Topics in high-dimensional geometry and optimal transport

Katarzyna Barbara Wyczesany

Department of Pure Mathematics and Mathematical Statistics

University of Cambridge

This dissertation is submitted for the degree of

Doctor of Philosophy

Sidney Sussex College October 2020

To my mum.

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared below and specified in the text.

It is not substantially the same as any that I have submitted, or is being concurrently submitted, for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. I further state that no substantial part of my dissertation has already been submitted, or is being concurrently submitted, for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution.

Chapters 1 and 2 consist of original work produced in collaboration with Prof. W. T. Gowers. Chapters 3 and 4 are the outcome of joint work with Prof. S. Artstein-Avidan and S. Sadovsky and present novel research, except when stated otherwise.

Katarzyna Barbara Wyczesany
October 2020

Topics in high-dimensional geometry and optimal transport

Katarzyna Barbara Wyczesany

Abstract

The first two chapters of this thesis are devoted to a question of Vitali Milman about the existence of well-complemented almost Euclidean subspaces of spaces uniformly isomorphic to ℓ_2^n. First, we show that there exist constants $\alpha, \epsilon > 0$ such that for every positive integer n there is a continuous odd function $\psi : S^m \to S^n$, with $m \geq \alpha n$, such that the ϵ-expansion of the image of ψ does not contain a great circle. We also show how this result is connected to the aforementioned conjecture, more precisely that it allows to build a counterexample to a variation of the question.

We then, in the second chapter, present an example of a normed space X of arbitrarily high dimension that is strongly 2-Euclidean but contains no 2-dimensional subspace that is strongly $(1 + \epsilon)$-Euclidean and strongly $(1 + \epsilon)$-complemented, where $\epsilon > 0$ is an absolute constant. This is a counterexample to the "strong" Milman problem.

The second part of this thesis involves topics related to optimal transport theory. The third chapter focuses on cost induced transforms. In particular, a family of order reversing isomorphisms $\mathcal{A}_t$, which are related to the polarity transform $\mathcal{A}$, is discussed. We prove that $\mathcal{A}_t$ is the unique, up to linear terms, order-reversing isomorphism on its image class.

In the last chapter, we give a new proof of the Rockafellar-Rüschendorf theorem about the existence of a potential for a given c-cyclically monotone set with a real-valued cost function. We then generalize the theorem to non-traditional cost functions, i.e. those which may also take the value $+\infty$, and prove that a necessary and sufficient condition for the existence of a potential is that of c-path-boundedness. Finally, we apply our theorem to show that for a continuous cost function and a compact, c-cyclically monotone set which is "bounded away from infinity" one gets a potential.

Acknowledgements

I am very grateful to many people for their encouragement and support during my time as a PhD student. First and foremost, I would like to thank my supervisor, Tim Gowers, for dedicating enormous amounts of his time to guide me with patience and kindness, and for sharing ideas with me so generously. I am indebted to Shiri Artstein-Avidan, who as my mentor and collaborator supported and believed in me. I thank her for many hours of discussion, without which the second part of this thesis would not have come to be. Many thanks to Martin Henk for all his help and words of encouragement when I needed them the most.

I would also like to thank my collaborator Shay Sadovsky for many interesting discussions and reading parts of this work. Big thanks to Alexandros Eskenazis for the advice and helpful comments on this thesis.

I am very thankful to the 2016 CCA cohort, in particular, to Laura, for the great company and support throughout our shared PhD journey. Big thanks to my many friends from the department and Sidney Sussex College, with a special thanks to Helen, for the magical Cambridge experience and distraction from work. I am also very grateful to my dear friend Harum for always brightening up my day.

My special thanks go to my best friend and boyfriend Sam, for his patience and support, especially in the final stages of my PhD. Lockdown would have been miserable without you.

Finally, I would like to thank my mum for her love, support and many many phone-calls over the years. I can never thank you enough.

Formal acknowledgements
This work was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/L016516/1 for the Cambridge Centre for Analysis (CCA).

Table of contents

1 High-dimensional tennis balls **1**

1.1 Background and Introduction 1

1.1.1 Reformulation of the problem 3

1.2 Concentration of measure and tennis balls 9

1.3 Constructing tennis balls 11

1.3.1 The definition of the tennis ball map 11

1.3.2 The definition of the set Γ 13

1.4 Proving that no great circle is contained in Γ_ϵ 16

1.4.1 Analysis of the points in a 2-dimensional subspace 18

1.5 Almost every point has an “atypical” image 22

1.6 Summary 29

2 A counterexample to a strengthening of Milman’s question **31**

2.1 Definition of the norm and initial observations 31

2.1.1 Characterization of ϵ-good points 32

2.2 Outline of the proof of Theorem 2.1.1 34

2.3 The set of sign vectors cannot be squeezed into a 4-dimensional subspace . 36

2.3.1 Case 1: $k = 1$ 38

2.3.2 Case 2: $2 \le k \le 4$ 39

2.3.3 Case 3: $k \ge 5$ 45

2.3.4 Choosing parameters 46

3 Cost functions and order reversing involutions **49**

3.1 Introduction and background 50

3.1.1 Cost induced transforms and their function classes 50

3.1.2 Connection with order-reversing isomorphisms 53

3.2 One parameter family of costs p_t 55

3.3 Uniqueness of $\mathcal{A}_t$ on $\mathcal{S}_t$ 62
3.3.1 Proof for the class $\mathcal{S}_t(\mathbb{R}^+)$ 63
3.3.2 Generalization to $\mathcal{S}_t(\mathbb{R}^n)$ 69
3.4 The associated subgradient mapping 71
3.4.1 A geometric interpretation for p_t 72

4 Existence of a potential for non-traditional costs 79
4.1 Introduction 80
4.2 The existence of a potential assuming regularity 84
4.2.1 Polar cost 84
4.2.2 The existence of a potential 86
4.3 A new proof of the Rockafellar-Rüschendorf theorem 97
4.4 Non-traditional cost functions 101
4.4.1 First proof of Theorem 4.4.3, assuming countability of the index set 104
4.4.2 Second proof of Theorem 4.4.3 110
4.4.3 Summary 111
4.5 Geometric conditions for continuous cost functions 111

References 115

Appendix A 117

Appendix B 121

Appendix C 125

Chapter 1

High-dimensional tennis balls

This chapter is based on joint work with W. T. Gowers and it is organised as follows. The first section serves as an introduction to some classical notions and results regarding geometry of finite-dimensional Banach spaces, which lays the ground for formulating a conjecture of V. Milman. We then consider a variant of the question which asks for a stronger conclusion. Other reformulations of the question are also considered.

In the second section, we relate the above mentioned question with our main result, which is the existence of a *high-dimensional tennis ball*, i.e. we prove the existence of constants $\alpha, \epsilon > 0$ such that for every large enough natural number n there is a continuous odd function $\psi : S^m \to S^n$, with $m = \alpha n$, such that the ϵ-expansion of the image of ψ does not contain a great circle.

In the third section a construction of a *tennis ball map* ψ is presented and an explanation of the main steps of the proof are outlined. We then, in sections four and five, present the full argument.

The final section is the summary of the proof and some final comments on the link between our construction and Milman's question.

1.1 Background and Introduction

In 1961, Dvoretzky [7] proved, answering an important question posed by Grothendieck [9], that there exists $N = N(k, \epsilon)$ such that for every normed space $(X, \|\cdot\|)$ of dimension $n \geq N$ one can find a k-dimensional subspace Y of X such that $d_{BM}(Y, \ell_2^k) \leq 1 + \epsilon$, where d_{BM} denotes the Banach-Mazur distance:

Definition 1.1.1. *Banach-Mazur distance between two Banach spaces X and Y is defined by*

$$d_{BM}(X,Y) = \inf\{\|T\|\|T^{-1}\| : \ T \in \mathrm{GL(X,Y)}\}.$$

The sharp dependence of N on k was established in 1971 by Milman [13], who also showed that the conclusion of Dvoretzky's theorem in fact holds for a random subspace (of appropriate dimension) with probability close to 1.

Let us call a normed space $X = (\mathbb{R}^n, \|\cdot\|)$ *C-Euclidean* if $d(X, \ell_2^n) \leq C$. A fairly straightforward use of Milman's method yields the following statement.

Theorem 1.1.2. *For every $C > 1$ and every $\epsilon > 0$ there exists $c > 0$ such that for every $n \in \mathbb{N}$, every n-dimensional C-Euclidean normed space X has a subspace Y of dimension at least cn which is $(1+\epsilon)$-Euclidean.*

In other words, under the additional hypothesis that X is C-Euclidean, one can obtain a linear dependence between the dimension of Y and the dimension of X.

Milman's proof of Dvoretzky's theorem exploited measure concentration and led to an explosion of activity in the theory of finite-dimensional normed spaces and to many striking results about the subspace structure of a normed space, often with surprisingly weak hypotheses on the space. Most of these results concerned the subspaces and not their relationship with the space itself. However, a desirable property for a subspace $Y \subset X$ is that it should be *complemented*. In an infinite-dimensional context, one says that Y is complemented if $Y = PX$ for a continuous projection P on X. In a finite-dimensional context, we need a more quantitative definition:

Definition 1.1.3. *Let $(X, \|\cdot\|)$ be an n-dimensional normed space. A subspace Y of X is called α-complemented if there exists a projection P such that $Y = PX$ and the operator norm of P is at most α, i.e.*

$$\|Px\| \leq \alpha \|x\| \quad \textit{for all} \quad x \in X.$$

There are several open problems about the existence of complemented subspaces. For example, it is not known whether there is a constant C and a function $f : \mathbb{N} \to \mathbb{N}$ that tends to infinity such that every n-dimensional normed space has a C-complemented subspace of dimension and codimension at least $f(n)$. (For a partial result in this direction, see [21].)

The results presented in the first and second chapter of this thesis originate from considering the following question of Milman.

Question 1.1.4. *Let $k \in \mathbb{N}$, $C \in \mathbb{R}$, and $\epsilon > 0$. Does there exist $N = N(C, k, \epsilon) \in \mathbb{N}$ such that every C-Euclidean normed space X of dimension $n \geq N$ has a k-dimensional subspace Y which is both $(1+\epsilon)$-Euclidean and $(1+\epsilon)$-complemented?*

We do not quite answer this question, but we give a negative answer to a question that is sufficiently close to Milman's that it seems highly unlikely that Milman's question has a positive answer. To describe our result, we introduce two further definitions.

Definition 1.1.5. *A normed space $X = (\mathbb{R}^n, \|\cdot\|)$ is called strongly C-Euclidean if there is a constant $t > 0$ such that $t|x| \leq \|x\| \leq tC|x|$ for every $x \in X$, where $|\cdot|$ denotes the standard Euclidean norm on $\mathbb{R}^n$.*

Definition 1.1.6. *A subspace Y of $X = (\mathbb{R}^n, \|\cdot\|)$ is called strongly α-complemented if the orthogonal projection $P_Y : X \to Y$ has norm at most α, i.e. $\|P_Y x\| \leq \alpha\|x\|$ for all $x \in X$.*

Let us note that the first definition is stronger than just being C-Euclidean, because instead of merely asking for any linear map T such that $|x| \leq \|Tx\| \leq C|x|$, we ask for T to be a multiple of the identity. Clearly, the second definition is also stronger than its previously defined counterpart 1.1.3 as we require that the projection is orthogonal.

These are natural strengthenings to consider, since Milman's proof of Theorem 1.1.2 in fact yields strongly $(1+\epsilon)$-Euclidean subspaces. Moreover, it seems quite unlikely that the Question 1.1.4 would have a positive answer unless the following question, which appears to be a little more approachable, also has a positive answer.

Question 1.1.7. *Let $k \in \mathbb{N}$, let $C \in \mathbb{R}$, and let $\epsilon > 0$. Does there exist $n \in \mathbb{N}$ such that every strongly C-Euclidean normed space X of dimension at least n has a k-dimensional subspace Y that is strongly $(1+\epsilon)$-Euclidean and strongly $(1+\epsilon)$-complemented?*

In the next chapter we will show that this question has a negative answer. However, we begin with an earlier result which is linked to the above question, but is also interesting in itself. In order to do so, we introduce a notion of *ϵ-good point.*

1.1.1 Reformulation of the problem

We shall now introduce a crucial definition that allows us to reformulate, in a convenient way, the condition that a subspace is both strongly $(1+\epsilon)$-Euclidean and strongly $(1+\epsilon)$-complemented. By $\langle \cdot , \cdot \rangle$ we denote the standard inner product.

Definition 1.1.8. *Let $X = (\mathbb{R}^n, \|\cdot\|)$ be a normed space and let $x \in X$. We say that x is ϵ-good if*

$$\langle x, z\rangle \leq (1+\epsilon)\frac{\|z\|}{\|x\|}|x|^2,$$

for every vector $z \in \mathbb{R}^n$.

To see what this means geometrically, consider the orthogonal projection P_x onto the 1-dimensional subspace of $\mathbb{R}^n$ generated by x. Writing x' for the normalized vector $x/|x|$, this has the formula

$$P_x z = \langle x', z\rangle x'.$$

Hence, the operator norm of P_x (as a map from X to X) is the maximum of the quantity

$$\frac{|\langle x', z\rangle|\|x'\|}{\|z\|} = \frac{|\langle x, z\rangle|\|x\|}{|x|^2\,\|z\|}$$

over all $z \in \mathbb{R}^n$. It follows that x is ϵ-good if and only if the orthogonal projection P_x has norm at most $1+\epsilon$ in the space $L(X)$ of linear operators from X to X. Clearly, the definition of ϵ-good point does not depend on the norm of the vector and hence it is enough to consider unit vectors. Let us write $S^n = \{x \in \mathbb{R}^{n+1} : |x| = 1\}$.

We now show that a subspace Y of a space X is strongly $(1+\epsilon)$-Euclidean and strongly $(1+\epsilon)$-complemented for some small ϵ if and only if every $y \in Y$ is δ-good for some small δ.

Lemma 1.1.9. *Let $X = (\mathbb{R}^n, \|\cdot\|)$ be a normed space and let $Y \subset X$ be a subspace. Then*

(i) if Y is strongly $(1+\epsilon)$-complemented and strongly $(1+\epsilon)$-Euclidean, then every $y \in Y$ is $(2\epsilon+\epsilon^2)$-good.

(ii) if $\epsilon \leq 1/9\pi^2$ and every point in Y is ϵ-good, then Y is strongly $(1+\epsilon)$-complemented and strongly $(1+3\pi\sqrt{\epsilon})$-Euclidean.

Before we prove the statement, let us note that this characterization is a reason why Question 1.1.7 is easier to attempt than the original one. Clearly, Question 1.1.7 is equivalent to

Question 1.1.10. *Let $\epsilon > 0$, $C \geq 1$ and $k \in \mathbb{N}$. Does there exist n such that if $\|\cdot\|$ is a norm on $\mathbb{R}^n$ such that $|x| \leq \|x\| \leq C|x|$ for every $x \in \mathbb{R}^n$, then the space $(\mathbb{R}^n, \|\cdot\|)$ has a subspace Y of dimension k such that every $y \in Y$ is ϵ-good?*

Proof of Lemma 1.1.9. Let P_Y be the orthogonal projection onto Y. If Y is strongly $(1+\epsilon)$-Euclidean and strongly $(1+\epsilon)$-complemented, then $\|P_Y x\| \leq (1+\epsilon)\|x\|$ for every $x \in X$

and there exists $\lambda \in \mathbb{R}$ such that $\lambda|y| \leq \|y\| \leq (1+\epsilon)\lambda|y|$ for every $y \in Y$. From this it follows that for every $y \in Y$ and every $x \in X$ we have

$$\langle y, x\rangle = \langle y, P_Y x\rangle \leq |y|\,|P_Y x| \leq |y|\,\frac{1}{\lambda}\|P_Y x\| \leq (1+\epsilon)\lambda\frac{|y|^2}{\|y\|}\frac{1}{\lambda}(1+\epsilon)\|x\| = (1+\epsilon)^2\frac{\|x\|}{\|y\|}|y|^2,$$

which implies that every point y in Y is $(2\epsilon+\epsilon^2)$-good, as claimed.

Conversely, assume that every point in Y is ϵ-good, so that for every $y \in Y$ and every $x \in X$ we have the inequality

$$\langle y, x\rangle \leq (1+\epsilon)\frac{\|x\|}{\|y\|}|y|^2.$$

Choose $x \in X$. Then $P_Y x \in Y$, so

$$|P_Y x|^2 = \langle P_Y x, P_Y x\rangle = \langle P_Y x, x\rangle \leq (1+\epsilon)\frac{\|x\|}{\|P_Y x\|}|P_Y x|^2,$$

and therefore $\|P_Y x\| \leq (1+\epsilon)\|x\|$. It follows that Y is strongly $(1+\epsilon)$-complemented.

Now assume for a contradiction that the subspace Y is not strongly $(1+a)$-Euclidean with $0 < a$. In particular, this means that we can find two unit vectors $y, w \in Y$ such that $\|y\| = \|w\|(1+a)$. Without loss of generality we may assume that $a \leq 1/2$, because if we can find unit vectors $y, w \in Y$ such that $\|y\| = \|w\|(1+a)$ then we can find y' such that $\|y'\| = \|w\|(1+a')$ for any $a' \leq a$.

Let us consider a sequence of unit vectors $w = x_0,\, x_1, \ldots, x_{m-1},\, x_m = y$ that are equally spaced along the shortest arc that joins w to y (which is unique, since w cannot equal $-y$). By the pigeonhole principle there exists i such that

$$\|x_i\|(1+a)^{1/m} \leq \|x_{i+1}\|.$$

We shall choose m in a way which ensures that x_i is a witness for x_{i+1} not being ϵ-good. Indeed, if we assume that m is at least $3\pi^2/a$ then since the angle between x_i and x_{i+1} is at most π/m we get that

$$\begin{aligned}\langle x_{i+1}, x_i\rangle\frac{\|x_{i+1}\|}{\|x_i\|\,|x_{i+1}|^2} &\geq \cos(\angle x_i x_{i+1})(1+a)^{1/m} \geq \left(1-\frac{\pi^2}{2m^2}\right)\left(1+\frac{a}{m}-\frac{a^2}{2m}\right)\\ &\geq 1+\frac{a}{m}-\frac{a^2m^2+\pi^2 m+\pi^2 a}{2m^3} \geq 1+\frac{a}{2m}.\end{aligned}$$

We used the fact that for $0 < k < 1$ and $a > 0$ we have that $(1+a)^k \geq 1 + ak - a^2\frac{k(1-k)}{2} \geq 1 + ak - \frac{a^2k}{2}$, and assumptions that $0 < a \leq 1/2$ and $m \geq 3\pi^2/a$.

It follows that the point x_{i+1} is not $\frac{a}{2m}$-good. Therefore, if every point is ϵ-good, we must have that $\frac{a}{\lceil 6\pi^2/a \rceil} \leq \epsilon$, which implies that $a \leq 3\pi\sqrt{\epsilon}$. Thus, we find that Y is strongly $(1 + 3\pi\sqrt{\epsilon})$-Euclidean, which completes the proof. $\square$

We now reformulate the definition of an ϵ-good point so that it can be applied not just to norms but to more general functions, with the aim of finding a generalization of Question 1.1.7 that does not rely on convexity.

Before we state the result, let us recall that a *support functional* of a norm $\|\cdot\|$ at x is any non-zero linear functional f such that for every y with $\|y\| \leq \|x\|$ we have $f(y) \leq f(x)$. Note that if the norm is differentiable, then writing $f(x)$ for $\|x\|$, we have that any multiple of $f'(x)$ is a support functional at x.

Proposition 1.1.11. *Let $(X, \|\cdot\|)$ be a normed space and suppose that $|x| \leq \|x\| \leq C|x|$ for every $x \in X$. For every $\delta > 0$ there exists $\epsilon > 0$ such that if $x \in X$ is any ϵ-good point, then there exist y, z such that $|x| = |y|$, $|x - y| < \delta|x|$, z is a support functional for y, and $|y - z| < \delta|x|$. Conversely, for every $\epsilon > 0$ there exists $\delta > 0$ such that x is an ϵ-good point if there exist y, z such that $|x - y| < \delta|x|$, z is a support functional for y, and $|y - z| < \delta|x|$.*

Proof. We shall do the second part first. Let $0 < \epsilon \leq 1$ and suppose that there exist y, z such that z is a support functional for y, and $|y - x|$ and $|z - y|$ are both at most $\delta|x|$.

Now let $w \in X$. Then

$$\langle w, x\rangle \leq \langle w, y\rangle + \delta|w||x| \leq \langle w, z\rangle + 2\delta|w||x|.$$

But z is a support functional for y, so

$$\langle w, z\rangle \leq \|w\|\,\|z\|^* = \|w\|\frac{\langle y, z\rangle}{\|y\|}$$

We also have that

$$\|y\| \geq \|x\| - C|x - y| \geq \|x\| - C\delta|x| \geq (1 - C\delta)\|x\|.$$

Finally, since $|x - z| \leq 2\delta|x|$ we have

$$\langle y, z\rangle \leq \langle x, z\rangle + \delta|x||z| \leq |x|^2 + 2\delta|x|^2 + \delta|x|^2(1 + 2\delta) \leq (1 + \delta)(1 + 2\delta)|x|^2.$$

Putting all this together, we find that

$$\langle w, x\rangle \leq \frac{(1+\delta)(1+2\delta)}{1-C\delta}\frac{\|w\|}{\|x\|}|x|^2 + 2\delta|w||x| \leq \left(\frac{(1+\delta)(1+2\delta)}{1-C\delta} + 2C\delta\right)\frac{\|w\|}{\|x\|}|x|^2.$$

It can be checked that if we set $\delta = \epsilon/5C$, then the factor in brackets is at most $1+\epsilon$.

For the other direction, assume that for all y such that $|y| = |x|$ and $|x-y| < \delta|x|$ we have that $|y-z| > \delta|x|$, where z is the support functional at y. We will choose z such that $|z| = |y|$.

We can assume that $|x| = 1$ and that for every unit vector y with $|y-x| < \delta$, we have that $|y-z| \geq \delta$. It follows that the component of z orthogonal to y has size at least $\delta/2$, which comes from the fact that we chose z with $|z| = |y| = 1$ and hence $\langle z, y\rangle = 1 - |z-y|^2/2 \leq 1-\delta^2/2$ and $|z - \langle y, z\rangle y|^2 = 1 - \langle y, z\rangle^2$. Further, we have that $|z| \geq 1$ for every y (since $\|y+w\| = \|y\| + \|w\| \geq \|y\| + |w|$ when w is a positive multiple of y), so for every unit vector y in the δ-neighborhood of x, the component of z orthogonal to y also has size at least $\delta/2$.

It follows that for any $\gamma < \delta$ we can find a path on the unit sphere that starts at x and ends at a point at distance at least γ from x such that the norm $\|\cdot\|$ decreases at a rate of at least $\delta/2$ along the path. This gives us a unit vector $\bar{y}$ such that $|\bar{y} - x| \leq \gamma$ and

$$\|\bar{y}\| \leq \|x\| - \gamma\delta/2 \leq \|x\|(1-\gamma\delta/2C).$$

It follows that $\langle x, \bar{y}\rangle > 1 - \gamma^2/2$, so

$$\langle x, \bar{y}\rangle > \frac{(1-\gamma^2/2)}{(1-\gamma\delta/2C)}\frac{\|\bar{y}\|}{\|x\|}|x|^2.$$

Setting $\gamma = \delta/2C$, we deduce that x is not $\delta^2/8C^2$-good. □

As mentioned before, if $f(x) = \|x\|$ is a differentiable function then $f'(x)$ is a multiple of the support functional at $x \in S^n$. A byproduct of the second part of the proof is that we may take $z = \frac{f'(y)}{|f'(y)|}$. Therefore, up to the dependence between ϵ and δ, we have that $x \in S^n$ is ϵ-good if and only if there exists $y \in S^n$ with $|x-y| < \delta$ and $|y - \frac{f'(y)}{|f'(y)|}| < \delta$.

Note that the latter condition is equivalent to $|f'(y)|(1 - \frac{\delta^2}{2}) < \langle y, f'(y)\rangle$ and since we have that $1 \leq |f'(x)| \leq C$, we get

$$|P_{y^\perp} f'(y)| = |f'(y) - \langle y, f'(y)\rangle y| \leq C(\delta^2 - \tfrac{\delta^4}{4})^{1/2} \leq C\delta,$$

where $P_{y^\perp}$ denotes the orthogonal projection onto the hyperplane orthogonal to y. Similarly one can easily find that the condition on $|P_{y^\perp} f'(y)|$ being small implies that $|y - \frac{f'(y)}{|f'(y)|}|$ is small (with some slightly changed constants). Therefore we get that for every δ there exists ϵ, such that if a point $x \in S^n$ is ϵ-good then there exists $y \in S^n$ with $|x - y| < \delta$ and $|P_{y^\perp} f'(y)| < \delta$.

Since it is enough to look only at unit vectors, if we consider a restriction of f, which is the function from the sphere (and call it again $f : S^n \to \mathbb{R}$), then the above condition means that the norm of the gradient of f at y must be small. For simplicity, instead of $\nabla_{S^n} f$, we will write f' for the gradient of a function from the sphere.

This motivates the following definition of a good* point for any differentiable function:

Definition 1.1.12. *Let $f : S^n \to \mathbb{R}$ be a differentiable function. We shall say that $x \in S^n$ is δ-good* if there exists $y \in S^n$ with $|x - y| < \delta$ and $|f'(y)| < \delta$.*

Clearly, due to Lemma 1.1.11 and the above remarks, we get that if f is a restriction of a norm, then definitions 1.1.12 and 1.1.8 are equivalent (up to the dependence between ϵ and δ).

We now note a simple fact about δ-good* points.

Lemma 1.1.13. *Let $f : S^n \to \mathbb{R}$ be a differentiable function and let $x \in S^n$ be a δ-good* point for f. Then if $\psi : S^n \to S^n$ is an invertible differentiable function such that both ψ and ψ^{-1} have Lipschitz constant at most α, then $\psi^{-1}(x)$ is an $\alpha\delta$-good* point for $f \circ \psi$.*

Proof. Assume that $\psi^{-1}(x)$ is not $\alpha\delta$-good* for $f \circ \psi$. For every $y \in S^n$ such that $d(x, y) < \delta$ we have that $d(\psi^{-1}(x), \psi^{-1}(y)) \le \alpha\delta$ and it follows that $|(f \circ \psi)'(\psi^{-1}(y))| \ge \alpha\delta$. But $(f \circ \psi)'(\psi^{-1}(y)) = \psi'(\psi^{-1}(y))^*(f'(y))$, and by the bi-Lipschitz property of ψ, this has magnitude at most $\alpha|f'(y)|$. Hence, we get that $|f'(y)| \ge \delta$, which means that x is not δ-good* and the result follows. $\square$

We conclude this section with a further question related to Milman's question.

Question 1.1.14. *Let $\delta > 0$ and let $k \in \mathbb{N}$. Does there exist n such that if $f : S^n \to [0, 1]$ is any differentiable even function, then there exists a subsphere S_Y of S^n of dimension k that consists only of δ-good* points?*

Note that when analyzing the connection between the above question and Question 1.1.7 one needs to take into account the dependence of δ and ϵ explained in Proposition 1.1.11. Nevertheless, if the answer to the Question 1.1.14 is yes (or if it is yes under the additional assumption that f is a Lipschitz function), then also the answer to the "strong" Milman's question 1.1.7 is yes.

1.2 Concentration of measure and tennis balls

Concentration of measure is a phenomenon extending an idea going back to the work of Lévy [12] concerning isoperimetric inequality on the sphere. It appears that the surface area of the ϵ-expansion of a set $A \subset S^n$, defined by $A_\epsilon = \{x \in S^n : d(x, A) < \epsilon\}$, is minimized for a given measure of A, when A is a spherical cap. Then a straightforward calculation shows that in high dimensions the mass of σ concentrates around the equator. More generally, for any measurable set A with $\sigma(A) \geq 1/2$, almost all points (in the sense of the measure σ) on S^n are within geodesic distance $n^{-1/2}$ from A.

Now, suppose we wish to find a counterexample to Question 1.1.7. As noted before, we may as well ask for the norm to be differentiable, and then from Proposition 1.1.11 follows that a point x will be ϵ-good if it is close to a point where the derivative is small. This implies that the set of points with small derivative must have very small measure, since otherwise by measure concentration its expansion will have measure very close to 1 and hence with high probability it will contain a random k-dimensional subspace.

This kind of observation already rules out many potential methods of constructing counterexamples. For instance, if one chooses a random collection of N unit vectors $u_1, \ldots, u_N$ for appropriate N and defines $\|x\|$ to be $\max_i |\langle x, u_i\rangle|$, then for almost all $x \in S^n$ the value of $\|x\|$ will be close to its minimum, which implies that x is ϵ-good.

However, there do exist norms that are C-equivalent to the Euclidean norm and have the property that the set of ϵ-good points has small measure. A simple example of such a norm is the weighted ℓ_2-norm given by the formula

$$\|x\|^2 = \sum_{i \leq n/2} 2x_i^2 + \sum_{i > n/2} x_i^2.$$

Letting A be the diagonal matrix with the first $n/2$ entries equal to 2 and the rest equal to 1, we can write the right-hand side as $\langle x, Ax\rangle$. When $x \neq 0$, the derivative of this norm at x is $Ax/\|x\|$, or if we regard the norm as a function defined on S^{n-1}, it is the projection of $Ax/\|x\|$ on to the subspace orthogonal to x. Thus, a point x is 0-good if and only if x is an eigenvector of A, which is the case if and only if it belongs to one of the two eigenspaces $\langle e_1, \ldots, e_{n/2}\rangle$ or $\langle e_{n/2+1}, \ldots, e_n\rangle$. It is a straightforward exercise to prove the more precise result that for every $\eta > 0$ there exists $\epsilon > 0$ such that x is ϵ-good only if the distance from x to one of these two subspaces is at most η. Since the set of such x has exponentially small measure, we have an example of a norm where it is not the case that almost all points are ϵ-good.

From the perspective of Milman's question, this may seem a dubious example, since the space is isometric to a Euclidean space, and therefore as far from a counterexample as it is possible to be. However, one can find examples (see Appendix A) to which this criticism does not apply.

The main point we wish to make here is that even this example gives us a strategy for building a counterexample to Question 1.1.14. Let S_1 and S_2 be the unit spheres of the two eigenspaces above, and suppose that we can find a function ψ satisfying the conditions of Lemma 1.1.13 and such that for some ϵ, the sets $\psi(S_1)_\epsilon$ and $\psi(S_2)_\epsilon$ do not contain the sphere of any 2-dimensional subspace. Then again by Lemma 1.1.13 there will exist $\eta > 0$ independent of n such that every point that is at least η away from $\psi(S_1) \cup \psi(S_2)$ is not η-good*. Also, if any 2-dimensional sphere contains a point close to $\psi(S_1)$ and a point close to $\psi(S_2)$, then since those sets are far apart (by the bi-Lipschitz property of ψ), it must also contain points far from both sets. Therefore, the function given by the composition $x \mapsto \|\psi^{-1}(x)\|$ is a counterexample.

The main result of this chapter is the following theorem

Theorem 1.2.1. *There exist constants $\alpha, \epsilon > 0$ and a bi-Lipschitz map $\psi : S^n \to S^n$ such that ψ preserves antipodal points and if X is a random linear subsphere of dimension $\lfloor \alpha n \rfloor$ then with positive probability $\psi(X)_\epsilon$ does not contain a great circle.*

Since it will appear that α provided by the theorem is significantly smaller than $1/2$, we replace the diagonal map A in the example above by a map that has roughly α^{-1} eigenspaces of dimension roughly αn with well-separated eigenvalues. In that way, we do indeed obtain a counterexample to Question 1.1.14.

We informally call the map ψ a *tennis ball map* and its image a *tennis ball* because in low dimensions it brings to mind the seam of a genuine tennis ball (though the resemblance is not perfect, since the seam of a genuine tennis ball is not centrally symmetric). More generally, we may want to call a subset of the sphere a tennis ball without referring to a function so let us introduce some useful terminology.

Definition 1.2.2. *An m-dimensional* topological subsphere *of S^n is the image of a continuous function $f : S^m \to S^n$ that preserves antipodal points.*

If an m-dimensional topological subsphere of S^n is the unit sphere of an $(m+1)$-dimensional subspace of $\mathbb{R}^n$, then we shall call it *linear*. We shall also refer to 1-dimensional linear subspheres as *great circles*.

Note that an m-dimensional topological subsphere is in a certain sense "genuinely m-dimensional", because it must intersect every $(n-m)$-dimensional linear subsphere. Indeed, assume that $f : S^m \to S^n$ is a continuous odd function whose image is the topological subsphere. Without loss of generality we can assume that the subspace is of the form $\{x : x_1 = \ldots = x_m = 0\}$ and let P_m be the projection to the first m coordinates. It follows by the Borsuk-Ulam theorem that we can find $x \in S^m$ such that $P_m f(x) = P_m f(-x)$, and since f is odd it follows that $P_m f(x) = 0$.

As a corollary from Theorem 1.2.1 we get the following

Theorem 1.2.3. *There exist constants $\alpha, \epsilon > 0$ such that for every n there is an $\lfloor \alpha n \rfloor$-dimensional topological subsphere X of S^n such that X_ϵ does not contain a great circle.*

1.3 Constructing tennis balls

Throughout this section it will be convenient to define the "standard" Euclidean norm on $\mathbb{R}^n$ by the formula

$$|x|_2^2 = n^{-1} \sum_{i=1}^{n} x_i^2. \tag{1.1}$$

The advantage of the factor n^{-1} on the right-hand side is that a typical coordinate of a random vector of norm 1 has order of magnitude 1 rather than order of magnitude $n^{-1/2}$. This norm is often called the L_2 norm on $\mathbb{R}^n$, and we write $L_2^n = (\mathbb{R}^n, |\cdot|_2)$. It is the Euclidean norm most commonly used in additive combinatorics. Following the standard terminology in that field, we shall sometimes write the right-hand side of the formula above as $\mathbb{E}_i x_i^2$. Moreover, from this point on S^{n-1} will denote a unit sphere in L_2^n.

1.3.1 The definition of the tennis ball map

We are aiming to prove Theorem 1.2.1, or in other words to prove that there exists a continuous map (in fact it will be bi-Lipschitz) $\psi : S^{n-1} \to S^{n-1}$ that preserves antipodal points, with the property that if X is a random $\lfloor \alpha n \rfloor$-dimensional subsphere of S^{n-1}, then with high probability $\psi(X)_\epsilon$ contains no linear subsphere of dimension 1. We shall achieve this by identifying a set $\Gamma \subset S^{n-1}$ such that with high probability $\psi(X \cap S^{n-1}) \subset \Gamma$, or equivalently $X \cap S^{n-1} \subset \psi^{-1}(\Gamma)$, and such that every great circle contains a point that does not belong to Γ_ϵ.

These properties are clearly in tension with each other: we need Γ to have small measure, or else its expansion Γ_ϵ will contain a great circle, but on the other hand we also need $\psi^{-1}(\Gamma)$

to have measure very close to 1, or else it will not contain almost all $\lfloor \alpha n \rfloor$-dimensional linear subspheres.

In order to resolve this tension, we define a map that takes "typical" vectors to highly "atypical" vectors. Let k be a large positive integer to be chosen later, let $\lambda > 1$, and define $s = \lambda^{1/2k}$. (The parameter λ will later be chosen to be 4, but we write most of the arguments in a slightly greater generality in order to emphasize certain flexibility in our construction and make the role of this parameter more explicit.)

Definition 1.3.1. *Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a strictly increasing continuous odd function such that for every integer m we have*

$$\varphi(s^{2mk-1}) = s^{2mk+1} \quad \text{and} \quad \varphi(s^{2mk+1}) = s^{2(m+1)k-1}.$$

Moreover, assume that $\varphi(s^{2k}x) = s^{2k}\varphi(x)$ holds for every real number x.

Abusing notation, we shall also, for a vector $x = (x_1, \ldots, x_n)$ in $\mathbb{R}^n$, write $\varphi(x)$ to denote the vector $(\varphi(x_1), \ldots, \varphi(x_n))$. Finally, the tennis ball map is a function $\psi : \mathbb{R}^n \to \mathbb{R}^n$, which is a normalized version of φ given by the formula

$$\psi(x) = \frac{\varphi(x)}{|\varphi(x)|}|x|.$$

Note that the property $\varphi(s^{2k}x) = s^{2k}\varphi(x)$ is there for cosmetic reasons only. As for the first, it tells us that the graph of φ has a kind of staircase shape with steps of sizes that grow exponentially (see Figure 1.1). Indeed, as x increases from s^{2mk-1} to s^{2mk+1}, which is only a small change proportionately speaking, $\varphi(x)$ increases from s^{2mk+1} to $s^{2(m+1)k-1}$, which is an increase by a factor of almost λ. Similarly, as x increases from s^{2mk+1} to $s^{2(m+1)k-1}$, which is about λ times as big, $\varphi(x)$ increases from $s^{2(m+1)k-1}$ to $s^{2(m+1)k+1}$, which is only a small increase.

In particular, writing A_m for the "wide" interval and B_m for the "narrow" interval, i.e.

$$A_m = [s^{2mk+1}, s^{2(m+1)k-1}] \quad \text{and} \quad B_m = [s^{2mk-1}, s^{2mk+1}] \tag{1.2}$$

we have that

$$\varphi(A_m) = B_{m+1}$$

for every m. Moreover, if by A we denote the union $\bigcup_m (A_m \cup (-A_m))$ and by B the union $\bigcup_m (B_m \cup (-B_m))$, then we also have that $\varphi(A) = B$.

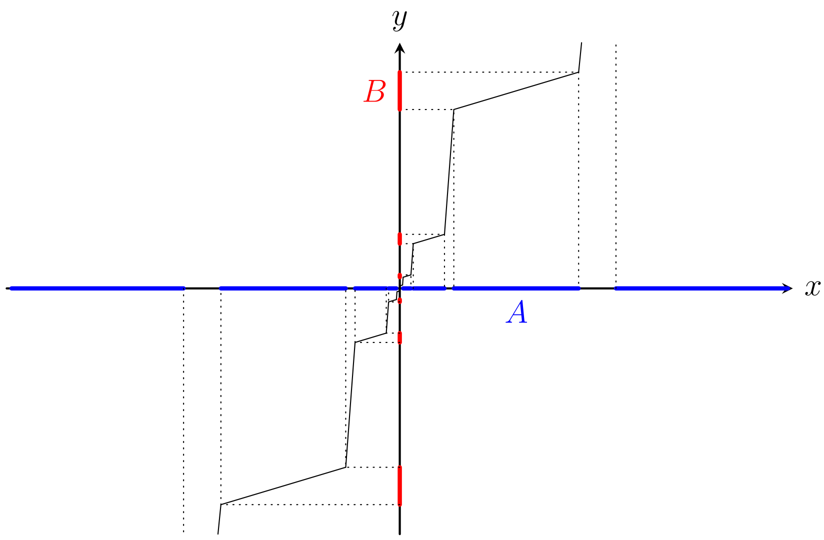

Fig. 1.1 An example of what the function φ may look like.

Note that if $x, y \in B$, then xy^{-1} belongs to an interval of the form $[s^{2mk-2}, s^{2mk+2}]$, or minus such an interval. In other words, we ensure that B is a "geometric-progression-like" set in the sense that BB^{-1} is not much larger than B itself. It follows that B^2B^{-2} consists of all points that belong to an interval of the form $[s^{2mk-4}, s^{2mk+4}]$. This fact will be useful later on. Another observation we shall use is that

$$|x| \le |\varphi(x)| \le \lambda|x| \tag{1.3}$$

for every x. This implies that ψ is a Lipschitz function. Similarly, we observe that the inverse of ϕ satisfies $\frac{1}{\lambda}|x| \le |\varphi^{-1}(x)| \le |x|$, and hence we get that ψ^{-1} is Lipschitz as well. This shows that the tennis ball map ψ is indeed bi-Lipschitz as claimed.

In what follows we shall show, that the tennis ball map ψ takes random linear subspheres of appropriate dimension to tennis balls.

1.3.2 The definition of the set Γ

Now that we have defined the tennis ball map, let us discuss the set Γ. The rough idea is that Γ is the set of $x \in S^{n-1}$ with almost all their coordinates in B. Since A has a small complement, one would expect almost all coordinates of a random vector to belong to A, and indeed this is the case. It forms the basis of a probabilistic argument that shows that with high probability every $x \in S^{n-1} \cap X$ has the property that almost every coordinate of

x belongs to A, which implies that almost every coordinate of $\varphi(x)$ belongs to B. Thus a "typical" vector (one with almost all coordinates in the large set A) is mapped to a highly "atypical" vector (one with almost all coordinates in the small set B).

This is a slight oversimplification, because of the normalization that replaces φ by ψ. The actual definition of Γ concerns the *ratios* of the coordinates rather than their actual values.

Now let us give some more details. For the purposes of this problem, it is more natural, when talking about a unit vector x, to attach a weight of x_i^2 to the ith coordinate. For example, the statement "almost every ratio $x_i x_j^{-1}$ belongs to BB^{-1}" should be interpreted as meaning that

$$\sum_{i,j} x_i^2 x_j^2 \, \mathbb{1}_{\{x_i x_j^{-1} \in BB^{-1}\}}(i,j) \geq (1-\epsilon) \sum_{i,j} x_i^2 x_j^2$$

for some small ϵ, and similarly for other statements about coordinates. More precisely, with every vector x we define an associated probability measure.

Definition 1.3.2. *For any $x \in \mathbb{R}^n \setminus \{0\}$ define its associated probability measure μ_x on $\{1,2,\ldots,n\}$ by the formula*

$$\mu_x(J) = \frac{|P_J x|^2}{|x|^2},$$

where P_J is the coordinate projection to the set J.

Let us now fix some useful notation. For a function $f : \{1,2,\ldots,n\} \to \mathbb{R}$ we write

$$\mathbb{E}_i^x f(i) = |x|^{-2} \sum_i x_i^2 f(i) \qquad \text{and} \qquad \mathbb{E}_{i,j}^x f(i,j) = |x|^{-4} \sum_{i,j} x_i^2 x_j^2 f(i,j),$$

while $\mathbb{E}_i$ or $\mathbb{E}_\theta$, as mentioned before, will be the usual average.

Then, given a property Q of integers in the set $\{1,2,\ldots,n\}$ we define

$$\mathbb{P}_i^x[Q(i)] = \mathbb{E}_i^x \mathbb{1}_Q(i) = \mu_x\{i : Q(i)\}.$$

Similarly, we shall write

$$\mathbb{P}_{i,j}^x[Q(i,j)] = \mathbb{E}_{i,j}^x \mathbb{1}_Q(i,j) = (\mu_x \times \mu_x)\{(i,j) : Q(i,j)\},$$

where $\mu_x \times \mu_x$ is the product measure.

