d-3)$-face $W_{i_2} \in \mathcal{C}_{i_2}$ that contains U_{i_1} and some $(d-3)$-face W'_{i_2} in $\mathcal{C}_{i_2}$ that contains the $(d-4)$-face $U_{i_2} \in \mathcal{C}_{i_2} \cap \mathcal{C}_{i_3}$.
- The path $L_{i_1} \cup L_{i_2}$ is a $(d-3, d-4)$-path in $\mathcal{C}_{i_1} \cup \mathcal{C}_{i_2}$ between W and W'_{i_2}.
- Continuing this process process, we get a path L_{i_p} in $\mathcal{C}_{i_p}$ between a $(d-3)$-face $W_{i_p} \in \mathcal{C}_{i_p}$ that contains $U_{i_{p-1}}$ and the $(d-3)$-face W' in $\mathcal{C}_{i_p}$.
- Concatenating all the paths $L_{i_1}, \ldots, L_{i_p}$, we obtain the path L, a $(d-3, d-4)$-path in $\mathcal{C}_{i_1} \cup \cdots \cup \mathcal{C}_{i_p} \subseteq \mathcal{C}^F$.

The existence of the path L between W and W' completes the proof of Claim 9. $\square$

We are now ready to complete the proof of (iii). The proof goes along the lines of the proof of Claim 9. We let $\mathcal{S}_{12} = \bigcup_{i=1}^{m} \mathcal{C}(J_i)$, where the facets $J_1, \dots, J_m$ are all the facets in P containing s_1 and s_2.

For every $i \in [1, m]$ we let $\mathcal{C}^{J_i}$ be the spanning strongly connected $(d-3)$-subcomplex in $\mathcal{A}_{12}^{J_i}$ given by Claim 9. And we let $\mathcal{C} := \bigcup \mathcal{C}^{J_i}$. Then $\mathcal{C}$ is a spanning $(d-3)$-subcomplex of $\mathcal{A}_{12}$; we show it is strongly connected.

If there are exactly two facets in $\mathcal{S}_{12}$, namely F_{12} and some other facet F, then the complex $\mathcal{A}_{12}$ coincides with the complex $\mathcal{A}_{12}^{F}$. The strong $(d-3)$-connectivity of $\mathcal{A}_{12}^{F}$ is then settled by Claim 9. Hence assume that there are more than two facets in $\mathcal{S}_{12}$; this implies that the smallest face containing s_1 and s_2 in $\mathcal{S}_{12}$ is at most $(d-3)$-dimensional.

Take any two $(d-3)$-faces W and W' in $\mathcal{C}$. If W and W' belong to the same facet F in $\mathcal{S}_{12}$, which is different from F_{12}, then W and W' are both in $\mathcal{A}_{12}^{F}$, and consequently, Claim 9 gives the desired $(d-3, d-4)$-path between W and W' in $\mathcal{A}_{12}^{F} \subseteq \mathcal{C}$. So assume there are distinct facets J and J' in $\mathcal{S}_{12}$ such that $W \subset J$ and $W' \subset J'$. Then $J \neq F_{12}$ and $J' \neq F_{12}$. By (ii), we can find a $(d-1, d-2)$-path $J_{i_1} \dots J_{i_q}$ in $\mathcal{S}_{12}$ between $J_{i_1} = J$ and $J_{i_q} = J'$ such that $J_{i_j} \neq F_{12}$ for $j \in [1, q]$.

We next show that, for each $\ell \in [1, q-1]$, there exists a $(d-4)$-face U_{i_ℓ} in $C^{J_{i_\ell}} \cap C^{J_{i_{\ell+1}}}$. As $J_{i_\ell}, J_{i_{\ell+1}} \neq F_{12}$, we obtain that $\mathbb{B}(J_{i_\ell} \cap J_{i_{\ell+1}}) - V(F_{12})$ is a nonempty, strongly connected $(d-3)$-subcomplex (Lemma 10.1.2); in particular, it contains a $(d-3)$-face K_{i_ℓ}. We pick U_{i_ℓ} in K_{i_ℓ}. On one hand, $J_{i_\ell} \cap J_{i_{\ell+1}} \not\subset F_1$, since $J_{i_\ell}, J_{i_{\ell+1}} \neq F_1$. On the other hand, $K_{i_\ell} \not\subset F_1$; otherwise $K_{i_\ell} = J_{i_\ell} \cap J_{i_{\ell+1}} \cap F_1$, a contradiction because $s_1 \notin K_{i_\ell}$ but $s_1 \in J_{i_\ell} \cap J_{i_{\ell+1}} \cap F_1$. As a consequence, $\mathbb{B}(K_{i_\ell}) - V(F_1)$ is a nonempty, strongly connected $(d-4)$-subcomplex (Lemma 10.1.2 again); in particular, it contains a desired $(d-4)$-face U_{i_ℓ}.

Let us get back to the $(d-3)$-faces $W \in \mathcal{C}^{J_{i_1}}$ and $W' \in \mathcal{C}^{J_{i_q}}$ and the path $J_{i_1} \dots J_{i_q}$ in $\mathcal{S}_{12}$. For every $\ell \in [1, q-1]$ consider a $(d-4)$-face U_{i_ℓ} in $\mathcal{C}^{J_{i_\ell}} \cap \mathcal{C}^{J_{i_{\ell+1}}}$. Using the strong $(d-3)$-connectivity of $\mathcal{C}^{J_{i_\ell}}$ and $\mathcal{C}^{J_{i_{\ell+1}}}$, we build the path L between W and W' by concatenating $(d-3, d-4)$-paths $L_{i_1}, \dots, L_{i_q}$.

- The path L_{i_1} is a $(d-3, d-4)$-path in $\mathcal{C}^{J_{i_1}}$ between W and some $(d-3)$-face W'_{i_1} in $\mathcal{C}^{J_{i_1}}$ that contains the $(d-4)$-face $U_{i_1} \in \mathcal{C}^{J_{i_1}} \cap \mathcal{C}^{J_{i_2}}$.
- The path L_{i_2} is a $(d-3, d-4)$-path in $\mathcal{C}^{J_{i_2}}$ between a $(d-3)$-face $W_{i_2} \in \mathcal{C}^{J_{i_2}}$ that contains U_{i_1} and some $(d-3)$-face W'_{i_2} in $\mathcal{C}^{J_{i_2}}$ that contains the $(d-4)$-face $U_{i_2} \in \mathcal{C}^{J_{i_2}} \cap \mathcal{C}^{J_{i_3}}$.
- The path $L_{i_1} \cup L_{i_2}$ is a $(d-3, d-4)$-path in $\mathcal{C}^{J_{i_1}} \cup \mathcal{C}^{J_{i_2}}$ between W and W'_{i_2}.
- Continuing this process process we get a path L_{i_q} in $\mathcal{C}^{J_{i_q}}$ between a $(d-3)$-face $W_{i_q} \in \mathcal{C}^{J_{i_q}}$ that contains $U_{i_{q-1}}$ and the $(d-3)$-face W' in $C^{J_{i_q}}$.
- Concatenating all the paths $L_{i_1}, \dots, L_{i_q}$, we obtain the path L, a $(d-3, d-4)$-path in $\mathcal{C}^{J_{i_1}} \cup \dots \cup \mathcal{C}^{J_{i_q}} \subseteq \mathcal{C} \subseteq \mathcal{A}_{12}$.

The path L between W and W' completes the proof of the lemma. $\square$

11.3 Linkedness of Cubical Polytopes

The aim of this section is to prove that, for every $d \neq 3$, a cubical d-polytope is $\lfloor (d+1)/2 \rfloor$-linked (Theorem 11.3.5).

We proceed with a simple lemma proved in [164, Sect. 3].

Lemma 11.3.1 ([164, Sect. 3])**.** *Let G be a $2k$-connected graph and let G' be a k-linked subgraph of G. Then G is k-linked.*

Theorem 11.1.8 in conjunction with Lemma 11.3.1 answers the question posed in [165, Question 5.4.12].

Proposition 11.3.2. *For every $d \geq 1$, a cubical d-polytope is $\lfloor d/2 \rfloor$-linked.*

Proof. Let P be a cubical d-polytope. The results for $d = 1, 2$ are trivial. The case of $d = 3$ follows from the connectivity of the graph of P, while the case of $d = 4$ follows from Proposition 9.0.8. For $d \geq 5$, since a facet of P is a $(d-1)$-cube with $d-1 \geq 4$, by Theorem 11.1.8 it is $\lfloor d/2 \rfloor$-linked. So Lemma 11.3.1 together with the d-connectivity of the graph of P gives the desired result. □

The remaining part of the section is devoted to proving Theorem 11.3.5 for odd $d \geq 5$. Since $\lfloor d/2 \rfloor = \lfloor (d+1)/2 \rfloor$ for even d, Proposition 11.3.2 trivially establishes Theorem 11.3.5 in this case.

Proposition 11.3.3. *Let F be a facet in the star $\mathcal{S}$ of a vertex in a cubical d-polytope. Then, for every $d \geq 2$, the antistar of F in $\mathcal{S}$ is $\lfloor (d-2)/2 \rfloor$-linked.*

Proof. Let $\mathcal{S}$ be the star of a vertex s in a cubical d-polytope and let F be a facet in the star $\mathcal{S}$. Let $\mathcal{A}$ denote the antistar of F in $\mathcal{S}$. The case of $d \leq 5$ follows from the strong $(d-2)$-connectivity of $\mathcal{A}$ (Proposition 10.2.2), which implies the $(d-2)$-connectivity of the graph of $\mathcal{A}$; refer to Proposition 9.0.1. So assume $d \geq 6$.

There is a $(d-2)$-face R in $\mathcal{A}$. Indeed, take a $(d-2)$-face R' in F containing s and consider the other facet F' in $\mathcal{S}$ containing R'; the $(d-2)$-face of F' disjoint from R' is the desired R. By Theorem 11.1.8 the ridge R is $\lfloor (d-1)/2 \rfloor$-linked but we only require it to be $\lfloor (d-2)/2 \rfloor$-linked. By Proposition 10.2.2 the graph of $\mathcal{A}$ is $(d-2)$-connected. Combining the linkedness of R and the connectivity of the graph of $\mathcal{A}$ settles the proposition by virtue of Lemma 11.3.1. □

For a pair of opposite facets $\{F, F^o\}$ in a cube, the restriction of the projection $\pi_{F^o} : Q_d \to F^o$ (Definition 11.1.1) to F is a bijection from $V(F)$ to $V(F^o)$. With the help of π, given the star $\mathcal{S}$ of a vertex s in a cubical polytope and a facet F in $\mathcal{S}$, we can define an injection from the vertices in F, except the vertex opposite to s, to the antistar of F in $\mathcal{S}$. Defining this injection is the purpose of Lemma 11.3.4.

Lemma 11.3.4. *Let F be a facet in the star $\mathcal{S}$ of a vertex s in a cubical d-polytope. Then there is an injective function, defined on the vertices of F except the vertex s^o opposite to s, that maps each such vertex in F to a neighbour in $V(\mathcal{S}) \setminus V(F)$.*

Proof. We construct the aforementioned injection f between $V(F) \setminus \{s^o\}$ and $V(\mathcal{S}) \setminus V(F)$ as follows. Let $R_1, \ldots, R_{d-1}$ be the $(d-2)$-faces of F containing s, and let $J_1, \ldots, J_{d-1}$ be the other facets of $\mathcal{S}$ containing $R_1, \ldots, R_{d-1}$, respectively. Every vertex in F other than s^o lies in $R_1 \cup \cdots \cup R_{d-1}$. Let R_i^o be the $(d-2)$-face in J_i that is opposite to R_i for $i \in [1, d-1]$. For every vertex v in $V(R_j) \setminus (V(R_1) \cup \cdots \cup V(R_{j-1}))$ define $f(v)$ as the projection π in J_j of v onto $V(R_j^o)$, namely $f(v) := \pi_{R_j^o}(v)$; observe that $\pi_{R_j^o}(v) \in$

$V(R_j^o) \setminus (V(R_1^o) \cup \cdots \cup V(R_{j-1}^o))$. Here R_{-1} and R_{-1}^o are empty sets. The function f is well defined as R_i and R_i^o are opposite $(d-2)$-cubes in the $(d-1)$-cube J_i.

To see that f is an injection, take distinct vertices $v_1, v_2 \in V(F) \setminus \{s^o\}$, where $v_1 \in V(R_i) \setminus (V(R_1) \cup \cdots \cup V(R_{i-1}))$ and $v_2 \in V(R_j) \setminus (V(R_1) \cup \cdots \cup V(R_{j-1}))$ for $i \leq j$. If $i = j$ then $f(v_1) = \pi_{R_i^o}(v_1) \neq \pi_{R_i^o}(v_2) = f(v_2)$. If instead $i < j$ then $f(v_1) \in V(R_i^o) \subseteq V(R_1^o) \cup \cdots \cup V(R_{j-1}^o)$, while $f(v_2) \notin V(R_1^o) \cup \cdots \cup V(R_{j-1}^o)$. □

We are now ready to prove our main result.

Theorem 11.3.5 (Linkedness of cubical polytopes)**.** *For every $d \neq 3$, a cubical d-polytope is $\lfloor (d+1)/2 \rfloor$-linked.*

Proof. Proposition 11.3.2 settled the case of even d, so we assume d is odd. In this section, we give a complete proof for $d \geq 7$ and proofs of several cases when $d = 5$. The case of $d = 5$ will be finalised in Section 11.3.1.

Let d be odd and $d \geq 5$. Let $k := (d+1)/2$. Let X be any set of $2k$ vertices in the graph G of a cubical d-polytope P. Recall the vertices in X are called terminals. Also let $Y := \{\{s_1, t_1\}, \ldots, \{s_k, t_k\}\}$ be an arbitrary labelling and pairing of the vertices of X. We aim to find a Y-linkage $\{L_1, \ldots, L_k\}$ in G where L_i joins the pair $\{s_i, t_i\}$ for $i = 1, \ldots, k$. Recall we say that a path is X-valid if it contains no inner vertex from X.

Further recall that, given a pair $\{F, F^o\}$ of opposite facets in a cube Q, for every vertex $z \in V(F)$ we denote by $z^p_{F^o}$ or $\pi^Q_{F^o}(z)$ the unique neighbour of z in F^o.

The initial idea of the proof is to reduce the analysis space from the whole polytope to a more manageable space, the star $\mathcal{S}_1$ of a terminal vertex in the boundary complex of P, say that of s_1. We do so by considering $2k-1$ disjoint paths $S_i := s_i - \mathcal{S}_1$ ($i \in [2,k]$) and $T_j := t_j - \mathcal{S}_1$ ($j \in [1,k]$) from the terminals into $\mathcal{S}_1$. Here we resort to the d-connectivity of G. In addition let $S_1 := s_1$. We then denote by $\bar{s}_i$ and $\bar{t}_j$ the intersection of the paths S_i and T_j with $\mathcal{S}_1$. Using the vertices s_1, $\bar{s}_i$ and $\bar{t}_i$ for $i \in [2,k]$, define sets $\bar{X}$ and $\bar{Y}$ in $\mathcal{S}_1$, counterparts to the sets X and Y of G. In an abuse of terminology, we also call terminals the vertices $\bar{s}_i$ and $\bar{t}_i$. In this way, the existence of a $\bar{Y}$-linkage $\{\bar{L}_1, \ldots, \bar{L}_k\}$ with $\bar{L}_i := \bar{s}_i - \bar{t}_i$ in $G(\mathcal{S}_1)$ implies the existence of a Y-linkage $\{L_1, \ldots, L_k\}$ in $G(P)$, since each path $\bar{L}_i$ ($i \in [1,k]$) can be extended with the paths S_i and T_i to obtain the corresponding path $L_i = s_i - t_i$.

The proof of Theorem 11.3.5 is long, so we outline the main ideas. We consider a facet F_1 of $\mathcal{S}_1$ containing $\bar{t}_1$ and having the largest possible number of terminals. In addition, we let $\mathcal{A}_1$ be the antistar of F_1 in $\mathcal{S}_1$ and let $\mathcal{L}_1$ be the link of s_1 in F_1. We decompose the proof into four cases based on the number of terminals in F_1.

Case 1. $|\bar{X} \cap V(F_1)| = d+1$ for $d \geq 7$.

Case 2. $|\bar{X} \cap V(F_1)| = d$ for $d \geq 5$.

Case 3. $3 \leq |\bar{X} \cap V(F_1)| \leq d-1$ for $d \geq 5$.

Case 4. $|\bar{X} \cap V(F_1)| = 2$ for $d \geq 5$.