We are now ready to define the set Γ.

Definition 1.3.3. *Let* $\Gamma \subset S^{n-1}$ *be a set given by*

$$\Gamma = \{y \in S^{n-1} : \ \mathbb{P}^y_{i,j}[y_i/y_j \in BB^{-1}] \geq 1 - 2\lambda^2\beta\}, \tag{1.4}$$

for a suitable choice of the parameter β*.*

Of no less importance will be a set Δ defined by

$$\Delta = \{x \in S^{n-1} : \ \mathbb{P}^x_i[x_i \in A] < 1 - \beta\}. \tag{1.5}$$

In this section we will be interested in its complement

$$\Delta^c = S^{n-1} \setminus \Delta = \{x \in S^{n-1} : \ \mathbb{P}^x_i[x_i \in A] \geq 1 - \beta\}$$

and this choice of notation was made for simplicity in further sections.

Proposition 1.3.4. *The image of the set* Δ^c *under the tennis ball map* ψ *is a subset of* Γ*, i.e. we have* $\psi(\Delta^c) \subset \Gamma$*.*

Proof. Let $x \in \Delta^c$ and note that if $\mathbb{P}^x_i[x_i \in A] \geq 1 - \beta$, then $\mathbb{P}^x_i[\varphi(x_i) \in B] \geq 1 - \beta$ by the definition of the function φ. Moreover, from (1.3) it follows that the function has the property that $\varphi(x)^2$ always lies between $|x|^2$ and $\lambda^2|x|^2$ and hence we have that $\mu_{\varphi(x)}(E) \leq \lambda^2\mu_x(E)$ for every vector $x \in \mathbb{R}^n$ and every set $E \subset \{1, 2, \ldots, n\}$, from which it follows (considering complements) that

$$\mathbb{P}^{\varphi(x)}_i[\varphi(x_i) \in B] > 1 - \lambda^2\beta.$$

Since this is true for every $x \in \Delta^c$ we get that

$$\varphi(\Delta^c) \subset \{y : \mathbb{P}^y_i[y_i \in B] > 1 - \lambda^2\beta\}. \tag{1.6}$$

In order to show the inclusion for ψ, which is essentially a normalised function φ, note that if we have y as above then

$$\mathbb{P}^y_{i,j}[y_i \in B \text{ and } y_j \in B] \geq 1 - 2\lambda^2\beta,$$

which in turn gives us

$$\mathbb{P}^y_{i,j}[y_i/y_j \in BB^{-1}] \geq 1 - 2\lambda^2\beta.$$

From this follows that such y belongs to Γ and since Γ is invariant under positive scalar multiples, using (1.6) we conclude that $\psi(\Delta^c) \subset \Gamma$. $\square$

In the next section we shall prove that the expansion Γ_ϵ contains no 1-dimensional linear subsphere, or great circle, and in the following one we shall prove that for suitable constant $\alpha > 0$, a random linear subsphere of dimension $\lfloor \alpha n \rfloor$ will be contained in Δ^c with high probability.

1.4 Proving that no great circle is contained in Γ_ϵ

We begin with a couple of lemmas that help us to describe the set Γ_ϵ. Recall that $|\cdot|_2$ is the L_2 norm defined in (1.1).

Lemma 1.4.1. *Let y, z be unit vectors in L_2^n. Assume that $|y - z|_2 \leq \epsilon$, and let E be a subset of $\{1, 2, \ldots, n\}$. Then $\left|\mathbb{P}_i^y[E] - \mathbb{P}_i^z[E]\right| \leq 2\epsilon$.*

Proof. The left-hand side is equal to $\frac{1}{n}||P_E y|^2 - |P_E z|^2| = |\mathbb{E}_i(y_i^2 - z_i^2)\mathbb{1}_E(i)| \leq \mathbb{E}_i|y_i^2 - z_i^2|$. Then by Cauchy-Schwarz inequality it follows that

$$\begin{aligned} \mathbb{E}_i|y_i^2 - z_i^2| &= \mathbb{E}_i|y_i - z_i|\,|y_i + z_i| \\ &\leq |y - z|_2\,|y + z|_2 \\ &\leq 2\epsilon, \end{aligned}$$

which proves the result. $\square$

Recall from §1.3.1 that BB^{-1} is the union of all intervals of the form $[s^{2mk-2}, s^{2mk+2}]$, and B^2B^{-2} is the union of all intervals of the form $[s^{2mk-4}, s^{2mk+4}]$, where $s = \lambda^{1/2k}$ was one of the parameters used to define the "staircase function" φ. It follows that if $t \in BB^{-1}$ and $u \notin B^2B^{-2}$, then $|t/u| \notin [s^{-2}, s^2]$, which implies in particular that $|t/u - 1| \geq 1 - s^{-2} =: \tau$.

Lemma 1.4.2. *Let $\tau = 1 - s^{-2}$ and $\epsilon < \tau^2$. Then for $z \in \Gamma_\epsilon$ we have*

$$\mathbb{P}_{i,j}^z[z_i/z_j \in B^2B^{-2}] \geq 1 - 2\lambda^2\beta - 6\epsilon.$$

Proof. Let $y \in \Gamma$ with $|y|_2 = 1$ be such that $|y - z|_2 \leq \epsilon$. Then $\mathbb{P}_{i,j}^y[y_i/y_j \in BB^{-1}] \geq 1 - 2\lambda^2\beta$, or equivalently

$$\mathbb{E}_{i,j} y_i^2 y_j^2 \mathbb{1}_{[y_i/y_j \in BB^{-1}]} \geq 1 - 2\lambda^2\beta.$$

By Lemma 1.4.1 (and recalling that $\mathbb{E}_i y_i^2 = 1$), it follows that

$$\mathbb{E}_{i,j}(y_i^2 - z_i^2)y_j^2 \mathbb{1}_{[y_i/y_j \in BB^{-1}]} \leq 2\epsilon$$

and hence combining both inequalities gives

$$\mathbb{E}_{i,j} z_i^2 y_j^2 \mathbb{1}_{[y_i/y_j \in BB^{-1}]} \geq 1 - 2\lambda^2\beta - 2\epsilon. \tag{1.7}$$

If the conclusion that $\mathbb{P}_{i,j}^z[z_i/z_j \in B^2B^{-2}] \geq 1 - 2\lambda^2\beta - 6\epsilon$ is not true, then we would have

$$\mathbb{E}_{i,j} z_i^2 z_j^2 \mathbb{1}_{[z_i/z_j \notin B^2B^{-2}]} \geq 2\lambda^2\beta + 6\epsilon,$$

and by again Lemma 1.4.1 in the same way, it follows that

$$\mathbb{E}_{i,j} z_i^2 y_j^2 \mathbb{1}_{[z_i/z_j \notin B^2B^{-2}]} \geq 2\lambda^2\beta + 4\epsilon. \tag{1.8}$$

By summing (1.7) and (1.8), and estimating trivially the probability of the union of the events by 1, we deduce that

$$\mathbb{E}_{i,j} z_i^2 y_j^2 \mathbb{1}_{[y_i/y_j \in BB^{-1} \text{ and } z_i/z_j \notin B^2B^{-2}]} \geq 2\epsilon.$$

As remarked before the lemma, if $y_i/y_j \in BB^{-1}$ and $z_i/z_j \notin B^2B^{-2}$, then $|\frac{y_i z_j}{y_j z_i} - 1| > \tau$. It follows that

$$\mathbb{E}_{i,j} z_i^2 y_j^2 \Big(\frac{y_i z_j}{y_j z_i} - 1\Big)^2 > 2\tau^2\epsilon.$$

But since $\mathbb{E}_i y_i^2 = \mathbb{E}_i z_i^2 = 1$ we have

$$\begin{aligned}\mathbb{E}_{i,j} z_i^2 y_j^2 \Big(\frac{y_i z_j}{y_j z_i} - 1\Big)^2 &= \mathbb{E}_{i,j}(y_i z_j - z_i y_j)^2 \\ &= \mathbb{E}_{i,j}(y_i^2 z_j^2 + y_j^2 z_i^2 - 2y_i y_j z_i z_j) \\ &= 2 - 2\langle y, z\rangle^2.\end{aligned}$$

Furthermore, if $2 - 2\langle y, z\rangle^2 > 2\tau^2\epsilon$, then $\langle y, z\rangle^2 < 1 - \tau^2\epsilon$, which implies that $\langle y, z\rangle < 1 - \tau^2\epsilon/2$, which in turn means that

$$|y - z|_2^2 = 2 - 2\langle y, z\rangle > \tau^2\epsilon,$$

and therefore that $|y - z|_2 > \tau\sqrt{\epsilon}$. However, since by assumption we have $\tau > \sqrt{\epsilon}$, this is a contradiction. $\square$

Corollary 1.4.3. *It follows directly from Lemma 1.4.2 that*

$$\Gamma_\epsilon \subset \{z : \mathbb{P}^z_{i,j}[z_i/z_j \in B^2 B^{-2}] \geq 1 - 2\lambda^2\beta - 6\epsilon\}.$$

This bigger set resembles Γ but is defined using slightly different parameters. We now turn to the proof that every great circle contains a point that does not belong to this slightly expanded Γ-like set.

1.4.1 Analysis of the points in a 2-dimensional subspace

Let Y be a 2-dimensional subspace of L_2^n and let $\{u, v\}$ be an orthonormal basis for Y. Then the unit sphere of Y consists of vectors $u\cos\theta + v\sin\theta$. The ith coordinate of such a vector, $u_i\cos\theta + v_i\sin\theta$ can be rewritten as $a_i\sin(\theta+\phi_i)$, where $a_i = \sqrt{u_i^2+v_i^2}$ and ϕ_i is chosen such that $a_i\sin\phi_i = u_i$ and $a_i\cos\phi_i = v_i$. By a denote the vector $(a_1,\dots,a_n)$ and note that $|a|_2 = 2$.

We would now like to prove that there are plenty of pairs (i,j) such that ϕ_i is not close to ϕ_j or $-\phi_j$.

Lemma 1.4.4. *With $a_1,\dots,a_n$ and $\phi_1,\dots,\phi_n$ as above, we have the inequality*

$$\mathbb{P}^a_{i,j}[\cos(2(\phi_i-\phi_j)) \leq 1/2] \geq 1/3.$$

Proof. Since u, v are fixed unit vectors (in L_2^n) we have that $|a|^2 = \sum_i(u_i^2+v_i^2) = 2n$ and hence $\mathbb{E}^a_i \sin^2(\theta+\phi_i) = |a|^{-2}\sum_i a_i^2\sin^2(\theta+\phi_i) = \frac{1}{2}$ for every θ. Therefore, we find on differentiating with respect to θ that

$$2\mathbb{E}^a_i \sin(\theta+\phi_i)\cos(\theta+\phi_i) = \mathbb{E}^a_i \sin(2\theta+2\phi_i) = 0$$

for every θ, and hence, on differentiating again, that

$$\mathbb{E}^a_i \cos(2\theta+2\phi_i) = 0$$

for every θ as well.

From that it follows that

$$\mathbb{E}^a_{i,j}\Big(\cos(2\theta+2\phi_i)\cos(2\theta+2\phi_j) + \sin(2\theta+2\phi_i)\sin(2\theta+2\phi_j)\Big) = \mathbb{E}^a_{i,j}\cos(2(\phi_i-\phi_j)) = 0.$$

Let F be the event that $\cos(2(\phi_i - \phi_j)) \leq 1/2$. We have seen that $\mathbb{E}^a_{i,j} \cos(2(\phi_i - \phi_j)) = 0$, and we also know that $\cos(2(\phi_i - \phi_j)) \in [-1, 1]$. So, using the total probability formula we get

$$0 = \mathbb{E}^a_{i,j} \cos(2(\phi_i - \phi_j)) \geq \frac{1}{2}\mathbb{P}^a_{i,j}[F^c] - \mathbb{P}^a_{i,j}[F] = \frac{1}{2} - \frac{3}{2}\mathbb{P}^a_{i,j}[F],$$

from which the desired inequality follows. □

Next, we need a technical lemma that will help us to show that if ϕ_i is not approximately $\pm\phi_j$, then $\sin(\theta + \phi_i)/\sin(\theta + \phi_j)$ is not often close to an element of some given geometric progression.

Lemma 1.4.5. *Let θ be chosen randomly from $[-\pi, \pi]$ and let $0 < a < b$. Then*

$$\mathbb{P}\left[a \leq \cot\theta \leq b\right] \leq \frac{b-a}{\pi(1+a^2)},$$

and the same bound holds for the probability that $\cot\theta \in [-b, -a]$.

Proof. Since cot is periodic with period π and is decreasing in the interval $(0, \pi)$, the probability in question is $(\cot^{-1} a - \cot^{-1} b)/\pi$. By the mean value theorem, $\cot^{-1} a - \cot^{-1} b$ is at most $|a - b|$ times the absolute value of the derivative of $\cot^{-1}$ at a. Since that derivative is $-1/(1 + a^2)$, the first result follows. The second then holds by symmetry. □

Recall once again that B^2B^{-2} is the set of all real numbers x such that

$$|x| \in [\lambda^m(1-\tau)^2, \lambda^m(1-\tau)^{-2}]$$

for some positive integer m and $1 - \tau = s^{-2}$.

The main point about the bound in the next lemma is not its exact form, but simply that it is $O(\xi)$ except when ϕ_i is close to ϕ_j or $\phi_j + \pi$.

Lemma 1.4.6. *Let $\xi = (1-\tau)^{-4} - 1$ and let $\theta \in [0, 2\pi]$ be chosen uniformly at random. Then*

$$\mathbb{P}\left[\frac{a_i \sin(\theta + \phi_i)}{a_j \sin(\theta + \phi_j)} \in B^2B^{-2}\right] \leq \frac{4\xi\lambda}{\pi(\lambda - 1)}(4 + |\cot(\phi_i - \phi_j)|).$$

Proof. The distribution of $\frac{a_i \sin(\theta+\phi_i)}{a_j \sin(\theta+\phi_j)}$ is the same as the distribution of

$$\frac{a_i \sin(\theta + \phi_i - \phi_j)}{a_j \sin(\theta)} = \frac{a_i}{a_j}\Big(\cos(\phi_i - \phi_j) + \sin(\phi_i - \phi_j)\cot\theta\Big).$$

Therefore, we are interested in the probability that $\cot\theta \in \frac{a_j}{a_i \sin(\phi_i - \phi_j)} B^2B^{-2} - \cot(\phi_i - \phi_j)$.

Let $t = |\frac{a_j(1-\tau)^2}{a_i \sin(\phi_i - \phi_j)}|$. Then

$$\frac{a_j}{a_i \sin(\phi_i - \phi_j)} B^2 B^{-2} = \bigcup_m \Big([t\lambda^m, t\lambda^m(1-\tau)^{-4}] \cup [-t\lambda^m(1-\tau)^{-4}, -t\lambda^m] \Big).$$

By Lemma 1.4.5, we get the bound

$$\mathbb{P}\Big[\cot\theta \in \Big[t\lambda^m - \cot(\phi_i - \phi_j), t\lambda^m(1-\tau)^{-4} - \cot(\phi_i - \phi_j)\Big]\Big] \leq \frac{\xi t\lambda^m}{\pi}$$

for all m, and in addition if $t\lambda^m \geq 2\cot(\phi_i - \phi_j)$, then since $1 - \tau < 1$ we have an upper bound of

$$\frac{\xi t\lambda^m}{\pi(1 + t^2\lambda^{2m}/4)} \leq \frac{4\xi}{\pi t\lambda^m}.$$

If $\cot(\phi_i - \phi_j) \geq 0$, then the probability that $\cot\theta + \cot(\phi_i - \phi_j)$ lies in the positive part of $\frac{a_j}{a_i \sin(\phi_i - \phi_j)} B^2 B^{-2}$ is therefore at most ξ/π multiplied by the sum

$$\sum_{t\lambda^m \leq S} t\lambda^m + 4 \sum_{t\lambda^m > S} \frac{1}{t\lambda^m},$$

where $S = \max\{2\cot(\phi_i - \phi_j), 1\}$. By the formula for the sum of a geometric progression, and recalling that $\lambda \geq \frac{3}{2}$, the first sum can be estimated by

$$\sum_{t\lambda^m \leq S} t\lambda^m = t \sum_{m=-\infty}^{m_0} \left(\frac{1}{\lambda}\right)^m = t\frac{(1/\lambda)^{-m_0}}{1 - \lambda^{-1}} \leq S\frac{\lambda}{\lambda - 1},$$

and similarly the second sum is at most $S^{-1}\frac{\lambda}{\lambda-1}$. Therefore, the total is at most

$$\frac{(S + 4S^{-1})\lambda}{\lambda - 1} \leq \frac{(5 + 2\cot(\phi_i - \phi_j))\lambda}{\lambda - 1}.$$

Therefore, we obtain an answer of at most $\xi\lambda(5 + 2\cot(\phi_i - \phi_j))/\pi(\lambda - 1)$.

If $\cot(\phi_i - \phi_j) < 0$, then in the same way we get that the probability that $\cot\theta \in \Big[t\lambda^m - \cot(\phi_i - \phi_j), t\lambda^m(1-\tau)^{-4} - \cot(\phi_i - \phi_j)\Big]$ is at most $\xi t\lambda^m/\pi$ for all m but now it is also at most $\xi/\pi t\lambda^m$ for all m. Using the first bound when $t\lambda^m \leq 1$ and the second when $t\lambda^m > 1$, we obtain an upper bound of at most

$$\frac{2\xi\lambda}{\pi(\lambda - 1)}.$$

Considering the negative part of B as well and combining these two estimates, we obtain the result stated. □

Corollary 1.4.7. *If $\tau \leq 10^{-4}$ and $\lambda \geq 3/2$, then in every 2-dimensional subspace of L_2^n there is a vector y such that*

$$\mathbb{P}^y_{i,j}[y_i/y_j \notin B^2B^{-2}] \geq 1/8.$$

Proof. Let a typical unit vector y in the subspace have ith coordinate $y_i = a_i \sin{(\theta + \phi_i)}$. We will bound the desired probability from below by adding an additional constraint. We will consider the probability that $\{\frac{y_i}{y_j} \notin B^2B^{-2}$ and $|\cot(\phi_i - \phi_j)| \leq 2\}$, which can be found by calculating the expected probability that $\{|\cot(\phi_i - \phi_j)| \leq 2\}$ and subtracting from it $\mathbb{P}^y_{i,j}\left[\frac{y_i}{y_j} \in B^2B^{-2} \text{ and } |\cot(\phi_i - \phi_j)| \leq 2\right]$. This is exactly what follows.

We have for each i, j that

$$\begin{aligned}
\mathbb{E}_\theta \sin^2(\theta + \phi_i)\sin^2(\theta + \phi_j) &= \frac{1}{4}\mathbb{E}_\theta\Big(\cos(\phi_i - \phi_j) - \cos(2\theta + \phi_i + \phi_j)\Big)^2 \\
&= \frac{1}{4}\Big(\cos^2(\phi_i - \phi_j) + \mathbb{E}_\theta \cos^2(2\theta + \phi_i + \phi_j)\Big) \\
&= \frac{1}{4}\Big(\cos^2(\phi_i - \phi_j) + \frac{1}{2}\Big) \geq \frac{1}{8}.
\end{aligned}$$

Here $\mathbb{E}_\theta$ is just the usual average. Recall that y is such that $|y| = 1$ and $y_i = a_i \sin(\theta + \phi_i)$ for some $\theta \in [0, 2\pi]$. For any event Q that depends on two coordinates i, j, we get

$$\begin{aligned}
\mathbb{E}_y\mathbb{P}^y_{i,j}[Q] &= \mathbb{E}_\theta\mathbb{E}_{i,j}a_i^2a_j^2\sin^2(\theta + \phi_i)\sin^2(\theta + \phi_j)\mathbb{1}_Q(i,j) \\
&\geq \frac{1}{8}\,\mathbb{E}_{i,j}a_i^2a_j^2\mathbb{1}_Q(i,j) \\
&= \frac{|a|^4}{8}\mathbb{E}^a_{i,j}\mathbb{1}_Q(i,j) \\
&= \frac{1}{2}\mathbb{P}^a_{i,j}[Q],
\end{aligned}$$

where we used that $|a|^2 = 2$.

It is easy to check the identity $\cot^2\alpha = \frac{2\cos^2\alpha}{1-\cos(2\alpha)}$, so if $\cos(2(\phi_i - \phi_j)) \leq 1/2$, then $|\cot(\phi_i - \phi_j)| \leq 2$. Therefore,

$$\begin{aligned}
\mathbb{E}_y\mathbb{P}^y_{i,j}[|\cot(\phi_i - \phi_j)| \leq 2] &\geq \mathbb{E}_y\mathbb{P}^y_{i,j}[\cos(2(\phi_i - \phi_j)) \leq 1/2] \\
&\geq \frac{1}{2}\,\mathbb{P}^a_{i,j}[\cos(2(\phi_i - \phi_j)) \leq 1/2] \;\geq\; \frac{1}{6},
\end{aligned}$$

where the last inequality follows from Lemma 1.4.4.

Now, by Lemma 1.4.6, if y is a random such vector, then for each i, j the probability that $y_i/y_j \in B^2B^{-2}$ is at most $\frac{4\xi\lambda}{\pi(\lambda-1)}\Big(4+|\cot(\phi_i-\phi_j)|\Big) \leq 4\xi\Big(4+|\cot(\phi_i-\phi_j)|\Big)$, where we used the fact that $\lambda \geq 3/2$. Note also that since $y_i^2 \leq a_i^2$ for each i, and $\mathbb{E}_i y_i^2 = \frac{1}{2}\mathbb{E}_i a_i^2$, we have that $\mathbb{P}^y_{i,j}[Q(i,j)] \leq 4\mathbb{P}^a_{i,j}[Q(i,j)]$ for every event $Q(i,j)$ that depends on two coordinates i, j. It follows that

$$\begin{aligned}\mathbb{E}_y\mathbb{P}^y_{i,j}[|\cot(\phi_i-\phi_j)| \leq 2 \text{ and } y_i/y_j \in B^2B^{-2}] \leq&\\ &\leq 4\mathbb{E}_y\mathbb{P}^a_{i,j}[|\cot(\phi_i-\phi_j)| \leq 2 \text{ and } y_i/y_j \in B^2B^{-2}]\\ &\leq 16\xi(4+2)\\ &= 96\xi.\end{aligned}$$

Together with the estimate in the previous paragraph, this implies that

$$\mathbb{P}^y_{i,j}[|\cot(\phi_i-\phi_j)| \leq 2 \text{ and } y_i/y_j \notin B] \geq 1/6 - 96\xi.$$

Since, ξ was defined as $(1-\tau)^{-4}-1$, it is straightforward to check that if $\tau \leq 10^{-4}$, then this is at least $1/8$, and the result follows. □

Corollary 1.4.8. *Provided that* $2\lambda^2\beta + 6\epsilon < 1/8$, *every great circle contains a point that does not belong to* Γ_ϵ.

Proof. The previous Corollary 1.4.8, applied to the subspace whose unit sphere is the great circle, gives us a point y such that $\mathbb{P}^y_{i,j}[y_i/y_j \notin B^2B^{-2}] \geq 1/8$. Since the event in square brackets is invariant under positive scalar multiples, we may assume that y is a unit vector and thus that it belongs to the great circle.

We showed earlier that if $z \in \Gamma_\epsilon$, then $\mathbb{P}^z_{i,j}[z_i/z_j \in B^2B^{-2}] \geq 1 - 2\lambda^2\beta - 6\epsilon$. Hence, if $2\lambda^2\beta + 6\epsilon < 1/8$, this implies that $y \notin \Gamma_\epsilon$, and we are done. □

1.5 Almost every point has an "atypical" image

In this section we want to show that there exists a subspace X of linear dimension such that $X \cap S^{n-1} \subset \Delta^c = \{x \in S^{n-1} : \mathbb{P}^x_i[x_i \in A] \geq 1-\beta\}$, so that $\psi(X \cap S^{n-1}) \subset \psi(\Delta^c) \subset \Gamma$, and indeed that for an appropriate constant $\alpha > 0$, almost all subspaces of dimension at most αn have this property. To this end, it will be sufficient to show that Δ has exponentially small measure. Note that

$$\mathbb{P}[x \in \Delta] = \mathbb{P}[\mathbb{P}_i^x[x_i \in A] < 1 - \beta] = \mathbb{P}[\mathbb{P}_i^x[x_i \in B] \geq \beta],$$

where B is, as before, the set

$$\bigcup_m (B_m \cup (-B_m)),$$

and $B_m = [s^{2mk-1}, s^{2mk+1}]$ for each integer m. Let $\eta = s - 1 > 0$ and as before let $\lambda = s^{2k}$. Then $B_m = [(1+\eta)^{-1}\lambda^m, (1+\eta)\lambda^m]$.

Let us say that a positive real number t is an *η-approximate power of* λ if there exists an integer m such that

$$(1+\eta)^{-1}\lambda^m \leq t \leq (1+\eta)\lambda^m.$$

For $\gamma \in [0,1]$ and $\xi \geq 0$ define Δ_γ^ξ by

$$\Delta_\gamma^\xi = \Big\{x \in \mathbb{R}^n : \mathbb{P}_i^x[|x_i| \text{ is a } \xi\text{-approximate power of } \lambda] \geq \gamma\Big\}. \tag{1.9}$$

As mentioned before, we shall end up taking $\lambda = 4$. For this reason, although Δ_γ^ξ depends on λ, we suppress this dependence in the notation. We shall be particularly interested in the set Δ_β^η, which restricted to S^{n-1} is equal to the set Δ defined in (1.5).

However, we shall also be interested in the set Δ_β^0, which we shall write simply as Δ_β. That is,

$$\Delta_\beta = \Big\{x \in \mathbb{R}^n : \mathbb{P}_i^x[|x_i| \text{ is a power of } \lambda] \geq \beta\Big\}. \tag{1.10}$$

Lemma 1.5.1. *If $y \in \Delta_\beta^\eta$ then there exists $x \in \Delta_{\beta(1-2\eta)}$ such that $|x-y| \leq \eta|y|$.*

Proof. We are given that $\mathbb{P}_i^y[|y_i|$ is an η-approximate power of $\lambda] \geq \beta$. Let J be the set of all i such that $|y_i|$ is an η-approximate power of λ. For each $i \in J$ let $|x_i|$ be the nearest power of λ to $|y_i|$ and let x_i have the same sign as y_i. For each $i \notin J$ let $x_i = y_i$. Then $|x_i - y_i| \leq \eta|y_i|$ for $i \in J$, so, writing P_J for the coordinate projection to J, we have that

$$|x-y|_2^2 = \frac{1}{n}\sum_{i\in J}|x_i - y_i|^2 \leq \eta^2 \frac{1}{n}\sum_{i \in J}|y_i|^2 = \eta^2|P_J y|_2^2 \leq \eta^2|y|_2^2.$$

We now need a lower bound for $|P_J x|^2/|x|^2$. We know that $|P_J y|^2 \geq \beta|y|^2$, and also that $|P_J x|^2 - |P_J y|^2 = |x|^2 - |y|^2$. We also have for each $i \in J$ that $(1+\eta)^{-2}y_i^2 \leq x_i^2 \leq (1+\eta)^2 y_i^2$, which implies that $(1+\eta)^{-2}|P_J y|^2 \le |P_J x|^2 \le (1+\eta)^2|P_J y|^2$. Therefore,

$$\frac{|P_J x|^2}{|x|^2} = \frac{\zeta|P_J y|^2}{|y - P_J y|^2 + \zeta|P_J y|^2}$$

for some $\zeta \in [(1+\eta)^{-2}, (1+\eta)^2]$. The right-hand side is minimized when $\zeta = (1+\eta)^{-2}$, and then it is at least $\zeta\beta \ge (1-2\eta)\beta$, which finishes the proof of the lemma. □

Our next aim is to prove an upper bound for the volume of the η-expansion of $\Delta_{\beta(1-2\eta)}$, which by the above lemma contains Δ_β^η. We shall do this in a series of simple steps.

Lemma 1.5.2. *For every constant $C > 1$, the number of sequences $(a_1, \dots, a_m)$ of positive integers that add up to at most Cm is at most $(Ce)^m$.*

Proof. For each sequence $a = (a_1, \dots, a_m)$ let C_a be the unit cube of points $x \in \mathbb{R}^n$ such that $a_i - 1 \le x_i < a_i$ for every i. Then if $a \ne b$, the unit cubes C_a and C_b are disjoint. Also, if a consists of positive integers and $\sum_i a_i \le Cm$, then the cube C_a is contained in the convex hull of the points 0 and Cme_i, where $e_1, \dots, e_m$ is the standard basis of $\mathbb{R}^m$. But this simplex has volume $(Cm)^m/m! \le (Ce)^m$. This proves the result. □

Corollary 1.5.3. *Let $\lambda, C > 1$ be real numbers and let m be a positive integer. Then the number of positive integer sequences $(a_1, \dots, a_m)$ such that $\lambda^{a_1} + \dots + \lambda^{a_m} \le Cm$ is at most $(e \log_\lambda C)^m$.*

Proof. If $\lambda^{a_1} + \dots + \lambda^{a_m} \le \lambda^a m$, then by Jensen's inequality $\lambda^{(a_1+\dots+a_m)/m} \le \lambda^a$, and hence $a_1 + \dots + a_m \le am$. Therefore, by Lemma 1.5.2 the number of such sequences is at most $(ea)^m$. By the hypothesis we have $a = \log_\lambda C$, and the result follows. □

Corollary 1.5.4. *Let $\lambda > 1$ be a real number, let m be a positive integer, and let $\eta > 0$. Let Ω be the set of all sequences $(x_1, \dots, x_m)$ such that $x_1^2 + \dots + x_m^2 \le C^2 m$ and each $|x_i|$ is a power of λ. Then there is an η-net of Ω of cardinality at most $(2e\log_\lambda(\lambda^2 C/\eta))^m$.*

Proof. Let $x \in \Omega$. For each i such that $|x_i| \le \eta/\lambda$, replace x_i by $\lambda^{-t}\,\mathrm{sign}(x_i)$, where t is chosen in such a way that $\eta/\lambda \le \lambda^{-t} < \eta$, and let the resulting vector be y. Then $|x_i - y_i| \le \eta$ for every i, so $|x - y|_2 \le \eta$. Now let Ω' consist of all vectors $x \in \Omega$ such that each $|x_i|$ is equal to λ^{a_i} for some integer a_i with $a_i \ge -t$. We have just shown that Ω' is an η-net of Ω.

The number of points in Ω' with positive coordinates is the number of integer sequences $(a_1, \dots, a_m)$ such that each a_i is at least $-t$ and $\lambda^{2a_1} + \dots + \lambda^{2a_m} \le C^2 m$. Rescaling, we see that is equal to the number of positive-integer sequences $(a_1, \dots, a_m)$ such that

$\lambda^{2a_1} + \cdots + \lambda^{2a_m} \leq \lambda^{2(t+1)}C^2 m$, which by Corollary 1.5.3 is at most $(e(t+1+\log_\lambda C))^m$. Since there are 2^m possible choices of signs, the size of Ω' is at most $(2e(t+1+\log_\lambda C))^m$. Noting that $t \leq \log_\lambda(\lambda/\eta) = 1 + \log_\lambda(1/\eta)$, we obtain the result. □

The important thing about the bound above is that the number we raise to the power m depends logarithmically on η. This shows that an η-net of Ω is *much* smaller than an η-net of the full sphere of radius C.

We shall need a lemma concerning the sizes of nets of unit balls. It is standard, but the version we give is less commonly used, so for convenience we include a proof. (The argument is essentially due to Rogers [17].)

Lemma 1.5.5. *Let X be an n-dimensional normed space with unit ball B_X and let $\delta > 0$. If n is sufficiently large, then X contains a δ-net of B_X of cardinality at most $2en\log(n)(1+\frac{1}{\delta})^n$.*

Proof. Let $\rho > 0$ be a small real number to be chosen later. (It will in fact depend on n.) Then a standard volume estimate shows that there is an ρ-net of B_X of size at most $(3/\rho)^n$. We shall now cover every point of this net with a union of balls of radius $\delta - \rho$ in order to obtain our δ-net, and then we will optimize over ρ.

To do this, let $\zeta = \delta - \rho$ and pick points $x_1, \ldots, x_N$ uniformly at random from $(1+\zeta)B_X$. If y is a point in the ρ-net, then the probability that y is not within any of the balls of radius ζ about the x_i is $(1 - (\frac{\zeta}{1+\zeta})^n)^N \leq \exp(-N(\frac{\zeta}{1+\zeta})^n)$. Therefore, we are done as long as

$$\left(\frac{3}{\rho}\right)^n \exp\left(-N\left(\frac{\zeta}{1+\zeta}\right)^n\right) < 1,$$

which is satisfied if $N > n\log(\frac{3}{\rho})(1+\frac{1}{\delta-\rho})^n$.

It can be checked that $1 + \frac{1}{\delta-\rho} = (1+\frac{1}{n})(1+\frac{1}{\delta})$ when $\rho = \delta(\frac{\delta+1}{n+\delta+1})$. For this value of ρ and for n is sufficiently large, we have that $\log(\frac{3}{\rho}) < \log(\frac{3n}{\delta})$, which is at most $\frac{3}{2}\log n$. We also have that $(1+\frac{1}{n})^n < \frac{4e}{3}$ when n is sufficiently large, and putting these estimates together we find that we can take N to be $2en\log n(1+\frac{1}{\delta})^n$, as claimed. □

Next, we need a simple technical lemma about the largest proportion of the unit sphere of L_2^n that can be covered by a ball of radius δ.

Lemma 1.5.6. *Let $B_\delta(x)$ be a closed ball of radius δ about a point x in L_2^n. If n is sufficiently large, then the probability that a random point of the unit sphere of L_2^n lies in $B_\delta(x)$ is at most $2\delta^n$.*

Proof. The intersection of $B_\delta(x)$ with the unit sphere is a spherical cap, and the measure of the spherical cap is maximized when the centre x of $B_\delta(x)$ is a vector of norm $\sqrt{1-\delta^2}$.

Define C to be the set of all y such that $|y|_2 \leq 1$ and $\left|x - \frac{y}{|y|_2}\right|_2 \leq \delta$. This is a convex hull of the spherical cap and the origin, and the proportion of its volume to the volume of the entire unit ball, is equal to the probability we are trying to estimate.

We are going to show that $B_\delta(x)$ contains the set $C \setminus (1-2\delta^2)C$. Indeed, we claim now that $B_\delta(x)$ contains all points y such that $\left|x - \frac{y}{|y|_2}\right|_2 \leq \delta$ and $1 \geq |y|_2 \geq 1 - 2\delta^2$. By the convexity of $B_\delta(x)$ it is sufficient to prove this when $|y|_2 = 1 - 2\delta^2$. The first assumption on y implies that

$$|x|^2 - \frac{2\langle x, y\rangle}{|y|_2} + n \leq \delta^2 n,$$

and therefore, since $|x|^2 = (1-\delta^2)n$, that

$$\langle x, y\rangle \geq (1-\delta^2)n|y|_2.$$

This implies that

$$\begin{aligned}
|x-y|_2^2 &= \frac{1}{n}(|x|^2 + |y|^2 - 2\langle x, y\rangle) \\
&\leq 1 - \delta^2 + (1-2\delta^2)^2 - 2(1-\delta^2)(1-2\delta^2) \\
&= \delta^2,
\end{aligned}$$

which proves the claim.

We have therefore shown that $B_\delta(x)$ contains the set $C \setminus (1-2\delta^2)C$. Now, since $(1-2\delta^2)C$ has volume $(1-2\delta^2)^n$ times that of C, if n is sufficiently large, then $B_\delta(x)$ contains at least half of C. The result follows, since the volume of $B_\delta(x)$ is δ^n times that of the unit sphere of L_2^n. □

Now let y be a vector in L_2^n supported on $J \subset \{1, \ldots, n\}$ of cardinality m and satisfying the inequality $|y|_2^2 \geq \beta(1-2\eta)$. Again let P_J be the coordinate projection to the set J and define

$$V_y = \{x \in S^{n-1} : P_J x = y\}.$$

We will next obtain an upper bound for the spherical volume of $(V_y)_\epsilon$, which is the ϵ-expansion of V_y.

Lemma 1.5.7. *Let $\delta > \eta > 0$. Then when n is sufficiently large, the probability that a random unit vector belongs to $(V_y)_\eta$ is at most $4e\delta^n n \log n\left(1 + \frac{\sqrt{1-\beta(1-2\eta)}}{\delta-\eta}\right)^{n-m}$.*

Proof. If we cover V_y by N balls of radius $\delta - \eta$, then the balls of radius δ with the same centres cover $(V_y)_\eta$, so by Lemma 1.5.6 the probability that a random unit vector lies in $(V_y)_\eta$ is at most $2N\delta^n$. But V_y is an $(n-m)$-dimensional sphere of radius at most $\sqrt{1-\beta(1-2\eta)}$, hence Lemma 1.5.5 implies that it can be covered by at most $N = 2en\log n\Big(1+\frac{\sqrt{1-\beta(1-2\eta)}}{\delta-\eta}\Big)^{n-m}$ balls of radius $\delta-\eta$. This implies the result. □

Theorem 1.5.8. *The probability that a random unit vector belongs to $(\Delta_{\beta(1-2\eta)})_\eta$ is exponentially small.*

Proof. Lemma 1.5.1 tells us that the set $\Delta_{\beta(1-2\eta)}$ is an η-net of the set of unit vectors in Δ_β^η. Moreover, the same is true if we restrict to vectors of norm at most $1+\eta \le 2$. For each $x \in \Delta_{\beta(1-2\eta)}$ there is a set $J \subset \{1,2,\dots,n\}$ such that $\mu_x(J) \ge \beta(1-2\eta)$ and $|x_i|$ is a power of λ for every $i \in J$. If $|J| = m$, then Corollary 1.5.4 implies that there is an η-net of size at most $\Big(2e\log_\lambda\Big(\sqrt{\frac{2n}{m}}\frac{\lambda^2}{\eta}\Big)\Big)^m$ of the set of vectors y such that $|y_i|$ is a power of λ for every $i \in J$, $y_i = 0$ for $i \notin J$, and $\sum_i y_i^2 \le 2n$.

Every unit vector in $\Delta_{\beta(1-2\eta)}$ lies in V_y for some such J and y. Therefore, summing over all J and all y in an η-net for each J and applying Lemma 1.5.7, we find that the probability that a random unit vector belongs to $(\Delta_{\beta(1-2\eta)})_\eta$ is at most

$$\sum_{m=1}^{n}\binom{n}{m}\Big(2e\log_\lambda\Big(\sqrt{\frac{2n}{m}}\frac{\lambda^2}{\eta}\Big)\Big)^m 4e\delta^n n\log n\Big(1+\frac{\sqrt{1-\beta(1-2\eta)}}{\delta-\eta}\Big)^{n-m}.$$

Now let us set $\lambda = 4$. Using the upper bound $\binom{n}{m} \le (en/m)^m$ and setting $\theta = m/n$, we can bound the previous expression above by

$$4en\log n\sum_{m=1}^{n}\Big(2e^2\delta\frac{1}{\theta}\log_4\Big(\frac{16\sqrt{2}}{\eta\sqrt{\theta}}\Big)\Big)^{\theta n}\Big(\delta+\frac{\delta\sqrt{1-\beta(1-2\eta)}}{\delta-\eta}\Big)^{(1-\theta)n}.$$

To prove that this is exponentially small, it is sufficient to show that

$$\Big(\frac{2e^2}{\log 4}\frac{\delta}{\theta}\log\Big(\frac{16\sqrt{2}}{\eta\sqrt{\theta}}\Big)\Big)^\theta\Big(\delta+\frac{\delta\sqrt{1-\beta(1-2\eta)}}{\delta-\eta}\Big)^{1-\theta} \tag{1.11}$$

is bounded above by a constant less than 1 as θ varies. First of all, note that in §1.4 we obtain the following bounds on parameters: $\eta \le 10^{-5}$ (since $\eta = s-1$ and we need $\tau = 1-s^{-2} \le 10^{-4}$), $\epsilon \le \tau^2$ and $\beta < \frac{1}{32}\left(\frac{1}{8}-6\epsilon\right)$. We need moreover, that $\delta > \eta$. Let us note that the above expression is a decreasing function of β and since $\epsilon \le 10^{-8}$ we can take $\beta = \frac{1}{257}$.

To begin with, we shall show that there exist constants for which the result holds and then we shall choose some particular values to obtain an upper bound for the maximum.

Let us consider the condition $\delta + \frac{\delta\sqrt{1-\beta(1-2\eta)}}{\delta-\eta} < 1$. For $\eta = c\delta$ this becomes

$$\delta + \frac{\sqrt{1-\beta(1-2\eta)}}{1-c} < 1,$$

and we can choose c such that $\frac{\sqrt{1-\beta(1-2\eta)}}{1-c}$ is less than $1-\beta/4$, and then if $\delta \le \beta/8$ the inequality holds. For the first part of expression (1.11), if we assume that $\theta \ge \sqrt{\delta}$ (and again take $\eta = c\delta$) we have that

$$\begin{aligned}\left(\frac{2e^2}{\log 4}\frac{\delta}{\theta}\log\Big(\frac{16\sqrt{2}}{\eta\sqrt{\theta}}\Big)\right)^{\theta} &\le \left(2e^2\sqrt{\delta}\Big(\log(16\sqrt{2}) + \log\frac{1}{c\delta} + \frac{1}{4}\log\frac{1}{\delta}\Big)\right)^{\theta}\\ &\le 2e^2\sqrt{\delta}\left(C_1 + \frac{5}{4}\log\frac{1}{\delta}\right)^{\theta},\end{aligned}$$

where C_1 is an absolute constant. Hence, we can choose δ small enough such that $2e^2\sqrt{\delta}\left(C_1 + \frac{5}{4}\log\frac{1}{\delta}\right) \le \frac{9}{10}$.

If $\theta < \sqrt{\delta}$, then we need to consider the whole expression (1.11). To begin with note that

$$\begin{aligned}\left(\frac{2e^2}{\log 4}\frac{\delta}{\theta}\log\Big(\frac{16\sqrt{2}}{\eta\sqrt{\theta}}\Big)\right)^{\theta} &\le \left(2e^2\frac{\delta}{\theta}(C_1 + \frac{5}{2}\log\frac{1}{\theta})\right)^{\theta} \le \left(\frac{C_2\delta\log\frac{1}{\theta}}{\theta}\right)^{\theta}\\ &\le (C_2\delta)^{\theta}\left(\frac{1}{\theta^2}\right)^{\theta} \le \left(\frac{1}{\theta^2}\right)^{\theta},\end{aligned}$$

for $\delta < 1/C_2$. Moreover, we have that $\log\left((\frac{1}{\theta})^{2\theta}\right) = 2\theta\log\frac{1}{\theta} \le 2\sqrt{\theta} \le 2\delta^{1/4}$. Therefore we can estimate $\left(\frac{1}{\theta^2}\right)^{\theta}$ by $1 + 4\delta^{1/4}$. Recalling that the right hand side part in (1.11) is at most $(1-\frac{\beta}{8})^{1-\theta} \le (1-\frac{\beta}{8})^{1-\sqrt{\delta}}$, we deduce that the expression (1.11) is less than 1 if we choose δ such that $(1+4\delta^{1/4})(1-\frac{\beta}{8})^{1-\sqrt{\delta}} < 1$. Finally we choose the smallest δ, so that it fulfils all the inequalities and hence the result follows.

One can check that if we choose $\eta = 10^{-12}, \delta = 10^{-6}$ and $\beta = \frac{1}{257}$, then the maximum is at most 0.9982. Therefore, the desired probability is at most $4en^2\log n(0.9982)^n$. □

Recall the set $\Delta = \Delta_\beta^\eta$ defined in (1.5). It follows that

Corollary 1.5.9. *There exists $\alpha > 0$ such that if n is sufficiently large, then the probability that a random subspace X of dimension at most αn contains a vector $x \in \Delta$, is exponentially small.*

Proof. Let $\sigma < 1$ be such that the probability that a random unit vector belongs to $(\Delta_{\beta(1-2\eta)})_\eta$ is at most σ^n when n is sufficiently large. Now choose $\alpha > 0$ such that $(1+\frac{1}{\eta})^\alpha < \sigma^{-1}$. Then for sufficiently large n, the unit sphere of any subspace of dimension at most αn has an η-net of size τ^{-n} for some τ with $\tau^{-1} < \sigma^{-1}$. If we take such a net and rotate it randomly, then the probability that any element of the net lands within η of $\Delta_{\beta(1-2\eta)}$ is exponentially small.

By Lemma 1.5.1 we have that $\Delta_\beta^\eta \subset (\Delta_{\beta(1-2\eta)})_\eta$ and hence it follows that the probability that a random subspace intersects Δ is exponentially small. □

Remark 1.5.10. *For the particular choice of constants made in the proof of Corollary 1.5.8, the condition we obtain in Corollary 1.5.9 is*

$$0.9982(4en^2 \log n)^{1/n} < \sigma.$$

The left hand side is a decreasing function of n with limit 0.9982*, so for n large enough the left hand side is less than* 0.999*. We then have that*

$$\alpha < \frac{-\ln(0.999)}{\ln(1+10^{10})} \approx 4.345 \times 10^{-5}.$$

Hence, for n large enough we can take $\alpha = 4.3 \times 10^{-5}$ in the construction.

1.6 Summary

We have now proved our main theorem, Theorem 1.2.1, but since the steps of the argument are somewhat scattered, let us briefly recall them. We defined a continuous odd bijection $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ and a normalized version $\psi : S^{n-1} \to S^{n-1}$. In (1.5) we defined a set Δ and in the last section §1.5 we proved that a random $\lfloor \alpha n \rfloor$-dimensional subspace lies in Δ^c with exponentially high probability.

We also defined a set $\Gamma \subset S^{n-1}$, see (1.4), and observed that $\psi(\Delta^c) \subset \Gamma$. Section 1.4 was devoted to the proof that no great circle is contained in the expansion Γ_ϵ.

Therefore, if we pick a random $\lfloor \alpha n \rfloor$-dimensional subsphere X, then with exponentially high probability we have that $X \cap S^{n-1} \subset \Delta^c$, and therefore that $\psi(X \cap S^{n-1}) \subset \psi(\Delta^c) \subset \Gamma$, from which it follows that $\psi(X)_\epsilon$ contains no great circle. This gives Theorem 1.2.1.

As explained at the end of the introduction, this also gives a counterexample to Question 1.1.14. However, there is a small technical issue in that our scalar function $\varphi : \mathbb{R} \to \mathbb{R}$ used in the construction is not differentiable at zero, and hence the function $\psi : S^{n-1} \to S^{n-1}$ is not differentiable at any point with a coordinate equal to zero.

This can be dealt with in the obvious way – by approximating φ by a function that is differentiable everywhere and equal to φ except inside a very small interval about 0. Unfortunately, there does not seem to be a quick and easy argument that this change does not give rise to a great circle that consists entirely of δ-good* points. However, it is straightforward to make small adjustments to our construction, the main one of which is to enlarge very slightly the set B so that it contains an interval about 0, and to show that the final result is indeed robust in the appropriate sense.

Thus, a question that can be thought of as a suitable Lipschitz version of Milman's original question has a negative answer. However, because the derivatives of the function ψ at two points x and y can be very different even when x and y are close, when we compose it with a weighted ℓ_2 norm as described in §1.2, we obtain a function that is not a norm, so we do not obtain a counterexample to Question 1.1.7 (or equivalently Question 1.1.10). To obtain such a counterexample, one would need much better control over the second derivative of ψ.

Chapter 2

A counterexample to a strengthening of Milman's question

In this chapter we describe further progress made towards answering Milman's Question 1.1.4, which was introduced in the first chapter of this thesis. We prove that the strengthening of the question, i.e. Question 1.1.7, has a negative answer. This clearly implies that also Question 1.1.14 has a negative answer. This is joint work with W. T. Gowers.

2.1 Definition of the norm and initial observations

This chapter will be devoted to proving the following theorem.

Theorem 2.1.1. *Let $P \in L(\mathbb{R}^n)$ be a random orthogonal projection of rank $n/2$ and let $A = I + P$. Define a norm on $\mathbb{R}^n$ by*

$$\|x\| = \langle x, Ax\rangle^{1/2} + \eta n^{-1/2}\|x\|_1, \tag{2.1}$$

where η is an absolute constant. Then there exists $\eta > 0$ such that the normed space $(\mathbb{R}^n, \|\cdot\|)$ is 2-Euclidean and with positive probability, for n large enough, does not contain any 2-dimensional subspace that is both strongly $(1+\epsilon)$-Euclidean and strongly $(1+\epsilon)$-complemented.

Clearly, the normed space $(\mathbb{R}^n, \|\cdot\|)$ defined in the above theorem constitutes a counterexample to Question 1.1.7, and hence also to Question 1.1.14.

Let us begin by some straightforward observations about the norm defined in (2.1). The first part of this norm is a weighted ℓ_2 norm with respect to a random orthonormal basis, where half the weights are 2 and half are 1, and the second is a multiple of the standard ℓ_1 norm. Note that $|x| \leq \|x\| \leq (\sqrt{2}+\eta)|x|$ holds for every x, where $|\cdot|$ denotes the standard Euclidean norm. Hence, as long as $\eta \leq 2-\sqrt{2}$, the norm defined in (2.1) is indeed strongly 2-Euclidean.

We will prove that with probability greater than zero (and in fact close to 1) there is no 2-dimensional subspace in $(\mathbb{R}^n, \|\cdot\|)$ that consists entirely of ϵ-good points (recall Definition 1.1.8 in the previous chapter, and its equivalence to the statement of the theorem captured in Lemma 1.1.9), for some absolute constant $\epsilon > 0$.

2.1.1 Characterization of ϵ-good points

The next lemma tells us what the support functionals are at a vector x. Let us use the notation $\operatorname{sign}(t)$ for the multivalued function from $\mathbb{R}$ to $\mathbb{R}$ that takes t to 1 if $t > 0$, to -1 if $t < 0$, and to any element of $[-1, 1]$ if $t = 0$. Then if $x \in \mathbb{R}^n$ we write $\operatorname{sign}(x)$ for the result of applying the multivalued function sign pointwise. Let us also write $\|x\|_A$ for $\langle x, Ax\rangle^{1/2}$.

Recall from the previous chapter that a *support functional* of a norm $\|\cdot\|$ at x is any non-zero linear functional f such that for every y with $\|y\| \leq \|x\|$ we have $f(y) \leq f(x)$.

Lemma 2.1.2. *The support functionals at x are multiples of $\frac{Ax}{\|x\|_A} + \eta n^{-1/2} \operatorname{sign}(x)$.*

Proof. Essentially this is just a question of calculating the derivative of the norm, except that where the derivative is not defined we may have to give it several values (just as one might say that the derivative of $|x|$ at zero is any element of $[-1, 1]$).

Let y be a sufficiently small vector. Then for any possible choice of $\operatorname{sign}(x)$, we have that

$$\|x+y\| \geq \|x\| + \frac{\langle Ax, y\rangle}{\|x\|_A} + \eta n^{-1/2}\langle \operatorname{sign}(x), y\rangle + o(y).$$

Therefore, if y is orthogonal to any value of $\frac{Ax}{\|x\|_A} + \eta n^{-1/2} \operatorname{sign}(x)$, we have that $\|x+y\| \geq \|x\| + o(y)$, from which it follows easily that $\frac{Ax}{\|x\|_A} + \eta n^{-1/2} \operatorname{sign}(x)$ is a support functional at x. □

From this and Proposition 1.1.11, introduced in the previous chapter, it follows that if a unit vector x is ϵ-good, then there exists a unit vector y and a scalar λ such that $|y - x| < \delta$ and

$$\left|\frac{Ay}{\|y\|_A} + \eta n^{-1/2} \operatorname{sign}(y) - \lambda y\right| < \delta,$$

where $\delta > 0$ depends on ϵ only.

Speaking more qualitatively, this implies that $\mathrm{sign}(y)$ is close to the subspace generated by Py and Qy, where $Q = I - P$. Indeed, recalling that $A = I + P$, we get that the above inequality is the same as

$$\frac{2-\lambda}{\|y\|_A} Py + \frac{1-\lambda}{\|y\|_A} Qy \approx_\delta \eta n^{-1/2}\,\mathrm{sign}(y),$$

where we write $u \approx_\delta v$ to mean that $|u - v| \leq \delta$. We will use this convenient notation throughout the chapter.

It follows that if we have a 2-dimensional subspace Y that consists entirely of ϵ-good points, then every unit vector $x \in Y$ is close to a vector y such that $\mathrm{sign}(y)$ is close to the subspace $PY + QY$, which has dimension at most 4. The main part of the proof will be based on this observation.

It will also be useful to note that all ϵ-good points must be close to PX or QX.

Lemma 2.1.3. *Let x be an ϵ-good vector in $(X, \|\cdot\|)$, such that $|x| = 1$. Then either $d(x, PX) \leq 3\delta + 2\eta$ or $d(x, QX) \leq 3\delta + 2\eta$, where δ is given by Proposition 1.1.11.*

Proof. As we have seen, there exists a unit vector y and $\lambda > 0$ such that $|x - y| \leq \delta$ and such that

$$\frac{Ay}{\|y\|_A} + \eta n^{-1/2}\,\mathrm{sign}(y) \approx_\delta \lambda y.$$

We have that

$$d(x, PX) \leq d(x, y) + d(y, PX) \leq \delta + |y - Py| = \delta + |Qy|$$

and similarly for $d(x, QX)$. Hence, our goal is to bound $\min\{|Py|, |Qy|\}$.

But $|Ay| \geq \|y\|_A$, and $\eta n^{-1/2}|\,\mathrm{sign}(y)| \leq \eta$, so $\lambda \geq 1 - \eta - \delta$ and

$$|Ay - \lambda\|y\|_A y| < (\delta + \eta)\|y\|_A \leq \sqrt{2}(\delta + \eta).$$

Thus, y is an approximate eigenvector of A and it remains to prove that an approximate eigenvector of A must be close to an eigenvector. (This is of course false for general linear maps.)

Since $P + Q = I$, we have $y = Py + Qy$ and $Ay = 2Py + Qy$, so if ν is any scalar, then

$$|Ay - \nu y|^2 = (2-\nu)^2|Py|^2 + (1-\nu)^2|Qy|^2.$$

Writing $2-\nu = a+1/2$ and $1-\nu = a-1/2$, one can rewrite the right-hand side as

$$\Big(a + \frac{|Py|^2 - |Qy|^2}{2}\Big)^2 + \frac{1}{4}\Big(1 - (|Py|^2 - |Qy|^2)^2\Big),$$

from which we see that if $|Ay - \nu y|^2 \le \tau$ then

$$(|Py|^2 - |Qy|^2)^2 \ge 1 - 4\tau,$$

which implies that either $|Py|^2 \le 2\tau$ or $|Qy|^2 \le 2\tau$. Since we may set $\tau = 2(\delta+\eta)^2$ the result follows. □

2.2 Outline of the proof of Theorem 2.1.1

Let us call a non-zero vector a *sign vector* if all its coordinates have the same absolute value. As before, we write X for $(\mathbb{R}^n, \|\cdot\|)$, though sometimes we abuse notation and use X to refer simply to $\mathbb{R}^n$.

In order to show that the norm defined in (2.1) indeed constitutes a counterexample to Question 1.1.7, i.e. there is no two-dimensional subspace of $(\mathbb{R}^n, \|\cdot\|)$ that consists entirely of ϵ-good points, we shall obtain a contradiction using statements of the following kind:

1. Every ϵ-good point is close to PX or QX.

2. With high probability, every point that is close to PX or QX has coordinates that take several different values.

3. If Y is a 2-dimensional subspace that consists entirely of ϵ-good points, then for every $x \in Y$ there exists x' close to x such that $\mathrm{sign}(x')$ is close to the subspace $PY + QY$.

4. The vectors $\mathrm{sign}(x')$ are not approximately contained in a 4-dimensional subspace. (This uses the fact that no vector in Y can be approximated by a vector with only few distinct coordinates.)

We have already proved statements 1 and 3 above. Next, we formulate and prove statement 2. We begin with a crude upper bound for the volume of the γ-expansion of the unit sphere of a subspace of dimension cn. (Much more accurate estimates exist, but for us a simple argument suffices.)

Lemma 2.2.1. *Let $\gamma > 0$ and assume that $n \geq 1 - \frac{\log(8\gamma)}{\log(2)}$. Then the probability that a random unit vector in $\mathbb{R}^n$ has distance at most γ from a given subspace $Y \subset X$ of dimension m is at most $24^n\gamma^{n-m}$.*

Proof. A spherical cap in S^{n-1} of Euclidean radius 2γ has volume at most $(4\gamma)^{n-1}$. Hence, for $n \geq 1 + \frac{\log(1/(8\gamma))}{\log(2)}$, the volume of the cap is at most $(8\gamma)^n$. We can also find a γ-net of Y of cardinality at most $(3/\gamma)^m$. But every point of S^{n-1} that is within γ of Y is within 2γ of a point in the γ-net, and from this the result follows. □

Lemma 2.2.2. *Let k be a positive integer, let $\gamma > 0$, and let Y be a random subspace of ℓ_2^n of dimension m. Then the probability that Y_γ contains a unit vector x with at most k distinct coordinates is at most $(3/\gamma)^k(48k)^n\gamma^{n-m}$.*

Proof. The number of partitions of $\{1, 2, \ldots, n\}$ into k sets is at most k^n, and for each partition $E_1, \ldots, E_k$ the set of vectors that are constant on each E_i is a k-dimensional subspace, so there is a γ-net of the unit sphere of this subspace of size at most $(3/\gamma)^k$.

The probability that Y_γ contains a vector with at most k distinct coordinates is at most the probability that $Y_{2\gamma}$ contains a point in one of these γ-nets, which is at most

$$(3/\gamma)^k(24k)^n(2\gamma)^{n-m} \leq (3/\gamma)^k(48k)^n\gamma^{n-m}$$

by Lemma 2.2.1 and a union bound. □

We present one more technical lemma that is similar to Lemma 2.2.2, and which will be an important part of the argument. Again we make no attempt to optimize bounds.

Lemma 2.2.3. *Let Z be a random subspace of ℓ_2^n of dimension m. Then the probability that Z_γ contains a point with support size at most r is at most $288^n\gamma^{n-m-r}$*

Proof. The number of sets of size at most r is $\binom{n}{r} \leq 2^n$. For each such set E the size of a γ-net of the unit sphere of the space of vectors supported on E is at most $(3/\gamma)^r$, and for each point in such a net the probability that it is in $Z_{2\gamma}$ is at most $24^n(2\gamma)^{n-m}$. Therefore, the probability we wish to bound is at most

$$2^n(3/\gamma)^r 24^n(2\gamma)^{n-m} \leq 288^n\gamma^{n-m-r},$$

which proves the lemma. □

The key point we shall need from the above lemma is that for any $c > 0$ there exists $\gamma > 0$ such that if $n - m - r \geq cn$, then the probability that Z_γ contains a point with small support is small. In particular, we have that

Corollary 2.2.4. *Let $X = \ell_2^n$ with $n \geq 2$, and let $P : X \to X$ be a random orthogonal projection of rank $n/2$. Denoting $Q = I - P$, we have that if $\gamma = 2^{-37}$ then the probability that either $PX_{2\gamma}$ or $QX_{2\gamma}$ contains a vector of support size at most $n/4$ is at most $\left(\frac{2}{3}\right)^n$.*

Proof. Applying Lemma 2.2.3, we find that the probability that $PX_{2\gamma}$ contains a vector of support size at most $n/4$ is at most $288^n(2\gamma)^{n/4} = (288/512)^n$. The same is true of $QX_{2\gamma}$ and the result follows with room to spare. □

For the remainder of this chapter, we shall assume that P has been chosen in such a way that neither $PX_{2\gamma}$ nor $QX_{2\gamma}$ contains a vector of support size at most $n/4$, and neither PX_γ nor QX_γ contains a vector with at most five distinct coordinates. By Lemma 2.2.2 and Corollary 2.2.4 such a P exists.

2.3 The set of sign vectors cannot be squeezed into a 4-dimensional subspace

Before we move to the heart of the argument, let us remark that as we move forward we will be dealing with many parameters. Since we do not wish to choose them straight away we make sure that it is easy to keep track of all the dependencies by having them labeled.

Recall that if Y is a 2-dimensional subspace that consists entirely of ϵ-good points, then for every unit vector $x \in Y$ there exists a unit vector y such that $|x - y| \leq \delta$, and

$$\eta n^{-1/2}\,\mathrm{sign}(y) \approx_\delta \lambda y + \frac{Ay}{\|y\|_A}.$$

From this it follows that

$$n^{-1/2}\,\mathrm{sign}(y) \approx_{\delta/\eta} \alpha_y Py + \beta_y Qy$$

for some constants α_y and β_y. Note that $|Ay| \leq \sqrt{2}\|y\|_A$, so from the first approximation and the triangle inequality we find that $|\lambda| \leq \sqrt{2} + \delta + \eta$. Therefore, provided that

$$\delta + \eta \leq 2 - \sqrt{2} \tag{2.2}$$

it follows that $|\lambda| \leq 2$. Since $\|y\|_A \geq 1$, it follows that $|\alpha_y|$ and $|\beta_y|$ are at most $4\eta^{-1}$. This bound will be important later.

Suppose now that

$$3\delta + 2\eta \leq \gamma. \tag{2.3}$$

If $x \in Y$, then by Lemma 2.1.3, either $d(x, PX)$ or $d(x, QX)$ is at most $3\delta + 2\eta$. Therefore, for every $y \in Y_\gamma$, either $d(y, PX)$ or $d(y, QX)$ is at most $3\delta + 2\eta + \gamma \leq 2\gamma$, so by he assumption made at the end of the previous section, y has support size at least $n/4$. That is, every vector in Y_γ has support size at least $n/4$. We shall use this property frequently in the rest of the section.

Now let us choose non-negative real numbers $r_1, \dots, r_n$ and phases $\phi_1, \dots, \phi_n \in [0, 2\pi)$ such that each unit vector in Y is equal to

$$x_\theta = (r_1 \sin(\theta + \phi_1), \dots, r_n \sin(\theta + \phi_n))$$

for some $\theta \in [0, 2\pi)$. Note that by looking at $\mathbb{E}_\theta |x_\theta|^2$ we find that $\sum_i r_i^2 = 2$.

We begin with a couple of simple technical lemmas.

Lemma 2.3.1. *Let $\delta, \xi > 0$, let x and y be two vectors in $\mathbb{R}^n$ such that $|x - y| \leq \delta$. Then the number of i such that $|x_i| \geq \xi n^{-1/2}$ and $\operatorname{sign}(x_i) \neq \operatorname{sign}(y_i)$ is at most $\xi^{-2}\delta^2 n$.*

Proof. For every i such that $|x_i| \geq \xi n^{-1/2}$ and $\operatorname{sign}(x_i) \neq \operatorname{sign}(y_i)$, we have that $|x_i - y_i|^2 \geq \xi^2 n^{-1}$. Since we also have that $\sum_i |x_i - y_i|^2 \leq \delta^2$, the claim follows. □

Let $\alpha > 0$ be a constant to be chosen later, and let

$$E = \{i : r_i \geq \alpha n^{-1/2}\}.$$

Further define $E_0 = \{1, 2, \dots, n\} \setminus E$, and let us write P_E for the coordinate projection onto E.

Lemma 2.3.2. *Let $0 < c, \xi < 1$ and let θ be chosen uniformly at random from $[0, 2\pi)$. Then with probability at least $1 - \frac{4\xi}{\alpha\pi c}$, the number of $i \in E$ such that $|(x_\theta)_i| < \xi n^{-1/2}$ is less than $c|E|$.*

Proof. Since $r_i \geq \alpha n^{-1/2}$ for every $i \in E$, and $(x_\theta)_i = r_i \sin(\theta + \phi_i)$, we have that for each $i \in E$,

$$\mathbb{P}\Big[|(x_\theta)_i| < \xi n^{-1/2}\Big] \leq \mathbb{P}\Big[|\sin(\theta + \phi_i)| < \xi/\alpha\Big] < \frac{4\xi}{\alpha\pi}.$$

Therefore, the expected number of $i \in E$ such that $|(x_\theta)_i| < \xi n^{-1/2}$ is less than $4\xi|E|/\alpha\pi$.

The result now follows from Markov's inequality. □

For choices of parameters that we shall make later, let us call θ and x_θ *typical* if

(i) the conclusion of Lemma 2.3.2 holds, and

(ii) there is no $i \in E$ with $\theta + \phi_i \in \{0, \pi\}$.

The second condition, which is there for convenience, holds with probability 1 and ensures that $\operatorname{sign}(x_\theta)_i \in \{-1, 1\}$ for every $i \in E$. Later we shall want to be sure that typical vectors exist, for which Lemma 2.3.2 tells us that a sufficient condition is the inequality

$$\frac{4\xi}{\alpha\pi c} < 1. \tag{2.4}$$

Let us now define Σ to be the set of all vectors of the form $n^{-1/2} P_E \operatorname{sign}(x_\theta)$ such that x_θ is a typical element of Y. Let $\beta > 0$ be another parameter to be chosen later, and let $V \subset \Sigma$ be the maximal centrally symmetric β-separated subset of Σ. Assume that V consists of $2k$ vectors (that is k antipodal pairs) and note that V is a β-net of Σ.

Recall that every vector in Y, and even in Y_γ, has support size at least $n/4$. It follows that $|E| \geq n/4$, and therefore that Σ consists of vectors in a sphere of radius $n^{-1/2}|E|^{1/2} \geq 1/2$. Therefore, as long as

$$\beta \leq 1 \tag{2.5}$$

we can choose any typical vector x_θ, and letting $v = n^{-1/2} P_E \operatorname{sign}(x_\theta)$ thereby obtain a β-separated subset $\{v, -v\}$ of Σ. This proves that $k \geq 1$.

We now consider three cases depending on the size of V.

2.3.1 Case 1: $k = 1$

Let $\zeta = 4\xi/\alpha\pi c$ and let $V = \{v, -v\}$. Since θ is typical with probability at least $1 - \zeta$, every closed interval of length greater than $2\pi\zeta$ contains a typical θ. If $n^{-1/2} P_E \operatorname{sign}(x_\theta) \approx_\beta v$, then $n^{-1/2} P_E \operatorname{sign}(-x_\theta) \approx_\beta -v$, so there is at least one θ such that $n^{-1/2} P_E \operatorname{sign}(x_\theta) \approx_\beta v$ and one such that $n^{-1/2} P_E \operatorname{sign}(x_\theta) \approx_\beta -v$. It follows that there exist typical unit vectors $x, y \in Y$ such that $|x - y| \leq 2\pi\zeta$ and such that $d(n^{-1/2} P_E \operatorname{sign}(x), v) \leq \beta$, and $d(n^{-1/2} P_E \operatorname{sign}(y), -v) \leq \beta$.

It follows that $n^{-1/2} P_E \operatorname{sign}(x)$ differs from v in at most $\beta^2 n/4$ coordinates, and $n^{-1/2} P_E \operatorname{sign}(y)$ differs from $-v$ in at most $\beta^2 n/4$ coordinates. Therefore, $P_E \operatorname{sign}(x)$ and $P_E \operatorname{sign}(y)$ are equal in at most $\beta^2 n/2$ coordinates. Moreover, by Lemma 2.3.1, the number of i for which $\operatorname{sign}(x_i) \neq \operatorname{sign}(y_i)$ and $|x_i| \geq \rho n^{-1/2}$ is at most $\rho^{-2}(2\pi\zeta)^2 n = \rho^{-2}\left(\frac{8\xi}{\alpha c}\right)^2 n$.

From these two facts it follows that the number of coordinates $i \in E$ for which $|x_i| \geq \rho n^{-1/2}$ is at most $(\rho^{-2}\left(\frac{8\xi}{\alpha c}\right)^2 + \beta^2/2)n$. Therefore, we find that x has distance at most ρ from a vector of support size at most $\left(\rho^{-2}\left(\frac{8\xi}{\alpha c}\right)^2 + \beta^2/2\right)n$.

If we choose parameters in such a way that

$$\rho \leq \gamma \tag{2.6}$$

and

$$\rho^{-2}\left(\frac{8\xi}{\alpha c}\right)^2 + \beta^2/2 < 1/4 \tag{2.7}$$

we obtain a contradiction with the fact that Y_γ does not contain a vector of support size less than $n/4$.

2.3.2 Case 2: $2 \leq k \leq 4$

Let $V = \{\pm v_1, \ldots, \pm v_k\}$. Since each v_i is of the form $n^{-1/2} P_E \operatorname{sign}(x_\theta)$ for some typical vector $x_\theta \in Y$, it takes values $\pm n^{-1/2}$ in E.

We now show that either this case can be reduced to the case $k = 1$ with β replaced by 48β or there is a subset V' of V consisting of at least two antipodal pairs such that V' is a 3κ-separated κ-net of Σ and $\kappa \leq 16\beta$.

Since V is a β-net of Σ, then if it is 3β-separated then we are done. If not, we can find $i \neq j$ such that $|v_i - v_j| \leq 3\beta$. Then we can remove $\pm v_j$ from V and we will still have a 4β-net. Similarly, if $V' = V \setminus \{\pm v_j\}$ is 12β-separated we are done, but if it contains two distinct elements v_i, v_j such that $|v_i - v_j| \leq 12\beta$, then again we can remove $\pm v_j$ and we will still have a 16β-net. Finally, if there are two distinct elements v_i, v_j with $|v_i - v_j| \leq 48\beta$, then we may remove $\pm v_j$ and end up with V' of the form $\{v, -v\}$ and we are back in case $k = 1$. However, now β is replaced by 48β, which we must allow for when choosing our parameters, so we need

$$\beta \leq 1/48. \tag{2.8}$$

If the process stops before we reach $k = 1$, then we have a 3κ-separated κ-net $V' = \{\pm v_1, \ldots, \pm v_m\}$ of Σ such that $2 \leq m \leq 4$ and $\kappa \leq 16\beta$ as claimed.

Recall that every interval of length greater than $2\pi\zeta$ contains a typical θ, and hence for every $\tau > 0$ there must exist θ, ϕ, and $v_i \neq \pm v_j$ such that

$$|\theta - \phi| \leq 2\pi\zeta + \tau,$$

$$|n^{-1/2} P_E \operatorname{sign}(x_\theta) - v_i| \leq \kappa,$$

and

$$|n^{-1/2} P_E \operatorname{sign}(x_\phi) - v_j| \leq \kappa.$$

Since $|v_i \pm v_j| \geq 3\kappa$, it follows that $n^{-1/2}|P_E \operatorname{sign}(x_\theta) - P_E \operatorname{sign}(x_\phi)|$ and $n^{-1/2}|P_E \operatorname{sign}(x_\theta) + P_E \operatorname{sign}(x_\phi)|$ are both at least κ. Now recall that for every $x \in Y$ we can find y with $|x - y| \leq \delta$ and such that

$$n^{-1/2} \operatorname{sign}(y) \approx_{\delta/\eta} \alpha_y P y + \beta_y Q y$$

for some constants α_y and β_y. Let us choose y_θ and y_ϕ that have this relationship with x_θ and x_ϕ, respectively.

By Lemma 2.3.1 and the assumption that x is typical, the number of coordinates in E such that $\operatorname{sign}(x_\theta)$ and $\operatorname{sign}(y_\theta)$ differ is at most $c|E| + \xi^{-2}\delta^2 n \leq (c + \xi^{-2}\delta^2)n$, so

$$n^{-1/2}|P_E \operatorname{sign}(x_\theta) - P_E \operatorname{sign}(y_\theta)| \leq 2(c + \xi^{-2}\delta^2)^{1/2},$$

and similarly for ϕ. If we choose parameters in such a way that

$$c + \xi^{-2}\delta^2 \leq \beta^4/256, \tag{2.9}$$

it follows that these distances are both at most $\beta^2/8$.

Let us write α_θ instead of α_{y_θ}, and similarly for ϕ. Then

$$n^{-1/2} \operatorname{sign}(y_\theta) \approx_{\delta/\eta} \alpha_\theta P y_\theta + \beta_\theta Q y_\theta,$$

and

$$n^{-1/2} \operatorname{sign}(y_\phi) \approx_{\delta/\eta} \alpha_\phi P y_\phi + \beta_\phi Q y_\phi.$$

Also, since $|\theta - \phi| \leq 2\pi\zeta + \tau$, for a $\tau > 0$ that we are free to choose, we have $|x_\theta - x_\phi| \leq 2\pi\zeta$ (because $|x_\theta - x_\phi| < |\theta - \phi|$), and therefore $|y_\theta - y_\phi| \leq 2\pi\zeta + 2\delta$. It follows that

$$|\alpha_\phi P y_\phi + \beta_\phi Q y_\phi - (\alpha_\phi P y_\theta + \beta_\phi Q y_\theta)| \leq 2(\pi\zeta + \delta)(|\alpha_\phi| + |\beta_\phi|).$$

Now recall that $|\alpha_\phi|$ and $|\beta_\phi|$ are both at most $4\eta^{-1}$. We therefore obtain the approximation

$$n^{-1/2}\,\mathrm{sign}(y_\phi) \approx_\sigma \alpha_\phi P y_\theta + \beta_\phi Q y_\theta,$$

where $\sigma = \delta/\eta + 16(\pi\zeta + \delta)/\eta = (16\pi\zeta + 17\delta)/\eta$.

It is convenient to encapsulate our knowledge so far as an approximate matrix equation

$$\begin{pmatrix} \alpha_\theta & \beta_\theta \\ \alpha_\phi & \beta_\phi \end{pmatrix} \begin{pmatrix} P y_\theta \\ Q y_\theta \end{pmatrix} \approx_\sigma n^{-1/2} \begin{pmatrix} \mathrm{sign}(y_\theta) \\ \mathrm{sign}(y_\phi) \end{pmatrix}$$

where by $\approx_\sigma$ in this context we mean that the approximation holds coordinatewise.

The rough idea of what we shall now do is as follows. Because $\mathrm{sign}(y_\theta)$ and $\mathrm{sign}(y_\phi)$ are not roughly proportional to each other, the matrix $\begin{pmatrix} \alpha_\theta & \beta_\theta \\ \alpha_\phi & \beta_\phi \end{pmatrix}$ is well-invertible, which allows us to deduce from the approximate matrix equation that $P y_\theta$ and $Q y_\theta$ can both be approximated by linear combinations of $n^{-1/2}\,\mathrm{sign}(y_\theta)$ and $n^{-1/2}\,\mathrm{sign}(y_\phi)$, with coefficients that are not too large. Therefore, y_θ can as well, which in turn implies that x_θ also can be approximated in such a way. But $\mathrm{sign}(y_\theta)$ and $\mathrm{sign}(y_\phi)$ take at most two values each on almost all of E, and x_θ is small outside E, so x_θ can be approximated by a vector whose coordinates have at most five distinct values. Then we can obtain a contradiction from Lemma 2.2.2.

To carry out this argument we begin by making precise the statement that $\mathrm{sign}(y_\theta)$ and $\mathrm{sign}(y_\phi)$ are not roughly proportional.

Lemma 2.3.3. *Let u and v be vectors in $\mathbb{R}^n$ that take values $\pm n^{-1/2}$ in a set E of size m. Suppose that there are r values in E with $u_i = v_i$ and s values with $u_i \neq v_i$. Then for every $\lambda \in \mathbb{R}$, $|u - \lambda v| \geq 2(rs/mn)^{1/2}$.*

Proof. We have that

$$\begin{aligned} n|u - \lambda v|^2 &\geq r(1-\lambda)^2 + s(1+\lambda)^2 \\ &= (1+\lambda^2)m + 2\lambda(s-r). \end{aligned}$$

This is minimized when $\lambda = (r-s)/m$, and the minimum works out to be $4rs/m$. The lemma follows on dividing both sides by n and taking the square root. □

We showed earlier that the distance between $n^{-1/2}P_E\,\mathrm{sign}(x_\theta)$ and $\pm n^{-1/2}P_E\,\mathrm{sign}(x_\phi)$ is at least κ, which by assumption is at least β. It follows further that the conditions of

Lemma 2.3.3 apply to $n^{-1/2}\operatorname{sign}(x_\theta)$ and $n^{-1/2}\operatorname{sign}(x_\phi)$ with both r and s at least $\beta^2 n/4$. Therefore, using the trivial bound $|E| \leq n$, we deduce that

$$n^{-1/2}|\operatorname{sign}(x_\theta) - \lambda\operatorname{sign}(x_\phi)| \geq \beta^2/2.$$

Therefore, using the fact that

$$n^{-1/2}|\operatorname{sign}(x_\theta) - \operatorname{sign}(y_\theta)| \leq \beta^2/8$$

and

$$n^{-1/2}|\operatorname{sign}(x_\phi) - \operatorname{sign}(y_\phi)| \leq \beta^2/8,$$

we find that

$$n^{-1/2}|\operatorname{sign}(y_\theta) - \lambda\operatorname{sign}(y_\phi)| \geq \frac{\beta^2}{2} - \frac{\beta^2}{8}(|\lambda|+1) \geq \beta^2/8$$

when $|\lambda| \leq 2$. In case $|\lambda| \geq 2$ we can instead use the bound

$$(1+\lambda^2)m + 2\lambda(s-r) \geq (|\lambda|-1)^2 m$$

and the fact that $m \geq \beta^2 n/2 \geq \beta^4 n/4$ to deduce that

$$n^{-1/2}|\operatorname{sign}(x_\theta) - \lambda\operatorname{sign}(x_\phi)| \geq \frac{\beta^2}{2}(|\lambda|-1),$$

from which it follows that

$$n^{-1/2}|\operatorname{sign}(y_\theta) - \lambda\operatorname{sign}(y_\phi)| \geq \frac{\beta^2}{2}(|\lambda|-1) - \frac{\beta^2}{8}(|\lambda|+1),$$

which is again at least $\beta^2/8$.

Now that we have shown in a precise sense that $n^{-1/2}\operatorname{sign}(y_\theta)$ and $n^{-1/2}\operatorname{sign}(y_\phi)$ are not approximately proportional to each other, we turn to deducing that the matrix $\begin{pmatrix}\alpha_\theta & \beta_\theta\\ \alpha_\phi & \beta_\phi\end{pmatrix}$ is well-invertible, by which we simply mean that its determinant is not too small.

Lemma 2.3.4. *Let $A = \begin{pmatrix}a & b\\ c & d\end{pmatrix}$ be a 2×2 real matrix, let x and y be vectors in a Euclidean space such that $\langle x, y\rangle = 0$, and let*

$$\begin{pmatrix}u\\ v\end{pmatrix} = \begin{pmatrix}a & b\\ c & d\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}.$$

Then there exists λ *such that* $|u - \lambda v| \leq \frac{|x|\,|y|\,|\det(A)|}{|v|}$

Proof. Consider first the case where x and y are unit vectors. Then

$$\begin{aligned}|u - \lambda v|^2 &= (a - \lambda c)^2 + (b - \lambda d)^2 \\ &= (c^2 + d^2)\lambda^2 - 2(ac + bd)\lambda + a^2 + b^2.\end{aligned}$$

This is minimized when $\lambda = \frac{ac+bd}{c^2+d^2}$, and the minimum is

$$a^2 + b^2 - \frac{(ac + bd)^2}{c^2 + d^2} = \frac{(ad - bc)^2}{c^2 + d^2} \ ,$$

which proves the result.

In the general case, we have that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a|x| & b|y| \\ c|x| & d|y| \end{pmatrix} \begin{pmatrix} x/|x| \\ y/|y| \end{pmatrix} \ .$$

Using the case for unit vectors, we deduce that there exists λ such that

$$|u - \lambda v|^2 \leq \frac{|x|^2|y|^2 \det(A)^2}{c^2|x|^2 + d^2|y|^2},$$

and again the result is proved. □

Let us now apply this lemma with $A = \begin{pmatrix} \alpha_\theta & \beta_\theta \\ \alpha_\phi & \beta_\phi \end{pmatrix}$, $x = Py_\theta$ and $y = Qy_\theta$. Let $\begin{pmatrix} u' \\ v' \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix}$ and let $u = n^{-1/2}\,\mathrm{sign}(y_\theta), v = n^{-1/2}\,\mathrm{sign}(y_\phi)$. The approximate matrix equation proved earlier states that $|u - u'| \leq \sigma$ and $|v - v'| \leq \sigma$. It follows from the lemma that there exists λ such that

$$\frac{|Py_\theta|\,|Qy_\theta|\,|\det(A)|}{|v'|} \geq |u' - \lambda v'| \geq |u - \lambda v| - (1 + |\lambda|)\sigma.$$

But we have shown that $n^{-1/2}|\,\mathrm{sign}(y_\theta) - \lambda\,\mathrm{sign}(y_\phi)| \geq \beta^2/8$ for every λ. Since we also know that $|Py_\theta| \leq 1$ and $|Qy_\theta| \leq 1$, it follows that

$$|\det(A)| \geq (\beta^2/8 - (1 + |\lambda|)\sigma)|v'|.$$

From the proof of the last lemma it follows that the minimum distance is achieved when $\lambda = \frac{\alpha_\phi\alpha_\theta|x|^2+\beta_\phi\beta_\theta|y|^2}{|v'|^2}$. Recall also from the beginning of the section that $\alpha_\phi, \alpha_\theta, \beta_\phi, \beta_\theta \leq 4\eta^{-1}$. Since $|v' - v| \leq \sigma$, $v = n^{-1/2}\operatorname{sign}(y_\phi)$, and every y_ϕ has support size at least $n/4$, we get that $|v'| \geq 1/2 - \sigma$. Hence

$$|\lambda| \leq \frac{16}{\eta^2(1/2-\sigma)^2}.$$

Assuming that

$$\sigma \leq 2^{-12}\beta^2\eta^2 \tag{2.10}$$

we may deduce that $\sigma(1 + |\lambda|) \leq \beta^2/16$ and hence that $|\det A| \geq (1/2 - \sigma)\beta^2/16 \geq \beta^2/64$.

Let us rewrite the approximate matrix equation as

$$\begin{pmatrix}\alpha_\theta & \beta_\theta\\ \alpha_\phi & \beta_\phi\end{pmatrix}\begin{pmatrix}Py_\theta\\ Qy_\theta\end{pmatrix} = \begin{pmatrix}u'\\ v'\end{pmatrix} \approx_\sigma n^{-1/2}\begin{pmatrix}\operatorname{sign}(y_\theta)\\ \operatorname{sign}(y_\phi)\end{pmatrix}.$$

Then

$$\begin{pmatrix}Py_\theta\\ Qy_\theta\end{pmatrix} = \det(A)^{-1}\begin{pmatrix}\beta_\phi & -\beta_\theta\\ -\alpha_\phi & \alpha_\theta\end{pmatrix}\begin{pmatrix}u'\\ v'\end{pmatrix}.$$

Since the coefficients of A have absolute value at most $4\eta^{-1}$, it follows that both Py_θ and Qy_θ are linear combinations of u' and v' with coefficients of absolute value at most $256\eta^{-1}\beta^{-2}$. Using again the fact that $|u' - n^{-1/2}\operatorname{sign}(y_\theta)|$ and $|v' - n^{-1/2}\operatorname{sign}(y_\phi)|$ are both at most σ, it follows that both Py_θ and Qy_θ can be approximated to within $512\eta^{-1}\beta^{-2}\sigma$ by the corresponding linear combinations of $n^{-1/2}\operatorname{sign}(y_\theta)$ and $n^{-1/2}\operatorname{sign}(y_\phi)$, and hence that y_θ can be approximated to within $1024\eta^{-1}\beta^{-2}\sigma$ by a linear combination of $n^{-1/2}\operatorname{sign}(y_\theta)$ and $n^{-1/2}\operatorname{sign}(y_\phi)$ with coefficients of absolute value at most $512\eta^{-1}\beta^{-2}$.

Now recall that

$$n^{-1/2}|P_E\operatorname{sign}(x_\theta) - P_E\operatorname{sign}(y_\theta)| \leq 2(c + \xi^{-2}\delta^2)^{1/2},$$

and similarly for ϕ, which implies in particular that $n^{-1/2}P_E\operatorname{sign}(y_\theta)$ and $n^{-1/2}P_E\operatorname{sign}(y_\phi)$ can be approximated to within $2(c + \xi^{-2}\delta^2)^{1/2}$ by vectors whose coordinates take just the values $\pm n^{-1/2}$ on E. This is because x_θ and x_ϕ are typical and that E is the set of indices $i \in \{1, \dots, n\}$ for which $r_i \geq \alpha n^{-1/2}$ and hence all coordinates of $P_E\operatorname{sign}(x_\theta), P_E\operatorname{sign}(x_\phi)$ are ± 1.

Putting together the bounds obtained in the last two paragraphs we get that $P_E y_\theta$ can be approximated by a linear combination of $n^{-1/2}P_E\operatorname{sign}(x_\theta)$ and $n^{-1/2}P_E\operatorname{sign}(x_\phi)$ to within $1024\eta^{-1}\beta^{-2}\sigma+2048\eta^{-1}\beta^{-2}(c+\xi^{-2}\delta^2)^{1/2}$. In other words, $P_E y_\theta$ can be approximated by a vector with at most 4 distinct coordinates.