Using the $(k-1)$-linkedness of F_1 (Theorem 11.1.8), we link as many pairs of terminals in F_1 as possible through disjoint $\bar{X}$-valid paths $\bar{L}_i := \bar{s}_i - \bar{t}_i$. Each path $\bar{L}_i$, which is fully contained in F_1, can be extended with the paths S_i and T_i to obtain paths $L_i := s_i - t_i$. For those terminals that cannot be linked in F_1, if possible we use the injection from $V(F_1)$ to $V(\mathcal{A}_1)$ granted by Lemma 11.3.4 to find a set $N_{\mathcal{A}_1}$ of pairwise distinct neighbours in $\mathcal{A}_1$ not in $\bar{X}$. Then, using the $(k-2)$-linkedness of $\mathcal{A}_1$ (Proposition 11.3.3), we link the corresponding pairs of terminals in $\mathcal{A}_1$ and vertices in $N_{\mathcal{A}_1}$ accordingly.

Each such path $\bar{L}_i := \bar{s}_i - \bar{t}_i$, now with a subpath in $\mathcal{A}_1$, can then be extended with the paths S_i and T_i to obtain paths $L_i := s_i - t_i$. This general scheme does not always work, as the vertex s_1^o opposite to s_1 in F_1 may not have an image in $\mathcal{A}_1$ under the aforementioned injection or the image of a vertex in F_1 under the injection may be a terminal. In those scenarios we resort to ad hoc methods, including linking corresponding pairs in the link of s_1 in F_1, which is $(k-1)$-linked by Proposition 11.1.9 and does not contain s_1 or s_1^o, or in ridges disjoint from F_1, which are $(k-1)$-linked by Theorem 11.1.8.

To aid the reader, each case is broken down into subcases highlighted in bold. With the exception of a portion of Case 1, all the cases work also for $d = 5$. The difficulty with $d = 5$ in Case 1 stems from the $(d-2)$-faces of P not being 2-linked (Corollary 9.0.6).

Case 1. $|\bar{X} \cap V(F_1)| = d + 1$ *for* $d \geq 7$.

Here we have that $V(\mathcal{A}_1) \cap \bar{X} = \emptyset$. This case is decomposed into three main subcases A, B and C, based on the nature of the vertex s_1^o opposite to s_1 in F_1, which is the only vertex in F_1 that does not have an image under the injection from F_1 to $\mathcal{A}_1$ defined in Lemma 11.3.4. We find it convenient to label the subcases with letters as we will return to these subcases in Section 11.3.1, where we deal with $d = 5$.

Subcase A. The vertex s_1^o opposite to s_1 in F_1 does not belong to $\bar{X}$. Let $\bar{X}' := \bar{X} \setminus \{\bar{t}_1\}$ and let $\bar{Y}' := \bar{Y} \setminus \{\{s_1, \bar{t}_1\}\}$. Since $|\bar{X}'| = d$, the strong $(k-1)$-linkedness of F_1 (Theorem 11.1.11) gives a $\bar{Y}'$-linkage $\{\bar{L}_2, \ldots, \bar{L}_k\}$ in the facet F_1 with each path $\bar{L}_i := \bar{s}_i - \bar{t}_i$ ($i \in [2, k]$) avoiding s_1. With the help of Lemma 11.3.4, we find pairwise distinct neighbours s_1' and $\bar{t}_1'$ in $\mathcal{A}_1$ of s_1 and $\bar{t}_1$, respectively. If none of the paths $\bar{L}_i$ touches $\bar{t}_1$, we find a path $\bar{L}_1 := s_1 - \bar{t}_1$ in $\mathcal{S}_1$ that contains a subpath in $\mathcal{A}_1$ between s_1' and $\bar{t}_1'$ (here use the connectivity of $\mathcal{A}_1$, Proposition 10.2.2), and we are home. Otherwise, assume that the path $\bar{L}_j$ contains $\bar{t}_1$. With the help of Lemma 11.3.4, find pairwise distinct neighbours $\bar{s}_j'$ and $\bar{t}_j'$ in $\mathcal{A}_1$ of $\bar{s}_j$ and $\bar{t}_j$, respectively, such that the vertices s_1', $\bar{t}_1'$, $\bar{s}_j'$ and $\bar{t}_j'$ are pairwise distinct. According to Proposition 11.3.3, the complex $\mathcal{A}_1$ is 2-linked for $d \geq 7$. Hence, we can find disjoint paths $\bar{L}_1' := s_1' - \bar{t}_1'$ and $\bar{L}_j' := \bar{s}_j' - \bar{t}_j'$ in $\mathcal{A}_1$, respectively; these paths naturally give rise to paths $\bar{L}_1 := s_1 - \bar{t}_1$ in $\mathcal{S}_1$ with $\bar{L}_1' \subset \bar{L}_1$ and $\bar{L}_j := \bar{s}_j - \bar{t}_j$ in $\mathcal{S}_1$ with $\bar{L}_j' \subset \bar{L}_j$. The desired Y-linkage is given by the following paths.

$$L_i := \begin{cases} s_1 \bar{L}_1 \bar{t}_1 T_1 t_1, & \text{for } i = 1; \\ s_i S_i \bar{s}_i \bar{L}_i \bar{t}_i T_i t_i, & \text{otherwise.} \end{cases}$$

Subcase B. The vertex s_1^o opposite to s_1 in F_1 belongs to $\bar{X}$ but is different from $\bar{t}_1$, say $s_1^o = \bar{s}_2$. First find a neighbour s_1' of s_1 and a neighbour $\bar{t}_1'$ of $\bar{t}_1$ in $\mathcal{A}_1$ using Lemma 11.3.4. There is a neighbour $\bar{s}_2^{F_1}$ of $\bar{s}_2$ in F_1 that is either $\bar{t}_2$ or a vertex not in $\bar{X}$: $\{s_1, \bar{s}_2\} \cap N_{F_1}(\bar{s}_2) = \emptyset$ and $|N_{F_1}(\bar{s}_2)| = d - 1$. Note that the link $\mathcal{L}_1$ of s_1 in F_1 contains all the vertices in F_1 except s_1 and $\bar{s}_2$.

Suppose $\bar{s}_2^{F_1} = \bar{t}_2$. Let $\bar{L}_2 := \bar{s}_2 \bar{t}_2$, and using the $(k-1)$-linkedness of $\mathcal{L}_1$ (see Proposition 11.1.9), find disjoint paths $\bar{t}_1 - \bar{t}_2$ and $\bar{L}_i := \bar{s}_i - \bar{t}_i$ for $i \in [3, k]$ in $\mathcal{L}_1$. Then define a path $\bar{L}_1 := s_1 - \bar{t}_1$ in $\mathcal{S}_1$ that contains a subpath in $\mathcal{A}_1$ between s_1' and $\bar{t}_1'$; here we use the connectivity of $\mathcal{A}_1$ (see Proposition 10.2.2). Combining the paths $\bar{L}_i$ with the paths S_i and T_i for $i \in [1, k]$ gives the desired Y-linkage.

Assume $\bar{s}_2^{F_1}$ is not in $\bar{X}$. Observe that $|\bar{X} \setminus \{s_1, \bar{s}_2\} \cup \{\bar{s}_2^{F_1}\}| = d$. Using the $(k-1)$-linkedness of $\mathcal{L}_1$ for $d \geq 7$ (Proposition 11.1.9), find in $\mathcal{L}_1$ disjoint paths $\bar{L}_2' := \bar{s}_2^{F_1} - \bar{t}_2$ and $\bar{L}_i' := \bar{s}_i - \bar{t}_i$ for $i \in [3, k]$. Since $\bar{t}_1$ is also in $\mathcal{L}_1$ it may happen that it lies in one of the paths $\bar{L}_i'$. If $\bar{t}_1$ does not belong to any of the paths $\bar{L}_i'$ for $i \in [2, k]$, then find a path $\bar{L}_1 := s_1 - \bar{t}_1$ in $\mathcal{S}_1$ that contains a subpath in $\mathcal{A}_1$ between s_1' and $\bar{t}_1'$ using the connectivity of $\mathcal{A}_1$ (see Proposition 10.2.2). The desired Y-linkage is then given by

$$L_i := \begin{cases} s_1 \bar{L}_1 \bar{t}_1 T_1 t_1, & \text{for } i = 1; \\ \bar{s}_2 \bar{s}_2^{F_1} \bar{L}_2' \bar{t}_2 T_2 t_2, & \text{for } i = 2; \\ s_i S_i \bar{s}_i \bar{L}_i' \bar{t}_i T_i t_i, & \text{otherwise.} \end{cases}$$

If $\bar{t}_1$ belongs to one of the paths $\bar{L}'_i$ with $i \in [2,k]$, say $\bar{L}'_j$, then consider in $\mathcal{A}_1$ a neighbour $\bar{t}'_j$ of $\bar{t}_j$ and, either a neighbour $\bar{s}'_j$ of $\bar{s}_j$ if $j \neq 2$ or a neighbour $\bar{s}'_2$ of $\bar{s}_2^{F_1}$. From Lemma 11.3.4 it follows that the vertices s'_1, $\bar{t}'_1$, $\bar{s}'_j$ and $\bar{t}'_j$ can be taken pairwise distinct. Since $\mathcal{A}_1$ is 2-linked for $d \geq 7$ (see Proposition 11.3.3), find in $\mathcal{A}_1$ a path $\bar{L}'_1$ between s'_1 and $\bar{t}'_1$ and a path $\bar{L}''_j$ between $\bar{s}'_j$ and $\bar{t}'_j$. As a consequence, we obtain in $\mathcal{S}_1$ a path $\bar{L}_1 := s_1 s'_1 \bar{L}'_1 \bar{t}'_1 \bar{t}_1$ and, either a path $\bar{L}_j := \bar{s}_j \bar{s}'_j \bar{L}''_j \bar{t}'_j \bar{t}_j$ if $j \neq 2$ or a path $\bar{L}_2 := \bar{s}_2 \bar{s}_2^{F_1} \bar{s}'_2 \bar{L}''_2 \bar{t}'_2 \bar{t}_2$. The desired Y-linkage is then as follows.

$$L_i := \begin{cases} s_1 \bar{L}_1 \bar{t}_1 T_1 t_1, & \text{for } i = 1; \\ s_i S_i \bar{s}_i \bar{L}_i \bar{t}_i T_i t_i, & \text{otherwise.} \end{cases}$$

Subcase C. The vertex opposite to s_1 in F_1 coincides with $\bar{t}_1$. Then $\bar{t}_1$ has no neighbour in $\mathcal{A}_1$. In fact, F_1 is the only facet in $\mathcal{S}_1$ containing $\bar{t}_1$.

If $\bar{t}_1$ has a neighbour $\bar{t}_1^{F_1}$ in F_1 that is not in $\bar{X}$, then reason as in the scenario in which $\bar{s}_2 = s_1^o$ and $\bar{s}_2$ has a neighbour not in $\bar{X}$. Indeed, using the $(k-1)$-linkedness of $\mathcal{L}_1$ (see Proposition 11.1.9), first find disjoint paths $\bar{L}_i := \bar{s}_i - \bar{t}_i$ in $\mathcal{L}_1$ for $i \in [2,k]$. Second, using Lemma 11.3.4, consider neighbours s'_1 and $\bar{t}'_1$ in $\mathcal{A}_1$ of s_1 and $\bar{t}_1^{F_1}$, respectively. If $\bar{t}_1^{F_1}$ doesn't belong to any path $\bar{L}_i$, then find a path $\bar{L}_1 := s_1 - \bar{t}_1$ that contains the edge $\bar{t}_1 \bar{t}_1^{F_1}$ and a subpath in $\mathcal{A}_1$ between s'_1 and $\bar{t}'_1$. The desired Y-linkage is then given by

$$L_i := \begin{cases} s_1 \bar{L}_1 \bar{t}_1 T_1 t_1, & \text{for } i = 1; \\ s_i S_i \bar{s}_i \bar{L}_i \bar{t}_i T_i t_i, & \text{otherwise.} \end{cases}$$

If $\bar{t}_1^{F_1}$ belongs to one of the paths $\bar{L}_i$ with $i \in [2,k]$, say $\bar{L}_j$, then consider in $\mathcal{A}_1$ a neighbour $\bar{s}'_j$ of $\bar{s}_j$ and a neighbour $\bar{t}'_j$ of $\bar{t}_j$. From Lemma 11.3.4, it follows that the vertices s'_1, $\bar{t}'_1$, $\bar{s}'_j$ and $\bar{t}'_j$ can be taken pairwise distinct. Find a path $\bar{L}_1 := s_1 - \bar{t}_1$ in $\mathcal{S}_1$ that contains the edge $\bar{t}_1 \bar{t}_1^{F_1}$ and a subpath $\bar{L}'_1$ in $\mathcal{A}_1$ between s'_1 and $\bar{t}'_1$, and a path $\bar{L}_j := \bar{s}_j - \bar{t}_j$ in $\mathcal{S}_1$ that contains a subpath $\bar{L}'_j$ in $\mathcal{A}_1$ between $\bar{s}'_j$ and $\bar{t}'_j$; the 2-linkedness of $\mathcal{A}_1$ for $d \geq 7$ ensures that the paths $\bar{L}'_1$ and $\bar{L}'_j$ can be taken disjoint, and so can the paths $\bar{L}_1$ and $\bar{L}_j$. The desired Y-linkage is as before:

$$L_i := \begin{cases} s_1 \bar{L}_1 \bar{t}_1 T_1 t_1, & \text{for } i = 1; \\ s_i S_i \bar{s}_i \bar{L}_i \bar{t}_i T_i t_i, & \text{otherwise.} \end{cases}$$

Finally, assume that all the $d-1$ neighbours of $\bar{t}_1$ in F_1, and thus in $\mathcal{S}_1$, belong to $\bar{X}$. We find it convenient to create a subcase for this configuration since the strategy we employ to deal with it also applies to the case of $d = 5$ (up to this point we have required $d \geq 7$).

> Let $d \geq 5$ and $|\bar{X} \cap V(F_1)| = d+1$. Assume $\operatorname{dist}_{F_1}(s_1, \bar{t}_1) = d-1$ and all the $d-1$ neighbours of $\bar{t}_1$ in F_1, and thus in $\mathcal{S}_1$, belong to $\bar{X}$. (C.2)

This is a complex configuration since every neighbour of $\bar{t}_1$ in F_1 belongs to $\bar{X}$ and s_1^o $(= \bar{t}_1)$ has no neighbour in $\mathcal{A}_1$. So let us appropriately call it *Configuration C.2*. As a consequence, whenever we can, we reduce Configuration C.2 to one of the previous configurations of Case 1 or to Cases 2-4. Let R be a $(d-2)$-face of F_1 containing $s_1^o = \bar{t}_1$, then $s_1 \notin R$. Denote by R_{F_1} the $(d-2)$-face of F_1 disjoint from R. Let J be the other facet of P containing R and let R_J denote the $(d-2)$-face of J disjoint from R. Then R_J is disjoint from F_1.

Consider again the paths S_i and T_j that bring the vertices s_i $(i \in [2,k])$ and t_j $(j \in [1,k])$ into $\mathcal{S}_1$. Also recall that the paths S_i and T_j intersect $\mathcal{S}_1$ at $\bar{s}_i$ and $\bar{t}_j$, respectively. Suppose at least one path S_i or T_j touches R_J. Observe that, since every $\bar{s}_i$ $(i \in [2,k])$ and $\bar{t}_j$ $(j \in [1,k])$ is in F_1, the intersection of R_J with a path S_i or T_j is not in $\mathcal{S}_1$. Pick one such path, say S_ℓ. Denote the intersection of S_ℓ and R_J

by $\bar{s}'_\ell$. If there are several paths S_ℓ, choose one with the distance in R between $\bar{s}'_\ell$ and $\pi^J_R(\bar{t}_1)$ as large as possible. From the definition of $\bar{s}_\ell$, it follows that $\bar{s}'_\ell$ is not in $\mathcal{S}_1$.