This in turn implies that $P_E x_\theta$ can be approximated to within $\delta + 1024\eta^{-1}\beta^{-2}\sigma + 2048\eta^{-1}\beta^{-2}(c+\xi^{-2}\delta^2)^{1/2}$ by such a vector. But $|x_\theta - P_E x_\theta| \leq \alpha$, so we end up with the conclusion that x_θ can be approximated to within $\alpha + \delta + 1024\eta^{-1}\beta^{-2}\sigma + 2048\eta^{-1}\beta^{-2}(c+\xi^{-2}\delta^2)^{1/2}$ by a vector that takes at most five distinct values (the fifth value being zero). Provided we have chosen our parameters in such a way that

$$\alpha + \delta + 1024\eta^{-1}\beta^{-2}\sigma + 2048\eta^{-1}\beta^{-2}(c+\xi^{-2}\delta^2)^{1/2} \leq \gamma \tag{2.11}$$

and since with our choice of γ (see Corollary 2.2.4) the expression $(3/\gamma)^5(240)^n\gamma^{n/2}$ is small, we obtain a contradiction to Lemma 2.2.2.

2.3.3 Case 3: $k \geq 5$

We begin with a simple lemma to estimate how well we can simultaneously approximate k orthonormal vectors by a $(k-1)$-dimensional subspace.

Lemma 2.3.5. *Let W be a $(k-1)$-dimensional subspace of $\mathbb{R}^n$ and let $u_1, \dots, u_k$ be an orthonormal sequence. Then there exists i such that $d(u_i, W) \geq k^{-1/2}$.*

Proof. Without loss of generality $k = n$. Now let v be a unit vector orthogonal to W. Then the orthogonal projection P_W to W is given by the formula $P_W(x) = x - \langle x, v\rangle v$, from which it follows that $d(x, W) = |\langle x, v\rangle|$. But $\sum_i \langle u_i, v\rangle^2 = 1$, so there must exist i such that $|\langle u_i, v\rangle| \geq k^{-1/2}$, which proves the lemma. □

Now, with the help of the assumption that $k \geq 5$, we prove that we cannot find a 4-dimensional subspace that approximately contains all the vectors in Σ. For convenience, let us reorder the coordinates in such a way that $E = \{1, 2, \dots, m\}$ and $0 \leq \phi_1 \leq \phi_2 \leq \dots \leq \phi_m < 2\pi$. Then for every typical vector x_θ, the set of i such that $(x_\theta)_i > 0$ is an interval mod m.

Lemma 2.3.6. *Let W be a 4-dimensional subspace of X. Then there is a vector $u = n^{-1/2} P_E \operatorname{sign}(x_\theta) \in \Sigma$ such that $d(u, W) \geq \beta/2\sqrt{5}$.*

Proof. Let $0 \leq \theta_1 < \dots < \theta_5 < \pi$ be such that $n^{-1/2} P_E \operatorname{sign}(x_{\theta_j}) \in \Sigma$ for $j = 1, 2, \dots, 5$ and write u_j for $n^{-1/2} P_E \operatorname{sign}(x_{\theta_j})$. For $j = 1, \dots, 5$, let $[a_j, b_j]$ be the interval mod m of i such that $(u_j)_i = n^{-1/2}$.

Note that $a_1 \geq \cdots \geq a_5$ and $b_1 \geq \cdots \geq b_5$, where here we refer to the cyclic ordering on the integers mod m. It follows that for each j, $(u_{j+1} - u_j)_i = 2n^{-1/2}$ on the interval $[a_{j+1}, a_j)$, $-2n^{-1/2}$ on the interval $(b_{j+1}, b_j]$, and zero everywhere else. In particular, the vectors of the form $u_{j+1} - u_j$ for $j = 1, 2, 3, 4$ and $u_1 + u_5$, are pairwise orthogonal.

By assumption Σ contains a β-separated set, so $|u_i \pm u_j| \geq \beta$ for every $i \neq j$. Setting $v_j = \frac{u_{j+1}-u_j}{|u_{j+1}-u_j|}$ for $j = 1, \ldots, 4$ and $v_5 = \frac{u_1+u_5}{|u_1+u_5|}$, we may deduce from Lemma 2.3.5 that $d(v_j, W) \geq 1/\sqrt{5}$ for some j. If $j = 1, \ldots, 4$ then this implies that $d(u_{j+1} - u_j, W) \geq \beta/\sqrt{5}$, which implies that either $d(u_{j+1}, W) \geq \beta/2\sqrt{5}$ or $d(u_j, W) \geq \beta/2\sqrt{5}$. Similarly, if $j = 5$ we get that either $d(u_1, W) \geq \beta/2\sqrt{5}$ or $d(u_5, W) \geq \beta/2\sqrt{5}$. This proves the lemma. □

Recall that (assuming appropriate choices of parameters (2.11))

$$n^{-1/2}|P_E \operatorname{sign}(x_\theta) - P_E \operatorname{sign}(y_\theta)| \leq \beta^2/8.$$

Recall also that

$$n^{-1/2} \operatorname{sign}(y_\theta) \approx_{\delta/\eta} \alpha_\theta P y_\theta + \beta_\theta Q y_\theta$$

for some coefficients $\alpha_\theta, \beta_\theta$ that have absolute values at most $4\eta^{-1}$. We also know that $|y_\theta - x_\theta| \leq \delta$. It follows that

$$n^{-1/2} P_E \operatorname{sign}(x_\theta) \approx_{\beta^2/8+\delta/\eta+8\delta/\eta} \alpha_\theta P_E P x_\theta + \beta_\theta P_E Q x_\theta.$$

It follows that the distance from $P_E \operatorname{sign}(x_\theta)$ to the subspace $P_E(PY + QY)$, which has dimension at most 4, is at most $\beta^2/8 + 9\delta\eta^{-1}$. If we pick parameters in such a way that

$$9\delta\eta^{-1} < \frac{\beta}{40}(4\sqrt{5} - 5\beta), \tag{2.12}$$

then this contradicts Lemma 2.3.6.

2.3.4 Choosing parameters

We conclude by showing that there exists a choice of parameters which fulfills all the conditions.

First, recall that we have already chosen γ in the Corollary 2.2.4 to be 2^{-37}. Further we see that $\beta = 2^{-6}$ satisfies (2.5) and (2.8). We can further choose $\eta = 2^{-40}$, $\alpha = 2^{-40}$, $\rho = 2^{-40}$ and then $c = 2^{-205}$, $\xi = 2^{-403}$ and finally $\delta = 2^{-506}$. It is easy to check that these parameters meet all the conditions.

This finishes the proof that the 2-Euclidean norm defined in (2.1), for $n \geq 34$, contains no two dimensional subspace which is both strongly $(1+\epsilon)$-complemented and strongly $(1+\epsilon)$-Euclidean with $\epsilon = \delta^2/8C^2 = 2^{-1017}$. The dependence of ϵ on δ was established in the proof of Lemma 1.1.11, while the condition on n comes from Lemma 2.2.1 and the fact that we use a 2γ-expansion.

Finally, we conclude that the Question 1.1.7 has a negative answer.

Chapter 3

Cost functions and order reversing involutions

This chapter is based on joint work with S. Artstein-Avidan and S. Sadovsky and is organised as follows. The first section will serve as an introduction to the basic notions in optimal transport theory. Mass transport problems are a classical object of study in mathematics, with a vast array of applications. The case when the cost function c is real-valued, namely $c(x, y) \neq +\infty$ for any two points, is well understood. However, non real-valued cost functions arise in a wide class of applications. They correspond to the case where some transportation schemes are simply prohibited. Infinite-valued costs appear naturally even in discrete settings, e.g. matchings.

The second section will mainly be concerned with a very specific one-parameter family of *non real-valued* cost functions, given by

$$p_t(x, y) = -\ln(\langle x, y\rangle - 1)_t,$$

where for a function f we define $(f(x))_t$ to be $f(x)$ if $f(x) \geq t$ and 0 otherwise. For the special value of the parameter $t = 0$, the cost is related to the important *polarity transform*, studied in for example [3] and [5]. The advantage in considering $t > 0$ is the boundedness of the corresponding cost function on the domain where it is finite (unlike in the case $t = 0$), which for instance allows one to use pre-existing theorems and show the existence of optimal transport plans.

Section 3 is devoted to the study of an order-reversing isomorphism $\mathcal{A}_t$, which we call the *truncated polarity transform* and which is associated with the cost p_t. Following [5], we show that on the induced p_t-class of functions, $\mathcal{A}_t$ is the unique (up to linear terms) order-reversing isomorphism.

In the final section of this chapter the notion of the *c-subgradient mapping* is introduced. It is an important concept in optimal transport theory and proving the existence of a function whose c-subgradient is the transport map will be the central question of the next chapter. Here, we present some general facts about the c-subgradient, and then develop further theory for the family p_t.

3.1 Introduction and background

3.1.1 Cost induced transforms and their function classes

In what follows, $c(x, y)$ will stand for a cost function which is allowed to attain the value $+\infty$ but not $-\infty$. We shall generally write $c : X \times Y \to (-\infty, \infty]$, but in this note it will always be the case that $X = Y = \mathbb{R}^n$.

Admissible pairs

A pair of functions $\varphi : X \to [-\infty, +\infty]$, $\psi : Y \to [-\infty, +\infty]$ will be called admissible if $\varphi(x) + \psi(y) \le c(x, y)$ for all $x \in X$ and $y \in Y$. In the case where $\varphi(x)$ and $\psi(y)$ are infinities of opposite signs, we shall use the convention $-\infty + \infty = -\infty$, so that there is *no* restriction on the cost in this case.

The cost induced transform

Given a function $\varphi : X \to [-\infty, +\infty]$, we may associate with it the largest ψ such that the two constitute an admissible pair. This defines a transform which is called the c-transform

Definition 3.1.1 (The c-transform)**.** *For $\varphi : X \to [-\infty, +\infty]$, its c-transform is defined to be the function $\varphi^c : Y \to \mathbb{R}$ given by*

$$\varphi^c(y) = \inf_x \left(c(x, y) - \varphi(x) \right).$$

Similarly for $\psi : Y \to [-\infty, +\infty]$, its c-transform is defined to be the function $\psi^c : X \to \mathbb{R}$ given by

$$\psi^c(x) = \inf_y \left(c(x, y) - \psi(y) \right).$$

Here if on the right hand side are infinities of opposite signs, which may occur (as $c \neq -\infty$) only if $\psi(y) = \infty$, we use the opposite convention, namely $-\infty + \infty = +\infty$, since when the cost $c(x,y)$ is infinite there is no restriction on the sum $\varphi(x)+\psi(y)$. In general one must be very careful with sums of opposite side infinities, as there is no obvious "rule of thumb" that can apply everywhere. To avoid arithmetic manipulations with infinite numbers, one may instead consider the infimum in the definition of φ^c to be taken only over those points x for which $c(x,y)<\infty$, and similarly for ψ^c.

Note that when $X=Y$ and $c(x,y)=c(y,x)$, the c-transforms in the Definition 3.1.1 amount to the same transform. We abuse notation slightly by using the same symbol for both transforms. This is because we usually assume that $X=Y=\mathbb{R}^n$ and that the cost is symmetric.

The following two facts follow directly from the definition.

Fact 3.1.2. *If φ and ψ are admissible then $\varphi^c \geq \psi$.*

Fact 3.1.3. *If $\varphi_1 \leq \varphi_2$ then $\varphi_1^c \geq \varphi_2^c$.*

A third simple fact is that on the image of the transform, that is, for functions of the form $\psi=\varphi^c$, the transform is an involution. In other words, we have

Fact 3.1.4. *For any φ we have that $\varphi^{ccc}=\varphi^c$.*

Proof. Indeed, the pair φ,φ^c is admissible and thus by Fact 3.1.2 we see that $\varphi\leq\varphi^{cc}$. We may now insert into this inequality the function φ^c, getting $\varphi^c\leq\varphi^{ccc}$. We may, instead, take the c transform of both side of this inequality, and by Fact 3.1.3 get that $\varphi^c\geq\varphi^{ccc}$. □

The c-class

Fact 3.1.4 implies that to each cost function there corresponds a natural class of functions on which the cost-induced transform is an order-reversing involution.

Definition 3.1.5 (The c-class)**.** *Given a cost function $c(x,y)$, the corresponding c-class is the image of the mapping $\varphi\mapsto\varphi^c$ on the class of all functions.*

It readily follows from the above facts and definition that

Fact 3.1.6. *A function φ is in the c-class if and only if $\varphi=\varphi^{cc}$.*

Among functions in the c-class, there is a subfamily which is of special importance, as in a sense it generates the class.

Definition 3.1.7 (Basic functions)**.** *A function of the form $\varphi(x)=c(x,y_0)+\beta$, for a fixed $y_0\in Y$ and a constant $\beta\in\mathbb{R}$, is called a basic function. When c is not a symmetric cost and X is different from Y, one must distinguish between X-basic functions of the form $\varphi(x)=c(x,y_0)+\beta$, and Y-basic functions of the form $\psi(y)=c(x_0,y)+\beta$.*

Fact 3.1.8. *The family of basic functions generated by a cost c belongs to the c-class.*

Proof. Consider the function $\psi_{y_0,\beta}: Y \to [-\infty, \infty]$ given by $\psi_{y_0,\beta}(y) = -\infty$ for $y \neq y_0$ and $\psi_{y_0,\beta}(y_0) = -\beta$. Then we have

$$\psi^c_{y_0,\beta}(x) = \inf_y \left(c(x,y) - \psi_{y_0,\beta}(y)\right) = c(x,y_0) + \beta.$$

Therefore, the function $\varphi(x) = c(x,y_0) + \beta$ is in the c-class. (One may also check directly that $\varphi = \varphi^{cc}$).

Note that we used the fact that $c(x,y) \neq -\infty$. Otherwise, under our convention, $-\infty + \infty = -\infty$ and the infimum could become $-\infty$ everywhere. □

It is useful to notice that the c-class is always closed under the operation of pointwise infimum.

Fact 3.1.9. *Assume $(\varphi_\alpha)_{\alpha\in I}$ is some family of functions which belong to the c-class, where I is some indexing set. Then the pointwise infimum*

$$\varphi(x) = \inf_{\alpha\in I} \varphi_\alpha(x)$$

is also in the c-class.

Proof. Indeed, by assumption there are functions ψ_α such that $\varphi_\alpha = \psi^c_\alpha$. Consider the function $\psi : Y \to [-\infty, \infty]$

$$\psi(y) = \sup_\alpha \psi_\alpha(y).$$

Even if ψ_α were in the c-class, we no longer claim this is the case for ψ. However, we may still consider its c-transform and we see that

$$\begin{aligned} \psi^c(x) &= \inf_y \left(c(x,y) - \psi(y)\right) = \inf_y \left(c(x,y) - \sup_\alpha \psi_\alpha(y)\right) \\ &= \inf_\alpha \inf_y \left(c(x,y) - \psi_\alpha(y)\right) = \inf_\alpha \psi^c(x) = \inf_\alpha \varphi_\alpha(x) = \varphi(x). \end{aligned}$$

□

It follows that the basic functions generate the c-class.

Fact 3.1.10. *Any function φ in the c-class is the infimum of basic functions.*

Whenever one encounters a class which is closed under pointwise infimum, one may define within this class the operation of $\hat{\sup}$ which is a replacement for the usual supremum, under which the class is not closed. For example, the class of convex bodies is closed under intersection, and the convex hull is the replacement for the union operation, which does not preserve convexity. The general definition is as follows:

Definition 3.1.11 (The operation $\hat{\sup}$)**.** *Let $\mathcal{S}$ be a class of functions which is closed under the operation of infimum. Given a family of functions $(\varphi_\alpha)_{\alpha\in I} \in \mathcal{S}$ for some index set I define*

$$\hat{\sup}(\varphi_\alpha : \alpha \in I) = \inf\{\varphi \in \mathcal{S} : \varphi \geq \varphi_\alpha \ \forall \alpha\}$$

This operation will prove itself important and we will come back to it after introducing order-reversing isomorphisms.

3.1.2 Connection with order-reversing isomorphisms

Duality is a classical notion in both geometry and analysis, and as it was pointed out in [4], it is derived from two abstract properties, namely the operation needs to be an *involution* and *order-reversing*. This allows one to define "duality" for functions.

Definition 3.1.12 (Order-reversing isomorphism)**.** *Let $\mathcal{S}$ be a class of functions. An order-reversing isomorphism is a bijective map $T : \mathcal{S} \to \mathcal{S}$ such that for every $\varphi, \psi \in \mathcal{S}$ we have*

(i) $TT\varphi = \varphi$,

(ii) $\varphi \leq \psi$ *if and only if* $T\varphi \geq T\psi$.

Before we move on to the connection between order-reversing isomorphisms and the c-transform let us state a general fact on the earlier defined $\hat{\sup}$ and order-isomorphisms.

Fact 3.1.13. *Let $\mathcal{S}$ be a class of functions which is closed under the operation of pointwise infimum. Let $T : \mathcal{S} \to \mathcal{S}$ be an order-reversing isomorphism. Then for $(\varphi_\alpha)_{\alpha\in I} \subset \mathcal{S}$ we have that*

$$T(\inf \varphi_\alpha) = \hat{\sup}(T\varphi_\alpha), \qquad T(\hat{\sup}\, \varphi_\alpha) = \inf(T\varphi_\alpha).$$

Proof. Indeed, by order reversal we know that since $\inf \varphi_\alpha \leq \varphi_\alpha$ for all α, we have that $T(\inf \varphi_\alpha) \geq T(\varphi_\alpha)$ for all α and it belongs the class $\mathcal{S}$. Hence, by the definition of $\hat{\sup}$, $T(\inf \varphi_\alpha) \geq \hat{\sup} T(\varphi_\alpha)$. On the other hand, as T is a bijection, there must be some function $\eta \in \mathcal{S}$ for which $T\eta = \hat{\sup} T(\varphi_\alpha)$. This means that $T\eta \geq T\varphi_\alpha$ for all α and by the fact that T is an order-reversing isomorphism this implies $\eta \leq \varphi_\alpha$ for all α and in particular $\eta \leq \inf_\alpha \varphi_\alpha$. We get that

$$T\eta \geq T(\inf_\alpha \varphi_\alpha) \geq \hat{\sup} T(\varphi_\alpha) = T\eta,$$

which implies that this is a line of equalities, and using injectivity, $\eta = \inf_\alpha \varphi_\alpha$ as claimed. $\square$

In a series of papers ([3],[4],[5]) S. Artstein-Avidan and V. Milman studied this concept of duality in two special classes of functions. They showed that on the class of lower-semicontinuous convex functions, which we will denote as $\mathrm{Cvx}(\mathbb{R}^n)$, any order-reversing involution must be, up to linear terms, the well known and important Legendre transform $\mathcal{L}$. They further considered the subclass $\mathrm{Cvx}_0(\mathbb{R}^n)$ of so-called *geometric convex functions*, which consists of non-negative lower-semicontinuous convex functions which take the value 0 at 0, and proved that on this subclass there are precisely two essentially different order-reversing isomorphisms, the Legendre transform $\mathcal{L}$ (for which $\mathrm{Cvx}_0(\mathbb{R}^n)$ is an invariant subclass) and the Polarity transform $\mathcal{A}$. Let us now demonstrate how these two transforms are connected to cost functions via cost induced transforms.

The Legendre transform

The Legendre transform is an extremely useful operation on functions, used throughout many areas in mathematics, see e.g. [16]. It is defined by

$$\mathcal{L}\varphi(y) = \sup_{x \in \mathbb{R}^n} \left(\langle x, y \rangle - \varphi(x) \right).$$

One may recover it by using one of the two following possibilities for cost functions: $c(x, y) = -\langle x, y \rangle$ or $\tilde{c}(x, y) = \frac{|x-y|^2}{2}$. Indeed,

$$\varphi^c(y) = \inf_x(-\langle x, y \rangle - \varphi(x)) = -\sup_x(\langle x, y \rangle - (-\varphi(x))) = -\mathcal{L}(-\varphi)(y).$$

One may easily check that the corresponding c-class is the class of upper-semicontinuous concave functions (which are allowed to attain the value $-\infty$). This also follows from the classical fact from convex geometry that $\mathcal{L}$ is an order-reversing involution on the class of lower-semicontinuous convex functions $\varphi : \mathbb{R}^n \to (-\infty, \infty]$.

Regarding the cost function $\tilde{c}$, we note that

$$\varphi^{\tilde{c}}(y) = \inf_x(|x|^2/2 - \langle x, y \rangle + |y|^2/2 - \varphi(x)) = |y|^2/2 + \inf_x(-\langle x, y \rangle - (\varphi(x) - |x|^2/2))$$

so that

$$\varphi^{\tilde{c}}(y) - |y|^2/2 = \inf_x(-\langle x, y\rangle - (\varphi(x) - |x|^2/2)) = (\varphi - |x|^2/2)^c.$$

This means that the $\tilde{c}$-class are simply functions in c-class with an added fixed function $-|x|^2/2$, and the induced transform is, up to this "change of appearance", the same.

Polarity transform

The $\mathcal{A}$-transform, which has recently received much attention [5, 6], is defined by

$$\mathcal{A}\varphi(y) = \sup_x \frac{(\langle x, y\rangle - 1)_+}{\varphi(x)},$$

where $f(x)_+ = \max\{f(x), 0\}$. It must be specified here that division by 0 of a positive number gives $+\infty$ whereas the fraction $\frac{0}{0}$ gives 0 (this corresponds, after taking a logarithm, to the fact that $-\infty + \infty = -\infty$ in this setting).

One may easily verify that the image of $\mathcal{A}$ is the set of previously mentioned geometric convex functions $\mathrm{Cvx}_0(\mathbb{R}^n)$, i.e. non-negative lower-semicontinuous convex functions which vanish at the origin (and are allowed to attain the value $+\infty$).

It turns out that $\mathcal{A}$ is induced by the cost function

$$p(x, y) = -\ln(\langle x, y\rangle - 1)_+.$$

Indeed, computing the p-transform we get that

$$\begin{aligned}\varphi^p(y) &= \inf_x(-\ln(\langle x, y\rangle - 1)_+ - \varphi(x)) = -\ln\left(\sup_x \frac{(\langle x, y\rangle - 1)_+}{e^{-\varphi(x)}}\right) \\ &= -\ln\left(\mathcal{A}(e^{-\varphi(\cdot)})(y)\right).\end{aligned}$$

Since the $\mathcal{A}$ transform is an order-reversing involution on the class of geometric convex functions, the p-class consists of $-\ln f$, where $f \in \mathrm{Cvx}_0(\mathbb{R}^n)$.

3.2 One parameter family of costs p_t

The Legendre transform and its associated cost function $c(x, y) = -\langle x, y\rangle$ have been thoroughly studied and are well understood (see [16] for an exposition). On the other hand, the cost function $p(x, y) = -\ln(\langle x, y\rangle - 1)_+$ associated with the Polarity transform does not fall into the standard setting because unlike most well studied cost functions it is not real valued. Therefore, many of the classical theorems from the optimal transport framework, which will be discussed in the next chapter, do not apply. Moreover, the region of finite cost $\{(x,y)\in\mathbb{R}^n\times\mathbb{R}^n : p(x,y)<\infty\}$ is an open set in $\mathbb{R}^{2n}$, which further complicates proving the existence of a transport plan with finite p cost.

For this reason we introduce a family of costs p_t with $t\geq 0$, defined as

$$p_t(x,y) = -\ln(\langle x,y\rangle - 1)_t,$$

where $f(x)_t = f(x)$ when $f(x)\geq t$ and 0 otherwise. It is clear that $p=p_0$ and that for $t>0$ the finite cost region is now a closed set (but it still is not real valued).

In this case the p_t-transform takes a form

$$\begin{aligned}\varphi^{p_t}(y) = \inf_x(-\ln(\langle x,y\rangle - 1)_t - \varphi(x)) &= -\ln\left(\sup_x \frac{(\langle x,y\rangle-1)_t}{e^{-\varphi(x)}}\right)\\ &= -\ln\left(\sup_{\{x:\langle x,y\rangle\geq 1+t\}} \frac{\langle x,y\rangle - 1}{e^{-\varphi(x)}}\right).\end{aligned}$$

This motivates defining the following definition for a "t-truncated polarity" transform $\mathcal{A}_t$.

Definition 3.2.1. *Let $t\geq 0$ and let $\varphi:\mathbb{R}^n\to[0,\infty]$. We define*

$$(\mathcal{A}_t\varphi)(y) = \sup_{\{x:\langle x,y\rangle\geq 1+t\}} \frac{\langle x,y\rangle - 1}{\varphi(x)}.$$

Similarly to the case $t=0$ we see that

$$e^{-\varphi^{p_t}} = \mathcal{A}_t e^{-\varphi}.$$

From this point onward we shall move back and forth between the logarithmic "cost level" and the multiplicative level. The main reason for this is that we already have some intuition for convex and geometric convex functions, but at the same time we have the general theory well developed for additive costs. We should thus keep a flexible mind in moving between these two levels: in one the functions are $\varphi:\mathbb{R}^n\to[-\infty,\infty]$ and in the other they are $f=e^{-\varphi}:\mathbb{R}^n\to[0,\infty]$.

Useful examples

To provide some intuition about the properties of the transform $\mathcal{A}_t$, we include a few simple examples here, and give the computations along with slightly more elaborate examples in Appendix B.

Example 3.2.2. *Let $t > 0$. For the convex indicator function $1_K^{\infty,s}$ which attains the value 0 on the interior of the convex body K, which is assumed to include the origin 0, the value $s \in [0, \infty]$ on its boundary, and the value $+\infty$ outside of K, we have that*

$$\mathcal{A}_t 1_K^{\infty,s} = 1_{(1+t)K^\circ}^{\infty,t/s}$$

where $t/0 = \infty$ and $t/\infty = 0$.

In particular we see that convex indicators of open convex sets are $\mathcal{A}_t$ functions as well.

Example 3.2.3. *For the truncated linear function $\ell_{y_0}(x) = \langle x, y_0 \rangle_+$ we have that*

$$\mathcal{A}_t \ell_{y_0} = \ell^\infty_{y_0/|y_0|^2}, \text{ and } \ell_{y_0} = \mathcal{A}_t \ell^\infty_{y_0/|y_0|^2}$$

where $\ell^\infty_{z_0}$ denotes the function which is ℓ_{z_0} on the ray $\mathbb{R}^+ z_0$ and $+\infty$ elsewhere. For the proof see Appendix B.

We have said in the general case that the basic functions for a given cost are functions of the form $\varphi(x) = c(x, y_0) + \beta$. In our particular case we will thus be dealing with functions

$$\varphi(x) = -\ln(\beta(\langle x, y \rangle - 1)_t),$$

where β is a non-negative constant. Correspondingly, in the multiplicative level, the basic functions are of the form $f_{y_0,\beta}(x) = \beta(\langle x, y_0 \rangle - 1)_t$, and based on their appearance we call these "escalator" functions. Let us fix the notation for this family of functions and define the *escalator function* $s_{a,b}$, where $a \in \mathbb{R}^n$ and $b \geq 0$, as

$$s_{a,b}(x) = b(\langle x, a \rangle - 1)_t.$$

In accordance with Fact 3.1.10, every function in the image of $\mathcal{A}_t$ can be expressed as a supremum of such escalator functions. Since the image of the transform $\mathcal{A}_t$, namely $e^{-\varphi}$ for φ in the p_t-class, will be one of our main objects to of study in this chapter, we give it a name and notation.

Definition 3.2.4. *We denote by $\mathcal{S}_t$ the class of functions which are in the image of $\mathcal{A}_t$. We call this class of functions the $\mathcal{A}_t$-class.*

We will also make use of the following special functions, which we call "knife"-functions: The *knife function* $k_{a,b}$, where $a \in \mathbb{R}^n$ and $b > 0$, is defined to be ∞ outside of the segment $[0, a] = \mathrm{conv}\{0, a\}$, and on the segment it is given by $k_{a,b}(x) = b|x|$. In other words,

$$k_{a,b}(x) = \max(1^{\infty}_{[0,a]}(x), b|x|),$$

where the function $1^{\infty}_{[0,a]}(x)$ is equal 0 if $x \in [0, a]$ and ∞ otherwise.

The following fact tells us that escalator functions and knife functions are *dual* in the $\mathcal{A}_t$ setting.

Fact 3.2.5. *The function $k_{a,b}$ is in $\mathcal{A}_t$-class, and it is the image (and pre-image) under $\mathcal{A}_t$ of the escalator function $s_{a,(|a|b)^{-1}}(y) = \frac{(\langle y,a\rangle-1)_t}{|a|b}$. In particular, $\mathcal{A}_t\mathcal{A}_t k_{a,b} = k_{a,b}$.*

Proof. Let us compute $\mathcal{A}_t s_{a,(|a|b)^{-1}}$. It follows from the definitions that

$$(\mathcal{A}_t s_{a,(|a|b)^{-1}})(y) \quad = \quad \sup_x \frac{(\langle x, y\rangle - 1)_t}{s_{a,(|a|b)^{-1}}(x)}.$$

So that if there is some x with $s_{a,(|a|b)^{-1}}(x) = 0$ and $\langle x, y\rangle - 1 \geq t$, then $\mathcal{A}_t s_{a,(|a|b)^{-1}} = \infty$. Since $s_{a,(|a|b)^{-1}}(x) = 0$ on the half-space $\langle x, a\rangle < 1 + t$, the only possibility for the two halfspaces not to intersect is if y is parallel to a and, furthermore, is in the segment $[0, a]$. Therefore, on the complement of this segment $\mathcal{A}_t s_{a,(|a|b)^{-1}}$ is $+\infty$. On the segment we have, for $y = \xi a$ with $\xi \in [0, 1]$, that

$$(\mathcal{A}_t s_{a,(|a|b)^{-1}})(\xi a) \quad = \quad \sup_x |a|b \frac{(\langle \xi a, x\rangle - 1)_t}{(\langle a, x\rangle - 1)_t}.$$

Clearly the expression is at most $|a|b\xi = b|y|$ (by replacing the numerator with $\xi(\langle a, x\rangle - 1)_t$), and by taking $x = Ra$ for $R \to \infty$ we get that the supremum is precisely $|a|b\xi = b|y|$, as claimed). This already implies that $\mathcal{A}_t s_{a,(|a|b)^{-1}} = k_{a,b}$ and so $k_{a,b}$ is in the image, and as escalator functions are the basic functions in the $\mathcal{A}_t$-class. This completes the proof.

For convenience we include the computation of $\mathcal{A}_t k_{a,b}$ as well. It follows from the definitions that

$$(\mathcal{A}_t k_{a,b})(y) \quad = \quad \sup_{x\in[0,a]} \frac{(\langle x, y\rangle - 1)_t}{|x|b}.$$

In the case where $\langle a, y\rangle < 1+t$ the supremum will be 0. On the half-space $\langle a, y\rangle \geq 1+t$ the supremum will be attained when $x = a$, in which case it is equal to $\frac{(\langle a,y\rangle - 1)_t}{|a|b}$, in other words, we get precisely $s_{a,(|a|b)^{-1}}$. □

The $\hat{\inf}$ operation

In the multiplicative level we know that $\mathcal{S}_t$ is closed under supremum, and therefore, analogously to Definition 3.1.11, we define the operation $\hat{\inf}$.

Definition 3.2.6 (The operation $\hat{\inf}$). *Let $(f_\alpha)_{\alpha\in I}$ be a family of $\mathcal{A}_t$-functions, for some index set I. Then their regularized infimum is defined by*

$$\hat{\inf}(f_\alpha : \alpha \in I) = \sup\{g \in \mathcal{S}_t : g \leq f_\alpha \ \forall\alpha\}$$

It follows from Fact 3.1.13 that

Fact 3.2.7. *For any family of $\mathcal{A}_t$-functions $(f_\alpha)_{\alpha\in I}$, it holds that*

$$\hat{\inf}(f_\alpha : \alpha \in I) = \mathcal{A}_t(\sup_\alpha \mathcal{A}_t f_\alpha) = \mathcal{A}_t\mathcal{A}_t(\inf_\alpha f_\alpha).$$

For the case $t = 0$, the operation $\hat{\inf}(f_\alpha : \alpha \in I)$ is the operation of taking the convex hull of the unions of the epi-graphs of the f_α's. The same is true for the Legendre transform (whereas for $\mathcal{L}$, the epi-graphs do not have to lie in the upper half-space and include the origin). Indeed, this is what occurs when one performs $\mathcal{A}_0\mathcal{A}_0$ on any (non-negative, vanishing at the origin) function. For $\mathcal{A}_t$ this is geometrically a different operation, and unfortunately does not have such a clear geometric illustration.

The function class

To understand better what is the class $\mathcal{S}_t$ of $\mathcal{A}_t$-functions, we establish a few other properties of this class. It has already been shown above that

Fact 3.2.8. *On $\mathcal{S}_t$, the transform $\mathcal{A}_t$ is an order-reversing involution. In general, $\mathcal{A}_t$ is order-reversing and $\mathcal{A}_t\mathcal{A}_t\varphi \leq \varphi$.*

Fact 3.2.9. *The functions in $\mathcal{S}_t$ are precisely those functions which can be written as the supremum of escalator functions.*

Fact 3.2.10. *The class $\mathcal{S}_t$ is closed under* max *(and* sup*). Moreover, one has that*

$$\mathcal{A}_t(\inf_{\alpha\in I}(\varphi_\alpha)) = \sup_{\alpha\in I}(\mathcal{A}_t\varphi_\alpha) \quad \text{and} \quad \mathcal{A}_t(\sup_{\alpha\in I}(\varphi_\alpha)) = \hat{\inf}_{\alpha\in I}(\mathcal{A}_t\varphi_\alpha)$$

The first new structural fact about $\mathcal{A}_t$ functions we would like to explain is that while they are not necessarily convex, the whole segment joining the point and the origin must also be in the epi-graph. In other words, we prove:

Lemma 3.2.11. *Any function in $\mathcal{S}_t$ can be written as the infimum of knife functions. In particular, an $\mathcal{A}_t$-function must be non-decreasing on every ray, and moreover, as a one-dimensional function on the ray, not only is the map $t \mapsto \varphi(tx)$ non-decreasing, but so is the map $t \mapsto \varphi(tx)/t$.*

Proof. It follows from Fact 3.2.5, 3.2.9 and 3.2.10 that any function $\varphi \in \mathcal{S}_t$ can be written as $\hat{\inf}$ of knife functions. It remains to show that for any z such that $\varphi(z) < \infty$, we have $k_{z,\varphi(z)} \geq \varphi$. In other words, we need to show that for $\lambda \in [0,1]$ if $y = \lambda z$ then $\varphi(y) \leq \lambda\varphi(z)$. Since the function φ belongs to the class, we may write that $\varphi = \mathcal{A}_t\psi$, and our claim is that

$$\sup_{\{x:\langle x,y\rangle \geq 1+t\}} \frac{\langle x,y\rangle - 1}{\psi(x)} \leq \lambda \sup_{\{x:\langle x,z\rangle \geq 1+t\}} \frac{\langle x,z\rangle - 1}{\psi(x)}.$$

Note that $\{x : \langle x,y\rangle \geq 1+t\} = \{x : \lambda\langle x,z\rangle \geq 1+t\} \subset \{x : \langle x,z\rangle \geq 1+t\}$, hence

$$\sup_{\{x:\langle x,y\rangle \geq 1+t\}} \frac{\langle x,y\rangle - 1}{\psi(x)} \leq \sup_{\{x:\langle x,z\rangle \geq 1+t\}} \frac{\langle x,y\rangle - 1}{\psi(x)} \leq \sup_{\{x:\langle x,z\rangle \geq 1+t\}} \lambda\frac{\langle x,z\rangle - 1}{\psi(x)},$$

which translates to the desired inequality. □

It is useful to know that while the class $\mathcal{S}_t$ clearly contains functions which are not geometric convex, it does contain all geometric convex functions. In fact, we show that

Lemma 3.2.12. *Let $0 \leq t_1 \leq t_2 < \infty$. Then*

$$\mathcal{S}_{t_1} \subset \mathcal{S}_{t_2}.$$

In particular, $\mathrm{Cvx}_0(\mathbb{R}^n) \subset \mathcal{S}_t$ for any $t \geq 0$.

Proof. It is enough to consider the case $t_1 < t_2$ as the equality case is trivial. By Facts 3.2.5 and 3.2.9 it suffices to show that all the functions of the form $\varphi(y) = (\mathcal{A}_{t_1}k_{a,b})(y)$, which is $\frac{\langle a,y\rangle - 1}{|a|b}$ for $\langle a,y\rangle \geq 1+t_1$ and 0 otherwise, are in $\mathcal{S}_{t_2}$. To see this we will find a family of "t_2-escalator" functions $\psi_\alpha \in \mathcal{S}_{t_2}$ such that $\varphi(y) = \sup_{\alpha\in I}\psi_\alpha(y)$ is a pointwise supremum, which by Fact 3.2.10 belongs to $\mathcal{S}_{t_2}$.

Define $\psi_\alpha(y) = \mathcal{A}_{t_2} k_{A_\alpha, B_\alpha}$, where $A_\alpha = \frac{1+t_2}{\alpha(1+t_1)} a$ and $B_\alpha = \frac{(1+t_1)t_2\alpha}{(1+t_2)((1+t_1)\alpha-1)} b$ with a parameter $\alpha \geq 1$. Again, using Fact 3.2.5 we see that

$$\psi_\alpha(y) = \frac{\frac{1+t_2}{\alpha(1+t_1)}\langle a, y\rangle - 1}{\frac{t_2}{(1+t_1)\alpha-1}|a|b}, \text{ when } \langle a, y\rangle \geq (1+t_1)\alpha$$

and 0 otherwise. Note that for y_α such that $\langle a, y_\alpha\rangle = (1+t_1)\alpha$ we have

$$\varphi(y_\alpha) = \frac{(1+t_1)\alpha - 1}{|a|b} = \psi_\alpha(y_\alpha).$$

If $1 \leq \alpha \leq \frac{1+t_2}{1+t_1}$ then we have that $\psi_\alpha(y) \leq \varphi(y)$ for all y. We have already shown that if y is such that $\langle a, y\rangle < (1+t_1)\alpha$ then $\psi_\alpha(y) = 0 \leq \varphi(y)$, and that if $\langle a, y\rangle = (1+t_1)\alpha$ then we have $\psi_\alpha(y) = \varphi(y)$. Finally, for y with $\langle a, y\rangle > (1+t_1)\alpha$, one can easily check that since $\alpha \leq \frac{1+t_2}{1+t_1}$ the gradient of $\psi_\alpha(y)$ has magnitude at most that of the gradient of $\varphi(y)$. Moreover, note that for $\alpha' = \frac{1+t_2}{1+t_1}$ we have that $\psi_{\alpha'}(y) = \varphi(y)$ for all y with $\langle a, y\rangle \geq 1 + t_2$. It follows that

$$\varphi(y) = \sup_{\{1 \leq \alpha \leq \frac{1+t_2}{1+t_1}\}} \psi_\alpha(y),$$

and hence the proof is finished. □

We have thus showed that for every $t \geq 0$ the class of functions $\mathcal{S}_t$ includes $\mathrm{Cvx}_0(\mathbb{R}^n)$. It turns out that for small positive t, these classes are not very different, in a sense explained in the next lemma.

Lemma 3.2.13. *Let $\varphi \in \mathcal{S}_t$. Then the sub-level-sets of φ are convex sets, and there exists some $\psi \in \mathrm{Cvx}_0(\mathbb{R}^n)$ such that for all y,*

$$(t+1)\psi\left(\frac{y}{t+1}\right) \leq \varphi(y) \leq \psi(y).$$

Proof. Note that the r-sub-level-set of $\psi = \sup_{\alpha \in I} \psi_\alpha$ is simply the intersection

$$K_r(\psi) = \{x : \psi(x) \leq r\} = \{x : \forall \alpha\,,\, \psi_\alpha(x) \leq r\} = \bigcap_{\alpha \in I} K_r(\psi_\alpha).$$

Combining this with Fact 3.2.9 and the fact that the level sets of functions of the form $\mathcal{A}_t k_{a,b}$ are half-spaces (which can be open or closed), we get that the level sets are convex.

For the second assertion let us notice that

$$\begin{aligned}(\mathcal{A}_t\varphi)(y) &= \sup_{\{x:\langle x,y\rangle\geq 1+t\}} \frac{\langle x,y\rangle-1}{\varphi(x)}\\ &= \sup_{\{x:\langle x,y\rangle\geq 1\}} \frac{\langle (1+t)x,y\rangle-1}{\varphi((1+t)x)}\\ &= \sup_{\{x:\langle x,y\rangle\geq 1\}} \left(\frac{\langle x,y\rangle-1}{\frac{1}{1+t}\varphi((1+t)x)}+\frac{t}{\varphi((1+t)x)}\right).\end{aligned}$$

Now let $\varphi_t(x)=\frac{1}{1+t}\varphi((1+t)x)$ and note that the above computation shows that

$$\mathcal{A}_t\varphi \geq \mathcal{A}\varphi_t.$$

Moreover, we have that

$$\mathcal{A}\varphi_t(y)=\sup_{\{x:\langle x,y\rangle\geq 1\}} \frac{\langle x,y\rangle-1}{\frac{1}{1+t}\varphi((1+t)x)}=(1+t)\mathcal{A}\varphi\Big(\frac{y}{1+t}\Big).$$

Finally, as we are taking the supremum over a smaller set, it is clear that

$$\mathcal{A}_t\varphi \leq \mathcal{A}\varphi.$$

It follows that

$$(1+t)\mathcal{A}\varphi\Big(\frac{y}{1+t}\Big)\leq \mathcal{A}_t\varphi(y)\leq \mathcal{A}\varphi(y),$$

which, since the image space of $\mathcal{A}_t$ and $\mathcal{A}$ are respectively $\mathcal{S}_t$ and $\mathrm{Cvx}_0(\mathbb{R}^n)$, concludes the proof. □

3.3 Uniqueness of $\mathcal{A}_t$ on $\mathcal{S}_t$

In this section we determine all order-isomorphisms on the class $\mathcal{S}_t$. We show that up to linear variants, for $t>0$, the only order-reversing transform on the function class $\mathcal{S}_t$ is $\mathcal{A}_t$. The case $t=0$, which was investigated in [5], is different as there are, up to linear variants, *two* order-reversing isomorphisms, one is $\mathcal{A}_0$ and the other is the Legendre transform, under which the class $\mathcal{S}_0$ is invariant.

Following [5] we will prove that

Theorem 3.3.1. *Let $t > 0$. Any order-reversing bijection $T : \mathcal{S}_t(\mathbb{R}^n) \to \mathcal{S}_t(\mathbb{R}^n)$ with $n \geq 1$ is of the form*
$$Tf = c(\mathcal{A}_t f) \circ B$$
where $B \in GL_n(\mathbb{R})$ and $c > 0$.

As we already know that $\mathcal{A}_t$ is an order-reversing isomorphism, determining all order-reversing isomorphisms is equivalent to determining all order-preserving ones, since an order-reversing isomorphism can be composed with $\mathcal{A}_t$ to form an order-preserving one. Thus, we may deduce Theorem 3.3.1 from the following:

Theorem 3.3.2. *Let $t > 0$. Any order-preserving bijection $T : \mathcal{S}_t(\mathbb{R}^n) \to \mathcal{S}_t(\mathbb{R}^n)$ with $n \geq 1$ is of the form*
$$Tf = cf \circ B$$
where $B \in GL_n(\mathbb{R})$ and $c > 0$.

We recall the definition of an order-preserving isomorphism on a class $\mathcal{S}$: it is a bijection $T : \mathcal{S} \to \mathcal{S}$ which satisfies that $f \leq g$ if and only if $Tf \leq Tg$. We begin with an observation regarding the action of an order-preserving map on linear functions and indicators.

Lemma 3.3.3. *Let $T : \mathcal{S}_t(\mathbb{R}^n) \to \mathcal{S}_t(\mathbb{R}^n)$ be an order-preserving isomorphism. Then the zero function $l_0(x) \equiv 0$ and the infinity function $1^{\infty}_{\{0\}}$ (which is 0 at the origin and $+\infty$ elsewhere) are fixed points of T.*

Proof. These are the maximal and minimal functions in the class. □

We will first consider the case of dimension 1, and in fact only consider functions defined on a ray emanating from the origin (as any function in $\mathcal{S}_t(\mathbb{R})$ is simply the concatenation of a function on $\mathbb{R}^+$ and a function on $\mathbb{R}^-$). The result for general dimension n will follow from this case in a similar way to the one used in the paper [5].

3.3.1 Proof for the class $\mathcal{S}_t(\mathbb{R}^+)$

By $f \not\gtrless g$ let us denote the fact that f and g are not comparable, i.e. there exist x, y such that $f(x) > g(x)$ and $f(y) < g(y)$. It will be useful to introduce a property (P) which was used also in the original [5] paper about the polarity transform $\mathcal{A}$.

Definition 3.3.4. *Let $\mathcal{S}$ be a partially ordered class of functions which is closed under the operation of pointwise maximum. We say a function $f \in \mathcal{S}$ satisfies property (P) if for any functions $g, h \in \mathcal{S}$ such that $h \not\gtrless f$ and $g \not\gtrless f$, we have that $\max(h, g) \not\gtrless f$.*

This definition is useful in understanding order-isomorphisms, and the next lemma explains how. It is Lemma 11 form [5] but we provide with the proof for completeness.

Lemma 3.3.5 (Lemma 11, [5])**.** *Let $T : \mathcal{S} \to \mathcal{S}$ be an order-preserving isomorphism. Then a function satisfying property (P) must be mapped to a function satisfying property (P).*

Proof. It follows in the same way as Fact 3.1.13 that for an order-preserving isomorphism on $\mathcal{S}$ we have that $T\max(h, g) = \max(Th, Tg)$. Assume that f satisfies (P), and let us show that so does Tf. Indeed, consider two functions, which we denote Tg and Th (which we may, as T is a bijection) and assume that they satisfy $Tg \not\geq Tf$ and $Th \not\geq Tf$. As T is an order-preserving we have that $g \not\geq f$ and $h \not\geq f$, and as f satisfies (P) this implies $\max(g, h) \not\geq f$. Using again that T is order-preserving, it follows that $T\max(g, h) \not\geq Tf$ and as $T\max(h, g) = \max(Th, Tg)$, the proof is complete. □

Our first step is thus determining the functions in $\mathcal{S}_t(\mathbb{R}^+)$ which satisfy (P), and this is the object of the next lemma.

Lemma 3.3.6. *The only functions in $\mathcal{S}_t(\mathbb{R}^+)$ which satisfy (P) are linear functions $l_a(x) = ax$ (for $a \geq 0$) and indicator functions $1^\infty_{[0,b]}(x)$ and $1^\infty_{[0,b)}(x)$.*

Proof. We start by showing that these function indeed satisfy (P).
(1) Assume $g \not\geq l_a$, then there is some $x_1 > 0$ for which $g(x_1) < ax_1$. By Lemma 3.2.11, this implies that $g(y) < ay$ for all $y \leq x_1$. Similarly if $h \not\geq l_a$ there is some $x_2 > 0$ for which $h(x_2) < ax_2$ and $h(y) < ay$ for all $y \leq x_2$. Therefore for all $y \leq \min(x_1, x_2)$ we have that $\max(g(y), h(y)) < ay = l_a(y)$ and so $\max(g, h) \not\geq l_a$ as needed.
(2) Assume $g \not\geq 1^\infty_{[0,a]}$. Then there is some $x_1 > a$ for which $g(x_1) < \infty$. Therefore, as functions in $\mathcal{S}_t$ are increasing on rays, for all $y \leq x_1$ we have $g(y) < \infty$. Similarly for $h \not\geq 1^\infty_{[0,a]}$ there is some $x_2 > a$ with $h(y) < \infty$ for all $y \leq x_2$. Therefore for all $y \leq \min(x_1, x_2)$ (which is $> a$) we have $\max(g(y), h(y)) < \infty$ and in particular $\max(g, h) \not\geq 1^\infty_{[0,a]}$, as needed.
(3) Assume $g \not\geq 1^\infty_{[0,a)}$. Then $g(a) < \infty$. Similarly if $h \not\geq 1^\infty_{[0,a)}$ then $h(a) < \infty$ and thus $\max(g(a), h(a)) < \infty$ and so $\max(g, h) \not\geq 1^\infty_{[0,a)}$.

Next, we would like to explain why a function satisfying (P) must be one of these three. Let f satisfy (P) and assume f attains some non-zero finite value. We claim that in such a case f must be linear. This will complete the proof as the only functions which attain only the values 0 and $+\infty$ in $\mathcal{S}_t$ are the two types of convex indicators mentioned above.

Assume that $0 < f(x_1) < \infty$. Denote $a = f(x_1)/x_1$, so that $f(x_1) = ax_1$. By Lemma 3.2.11 we know that the knife function $k_{x_1,a} \geq f$ and we have that $k_{x_1,a} = \max(1^\infty_{[0,x_1]}, l_a)$. Since $f(x_1) > 0$ we know that $1^\infty_{[0,x_1]} \not\geq f$. Therefore, as f satisfies (P), it must be that

$l_a \geq f$. However, as these two function intersect at $(x_1, f(x_1))$, by Lemma 3.2.11, for any $x > x_1$ we must also have that $f(x) = l_a(x)$ otherwise the knife function $k_{x,f(x)}$ would not be above the point $(x_1, f(x_1))$ on the epi-graph of f.

Next we need to explain why also in the segment $[0, x_1]$ the function must be linear. To this end, assume there is some point $0 < x_2 < x_1$ for which $0 < f(x_2) < ax_2$ (recall we have already shown that $l_a \geq f$, so this is the only other option). Using once again Lemma 3.2.11 we know that $k_{x_2,f(x_2)} \geq f$ so by the same argument as above, applied this time to x_2 and $a' = f(x_2)/x_2$, we get that $l_{a'} \geq f$ but this is clearly not true at the point x_1, which is a contradiction. Therefore on the segment $[0, x_1]$ we must have that either $f(x) = ax$, or $f(x) = 0$.

All we are left with is the option that f vanishes on some segment $[0, x]$ (or $[0, x)$) for $x > 0$ and is then linear. However, these functions are not in the class $\mathcal{S}_t$. Indeed, such a function is the pointwise minimum of $1^{\infty}_{[0,b]}$ and ℓ_a (or $1_{[0,b)}$ and ℓ_a, for which computations are analogous) and for it to be in the class $\mathcal{S}_t$ it would need to be the case that

$$\min(1^{\infty}_{[0,b]}, \ell_a) = \mathcal{A}_t\left(\mathcal{A}_t\left(\min(1^{\infty}_{[0,b]}, \ell_a)\right)\right).$$

But, using Fact 3.2.10 and Examples 3.2.2 and 3.2.3 we see that

$$\mathcal{A}_t\left(\mathcal{A}_t\left(\min(1^{\infty}_{[0,b]}, \ell_a)\right)\right) = \mathcal{A}_t\left(\max(\mathcal{A}_t 1^{\infty}_{[0,b]}, \mathcal{A}_t \ell_a)\right) = \mathcal{A}_t(\max(1^{\infty}_{[0,(1+t)/b)}, \ell_{1/a})$$

and it can be easily computed that this is equal to

$$\begin{cases} 0 & \text{for } x \in [0, b] \\ ax - \frac{ab}{1+t} & \text{for } x > b. \end{cases}$$

As this is clearly a different function, it follows that f cannot vanish on any segment $[0, x]$ for $x > 0$, which completes the proof of the Lemma. □

Lemma 3.3.7. *Let $T : \mathcal{S}_t(\mathbb{R}^+) \to \mathcal{S}_t(\mathbb{R}^+)$ be an order-preserving isomorphism. Then all linear functions $\ell_b(x) := bx$ (where $b > 0$) are mapped to linear functions, all convex closed-interval indicators $1^{\infty}_{[0,a]}$ are mapped to convex closed-interval indicators, and all convex half-open-interval indicators $1^{\infty}_{[0,a)}$ are mapped to convex half-open-interval indicators. In particular, there exist $u, w : \mathbb{R}_+ \to \mathbb{R}_+$ increasing bijections, such that $T(1^{\infty}_{[0,a]}) = 1^{\infty}_{[0,u(a)]}$ and $T(\ell_b) = \ell_{w(b)}$.*

Proof. By Lemma 3.3.5 and Lemma 3.3.6 we know that the class of all linear functions and all convex indicators (both closed and half-open) is invariant. Note that this class is in fact the union of two chains (with respect to the partial order). Indeed, the set of linear functions is ordered (according to the usual order of their slopes) and the set of indicators is ordered in that if $a < b$ then $1^\infty_{[0,a]} \leq 1^\infty_{[0,b)} \leq 1^\infty_{[0,b]}$. Moreover, no function in this class is between $1^\infty_{[0,b)}$ and $1^\infty_{[0,b]}$. Moreover, for $a > 0, b > 0$, ℓ_a and $1_{[0,b]}$ are not comparable, and neither are ℓ_a and $1_{[0,b)}$.

Note that an order-preserving isomorphism maps these chains to themselves. Indeed, since the chain of $(\ell_a)_{a>0}$ has the property that between any two elements there is a third one, it cannot be mapped to the chain of convex indicators. Thus it is invariant. Moreover, the chain of convex indicators is invariant. Finally, if a convex half-open-interval indicator $1^\infty_{[0,b)}$ was mapped to a closed one $1^\infty_{[0,c]}$, the image of $1^\infty_{[0,b]}$ would have to be an element φ of this class such that above $1^\infty_{[0,c]}$ and below φ there are no other elements, but no such element φ exists. Therefore, half-open-interval indicators are mapped to half-open-interval indicators and closed-interval indicators are mapped to closed-interval indicators, as claimed.

The functions $u, w : \mathbb{R}^+ \to \mathbb{R}^+$ defined by $T(1^\infty_{[0,a]}) = 1^\infty_{[0,u(a)]}$ and $T(\ell_b) = \ell_{w(b)}$ are increasing, by the order-preserving property of T. As their image is all of $\mathbb{R}^+$ (since T is an isomorophism) it is a bijection. □

Lemma 3.3.8. *Let $T : \mathcal{S}_t(\mathbb{R}^+) \to \mathcal{S}_t(\mathbb{R}^+)$ be an order-preserving isomorphism. Then knife functions are mapped to knife functions and escalators to escalators. Moreover, if we denote by $u, w : \mathbb{R}^+ \to \mathbb{R}^+$ the functions defined by*

$$T(1^\infty_{[0,a]}) = 1^\infty_{[0,u(a)]} \quad \text{and} \quad T(\ell_b) = \ell_{w(b)},$$

then $T(1^\infty_{[0,a)}) = 1^\infty_{[0,u(a))}$, $T(k_{a,b}) = k_{u(a),w(b)}$ and $T(s_{a,b}) = s_{\alpha,\beta}$ with $\alpha = \frac{1+t}{u((1+t)a^{-1})}$ and $\beta = \frac{w(ab)u((1+t)a^{-1})}{1+t}$.

Proof. By the same argument as above $T(1^\infty_{[0,a)})$ is the largest element in the invariant class of (P) property which is below $T(1^\infty_{[0,a]}) = 1^\infty_{[0,u(a)]}$, and this element is $1^\infty_{[0,u(a))}$.

Observe that $k_{a,b} = \max(1^\infty_{[0,a]}, \ell_b)$ and using that T is order-preserving it maps $\max(1^\infty_{[0,a]}, \ell_b)$ to $\max(T1^\infty_{[0,a]}, T\ell_b) = \max(1^\infty_{[0,u(a)]}, \ell_{w(b)})$.

Finally, Facts 3.2.5 and 3.2.10 show that the escalator function can be written as $s_{a,b} = \mathcal{A}_t(k_{a,(ab)^{-1}}) = \mathcal{A}_t(\max(1^\infty_{[0,a]}, \ell_{(ab)^{-1}}) = \hat{\inf}\{1_{[0,\frac{1+t}{a})}, \ell_{ab}\}$. Then as before we see that T maps $\hat{\inf}(1^\infty_{(0,\frac{(1+t)}{a})}, \ell_{ab})$ to

$$\hat{\inf}(1^\infty_{(0,u(\frac{(1+t)}{a}))}, \ell_{w(ab)}) = s_{\frac{1+t}{u((1+t)a^{-1})}, \frac{w(ab)u((1+t)a^{-1})}{1+t}}.$$

□

We next claim that the maps u and w given in Lemma 3.3.8 are of a special linear form.

Lemma 3.3.9. *Let $T : \mathcal{S}_t(\mathbb{R}^+) \to \mathcal{S}_t(\mathbb{R}^+)$ be an order-preserving isomorphism, and assume that $T(1^{\infty}_{[0,a]}) = 1^{\infty}_{[0,u(a)]}$ and $T(l_b) = l_{w(b)}$ for some increasing bijections $u, w : \mathbb{R}_+ \to \mathbb{R}_+$. Then there exist $\alpha, \beta \in \mathbb{R}^+$ such that $u(a) = \alpha a$ and $w(b) = \beta b$, in particular, $T(k_{a,b}) = k_{\alpha a,\beta b}$.*

Proof. Let $k_{a,b}$ be some knife function with $a, b > 0$, and note that any escalator function of the form
$$s_\lambda(x) := s_{\frac{1+t}{\lambda a},\frac{\lambda ab}{1+t-\lambda}}(x) = \frac{\lambda ab}{1+t-\lambda}\left(\frac{1+t}{\lambda a}\cdot x - 1\right)_t$$
where $0 < \lambda \le 1$, is strictly below $k_{a,b}$, i.e. $s_\lambda \le k_{a,b}$. Indeed, $s_\lambda(x) = 0$ for all $x < \lambda a$ and for $x \ge \lambda a$ it is an increasing function with $s_\lambda(a) = ab = k_{a,b}(a)$.

Since T is order-preserving we must have $Ts_\lambda \le Tk_{a,b}$ and as the two functions intersected at a point (a, ab) we must have that for any $\lambda \in (0, 1]$,
$$Ts_\lambda(u(a)) = Tk_{a,b}(u(a)).$$
Now, using Lemma 3.3.8, we get that $Tk_{a,b}(u(a)) = k_{u(a),w(b)}(u(a)) = u(a)w(b)$ and
$$Ts_\lambda(u(a)) = s_{\frac{1+t}{u(\lambda a)}(u(a)),\frac{u(\lambda a)w\left(\frac{(1+t)b}{1+t-\lambda}\right)}{1+t}}(u(a)) = \frac{w\left(\frac{(1+t)b}{1+t-\lambda}\right)}{1+t}\left((1+t)u(a) - u(\lambda a)\right).$$
As a result the equality $Ts_\lambda(u(a)) = Tk_{a,b}(u(a))$ can be written as
$$\left(u(a) - \frac{u(\lambda a)}{1+t}\right) w\left(\frac{(1+t)b}{1+t-\lambda}\right) = u(a)w(b).$$
Since $u(a)w(b) \neq 0$ (as $k_{a,b}$ is not the function $1^{\infty}_{\{0\}}$), after rearranging we see that the ratio $w(b)/w(\frac{b}{1-\frac{\lambda}{1+t}})$ does not depend on b and similarly $u(\lambda a)/u(a)$ does not depend on a. Indeed, we get that equality
$$1 - \frac{u(\lambda a)}{(1+t)u(a)} = \frac{w(b)}{w\left(\frac{b}{1-\frac{\lambda}{1+t}}\right)} \tag{3.1}$$
needs to hold for all a, b and hence be constant. Note that neither $w(1)$ or $u(1)$ is equal zero, so we may write

$$w\left(\frac{b}{1-\frac{\lambda}{1+t}}\right) = w(b)\frac{w\left(\frac{1}{1-\frac{\lambda}{1+t}}\right)}{w(1)} \quad \text{and} \quad u(\lambda a) = u(a)\frac{u(\lambda)}{u(1)}.$$

Note, that these equalities formally hold for $\lambda \in (0,1]$ but can be extended. In the case of u, for $\lambda > 1$ let $a' = \lambda a$ then

$$\frac{u(\lambda a)}{u(a)} = \left(\frac{u((1/\lambda)a')}{u(a')}\right)^{-1} = \left(\frac{u(1/\lambda)}{u(1)}\right)^{-1} = \frac{u((1/\lambda)\lambda)}{u(1/\lambda)} = \frac{u(\lambda)}{u(1)},$$

where we used the previous identity twice, with $\frac{1}{\lambda} \in (0,1)$. Hence, we can allow all $\lambda \in (0,+\infty)$. In the case of w, we see that the parameter $\eta := \frac{1}{1-\frac{\lambda}{1+t}}$ belongs to the interval $(1, 1+\frac{1}{t}]$ and also that the equality holds trivially for $\eta = 1$. Moreover, arguing in the same way as for u, we see that in fact the equality holds for all $\eta \in [\frac{t}{1+t}, \frac{1+t}{t}]$.

Let $U(x) = u(x)/u(1)$ and $W(x) = w(x)/w(1)$. Clearly they are non-negative monotone and continuous functions (as u and w are increasing bijections from $\mathbb{R}_+$ to $\mathbb{R}_+$), and we see that

$$\begin{aligned} &\forall x \in \mathbb{R}_+ \ \forall y \in (0,\infty) && U(xy) = U(x)U(y) \\ &\forall x \in \mathbb{R}_+ \ \forall y \in [\tfrac{t}{1+t}, \tfrac{1+t}{t}] && W(xy) = W(x)W(y). \end{aligned}$$

These are multiplicative Cauchy functional equations, where the first one is classical and the second has restricted domain. We include a proof for this kind of restricted domain in the appendix (Lemma B.0.3). It follows that there exist constants $c, d \in \mathbb{R}$ such that $U(x) = x^c$ and $W(x) = x^d$. Thus, letting $u(1) = \alpha, w(1) = \beta$ with $\alpha, \beta > 0$ we have

$$u(x) = \alpha x^c, \quad w(x) = \beta x^d.$$

Plugging this into the equation (3.1) gives for all $\lambda \in (0,1]$ that

$$1 - \frac{\lambda^c}{1+t} = \left(1 - \frac{\lambda}{1+t}\right)^d.$$

This equation can hold for all $0 < \lambda \leq 1$ only if $c = d = 1$, and so $u(a) = \alpha a$ and $w(b) = \beta b$. $\square$

Theorem 3.3.10. *An order-preserving isomorphism $T : \mathcal{S}_t(\mathbb{R}^+) \to \mathcal{S}_t(\mathbb{R}^+)$ must behave like the identity, i.e. for any function $f \in \mathcal{S}_t(\mathbb{R}^+)$ there exist $\alpha, \beta \in \mathbb{R}^+$ such that*

$$Tf(x) = \beta f\left(\frac{x}{\alpha}\right).$$

Proof. It is clear from the Lemma 3.2.11 that every function in $\mathcal{S}_t$ can be written as the infimum of knife functions and that indeed, $f(x) = \inf_y (k_{y,f(y)/y})(x)$. Then by Lemma 3.3.9 we get

$$\begin{aligned} Tf(x) &= T\Big(\inf_y (k_{y,f(y)/y})\Big)(x) = \inf_y \Big(Tk_{y,f(y)/y}\Big)(x) \\ &= \inf_y \Big(k_{\alpha y,\beta f(y)/y}\Big)(x) = \inf_z \Big(k_{z,\beta f(\frac{z}{a})/\frac{z}{a}}\Big)(x) \\ &= \beta f\Big(\tfrac{x}{\alpha}\Big) \end{aligned}$$

as claimed. □

Corollary 3.3.11. *The only order-reversing isomorphism on the class $\mathcal{S}_t(\mathbb{R}_+)$ is, up to a linear transformation, $\mathcal{A}_t$.*

3.3.2 Generalization to $\mathcal{S}_t(\mathbb{R}^n)$

Having understood how an order-preserving involution behaves on $\mathcal{S}_t(\mathbb{R}^+)$ we are ready to consider higher dimensions. Let us first recall a result from [5, Lemma 9] about order-preserving isomorphisms on the class $\mathcal{S}_0$. One may check that, as the proof only uses the properties of the map T and the support of functions in the class, it is enough to use Lemma 3.2.13 to see that the exact same result extends to the functions in $\mathcal{S}_t$ for $t > 0$.

Lemma 3.3.12. *If $T : \mathcal{S}_t(\mathbb{R}^n) \to \mathcal{S}_t(\mathbb{R}^n)$ is an order-preserving isomorphism, then there is a bijection $\Phi : S^{n-1} \to S^{n-1}$ such that any function supported on $\mathbb{R}^+ y$ is mapped to a function supported on $\mathbb{R}^+ z$ for $z = \Phi(y)$. Moreover, the function $1^\infty_{\mathbb{R}^+ y}$ is mapped to the function $1^\infty_{\mathbb{R}^+ z}$.*

Therefore we know that any order-preserving bijection must act ray-wise, and in this subsection we will determine how all the rays fit together. This proof again follows almost identically [5, Subsection 6.7.].

Proof of Theorem 3.3.2. Let T be an order-preserving isomorphism on $\mathbb{R}^n$, $n \geq 2$. From Lemma 3.3.12 we have a function $\Phi : S^{n-1} \to S^{n-1}$ such that the function supported on $\mathbb{R}^+ y$ is mapped to the function supported on $\mathbb{R}^+ z$ for $z = \Phi(y)$ and $1^\infty_{\mathbb{R}^+ y}$ is mapped to the function $1^\infty_{\mathbb{R}^+z}$. Hence, for each $y \in S^{n-1}$ we have a mapping from $\mathcal{S}_t(\mathbb{R}^+y)$ (which is isometric, say via I_1, to $\mathcal{S}_t(\mathbb{R}^+)$) to $\mathcal{S}_t(\mathbb{R}^+\Phi(y))$ (again isometric, say via I_2, to $\mathcal{S}_t(\mathbb{R}^+)$). Hence, from Theorem 3.3.10 we find that for each ray $\mathbb{R}^+y$ there are constants α_y, β_y such that if $x \in \mathbb{R}^+y$ then

$$\Big((I_2 \circ T \circ I_1^{-1})(f)\Big)(x) = \beta_y f\left(\frac{x}{\alpha_y}\right).$$

Using this, let us now check how T acts on indicator functions in $\mathcal{S}_t(\mathbb{R}^n)$. To this end, define a function $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ by $T(1^\infty_{[0,x]}) = 1^\infty_{[0,\varphi(x)]}$, and it is well defined due to its connection to Φ. This means that for any convex set $0 \in K \subset \mathbb{R}^n$ we have

$$T1^\infty_K = T\Big(\hat{\inf}_{x\in K}(1^\infty_{[0,x]})\Big) = \hat{\inf}_{x\in K}(T1^\infty_{[0,x]}) = \hat{\inf}_{x\in K}(1^\infty_{[0,\varphi(x)]}) = 1^\infty_{\varphi(K)}.$$

Therefore, since T preserves convexity as a bijective interval preserving map, it follows that also φ preserves convexity (see [2]). Hence, by well known result due to Schneider [19], we conclude that φ is linear, i.e. $\varphi(x) = Bx$ for $B \in \mathrm{GL}_n(\mathbb{R})$.

Now let us consider a function $f = \hat{\inf}(k_{x,b}, k_{x',b'}) \in \mathcal{S}_t$. We have that $f|_{[0,x]} = k_{x,b}$ and $f|_{[0,x']} = k_{x',b'}$ for $x \neq x'$. Then by Lemma 3.2.13 follows, that f needs to be finite on the whole set $\mathrm{conv}\{0, x, x'\}$, as it is contained in the $\max(xb, x'b')$-level set, but it will be infinite otherwise. Moreover, we can find a linear function ℓ which coincides with $k_{x,b}$ and $k_{x',b'}$ on their domains and hence on $\mathrm{conv}\{0, x, x'\}$ can be expressed as $\ell(t(\lambda x + (1-\lambda)x')) = t(\lambda b + (1-\lambda)b')$ for $\lambda, t \in [0, 1]$. Since the class $\mathcal{S}_t$ is closed under supremum, we get that $\hat{\inf}(k_{x,b}, k_{x',b'}) = \max\{\ell, 1^\infty_{\mathrm{conv}\{0,x,x'\}}\}$.

Therefore the largest knife function with domain $[0, \lambda x + (1-\lambda)x']$, which is smaller than $\hat{\inf}(k_{x,b}, k_{x',b'})$ (and in this case is equal to it) is $k_{\lambda x+(1-\lambda)x', \lambda b+(1-\lambda)b'}$ and we have that

$$k_{\lambda x+(1-\lambda)x',\lambda b+(1-\lambda)b'} = \lambda k_{x,b} + (1-\lambda)k_{x',b'}. \tag{3.2}$$

The three knife functions $k_{x,b}$, $k_{x',b'}$, $k_{\lambda x+(1-\lambda)x',\lambda b+(1-\lambda)b'}$ are mapped by T to knife functions with bases $[0, Bx], [0, Bx']$ and $[0, \lambda Bx + (1-\lambda)Bx']$ and slopes which a priori depend on the base, hence we denote them by $u(x, b), u(x', b')$ and $u(\lambda x + (1-\lambda)x', \lambda b + (1-\lambda)b')$, respectively. Then the equality 3.2 gives us

$$u(\lambda(x, b) + (1-\lambda)(x', b')) = \lambda u(x, b) + (1-\lambda)u(x', b'),$$

which means that u is linear. Noting that $u(x, 0) = 0$ for all x we get that $u(x, b) = \beta_0 b$. This yields that the the function f is mapped to $\beta_0 f \circ B$, and the result can be extended to all $\mathcal{S}_t$ class functions due to Lemma 3.2.11. □

3.4 The associated subgradient mapping

It will be useful, before we consider some special cases, to recall and study a notion of a *c-subgradient mapping* for a general cost function $c : X \times Y \to (-\infty, \infty]$. As we have seen already, a cost function induces an order-reversing transformation $\varphi \to \varphi^c$ and a class of functions which we called the c-class, namely all functions which are images of this mapping, on which the mapping is an order-reversing isomorphism. No less important is the c-subgradient mapping, which is induced by the cost function and is a set valued mapping from X to Y (or vice versa).

Definition 3.4.1 (c-subgradient)**.** *Given $\varphi : X \to [-\infty, \infty]$ which is a c-class function, we consider the subset $\partial^c\varphi \subset X \times Y$, and its sections $\partial^c\varphi(x) = \{y \in Y : (x, y) \in \partial^c\varphi\}$ and $\partial^c\varphi^c(y) = \{x \in X : (x, y) \in \partial^c\varphi\}$, where*

$$\partial^c\varphi = \{(x, y) : \varphi(x) + \varphi^c(y) = c(x, y) < \infty\}.$$

We call the section $\partial^c\varphi(x)$ the c-subgradient of φ at $x \in X$, and analogously $\partial^c\varphi^c(y)$ the c-subgradient of φ^c at $y \in Y$.

By the symmetry of the definition, we see that

$$x \in \partial^c\varphi^c(y) \quad \Leftrightarrow \quad y \in \partial^c\varphi(x).$$

Some examples are in order here, in particular for $c(x, y) = -\langle x, y\rangle$ it is clear that $\partial^c\varphi$ is the classical superdifferential of a concave function φ. Indeed, by the definition of the c-transform, $\varphi(x) + \varphi^c(y) = -\langle x, y\rangle$ holds if and only if for all $z \in X$ we have that $-\langle x, y\rangle \le \varphi(x) - \varphi(z) - \langle z, y\rangle$.

For p_0, the polar cost, the c-subgradient is the polar-subgradient of the geometric convex function $e^{-\varphi}$ which was discussed in [6]. We shall get back to this case when we discuss the c-subgradient for the class of costs p_t below.

First, let us establish some general properties of c-subgradients.

Lemma 3.4.2. *Let $c : X \times Y \to (-\infty, \infty]$ be a cost function and consider a basic function $\varphi(x) = c(x, y_0) + \beta$ for some $y_0 \in Y$. Then, if $c(x, y_0) < \infty$, the c-subgradient of φ at x contains y_0, i.e. $y_0 \in \partial^c\varphi(x)$.*

Proof. Indeed, let φ be as in the statement. From the definition it follows that $y \in \partial^c\varphi(x)$ is and only if $c(x, y) < \infty$ and

$$c(x, y) - \varphi(x) = \varphi^c(y) = \inf_z(c(z, y) - \varphi(z)).$$

Plugging in the definition of φ we see that $y = y_0$ always satisfies the equality. $\square$

Since all functions in the c-class can be built using basic functions, we can see that in fact the c-subgradient of a function φ at a point x has a very simple geometric meaning: As the function φ is in the c-class, it is the image, under the c-transform, of another c-class function $\psi = \varphi^c$. Therefore, it can be written as an infimum over basic functions as follows:

$$\varphi(x) = \inf_y \left(c(x,y) - \varphi^c(y)\right).$$

All the functions on the right hand side lie above φ. If any one of the basic functions (indexed by y) on the right hand side is tangent to φ at the point x, then the pair (x,y) belongs to $\partial^c\varphi$, and $y \in \partial^c\varphi(x)$. In other words

Lemma 3.4.3. *Let φ be a c-class function, and $x \in X$ and assume that $\varphi(x) < \infty$. Then $y_0 \in \partial^c\varphi(x)$ if and only if $c(x,y_0) < \infty$ and the function $\ell(z) = c(z,y_0) - c(x,y_0) + \varphi(x)$ satisfies*

$$\ell(z) \geq \varphi(z) \quad \textit{for all} \quad z \in X.$$

Proof. By the definition we have that $y_0 \in \partial^c\varphi(x)$ if and only if

$$\varphi(x) + \varphi^c(y_0) = c(x,y_0) < \infty.$$

Using the definition of the c-transform we see that

$$\varphi(x) = c(x,y_0) - \varphi^c(y_0) = \sup_z (c(x,y_0) - c(z,y_0) + \varphi(z)),$$

which holds if and only if for all z we have $c(z,y_0) - c(x,y_0) + \varphi(x) \geq \varphi(z)$. $\square$

Looking again at Lemma 3.4.3 we see that if we are after finding all $y \in \partial^c\varphi(x)$ we are in fact trying to find all y for which $z \mapsto c(z,y) - \varphi(z)$ has a minimum which is attained at x.

3.4.1 A geometric interpretation for p_t

In the case of p_t it is convenient to consider the multiplicative setting, in which the role of the transform $\mathcal{A}_t$ is more apparent. Therefore, let us note that the definition of the p_t-subgradient can be rewritten, for $\varphi \in \mathcal{S}_t$, as

$$\partial^t\varphi = \{(x,y) : \varphi(x)\mathcal{A}_t\varphi(y) = \langle x,y\rangle - 1 \geq t\}.$$

For simplicity we write ∂^t rather than ∂^{p_t}, and in the case $t = 0$ we denote the *polar subgradient* by ∂°. Note that we are slightly abusing notation as here $\varphi \in \mathcal{S}_t$ while in Definition 3.4.1 we have a p_t-subgradient for a p_t-class function, i.e. minus the logarithm of a $\mathcal{S}_t$ function. However, it should be clear from the context whether we are dealing with the additive or multiplicative level.

We can rewrite Lemma 3.4.3, using the following notation $Z_\varphi = \{x : \varphi(x) = 0\}$ and $\mathrm{dom}(\varphi) = \{x : \varphi(x) < \infty\}$, as

Lemma 3.4.4. *Fix $t \geq 0$. Let $\varphi = \mathcal{A}_t\psi$ and $\psi = \mathcal{A}_t\varphi$, and let $x \in \mathrm{dom}(\varphi) \setminus Z_\varphi$. Then $y_0 \in \partial^t\varphi(x)$ if and only if the function $\ell(z) = \frac{1}{\psi(y_0)}\left(\langle z, y_0\rangle - 1\right)_t$ satisfies $\ell(z) \leq \varphi(z)$ for all z and $\ell(x) = \varphi(x) \neq 0$.*

Even though this lemma is a corollary, let us provide a direct proof.

Proof. Let $\varphi \in \mathcal{S}_t$ and assume that the function $\ell(z) = \frac{1}{\psi(y_0)}\left(\langle z, y_0\rangle - 1\right)_t$ with $\psi = \mathcal{A}_t\varphi$, is such that $\ell(z) \leq \varphi(z)$ for all z and that the equality holds at x. In order to prove that $y_0 \in \partial^t\varphi(x)$ it is enough to show that $\varphi(x)(\mathcal{A}_t\varphi)(y_0) = (\langle x, y_0\rangle - 1)_t$. Note that since $\varphi(x) \neq 0$ we have that $\langle x, y_0\rangle \geq 1 + t$. Writing out the definition of $\mathcal{A}_t$ we see that this is equivalent to

$$\sup_{\{z:\langle z,y_0\rangle \geq 1+t\}} \frac{1}{\varphi(z)}(\langle z, y_0\rangle - 1) = \frac{1}{\varphi(x)}(\langle x, y_0\rangle - 1).$$

The inequality "$\geq$" always holds, as x participates in the supremum, and therefore it remains to show that for every z with $\langle z, y_0\rangle \geq 1 + t$ we have

$$\frac{1}{\varphi(z)}(\langle z, y_0\rangle - 1) \leq \frac{1}{\varphi(x)}(\langle x, y_0\rangle - 1).$$

But since $\ell(x) = \varphi(x)$, we know that $\frac{1}{\psi(y_0)}(\langle x, y_0\rangle - 1) = \varphi(x)$ or, in other words, that $\psi(y_0) = \frac{\langle x,y_0\rangle - 1}{\varphi(x)}$. So, rearranging the inequality above, we see that we need to show that

$$\frac{1}{\psi(y_0)}(\langle z, y_0\rangle - 1)_t \leq \varphi(z)$$

which is always true by the assumption that $\ell(z) \leq \varphi(z)$. It follows that $y_0 \in \partial^t\varphi(x)$ as claimed.

For the other direction, assume that $y_0 \in \partial^t\varphi(x)$, which means that $\langle x, y_0\rangle \geq 1 + t$ and that $\varphi(x)\psi(y_0) = \langle x, y_0\rangle - 1 \geq t$. From this follows that

$$\varphi(x) = \frac{1}{\psi(y_0)}\left(\langle x, y_0\rangle - 1\right)_t = \ell(x).$$

To show that $\ell(z) \le \varphi(z)$, we use the definition of $\mathcal{A}_t$ and we get that

$$\varphi(z) = (\mathcal{A}_t\psi)(z) \ge \frac{1}{\psi(y_0)} \left(\langle z, y_0\rangle - 1\right)_t = \ell(z).$$

□

In case of the polar cost p_0 and the associated polarity transform $\mathcal{A}$, there is a straightforward connection between the polar subgradient and the classical subgradient. This is captured by the following lemma, which is a version of [6, Lemma 3.3].

Lemma 3.4.5. *Let $\varphi \in \mathrm{Cvx}_0(\mathbb{R}^n)$ and let $x \in \mathrm{dom}(\varphi) \setminus Z_\varphi$. Then*

(i) for any $z \in \partial\varphi(x)$ such that $\langle x, z\rangle \ne \varphi(x)$, we have that $y = \frac{z}{\langle x,z\rangle - \varphi(x)} \in \partial^\circ\varphi(x)$,

(ii) for any $y \in \partial^\circ\varphi(x)$ there exists some $z \in \partial\varphi(x)$ such that $\langle x, z\rangle \ne \varphi(x)$ and such that $y = \frac{z}{\langle x,z\rangle - \varphi(x)}$.

Proof. *(i)* To show that a point y lies in the polar gradient at x we need that $\varphi(x)(\mathcal{A}\varphi)(y) = \langle x, y\rangle - 1 > 0$, which is equivalent to $\langle x, y\rangle - 1 > 0$ together with the fact that for every w with $\langle w, y\rangle > 1$ and $\varphi(w) > 0$, we have

$$\frac{\langle w, y\rangle - 1}{\varphi(w)} \le \frac{\langle x, y\rangle - 1}{\varphi(x)}.$$

Recall that our y is $\frac{z}{\langle x,z\rangle - \varphi(x)}$ for some $z \in \partial\varphi$ and that it is well defined as $\varphi(x) \ne \langle x, z\rangle$. We rearrange the inequality we wish to prove, plugging in the definition of y, to get

$$\frac{\langle w, \frac{z}{\langle x,z\rangle - \varphi(x)}\rangle - 1}{\varphi(w)} \le \frac{\langle x, \frac{z}{\langle x,z\rangle - \varphi(x)}\rangle - 1}{\varphi(x)} = \frac{1}{\langle x, z\rangle - \varphi(x)}$$

which is the same as

$$\varphi(x) + \langle w - x, z\rangle \le \varphi(w)$$

and the latter is clearly true as $z \in \partial\varphi(x)$.
(ii) Given $y \in \partial^\circ\varphi(x)$ it follows from the definition that $\langle x, y\rangle > 1$. Consider

$$z = \frac{y\varphi(x)}{\langle y, x\rangle - 1},$$

which is well defined, and also implies that $y = \frac{z}{\langle x,z\rangle - \varphi(x)}$. We need to show that $z \in \partial\varphi(x)$ and $\langle z, x\rangle \ne \varphi(x)$.

The latter follows easily since $\langle z,x\rangle = \varphi(x)\Big(1+\frac{1}{\langle x,y\rangle-1}\Big)$ and once again that $\langle x,y\rangle > 1$. For the former, we use as before that $y \in \partial^\circ\varphi(x)$ to conclude that for any w with $\langle w,y\rangle > 1$ and $\varphi(w) > 0$ we have

$$\frac{\langle w,y\rangle - 1}{\varphi(w)} \leq \frac{\langle x,y\rangle - 1}{\varphi(x)}.$$

Plugging in y and rearranging, we get that

$$\varphi(x) + \langle w-x, z\rangle \leq \varphi(w)$$

for any w such that $\langle w,z\rangle > \langle x,z\rangle - \varphi(x)$ and $\varphi(w) > 0$. In the case when w is such that $\langle w,z\rangle \leq \langle x,z\rangle - \varphi(x)$, this actually means that $\varphi(x) + \langle w-x,z\rangle \leq 0$ and since the geometric convex functions are non-negative the desired inequality trivially follows.

It remains to consider the case when $w \in Z_\varphi$, i.e. when $\varphi(w) = 0$. But then, plugging in the previously defined z, we have that the inequality defining the subgradient of φ at x becomes just

$$\langle w,y\rangle \leq 1.$$

That is, we need to show that $\partial^\circ\varphi(x)$ is contained in the polar set of Z_φ. Indeed, $y \in \partial^\circ\varphi(x)$ implies in particular that $y \in \mathrm{dom}(\mathcal{A}\varphi)$ (since the value of $\mathcal{A}\varphi(y) = \frac{\langle x,y\rangle-1}{\varphi(x)} < \infty$) and it follows from the definition of $\mathcal{A}$ that $\mathrm{dom}(\mathcal{A}\varphi) \subset Z_\varphi^\circ$, which completes the proof. $\square$

The next proposition shows, rather surprisingly, that even though the polar subgradient at any point x is a convex set in Y, this is not true for the p_t-subgradient with $t > 0$.

Proposition 3.4.6. *Let $\varphi : \mathbb{R}^n \to [0,\infty]$ be an $\mathcal{A}_t$-class function and let $x \in \mathrm{dom}(\varphi) \setminus Z_\varphi$ such that $\partial^t\varphi(x)$ is non-empty. If $t = 0$, then the set $\partial^\circ\varphi(x)$ is convex. Furthermore, if $n = 1$ then for any $t > 0$ the set $\partial^t\varphi(x)$ is convex. However, for $n \geq 2$ and $t > 0$, there exists $s \in (0,1)$ and an $\mathcal{A}_t$-class function φ such that $y_0 \neq y_1 \in \partial^t\varphi(x)$ but $(1-s)y_0 + sy_1 \notin \partial^t\varphi(x)$.*

Proof. For $t \geq 0$, let $\varphi \in \mathcal{S}_t$ and let $\psi = \mathcal{A}_t\varphi$. Assume that $y_0 \neq y_1$ and that $y_i \in \partial^t\varphi(x)$ for $i = 0,1$. We will check when, for all $s \in (0,1)$ we have that $(1-s)y_0 + sy_1 \in \partial^t\varphi(x)$.

Without loss of generality we can assume that $\langle x,y_0\rangle \leq \langle x,y_1\rangle$. It follows from the definition of the subgradient that $\langle x,y_i\rangle \geq 1+t$ for $i = 0,1$. By Lemma 3.4.4 there exist functions

$$\ell_i(z) = \frac{1}{\psi(y_i)}(\langle z,y_i\rangle - 1)_t$$

such that $\ell_i(z) \leq \varphi(z)$ for all z and $\ell_i(x) = \varphi(x)$.

For $s \in (0,1)$ define $y_s = (1-s)y_0 + sy_1$ and consider the functions

$$\ell_s(z) = \frac{1}{\psi(y_s)}(\langle z, y_s\rangle - 1)_t.$$

If for all $0 < s < 1$, we have that $\ell_s(z) \le \varphi(z)$ for all z and $\ell_s(x) = \varphi(x)$, then $\partial^t\varphi(x)$ is a convex set.