We first find an $\bar{X}$-valid path M_ℓ on J from $\bar{s}'_\ell$ to $\mathcal{S}_1$ that is disjoint from the other paths S_i and T_j; we do so in order to replace the subpath $\bar{s}'_\ell S_\ell \bar{s}_\ell$ of S_ℓ with M_ℓ, thereby obtaining a new terminal $\{\bar{s}_\ell\} := V(M_\ell) \cap V(\mathcal{S}_1)$ that replaces the old $\bar{s}_\ell$. Partition the graph $G(R_J)$ into two subgraphs G_{bad} and G_{good} such that G_{bad} contains the neighbours of the terminals in R, namely $V(G_{\text{bad}}) = \pi^J_{R_J}(\bar{X} \cap V(R))$ and $V(G_{\text{good}}) = V(R_J) \setminus V(G_{\text{bad}})$. See Fig. 11.2(a). If $\bar{s}'_\ell \in G_{\text{good}}$, then M_ℓ consists of $\bar{s}'_\ell \pi^J_R(\bar{s}'_\ell)$. Otherwise, $\text{dist}_R(\bar{s}'_\ell, \pi^J_R(\bar{t}_1)) \le 1$, and the path M_ℓ traverses R_J from $\bar{s}'_\ell \in G_{\text{bad}}$ towards G_{good}, reaching a vertex $\bar{s}''_\ell$ that is either in $V(G_{\text{good}}) \setminus V(\mathcal{S}_1)$, in which case its projection $\pi^J_R(\bar{s}''_\ell)$ onto R is not in $\bar{X}$ and $M_\ell := s'_\ell \bar{s}''_\ell \pi^J_R(\bar{s}''_\ell)$, or in $V(\mathcal{S}_1)$, in which case $M_\ell := s'_\ell \bar{s}''_\ell$. Observe that the path M_ℓ cannot meet another path S_i or T_j in $G_{\text{bad}} \cup G_{\text{good}}$, say S_m, because in that case $\text{dist}_R(\bar{s}'_m, \pi^J_R(\bar{t}_1)) > \text{dist}_R(\bar{s}'_\ell, \pi^J_R(\bar{t}_1))$ and we should have chosen S_m instead of S_ℓ. Let $\bar{s}_\ell$ be the intersection of M_ℓ with $\mathcal{S}_1$; this is the new $\bar{s}_\ell$, and so the previous $\bar{s}_\ell$ is disregarded as a terminal. Then this new $\bar{s}_\ell$ is either in $\mathcal{A}_1$ or in R but at distance at least two from the vertex s_1^o opposite to s_1 in F_1.

If S_ℓ coincides with T_1, then there is a reduction either to a different configuration of Case 1 or to Cases 2-4, depending on whether the new vertex $\bar{t}_1$ is in F_1 or in $\mathcal{A}_1$. If the new $\bar{t}_1$ is in $\mathcal{A}_1$, we need to choose a new facet F_1 in $\mathcal{S}_1$, one containing the new $\bar{t}_1$ and a largest number of terminals. This new facet F_1 cannot contain the $d-1$ neighbours of the old $\bar{t}_1$ in the old F_1, thereby implying that the number of terminals in this new F_1 would be at most d and giving a reduction to Cases 2-4. Indeed, the intersection between the new and the old F_1 is at most $(d-2)$-dimensional and no $(d-2)$-dimensional face of the old F_1 contains all the $d-1$ neighbours of the old $\bar{t}_1$. Refer to Fig. 11.2(a) for a depiction of this case. If the aforementioned path S_ℓ is different from T_1 then there is a reduction to Case 1 or Case 2, depending on whether the new vertex $\bar{s}_\ell$ is in F_1 or in $\mathcal{A}_1$.

Consequently, we can assume that, for any ridge R of F_1 that contains $\bar{t}_1$, the aforementioned ridge R_J in the facet J is disjoint from all the paths S_i and T_j.

Consider the vertex $\bar{t}_1$ in F_1, an aforementioned ridge R, and the corresponding facet J and ridge R_J. There is a unique neighbour of $\bar{t}_1$ in R_{F_1}, say $\bar{s}_k$, while every other neighbour of $\bar{t}_1$ in F_1 is in R. There is a path $\bar{L}'_k := \bar{s}_k - \pi^{F_1}_{R_{F_1}}(\bar{t}_k)$ in R_{F_1} that avoids s_1, since R_{F_1} is $(d-2)$-connected. Let $\bar{X}^p := \pi^J_{R_J}(\bar{X} \setminus \{s_1, \bar{s}_k, \bar{t}_k\})$. See Fig. 11.2(b). Let $s_1^{pp} := \pi^J_{R_J}(\pi^{F_1}_R(s_1))$. The $d-1$ vertices in $\bar{X}^p \cup \{s_1^{pp}\}$ can be linked in R_J (see Theorem 11.1.8) by a linkage $\{\bar{L}'_1, \ldots, \bar{L}'_{k-1}\}$. Observe that, for the special case of $d=5$ where R_J is a 3-cube, we do not get the configuration preventing the 2-linkedness of R_J (Configuration F; see Definition 9.0.7) since $\text{dist}_{R_J}(s_1^{pp}, \pi^J_{R_J}(\bar{t}_1)) = 3$. The linkage $\{\bar{L}'_1, \ldots, \bar{L}'_{k-1}\}$ together with the two-path $\bar{L}_k := \bar{s}_k \pi^{F_1}_{R_{F_1}}(\bar{t}_k)\bar{t}_k$ can be extended to a linkage $\{\bar{L}_1, \ldots, \bar{L}_k\}$ given by

$$\bar{L}_i := \begin{cases} s_1 \pi^{F_1}_R(s_1) s_1^{pp} \bar{L}'_1 \pi^J_{R_J}(\bar{t}_1)\bar{t}_1, & \text{for } i = 1; \\ \bar{s}_i \pi^J_{R_J}(\bar{s}_i) \bar{L}'_i \pi^J_{R_J}(\bar{t}_i)\bar{t}_i, & \text{for } i \in [2, k-1]; \\ \bar{s}_k \pi^{F_1}_{R_{F_1}}(\bar{t}_k)\bar{t}_k, & \text{for } i = k. \end{cases}$$

Concatenating the paths S_i $(i \in [2,k])$ and T_j $(j \in [1,k])$ with the linkage $\{\bar{L}_1, \ldots, \bar{L}_k\}$ gives the desired Y-linkage.

Case 2. $|\bar{X} \cap V(F_1)| = d$ *for* $d \ge 5$.

Without loss of generality, assume that $\bar{t}_2 \notin V(F_1)$.

Suppose first that $\text{dist}_{F_1}(\bar{s}_2, s_1) < d-1$. According to Lemma 11.3.4, there exists a neighbour $\bar{s}'_2$ of $\bar{s}_2$ in $\mathcal{A}_1$. With the use of the strong $(k-1)$-linkedness of F_1 (Theorem 11.1.11), find disjoint paths $\bar{L}_1 := s_1 - \bar{t}_1$ and $\bar{L}_i := \bar{s}_i - \bar{t}_i$ $(i \in [3,k])$ in F_1, each avoiding $\bar{s}_2$. Find a path $\bar{L}_2$ in $\mathcal{S}_1$ between $\bar{t}_2$ and $\bar{s}_2$ that consists of the edge $\bar{s}_2\bar{s}'_2$ and a subpath in $\mathcal{A}_1$ between $\bar{t}_2$ and $\bar{s}'_2$, using the connectivity of

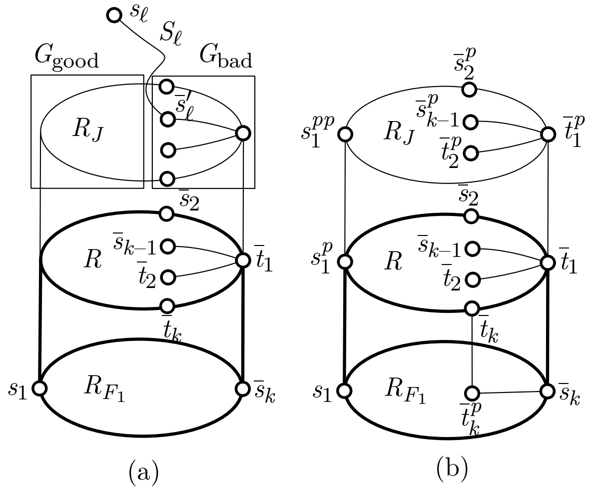


Figure 11.2: Auxiliary figure for Case 1 of Theorem 11.3.5, where the facet F_1 is highlighted in bold. (a) A configuration where a path S_i or T_j touches R_J. (b) A configuration where no path S_i or T_j touches R_J.

$\mathcal{A}_1$ (see Proposition 10.2.2). Combining the paths $\bar{L}_i$ ($i \in [1,k]$) with the paths S_i ($i \in [2,k]$) and T_j ($j \in [1,k]$) gives the desired Y-linkage.

Now assume $\operatorname{dist}_{F_1}(\bar{s}_2, s_1) = d-1$. Since $2k-1 = d$ and there are $d-1$ pairs of opposite $(d-2)$-faces in F_1, by Lemma 11.1.2 there exists a pair $\{R, R^o\}$ of opposite ridges of F_1 that is not associated with the set $\bar{X}_{s_2} := (\bar{X} \cap V(F_1)) \setminus \{\bar{s}_2\}$, whose cardinality is $d-1$. Assume $\bar{s}_2 \in R$.

Suppose all the neighbours of $\bar{s}_2$ in R are in $\bar{X}$; that is, $N_R(\bar{s}_2) = \bar{X} \setminus \{s_1, \bar{s}_2, \bar{t}_2\}$. The projection $\pi^{F_1}_{R^o}(\bar{s}_2)$ of $\bar{s}_2$ onto R^o is not in $\bar{X}$ since s_1 is the only terminal in R^o and $\operatorname{dist}_{F_1}(\bar{s}_2, s_1) = d-1 \geq 2$. Next find disjoint paths $\bar{L}_i := \bar{s}_i - \bar{t}_i$ for $i \in [3,k]$ in R that do not touch $\bar{s}_2$ or $\bar{t}_1$, using the $(d-1)/2$-linkedness of R if $d \geq 7$ or the 3-connectivity of R if $d = 5$. With the help of Lemma 11.3.4, find a neighbour u of $\pi^{F_1}_{R^o}(\bar{s}_2)$ in $\mathcal{A}_1$, and with the connectivity of $\mathcal{A}_1$, a path $\bar{L}_2$ between $\bar{s}_2$ and $\bar{t}_2$ that consists of the length-two path $\bar{s}_2\pi^{F_1}_{R^o}(\bar{s}_2)u$ and a subpath in $\mathcal{A}_1$ between u and $\bar{t}_2$. Finally, find a path $\bar{L}_1$ in F_1 between s_1 and $\bar{t}_1$ that consists of the edge $\bar{t}_1\pi^{F_1}_{R^o}(\bar{t}_1)$ and a subpath in R^o disjoint from $\pi^{F_1}_{R^o}(\bar{s}_2)$ (here use the 2-connectivity of R^o). Combining the paths $\bar{L}_i$ ($i \in [1,k]$) with the paths S_i ($i \in [2,k]$) and T_j ($j \in [1,k]$) we obtain the desired Y-linkage.

Thus assume there exists a neighbour $\bar{s}'_2$ of $\bar{s}_2$ in $V(R) \setminus \bar{X}$. Let $\bar{X}_{R^o} := \pi^{F_1}_{R^o}(\bar{X} \setminus \{\bar{s}_2, \bar{t}_2\})$. Find a path $\bar{L}_2$ between $\bar{t}_2$ and $\bar{s}_2$ that consists of the edge $\bar{s}_2\bar{s}'_2$ and a subpath in $\mathcal{A}_1$ between $\bar{t}_2$ and a neighbour $\bar{s}''_2$ of $\bar{s}'_2$ in $\mathcal{A}_1$ (see Lemma 11.3.4). In the case of $d \geq 7$, find disjoint paths $\bar{L}_i := \pi^{F_1}_{R^o}(\bar{s}_i) - \pi^{F_1}_{R^o}(\bar{t}_i)$ ($i \in [1,k]$ and $i \neq 2$) in R^o linking the $d-1$ vertices in $\bar{X}_{R^o}$ using the $(k-1)$-linkedness of R^o; add the edge $\pi^{F_1}_{R^o}(\bar{t}_i)\bar{t}_i$ to $\bar{L}_i$ if $\bar{t}_i \in R$ or the edge $\pi^{F_1}_{R^o}(\bar{s}_i)\bar{s}_i$ to $\bar{L}_i$ if $\bar{s}_i \in R$. Extend the disjoint paths $\bar{L}_i$ ($i \in [1,k]$) to a Y-linkage.

The case of $d = 5$ requires a bit more care. If the four vertices in $\bar{X}_{R^o}$—namely, s_1, $\pi^{F_1}_{R^o}(s_3)$, $\pi^{F_1}_{R^o}(t_3)$ and $\pi^{F_1}_{R^o}(t_1)$—do not form a Configuration F (Definition 9.0.7), then the same reasoning as in the case of $d = 7$ applies. Thus assume otherwise; that is, they are in cyclic order $s_1\pi^{F_1}_{R^o}(s_3)\pi^{F_1}_{R^o}(t_1)\pi^{F_1}_{R^o}(t_3)$ in a 2-face of R^o. This in turn implies that $\pi^{F_1}_{R}(s_3) \notin \{\bar{s}_2, \bar{s}'_2\}$ and $\pi^{F_1}_{R}(t_3) \notin \{\bar{s}_2, \bar{s}'_2\}$, since $\operatorname{dist}_{F_1}(s_1, \bar{s}_2) = 4$.

Find a path $\bar{L}'_3$ in R between $\pi^{F_1}_{R}(s_3)$ and $\pi^{F_1}_{R}(t_3)$ such that $\bar{L}'_3$ is disjoint from both $\bar{s}_2$ and $\bar{s}'_2$ and disjoint from $\bar{t}_1$ if $\bar{t}_1 \in R$; here use Corollary 11.1.4, which ensures that the vertices $\bar{s}_2$, $\bar{s}'_2$ and $\bar{t}_1$, if they are all in R, cannot separate $\pi^{F_1}_{R}(s_3)$ from $\pi^{F_1}_{R}(t_3)$ in R, since a separator of size three in R must be an independent set. Extend the path $\bar{L}'_3$ in R to a path $\bar{L}_3 := \bar{s}_3 - \bar{t}_3$ in F_1, if necessary. Find a path $\bar{L}'_1 := s_1 - \pi^{F_1}_{R^o}(t_1)$ in R^o disjoint from $\pi^{F_1}_{R^o}(s_3)$ and $\pi^{F_1}_{R^o}(t_3)$, using the 3-connectivity of R^o, and extend

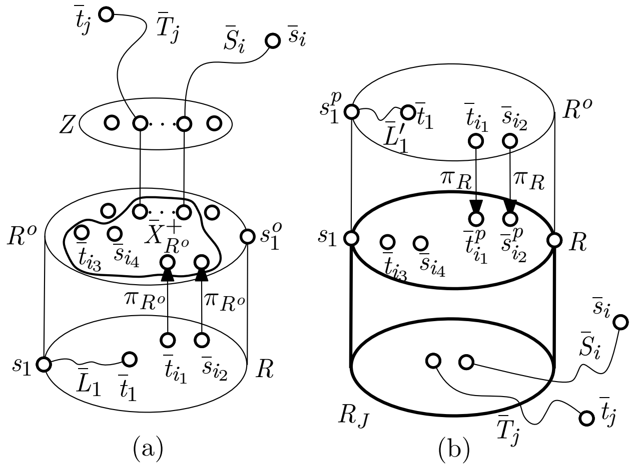


Figure 11.3: Auxiliary figure for Case 3 of Theorem 11.3.5. (a) A configuration where $\bar{t}_1 \in R$ and the subset $\bar{X}^+_{R^o}$ of R^o is highlighted in bold. (b) A configuration where $\bar{t}_1 \in R^o$ and the facet J is highlighted in bold.

it to a path $\bar{L}_1 := s_1 - \bar{t}_1$ in F_1, if necessary. Using the paths S_i ($i = 2, 3$) and T_j ($j = 1, 2, 3$) extend the linkage $\{\bar{L}_1, \bar{L}_2, \bar{L}_3\}$ to a Y-linkage. This completes the proof of Case 2.

Case 3. $3 \le |\bar{X} \cap V(F_1)| \le d - 1$ *for* $d \ge 5$.

Since $2k - 1 = d$ and there are $d - 1$ pairs of opposite facets in F_1, by Lemma 11.1.2 there exists a pair $\{R, R^o\}$ of opposite ridges of F_1 that is not associated with $\bar{X} \cap V(F_1)$. Assume $s_1 \in R$. We consider two subcases according to whether $\bar{t}_1 \in R$ or $\bar{t}_1 \in R^o$.