Let us first note that

$$\psi(y_s) = \mathcal{A}_t\varphi(y_s) = \sup_{\{z:\langle z,y_s\rangle \ge 1+t\}} \left((1-s)\frac{\langle z, y_0\rangle - 1}{\varphi(z)} + s\frac{\langle z, y_1\rangle - 1}{\varphi(z)} \right).$$

Since both of the functions on the right hand side achieve the supremum at x (this follows from the fact that $y_0, y_1 \in \partial^t\varphi(x)$ and $\langle x, y_s\rangle = (1-s)\langle x, y_0\rangle + s\langle x, y_1\rangle \ge 1+t$), we get that

$$\psi(y_s) = (1-s)\psi(y_0) + s\psi(y_1).$$

Now note that since both ℓ_i are below φ we have that $\max(\ell_0(z), \ell_1(z)) \le \varphi(z)$ for all z. Since the $\mathcal{S}_t$ class is closed under maximum, and therefore a priori φ could be equal to the $\max_i(\ell_i)$, we need to show that $\ell_s(z) \le \max(\ell_0(z), \ell_1(z))$. We need to consider four cases
(i) Assume that z is such that $\max(\langle z, y_0\rangle, \langle z, y_1\rangle) < 1+t$. It is clear that then also $\langle z, y_s\rangle < 1+t$ and hence $\ell_s(z) = \ell_0(z) = \ell_1(z) = 0$.
(ii) Assume that z is such that $\min(\langle z, y_0\rangle, \langle z, y_1\rangle) \ge 1+t$. Then consider a function

$$\ell_\lambda(z) = (1-\lambda)\ell_0(z) + \lambda\ell_1(z).$$

It is clear that for all $\lambda \in (0,1)$ we have that $\ell_\lambda(z) \le \max(\ell_0(z), \ell_1(z))$. For our z, plugging in the definitions of ℓ_i we get that

$$\ell_\lambda(z) = \left(\frac{(1-\lambda)}{\psi(y_0)} + \frac{\lambda}{\psi(y_1)}\right) \cdot \left(\langle z, \frac{\frac{(1-\lambda)y_0}{\psi(y_0)} + \frac{\lambda y_1}{\psi(y_1)}}{\frac{(1-\lambda)}{\psi(y_0)} + \frac{\lambda}{\psi(y_1)}}\rangle - 1\right)$$

Hence, taking $s = s_\lambda = \frac{\lambda\psi(y_0)}{(1-\lambda)\psi(y_1)+\lambda\psi(y_0)}$ (considering all $\lambda \in (0,1)$ this is just a reparametrization) we see that $\ell_s(z) = \ell_\lambda(z)$ and the claim follows.
(iii) Assume that z is such that $\langle z, y_0\rangle \ge 1+t$ and $\langle z, y_1\rangle < 1+t$. Then the $\max_{i=0,1}(\ell_i) = \ell_0$. As in the previous case, let $s = \frac{\lambda\psi(y_0)}{(1-\lambda)\psi(y_1)+\lambda\psi(y_0)}$. Then $\ell_0(z) - \ell_s(z)$ is equal to

$$\frac{1}{\psi(y_0)}(\langle z, y_0\rangle - 1) - \frac{(1-\lambda)\psi(y_1)+\lambda\psi(y_0)}{\psi(y_0)\psi(y_1)}\Big(\langle z, \frac{(1-\lambda)\psi(y_1)}{(1-\lambda)\psi(y_1)+\lambda\psi(y_0)}y_0 + \frac{\lambda\psi(y_0)}{(1-\lambda)\psi(y_1)+\lambda\psi(y_0)}y_1\rangle - 1\Big),$$

which is the same as

$$\begin{aligned}&\langle z, (\tfrac{1}{\psi(y_0)} - \tfrac{1-\lambda}{\psi(y_0)})y_0\rangle - \langle z, \tfrac{\lambda}{\psi(y_1)}y_1\rangle + \tfrac{(1-\lambda)\psi(y_1)+\lambda\psi(y_0)}{\psi(y_0)\psi(y_1)} - \tfrac{1}{\psi(y_0)}\\ &= \tfrac{\lambda}{\psi(y_0)}(\langle z, y_0\rangle - 1) - \tfrac{\lambda}{\psi(y_1)}(\langle z, y_1\rangle - 1).\end{aligned}$$

This expression is non-negative if and only if $\frac{\psi(y_1)}{\psi(y_0)} \geq \frac{\langle z,y_1\rangle-1}{\langle z,y_0\rangle-1}$. But $\psi(y_i) = \frac{\langle x,y_i\rangle-1}{\varphi(x)}$, since $y_i \in \partial^t\varphi(x)$. Therefore, we get that $\ell_0(z) \geq \ell_s(z)$ for z such that $\langle z, y_0\rangle \geq 1+t$ and $\langle z, y_1\rangle < 1+t$, if and only if

$$\frac{\langle x, y_1\rangle - 1}{\langle x, y_0\rangle - 1} \geq \frac{\langle z, y_1\rangle - 1}{\langle z, y_0\rangle - 1}.$$

But by the restriction on z, it follows that $\frac{\langle z,y_1\rangle-1}{\langle z,y_0\rangle-1} < \frac{t}{t} = 1$ and since $\langle x, y_0\rangle \leq \langle x, y_1\rangle$ we get that $\frac{\langle x,y_1\rangle-1}{\langle x,y_0\rangle-1} \geq 1$. Hence indeed $\ell_0(z) \geq \ell_s(z)$.
(iv) Finally, if z is such that $\langle z, y_0\rangle < 1+t$ and $\langle z, y_1\rangle \geq 1+t$. Similarly, we consider the difference $\ell_1(z) - \ell_s(z)$ which turns out to be equal to

$$\tfrac{1-\lambda}{\psi(y_1)}(\langle z, y_1\rangle - 1) - \tfrac{1-\lambda}{\psi(y_0)}(\langle z, y_0\rangle - 1).$$

This expression is non-negative if and only if $\frac{\psi(y_0)}{\psi(y_1)} \geq \frac{\langle z,y_0\rangle-1}{\langle z,y_1\rangle-1}$, which is the same as $\frac{\langle x,y_0\rangle-1}{\langle x,y_1\rangle-1} \geq \frac{\langle z,y_0\rangle-1}{\langle z,y_1\rangle-1}$. Since the denominators are non-negative, it turns out that this condition is equivalent to

$$\langle z, y_1 - y_0\rangle(\langle x, y_0\rangle - 1) + \langle x, y_1 - y_0\rangle(1 - \langle z, y_0\rangle) \geq 0. \tag{3.3}$$

Note that the first component is non-negative since $\langle z, y_1 - y_0\rangle \geq 0$ and the expression in the bracket is at least $t \geq 0$. Moreover, it is true that $\langle x, y_1 - y_0\rangle \geq 0$. However, the expression $1 - \langle z, y_0\rangle$, since we have that $\langle z, y_0\rangle < 1+t$, is greater or equal to $-t$. Therefore, if $t = 0$ then the claim follows and $\partial^\circ\varphi(x)$ is convex. Moreover, it can be easily checked that in $\mathbb{R}$ the same is true for any $t > 0$.

However, in $\mathbb{R}^n$ with $n \geq 2$ one can choose x, y_0, y_1 and z such that the inequality (3.3) is violated. In other words, there exists a function $\varphi(z) = \max\{\ell_0, \ell_1\}$, whose p_t-subgradient at x is not convex. Indeed, it is enough to choose x, y_0, y_1, z in such a way that the following inequalities hold: $\langle x, y_1\rangle \geq \langle x, y_0\rangle \geq 1+t$, $\langle z, y_1\rangle \geq 1+t$, $\langle z, y_0\rangle < 1+t$ and $\langle z, y_1 - y_0\rangle(\langle x, y_0\rangle - 1) + \langle x, y_1 - y_0\rangle(1 - \langle z, y_0\rangle) < 0$, and if the dimension is greater than or equal 2 then there are enough degrees of freedom to do so. $\square$

Chapter 4

Existence of a potential for non-traditional costs

In this chapter, we focus on finding conditions which guarantee the existence of a potential for non-traditional cost functions, i.e. costs that may assume the value $+\infty$. We are motivated by the example of the *polar cost*

$$p(x, y) = -\ln(\langle x, y\rangle - 1)_+$$

introduced in the previous chapter (where $(f(x))_+ = \max(f(x), 0)$).

We show that the classical assumption of c-cyclic monotonicity, which is sufficient to guarantee the existence of a potential in the case of real-valued costs, is not enough for non-traditional costs. We prove that in this case, the necessary and sufficient condition is that of *c-path-boundedness*. The results presented in this chapter are joint work with S. Artstein-Avidan and S. Sadovsky.

In the first section we formulate the problem and present the classical Rockafellar-Rüschendorf theorem for real-valued cost functions. We provide the well known proof for comparison with our new method.

We will then, in the second section, undertake a new approach by assuming strong regularity conditions on the c-cyclically monotone set. This approach is motivated by features of the family of costs p_t introduced in the previous chapter.

The third section contains a new proof of the Rockafellar-Rüschendorf theorem based on our result on the solvability of a particular system of (infinitely many) linear inequalities. In the fourth section we present our main result:

Theorem. *Let $c : X \times Y \to (-\infty, +\infty]$ be a cost function, and let $G \subset X \times Y$. Then there exists a c-class function $\varphi : X \to [-\infty, +\infty]$ such that $G \subset \partial^c \varphi$ if and only if G is c-path-bounded, i.e. for any $(x_s, y_s), (x_f, y_f) \in G$ there exists some constant $M = M((x_s, y_s), (x_f, y_f))$ such that for any $m \in \mathbb{N}$ and any $(x_i, y_i) \in G$, $i = 2, \ldots, m-1$, letting $(x_s, y_s) = (x_1, y_1)$ and $(x_f, y_f) = (x_m, y_m)$, we have*

$$\sum_{i=1}^{m-1} \Big(c(x_i, y_i) - c(x_{i+1}, y_i) \Big) \leq M.$$

We then apply this theorem to continuous cost functions, and give a more geometric characterization of the necessary conditions on the set G for a potential to exist.

4.1 Introduction

When considering a mass transportation problems, given a cost function one is interested in the existence of transport plans with finite cost, and in particular the optimal one. Given two probability measures μ, ν on X and Y respectively, a *transport plan* is a measure $\pi \in P(X \times Y)$ with marginals μ, ν respectively, i.e. such that for every ν-measurable set A and μ-measurable set B we have that $\pi(X \times A) = \nu(A)$, $\pi(B \times Y) = \mu(B)$. Necessary conditions for the existence of transport plans with finite cost were characterised by Strassen [20].

We say that the transport plan π_0 is *optimal* with respect to a cost function c if

$$\inf_{\pi \in P(X \times Y)} \int_{X \times Y} c(x, y) d\pi(x, y)$$

is achieved at π_0. Finding optimal transport plans is called the Monge-Kantorovitch problem (for an overview see e.g. [23]). It has been shown by Pratelli [14], that if the cost function is continuous but possibly $+\infty$ valued then a transport plan is optimal if and only if it is concentrated in a *c-cyclically monotone* set. In the same paper, it was further shown that the same is true for an arbitrary lower semi-continuous cost function in the case of atomic measures. Let us introduce the relevant definition.

Definition 4.1.1. *(c-cyclic monotonicity) Let $c : X \times Y \to (-\infty, +\infty]$ be a cost function. The set $G \subseteq X \times Y$ is called c-cyclically monotone if for any $(x, y) \in G$ we have that*

$c(x,y) < \infty$, and for any $N \in \mathbb{N}$ and any $\{(x_i, y_i)\}_{i=1}^N \subset G$ we have that

$$\sum_{i=1}^{N} c(x_i, y_i) \leq \sum_{i=1}^{N} c(x_i, y_{\sigma(i)}),$$

holds for any permutation $\sigma \in S_N$.

This notion was first introduced by Rockafellar [15] as a generalization of monotonicity in dimension one. The author was concerned with the cost $c(x,y) = \frac{|x-y|^2}{2}$, or equivalently $c(x,y) = -\langle x, y\rangle$, which is why we refer to both of them as *quadratic cost*. The two costs are equivalent in the sense that the c-cyclic monotonicity condition, as well as the optimality of a plan, are the same for both of them. Indeed, plugging them into the definition, it is easy to see that terms $\frac{|x|^2}{2}, \frac{|y|^2}{2}$ cancel and the remaining ones are precisely the mixing terms $-\langle x, y\rangle$.

In what follows, when we refer to c-cyclic monotonicity with respect to the quadratic cost we will just say *cyclic monotonicity*, and if we refer to the cyclic monotonicity induced by the polar cost p, we will say *polar cyclic monotonicity.*

Knowing that optimal plans are concentrated on c-cyclically monotone sets, it is natural to ask when such plans can be expressed by a point map $T : X \to Y$. The condition that T transports μ to ν becomes $\mu(T^{-1}(A)) = \nu(A)$ for all ν-measurable sets $A \subset Y$. In particular, a *transport map* is a transport plan concentrated on the graph of a mapping $T : X \to Y$. Moreover, similarly to that of a transport plan, the optimality of a transport map means that it minimizes the expression

$$\int_X c(x, Tx) d\mu(x).$$

Turns out that there is a connection between such a map T and a so-called *potential.* Recall from the previous chapter the notion of the c-subgradient mapping. We will be interested in

Definition 4.1.2. *A function $\varphi : X \to [-\infty, +\infty]$ in the c-class is called a potential of the set $\{(x_i, y_i)\}_{i\in I} \subset X \times Y$, with respect to the cost c, if*

$$y_i \in \partial^c \varphi(x_i) \quad \forall i \in I.$$

The main question addressed in this chapter is to determine when, for a given set $G = \{(x_i, y_i) : \ i \in I\}$, there exists a potential φ such that $G \subset \partial^c\varphi$. First note that

Fact 4.1.3. *For any cost function c and any function φ in the c-class the set*

$$\partial^c\varphi = \{(x, \partial^c\varphi(x)) : \ x \in X\} \subset X \times Y$$

is c-cyclically monotone.

Proof. Indeed, if $y_i \in \partial^c\varphi(x_i)$ then for all z we have $\varphi(z) \leq c(z, y_i) - c(x_i, y_i) + \varphi(x_i)$. Having chosen any permutation σ on the set of the indices $\{1, \dots, N\}$ we get

$$\sum_{i=1}^{N} \Big(c(x_i, y_i) - c(x_{\sigma(i)}, y_i) \Big) \leq \sum_{i=1}^{N} \Big(\varphi(x_i) - \varphi(x_{\sigma(i)}) \Big) = 0.$$

□

It was shown by Rockafellar in the case of quadratic cost [15], and Rüschendorf [18] in the case of a general but real-valued cost, that the reverse implication is also true. We shall present the proof in order to draw attention to the fact that it relies on the cost being finite valued, and therefore one cannot apply it to, for example, the polar cost.

Theorem 4.1.4 (Rockafellar-Rüschendorf)**.** *Let $c : X \times Y \to \mathbb{R}$ be a real-valued cost function and $G \subset X \times Y$ a c-cyclically monotone set. Then there exists a c-class function $\varphi : X \to [-\infty, +\infty]$ such that $G \subset (x, \partial^c\varphi(x))$.*

Proof. Fix some element $(x_0, y_0) \in G$. We will make sure $\varphi(x_0) = 0$. Define

$$\varphi(x) = \inf\{c(x, y_m) - c(x_0, y_0) + \sum_{i=1}^{m} (c(x_i, y_{i-1}) - c(x_i, y_i))\}.$$

Here the infimum runs over all $m \in \mathbb{N} \cup \{0\}$ and all m-tuples $(x_i, y_i) \in G,\ i = 1, \dots, m$.

The first observation is that the function φ is in the c-class. Indeed, it follows from the fact that the c-class is closed under infimum, and that each of the functions over which we infimize is simply of the form $c(x, y_m) + \beta$ and therefore in the c-class.

The second observation for φ is that $\varphi(x_0) = 0$. By picking $m = 0$ it is clear that $\varphi(x_0) \leq 0$. On the other hand, the c-cyclic monotonicity implies that

$$\sum_{i=0}^{m} c(x_i, y_i) \leq c(x_0, y_m) + \sum_{i=1}^{m} c(x_i, y_{i-1})$$

for any choice of $(x_i, y_i)_{i=1}^m \subset G$, hence the expression in the infimum evaluated at x_0, is at least 0. It is in this step, where we cannot omit the assumption that $c(x, y) < \infty$.

It remains to show that if $(x, y) \in G$ then $\varphi(x) + \varphi^c(y) = c(x, y) < \infty$. By the definition of the c-transform it is enough to prove that for $(x, y) \in G$ we have that $\varphi(x) \in \mathbb{R}$ and that for any $z \in X$ we have

$$c(x, y) - \varphi(x) \leq c(z, y) - \varphi(z). \tag{4.1}$$

Clearly, $\varphi(x) \leq c(x, y_0) - c(x_0, y_0)$ and hence finite. Moreover, if φ satisfies the above inequality then in particular $c(x, y) - \varphi(x) \leq c(x_0, y) - \varphi(x_0)$ and therefore we get $c(x, y) - c(x_0, y) \leq \varphi(x)$, which again is finite.

It remains to show (4.1). To this end take some $t > \varphi(x)$ and note that since $\varphi(x)$ is defined as an infimum we can find some $m \in \mathbb{N}$ and some $(x_i, y_i)_{i=1}^m \subset G$ such that

$$t > c(x, y_m) - c(x_0, y_0) + \sum_{i=1}^{m} (c(x_i, y_{i-1}) - c(x_i, y_i)). \tag{4.2}$$

Let us add $(x, y) \in G$ to the previously chosen collection $(x_i, y_i)_{i=1}^m \subset G$ and use the new $(m + 1)$-tuple for an upper bound of $\varphi(z)$. This means that for every $z \in X$ we get

$$\varphi(z) \leq c(z, y) - c(x_0, y_0) + \sum_{i=1}^{m} \Big(c(x_i, y_{i-1}) - c(x_i, y_i) \Big) + c(x, y_m) - c(x, y).$$

Using (4.2) it follows that

$$\varphi(z) - c(z, y) < t - c(x, y)$$

and since this is true for any $t > \varphi(x)$, we arrive at

$$\varphi(z) - c(z, y) \leq \varphi(x) - c(x, y),$$

as claimed. This means that indeed $(x, y) \in \partial^c \varphi$. $\square$

Interestingly, there is only one place in the proof where the c-cyclic monotonicity is used, and that is to show that at x_0 the function φ is finite.

Turns out that in the case of non-traditional cost functions the c-cyclic monotonicity is no longer a sufficient condition for the existence of a potential, and the function defined in the above proof may be infinite for $x \in P_X G$, where P_X denotes the orthogonal projection onto X. One can easily construct an example of a polar cyclically monotone set A, for which there is no function φ such that $A \subset \partial^\circ \varphi$ (see Example C.0.1 in Appendix C).

4.2 The existence of a potential assuming regularity

In what follows we shall adopt the idea that for the quadratic cost the c-class consists of all upper-semicontinuous concave functions and that we have a differentiable way of characterizing them. Similarly, in the case of the polar cost the c-class consists of all the functions that are (minus logarithm of) geometric convex functions and hence a similar characterization will be exploited. Let us recall that the class $\mathrm{Cvx}_0(\mathbb{R}^n)$ of geometric convex functions consists of all non-negative lower-semicontinuous convex functions which vanish at the origin.

We subsequently found that Rockafellar wrote the following in [15]:

> "The cyclic monotonicity condition can be viewed heuristically as a discrete substitute for two classical conditions: that a smooth convex function has a positive semi-definite second differential, and that all circuit integrals of an integrable vector field must vanish."

Of course this statement is specific to the quadratic cost, but as mentioned above, we can use a version of it to understand better the polar cost.

We shall show that indeed when considering a polar cyclically monotone set $G = \{(x_i, y(x_i))\}_{i\in I}$ with a differentiable function $y : \mathbb{R}^n \to \mathbb{R}^n$ such that for some $\epsilon > 0$ it holds that $\langle x_i, y(x_i)\rangle \geq 1 + \epsilon$ for all $i \in I$, then if in addition the set $\{x \in \mathbb{R}^n : \ x \in P_X G\}$ is a finite union of contractible open sets, it follows that a potential exists.

4.2.1 Polar cost

In the rest of this section we shall focus solely on the polar cost and the polar subgradient. As we shall see, the method used in this special case leads us to a very general method.

Recall that the cost function we are primarily interested in is defined by

$$p(x, y) = -\ln(\langle x, y\rangle - 1)_+$$

and that the main question is, when, for a given p-cyclically monotone set $G = \{(x_i, y_i)\}_{i\in I}$ there exists a potential $\varphi \in \mathrm{Cvx}_0(\mathbb{R}^n)$ such that for every $x_i \in P_X G$ we have

$$y_i \in \partial^\circ\varphi(x_i).$$

We will consider the case when G is a graph of some function, that is $G = (x, y(x))$ for $x \in P_X G$. Assuming further regularity on $y(x)$ we will make use of the connection between the polar gradient and the classical gradient, which was obtained in [6] and which we presented as Lemma 3.4.5. It states that if φ is a differentiable function and $y(x) = \nabla^\circ\varphi(x)$ then defining $z(x) = \nabla\varphi(x)$ we have that

$$y(x) = \frac{z(x)}{\langle z(x), x\rangle - \varphi(x)}. \tag{4.3}$$

We write ∇ and ∇° instead of ∂, ∂° to emphasize the assumption that φ is differentiable, and in that case the notions of gradient and subgradient coincide.

Rearranging and substituting in the definition of $z(x)$ we see that (4.3) is equivalent to

$$\nabla\varphi(x) = (\langle\nabla\varphi(x), x\rangle - \varphi(x))y(x).$$

By taking a scalar product of both sides of this equality with x and rearranging again we get that

$$\nabla\varphi(x) = \frac{\varphi(x)}{\langle y(x), x\rangle - 1} y(x),$$

which can also be written as

$$\nabla \ln \varphi(x) = \frac{y(x)}{\langle y(x), x\rangle - 1}. \tag{4.4}$$

Note that this is a system of first order linear partial differential equations.

Let us assume that a solution φ exists and is twice differentiable. Then the condition for it to be convex is that its Hessian, which we denote as $D\nabla\varphi$, is a positive semi-definite matrix. To establish this let $\psi = \ln\varphi$. Then

$$\nabla\varphi = e^\psi \nabla\psi$$

and

$$D\nabla\varphi = e^\psi(D\nabla\psi + \nabla\psi \otimes \nabla\psi), \tag{4.5}$$

where $v \otimes w$ denotes the tensor product. On the other hand, by the definition of ψ and using (4.4), we get

$$D\nabla\psi = \frac{Dy(x)}{\langle y(x), x\rangle - 1} - \frac{y(x) \otimes (x^T Dy(x) + y(x))}{(\langle y(x), x\rangle - 1)^2}.$$

Plugging this into (4.5) and using that $\nabla\psi = \frac{y(x)}{\langle y(x),x\rangle - 1}$, we can rewrite the condition $D\nabla\varphi \geq 0$, which is our notation for a matrix being positive semi-definite, in an equivalent way as

$$\frac{1}{(\langle x, y(x)\rangle - 1)^2}\Big((\langle x, y(x)\rangle - 1)I - y(x) \otimes x\Big)Dy(x) \geq 0.$$

Simplifying the expression we get

$$M := \left(I - \frac{y(x) \otimes x}{\langle x, y(x)\rangle - 1}\right) Dy(x) \geq 0 \tag{4.6}$$

4.2.2 The existence of a potential

Convex domain

We next show that the condition (4.6) is also sufficient to recover a potential in the case of a convex domain.

Proposition 4.2.1. *Let $\mathcal{U} \subset \mathbb{R}^n$ be a contractible open set, $y(x) : \mathcal{U} \to \mathbb{R}^n$ a continuously differentiable function such that the set $G = (x, y(x))$ is polar cyclically monotone. Assume further that the matrix $M = \Big(I - \frac{y(x)\otimes x}{\langle x,y(x)\rangle - 1}\Big) Dy(x)$ is symmetric. Then there exists a function $\varphi(x)$ such that $\nabla \ln \varphi(x) = \frac{y(x)}{\langle y(x),x\rangle - 1}$ for all $x \in \mathcal{U}$. Moreover, if M is positive semi-definite and $\mathcal{U}$ is convex, then the function $\varphi \in \mathrm{Cvx}_0(\mathbb{R}^n)$.*

Proof. Given $y(x)$ for all $x \in \mathcal{U}$, define

$$z(x) = \frac{y(x)}{\langle x, y(x)\rangle - 1}.$$

Note that by polar cyclic monotonicity of the set $G = (x, y(x))$ we know that $\langle x, y(x)\rangle - 1 > 0$ for all $x \in \mathcal{U}$. Therefore $z(x)$ is well defined and is a C^1 function on $\mathcal{U}$. We now show that the Jacobian $Dz(x)$ is a symmetric matrix. We check that

$$Dz(x) = \frac{Dy(x)}{\langle x, y(x)\rangle - 1} - \frac{y(x) \otimes (x^T Dy(x) + y(x))}{(\langle x, y(x)\rangle - 1)^2},$$

which is the same as

$$\frac{1}{\langle x, y(x)\rangle - 1}\Big(I - \frac{y(x) \otimes x}{\langle x, y(x)\rangle - 1}\Big)Dy - z(x) \otimes z(x).$$

By assumption the first term is symmetric and so, clearly, is the second. Hence, $Dz(x)$ is symmetric and as such it corresponds to a closed 1-form. Indeed, consider

$$d\omega = \sum_{i=1}^{n} z_i(x)dx_i = \sum_{i,j} \frac{\partial z_i(x)}{\partial x_j} dx_j \wedge dx_i = \sum_{i<j} \left(\frac{\partial z_i(x)}{\partial x_j} - \frac{\partial z_j(x)}{\partial x_i} \right) dx_j \wedge dx_i = 0,$$

where the last equality follows from symmetry of Dz and means that $d\omega$ is a closed form. Further, it is a classical fact (the Poincaré Lemma, see e.g. [11]) that on a contractible subset of $\mathbb{R}^n$ every closed 1-form is exact. This implies that there exists a function ψ such that $z(x) = \nabla\psi(x)$.

Now assume that $\mathcal{U}$ is convex. Let $\varphi(x) = e^{\psi(x)}$, and note that then $\nabla \ln \varphi(x) = \nabla\psi(x) = z(x)$ as required. Moreover, $D\nabla\varphi = e^{\psi}(\nabla\psi \otimes \nabla\psi + D\nabla\psi)$ which is exactly the matrix we assumed to be positive definite (up to multiplying by a positive number), so φ is a convex function.

In order to see that $\varphi \in \mathrm{Cvx}_0(\mathbb{R}^n)$ let us recall that $z(x) = \frac{y(x)}{\langle y(x),x\rangle - 1}$. Taking a scalar product with x on both sides we get that $\langle z(x), x\rangle = \frac{\langle y(x),x\rangle}{\langle y(x),x\rangle - 1} > 1$. Since we also have that $z = \nabla\psi = \frac{\nabla\varphi}{\varphi}$ we can conclude that

$$\langle \nabla\varphi(x), x\rangle \geq \varphi(x)$$

for all $x \in \mathcal{U}$. Moreover, since $\varphi = e^{\psi}$ it is a non-negative function. Then the fact that φ can be extended to a function in $\mathrm{Cvx}_0(\mathbb{R}^n)$ follows from the next Lemma 4.2.2. $\square$

Lemma 4.2.2. *Let φ be a non-negative convex function defined on a convex set $\mathcal{U} \subset \mathbb{R}^n$. Assume that $\langle \nabla\varphi(x), x\rangle \geq \varphi(x)$ for all $x \in \mathcal{U}$. Then there exist a geometric convex function $\Phi \in \mathrm{Cvx}_0(\mathbb{R}^n)$ defined for every $x \in \mathbb{R}^n$ and such that $\Phi|_{\mathcal{U}} = \varphi$.*

Proof. For any $x \in \mathbb{R}^n$ let us define the function

$$\Phi(x) = \max\{\sup_{u\in\mathcal{U}}\{\langle \nabla\varphi(u), x-u\rangle + \varphi(u)\}, 0\}.$$

It is clearly non-negative and as a supremum of linear functions Φ is convex. Moreover, since φ was convex we indeed have that $\Phi|_{\mathcal{U}} = \varphi$, and by the assumption that $\langle \nabla\varphi(u), u\rangle \geq \varphi(u)$ for all $u \in \mathcal{U}$, the expression $\sup_{u\in\mathcal{U}}\{\langle \nabla\varphi(u), x-u\rangle + \varphi(u)\}$ evaluated at $x = 0$ is less than or equal to 0. It follows that the function $\Phi(x)$ is the geometric convex function we needed. $\square$

Let us now show that the regularity and the polar cyclic monotonicity of a set $G = \{(x, y(x)) : x \in \mathcal{U}\}$ indeed imply that M is positive semi-definite.

Lemma 4.2.3. *Let $G = (x, y(x)) \subset \mathbb{R}^n \times \mathbb{R}^n$ be a polar cyclically monotone set defined for all $x \in \mathcal{U}$, where $\mathcal{U} \subset \mathbb{R}^n$ is an open set. Assume further that at the point x_0 the function $y(x)$ is differentiable. Then the matrix in (4.6) at $x = x_0$ is symmetric and positive semi-definite. That is,*

$$w^T\Big(I - \frac{y(x_0) \otimes x_0}{\langle x_0, y(x_0)\rangle - 1}\Big)Dy(x_0)w \geq 0$$

for all $w \in \mathbb{R}^n$.

Proof. For $x_0 \in \mathcal{U}$ let us define a family of points $x_i = x_0 + hv_i$ for all $1 \leq i \leq m$, where $v_i \in \mathbb{R}^n$ and $h \in \mathbb{R}$ is small enough such that $x_i \in \mathcal{U}$ for all $i = 1, \ldots, m$. Then from polar cyclic monotonicity follows that $\langle x_i, y(x_i)\rangle > 1$, and we have that

$$y(x_i) = y(x_0) + h\frac{\partial y}{\partial v_i}(x_0) + \xi_i,$$

where $\xi_i = o(h)$. Since all the partial derivatives are at x_0 we will suppress it in the notation, and we will write y_i instead of $y(x_i)$. Fix $t > 0$ such that for all $1 \leq i \leq m$ we have that $\langle x_i, y_i\rangle > 1 + t$ and we can assume that $h \ll t$ so that all the terms in the polar cyclic monotone condition

$$\prod_{i=1}^{m}(\langle x_i, y_i\rangle - 1) \geq \prod_{i=1}^{m}(\langle x_i, y_{i+1}\rangle - 1)$$

are positive (where we identify y_{m+1} with y_1). Substituting in the definitions of x_i, y_i we get

$$\prod_{i=1}^{m}\Big((\langle x_0, y_0\rangle - 1) + \langle x_0, \xi_i\rangle + \langle hv_i, \xi_i\rangle + \langle x_0, h\frac{\partial y}{\partial v_i}\rangle + \langle hv_i, y_0\rangle + \langle hv_i, h\frac{\partial y}{\partial v_i}\rangle\Big) \geq$$
$$\prod_{i=1}^{m}\Big((\langle x_0, y_0\rangle - 1) + \langle x_0, \xi_{i+1}\rangle + \langle hv_i, \xi_{i+1}\rangle + \langle x_0, h\frac{\partial y}{\partial v_{i+1}}\rangle + \langle hv_i, y_0\rangle + \langle hv_i, h\frac{\partial y}{\partial v_{i+1}}\rangle\Big).$$

Let us multiply both sides of the above inequality by h^{-2} and check what the limit is as $h \to 0$. First note that the terms with h to the power less than two cancel out. Then even more terms vanish when taking the limit and we arrive at

$$(\langle x_0, y_0\rangle - 1)\sum_{i=1}^{m}\langle v_i, \frac{\partial y}{\partial v_i}\rangle + \sum_{i=1}^{m}\sum_{j\neq i}\langle v_i, y_0\rangle\langle x_0, \frac{\partial y}{\partial v_j}\rangle \geq$$
$$(\langle x_0, y_0\rangle - 1)\sum_{i=1}^{m}\langle v_i, \frac{\partial y}{\partial v_{i+1}}\rangle + \sum_{i=1}^{m}\sum_{j\neq i}\langle v_i, y_0\rangle\langle x_0, \frac{\partial y}{\partial v_{j+1}}\rangle.$$

If the double sum on the right was on all indices, not just $i \neq j$, we would have a cancellation, so we may instead write

$$\begin{aligned}&(\langle x_0, y_0\rangle - 1)\sum_{i=1}^{m}\langle v_i, \frac{\partial y}{\partial v_i}\rangle - \sum_{i=1}^{m}\langle v_i, y_0\rangle\langle x_0, \frac{\partial y}{\partial v_i}\rangle \geq \\ &(\langle x_0, y_0\rangle - 1)\sum_{i=1}^{m}\langle v_i, \frac{\partial y}{\partial v_{i+1}}\rangle - \sum_{i=1}^{m}\langle v_i, y_0\rangle\langle x_0, \frac{\partial y}{\partial v_{i+1}}\rangle\end{aligned} \tag{4.7}$$

On the other hand, to prove that the matrix $M = (I - \frac{yx^T}{\langle x,y\rangle - 1})Dy$ is positive semi-definite it is enough to show that the linear map $v \mapsto Mv$ is cyclically monotone, namely

$$\sum\langle v_i, Mv_i\rangle \geq \sum\langle v_i, Mv_{i+1}\rangle.$$

Indeed, it then follows from the Rockafellar's theorem 4.1.4 and the fact that $v \mapsto Mx$ is a linear function, that the potential is a convex quadratic, and a convex quadratic is exactly a symmetric positive semi definite matrix.

To show that the map $v \mapsto Mv$ is cyclically monotone note that by the definition (4.6) of the matrix M we have

$$\langle v_i, Mv_i\rangle = \langle v_i, \frac{\partial y}{\partial v_i}\rangle - \langle v_i, y_0\rangle\langle x_0, \frac{\partial y}{\partial v_i}\rangle \cdot \frac{1}{\langle x_0, y_0\rangle - 1},$$

and similarly for the second term. But putting those terms into the inequality above we see that this is exactly the same as (4.7), which finishes the proof. □

Contractible domain

Next, we relax the assumption that the domain is convex and prove the result for a contractible domain. It turns out that we can show convexity of the recovered function (and its extension) just by using c-cyclic monotonicity.

Note that since the domain we are going to consider is open and contractible, it follows that it is polygonally path-connected. This will be important, when we make use of the next proposition. Recall that $(f(x))_+ = \max(f(x), 0)$.

Proposition 4.2.4. *Let G be a polar cyclically monotone set, and assume that for some continuously differentiable function $\varphi : \mathcal{U} \to \mathbb{R}^n$, we have $(u, \nabla^\circ\varphi(u)) = (u, y(u)) = G$. Assume that there exists $\epsilon > 0$ such that $\langle u, y(u)\rangle \geq 1 + \epsilon$, for all $u \in \mathcal{U}$. Assume further that for any two points $u_i, v_i \in \mathcal{U}$ there are polygonal paths $\gamma_i : [a, b] \to \mathcal{U}$ with*

$\gamma_i(a) = u_i, \gamma_i(b) = v_i$, for $i = 1, \dots, m$. Then

$$\frac{\varphi(u_1)}{\varphi(v_1)} \cdot \frac{\langle y(v_1), v_1\rangle - 1}{(\langle y(v_1), u_1\rangle - 1)_+} \geq 1,$$

$$\frac{\varphi(u_1)}{\varphi(v_1)} \cdot \frac{\varphi(u_2)}{\varphi(v_2)} \cdot \frac{\langle y(v_1), v_1\rangle - 1}{(\langle y(v_1), u_2\rangle - 1)_+} \cdot \frac{\langle y(v_2), v_2\rangle - 1}{(\langle y(v_2), u_1\rangle - 1)_+} \geq 1,$$

and, generally,

$$\prod_{i=1}^{m} \frac{\varphi(u_i)}{\varphi(v_i)} \cdot \prod_{i=1}^{m} \frac{\langle y(v_i), v_i\rangle - 1}{(\langle y(v_i), u_{i+1}\rangle - 1)_+} \geq 1,$$

where we let $u_{m+1} = u_1$.

Before we prove the above proposition, we need a technical lemma.

Lemma 4.2.5. *Let $G = (x, y(x))$ be a polar cyclically monotone set such that $P_X G = \mathcal{U}$. Assume that $y : \mathcal{U} \to \mathbb{R}^n$ is continuous and fix $u \in \mathcal{U}$. Let w be a vector such that for every $t \in [0, 1]$ we have $u + tw \in \mathcal{U}$. Moreover, assume that there is $\epsilon > 0$ such that $\langle y(u + tw), u + tw\rangle \geq 1 + \epsilon$ for all $t \in [0, 1]$. Then*

$$\int_0^1 \frac{\langle y(u + tw), w\rangle}{\langle y(u + tw), u + tw\rangle - 1} dt = \lim_{n\to\infty} \sum_{i=1}^{n} \ln\left(\frac{\langle y(u + \frac{i}{n}w), u + \frac{i}{n}w\rangle - 1}{\langle y(u + \frac{i}{n}w), u + \frac{i-1}{n}w\rangle - 1}\right).$$

Proof. By the definition of Riemann integral, the left hand side (which by the assumption that for all $t \in [0, 1]$ we have $\langle y(u + tw), u + tw\rangle \geq 1 + \epsilon$ is real valued and continuous, and hence Riemann integrable) is equal to

$$\lim_{n\to\infty} \sum_{i=1}^{n} \frac{\langle y(u + \frac{i}{n}w, \frac{1}{n}w\rangle}{\langle y(u + \frac{i}{n}w), u + \frac{i}{n}w\rangle - 1}.$$

Our goal is to show that the absolute value of the difference of the partial sums tends to zero as n tends to infinity. Note that since y is continuous and we have $\langle y(u+tw), u+tw\rangle \geq 1+\epsilon$ for all $t \in [0, 1]$, there exists n_0 such that for all $n \geq n_0$ we have that $\langle y(u+\frac{i}{n}w), u+\frac{i-1}{n}w\rangle \geq 1 + \epsilon/2$. Furthermore, we have

$$\ln\left(\frac{\langle y(u + \frac{i}{n}w), u + \frac{i}{n}w\rangle - 1}{\langle y(u + \frac{i}{n}w), u + \frac{i-1}{n}w\rangle - 1}\right) = \ln\left(1 + \frac{\langle y(u + \frac{i}{n}w), \frac{1}{n}w\rangle}{\langle y(u + \frac{i}{n}w), u + \frac{i-1}{n}w\rangle - 1}\right)$$

and that $\frac{\langle y(u+\frac{i}{n}w), \frac{1}{n}w\rangle}{\langle y(u+\frac{i}{n}w), u+\frac{i-1}{n}w\rangle - 1} = O\left(n^{-1}\right)$, since as we mentioned the denominator is bounded below by $\epsilon/2 > 0$. Therefore, using Taylor expansion of $\ln(1 + x) = x - O(x^2)$ it is enough to show that

$$\left|\sum_{i=1}^{n}\left(\frac{\langle y(u+\frac{i}{n}w),\frac{1}{n}w\rangle}{\langle y(u+\frac{i}{n}w),u+\frac{i-1}{n}w\rangle-1}+O\left(n^{-2}\right)-\frac{\langle y(u+\frac{i}{n})w,\frac{1}{n}w\rangle}{\langle y(u+\frac{i}{n}w),u+\frac{i}{n}w\rangle-1}\right)\right|\to 0$$

as $n\to\infty$. Putting the terms together we see that we are summing the following expressions

$$\frac{\langle y(u+\frac{i}{n}w),\frac{1}{n}w\rangle^2}{(\langle y(u+\frac{i}{n}w),u+\frac{i-1}{n}w\rangle-1)(\langle y(u_2+\frac{i}{n}w),u_2+\frac{i}{n}w\rangle-1)}+O(n^{-2})$$

and that these are also $O(n^{-2})$. Therefore, we have that $\sum_{i=1}^{n}O\left(n^{-2}\right)=O(n^{-1})\to 0$, which completes the proof. □

Proof of Proposition 4.2.4. By assumption there is a polygonal path γ connecting u_1 and v_1. Moreover, we can assume that there is $m\in\mathbb{N}$ and $v_1=z_0,z_1,\dots,z_m=u_1$ in $\mathcal{U}$, such that γ decomposes into m linear segments γ_i, say defined on $[0,1]$ and such that $\gamma_i(t)=(1-t)z_{i-1}+tz_i$ for $i=1,\dots,m$. Hence $\gamma_i(1)=z_i=\gamma_{i+1}(0)$ and we can define γ on $[0,m]$, i.e. for $k\in\{1,\dots,m\}$ and $s\in[k-1,k)$ let $\gamma(s)=\gamma_k(s-(k-1))$ and $\gamma(m)=u_1$.

Note that by the fundamental theorem of multivariate calculus we have that

$$\begin{aligned}\ln(\varphi(v_1))-\ln(\varphi(u_1))&=\sum_{i=1}^{m}\ln(\varphi(z_{i-1}))-\ln(\varphi(z_i))\\&=\sum_{i=1}^{m}\int_0^1\langle\nabla(\ln(\varphi(z_{i-1}+t(z_i-z_{i-1}))),z_i-z_{i-1}\rangle dt\\&=\sum_{i=1}^{m}\int_0^1\frac{\langle y(z_{i-1}+t(z_i-z_{i-1})),z_i-z_{i-1}\rangle}{\langle z_{i-1}+t(z_i-z_{i-1})),z_{i-1}+t(z_i-z_{i-1})\rangle-1}dt\end{aligned}$$

Applying Lemma 4.2.5 with $w_i=z_i-z_{i-1}$ and after summing all the components the first claim follows by polar cyclic monotonicity as the terms $\ln(\langle y(v_1),u_1\rangle-1)-\ln(\langle y(v_1),v_1\rangle-1)$ complete the cycle.

Similarly, we see that one can extend the cycle and go through the other points and the remaining claims follow. □

We are now ready to show that on a contractible open domain there exists a potential.

Proposition 4.2.6. *Let $\mathcal{U}\subset\mathbb{R}^n$ be contractible and open, $y:\mathcal{U}\to\mathbb{R}^n$ be a continuously differentiable function and such that $G=(x,y(x))$ is polar cyclically monotone. Moreover, assume that there exists $\epsilon>0$ such that $\langle x,y(x)\rangle\geq 1+\epsilon$ for all $x\in\mathcal{U}$. Then there exists a geometric convex function $\Phi:\mathbb{R}^n\to\mathbb{R}$ such that $\nabla^\circ\Phi(x)=y(x)$ for all $x\in\mathcal{U}$.*

Proof. Since $\mathcal{U}$ is a contractible and open set, the existence of a function $\varphi : \mathcal{U} \to \mathbb{R}$ follows from Lemma 4.2.1 and 4.2.3. It remains to show that despite $\mathcal{U}$ not being convex we can find a geometric convex extension of φ to $\mathbb{R}^n$. To this end, for any $x \in \mathbb{R}^n$ define

$$\Phi(x) = \sup_{u \in \mathcal{U}} \left(\frac{\varphi(u)}{\langle y(u), u \rangle - 1} (\langle y(u), x \rangle - 1)_+ \right).$$

Since $(y(u), u) \in G$ we know that $\langle y(u), u \rangle > 1$ and hence the function is well defined. Also, note that instead of writing the extension as the supremum of the hyperplanes $\varphi(u) + \langle \nabla \varphi(u), x - u \rangle$ (and then taking the maximum with 0), we just write it as the supremum over all basic functions and the resulting function is the same.