Suppose first that $\bar{t}_1 \in R$. The $(d-2)$-connectivity of R ensures the existence of an X-valid path $\bar{L}_1 := s_1 - \bar{t}_1$ in R. Let $\bar{X}_{R^o} := \pi^{F_1}_{R^o}(\bar{X} \setminus \{s_1, \bar{t}_1\})$. Then $1 \le |\bar{X}_{R^o}| \le d - 3$. Let s_1^o be the vertex opposite to s_1 in F_1; recall that s_1^o has no neighbour in $\mathcal{A}_1$. Let Z be a set of $|V(\mathcal{A}_1) \cap \bar{X}|$ pairwise distinct vertices in $\mathcal{A}_1$ adjacent to vertices in $V(R^o) \setminus (\bar{X}_{R^o} \cup \{s_1^o\})$; use Lemma 11.3.4 to obtain Z. Then $|Z| = |V(\mathcal{A}_1) \cap \bar{X}| \le d - 2$. To see that we can accommodate the set Z in $V(R^o) \setminus (\bar{X}_{R^o} \cup \{s_1^o\})$ observe that, for $d \ge 5$ and $|\bar{X}_{R^o}| \le d - 3$, we get

$$|V(R^o) \setminus (\bar{X}_{R^o} \cup \{s_1^o\})| \ge 2^{d-2} - (d-3) - 1 \ge d - 2.$$

Using the $(d-2)$-connectivity of $\mathcal{A}_1$ (Proposition 10.2.2) and Menger's theorem, find disjoint paths $\bar{S}_i$ and $\bar{T}_j$ ($i, j \ne 1$) between R^o and terminals $\bar{s}_i$ and $\bar{t}_j$ in $\mathcal{A}_1$, respectively, which pass through Z. Essentially, each path $\bar{S}_i = \bar{s}_i - R^o$ with $\bar{s}_i \in \mathcal{A}_1$ consists of a path between $\bar{s}_i$ and a vertex z_i in Z plus an edge $z_i\bar{z}_i$ with $\bar{z}_i \in V(R^o) \setminus (\bar{X}_{R^o} \cup \{s_1^o\})$; the same applies to each path $\bar{T}_j = \bar{t}_j - R^o$ with $\bar{t}_j \in \mathcal{A}_1$. If $\bar{s}_i$ or $\bar{t}_j$ is already in R^o, let $\bar{S}_i := \bar{s}_i$ or $\bar{T}_j := \bar{t}_j$, accordingly. If instead $\bar{s}_i$ or $\bar{t}_j$ is in R, let $\bar{S}_i$ be the edge $\bar{s}_i\pi^{F_1}_{R^o}(\bar{s}_i)$ or let $\bar{T}_j$ be the edge $\bar{t}_j\pi^{F_1}_{R^o}(\bar{t}_j)$. Let $\bar{X}^+_{R^o}$ be the intersections of R^o and the paths $\bar{S}_i$ and $\bar{T}_j$ ($i, j \ne 1$). Consequently, the paths $\bar{S}_i$ and $\bar{T}_i$ for $i \in [2, k]$ are pairwise disjoint. The set $\bar{X}^+_{R^o}$, with cardinality $d - 1$, can be linked through paths $\bar{L}_i$ ($i \in [2, k]$) by the $(k-1)$-linkedness of R^o (Theorem 11.1.8). See Fig. 11.3(a) for a depiction of this configuration. In this case, the desired Y-linkage is given by the following paths.

$$L_i := \begin{cases} s_1\bar{L}_1\bar{t}_1T_1t_1, & \text{for } i = 1; \\ s_iS_i\bar{s}_i\bar{S}_i\bar{L}_i\bar{T}_i\bar{t}_iT_it_i, & \text{otherwise.} \end{cases}$$

Some comments for $d = 5$ are in order. By virtue of Proposition 9.0.5 we need to make sure that the four vertices in $\bar{X}^+_{R^o}$ do not lie in the forbidden configuration described in Proposition 9.0.5, Configuration

F (Definition 9.0.7). To ensure this, we need to be a bit more careful when selecting the vertices in Z. Indeed, if there are already two vertices in $\bar{X}_{R^o}$ at distance three in R^o, no care is needed when selecting Z, so proceed as in the case of $d \geq 7$. Otherwise, pick a vertex $u \in V(R^o) \setminus (\bar{X}_{R^o} \cup \{s_1^o\})$ such that u is the unique vertex in R^o with $\text{dist}_{R^o}(u, x) = 3$ for a vertex $x \in \bar{X}_{R^o}$; this vertex x exists because $|\bar{X} \cap V(F_1)| \geq 3$. Then $u \notin \bar{X}_{R^o}$. Selecting such a $u \neq s_1^o$ is always possible because s_1^o is not at distance three in R^o from *any* vertex in $\bar{X}_{R^o}$: the unique vertex in R^o at distance three from s_1^o is $\pi_{R^o}^{F_1}(s_1)$, and $\pi_{R^o}^{F_1}(s_1) \notin \bar{X}$ because the pair $\{R, R^o\}$ is not associated with $\bar{X} \cap V(F_1)$. Once u is selected, let Z contain a neighbour z of u. In this way, some path $\bar{S}_i$ or $\bar{T}_j$ bringing terminals $\bar{s}_i$ or $\bar{t}_j$ in $\mathcal{A}_1$ into R^o through Z would use the vertex z, thereby ensuring that x and u would be both in $\bar{X}_{R^o}^+$, avoiding the forbidden configuration of Proposition 9.0.5.

Suppose now that $\bar{t}_1 \in R^o$. Let $\bar{X}_R := \pi_R^{F_1}((\bar{X} \setminus \{\bar{t}_1\}) \cap V(F_1))$. There are at most $d-2$ terminal vertices in R^o. Therefore, the $(d-2)$-connectivity of R^o ensures the existence of a $\bar{X}$-valid $\pi_{R^o}^{F_1}(s_1) - \bar{t}_1$ path $\bar{L}_1'$ in R^o. Then let $\bar{L}_1 := s_1 \pi_{R^o}^{F_1}(s_1) \bar{L}_1' \bar{t}_1$. Let J be the other facet in $\mathcal{S}_1$ containing R and let R_J be the $(d-2)$-face of J disjoint from R. Then $R_J \subset \mathcal{A}_1$. Since there are at most $d-2$ terminals in $\mathcal{A}_1$ and since $\mathcal{A}_1$ is $(d-2)$-connected (Proposition 10.2.2), we can find corresponding disjoint paths $\bar{S}_i$ and $\bar{T}_j$ bringing the terminals in $\mathcal{A}_1$ to R_J. For terminals $\bar{s}_i$ and $\bar{t}_j$ in $\bar{X} \cap V(R)$, let $\bar{S}_i := \bar{s}_i$ and $\bar{T}_j := \bar{t}_j$ for $i, j \neq 1$, while for terminals $\bar{s}_i$ and $\bar{t}_j$ in $\bar{X} \cap V(R^o)$, let $\bar{S}_i := \bar{s}_i \pi_R^{F_1}(\bar{s}_i)$ and $\bar{T}_j := \bar{t}_j \pi_R^{F_1}(\bar{t}_j)$ for $i, j \neq 1$. Let $\bar{X}_J$ be the set of the intersections of the paths $\bar{S}_i$ and $\bar{T}_j$ with J plus the vertex s_1. Then $\bar{X}_J \subset V(J)$ and $|\bar{X}_J| = d$ (since $\bar{t}_1 \in R^o$). Let $\{\hat{s}_i\} := V(\bar{S}_i) \cap V(J)$ and $\{\hat{t}_i\} := V(\bar{T}_i) \cap V(J)$ for $i \neq 1$. Resorting to the strong $(k-1)$-linkedness of the facet J (Theorem 11.1.11), we obtain $k-1$ disjoint paths $\bar{L}_i := \hat{s}_i - \hat{t}_i$ for $i \neq 1$ that correspondingly link $\bar{X}_J$ in J, with all the paths avoiding s_1. See Fig. 11.3(b) for a depiction of this configuration. In this case, the desired Y-linkage is given by the following paths.

$$L_i := \begin{cases} s_1 \bar{L}_1 \bar{t}_1 T_1 t_1, & \text{for } i = 1; \\ s_i S_i \bar{s}_i \bar{S}_i \bar{L}_i \bar{T}_i \bar{t}_i T_i t_i, & \text{otherwise.} \end{cases}$$

Case 4. $|\bar{X} \cap V(F_1)| = 2$ *for* $d \geq 5$.

In this case, we have that $|V(\mathcal{A}_1) \cap \bar{X}| = d-1$. Denote by $\mathcal{S}_{12}$ the star of $\bar{s}_2$ in $\mathcal{S}_1$; that is, the complex formed by the facets of P containing s_1 and $\bar{s}_2$. Denote by $G(\mathcal{S}_{12})$ the graph of $\mathcal{S}_{12}$ and by $\varGamma_{12}$ the subgraph of $G(\mathcal{S}_{12})$ and $G(\mathcal{A}_1)$ that is induced by $V(\mathcal{S}_{12}) \setminus V(F_1)$.

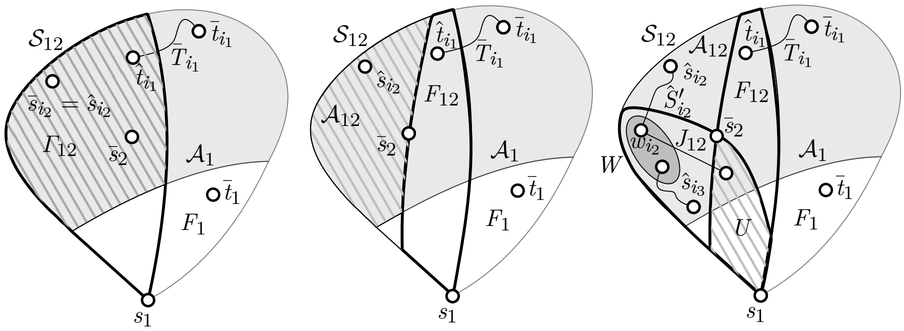
The initial strategy for this case is to bring the terminals in $\bar{X} \setminus V(F_1)$ into $\varGamma_{12}$. Excluding $\bar{s}_2$, there are exactly $d-2$ terminals in $V(\mathcal{A}_1)$. Denote by $\bar{S}_i$ an $\bar{X}$-valid path in $\mathcal{A}_1$ from the terminal $\bar{s}_i \in \mathcal{A}_1$ to $\varGamma_{12}$ that is disjoint from $\bar{X} \setminus \{\bar{s}_i\}$, and denote by $\hat{s}_i$ the intersection of $\bar{S}_i$ and $\varGamma_{12}$. Similarly, define $\bar{T}_j$ and $\hat{t}_j$. The existence of these $d-2$ pairwise disjoint $\bar{X}$-valid paths $\bar{S}_i$ and $\bar{T}_j$ is ensured by the $(d-2)$-connectivity of the graph $G(\mathcal{A}_1)$ of $\mathcal{A}_1$, which in turn is guaranteed by Proposition 10.2.2. It is important to note that each path $\bar{S}_i$ or $\bar{T}_j$ touches $\mathcal{S}_{12}$ at a vertex other than $\bar{s}_2$. Every terminal vertex $\bar{x}$ already in $\varGamma_{12}$ is also denoted by $\hat{x}$, and the corresponding path $\bar{S}_i$ or $\bar{T}_j$ consists only of the vertex $\hat{x}$. We also let $\hat{s}_2$ denote $\bar{s}_2$. The set of vertices $\hat{x}$ is accordingly denoted by $\hat{X}$. Then $|\hat{X}| = d-1$. Continuing with our abuse of terminology, since there is no potential for confusion, we call the vertices in $\hat{X}$ terminals as well. Figure 11.4(a) depicts this configuration.

Pick a facet F_{12} in $\mathcal{S}_{12}$ that contains $\hat{t}_2$. An important point is that $\bar{t}_1$ is not in F_{12}; otherwise F_{12} would contain s_1,$\hat{s}_2$ and $\bar{t}_1$ and it should have been chosen instead of F_1. The second part of the strategy is to bring the $d-1$ terminal vertices $\hat{x} \in \varGamma_{12}$ into the facet F_{12} so that they can be linked there.

Suppose $\mathcal{S}_{12}$ **only consists of** F_{12}. Then

$$\hat{X} = \{\hat{s}_2, \ldots, \hat{s}_k, \hat{t}_2, \ldots, \hat{t}_k\} \subset V(F_{12}).$$

As a consequence, with the help of the strong $(k-1)$-linkedness of F_{12} (Theorem 11.1.11), we can link the pairs $\{\hat{s}_i, \hat{t}_i\}$ for $i \in [2, k]$ in F_{12} through disjoint paths $\hat{L}_i$, all avoiding s_1. The paths $\hat{L}_i$ concatenated with the paths S_i, $\bar{S}_i$, T_i and $\bar{T}_i$ for $i \in [2, k]$ give a $(Y \setminus \{s_1, t_1\})$-linkage $\{L_2, \ldots, L_k\}$. To see the existence of a path $\bar{L}_1$ between s_1 and $\bar{t}_1$ that is disjoint from all the paths in $\{L_2, \ldots, L_k\}$, note that



(a) (b) (c)

Figure 11.4: Auxiliary figure for Case 4 of Theorem 11.3.5: A representation of $\mathcal{S}_1$ where the complex $\mathcal{S}_{12}$ is highlighted in bold, the facet F_1 is tiled in a falling pattern and the complex $\mathcal{A}_1$ is coloured in grey. (a) A configuration where the subgraph Γ_{12} is tiled in a falling pattern. (b) A depiction of $\mathcal{S}_{12}$ with more than one facet; the facet F_{12} is highlighted in bold and the complex $\mathcal{A}_{12}$ is highlighted in a falling pattern. (c) A depiction of $\mathcal{S}_{12}$ with more than one facet; the facets F_{12} and J_{12} are highlighted in bold and their intersection U is highlighted in a falling pattern. The set W in J_{12} is coloured in grey.

the intersection of F_{12} and F_1 is at most a $(d-2)$-face containing s_1 (but not t_1), which is contained in a $(d-2)$-face R of F_1 containing s_1 but not t_1 (Remark 11.2.2). Find a path $\bar{L}'_1$ in R^o, the ridge of F_1 disjoint from R and containing t_1, between $\pi_{R^o}^{F_1}(s_1)$ and $\bar{t}_1$ and let $\bar{L}_1 := s_1\pi_{R^o}^{F_1}(s_1)\bar{L}'_1\bar{t}_1$. Hence the desired Y-linkage is as follows.

$$L_i := \begin{cases} s_1\bar{L}_1\bar{t}_1T_1t_1, & \text{for } i=1;\\ s_iS_i\bar{s}_i\bar{S}_i\hat{s}_i\hat{L}_i\hat{t}_i\bar{T}_i\bar{t}_iT_it_i, & \text{otherwise.}\end{cases}$$

Assume $\mathcal{S}_{12}$ has more than one facet. Recall that

$$\hat{X} = \{\hat{s}_2, \ldots, \hat{s}_k, \hat{t}_2, \ldots, \hat{t}_k\} \subset V(\Gamma_{12}).$$

We denote by $\mathcal{A}_{12}$ the complex of $\mathcal{S}_{12}$ induced by $V(\mathcal{S}_{12}) \setminus (V(F_1) \cup V(F_{12}))$. Then the graph $G(\mathcal{A}_{12})$ of $\mathcal{A}_{12}$ coincides with the subgraph of Γ_{12} induced by $V(\Gamma_{12}) \setminus V(F_{12})$. Figure 11.4(b) depicts this configuration.

Our strategy is first to bring the $d-3$ terminal vertices $\hat{x}$ in Γ_{12} other than $\hat{s}_2$ and $\hat{t}_2$ into $F_{12} \setminus F_1$ through disjoint paths $\hat{S}_i$ and $\hat{T}_j$, without touching $\hat{s}_2$ and $\hat{t}_2$. Second, denoting by $\tilde{s}_i$ and $\tilde{t}_j$ the intersection of $\hat{S}_i$ and $\hat{T}_j$ with $V(F_{12}) \setminus V(F_1)$, respectively, we link the pairs $\{\tilde{s}_i, \tilde{t}_i\}$ for $i = [2, k]$ in F_{12} through disjoint paths $\tilde{L}_i$, without touching s_1; here we resort to the strong $(k-1)$-linkedness of F_{12}. And finally, we find a path $\bar{L}_1 := s_1 - \bar{t}_1$ in F_1, which is disjoint from the paths $\tilde{L}_i$ with $i \in [2, k]$. We develop these ideas below.

From Lemma 11.2.3(iii), it follows that $\mathcal{A}_{12}$ is nonempty and contains a spanning strongly connected $(d-3)$-subcomplex, thereby implying, by Proposition 9.0.1, that

$$G(\mathcal{A}_{12}) \text{ is } (d-3)\text{-connected.}$$

Since $\mathcal{S}_{12}$ contains more than one facet, there is a $(d-2)$-face U in F_{12} that contains s_1 and $\hat{s}_2$ $(= \bar{s}_2)$; see Remark 11.2.1. Denote by J_{12} the other facet in $\mathcal{S}_{12}$ containing U and by U_J the $(d-2)$-face in J_{12} disjoint from U, and as a consequence, disjoint from F_{12}.

Denote by $\mathcal{C}_U$ the subcomplex of U induced by $V(U) \setminus V(F_1)$, namely the antistar of $U \cap F_1$ in U. Since $\hat{s}_2 \in V(U) \setminus V(F_1)$, it follows that $\mathcal{C}_U$ is nonempty, and thus, thanks to Lemma 10.1.2, that it is a strongly connected $(d-3)$-complex. Then, from C_U containing a $(d-3)$-face it follows that

$$|V(\mathcal{C}_U)| = |V(U) \setminus V(F_1))| \geq 2^{d-3} \geq d-1 \text{ for } d \geq 5. \tag{11.1}$$

Moreover, denote by $\mathcal{C}_{U_J}$ the subcomplex of U_J induced by $V(U_J)\setminus V(F_1)$. If $U_J \cap F_1 = \emptyset$ then $\mathcal{C}_{U_J} = U_J$. Otherwise $\mathcal{C}_{U_J}$ is the antistar of $U_J \cap F_1$ in U_J, and since $U \cap F_1 \neq \emptyset$ (s_1 is in both), it follows that $U_J \not\subseteq F_1$, and thus, that $\mathcal{C}_{U_J}$ is nonempty. Put differently, the vertex in J_{12} opposite to s_1 is not in U, since $s_1 \in U$, nor is it in F_1, and so it must be in $\mathcal{C}_{U_J}$. Therefore, according to Lemma 10.1.2, $\mathcal{C}_{U_J}$ is a strongly connected $(d-3)$-complex. Hence, in both instances,

$$|V(\mathcal{C}_{U_J})| = |V(U_J) \setminus V(F_1))| \geq 2^{d-3} \geq d-1 \text{ for } d \geq 5. \tag{11.2}$$

Recall that we want to bring every vertex in the set $\hat{X}$, which is contained in Γ_{12}, into $F_{12} \setminus F_1$. We construct $|\hat{X} \cap V(\mathcal{A}_{12})|$ pairwise disjoint paths $\hat{S}_i$ and $\hat{T}_j$ from $\hat{s}_i \in \mathcal{A}_{12}$ and $\hat{t}_j \in \mathcal{A}_{12}$, respectively, to $V(F_{12}) \setminus V(F_1)$ as follows. Pick a set

$$W \subset V(\mathcal{C}_{U_J}) \setminus \pi_{U_J}^{J_{12}}\left((\hat{X} \cup \{s_1\}) \cap U\right)$$

of $|\hat{X} \cap V(\mathcal{A}_{12})|$ vertices in $\mathcal{C}_{U_J}$. In other words, the vertices in W are in $\mathcal{C}_{U_J}$ and are not projections of the vertices in $(\hat{X} \cup \{s_1\}) \cap U$ onto U_J. We show that the set W exists, which amounts to showing that $\mathcal{C}_{U_J}$ has enough vertices to accommodate W.

First note that

$$\begin{aligned} |\hat{X} \cap V(\mathcal{A}_{12})| + |(\hat{X} \cup \{s_1\}) \cap V(F_{12})| = |\hat{X} \cup \{s_1\}| = d, \\ (\hat{X} \cup \{s_1\}) \cap V(U) \subseteq (\hat{X} \cup \{s_1\}) \cap V(F_{12}). \end{aligned} \tag{11.3}$$

If $U_J \cap F_1 = \emptyset$ then $\mathcal{C}_{U_J} = U_J$, and Eq. (11.3), together with $|V(U_J)| = 2^{d-2} \geq d$ for $d \geq 5$, give the following chain of inequalities

$$\begin{aligned} \left|V(C_{U_J}) \setminus \pi_{U_J}^{J_{12}}\left((\hat{X} \cup \{s_1\}) \cap V(U)\right)\right| &\geq d - \left|(\hat{X} \cup \{s_1\}) \cap V(U)\right| \\ &\geq \left|\hat{X} \cup \{s_1\}\right| - \left|(\hat{X} \cup \{s_1\}) \cap V(F_{12})\right| = \left|\hat{X} \cap V(\mathcal{A}_{12})\right| = |W|, \end{aligned}$$

as desired.

Suppose now $U_J \cap F_1 \neq \emptyset$. Since $s_1 \in U \cap F_1$ and $J_{12} = \operatorname{conv}\{U \cup U_J\}$, the cube $J_{12} \cap F_1$ has opposite facets $U_J \cap F_1$ and $U \cap F_1$. From $s_1 \in U \cap F_1$ it follows that $\pi_{U_J}^{J_{12}}(s_1) \in U_J \cap F_1$, and thus, that $\pi_{U_J}^{J_{12}}(s_1) \notin \mathcal{C}_{U_J}$; here we use the following remark.

Remark 11.3.6. Let (K, K^o) be opposite facets in a cube Q and let B be a proper face of Q such that $B \cap K \neq \emptyset$ and $B \cap K^o \neq \emptyset$. Then $\pi_{K^o}^{Q}(B \cap K) = B \cap K^o$.

Since $\pi_{U_J}^{J_{12}}(s_1) \notin \mathcal{C}_{U_J}$, using Eqs. (11.2) and (11.3) we get

$$\begin{aligned} \left|V(C_{U_J}) \setminus \pi_{U_J}^{J_{12}}\left((\hat{X} \cup \{s_1\}) \cap V(U)\right)\right| &= \left|V(C_{U_J}) \setminus \pi_{U_J}^{J_{12}}\left(\hat{X} \cap V(U)\right)\right| \\ &\geq d - 1 - \left|\hat{X} \cap V(U)\right| \geq \left|\hat{X}\right| - \left|\hat{X} \cap V(F_{12})\right| = \left|\hat{X} \cap V(\mathcal{A}_{12})\right| = |W|. \end{aligned}$$

In this way we have shown that $\mathcal{C}_{U_J}$ can accommodate the set W.

There are at most $d-3$ vertices $\hat{x}$ in $\hat{X} \cap V(\mathcal{A}_{12})$ since $\hat{s}_2$ and $\hat{t}_2$ are already in $V(F_{12}) \setminus V(F_1)$. Since $G(\mathcal{A}_{12})$ is $(d-3)$-connected, we can find $|W| = |\hat{X} \cap V(\mathcal{A}_{12})|$ pairwise disjoint paths $\hat{S}_i'$ and $\hat{T}_j'$ in $\mathcal{A}_{12}$ from the terminals $\hat{s}_i$ and $\hat{t}_j$ in $\hat{X} \cap V(\mathcal{A}_{12})$ to W. The path $\hat{S}_i$ then consists of the subpath $\hat{S}_i' := \hat{s}_i - w_i$ with $w_i \in W$ plus the edge $w_i \pi_U^{J_{12}}(w_i)$; from the choice of W it follows that $\pi_U^{J_{12}}(w_i) \notin \hat{X} \cup \{s_1\}$. The paths $\hat{T}_j'$ and $\hat{T}_j$ are defined analogously. Figure 11.4(c) depicts this configuration.

Denote by $\tilde{s}_i$ the intersection of $\hat{S}_i$ and $V(F_{12}) \setminus V(F_1)$; similarly, define $\tilde{t}_j$. Every terminal vertex $\hat{x}$ already in F_{12} is also denoted by $\tilde{x}$, and in this case we let $\hat{S}_i$ or $\hat{T}_j$ be the vertex $\tilde{x}$.

Now F_{12} contains the pairs $(\tilde{s}_i, \tilde{t}_i)$ for $i \in [2, k]$ and the terminal s_1, as desired. Link these pairs in F_{12} through disjoint paths $\tilde{L}_i$, each avoiding s_1, with the use of the strong $(k-1)$-linkedness of F_{12} (see Theorem 11.1.11). The paths $\tilde{L}_i$ concatenated with the paths S_i, $\bar{S}_i$, $\hat{S}_i$, T_i, $\bar{T}_i$ and $\hat{T}_i$ for $i \in [2, k]$ give a $(Y \setminus \{s_1, t_1\})$-linkage $\{L_2, \ldots, L_k\}$.

We now show how to find the path $\bar{L}_1$ in F_1 between s_1 and $\bar{t}_1$. The paths $\bar{S}_i$ and $\bar{T}_i$ $(i \in [2, k])$ lie in $\mathcal{A}_1$, while the paths $\hat{S}_i$ and $\hat{T}_i$ $(i \in [2, k])$ lie in Γ_{12}, thus, in $\mathcal{A}_1$ as well. As a consequence, the only paths that could possibly touch F_1 are the paths $\tilde{L}_i$ $(i \in [2, k])$, and they all lie within F_{12}. As in the case of $\mathcal{S}_{12}$ having exactly one facet, the intersection of F_{12} and F_1 is at most a $(d-2)$-face containing s_1 but not $\bar{t}_1$, which is contained in a $(d-2)$-face R of F_1 containing s_1 but not t_1. Find a path $\bar{L}_1'$ in R^o between $\pi_{R^o}^{F_1}(s_1)$ and $\bar{t}_1$ and let $\bar{L}_1 := s_1 \pi_{R^o}^{F_1}(s_1) \bar{L}_1' \bar{t}_1$. Hence the desired Y-linkage is as follows.

$$L_i := \begin{cases} s_1 \bar{L}_1 \bar{t}_1 T_1 t_1, & \text{for } i = 1; \\ s_i S_i \bar{s}_i \bar{S}_i \hat{s}_i \hat{S}_i \tilde{s}_i \tilde{L}_i \tilde{t}_i \hat{T}_i \hat{t}_i \bar{T}_i \bar{t}_i T_i t_i, & \text{otherwise.} \end{cases}$$

This completes the proof of the case and of the theorem for $d \geq 7$. □

The missing case of $d = 5$ is settled in the next subsection.

11.3.1 Proof of Theorem 11.3.5 for $d = 5$

We keep the same notation as in the first part of the proof in Theorem 11.3.5. Let $X := \{s_1, s_2, s_3, t_1, t_2, t_3\}$ be any set of six vertices in the graph G of a cubical 5-polytope P. We call the vertices in X terminals. Also let $Y := \{\{s_1, t_1\}, \{s_2, t_2\}, \{s_3, t_3\}\}$. We aim to find a Y-linkage $\{L_1, L_2, L_3\}$ in G where L_i joins the pair $\{s_i, t_i\}$ for $i = 1, 2, 3$. We say that a path is X-valid if it contains no inner vertex from X.

As in the proof of Theorem 11.3.5, denote by $\mathcal{S}_1$ the star of s_1 in $\mathbb{B}(P)$. From the 5-connectivity of G follows the existence of five disjoint paths $S_i := s_i - \mathcal{S}_1$ $(i \in [2, 3])$ and $T_j := t_j - \mathcal{S}_1$ $(j \in [1, 3])$. Further denote by $\bar{s}_i$ and $\bar{t}_j$ the intersection of the paths S_i and T_j with $\mathcal{S}_1$. Using the vertices s_1, $\bar{s}_i$ and $\bar{t}_i$ for $i = 2, 3$, define sets $\bar{X}$ and $\bar{Y}$, counterparts to the sets X and Y of G. In an abuse of terminology, we call terminals the vertices $\bar{s}_i$ and $\bar{t}_i$. Let F_1 be a facet of $\mathcal{S}_1$ containing $\bar{t}_1$ and a largest number of terminals. Let $\mathcal{A}_1$ denote the antistar of F_1 in $\mathcal{S}_1$ and let $\mathcal{L}_1$ denote the link of s_1 in F_1.

From Proposition 9.0.5 it follows that, if a 3-cube contains exactly four terminals, then the only configuration in which the two pairs $\{s_{i_1}, t_{i_1}\}$ and $\{s_{i_2}, t_{i_2}\}$ cannot be linked in the 3-cube occurs when the terminals appear in cyclic order $s_{i_1} s_{i_2} t_{i_1} t_{i_2}$ in a 2-face of the 3-cube. We called this forbidden configuration Configuration F (Definition 9.0.7). If a 3-cube contains more than four terminals, then we also call *Configuration F* any configuration in which four terminals, say s_{i_1}, t_{i_1}, x_1 and x_2, lie in a 2-face and are arranged in cyclic order $s_{i_1} x_1 t_{i_1} x_2$; here $\{x_1, x_2\}$ may or may not belong to Y. The reason for calling this configuration forbidden is that the terminals $\{s_{i_1}, t_{i_1}\}$ cannot be linked via an X-valid path in the 3-cube.

Proposition 11.3.7. *A cubical 5-polytope is 3-linked.*

Proof. We consider three cases based on the number of terminals in F_1. The cases of $|\bar{X} \cap V(F_1)| = 5, 4, 3, 2$ were settled in Cases 2-4 of Theorem 11.3.5. Consequently, here we focus on the case of $|\bar{X} \cap V(F_1)| = 6$ and its subcases A, B and C. We find it advantageous to settle the subcases A and B at the same time.

Throughout the section we denote by s_1^o the vertex of F_1 opposite to s_1.

Subcases A and B. The vertex s_1^o opposite to s_1 in F_1 either does not belong to $\bar{X}$ or belongs to $\bar{X}$ but is different from $\bar{t}_1$.

There is a 3-face R of F_1 containing both s_1 and $\bar{t}_1$. Let J_1 be the other facet in $\mathcal{S}_1$ containing R. Denote by R_J and R_F the ridges in J_1 and F_1, respectively, that are disjoint from R. Then $s_1^o \in R_F$. We need the following claim.

Claim 10. *If a 3-cube R contains three pairs of terminals, there must exist two pairs of terminals in R that are not arranged in Configuration F; that is, two pairs, say $\{s_1, \bar{t}_1\}$ and $\{\bar{s}_2, \bar{t}_2\}$, are not arranged in cyclic order $s_1\bar{s}_2\bar{t}_1\bar{t}_2$ in a 2-face.*

Proof. If the cube does not contain any Configuration F, we are done. So suppose it does, namely $s_1x_1\bar{t}_1x_2$ with $x_1, x_2 \in \bar{X}$. Without loss of generality assume that $\bar{s}_2 \notin \{x_1, x_2\}$. Then $\bar{s}_2$ cannot be adjacent to both $\bar{s}_1$ and $\bar{t}_1$, since the bipartite graph $K_{2,3}$ is not a subgraph of $G(Q_3)$ (Remark 11.1.5). Thus the pairs $\{\bar{s}_1, \bar{t}_1\}$ and $\{\bar{s}_2, \bar{t}_2\}$ do not exemplify a Configuration F. □

Suppose all the six terminals are in the 3-face R. By virtue of Claim 10 we may assume $\{s_1, \bar{t}_1\}$ and $\{\bar{s}_2, \bar{t}_2\}$ are not arranged in Configuration F in R. Proposition 9.0.5 ensures that the pairs $\{\pi^{J_1}_{R_J}(s_1), \pi^{J_1}_{R_J}(\bar{t}_1)\}$ and $\{\pi^{J_1}_{R_J}(\bar{s}_2), \pi^{J_1}_{R_J}(\bar{t}_2)\}$ in R_J can be linked in R_J through disjoint paths $\bar{L}_1'$ and $\bar{L}_2'$, since they do not lie in Configuration F in R_J: if two pairs do not lie in Configuration F within a facet, then neither do their projections onto the opposite facet. Moreover, by the connectivity of R_F, there is a path $\bar{L}_3'$ in R_F linking the pair $\{\pi^{F_1}_{R_F}(\bar{s}_3), \pi^{F_1}_{R_F}(\bar{t}_3)\}$. The linkage $\{\bar{L}_1', \bar{L}_2', \bar{L}_3'\}$ can naturally be extended to a $\bar{Y}$-linkage, which combined with the paths S_i $(i = 2, 3)$ and T_j $(j = 1, 2, 3)$ can be extended to the required Y-linkage.

Suppose that R contains a pair $\{\bar{s}_i, \bar{t}_i\}$ for $i = 2, 3$, say $\{\bar{s}_2, \bar{t}_2\}$. There are at most five terminals in R, and consequently, applying Lemma 11.1.7 to the polytope F_1 and its facet R, we obtain an $\bar{X}$-valid path $\bar{L}_1 := s_1 - \bar{t}_1$ in R or an $\bar{X}$-valid path $\bar{L}_2 := \bar{s}_2 - \bar{t}_2$ in R. For the sake of concreteness, say an $\bar{X}$-valid path $\bar{L}_2$ exists in R. From the connectivity of R_F and R_J follows the existence of a path $\bar{L}_3'$ in R_F between $\pi^{F_1}_{R_F}(\bar{s}_3)$ and $\pi^{F_1}_{R_F}(\bar{t}_3)$, and of a path $\bar{L}_1'$ in R_J between $\pi^{J_1}_{R_J}(s_1)$ and $\pi^{J_1}_{R_J}(\bar{t}_1)$. The linkage $\{\bar{L}_1', \bar{L}_2', \bar{L}_3'\}$ can be naturally extended to a linkage $\{s_1 - \bar{t}_1, \bar{s}_2 - \bar{t}_2, \bar{s}_3 - \bar{t}_3\}$ in $\mathcal{S}_1$, which, in turn, combined with the paths S_i $(i = 2, 3)$ and T_j $(j = 1, 2, 3)$, gives rise to a Y-linkage, as desired.