Clearly we have that $\Phi \in \mathrm{Cvx}_0(\mathbb{R}^n)$. It remains to show that $\Phi(u) = \varphi(u)$ for all $u \in \mathcal{U}$. This means that if we take $x = u \in \mathcal{U}$ then the supremum is achieved at the same u. In other words, it is enough to show, since the functions are non-negative, that

$$\varphi(u) \left(\frac{\varphi(v)}{\langle y(v), v \rangle - 1} (\langle y(v), u \rangle - 1)_+ \right)^{-1} \geq 1$$

holds for any pair of points $u, v \in \mathcal{U}$. First note that if $\langle y(v), u \rangle < 1$ then the left hand side becomes $+\infty$ and the inequality is satisfied. Otherwise we can rewrite this inequality as

$$\frac{\varphi(u)}{\varphi(v)} \cdot \frac{\langle y(v), v \rangle - 1}{\langle y(v), u \rangle - 1} \geq 1,$$

which holds by the first inequality in the statement of Proposition 4.2.4 (which we may use as any contractible set is polygonally path-connected), and finishes the proof. □

A domain is a union of two contractible sets

Having seen that the polar cyclic monotonicity implies that the recovered function is indeed geometric convex it is natural to try to use it in a similar way in the case where the domain is a union of contractible open sets. We illustrate this in the case of two domains, and then the case of more than two domains will be solved in a similar manner.

Proposition 4.2.7. *Let $\mathcal{U}_1, \mathcal{U}_2$ be two contractible open subsets of $\mathbb{R}^n$ such that $\mathcal{U}_1 \cap \mathcal{U}_2 = \emptyset$. Let $y : \mathcal{U}_1 \cup \mathcal{U}_2 \to \mathbb{R}^n$ be continuously differentiable on its domain. Moreover, assume that $G = (x, y(x))$ is polar cyclically monotone and that there exists $\epsilon > 0$ such that $\langle x, y(x) \rangle \geq 1 + \epsilon$ for all $x \in \mathcal{U}_1 \cup \mathcal{U}_2$. Then there exist a geometric convex function $\Phi : \mathbb{R}^n \to \mathbb{R}$ such that $\nabla^\circ \Phi|_{\mathcal{U}_1 \cup \mathcal{U}_2} = y(x)$.*

Proof. By Proposition 4.2.6 we know that there exist functions φ_1 of $\mathcal{U}_1$ and φ_2 on $\mathcal{U}_2$ such that their polar gradient is equal to y on $\mathcal{U}_1$ and $\mathcal{U}_2$ respectively. These two functions are chosen up to a multiplicative constant, which before was not relevant, but here will allow us to "fit" those two functions together.

Fix φ_1 arbitrarily and let the second function be $\eta\varphi_2$ for some $\eta > 0$ to be chosen later.

From the proof of Proposition 4.2.6 we know that each of φ_1, φ_2 can be extended to a geometric convex function on the whole $\mathbb{R}^n$; for $i = 1, 2$ define

$$\Phi_i(x) = \sup_{u \in \mathcal{U}_i} \left(\frac{\varphi_i(u)}{\langle y(u), u\rangle - 1} (\langle y(u), x\rangle - 1)_+ \right).$$

We next show that there exist a constant $\eta > 0$ such that the function

$$\Phi(x) = \max\Big(\Phi_1(x), \eta\Phi_2(x)\Big)$$

is the required potential. Since Φ clearly belongs to $\mathrm{Cvx}_0(\mathbb{R}^n)$, it is enough to show that $\frac{\Phi_1(x)}{\eta\Phi_2(x)} \geq 1$ for all $x \in \mathcal{U}_1$ and $\frac{\Phi_1(x)}{\eta\Phi_2(x)} \leq 1$ for all $x \in \mathcal{U}_2$. This means that we need to show that for every $u_1, u_2 \in \mathcal{U}_1$ and $v_1, v_2 \in \mathcal{U}_2$ we have

$$\frac{\varphi_1(u_1)}{\frac{\eta\varphi_2(v_1)}{\langle y(v_1), v_1\rangle - 1}(\langle y(v_1), u_1\rangle - 1)} \geq 1 \quad \text{and} \quad 1 \geq \frac{\frac{\varphi_1(u_2)}{\langle y(u_2), u_2\rangle - 1}(\langle y(u_2), v_2\rangle - 1)}{\eta\varphi_2(v_2)}.$$

This is the same as

$$\frac{\varphi_1(u_1)(\langle y(v_1), v_1\rangle - 1)}{\varphi_2(v_1)(\langle y(v_1), u_1\rangle - 1)} \geq \eta \quad \text{and} \quad \eta \geq \frac{\varphi_1(u_2)(\langle y(u_2), v_2\rangle - 1)}{\varphi_2(v_2)(\langle y(u_2), u_2\rangle - 1)}.$$

Since we only need to show that there is some constant η fulfilling the above inequalities, it is enough to show that

$$\frac{\varphi_1(u_1)}{\varphi_1(u_2)} \cdot \frac{\varphi_2(v_2)}{\varphi_2(v_1)} \cdot \frac{\langle y(v_1), v_1\rangle - 1}{\langle y(v_1), u_1\rangle - 1} \cdot \frac{\langle y(u_2), u_2\rangle - 1}{\langle y(u_2), v_2\rangle - 1} \geq 1.$$

This follows from Proposition 4.2.4. Hence a required constant η can be found. □

A domain is a union of m contractible sets

It is clear from the above proof that in order to generalize Lemma 4.2.7 to the domain which is a union of m contractible open sets, we will need to show that there exists a joint solution to a family of inequalities. This will guarantee that the function Φ is indeed a potential, i.e. that the supremum is achieved on the relevant function.

So far it has been convenient to favour the multiplicative setting but from here onwards we switch to the additive level where the role of the cost function $p(x,y) = -\ln(\langle x,y\rangle - 1)_+$ is prominent. Given m contractible open sets $\mathcal{U}_i$, we use Proposition 4.2.6 to find $\varphi_i \in \mathrm{Cvx}_0(\mathbb{R}^n)$ which are defined up to a multiplicative constant. We then define functions $f_i = -\ln(\varphi_i)$, which now belong to the p-class and are defined up to an additive constant, say η_i. Hence, the condition for joining any two functions f_i and f_j for $i \neq j$ becomes

$$f_i(u_i) + \eta_i \leq f_j(u_j) + \eta_j + \ln(\langle y(u_j), u_j\rangle - 1) - \ln(\langle y(u_j), u_i\rangle - 1)_+$$

for all $u_i \in \mathcal{U}_i$ and $u_j \in \mathcal{U}_j$, and the second inequality with exchanged i, j. Therefore, in order to connect m such functions in a way that gives us a potential, one needs to find a simultaneous solution $\eta \in \mathbb{R}^m$ of $m(m-1)$ inequalities of the form

$$f_i(u_i) - f_j(u_j) + p(u_j, y(u_j)) - p(u_i, y(u_j)) \leq \eta_j - \eta_i.$$

Note, that the cost function p takes values in $(-\infty, +\infty]$ and that since we have $(u_j, y(u_j)) \in G$ it follows that $p(u_j, y(u_j)) \in \mathbb{R}$. However, for $i \neq j$ we may have that $p(u_i, y(u_j)) = +\infty$ and hence the left hand side can attain values in $\mathbb{R} \cup \{-\infty\}$.

The next lemma explains when such systems of linear inequalities have a solution, and is of at most importance for understanding the role of c-cyclic monotonicity for general costs, which we study in the next section.

Lemma 4.2.8. *Let $\alpha_{i,j} \in [-\infty, +\infty)$ for $i \neq j$ and $i,\, j \in \{1, \dots, m\}$. Then the following are equivalent:*

(i) There exists a vector $\eta \in \mathbb{R}^m$ such that for all i, j we have $\alpha_{i,j} \leq \eta_i - \eta_j$.

(ii) For any any permutation σ of $\{1, \dots, m\}$ we have that $\sum_{i=1}^m \alpha_{i,\sigma(i)} \leq 0$, where we let $\alpha_{i,i} = 0$.

Proof. Clearly (a) implies (b), by summing over the pairs $(i, \sigma(i))$, as explained above we have that

$$\sum_{i=1}^m \alpha_{i,\sigma(i)} \leq \sum_{i=1}^m \eta_i - \eta_{\sigma(i)} = 0.$$

For the other direction, we shall use induction. Note that we are given a set of at most $m(m-1)$ inequalities and we would like to show they have a joint solution. Without loss of generality we may assume that $\sum \eta_i = 0$, so in fact the $m(m-1)$ inequalities are on a vector which is essentially in $\mathbb{R}^{m-1}$. By Helly's theorem [10], it is enough to make sure that any m of these inequalities have a joint solution.

To use induction on m, first note that in the case when $m = 1$ there is no condition. We may now assume that the implication is true for $(m-1)$. Then, given a subset P of m pairs $\{(i,j)\}$, if there is an index within $\{1,\ldots,m\}$ which does not appear in any of the pairs, we can discard it from the collection and by induction we know that the intersection of the corresponding half-spaces in not empty. We may thus assume that the family of m inequalities which we are trying to simultaneously satisfy include each of the integers in $\{1,\ldots,m\}$ at least once.

Further, we argue that if one of these integers, say k, appears only once then the constraint on η_k is one sided. In particular, we may consider the vector η without its k-th coordinate, solve the system on inequalities using the induction assumption and then solve the single remaining inequality for η_k, which as a single inequality in a single variable always has a solution. Moreover, the same argument applies if one of the integers appears any number of times but only as the first (resp. only as the second) in any pair $(i,j) \in P$.

We may thus assume, without loss of generality, that each of the indices in $\{1,\ldots,m\}$ appears at least once as a first index and at least once as a second index in P. But P is a set of m pairs, and so there are precisely m appearances as a first index and m as a second, which means that each integer appears once as a left and once as a right index. In other words, the pairs in P form a permutation of $\{1,\ldots,m\}$. We may decompose the permutation into its disjoint cycles. If there are more than one, then again by induction we have a solution. If not, then we have a cyclic permutation σ on $\{1,\ldots,m\}$ and we want to find a joint solution to the inequalities

$$\alpha_{i,\sigma(i)} \le \eta_i - \eta_{\sigma(i)}.$$

Fix η_1 arbitrarily, and define recursively $\eta_{\sigma^n(1)} = \eta_{\sigma^{n-1}(1)} - \alpha_{\sigma^{n-1}(1),\sigma^n(1)}$. Since $\sigma^m(1) = 1$, we have that after m steps the vector η is defined and all inequalities are satisfied (as equalities, in fact) except possibly the last one:

$$\alpha_{\sigma^{m-1}(1),1} \le \eta_{\sigma^{m-1}(1)} - \eta_1. \tag{4.8}$$

Let $k = \sigma^{m-1}(1)$ and note that since we have defined everything explicitly, we have

$$\begin{aligned}\eta_k - \eta_1 = \eta_{\sigma^{-1}(k)} - \alpha_{\sigma^{-1}(k),k} - \eta_1 &= \eta_{\sigma^{-i}(k)} - \sum_{j=0}^{i-1} \alpha_{\sigma^{-(j+1)}(k),\sigma^{-j}(k)} - \eta_1 \\ &= \eta_1 - \sum_{j=0}^{m-2} \alpha_{\sigma^{-(j+1)}(k),\sigma^{-j}(k)} - \eta_1.\end{aligned}$$

Canceling η_1 and noting that $\alpha_{\sigma^{m-1}(1),1} = \alpha_{k,\sigma^{-(m-1)}(k)}$, the condition (4.8) becomes

$$\sum_{i=1}^{m} \alpha_{i,\sigma(i)} \leq 0$$

and holds by our assumption (b). Therefore any m inequalities can be satisfied simultaneously and since the problem is essentially $(m-1)$-dimensional (the condition that the sum is 0 can be satisfied simply by adding a constant to all coordinates of the vector), it follows from Helly's theorem that if (b) holds then the system of inequalities (a) has a solution. $\square$

Using this, we prove that indeed the theorem holds for unions of finitely many contractible open sets.

Theorem 4.2.9. *Let $\mathcal{U} = \cup_{i=1}^m \mathcal{U}_i$ where $\mathcal{U}_i$ are pairwise disjoint, contractible and open subsets of $\mathbb{R}^n$ and let $y : \mathcal{U} \to \mathbb{R}^n$ be a continuously differentiable function such that the graph $G = (x, y(x))$ is polar cyclically monotone. Assume that there exists $\epsilon > 0$ such that $\langle x, y(x)\rangle \geq 1 + \epsilon$ for all $x \in \bigcup_{i=1}^m \mathcal{U}_i$. Then there exists a geometric convex function $\Phi : \mathbb{R}^n \to \mathbb{R}$ such that $\nabla^\circ \Phi(x)|_{\bigcup_{i=1}^m \mathcal{U}_i} = y(x)$ for all $x \in \mathcal{U}$.*

Proof. As mentioned before, we use Proposition 4.2.6 to find functions φ_i defined on $\mathcal{U}_i$ such that $\nabla^\circ \varphi|_{\mathcal{U}_i} = y|_{\mathcal{U}_i}$, which are defined up to a multiplicative constant. Define functions

$$f_i(x) = -\ln(\varphi_i(x)) + \eta_i$$

for $x \in \mathcal{U}_i$ and note that now $\nabla^\circ f_i(x) = y(x)$. Define

$$-\ln(\Phi(x)) = \inf_{u \in \mathcal{U}_i} \left(f_i(u) + \eta_i + \ln(\langle y(u), u\rangle - 1) - \ln(\langle y(u), x\rangle - 1)_+\right).$$

Then $\nabla^\circ(-\ln \Phi(x)) = y(x)$ if and only if for any i, j, any $u_i \in \mathcal{U}_i$ and $v_j \in \mathcal{U}_j$, we have that

$$f_j(v_j) + \eta_j \leq f_i(u_i) + \eta_i - \ln\left(\frac{\langle v_j, y(u_i)\rangle - 1}{\langle u_i, y(u_i)\rangle - 1}\right)_+$$

Rewriting, this is equivalent to proving that there exists some $\{\eta_i\}_{i=1}^m \in \mathbb{R}^m$ such that for all $i \neq j$, we have

$$\sup_{u_i \in \mathcal{U}_i, v_j \in \mathcal{U}_j} f_j(v_j) - f_i(u_i) + \ln \left(\frac{\langle v_j, y(u_i \rangle - 1}{\langle u_i, y(u_i) \rangle - 1} \right)_+ \leq \eta_i - \eta_j.$$

In the case of $m = 2$ this amounted to an inequality from above and from below on $\eta_1 - \eta_2$ (and we assumed one of them to be 0 without loss of generality). Now we have $m(m-1)$ inequalities, and the way we have wrote them corresponds to Lemma 4.2.8, with

$$\alpha_{i,j} = \sup_{u_i \in \mathcal{U}_i, v_j \in \mathcal{U}_j} f_j(v_j) - f_i(u_i) + \ln \left(\frac{\langle v_j, y(u_i \rangle - 1}{\langle u_i, y(u_i) \rangle - 1} \right)_+ .$$

The lemma implies that a solution $\eta = \{\eta_i\}_{i=1}^m$ exists if and only if for any permutation σ of $\{1, \ldots, m\}$ we have $\sum_{i=1}^m \alpha_{i,\sigma(i)} \leq 0$.

To show this, consider a permutation σ, which without loss of generality (by renaming the indices) is given by $1 \to 2 \to 3 \to \cdots \to k \to 1$. Hence, we need to show that $\sum_{i=1}^{k-1} \alpha_{i,i+1} + \alpha_{k,1} \leq 0$, which after substituting in the definition of $\alpha_{i,j}$ becomes

$$\sum_{i=1}^{k} \sup_{u_i \in \mathcal{U}_i, v_{i+1} \in \mathcal{U}_{i+1}} f_{i+1}(v_{i+1}) - f_i(u_i) + \ln \left(\frac{\langle v_{i+1}, y(u_i \rangle - 1}{\langle u_i, y(u_i) \rangle - 1} \right)_+ \leq 0$$

where we let $v_{k+1} = v_1$.

To show the correspondence with Proposition 4.2.4, we go back to the multiplicative setting. Note that clearly the above assertion is satisfied if and only if for any $u_i \in \mathcal{U}_i$ and $v_i \in \mathcal{U}_i$ we have

$$\prod_{i=1}^{k} \frac{\varphi(v_{i+1})}{\varphi(u_i)} \frac{\langle u_i, y(u_i) \rangle - 1}{(\langle v_{i+1}, y(u_i \rangle - 1)_+} \geq 1,$$

where $v_{k+1} = v_1$ and we let φ stand for any of the φ_i when applied to a point in $\mathcal{U}_i$. Up to slightly changed notation this is precisely the conclusion of Proposition 4.2.4 and the proof is complete. □

4.3 A new proof of the Rockafellar-Rüschendorf theorem

The observation about how to "patch" functions defined on any finite number of regions has led us to a very general observation. When, for a given c-cyclically monotone set

$G = \{(x_i, y_i)\}_{i=1}^m$, one looks for a potential, i.e. a c-class function $\varphi : X \to [-\infty, +\infty]$ such that for all $i = 1, \ldots, m$ satisfies

$$y_i \in \partial^c \varphi(x_i),$$

it is very natural to consider a function

$$\varphi_\eta(x) = \min_{j \in \{1, \ldots, m\}} \{c(x, y_j) - \eta_j\},$$

depending on the vector $\eta \in \mathbb{R}^m$. It is clear that such a function, as a pointwise minimum of basic functions, belongs to the c-class. Then the question becomes whether there exists a choice of η such that

$$\arg \min_{j \in [m]} \{c(x_i, y_j) - \eta_j\} = i$$

for $i = 1, \ldots, m$. In other words, finding a potential for the finite cyclically monotone set $G = \{(x_i, y_i)\}_{i \in I}$ amounts to showing there exists a solution η to a family of inequalities

$$c(x_i, y_i) - \eta_i \leq c(x_i, y_j) - \eta_j.$$

We rewrite them as

$$c(x_i, y_i) - c(x_i, y_j) \leq \eta_i - \eta_j.$$

As we have pointed out already, since $(x_i, y_i) \in G$ is c-cyclically monotone, the left hand side is never $+\infty$. However, it may be that $c(x_i, y_j) = +\infty$ in which case the left hand side is $-\infty$ and the inequality holds trivially. This seemingly technical point will be of huge significance when we consider infinite families of inequalities.

More formally we have the following theorem (where in fact we solve for $x_i \in \partial^c \varphi^c(y_i)$, rather than the equivalent $y_i \in \partial^c \varphi(x_i)$ as in the example above, as it turned out to be more convenient).

Theorem 4.3.1. *Let $c : X \times Y \to (-\infty, +\infty]$ be a cost function and let $G = \{(x_i, y_i) : i \in I\} \subset X \times Y$. There exists a potential $\varphi : X \to [-\infty, \infty]$ with $G \subset \partial^c \varphi$ if and only if the following system of inequalities*

$$c(x, y) - c(z, y) \leq \varphi(x) - \varphi(z),$$

has a solution $\varphi : P_X G \to \mathbb{R}$, where $(x, y) \in G$ and $z \in P_X G$.

Proof. Assume that there exists a potential $\varphi : X \to [-\infty, \infty]$ such that $G \subset \partial^c \varphi$. We may restrict φ to $P_X G$, and there φ must attain only finite values because $G \subset (x, \partial^c \varphi(x))$

means in particular that $\varphi(x) + \varphi^c(y) = c(x, y) < \infty$. For every $z \in P_X G$ we have

$$\varphi(z) = \inf_{w \in Y} \Big(c(z, w) - \varphi^c(w) \Big) \leq c(z, y) - \varphi^c(y),$$

and at the same time, since $(x, y) \in \partial^c \varphi$,

$$\varphi(x) = \inf_{w \in Y} \Big(c(x, w) - \varphi^c(w) \Big) = c(x, y) - \varphi^c(y).$$

Taking the difference of the two equations we get

$$c(x, y) - c(z, y) \leq \varphi(x) - \varphi(z).$$

For the other direction, assume we have a solution to the system of inequalities. We would like to extend it into some c-class function defined on X. To this end we take

$$\Phi(z) = \inf\{c(z, y) - c(x, y) + \varphi(x)\},$$

where the infimum runs over all pairs $(x, y) \in G$. We need to show that the function Φ, which is clearly in the c-class, is indeed an extension of the original function $\varphi : P_X G \to \mathbb{R}$ and that it is a potential, namely $G \subset \partial^c \Phi$.

Since φ is a solution of the system of inequalities it follows that for any $z \in P_X G$ we have

$$\varphi(z) \leq \varphi(x) + c(z, y) - c(x, y)$$

and so the $\arg\inf$ of Φ is attained at z itself. In particular, we get that Φ is indeed an extension of the original solution, and that if $(z, w) \in G$ then $w \in \partial^c \Phi(z)$, as required. $\square$

The above theorem, while very simple in nature, reduces the question of finding a potential to the question of determining when a set of linear inequalities has a solution. The index set for the inequalities are pairs $((x, y), z) \in G \times P_X G$ (or, equivalently, pairs $((x, y), (z, w)) \in G \times G$, where we ignore w as it does not appear in the inequalities). The vector we are looking for is indexed by $P_X G$, and we denote it $(\varphi(x))_{x \in P_X G}$. Even though it might seem that we want to have $(\varphi(x, y))_{(x,y) \in G}$, such notation may be misleading as while this seems to allow multi-valued φ, note that if (x, y) and (x, y') are both in G then

$$c(x, y) - c(x, y) \leq \varphi(x, y) - \varphi(x, y')$$

and

$$c(x,y') - c(x,y') \le \varphi(x,y') - \varphi(x,y)$$

which means

$$\varphi(x,y) = \varphi(x,y').$$

In other words, even if we do index the vector by $(x,y) \in G$ and not $x \in P_X G$, we still get that the solution vector depends only on the first coordinate.

Having solved the finite case (Lemma 4.2.8), it is natural to see how this method works when dealing with an infinite set G. Rockafellar's theorem and its extension by Rüschendorf works very well in sets of arbitrary cardinality, but fails when the cost can assume infinite values. Indeed, as we shall explore in depth below - when the cost can be infinite there is a compactness which is lost, and extra conditions must be added for the mere existence of a potential. However, before considering these additional conditions, it will be instructive to see how the above technique can be used to give an alternative proof of the Rockafellar-Rüschendorf theorem.

Proposition 4.3.2. *Let $\alpha_{i,j} \in \mathbb{R}$, where $i,j \in I$, for some index set I. The system of inequalities*

$$\alpha_{i,j} \le \eta_i - \eta_j$$

has a solution if and only if for any finite subset $J \subset I$ and any permutation σ on J we have

$$\sum_{i\in J} \alpha_{i,\sigma(i)} \le 0.$$

Proof. One direction is trivial and follows from Lemma 4.2.8. For the other direction, fix some $i_0 \in I$ and let $\eta_{i_0} = 0$. If the family of inequalities has a solution then it must belong to the space

$$X = \prod_{i\in I\setminus\{i_0\}} [\alpha_{i,i_0}, -\alpha_{i_0,i}] = \prod_{i\in I} X_i,$$

since among the inequalities we will have $\alpha_{i,i_0} \le \eta_i - \eta_{i_0} \le -\alpha_{i_0,i}$. Note that X is non-empty since for every $i \in I \setminus \{i_0\}$ the corresponding interval is non-empty, which follows by assumption of solvability for any finite subset applied to $\{i_0, i\}$.

From Tychonoff's theorem [22], which states that the product of any collection of compact topological spaces is compact with respect to the product topology, it follows that X is compact with respect to the product topology. Recall that the base for the open sets is the set of Cartesian products $\prod_{i\in I} U_i$, where U_i is an open set in X_i and $U_i = X_i$ for all but finitely many i.

The set of elements $v \in X$ for which a certain inequality $\alpha_{i,j} \leq \eta_i - \eta_j$ is not satisfied is clearly an open set, as it is an open set in the two participating coordinates and it is the full interval in all the other components, call it $V_{i,j}$.

Towards a contradiction, assume that there is no solution to the system of inequalities. This means that for every point in the product space X at least one of the inequalities is not satisfied and hence every point in X lies in at least one of the open sets $(V_{i,j})_{i,j \in I \times I}$. We thus have a cover of the space X, and by compactness there exists a finite subcover. This means that only a finite number of indices participate in the corresponding inequalities. But by assumption, again using Lemma 4.2.8, we get that for every finite subset of the inequalities there exists a solution and we get a contradiction to the covering property of the finite subcover. We conclude that the original collection was not a cover, which means that in fact there is a solution to the infinite family of inequalities.

□

As a corollary we have a new and simple proof for the Rockafellar-Rüschendorf theorem 4.1.4 for real valued costs. An advantage, as we shall see below, of having such a simple proof is that one can see what goes wrong in the case of general costs, and what needs to be added for the theorem to work.

4.4 Non-traditional cost functions

Let us examine what fails in the proof when the cost function is allowed to attain the value $+\infty$ (we never allow $-\infty$). Recall that if $G = \{(x_i, y_i) : i \in I\}$ is c-cyclically monotone then $c(x_i, y_i) < \infty$, but now we may get infinite values for "mixed pairs", i.e. it may be that $c(x_i, y_j) = +\infty$. If this is the case then the pair (i, j) does not give an actual inequality, that is the corresponding $\alpha_{i,j}$ is equal to $-\infty$ and so there is no constraint. In such a case Proposition 4.3.2 fails, owing to the lost compactness, and one may easily construct an infinite family of inequalities such that any finite sub-family has a solution but there is no joint solution for the entire family.

We note that there is an obvious necessary condition for the existence of a joint solution to such an infinite set of inequalities, which does not follow merely from the condition of the existence of a solution for any finite subset (for example consider: $\eta_i - \eta_1 \geq i$ and $\eta_2 - \eta_i \geq 0$). It is also not enough to assume $\alpha_{i,j}$ are uniformly bounded, as one may construct a counterexample satisfying that as well.

In fact, examining the counterexamples one sees that if

$$\alpha_{i,j} \leq \eta_i - \eta_j,$$

is a family of inequalities that has a solution, then we must have that for any $i, j \in I$ there exists some $M = M(i,j)$ such that for any m and any $\{i_k\}_{k=1}^m$,

$$\alpha_{i,i_1} + \sum_{k=1}^{m-1} \alpha_{i_k,i_{k+1}} + \alpha_{i_m,j} \leq M. \tag{4.9}$$

Indeed, the existence of a joint solution implies that

$$\eta_i - \eta_j = \eta_i - \eta_{i_1} + \sum_{k=1}^{m-1} \left(\eta_{i_k} - \eta_{i_{k+1}}\right) + \eta_{i_m} - \eta_j \geq \alpha_{i,i_1} + \sum_{k=1}^{m-1} \alpha_{i_k,i_{k+1}} + \alpha_{i_m,j}.$$

Our main observation is that the condition (4.9) is sufficient for the existence of a joint solution to the infinite family of inequalities. First, however, note that

Remark 4.4.1. *The condition* (4.9) *implies that for any finite subset $J \subset I$ and any permutation σ on J we have that $\sum_{i\in J} \alpha_{i,\sigma(i)} \leq 0$. Indeed, if this was not true, then by summing multiple times over the same cycle one could get an arbitrarily large sum, i.e. contradict condition* (4.9).

Our main theorem is the following

Theorem 4.4.2. *Let $\alpha_{i,j} \in [-\infty, \infty)$ for $i, j \in I$. Then the system of inequalities*

$$\alpha_{i,j} \leq \eta_i - \eta_j, \quad i, j \in I \tag{4.10}$$

has a solution if and only if for any i, $j \in I$ there exists some constant $M = M(i,j)$ such that for any $m \in \mathbb{N}$ and any $i_2, \dots, i_{m-1}$, letting $i = i_1$ and $j = i_m$, we have that $\sum_{k=1}^{m-1} \alpha_{i_k,i_{k+1}} \leq M$.

Instead of proving Theorem 4.4.2 directly, we shall use the following seemingly weaker result.

Theorem 4.4.3. *Let $a_{i,j} \in [-\infty, \infty)$ for $i, j \in I$. Assume further that for any $m \in \mathbb{N}$ and any $i_1, \dots, i_m$ we have $a_{i_1,i_m} \geq \sum_{k=1}^{m-1} a_{i_k,i_{k+1}}$. Then the system of inequalities*

$$a_{i,j} \leq \eta_i - \eta_j, \quad i, j \in I$$

has a solution.

We begin by assuming that Theorem 4.4.3 holds and show that it implies Theorem 4.4.2. After this reduction, we prove Theorem 4.4.3.

Proof of Theorem 4.4.2. The "only if" part is easy. Indeed if the solution to the whole family exists, then condition (4.9) follows by summing the relevant inequalities and we get that $M(i,j) = \eta_i - \eta_j$ provides the bound.

For the other direction we shall assume that for any $i, j \in I$ there exists some $M(i,j)$ such that for any m and any $\{i_k\}_{k=2}^{m-1}$, letting $i = i_1$ and $j = i_m$, we have

$$\sum_{k=1}^{m-1} \alpha_{i_k,i_{k+1}} \leq M(i,j).$$

Using this condition, we get that the numbers

$$a_{i,j} = a_{i_1,i_m} := \sup\left(\sum_{k=1}^{m-1} \alpha_{i_k,i_{k+1}} :\ m \geq 2,\ i_2, \dots, i_{m-1} \in I\right),$$

are finite.

Clearly, we have that $a_{i,j} \geq \alpha_{i,j}$ so if the family of inequalities

$$a_{i,j} \leq \eta_i - \eta_j \tag{4.11}$$

has a solution, then so does (4.10). On the other hand, since (4.11) is obtained by summing inequalities from (4.10), we get that the two systems are equivalent, i.e. each is solvable if and only if the other one is.

Now, in order to use Theorem 4.4.3 to conclude that (4.11) has a solution it remains to show that for any $m \geq 2$ and any set of indices $i_1, i_2, \dots, i_m \in I$ we have that $a_{i_1,i_m} \geq \sum_{k=1}^{m-1} a_{i_k,i_{k+1}}$.

To this end fix $\epsilon > 0$. Then for any $k \in \{1, \dots, m-1\}$ there exists m_k and $i_2^{(k)}, \dots, i_{m_k-1}^{(k)}$ such that, letting $i_k = i_1^{(k)}$ and $i_{k+1} = i_{m_k}^{(k)}$, we have

$$a_{i_k,i_{k+1}} \leq \sum_{l=1}^{m_k-1} \alpha_{i_l^{(k)},i_{l+1}^{(k)}} + \epsilon/m.$$

Thus, we identified a finite set of indices $J \subset I$, given by

$$J = \{i_1^{(k)}, i_2^{(k)} \dots, i_{m_k}^{(k)} :\ k \in \{1, \dots, m-1\},\ m \geq 2\},$$

which is naturally arranged as a path from $i_1 = i$ to $i_m = j$. Using the definition of $a_{i,j}$, and since this path participates in the supremum, we get that

$$\sum_{k=1}^{m-1} a_{i_k,i_{k+1}} \leq a_{i_1,i_m} + \epsilon.$$

As this holds for an arbitrary ϵ, we get the claim and the proof of the reduction is complete. □

4.4.1 First proof of Theorem 4.4.3, assuming countability of the index set

We will give two proofs of Theorem 4.4.3, and they will be presented in chronological order.

The first proof requires an assumption that the set of indices I is countable, and it uses the following result of Ky Fan [8]:

Theorem 4.4.4 (Ky Fan)**.** *Let E be a locally convex, real Hausdorff vector space. Let $x_\nu \in E$ be an indexed set of vectors with indices $\nu \in I$, and $\alpha_\nu \in \mathbb{R}$. Then there exists $f \in E^*$, i.e. f is a continuous linear functional on E, such that the system of inequalities*

$$f(x_\nu) \geq \alpha_\nu \quad \textit{for} \quad \nu \in I$$

is satisfied if and only if the point $(0,1) \in E \times \mathbb{R}$ does not belong to the closed convex cone $C \subset E \times \mathbb{R}$ spanned by the elements $\{(x_\nu, \alpha_\nu)\}_{\nu \in I}$.

We shall be using it for the space $\mathbb{R}^I$ with the box topology. It is clearly Hausdorff and locally convex, hence it remains to show that the point $(0,1)$ has a neighborhood separated from the cone generated by the vectors $(e_i - e_j, a_{i,j})$. In the box topology, open neighborhoods of $(0,1)$ are of the form $\prod_{i \in I}(-\epsilon_i, \epsilon_i) \times (1-\epsilon, 1+\epsilon)$.

The neighborhood we pick will be of the form $\prod_{i \in I}(-\epsilon_i, \epsilon_i) \times (1/2, \infty)$, where the sequence ϵ_i will be chosen in a way which depends only on $\{a_{i,j}\}_{i,j \in I}$.

However, before we move on to that, we include the proof.

Proof of Ky Fan's Theorem [8]. Assume that for some $f \in E^*$ the system of inequalities $f(x_\nu) \geq \alpha_\nu$ for $\nu \in I$, is satisfied. Then we may define a continuous linear functional φ on $E \times \mathbb{R}$ by

$$\varphi(x, \alpha) = f(x) - \alpha.$$

It follows that $\varphi(x_\nu, \alpha_\nu) \geq 0$ for every $(x_\nu, \alpha_\nu) \in E \times \mathbb{R}$, $\nu \in I$.

Let C be the closed convex cone in $E \times \mathbb{R}$ spanned by $\{(x_\nu, \alpha_\nu)\}_{\nu \in I}$. This means that for every $\lambda > 0$ we have that $\lambda C \subset C$. From this and the previous paragraph it follows that

$$\varphi(x, \alpha) \geq 0 \quad \text{for all} \quad (x, \alpha) \in C.$$

But $\varphi(0,1) = f(0) - 1 = -1$ which implies that $(0,1) \notin C$.

For the other direction, assume that C is the cone defined above and that $(0,1) \notin C$. By the Hahn-Banach separation theorem, in the locally convex topological vector space $E \times \mathbb{R}$, there is a hyperplane which strictly separates the closed convex cone C and the point $(0,1)$. This means that there exists $\varphi = (f, a) \in E^* \times \mathbb{R}$ and a real number b such that

$$\varphi(x, \alpha) = f(x) + a\alpha > b \quad \text{for all } (x, \alpha) \in C$$

and

$$\varphi(0,1) = f(0) + a = a < b.$$

Again, using the fact that for every $\lambda > 0$, $\lambda C \subset C$, the first inequality implies that

$$f(x) + a\alpha \geq 0 \quad \text{for all} \quad (x, \alpha) \in C.$$

Since clearly $(0,0) \in C$, using again the first inequality we get that $b < 0$. Then the second inequality implies that $a < 0$. This in turn implies that we may rewrite the system above as

$$-\frac{1}{a} f(x) - \alpha \geq 0 \quad \text{for all} \quad (x, \alpha) \in C,$$

and hence we conclude that the family of inequalities $-\frac{1}{a} f(x_\nu) \geq \alpha_\nu$, has a solution.

$\square$

Graph theoretical component of the proof

In a way which will become clear in the next sections, we shall make use of a decomposition of a nearly balanced weighted acyclic directed graph into weighted paths. A directed graph is called *acyclic* if there are no directed cycles in the graph. We call a weighted graph *nearly balanced* if we have some control over the difference between the in-coming and out-going total weight in each vertex. We assume that all the weights are non-negative.

Proposition 4.4.5. *Let $\Gamma = (V, E, (w_e)_{e\in E})$ be a finite directed weighted acyclic graph. Assume that it is almost balanced in the following sense: for some fixed vector $(\epsilon_v)_{v\in V}$ with non-negative entries, we have for every vertex $v \in V$ that*

$$\sum_{(x,v)\in E} w_{(x,v)} - \sum_{(v,y)\in E} w_{(v,y)} \in [-\epsilon_v, \epsilon_v].$$

Then there exists a weighted decomposition of Γ into paths $P_k = v_1^{(k)} \to v_2^{(k)} \to \cdots \to v_{m_k}^{(k)}$ with equal weights μ_k for each edge in P_k, such that for any $e \in E$ we have

$$w_e = \sum_{k\colon e\in P_k} \mu_k \quad \text{and} \quad \sum_k \mu_k \leq \frac{1}{2}\sum_{v\in V} \epsilon_v. \tag{4.12}$$

Moreover, it holds individually for each $v \in V$ that

$$\sum_{\{k\colon v=s_k \ \text{or}\ v=f_k\}} \mu_k \leq \epsilon_v. \tag{4.13}$$

Here we let $s_k = v_1^{(k)}$ denote the starting vertex of the path P_k and $f_k = v_{m_k}^{(k)}$ denote its end point. In fact, if $\sum_{(x,v)\in E} w_{(x,v)} \leq \sum_{(v,y)\in E} w_{(v,y)}$ then $v \neq f_k$ for any k and if $\sum_{(x,v)\in E} w_{(x,v)} \geq \sum_{(v,y)\in E} w_{(v,y)}$ then $v \neq s_k$ for any k.

Remark 4.4.6. *Note that* (4.13) *implies the second part of* (4.12)*, which can be seen by summing the inequalities in* (4.13) *over all $v \in V$. Then the right hand side becomes $\sum_v \epsilon_v$, while the left hand side is*

$$\sum_v \sum_{\{k:s_k=v \ \text{or} f_k=v\}} \mu_k,$$

and since every path has precisely one starting point s_k and one final point f_k, we get that each μ_k was summed twice, that is we get exactly $2\sum \mu_k$.

Proof of Proposition 4.4.5. Note that if $|V| = 2$ then the claim is trivial as we can use just one path. We then have that $\epsilon_1 = \epsilon_2 = w_{(1,2)} = \mu_1$, where we used 1 and 2 as labels of the vertices, and assumed the only edge is $(1, 2)$.

We shall use induction on the number of edges. If V is any set and $|E| = 1$ then the situation is exactly as in the first case we considered, and there is nothing to prove.

Assume we know the claim for $|E| < k$ and we are given Γ with $|E| = k$. Consider the edge $e_* = (x, y)$ with smallest weight w_* and pick a maximal path P which includes it (maximal in the sense that it cannot be extended to a longer path), say $s = v_1 \to v_2 \to \cdots \to v_m = f$. Maximality implies that there is no outgoing edge from its end vertex

$f = v_m$, and no edge going into its start vertex $s = v_1$. In particular, the "almost balanced" restriction on s reads $\sum_{(s,y)\in E} w_{(s,y)} \le \epsilon_s$ and on f reads $\sum_{(x,f)\in E} w_{(x,f)} \le \epsilon_f$. Moreover, since w_* was a minimal weight in the whole graph it follows that $\epsilon_s \ge w_*$ and $\epsilon_f \ge w_*$.

Define Γ' to be a graph with the same vertices V and edges $e \in E$, whose weights are defined as

$$w'_e = \begin{cases} w_e - w_* & \text{if } e \in P, \\ w_e & \text{otherwise.} \end{cases}$$

Since w_* was chosen as the minimal weight, we see that all new weights remain non-negative. Note that the edge $e_* = (x, y)$ now has weight zero and thus can be omitted. Therefore, the graph Γ' is a directed weighted acyclic graph with $k - 1$ edges. It also satisfies the almost-balanced condition, with the new vector ϵ'_v given by

$$\epsilon'_s = \epsilon_s - w_*, \;\; \epsilon'_f = \epsilon_f - w_*, \;\text{ and }\; \epsilon'_v = \epsilon_v \;\text{ for }\; v \in V \setminus \{s, f\}.$$

Note that $\sum_{v\in V} \epsilon'_v = \sum_{v\in V} \epsilon_v - 2w_*$.

By the induction assumption, the new graph Γ' has a weighted decomposition: that is, we can find paths $(P_k)_{k\in S}$ and weights μ_k such that $\sum_{\{k:\, e\in P_k\}} \mu_k = w'_e$ and for every $v \in V$, we have

$$\sum_{\{k:\, s_k=v \text{ or } f_k=v\}} \mu_k \le \epsilon'_v.$$

We add to the collection the path P with a weight w_* on each edge. We claim that this constitutes the desired weighted decomposition of Γ.

Indeed, if we compute $\sum_{\{k:\, s_k=v \text{ or } f_k=v\}} \mu_k$ for a vertex which is neither s nor f, i.e. not an end point of P, we get the same result as in Γ' and hence it is at most $\epsilon'_v = \epsilon_v$. If we compute the sum for $v = s$ or $v = f$, we get the sum in Γ' with added w_*, which is thus bounded by $\epsilon'_v + w_* = \epsilon_v$, as needed.

Finally, by construction, a vertex can be chosen as a starting vertex s_k for some path P_k only if, after equal weights were removed from its in-going and out-going edges, there was no weight left in the in-going edges. In other words, only if $\sum_{(x,v)\in E} w_{(x,v)} \le \sum_{(v,y)\in E} w_{(v,y)}$. Similarly, a vertex can be chosen as f_k for some path P_k only if $\sum_{(x,v)\in E} w_{(x,v)} \ge \sum_{(v,y)\in E} w_{(v,y)}$, which completes the proof. □

We are now ready to prove our main theorem.

Proof of Theorem 4.4.3. We begin by defining the neighborhood. Using that I is countable, we fix an ordering of it, and for every $i \in I$ we define

$$\epsilon_i = \frac{1}{5} 2^{-i} \frac{1}{\max\{|a_{k,j}| < \infty : k, j \leq i\} + 1}.$$

Note that $\epsilon_i > 0$ for every i.

Then, towards a contradiction, assume that the set $\prod_{i \in I}(-\epsilon_i, \epsilon_i) \times (1/2, \infty)$, which is an open neighborhood of $(0, 1) \in \mathbb{R}^I \times \mathbb{R}$ in the box topology, intersects with the closed convex cone generated by the vectors $(e_i - e_j, a_{i,j})$. This means that there is some finite $J \subset I$, and a positive combination $\sum_{i,j \in J} \lambda_{i,j}(e_i - e_j, a_{i,j})$ which is inside this neighborhood. This condition amounts to

$$\sum_{j \in J} \lambda_{i,j} - \lambda_{j,i} \in (-\epsilon_i, \epsilon_i) \ \ \forall i \in J \quad \text{and} \quad \sum_{i,j \in J} \lambda_{i,j} a_{i,j} \geq 1/2. \tag{4.14}$$

Let Λ be the matrix with entries $\lambda_{i,j}$. Note that subtracting any positive multiple of a permutation matrix from the matrix Λ has no effect on the sum on the left and only increases the sum on the right (by c-cyclic monotonicity). Thus we may assume without loss of generality that the matrix Λ contains no positive multiple of a permutation matrix.