Suppose that the ridge R contains no other pair from $\bar{Y}$ and that the ridge R_F contains a pair $(\bar{s}_i, \bar{t}_i)$ $(i = 2, 3)$. Without loss of generality assume $\bar{s}_2$ and $\bar{t}_2$ are in R_F.

First suppose that $s_3 \in R$, which implies that $\bar{t}_3 \in R_F$. Further suppose that there is a path M_3 of length at most two from $\bar{t}_3$ to R that is disjoint from $\bar{X} \setminus \{\bar{s}_3, \bar{t}_3\}$. Let $\{\bar{t}_3'\} := V(M_3) \cap V(R)$. Use the 2-linkedness of J_1 (Proposition 9.0.8) to find disjoint paths $\bar{L}_1 := s_1 - \bar{t}_1$ and $\bar{L}_3' := \bar{s}_3 - \bar{t}_3'$ in J_1. Let $\bar{L}_3 := \bar{s}_3\bar{L}_3'\bar{t}_3'M_3\bar{t}_3$. Use the 3-connectivity of R_F to find an $\bar{X}$-valid path $\bar{L}_2 := \bar{s}_2 - \bar{t}_2$ in R_F; note that $|V(M_3) \cap V(R_F)| \le 2$. The linkage $\{\bar{L}_1, \bar{L}_2, \bar{L}_3\}$ can be naturally extended to a Y-linkage. Now suppose there is no such path M_3 from $\bar{t}_3$ to R. Then, both $\bar{s}_2$ and $\bar{t}_2$ must be neighbours of $\bar{t}_3$ in R_F; the projection $\pi^{F_1}_{R}(\bar{t}_3)$ is in $\{s_1, \bar{t}_1\}$, say $\pi^{F_1}_{R}(\bar{t}_3) = t_1$; and the projection $\pi^{F_1}_{R_F}(s_1)$ is a neighbour of $\bar{t}_3$ in R_F. This configuration implies that s_1 and $\bar{t}_1$ are adjacent in R. Let $\bar{L}_1 := s_1\bar{t}_1$. Find a path $\bar{L}_2 := \bar{s}_2 - \bar{t}_2$ in R_F that is disjoint from $\pi^{F_1}_{R_F}(s_1)$ and t_3, using the 3-connectivity of R_F. Then, with the help of Lemma 11.3.4, find a neighbour $\bar{s}_3'$ in $\mathcal{A}_1$ of $\bar{s}_3$, and a neighbour $\bar{t}_3' \neq \bar{s}_3'$ in $\mathcal{A}_1$ of $\bar{t}_3$; note that, since $\operatorname{dist}_{F_1}(s_1, \bar{t}_3) \le 2$, $\bar{t}_3 \neq s_1^o$. Find a path $\bar{L}_3$ in $\mathcal{S}_1$ between $\bar{s}_3$ and $\bar{t}_3$ that contains a subpath in $\mathcal{A}_1$ between $\bar{s}_3'$ and $\bar{t}_3'$; here use the connectivity of $\mathcal{A}_1$ (Proposition 10.2.2). The linkage $\{\bar{L}_1, \bar{L}_2, \bar{L}_3\}$ can be naturally extended to a Y-linkage.

Assume that $\bar{s}_3 \in R_F$; by symmetry we can further assume that $\bar{t}_3 \in R_F$. The connectivity of R ensures the existence of a path $\bar{L}_1 := \bar{s}_1 - \bar{t}_1$ therein. In the case of $s_1^o \in \bar{X}$, without loss of generality, assume $s_1^o = \bar{s}_2$. The 3-connectivity of R_F ensures the existence of an $\bar{X}$-valid path $\bar{s}_2 - \bar{t}_2$ therein. Use Lemma 11.3.4 to find pairwise distinct neighbours $\bar{s}_3'$ of $\bar{s}_3$ and $\bar{t}_3'$ of $\bar{t}_3$ in $\mathcal{A}_1$; these exist since $\bar{s}_3 \neq s_1^o$ and $\bar{t}_3 \neq s_1^o$. Using the connectivity of $\mathcal{A}_1$ (Proposition 10.2.2), find a path $\bar{L}_3 := \bar{s}_3 - \bar{t}_3$ in $\mathcal{S}_1$ that contains a subpath $\bar{s}_3' - \bar{t}_3'$ in $\mathcal{A}_1$. Extend the linkage $\{\bar{L}_1, \bar{L}_2, \bar{L}_3\}$ to a Y-linkage.

Assume neither R nor R_F contains a pair $\{\bar{s}_i, \bar{t}_i\}$ $(i = 2, 3)$. Without loss of generality assume that $\bar{s}_2, \bar{s}_3 \in R$, that $\bar{t}_2, \bar{t}_3 \in R_F$ and that $\bar{t}_2 \neq s_1^o$.

First suppose that there exists a path M_3 in F_1 from $\bar{s}_3$ to R_F that is of length at most two and is disjoint from $\bar{X} \setminus \{\bar{s}_3, \bar{t}_3\}$. Find pairwise distinct neighbours $\bar{s}_2'$ and $\bar{t}_2'$ of $\bar{s}_2$ and $\bar{t}_2$, respectively, in $\mathcal{A}_1$, and a path $\bar{L}_2 := \bar{s}_2 - \bar{t}_2$ in $\mathcal{S}_1$ that contains a subpath $\bar{s}_2' - \bar{t}_2'$ in $\mathcal{A}_1$ (using the connectivity of $\mathcal{A}_1$). Let $\{\hat{s}_3\} := V(M_3) \cap V(R_F)$. Using the 3-connectivity of R_F link the pair $\{\hat{s}_3, \bar{t}_3\}$ in R_F through a path $\bar{L}_3$. Since Corollary 11.1.4 ensures that any separator of size three in a 3-cube must be independent, we can find an $\bar{L}_1 := s_1 - \bar{t}_1$ in R that is disjoint from s_2 and $V(M_3) \cap V(R)$; the set $V(M_3) \cap V(R)$ has either cardinality one or contains an edge. Extend the linkage $\{\bar{L}_1, \bar{L}_2, \bar{L}_3\}$ to a required Y-linkage.

Assume that there is no such path M_3. In this case, the neighbours of $\bar{s}_3$ in F_1 are $s_1, \bar{t}_1, \bar{s}_2$ from R and $\bar{t}_2$ from R_F. Use Lemma 11.3.4, to find a neighbour $\bar{s}_3'$ of $\bar{s}_3$ in $\mathcal{A}_1$ and a neighbour $\bar{t}_3'$ either of $\bar{t}_3$, if $\bar{t}_3 \neq s_1^o$, or of a neighbour u of $\bar{t}_3$ in R_F (different from $\bar{t}_2$), if $\bar{t}_3 = s_1^o$. Let B_3 be the path of length at most two from $\bar{t}_3$ to $\mathcal{A}_1$: $B_3 := \bar{t}_3\bar{t}_3'$ if $\bar{t}_3 \neq s_1^o$ and $B_3 := \bar{t}_3 u \bar{t}_3'$ if $\bar{t}_3 = s_1^o$. Find a path $\bar{L}_3$ in $\mathcal{S}_1$ between $\bar{s}_3$ and $\bar{t}_3$ that contains a subpath in $\mathcal{A}_1$ between $\bar{s}_3'$ and $\bar{t}_3'$; here use the connectivity of $\mathcal{A}_1$ (Proposition 10.2.2). We next find a path M_2 in F_1 from $\bar{s}_2$ to R_F that is of length at most two and is disjoint from $V(B_3) \cup \{s_1, \bar{t}_1, \bar{t}_3\}$. There are exactly four disjoint such $\bar{s}_2 - R_F$ paths of length at most two, one through each of the neighbours of $\bar{s}_2$ in F_1. One such path is $\bar{s}_2\bar{s}_3\bar{t}_2$. Among the remaining three $\bar{s}_2 - R_F$ paths, since none of them contains s_1 or $\bar{t}_1$ and since $|V(B_3) \cap V(R_F)| \leq 2$, the existence of a path M_2 is ensured. Let $\hat{s}_2 := V(M_2) \cap V(R_F)$. Find a path $\bar{L}_2' := \hat{s}_2 - \bar{t}_2$ in R_F that is disjoint from $V(B_3)$, using the 3-connectivity of R_F. Let $\bar{L}_2 := \bar{s}_2 M_2 \hat{s}_2 \bar{L}_2' \bar{t}_2$. Since the vertices in $V(M_2) \cap V(R) \cup \{\bar{s}_3\}$ cannot separate s_1 from $\bar{t}_1$ in R (Corollary 11.1.4), find a path $\bar{L}_1 := s_1 - \bar{t}_1$ in R disjoint from $V(M_2) \cap V(R) \cup \{\bar{s}_3\}$; the set $V(M_2)$ has cardinality one or contains one edge. Extend the linkage $\{\bar{L}_1, \bar{L}_2, \bar{L}_3\}$ to a required Y-linkage.

Subcase C. The vertex opposite to s_1 in F_1 coincides with $\bar{t}_1$.

If all the four neighbours of $\bar{t}_1$ in F_1 belong to $\bar{X}$, we have what we dubbed Configuration C.2 in Case 1 of Theorem 11.3.5. We deal with Configuration C.2 here as we did then. For the details please refer to the proof of (C.2) in Theorem 11.3.5.

Hence we may suppose that the vertex opposite to s_1 in F_1 coincides with $\bar{t}_1$ and that $\bar{t}_1$ has a neighbour $\bar{t}_1'$ not in $\bar{X}$. We reason as in Subcases A and B. We give the details for the sake of completeness.

Let R denote the 3-face in F_1 containing both s_1 and $\bar{t}_1'$; $\operatorname{dist}_R(s_1, \bar{t}_1') = 3$. Let R_F be the 3-face of F_1 disjoint from R. Let J_1 be the other facet in $\mathcal{S}_1$ containing R and let R_J be the 3-face of J_1 disjoint from R.

Suppose R contains a pair $\{\bar{s}_i, \bar{t}_i\}$ $(i = 2, 3)$, say $(\bar{s}_2, \bar{t}_2)$. There are at most five terminals in R. Since the smallest face in R containing s_1 and $\bar{t}_1'$ is 3-dimensional, the pairs $\{\pi_{R_J}^{J_1}(s_1), \pi_{R_J}^{J_1}(\bar{t}_1')\}$ and $\{\pi_{R_J}^{J_1}(\bar{s}_2), \pi_{R_J}^{J_1}(\bar{t}_2)\}$ do not lie in Configuration F in R_J, and so they can be linked therein through disjoint paths $\bar{L}_1'$ and $\bar{L}_2'$ thanks to Proposition 9.0.5. Let $\bar{L}_1 := s_1 \pi_{R_J}^{J_1}(s_1) \bar{L}_1' \pi_{R_J}^{J_1}(\bar{t}_1') \bar{t}_1' \bar{t}_1$ and $\bar{L}_2 := \bar{s}_2 \pi_{R_J}^{J_1}(s_2) \bar{L}_2' \pi_{R_J}^{J_1}(\bar{t}_2) \bar{t}_2$. From the 3-connectivity of R_F follows the existence of a path $\bar{L}_3'$ in R_F between $\pi_{R_F}^{F_1}(\bar{s}_3)$ and $\pi_{R_F}^{F_1}(\bar{t}_3)$ that avoids $\bar{t}_1$; note that $\bar{s}_3$ may coincide with $\pi_{R_F}^{F_1}(\bar{s}_3)$, and so may $\bar{t}_3$ with $\pi_{R_F}^{F_1}(\bar{t}_3)$. Let $\bar{L}_3 := \bar{s}_3 \pi_{R_F}^{F_1}(\bar{s}_3) \bar{L}_3' \pi_{R_F}^{F_1}(\bar{t}_3) \bar{t}_3$. The linkage $\{\bar{L}_1, \bar{L}_2, \bar{L}_3\}$ can be naturally extended to a $\bar{Y}$-linkage, which in turn can be extended to a Y-linkage.

Suppose that the ridge R contains no pair $\{\bar{s}_i, \bar{t}_i\}$ $(i = 2, 3)$ and that the ridge R_F contains a pair $\{\bar{s}_i, \bar{t}_i\}$ $(i = 2, 3)$, say $\{\bar{s}_2, \bar{t}_2\}$. Then, there are at most five terminals in R_F. If there are at most four terminals in R_F, the 3-connectivity of R_F ensures the existence of an $\bar{X}$-valid path $\bar{L}_2 := \bar{s}_2 - \bar{t}_2$ in R_F; if there are exactly five terminals in R_F, applying Lemma 11.1.7 to the polytope F_1 and its facet R_F gives either an $\bar{X}$-valid path $\bar{L}_2 := \bar{s}_2 - \bar{t}_2$ or an $\bar{X}$-valid path $\bar{L}_3 := \bar{s}_3 - \bar{t}_3$ in R_F. As a result, without loss of generality, we can assume there is an $\bar{X}$-valid path $\bar{L}_2 := \bar{s}_2 = \bar{t}_2$ in R_F. Then, with the help of Lemma 11.3.4, find pairwise distinct neighbours $\bar{s}_3'$ and $\bar{t}_3'$ in $\mathcal{A}_1$ of $\bar{s}_3$ and $\bar{t}_3$, respectively, and a path $\bar{L}_3$ in $\mathcal{S}_1$ between $\bar{s}_3$ and $\bar{t}_3$ that contains a subpath in $\mathcal{A}_1$ between $\bar{s}_3'$ and $\bar{t}_3'$; here use the connectivity of $\mathcal{A}_1$ (Proposition 10.2.2). In addition, let $\bar{L}_1$ be a path in F_1 between s_1 and $\bar{t}_1$ that uses a subpath in R between s_1 and $\bar{t}_1'$; here use the 3-connectivity of R to avoid any terminal in R. The linkage $\{\bar{L}_1, \bar{L}_2, \bar{L}_3\}$ can be naturally extended to a Y-linkage.

Assume neither R nor R_{F_1} contains a pair $\{\bar{s}_i, \bar{t}_i\}$ ($i = 2, 3$). Without loss of generality, we can assume $\bar{s}_2, \bar{s}_3 \in R$ and $\bar{t}_2, \bar{t}_3 \in R_F$.

For *some* $i = 2, 3$, there exists a path M_i in F_1 from $\bar{s}_i$ to R_F that is of length at most two and is disjoint from $\bar{t}_1'$ and $\bar{X} \setminus \{\bar{s}_i, \bar{t}_i\}$. Suppose there is no such $\bar{s}_3 - R_F$ path. Then the neighbours of $\bar{s}_3$ in F_1 would be $s_1, \bar{t}_1', \bar{s}_2$ from R and $\bar{t}_2$ from R_F. But, since there are exactly four $\bar{s}_2 - R_F$ paths of length at most two in F_1 and since the vertex $\bar{s}_2$ could not be adjacent to $\{s_1, \bar{t}_1'\}$, the existence of such an $\bar{s}_2 - R_F$ path would be guaranteed. Hence assume the existence of such a path $M_3 := \bar{s}_3 - R_F$. Let $\hat{s}_3 := V(M_3) \cap V(R_F)$. Find an $\bar{X}$-valid path $\bar{L}_3' := \hat{s}_3 - \bar{t}_3$ in R_F using its 3-connectivity. Let $\bar{L}_3 := \bar{s}_3 M_3 \hat{s}_3 \bar{L}_3' \bar{t}_3$. Then find pairwise distinct neighbours $\bar{s}_2'$ and $\bar{t}_2'$ of $\bar{s}_2$ and $\bar{t}_2$, respectively, in $\mathcal{A}_1$, and a path $\bar{L}_2 := \bar{s}_2 - \bar{t}_2$ in $\mathcal{S}_1$ that contains a subpath $\bar{s}_2' - \bar{t}_2'$ in $\mathcal{A}_1$ (using the connectivity of $\mathcal{A}_1$). Since Corollary 11.1.4 ensures that any separator of size three in a 3-cube must be independent, we can find an $\bar{L}_1' := s_1 - \bar{t}_1'$ in R that is disjoint from $\bar{s}_2$ and $V(M_3) \cap V(R)$; the set $V(M_3) \cap V(R)$ has either cardinality one or contains an edge. Let $\bar{L}_1 := s_1 \bar{L}_1' \bar{t}_1' \bar{t}_1$. Extend the linkage $\{\bar{L}_1, \bar{L}_2, \bar{L}_3\}$ to a required Y-linkage.

And finally, the proof of the proposition is complete and so is the proof of Theorem 11.3.5. □

11.4 Strong Linkedness of Cubical Polytopes

The property of strong linkedness, see Theorems 11.1.10 and 11.1.11, also holds for cubical polytopes.

Theorem 11.4.1 (Strong linkedness of cubical polytopes). *For every $d \neq 3$, a cubical d-polytope is strongly $\lfloor (d+1)/2 \rfloor$-linked.*

Proof. Let P be a cubical d-polytope. For odd d Theorems 11.3.5 and 11.4.1 are equivalent. So assume $d = 2k$. Let X be a set of $d+1$ vertices in P. Arbitrarily pair $2k$ vertices in X to obtain $Y := \{\{s_1, t_1\}, \ldots, \{s_k, t_k\}\}$. Let x be the vertex of X not paired in Y. We find a Y-linkage $\{L_1, \ldots, L_k\}$ where each path L_i joins the pair $\{s_i, t_i\}$ and avoids the vertex x.

Using the d-connectivity of $G(P)$ and Menger's theorem, bring the $d = 2k$ terminals in $X \setminus \{x\}$ to the link of x in the boundary complex of P through $2k$ disjoint paths L_{s_i} and L_{t_i} for $i \in [1, k]$. Let $s_i' := V(L_{s_i}) \cap \operatorname{link}(x)$ and $t_i' := V(L_{t_i}) \cap \operatorname{link}(x)$ for $i \in [1, k]$. Thanks to Proposition 9.0.11, the link of x is combinatorially equivalent to a cubical $(d-1)$-polytope, which is $d/2$-linked by Theorem 11.3.5. Using the $d/2$-linkedness of $\operatorname{link}(x)$, find disjoint paths $L_i' := s_i' - t_i'$ in $\operatorname{link}(x)$. Observe that all these k paths $\{L_1', \ldots, L_k'\}$ avoid x. Extend each path L_i' with L_{s_i} and L_{t_i} to form a path $L_i := s_i - t_i$ for $i \in [1, k]$. The paths $\{L_1, \ldots, L_k\}$ forms the desired Y-linkage. □

Bibliography

[1] Ron M. Adin, Daniel Kalmanovich, and Eran Nevo. On the cone of f-vectors of cubical polytopes. *Discrete Math.*, 184(1–3):25–48, 1998.

[2] Hédy Attouch, Jérôme Bolte, Patrick Redont, and Antoine Soubeyran. Proximal alternating minimization and projection methods for nonconvex problems: an approach based on the Kurdyka–Łojasiewicz inequality. *Math. Oper. Res.*, 35(2):438–457, 2010.

[3] Didier Aussel. Subdifferential properties of quasiconvex and pseudoconvex functions: Unified approach. *Optimization*, 97(1):29–45, 1998.

[4] Didier Aussel, Jean-Noël Corvellec, and Marc Lassonde. Subdifferential characterization of quasiconvexity and convexity. *J. Convex Anal.*, 1:195–201, 1994.

[5] Didier Aussel, Jean-Noël Corvellec, and Marc Lassonde. Mean-value property and subdifferential criteria for lower semicontinuous functions. *Trans. Amer. Math. Soc.*, 347:4147–4161, 1995.

[6] Dominique Azé. A survey on error bounds for lower semicontinuous functions. In *Proceedings of 2003 MODE-SMAI Conference*, volume 13 of *ESAIM Proc.*, pages 1–17. EDP Sci., Les Ulis, 2003.

[7] Dominique Azé and Jean-Noël Corvellec. Nonlinear error bounds via a change of function. *J. Optim. Theory Appl.*, 172(1):9–32, 2017.

[8] Andrew Bakan, Frank Deutsch, and Wu Li. Strong CHIP, normality, and linear regularity of convex sets. *Trans. Amer. Math. Soc.*, 357(10):3831–3863, 2005.

[9] Michel L. Balinski. On the graph structure of convex polyhedra in n-space. *Pacific J. Math.*, 11:431–434, 1961.

[10] Emmanuel N. Barron, Rafal K. Goebel, and Robert R. Jensen. Function which are quasiconvex under linear perturbations. *SIAM J. Optim.*, 22:1089–1108, 2012.

[11] Emmanuel N. Barron, Rafal K. Goebel, and Robert R. Jensen. The quasiconvex envelope through first-order partial differential equations which characterize quasiconvexity of nonsmooth functions. *Discrete Contin. Dyn. Syst. Ser. B*, 17:1693–1706, 2012.

[12] Heinz H. Bauschke and Jonathan M. Borwein. On the convergence of von Neumann's alternating projection algorithm for two sets. *Set-Valued Anal.*, 1(2):185–212, 1993.

[13] Heinz H. Bauschke and Jonathan M. Borwein. On projection algorithms for solving convex feasibility problems. *SIAM Rev.*, 38(3):367–426, 1996.

[14] Heinz H. Bauschke, Jonathan M. Borwein, and Wu Li. Strong conical hull intersection property, bounded linear regularity, Jameson's property (G), and error bounds in convex optimization. *Math. Program., Ser. A*, 86(1):135–160, 1999.

[15] Heinz H. Bauschke, Jonathan M. Borwein, and Paul Tseng. Bounded linear regularity, strong CHIP, and CHIP are distinct properties. *J. Convex Anal.*, 7(2):395–412, 2000.

[16] Heinz H. Bauschke, Minh N. Dao, Dominikus Noll, and Hung M. Phan. Proximal point algorithm, Douglas–Rachford algorithm and alternating projections: a case study. *J. Convex Anal.*, 23(1):237–261, 2016.

[17] Heinz H. Bauschke, D. Russell Luke, Hung M. Phan, and Xianfu Wang. Restricted normal cones and the method of alternating projections: applications. *Set-Valued Var. Anal.*, 21(3):475–501, 2013.

[18] Heinz H. Bauschke, Russell D. Luke, Hung M. Phan, and Xianfu Wang. Restricted normal cones and the method of alternating projections: theory. *Set-Valued Anal.*, 21(3):431–473, 2013.

[19] Ewa M. Bednarczuk and Monika Syga. On lagrange duality for several classes of nonconvex optimization problems. In Hoai An Le Thi, Hoai Minh Le, and Tao Pham Dinh, editors, *Optimization of Complex Systems: Theory, Models, Algorithms and Applications*, pages 175–181, Cham, 2020. Springer International Publishing.

[20] Mira Bivas, Mikhail Krastanov, and Nadezhda Ribarska. On tangential transversality. *J. Math. Anal. Appl.*, 2019.

[21] G Blind and R. Blind. The almost simple cubical polytope. *Proc. Amer. Math. Soc.*, 147:1851–1866, 2019.

[22] Béla Bollobás and Andrew Thomason. Highly linked graphs. *Combinatorica*, 16(3):313–320, 1996.

[23] Jonathan M. Borwein, Regina S. Burachik, and Liangjin Yao. Conditions for zero duality gap in convex progamming. *J. Nonlinear Convex Anal.*, 15(1):167–190, 2014.

[24] Jonathan M. Borwein and Alejandro Jofré. A nonconvex separation property in Banach spaces. *Math. Methods Oper. Res.*, 48(2):169–179, 1998.

[25] Jonathan M. Borwein and Qiji J. Zhu. *Techniques of Variational Analysis.* Springer, New York, 2005.

[26] Haim Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations.* Springer, USA, 2011.

[27] Gunnar Brinkmann, Sam Greenberg, Catherine Greenhill, Brendan D. McKay, Robin Thomas, and Paul Wollan. Generation of simple quadrangulations of the sphere. *Discrete Math.*, 306(1):33–54, 2005.

[28] Hoa T. Bui, Nguyen Duy Cuong, and Alexander Y. Kruger. Transversality of collections of sets: Metric characterizations. *Preprint, Optimization Online: 2019-07-7275*, 2019.

[29] Hoa T. Bui and Alexander Y. Kruger. About extensions of the extremal principle. *Vietnam J. Math.*, 46(2):215–242, 2018.

[30] Hoa T. Bui and Alexander Y. Kruger. Extremality, stationarity and generalized separation of collections of sets. *J. Optim. Theory Appl.*, 182(1):211–264, 2019.

[31] Hoa T. Bui, Scott B. Lindstrom, and Vera Roshchina. VADU 2018 open problem session. *Preprint, arXiv*, 1806.01163, 2018.

[32] Hoa T. Bui, Guillermo Pineda-Villavicencio, and Julien Ugon. The linkedness of cubical polytopes. *arXiv:1802.09230*, 2018.

[33] Hoa T. Bui, Guillermo Pineda-Villavicencio, and Julien Ugon. Connectivity of cubical polytopes. *J. Combin. Theory Ser. A*, 169:105–126, 2020.

[34] R. S. Burachik and A. M. Rubinov. Abstract convexity and augmented lagrangians. *SIAM J. Optim.*, 18(2):413–436, 2007.

[35] Regina S. Burachik and Alexander M. Rubinov. On abstract convexity and set valued analysis. *J. Nonlinear Convex Anal.*, 9(1):105–123, 2008.

[36] R. Cibulka, M. Fabian, and Alexander Y. Kruger. On semiregularity of mappings. *J. Math. Anal. Appl.*, 473(2):811–836, 2019.

[37] Frank H. Clarke. *Optimization and Nonsmooth Analysis.* John Wiley & Sons Inc., New York, 1983.

[38] Rafael Correa, Alejandro Jofré, and Lionel Thibault. Characterization of lower semicontinuous convex functions. *Proc. Amer. Math. Soc.*, 116:67–72, 1992.

[39] Nguyen Duy Cuong and Alexander Y. Kruger. Nonlinear semitransversality and transversality of collections of sets: Dual space characterizations. *Preprint*, 2019.

[40] Nguyen Duy Cuong and Alexander Y. Kruger. Nonlinear subtransversality of collections of sets: Dual space characterizations. *Preprint*, 2019.

[41] Nguyen Duy Cuong and Alexander Y. Kruger. Nonlinear transversality of collections of sets: Dual space sufficient characterizations. *Preprint, Optimization Online: 2019-07-7313*, 2019.

[42] Nguyen Duy Cuong and Alexander Y. Kruger. Nonlinear transversality of collections of sets: Primal space characterizations. *Preprint, arXiv: 1902.06186*, 2019.

[43] Nguyen Duy Cuong and Alexander Y. Kruger. Nonlinear transversality of collections of sets: Dual space necessary characterizations. *J. Convex Anal.*, 27, 2020.

[44] M.H. Daryaei and Hossein Mohebi. Abstract convexity of extended real-valued icr functions. *Optimization*, 62:835–855, 2013.

[45] Ennio De Giorgi, Antonio Marino, and Mario Tosques. Evolution problerns in in metric spaces and steepest descent curves. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8)*, 68(3):180–187, 1980. In Italian. English translation: Ennio De Giorgi, Selected Papers, Springer, Berlin 2006, 527–533.

[46] Reinhard Diestel. Graph Theory*, 5th ed., Graduate Texts in Mathematics*, volume 73. Springer-Verlag, Berlin, 2017.

[47] Luc T. Dinh. Characterizations of quasiconvex functions. *Bull. Australian. Math. Soc.*, 48:393–406, 1993.

[48] AliReza Doagooei and Hossein Mohebi. Monotonic analysis over ordered topological vector spaces: Iv. *J. Glob. Optim.*, 45.

[49] Szymon Dolecki. Tangency and differentiation: some applications of convergence theory. *Ann. Mat. Pura Appl. (4)*, 130:223–255, 1982.

[50] Maksim V. Dolgopolik. New global optimality conditions for nonsmooth DC optimization problems. *Preprint, arXiv:*, 1808.03590:1–29, 2018.

[51] Asen L. Dontchev and Tyrrel R.l Rockafellar. *Implicit Functions and Solution Mappings. A View from Variational Analysis.* Springer Series in Operations Research and Financial Engineering. Springer, New York, 2 edition, 2014.

[52] Dmitriy Drusvyatskiy, Alexander D. Ioffe, and Adrian S. Lewis. Transversality and alternating projections for nonconvex sets. *Found. Comput. Math.*, 15(6):1637–1651, 2015.

[53] Dmitriy Drusvyatskiy and Adrian S. Lewis. Inexact alternating projections on nonconvex sets. *Vietnam J. Math.*, 47(3):669–681, 2019.

[54] Alexander Y. Dubovitskii and Alexei A. Miljutin. Extremum problems in the presence of restrictions. *USSR Comput. Maths. Math. Phys.*, 5:1–80, 1965.

[55] Joydeep Dutta, Enrique J. Martínez-Legaz, and Alexander M. Rubinov. Monotonic analysis over cones: I. *Optimization*, 53:165–177, 2004.

[56] Joydeep Dutta, Enrique J. Martínez-Legaz, and Alexander M. Rubinov. Monotonic analysis over cones: Ii. *Optimization*, 53:529–547, 2004.

[57] Andrew C. Eberhard and Hossein Mohebi. Maximal abstract monotonicity and generalized fenchel's conjugation formulas. *Set-Valued Anal.*, 18:79–108, 2010.

[58] I. Ekeland. On the variational principle. *J. Math. Anal. Appl.*, 47:324–353, 1974.

[59] Marián Fabian. Subdifferentiability and trustworthiness in the light of a new variational principle of Borwein and Preiss. *Acta Univ. Carolinae*, 30:51–56, 1989.

[60] S. Gallivan. Disjoint edge paths between given vertices of a convex polytope. *J. Combin. Theory Ser. A*, 39(1):112–115, 1985.

[61] Jacob E. Goodman, Joseph O'Rourke, and Csaba D. Tóth. Handbook of discrete and computational geometry*, 1st ed., Discrete Mathematics and its Applications (Boca Raton).* Chapman & Hall/CRC, Boca Raton, 1997.

[62] Branko Grünbaum. Convex polytopes*, 2nd ed., Graduate Texts in Mathematics*, volume 221. Springer-Verlag, New York, 2003.

[63] Victor Guillemin and Alan Pollack. *Differential Topology.* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1974.

[64] H. Hartwig. On generalized convex function, optimization. *Set-Valued Var. Anal.*, 60(3):49–60, 1983.

[65] Robert Hesse and Russell D. Luke. Nonconvex notions of regularity and convergence of fundamental algorithms for feasibility problems. *SIAM J. Optim.*, 23(4):2397–2419, 2013.

[66] Morris W. Hirsch. *Differential Topology.* Springer-Verlag, New York-Heidelberg, 1976. Graduate Texts in Mathematics, No. 33.

[67] Phu X. Hoang and An T. Pham. Stable generalization of convex functions. *Optimization*, 38:309–318, 1996.

[68] Alexander D. Ioffe. Approximate subdifferentials and applications. III. The metric theory. *Mathematika*, 36(1):1–38, 1989.

[69] Alexander D. Ioffe. Fuzzy principles and characterization of trustworthiness. *Set-Valued Anal.*, 6:265–276, 1998.

[70] Alexander D. Ioffe. Metric regularity and subdifferential calculus. *Russian Math. Surveys*, 55:501–558, 2000.

[71] Alexander D. Ioffe. Abstract convexity and non-smooth analysis. *Adv. Math. Econ*, 3:45–61, 2001.

[72] Alexander D. Ioffe. Metric regularity – a survey. Part I. Theory. *J. Aust. Math. Soc.*, 101(2):188–243, 2016.

[73] Alexander D. Ioffe. Metric regularity – a survey. Part II. Applications. *J. Aust. Math. Soc.*, 101(3):376–417, 2016.

[74] Alexander D. Ioffe. Transversality in variational analysis. *J. Optim. Theory Appl.*, 174(12):343–366, 2017.

[75] Alexander D. Ioffe. *Variational Analysis of Regular Mappings. Theory and Applications.* Springer Monographs in Mathematics. Springer, 2017.

[76] Alexander D. Ioffe and Vladimir M. Tikhomirov. *Theory of Extremal Problems*, volume 6 of *Studies in Mathematics and Its Applications.* North-Holland Publishing Co., Amsterdam, 1979.

[77] Singer Ivan. *Abstract convex analysis.* Wiley-Interscience, New York, 2006.

[78] Jeya Jeyakumar, Alexander M. Rubinov, and Zhen-Yu Wu. Generalized fenchel's conjugation formulas and duality for abstract convex functions. *J. Optim. Theory Appl.*, 132:441–458, 2007.

[79] M. Joswig and Günter M. Ziegler. Neighborly cubical polytopes. *Discrete Comput. Geom.*, 24(2–3):325–344, 2000.

[80] Ken-ichi Kawarabayashi, Alexandr V. Kostochka, and Gexin Yu. On sufficient degree conditions for a graph to be k-linked. *Comb. Probab. Comput.*, 15(5):685–694, 2006.