Choose an order on the elements of J and consider them as vertices of a weighted directed graph Γ, where we define the weights to be $w_{(i,j)} = \lambda_{i,j}$. The assumption that Λ contains no positive multiple of a permutation matrix implies that the graph Γ is acyclic. Moreover, the first condition in (4.14) means that the graph Γ is almost balanced, up to the weights (ϵ_i). Using Proposition 4.4.5 we find paths $P_k = v_{i_1}^{(k)} \to v_{i_2}^{(k)} \to \cdots \to v_{i_{m_k}}^{(k)}$ and weights μ_k that cover the graph Γ and satisfy for every vertex v_i that

$$\sum_{\{k :\, v_{i_1}^{(k)} = v_i \text{ or } v_{i_{m_k}}^{(k)} = v_i\}} \mu_k \leq \epsilon_i.$$

Denote by A_k the adjacency matrix of the path P_k, so that $\Lambda = \sum \mu_k A_k$. Then

$$\sum_{i,j \in J} \lambda_{i,j} a_{i,j} = \sum_k \mu_k \sum_{i,j \in J} (A_k)_{i,j} a_{i,j}.$$

Moreover, by the assumption on $a_{i,j}$ and since $(A_k)_{i,j} \in \{0, 1\}$ and are indicating a path we get that

$$\sum_{i,j \in J} (A_k)_{i,j} a_{i,j} \leq a_{v_{i_1}^{(k)}, v_{i_{m_k}}^{(k)}}$$

where we used $v_{i_1}^{(k)}, v_{i_{m_k}}^{(k)}$ to denote the first and last element in the path P_k. Therefore,

$$\sum_{i,j\in J} \lambda_{i,j} a_{i,j} \leq \sum_k \mu_k a_{v_{i_1}^{(k)}, v_{i_{m_k}}^{(k)}} =: S.$$

Let us now decompose the sum S according to the start and end vector of the path P_k in the following way.

$$S = \sum_{l=1}^{\infty} \sum_{\{k:\, \max(v_{i_1}^{(k)}, v_{i_{m_k}}^{(k)})=l\}} \mu_k a_{v_{i_1}^{(k)}, v_{i_{m_k}}^{(k)}}.$$

Indeed, each path is summed exactly once, according to the quantity $l = \max(v_{i_1}^{(k)}, v_{i_{m_k}}^{(k)})$ and the earlier fixed order of elements in J. Note that in fact the first sum is finite since J is finite.

From the definition of ϵ_i, it follows that

$$a_{k,j} \leq \frac{1}{5\epsilon_{\max\{j,k\}}} 2^{-\max\{j,k\}}.$$

Using this estimate for the sum S, we get

$$S \leq \sum_{l=1}^{\infty} \frac{1}{5\epsilon_l} 2^{-l} \sum_{\{k:\, \max(v_{i_1}^{(k)}, v_{i_{m_k}}^{(k)})=l\}} \mu_k.$$

Recall that the set $\{k : \max(v_{i_1}^{(k)}, v_{i_{m_k}}^{(k)}) = l\}$ is the set of all paths which either start or terminate at l, but their other endpoint (start or end point) is below l. From Proposition 4.4.5 we know that

$$\sum_{\{k:\, v_{i_1}^{(k)}=l \text{ or } v_{i_{m_k}}^{(k)}=l\}} \mu_k \leq \epsilon_l,$$

so in particular

$$\sum_{\{k:\, \max(v_{i_1}^{(k)}, v_{i_{m_k}}^{(k)})=l\}} \mu_k \leq \epsilon_l.$$

We thus may conclude that

$$S \leq \sum_{l=1}^{\infty} \frac{1}{5\epsilon_l} 2^{-l} \epsilon_l = 1/5 < 1/2$$

which is a contradiction to the assumption and the proof is complete. $\square$

4.4.2 Second proof of Theorem 4.4.3

The second proof is very different in nature and it does not require any information about the index set I. It uses Zorn's Lemma [24].

Proof. Assume that for given $a_{i,j} \in [-\infty, +\infty)$ with $i, j \in I$, for any m and any $i_1, i_2, \ldots, i_m$ it holds that $a_{i_1,i_m} \geq \sum_{k=1}^{m-1} a_{i_k,i_{k+1}}$. We want to show that then the system of inequalities

$$a_{i,j} \leq \eta_i - \eta_j, \quad i, j \in I$$

has a solution.

Define a partially ordered set as all pairs (J, f_J), where $J \subset I$ and $f_J : J \to \mathbb{R}$ are such that for any $i, j \in J$ we have $a_{i,j} \leq f_J(i) - f_J(j)$. The set defined in this way is non-empty, which follows from Lemma 4.2.8 by considering any finite subset J. We use the standard partial order, i.e. we say that $(J, f_J) \leq (K, f_K)$ if $J \subset K$ and $f_K|_J = f_J$.

First, let us notice that every chain has an upper bound. Assume that (J_α, f_{J_α}) is a chain (namely any two elements are comparable). Consider $J = \cup_\alpha J_\alpha$ and $f_J = \cup_\alpha f_{J_\alpha}$. This function is well defined because of the chain properties, and it belongs to the chain because if $i, j \in J$ then for some α we have $(i, j) \in J_\alpha$ and hence $f_J|_{J_\alpha}$ satisfies the inequality on (i, j) and so does f_J. Finally, (J, f_J) is clearly an upper bound for the chain. Therefore, we have shown that every chain has an upper bound, and it follows from Zorn's lemma that there is a maximal element. Denote the maximal element by (J_*, f_*).

Assume towards a contradiction that $J_* \neq I$, that is, there exists some element $i_0 \in I$ such that $i_0 \notin J_*$. If f_* can be extended to be defined on i_0, in such a way that all inequalities with indices of the form (i_0, j) and (j, i_0) with $j \in J_*$ hold, we will contradict maximality and complete the proof. Note that the inequalities that need to be satisfied are

$$a_{i_0,j} \leq f(i_0) - f_*(j) \quad \text{and} \quad a_{j,i_0} \leq f_*(j) - f(i_0).$$

That is, f is to be defined in such a way that $f|_{J_*} = f_*$ and for $f(i_0)$ we have

$$\sup_{j \in J_*}(a_{i_0,j} + f_*(j)) \leq f(i_0) \leq \inf_{j \in J_*}(f_*(j) - a_{j,i_0}).$$

For such an element to exist, we must show that

$$\sup_{j \in J_*}(a_{i_0,j} + f_*(j)) \leq \inf_{j \in J_*}(f_*(j) - a_{j,i_0}), \tag{4.15}$$

or, in other words, that for any $j, k \in J_*$ we have

$$a_{i_0,j} + f_*(j) \leq f_*(k) - a_{k,i_0}.$$

We can rewrite the condition as $a_{i_0,j} + a_{k,i_0} \leq f_*(k) - f_*(j)$. Under our assumptions $a_{k,j} \geq a_{k,i_0} + a_{i_0,j}$. Since f_* already satisfies the inequality $a_{k,j} \leq f_*(k) - f_*(j)$, we know the above inequality holds for any $j, k \in J_*$, and so the inequality (4.15) holds and we may extend the function f_*. This is a contradiction to maximality, and we conclude that $J_* = I$, and hence we have found a solution to the full system of inequalities. □

4.4.3 Summary

Let us recap, in the language of our original question, what we have shown so far. Before we present the theorem, let us recall that $\alpha_{i,j}$ in Theorem 4.4.2, in the transportation setting, correspond to $c(x_i, y_i) - c(x_j, y_i)$. We then make the following definition

Definition 4.4.7. *The set $G \subset X \times Y$ is called c-path-bounded if it satisfies that for any $(x, y), (z, w) \in G$ there exists some constant $M = M((x, y), (z, w))$ such that for any $m \in \mathbb{N}$ and any $\{(x_i, y_i)\}_{i=2}^{m-1} \in G$, letting $(x, y) = (x_1, y_1)$ and $(z, w) = (x_m, y_m)$, we have*

$$\sum_{i=1}^{m-1} \Big(c(x_i, y_i) - c(x_{i+1}, y_i) \Big) \leq M.$$

Note that in light of Remark 4.4.1, a c-path-bounded set G is c-cyclically monotone. Hence, with this definition the Theorem 4.4.2 reads

Theorem 4.4.8. *Let $c : X \times Y \to (-\infty, +\infty]$ be a cost function, and let $G \subset X \times Y$. Then there exists a potential $\varphi : X \to [-\infty, \infty]$ with $G \subset \partial^c \varphi$ if and only if G is c-path-bounded.*

4.5 Geometric conditions for continuous cost functions

In order to make use of Theorem 4.4.8, one needs to show that the condition regarding the existence of M in the definition of c-path-boundedness is satisfied. In what follows, we will only consider continuous cost functions, one example of which is the polar cost $p(x, y)$.

Let $c : \mathbb{R}^n \times \mathbb{R}^n \to (-\infty, +\infty]$ be a continuous cost function. For $T < \infty$ define

$$S_T = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : c(x, y) \leq T\}.$$

We prove the following:

Theorem 4.5.1. *Let $G \subset \mathbb{R}^n \times \mathbb{R}^n$ be a bounded and c-cyclically monotone set with respect to a continuous cost function c. Moreover, assume that there is $T > 0$ such that $G \subset S_T$, namely $c(x, y) \leq T$ for all $(x, y) \in G$. Then there exists a potential φ such that $G \subset \partial^c\varphi$.*

Proof. By Theorem 4.4.8 it is enough to show that G is c-path-bounded. To do so, we consider a directed graph with vertex set G, in which there is an edge from (x, y) to (z, w) if $c(z, y) < \infty$. On the vertex set of this graph we define an equivalence relation, where two points are equivalent if there exists a directed cycle passing through both of them (or, equivalently, if there is a directed path from each of the points to the other). This is clearly an equivalence relation.

If the points (x_s, y_s) and (x_f, y_f) are in the same equivalence class, then we see that there is a constant M as required above, since we may complete any path from (x_s, y_s) to (x_f, y_f) to a cycle using some fixed path from (x_f, y_f) to (x_s, y_s), say $(z_1, w_1), \ldots (z_k, w_k)$, and using c-cyclic monotonicity we have for any m and any $(x_i, y_i)_{i=1}^m \subset G$ that

$$\begin{aligned} &c(x_s, y_s) - c(x_1, y_s) + \sum_{i=1}^{m-1} \Big(c(x_i, y_i) - c(x_{i+1}, y_i)\Big) + c(x_m, y_m) - c(x_f, y_m) \\ &\quad \leq - \left[c(x_f, y_f) - c(z_1, y_f) + \sum_{i=1}^{k-1} \Big(c(z_i, w_i) - c(z_{i+1}, w_i)\Big) + c(z_k, w_k) - c(x_s, w_k)\right]. \end{aligned}$$

We thus only need to address pairs that lie in different equivalence classes.

Let us first observe that under the assumptions we have made, there are only finitely many equivalence classes. Indeed, by continuity of c on the compact set $\overline{G} \subset S_T$, we can find some $\delta > 0$ such that if $\|(z, w) - (x, y)\| < \delta$ and (x, y), $(z, w) \in G$ then $c(x, w)$, $c(z, y) < T + 1$. In particular, any two points in G in distance less than δ belong to the same equivalence class. Using compactness of $\overline{G}$ again, we see there can be no more than a finite number of different equivalence classes. So, from here onwards we assume there are at most k_0 equivalence classes.

Next, we claim that within one equivalence class, call it $[v]$, there exists a uniform bound M such that for any two points (x_s, y_s) and (x_f, y_f) in this class, $M(s, f) \leq M$. To see this define the function $\Psi : [v] \times [v] \to \mathbb{R}$ to be the supremum over any path from the first given point to the second one. In other words, for any m and any $(z_i, w_i)_{i=1}^m \in G$ we define a function $\psi((x_s, y_s), (x_f, y_f))$ by

$$\sup \left\{ c(x_s, y_s) - c(x_1, y_s) + \sum_{k=1}^{m-1} (c(x_k, y_k) - c(x_{k+1}, y_k)) + c(x_m, y_m) - c(x_f, y_m) \right\}.$$

We have already shown, using c-cyclic monotonicity, that ψ is finite and now we claim that it is bounded. To this end, it will suffice to show that ψ is uniformly continuous, and in particular, can be extended continuously to the closure of the equivalence class $[v]$.

Assume that $\|(x_s, y_s) - (x'_s, y'_s)\| \le \delta$ and $\|(x_f, y_f) - (x'_f, y'_f)\| \le \delta$. Let $\psi((x_s, y_s), (x_f, y_f)) = M_1$. This means that any path between (x_s, y_s) and (x_f, y_f) has "total cost" at most M_1. Given some path (x_i, y_i) joining (x'_s, y'_s) and (x'_f, y'_f), we may add to it the two points (x_s, y_s) and (x_f, y_f) as the first and last points, getting a new path, between (x_s, y_s) and (x_f, y_f), and hence its total cost is bounded by M_1. We thus see that

$$\psi((x'_s, y'_s), (x'_f, y'_f)) + c(x_s, y_s) - c(x'_s, y_s) + c(x'_f, y'_f) - c(x_f, y'_f) \le M_1.$$

Since c is continuous, on a compact set it is uniformly continuous and hence there exists a small enough $\delta = \delta(\epsilon)$ such that $|c(x_s, y_s) - c(x'_s, y_s)| < \epsilon/2$ and $\left|c(x'_f, y'_f) - c(x_f, y'_f)\right| < \epsilon/2$ and we get

$$\psi((x'_s, y'_s), (x'_f, y'_f)) - \psi((x_s, y_s), (x_f, y_f)) \le \epsilon.$$

By symmetry, the claim follows, and we get that within one equivalence class, the "total cost" of any path is bounded from above with a uniform bound.

Finally, when considering points in different equivalence classes, observe that any path joining these points can be split into at most k_0 paths, each within one of the equivalence classes, and at most $k_0 - 1$ extra "single steps", each joining a pair of equivalence classes. This is because between equivalence classes there are "finite valued paths" only in one direction.

For example, if $k_0 = 2$ then every path between $(x_s, y_s) \in [v_1]$ and $(x_f, y_f) \in [v_2]$ can be separated into three components: the first lies only in $[v_1]$, the second is just one edge, joining a point in $[v_1]$ to a point in $[v_2]$, and the third lies in $[v_2]$. More precisely, letting (z, w) be the last vertex of the path lying in $[v_1]$ and (z', w') be the first vertex of the path lying in $[v_2]$, we have

$$\begin{aligned}
&c(x_s, y_s) - c(x_2, y_s) + \sum_{i=2}^{m-1} \Big(c(x_i, y_i) - c(x_{i+1}, y_i)\Big) + c(x_m, y_m) - c(x_f, y_m) \\
&\le \left[c(x_s, y_s) - c(x_2, y_s) + \sum_{i=2}^{m_1-1} \Big(c(x_i, y_i) - c(x_{i+1}, y_i)\Big) + c(x_{m_1}, y_{m_1}) - c(z, y_{m_1})\right] + c(z, w) \\
&- c(z', w) + \left[c(z', w') - c(x_{m_1+3}, w) + \sum_{i=m_1+3}^{m-1} \Big(c(x_i, y_i) - c(x_{i+1}, y_i)\Big) + c(x_m, y_m) - c(x_f, y_m)\right].
\end{aligned}$$

Note that each of the "single steps" joining two different equivalence classes is bounded by

$$c(z,w) - c(z',w) \leq T - \inf\{c(x,w) : (x,y) \in G, (z,w) \in G\} =: M_2,$$

where the infimum is bounded by continuity of c and compactness of $\overline{G}$. To sum up, any path in $G \subset S_T$ has an upper bound on total cost given by $\sum_{i=1}^{k_0} M([v_i]) + (k_0 - 1)M_2$, which are the uniform bounds on the finite number of equivalence classes and the steps between classes. As a result we have a uniform bound for any path with any beginning and end point in G. Thus, Theorem 4.4.8 yields the existence of a c-class function φ such that $G \subset \partial^c \varphi$. □

Remark 4.5.2. *Let us note that we have also shown that if all the points in G are in one equivalence class then, even if G is not a subset of S_T – i.e. there may be a sequence of points $(x_i, y_i) \in G$ such that $c(x_i, y_i) \to \infty$ – there exists a potential. In particular, this will be the case if $P_X G$ is a contractible set.*

References

[1] Artstein, S. (2002). Proportional concentration phenomena on the sphere. *Israel Journal of Mathematics*, 132(1):337–358.

[2] Artstein-Avidan, S., Florentin, D., and Milman, V. (2011). Order isomorphisms in windows. *Electronic Research Announcements*, 18:112–118.

[3] Artstein-Avidan, S. and Milman, V. (2008). A new duality transform. *Comptes Rendus Mathematique*, 346(21-22):1143–1148.

[4] Artstein-Avidan, S. and Milman, V. (2009). The concept of duality in convex analysis, and the characterization of the legendre transform. *Annals of mathematics*, pages 661–674.

[5] Artstein-Avidan, S. and Milman, V. (2011). Hidden structures in the class of convex functions and a new duality transform. *Journal of the European Mathematical Society*, 13(4):975–1004.

[6] Artstein-Avidan, S. and Rubinstein, Y. A. (2017). Differential analysis of polarity: Polar Hamilton-Jacobi, conservation laws, and Monge Ampère equations. *Journal d'Analyse Mathématique*, 132(1):133–156.

[7] Dvoretzky, A. (1961). Some results on convex bodies and Banach spaces. In *Proc. Internat. Sympos. Linear Spaces (Jerusalem, 1960)*, pages 123–160. Academic Press.

[8] Fan, K. (1968). On infinite systems of linear inequalities. *Journal of Mathematical Analysis and Applications*, 21(3):475–478.

[9] Grothendieck, A. (1998). Sur certaines classes de suites dans les espaces de banach, et le théorème de dvoretzky-rogers (fac-simile, boletim da sociedade de matemática de são paulo, vol. 8, 1958). *Resenhas do Instituto de Matemática e Estatística da Universidade de São Paulo*, 3(4):447–474.

[10] Helly, E. (1923). Über Mengen konvexer Körper mit gemeinschaftlichen Punkte. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 32:175–176.

[11] Lee, J. M. (2013). Smooth manifolds. In *Introduction to Smooth Manifolds*, pages 1–31. Springer.

[12] Lévy, P. (1951). *Problèmes concrets d'analyse fonctionnelle.* Gauthier-Villars, Paris.

[13] Milman, V. (1971). New proof of the theorem of A. Dvoretzky on intersections of convex bodies. *Functional Analysis and Its Applications*, 5(4):288–295.

[14] Pratelli, A. (2008). On the sufficiency of c-cyclical monotonicity for optimality of transport plans. *Mathematische Zeitschrift*, 258(3):677–690.

[15] Rockafellar, R. (1966). Characterization of the subdifferentials of convex functions. *Pacific Journal of Mathematics*, 17(3):497–510.

[16] Rockafellar, R. T. (1970). *Convex analysis*. Number 28. Princeton university press.

[17] Rogers, C. A. (1957). A note on coverings. *Mathematika*, 4(1):1–6.

[18] Rüschendorf, L. (1996). On c-optimal random variables. *Statistics & probability letters*, 27(3):267–270.

[19] Schneider, R. (2008). The endomorphisms of the lattice of closed convex cones. *Beitr. Algebra Geom*, 49:541–547.

[20] Strassen, V. et al. (1965). The existence of probability measures with given marginals. *The Annals of Mathematical Statistics*, 36(2):423–439.

[21] Szarek, S. and Tomczak-Jaegermann, N. (2009). On the nontrivial projection problem. *Advances in Mathematics*, 221:331–342.

[22] Tychonoff, A. (1930). Über die topologische Erweiterung von Räumen. *Mathematische Annalen*, 102(1):544–561.

[23] Villani, C. (2008). *Optimal transport: old and new*, volume 338. Springer Science & Business Media.

[24] Zorn, M. (1935). A remark on method in transfinite algebra. *Bulletin of the American Mathematical Society*, 41(10):667–670.

Appendix A

Norm with 'few' good points

The example we gave earlier of a norm such that the set of ϵ-good points has small measure was just the Euclidean norm in a non-standard position. From the perspective of Milman's question, this is an unsatisfactory example, since it is isometric to the usual Euclidean norm. In particular, we have not ruled out that for any norm that is C-Euclidean, there is some position (that is, invertible linear transformation of the norm) such that almost all points are ϵ-good, which would straightforwardly imply a positive answer to Milman's question.

We present here strong evidence that such a result cannot be obtained, by giving an example of a C-Euclidean norm such that the standard basis is 1-symmetric and the measure of ϵ-good points is exponentially small. We do not prove that the same is true for all norms that are equivalent to this one, but it seems highly unlikely that a non-standard position of a highly symmetric unit ball would have far more good points than the standard one.

The norm we take is the mixed ℓ_2/ℓ_∞ norm on $\mathbb{R}^n$ defined by the formula

$$\|x\| = \max_{|I|=m} |P_I x|,$$

where P_I is the coordinate projection to the set of indices I. If $m = cn$, then it is simple to see that $|x| \geq \|x\| \geq c^{1/2}|x|$ for every x, so this norm is $c^{-1/2}$-Euclidean.

Lemma A.0.1. *The measure of the set of ϵ-good points $x \in S^n$ with respect to the norm $\|\cdot\|$ is at most $(12\epsilon^{\frac{1-\lambda}{2}})^n \frac{1}{(1-\sqrt{\epsilon}-\lambda)\sqrt{n}}$, where $\lambda = \frac{m+1}{n}$.*

Proof. Let x be an ϵ-good point and suppose that $x_1 \geq x_2 \geq \cdots \geq x_n \geq 0$. Let y be the vector with coordinates $(x_1, x_2, \ldots, x_m, x_m, \ldots, x_m)$, and note that $\|y\| = \|x\|$. Since x is ϵ-good, y is not a witness for x being ϵ-bad, and therefore

$$\langle x, y\rangle \leq (1+\epsilon)|x|^2.$$

Expanding both sides, we find that

$$\sum_{i\leq m} x_i^2 + \sum_{i>m} x_i x_m \leq (1+\epsilon)\sum_i x_i^2$$

and therefore that

$$\sum_{i>m} x_i(x_m - x_i) \leq \epsilon \sum_i x_i^2.$$

Let $A = \{i > m : x_i \geq x_m/2\}$ and let $u = P_A y$. In other words, $u_i = x_m$ if $x_i \geq x_m/2$ and 0 otherwise. Then $(x_i - u_i)^2 \leq x_i(x_i - u_i)$ for every i. If we also let v be the coordinate projection of x to the set of coordinates greater than m, then this implies that $|v - u|^2 \leq \epsilon|x|^2 = \epsilon$.

Hence, for every ϵ-good unit vector x we can find a set J that consists of $n-m$ coordinates, a positive real number λ, and a vector y such that $|x - y| \leq \sqrt{\epsilon}$ and $y_i \in \{0, \lambda, -\lambda\}$ for every $i \in J$.

It is therefore sufficient to prove that the $\sqrt{\epsilon}$-expansion of the set of such y intersects the unit sphere in a set of small measure.

To do this, let us first fix a set J and a subset $A \subset J$ and a set of signs $\epsilon_i = \pm 1$, one for each $i \in A$, and consider the set of all y such that $\epsilon_i y_i$ is constant on A and $y_i = 0$ on $J \setminus A$. This is a subspace of $\mathbb{R}^n$ of dimension $m+1$, so by standard estimates (see e.g. [1]) its $\sqrt{\epsilon}$-expansion intersects the unit sphere in a set of measure at most

$$\mu((E_{m+1})_{\sqrt{\epsilon}}) \simeq \frac{1}{\sqrt{n\pi}} \frac{\sqrt{\lambda(1-\lambda)}}{(1-\lambda) - \sin^2(\sqrt{\epsilon})} e^{-\frac{n}{2}u(\lambda,\epsilon)} \leq \frac{1}{\sqrt{n\pi}} \frac{1}{cos^2(\sqrt{\epsilon}) - \lambda} e^{-\frac{n}{2}u(\lambda,\epsilon)},$$

where $\lambda = \frac{m+1}{n}$ and $u(\lambda,\epsilon) = (1-\lambda)\log\frac{1-\lambda}{\sin^2(\sqrt{\epsilon})} + \lambda \log \frac{\lambda}{\cos^2(\sqrt{\epsilon})}$. This function can be estimated from below by

$$u(\lambda,\epsilon) = \log\left(\frac{1-\lambda}{\epsilon}\left(\frac{\lambda}{1-\lambda}\right)^{\lambda} (\tan\sqrt{\epsilon})^{2\lambda}\right) \geq \log\left(\epsilon^{\lambda-1}\frac{\lambda^\lambda}{(1-\lambda)^{\lambda-1}}\right).$$

As a result we get

$$\begin{aligned}\mu((E_{m+1})_{\sqrt{\epsilon}}) &\leq \frac{1}{\sqrt{n}(1-\sqrt{\epsilon}-\lambda)}\left(\epsilon^{\lambda-1}\frac{\lambda^\lambda}{(1-\lambda)^{\lambda-1}}\right)^{-n/2} \\ &\leq \frac{1}{\sqrt{n}(1-\sqrt{\epsilon}-\lambda)}\left(\frac{\epsilon^{1-\lambda}}{(1-\lambda)^{1-\lambda}\lambda^\lambda}\right)^{n/2}.\end{aligned}$$

The number of ways of choosing A, J and the signs is $\binom{n}{\lambda n}3^{(1-\lambda)n}$. Therefore, for a fixed $\lambda = (m+1)/n$ we get that the measure of good points is at most

$$\begin{aligned}
&\binom{n}{\lambda n}3^{(1-\lambda)n}\frac{1}{\sqrt{n}(1-\sqrt{\epsilon}-\lambda)}\left(\frac{\epsilon^{1-\lambda}}{(1-\lambda)^{1-\lambda}\lambda^{\lambda}}\right)^{n/2} \\
&\qquad\leq \left(\frac{e}{\lambda}\right)^{\lambda n}\frac{1}{\sqrt{n}(1-\sqrt{\epsilon}-\lambda)}\left(\frac{\epsilon^{1-\lambda}}{(1-\lambda)^{1-\lambda}\lambda^{\lambda}}\right)^{n/2}3^{(1-\lambda)n} \\
&\qquad\leq \frac{1}{\sqrt{n}(1-\sqrt{\epsilon}-\lambda)}12^{n}\epsilon^{(1-\lambda)n/2},
\end{aligned}$$

since the function $x^{3x/2}(1-x)^{(1-x)/2}$ is bounded below for all $x \in [0,1]$. Hence, the measure of ϵ-good points is at most $(12\epsilon^{\frac{1-\lambda}{2}})^n\frac{1}{(1-\sqrt{\epsilon}-\lambda)\sqrt{n}}$, and for a fixed epsilon and $\lambda < \min\{1-2\sqrt{\epsilon}, 1-2\frac{\log 12}{\log\frac{1}{\epsilon}}\}$, this tends to 0 as the dimension increases. $\square$

Appendix B

Proof of Example 3.2.2. We want to show that $\mathcal{A}_t 1_K^{\infty,s} = 1_{(1+t)K^\circ}^{\infty,t/s}$ for a convex set $K \subset \mathbb{R}^n$ and such that $0 \in \mathrm{int}(K)$. To this end let $\varphi = 1_K^{\infty,s}$, then by the definition of the transform we have

$$\mathcal{A}_t\varphi(y) \quad = \quad \sup_{\{x:\langle x,y\rangle \geq 1+t\}} \frac{\langle x,y\rangle - 1}{1_K^{\infty,s}(x)}.$$

Note that $1_K^{\infty,s}$ can only attain three different values $0, s, +\infty$. It follows that if for all $x \in \mathrm{int}(K)$ we have that $\langle x,y\rangle < 1+t$ then we must have $y \in \mathrm{int}((1+t)K^\circ)$ and $\mathcal{A}_t\varphi(y) = 0$. On the other hand, if y is such that there exists $x \in \mathrm{int}(K)$ such that $\langle x,y\rangle \geq 1+t$ then $y \in \mathbb{R}^n \setminus (1+t)K^\circ$ and the supremum becomes $+\infty$. In the case when y is on the boundary of K, i.e. $y \in \partial(1+t)K^\circ$, there is at least one $x \in \partial K$ for which $\langle x,y\rangle = 1+t$, therefore the numerator becomes t and the denominator s, as claimed. $\square$

Proof of Example 3.2.3. We want to prove that $\mathcal{A}_t\ell_{y_0} = \ell^\infty_{y_0/|y_0|^2}$ holds for a truncated linear function $\ell_{y_0}(x) = \langle x, y_0\rangle_+$. Indeed,

$$(\mathcal{A}_t\varphi)(y) \quad = \quad \sup_{\{x:\langle x,y\rangle \geq 1+t\}} \frac{\langle x,y\rangle - 1}{\langle x,y_0\rangle_+} = \sup_{\{x:\langle x,y\rangle \geq 1+t\}} \langle \frac{x}{\langle x,y_0\rangle_+}, y\rangle - \frac{1}{\langle x,y_0\rangle_+}$$

If y is not on the line $\mathbb{R}y_0$, then choose x to be $M(y - \frac{\langle y,y_0\rangle}{|y_0|^2}y_0)$ which is non-zero and M is large, in particular so that $\langle x,y\rangle \geq 1+t$. Note that such chosen x is orthogonal to y_0 and hence the denominator vanishes. It follows that $\mathcal{A}_t\ell_{y_0} = +\infty$ for any $y \notin \mathbb{R}y_0$. If y belongs to $\mathbb{R}_+y_0$, i.e. there exists $s > 0$ such that $y = sy_0$ we have that

$$(\mathcal{A}_t\varphi)(y) \quad = \quad \sup_{\{x:s\langle x,y_0\rangle \geq 1+t\}} s\frac{\langle x,y_0\rangle - \frac{1}{s}}{\langle x,y_0\rangle_+}.$$

Clearly the supremum is at most s, but by taking $x = My_0$ for $M \to \infty$ one gets that indeed $\mathcal{A}_t\varphi(sy_0) = s$. If $y = sy_0$ for $s < 0$ we get that the denominator is 0, as we supremize over all x such that $\langle x, y_0\rangle \le \frac{1+t}{s} < 0$, hence the supremum is again $+\infty$.

For the reverse direction, i.e. to show that $\ell_{y_0}(x)$ is equal to

$$\mathcal{A}_t\ell^\infty_{y_0/|y_0|^2}(x) = \sup_{\{z:\langle z,x\rangle\ge 1+t\}} \frac{\langle x, z\rangle - 1}{\ell^\infty_{y_0/|y_0|^2}(z)},$$

note that if x is such that for all $z = s\frac{y_0}{|y_0|^2}$ with $s > 0$ we have that $\langle x, z\rangle < 1+t$, then $\langle x, \frac{y_0}{|y_0|^2}\rangle \le 0$ and the denominator becomes $+\infty$ and hence $\mathcal{A}_t\ell^\infty_{y_0/|y_0|^2}(x) = 0$. If however, $\langle x, \frac{y_0}{|y_0|^2}\rangle > 0$ then we can choose z to be $s\frac{y_0}{|y_0|^2}$ for any $s > 0$. This means that the supremum is in fact over all $s > 0$ and we get

$$\sup_{s>0} \frac{s\langle x, \frac{y_0}{|y_0|^2}\rangle - 1}{s\langle \frac{y_0}{|y_0|^2}, \frac{y_0}{|y_0|^2}\rangle} = \langle x, y_0\rangle,$$

which precisely means that ℓ_{y_0} is the $\mathcal{A}_t$ image of $\ell^\infty_{y_0/|y_0|^2}$. □

Example B.0.1. *Let $\varphi(x) = \|x\|_K$, be a norm. Then*

$$(\mathcal{A}_t\|\cdot\|_K)(y) = \mathcal{A}(\|\cdot\|_K)(y) = h_K(y),$$

where h_K is a support function of K.

Proof. From the definition, for every $t \ge 0$ we have that

$$\begin{aligned}(\mathcal{A}_t\|\cdot\|_K)(y) &= \sup_{\{\langle x,y\rangle\ge 1+t\}} \frac{\langle x, y\rangle - 1}{\|x\|} = \sup_{u\in S^{n-1}} \sup_{\{s>0:\ \langle u,y\rangle\ge\frac{1+t}{s}\}} \frac{s\langle u, y\rangle - 1}{s\|u\|_K} \\ &= \sup_{u\in S^{n-1}} \frac{\langle u, y\rangle}{\|u\|_K} = h_K(y).\end{aligned}$$

□

Example B.0.2. *For $p > 1$ define $\varphi(x) = \frac{\|x\|_K^p}{p}$, where $\|\cdot\|_K$ is a norm. Then*

$$\mathcal{A}_t\varphi(y) = \begin{cases} \left(\frac{p-1}{p}\right)^{p-1} h_K^p(y) & \textit{when } \frac{1}{p-1} \ge t \\ \frac{pt}{(1+t)^p} h_K^p(y) & \textit{otherwise.}\end{cases}$$

In particular, for $p = 2$ and the ℓ_2 norm we get that $\mathcal{A}_t\left(\frac{|\cdot|^2}{2}\right)(y)$ is equal to $\mathcal{A}\left(\frac{|\cdot|^2}{2}\right)(y)$ for $t \le 1$ and is different otherwise.

Proof. By the definition we have

$$\mathcal{A}_t\varphi(y) = \sup_{\{x:\langle x,y\rangle \geq 1+t\}} p\frac{\langle x,y\rangle - 1}{\|x\|_K^p} = \sup_{u\in S^{n-1}} \sup_{\{s>0:\ \langle u,y\rangle \geq \frac{1+t}{s}\}} p\frac{s\langle u,y\rangle - 1}{s^p\|u\|_K^p}.$$

Define $f(s) = s^{1-p}\langle u,y\rangle - s^{-p}$. Then, $f'(s_0) = 0$ if and only if $s_0 = \frac{p}{(p-1)\langle u,y\rangle}$ and it is easy to check that since $p > 1$ we have that $f''(s_0) \leq 0$. Clearly, since we need $\langle u,y\rangle \geq \frac{1+t}{s}$, we will only consider u with $\langle u,y\rangle > 0$ hence we see that $s_0 > 0$ and that the condition $\langle u,y\rangle \geq \frac{1+t}{s_0}$ becomes $\frac{1}{p-1} \geq t$.

Therefore, if $\frac{1}{p-1} \geq t$ we plug in s_0 and since $\sup_{u\in S^{n-1}} \frac{\langle u,y\rangle^p}{\|u\|_K^p} = h_K^p(y)$, we indeed get $\left(\frac{p-1}{p}\right)^{p-1} h_K^p(y)$. If $t > \frac{1}{p-1}$ then by monotonicity of f, instead of s_0, we take the smallest value s which fulfills the condition, i.e. $s_1 = \frac{1+t}{\langle u,y\rangle}$ and by plugging it in the supremum we get the desired result. □

Lemma B.0.3 (A restricted Cauchy functional equation)**.** *Assume that a continuous function* $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ *satisfies*

$$\varphi(xy) = \varphi(x)\varphi(y)$$

for all $x \in \mathbb{R}_+$ *and for all* $y \in [\alpha,\beta]$ *with* $0 < \alpha < 1 < \beta$*. Then,* $\varphi(x) = x^c$ *for some* $c \in \mathbb{R}$*.*

Proof. Let $f(x) = \ln(\varphi(e^x))$, we must show that given the condition

$$f(x+y) = f(x) + f(y)$$

for all $x \in \mathbb{R}$, $y \in [\ln\alpha, \ln\beta] = [a,b]$, we must have $f(x) = cx$ for some $c \in \mathbb{R}$. Note that $a < 0 < b$, and also that since φ is a continuous function, so is f and it is enough to prove the result for $x \in \mathbb{Q}$ and $y \in [a,b]\cap\mathbb{Q} := A$ and use the fact that rational numbers are dense.

Consider $y = 0 \in A$ and note that for any x we have $f(x) = f(x) + f(0)$, and hence $f(0) = 0$. Moreover, for any $n \in \mathbb{N}$ and $y \in [a,b]$, we have that $f(ny) = f((n-1)y + y) = nf(y)$. Now, for any $n \in \mathbb{N}$ we have that if $y \in A$ then so is $\frac{y}{n}$, hence substituting and multiplying the previous equality by $\frac{m}{n}$, for some $m \in \mathbb{N}$, we get

$$\frac{m}{n}f(y) = mf\left(\frac{y}{n}\right) = f\left(\frac{m}{n}y\right),$$

where the last equality follows from the previously established one. Now, for any $x \in \mathbb{Q}_+$ we can find $m,n \in \mathbb{N}$ and $\frac{p}{q} = y \in [0,b]\cap\mathbb{Q}$ with $p,q \in \mathbb{N}$, such that $x = \frac{m}{n}y$. It follows

that
$$f(x) = f\left(\frac{m}{n}y\right) = \frac{m}{n}f(y) = \frac{m}{n}\frac{p}{q}f\left(\frac{q}{p}y\right) = xf(1).$$
If $x < 0$ then choose $y \in [0, b] \cap \mathbb{Q}$ and $k \in \mathbb{N}$ such that $x + ky = 0$. Then

$$0 = f(0) = f(x + ky) = f(x) + kf(y) = f(x) + f(ky) = f(x) + f(-x).$$

Combining this with the previous identity we get for $x \in \mathbb{Q}_-$ that

$$f(x) = -f(-x) = -(-xf(1)) = xf(1).$$

Setting $c = f(1)$, this finishes the proof. $\square$

Appendix C

Example C.0.1. *Consider the set*

$$A = \{(x,y): \ x \in (\tfrac{1}{2}, 2), \ y = \tfrac{5}{2} - x\} \cup \{(3, \tfrac{1}{2})\} \subset \mathbb{R} \times \mathbb{R}.$$

This is a p_0-cyclically monotone set since for every point $(x,y) \in A$ we have $\langle x, y\rangle > 1$ and it is a graph of non-increasing function on its domain (for the characterization of polar cyclically monotone sets in $\mathbb{R}$, see the subsection below in the Appendix C).

However, there is no potential, i.e. $\varphi \in \mathrm{Cvx}_0(\mathbb{R})$ such that $A \subset \partial^\circ \varphi$. First consider $x \in (\frac{1}{2}, 2)$. By Lemma 3.4.5 follows that if $\varphi \in \mathrm{Cvx}_0$, such that $y \in \partial^\circ \varphi(x)$, exists then there is $z \in \partial\varphi(x)$ such that $\frac{z}{xz - \varphi(x)} = y$. This gives us the equation $\frac{5}{2} - x = \frac{z}{xz-\varphi(x)}$, which in turn due to the continuity of y for $\frac{1}{2} < x < 2$, is equal to

$$\frac{\varphi'(x)}{\varphi(x)} = \frac{5/2 - x}{(2-x)(x-1/2)}.$$

Finally, we get $\varphi(x) = \left(\frac{(1-2x)^4}{2-x}\right)^{1/3}$. But note that this function, which is well defined on $(\frac{1}{2}, 2)$, does not have a finite extension, for $x > 2$, which remains a geometric convex function. Indeed, as $x \to 2^-$ the function tends to $+\infty$. But this means that φ has to be $+\infty$ for all $x > 2$ and hence there is no function $\psi \in \mathrm{Cvx}_0$ such that $\psi|_{(1/2,2)} = \varphi|_{(1/2,2)}$ and such that the point $(3, 1/2)$ belongs to $\partial^\circ \psi$.

Polar cyclically monotone sets in $\mathbb{R}$

We claim that in dimension 1 it is very easy to characterize polar cyclically monotone sets. We first reduce the question to two separate questions in quadrants:

Lemma C.0.2. *A set $A \subseteq \mathbb{R} \times \mathbb{R}$ is polar cyclically monotone if and only if $A_1 = A \cap \mathbb{R}^+ \times \mathbb{R}^+$ and $A_2 = A \cap \mathbb{R}^- \times \mathbb{R}^-$ are both polar cyclically monotone and $A = A_1 \cup A_2$.*

Proof. Since for every pair $(x, y) \in A$ we must have $c(x, y) < \infty$, we get that $A \subset \{(x, y) \in \mathbb{R}^2 : \ xy > 1\}$. Hence, we may write that $A = A_1 \cup A_2$ where $A_1 = A \cap (\mathbb{R}^+ \times \mathbb{R}^+)$ and $A_2 = A \cap (\mathbb{R}^- \times \mathbb{R}^-)$. Clearly if A is c-cyclically monotone then by the definition so are A_1 and A_2. On the other hand, if both A_1 and A_2 are c-cyclically monotone then for any permutation for which $x_i y_{\sigma(i)} < 0$ the total cost will be infinite and the desired inequality will hold. We thus only need to consider permutations for which the sets $\{i : x_i < 0\}$ and $\{i : x_i > 0\}$ are invariant, and then the condition on A is the sum of the conditions on A_1 and A_2. □

Lemma C.0.3. *A set $A \subseteq \mathbb{R}^+ \times \mathbb{R}^+$ is polar cyclically monotone if and only if A lies on a graph of non-increasing function.*

Proof. It is easy to see that the c-cyclic monotonicity condition with respect to the polar cost is equivalent to

$$\prod_{i=1}^{N} (x_i y_i - 1) \geq \prod_{i=1}^{N} (x_i y_{\sigma(i)} - 1).$$

First, we show that this is true for $x_1 \leq x_2 \leq \ldots \leq x_N$ if $y_1 \geq y_2 \geq \ldots \geq y_N$. We will show it by induction on N. Let $N = 2$ and $(x_1, y_1), (x_2, y_2) \in A$ with $x_1 \leq x_2$. Then the following condition needs to be fulfilled

$$(x_1 y_1 - 1)(x_2 y_2 - 1) \geq (x_1 y_2 - 1)(x_2 y_1 - 1).$$

After rearranging one can see that this is equivalent to

$$(x_2 - x_1)(y_1 - y_2) \geq 0,$$

and since first bracket is non-negative we get $y_1 \geq y_2$.

Now assume that the claim is true for $N = k$ and we will show it for $k + 1$. If $\sigma(k+1) = k+1$ then we are done. If not and $\sigma(k+1) = j \neq k+1$, then assume $\sigma(i) = k+1$. We then know that $i < k + 1$ and $j < k + 1$ and therefore

$$x_i \leq x_{k+1}, \text{ and } \ y_j \geq y_{k+1}.$$

By the base case it follows that

$$(x_i y_j - 1)(x_{k+1} y_{k+1} - 1) \geq (x_i y_{k+1} - 1)(x_{k+1} y_j - 1).$$

Now, define a permutation σ^* on the first k elements by $\sigma^*(x) = \sigma(x)$ if $x \neq i$ and $\sigma^*(i) = j$. We then have

$$\prod_{i=1}^{k}(x_i y_i - 1) \geq \prod_{i=1}^{k}(x_1 y_{\sigma^*(i)} - 1).$$

Multiplying both sides of this inequality by $(x_{k+1}y_{k+1} - 1)$ and using the previous one, proves the claim.

Finally, there is just one c-cyclically monotone set for given $\{x_i\}_{i=1}^N, \{y_i\}_{i=1}^N$ and also there is just one decreasing rearrangement, so they must coincide.

$\square$