[81] Phan Q. Khanh, Alexander Y. Kruger, and Nguyen H. Thao. An induction theorem and nonlinear regularity models. *SIAM J. Optim.*, 25(4):2561–2588, 2015.

[82] Diethard Klatte and Bernd Kummer. *Nonsmooth Equations in Optimization. Regularity, Calculus, Methods and Applications*, volume 60 of *Nonconvex Optimization and its Applications.* Kluwer Academic Publishers, Dordrecht, 2002.

[83] Alexander Y. Kruger. *Generalized differentials of nonsmooth functions.* VINITI no. 1332-81. Minsk, 1981. 67 pp. In Russian. Available at https://asterius.ballarat.edu.au/akruger/research/publications.html.

[84] Alexander Y. Kruger. *ε-semidifferentials and ε-normal elements.* VINITI no. 1331-81. Minsk, 1981. 76 pp. In Russian. Available at https://asterius.ballarat.edu.au/akruger/research/publications.html.

[85] Alexander Y. Kruger. About extremality of systems of sets. *Dokl. Nats. Akad. Nauk Belarusi*, 42(1):24–28, 1998. In Russian. Available at https://asterius.ballarat.edu.au/akruger/research/publications.html.

[86] Alexander Y. Kruger. Strict (ε, δ)-semidifferentials and extremality of sets and functions. *Dokl. Nats. Akad. Nauk Belarusi*, 44(2):19–22, 2000. In Russian. Available at https://asterius.ballarat.edu.au/akruger/research/publications.html.

[87] Alexander Y. Kruger. Strict (ε, δ)-subdifferentials and extremality conditions. *Optimization*, 51(3):539–554, 2002.

[88] Alexander Y. Kruger. On Fréchet subdifferentials. *J. Math. Sci.*, 116(3):3325–3358, 2003.

[89] Alexander Y. Kruger. Weak stationarity: eliminating the gap between necessary and sufficient conditions. *Optimization*, 53(2):147–164, 2004.

[90] Alexander Y. Kruger. Stationarity and regularity of set systems. *Pac. J. Optim.*, 1(1):101–126, 2005.

[91] Alexander Y. Kruger. About regularity of collections of sets. *Set-Valued Anal.*, 14(2):187–206, 2006.

[92] Alexander Y. Kruger. About stationarity and regularity in variational analysis. *Taiwan. J. Math.*, 13(6A):1737–1785, 2009.

[93] Alexander Y. Kruger. Error bounds and Hölder metric subregularity. *Set-Valued Var. Anal.*, 23(4):705–736, 2015.

[94] Alexander Y. Kruger. Error bounds and metric subregularity. *Optimization*, 64(1):49–79, 2015.

[95] Alexander Y. Kruger. Nonlinear metric subregularity. *J. Optim. Theory Appl.*, 171(3):820–855, 2016.

[96] Alexander Y. Kruger. About intrinsic transversality of pairs of sets. *Set-Valued Var. Anal.*, 26(1):111–142, 2018.

[97] Alexander Y. Kruger and Marco A. López. Stationarity and regularity of infinite collections of sets. *J. Optim. Theory Appl.*, 154(2):339–369, 2012.

[98] Alexander Y. Kruger and Marco A. López. Stationarity and regularity of infinite collections of sets. Applications to infinitely constrained optimization. *J. Optim. Theory Appl.*, 155(2):390–416, 2012.

[99] Alexander Y. Kruger, D. Russell Luke, and Nguyen H. Thao. About subtransversality of collections of sets. *Set-Valued Var. Anal.*, 25(4):701–729, 2017.

[100] Alexander Y. Kruger, D. Russell Luke, and Nguyen H. Thao. Set regularities and feasibility problems. *Math. Program., Ser. B*, 168(1-2):279–311, 2018.

[101] Alexander Y. Kruger and Boris S. Mordukhovich. New necessary optimality conditions in problems of nondifferentiable programming. In *Numerical Methods of Nonlinear Programming*, pages 116–119. Kharkov, 1979. In Russian. Available at https://asterius.ballarat.edu.au/akruger/research/publications.html.

[102] Alexander Y. Kruger and Boris S. Mordukhovich. Extremal points and the Euler equation in nonsmooth optimization problems. *Dokl. Akad. Nauk BSSR*, 24(8):684–687, 1980. In Russian. Available at https://asterius.ballarat.edu.au/akruger/research/publications.html.

[103] Alexander Y. Kruger and Boris S. Mordukhovich. *Generalized normals and derivatives and necessary conditions for an extremum in problems of nondifferentiable programming. II.* VINITI no. 494-80, 60 pp. Minsk, 1980. In Russian. Available at https://asterius.ballarat.edu.au/akruger/research/publications.html.

[104] Alexander Y. Kruger and Nguyen H. Thao. About uniform regularity of collections of sets. *Serdica Math. J.*, 39:287–312, 2013.

[105] Alexander Y. Kruger and Nguyen H. Thao. About $[q]$-regularity properties of collections of sets. *J. Math. Anal. Appl.*, 416(2):471–496, 2014.

[106] Alexander Y. Kruger and Nguyen H. Thao. Quantitative characterizations of regularity properties of collections of sets. *J. Optim. Theory Appl.*, 164(1):41–67, 2015.

[107] Alexander Y. Kruger and Nguyen H. Thao. Regularity of collections of sets and convergence of inexact alternating projections. *J. Convex Anal.*, 23(3):823–847, 2016.

[108] D. G. Larman and P. Mani. On the existence of certain configurations within graphs and the 1-skeletons of polytopes. *Proc. London Math. Soc.*, 20(3):144–160, 1970.

[109] Adrian S. Lewis, Russell D. Luke, and Jerome Malick. Local linear convergence for alternating and averaged nonconvex projections. *Found. Comput. Math.*, 9(4):485–513, 2009.

[110] Adrian S. Lewis and Jerome Malick. Alternating projections on manifolds. *Math. Oper. Res.*, 33(1):216–234, 2008.

[111] Chong Li, Kung Fu Ng, and Ting Kei Pong. The SECQ, linear regularity, and the strong CHIP for an infinite system of closed convex sets in normed linear spaces. *SIAM J. Optim.*, 18(2):643–665, 2007.

[112] Guoyin Li, Kung Fu Ng, and Xi Yin Zheng. Unified approach to some geometric results in variational analysis. *J. Funct. Anal.*, 248(2):317–343, 2007.

[113] Guoyin Li, Chunming Tang, Gaohang Yu, and Zengxin Wei. On a separation principle for nonconvex sets. *Set-Valued Anal.*, 16:851–860, 2008.

[114] Russell D. Luke, Marc Teboulle, and Nguyen H. Thao. Necessary conditions for linear convergence of iterated expansive, set-valued mappings. *Math. Program., Ser. A*. doi.org/10.1007/s10107-018-1343-8.

[115] B. Martos. *Nonlinear Programming-Theory and Methods.* Akaderniai Kiado, Budapest, 1975.

[116] Peter McMullen. Transforms, diagrams and representations. *Contributions to Geometry: Proceedings of the Geometry-Symposium held in Siegen from 28 June 1978 to 1 July 1978 (Jürgen Tölke and Jörg M. Wills, eds.), Birkhäuser Basel*, 20(3):92–130, 1979.

[117] K. Menger. Zur allgemeinen kurventheorie. *Fund. Math.*, 10(1):96–115, 1927.

[118] Gábor Mészáros. On linkedness in the Cartesian product of graphs. *Period. Math. Hungar.*, 72(2):130–138, 2016.

[119] Hossein Mohebi and M. Samet. Abstract convexity of topical functions. *J. Glob. Optim.*, 58:365–357, 2014.

[120] Boris S. Mordukhovich. *Variational Analysis and Generalized Differentiation. I: Basic Theory*, volume 330 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Berlin, 2006.

[121] Angelia Nedić, Asuman Ozdaglar, and Alexander M. Rubinov. Abstract convexity for nonconvex optimization duality. *Optimization*, 56:655–674, 2007.

[122] Angelia Nedić, Asuman Ozdaglar, and Alexander M. Rubinov. Monotonic analysis over cones: Iii. *J. Convex Anal.*, 15:581–592, 2008.

[123] Kung Fu Ng and Wei Hong Yang. Regularities and their relations to error bounds. *Math. Program., Ser. A*, 99(3):521–538, 2004.

[124] Kung Fu Ng and Rui Zang. Linear regularity and ϕ-regularity of nonconvex sets. *J. Math. Anal. Appl.*, 328(1):257–280, 2007.

[125] Huynh V. Ngai and Michel Théra. Metric inequality, subdifferential calculus and applications. *Set-Valued Anal.*, 9(1–2):187–216, 2001.

[126] Huynh Van Ngai and Michel Théra. Error bounds in metric spaces and application to the perturbation stability of metric regularity. *SIAM J. Optim.*, 19(1):1–20, 2008.

[127] Thao H. Nguyen, Hoa T. Bui, Nguyen Duy Cuong, and Michel Verhaegen. Some new characterizations of intrinsic transversality in Hilbert spaces. *Preprint, arXiv:1812.11893*, pages 1–26, 2018.

[128] Dominikus Noll and Aude Rondepierre. On local convergence of the method of alternating projections. *Found. Comput. Math.*, 16(2):425–455, 2016.

[129] Jean-Paul Penot. *Calculus Without Derivatives*, volume 266 of *Graduate Texts in Mathematics*. Springer, New York, 2013.

[130] Jean-Paul Penot. *Analysis–from Concepts to Applications*. Universitext. Springer, Cham, 2016.

[131] Paul J. Penot and Quang H. Pham. Generalized convexity of functions and generalized monotonicity of set-valued maps. *J. Optim. Theory Appl.*, 92:343–356, 1997.

[132] Micha A. Perles and Nagabhushana Prabhu. A property of graphs of convex polytopes. *J. Combin. Theory Ser. A*, 62(1):155–157, 1993.

[133] Julian Pfeifle, Vincent Pilaud, and Francisco Santos. Polytopality and Cartesian products of graphs. *Israel J. Math.*, 192(1):121–141, 2012.

[134] An T. Pham. Stability of generalized monotone maps with respect to their characterizations. *Optimization*, 55:289–299, 2006.

[135] Hung M. Phan. Linear convergence of the douglas-rachford method for two closed sets. *Optimization*, 65:369–385, 2016.

[136] Robert R. Phelps. *Convex Functions, Monotone Operators and Differentiability*, volume 1364 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2nd edition, 1993.

[137] René A. Poliquin. Subgradient monotonicity and convex functions. *Nonlinear Anal.*, 14:305–317, 1990.

[138] Mark Ramras. Minimum cutsets in hypercubes. *Discrete Math.*, 289(1):193–198, 2004.

[139] Jürgen Richter-Gebert. Realization spaces of polytopes. Springer, Berlin, 2006.

[140] Tyrrell R. Rockafelar. On the maximal monotonicity of subdifferential mappings. *Pacific J. Math.*, 33(1):209–216, 1970.

[141] Tyrrell R. Rockafellar. *Convex Analysis.* Princeton Mathematical Series, Vol. 28. Princeton University Press, Princeton, NJ, 1970.

[142] Tyrrell R. Rockafellar. Directionally Lipschitzian functions and subdifferential calculus. *Proc. London Math. Soc. (3)*, 39(2):331–355, 1979.

[143] Tyrrell R. Rockafellar and Roger J.-B. Wets. *Variational Analysis.* Springer, Berlin, 1998.

[144] Alexander Rubinov. *Abstract convexity and global optimization*, volume 44 of *Nonconvex Optimization and its Applications.* Kluwer Academic Publishers, Dordrecht, 2000.

[145] Alexander M. Rubinov. *Abstract Convexity and Global Optimization.* Kluwer Academic Publishers, Dordrecht, 2000.

[146] Alexander M. Rubinov. Abstract convexity, global optimization and data classification. *Opsearch*, 38(3):247–265, 2001.

[147] Alexander M. Rubinov and Zari Dzalilov. Abstract convexity of positively homogeneous functions. *J. Stat. Manag. Syst.*, 5:1–20, 2002.

[148] Alexander M. Rubinov and Amos Uderzo. On global optimality conditions via separation functions. *J. Optim. Theory Appl.*, 109:345–370, 2001.

[149] Alexander M. Rubinov and Zhen Y. Wu. Optimality conditions in global optimization and their applications. *Math. Program., Ser. B*, 120:101–123, 2009.

[150] G. T. Sallee. Incidence graphs of convex polytopes. *J. Combinatorial Theory*, 2:466–506, 1967.

[151] Raman Sanyal and Günter M. Ziegler. Construction and analysis of projected deformed products. *Discrete Comput. Geom.*, 43(2):412–435, 2010.

[152] Siegfried Schaible and W. T. Ziernba. *On Generalized Concavity in Optimization and Economics.* Academic Press, New York-London-Toronto-Sydney-San Francisco, 1981.

[153] P. D. Seymour. Disjoint paths in graphs. *Discrete Math.*, 29(3):293–309, 1980.

[154] A.P. Shveidel. Abstract convex sets with respect to the class of general min-type functions. *Optimization*, 52:571–579, 2003.

[155] Alexander S. Strekalovsky. Global optimality conditions and exact penalization. *Optim. Lett.*, Nov 2017.

[156] Monika Syga. Minimax theorems for ϕ-convex functions: sufficient and necessary conditions. *Optimization*, 65(3):635–649, 2016.

[157] Monika Syga. Minimax theorems for extended real-valued abstract convex–concave functions. *J. Optim. Theory Appl.*, 176:306–318, 2018.

[158] Nguyen H. Thao. *About Regularity Properties in Variational Analysis and Applications in Optimization.* PhD thesis, Federation University Australia, Ballarat, 2015.

[159] Nguyen H. Thao. A convergent relaxation of the douglas-rachford algorithm. *Comput. Optim. Appl.*, 70(3):841–863, 2018.

[160] Nguyen H. Thao, Oleg Soloviev, and Michel Verhaegen. Convex combination of alternating projection and douglas–rachford algorithm for phase retrieval. *Preprint*, 2018.

[161] C. Thomas. 2-linked graphs. *European J. Combin.*, 1(4):371–378, 1980.

[162] Robin Thomas and Paul Wollan. An improved linear edge bound for graph linkages. *European J. Combin.*, 26(3-4):309–324, 2005.

[163] Nguyen Thi Quynh Trang. A note on an approximate mean value theorem for fréchet subgradients. *J. Nonlinear Convex Anal.*, 75:380–383, 2012.

[164] Axel Werner and Ronald F. Wotzlaw. On linkages in polytope graphs. *Adv. Geom.*, 11(3):411–427, 2011.

[165] Ronald F. Wotzlaw. On linkages in polytope graphs. *Adv. Geom.*, 11(3):411–427, 2011.

[166] Chaoli Yao and Shengjiei Li. Vector topical function, abstract convexity and image space analysis. *J. Optim. Theory Appl.*, 177:717–742, 2018.

[167] J. J. Ye and X. Y. Ye. Necessary optimality conditions for optimization problems with variational inequality constraints. *Math. Oper. Res.*, 22(4):977–997, 1997.

[168] Constantin Zălinescu. *Convex Analysis in General Vector Spaces.* World Scientific Publishing Co. Inc., River Edge, NJ, 2002.

[169] Xi Yin Zheng and Kung Fu Ng. Perturbation analysis of error bounds for systems of conic linear inequalities in Banach spaces. *SIAM J. Optim.*, 15(4):1026–1041, 2005.

[170] Xi Yin Zheng and Kung Fu Ng. The Lagrange multiplier rule for multifunctions in Banach spaces. *SIAM J. Optim.*, 17(4):1154–1175, 2006.

[171] Xi Yin Zheng and Kung Fu Ng. Linear regularity for a collection of subsmooth sets in Banach spaces. *SIAM J. Optim.*, 19(1):62–76, 2008.

[172] Xi Yin Zheng and Kung Fu Ng. A unified separation theorem for closed sets in a Banach space and optimality conditions for vector optimization. *SIAM J. Optim.*, 21(3):886–911, 2011.

[173] Xi Yin Zheng, Zhou Wei, and Jen-Chih Yao. Uniform subsmoothness and linear regularity for a collection of infinitely many closed sets. *Nonlinear Anal.*, 73(2):413–430, 2010.

[174] Xi Yin Zheng, Zhuowei Yang, and Jiangqian Zou. Exact separation theorem for closed sets in Asplund spaces. *Optimization*, 66(7):1065–1077, 2017.

[175] Günter M. Ziegler. Lectures on polytopes*, Graduate Texts in Mathematics*, volume 152. Springer-Verlag, New York, 1995.

[176] Günter M. Ziegler. Lectures on 0/1-polytopes. *Polytopes–combinatorics and computation (Oberwolfach, 1997), DMV Sem.*, 229:1–41, 2000.