MODULATION SPACES, BMO AND THE ZAK TRANSFORM, AND MINIMIZING IPH FUNCTIONS OVER THE UNIT SIMPLEX

A Thesis
Presented to
The Academic Faculty

by

Ramazan Tinaztepe

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Mathematics

Georgia Institute of Technology
August 2010

MODULATION SPACES, BMO AND THE ZAK TRANSFORM, AND MINIMIZING IPH FUNCTIONS OVER THE UNIT SIMPLEX

Approved by:

Professor Christopher Heil, Advisor
School of Mathematics
Georgia Institute of Technology

Professor Doron Lubinsky
School of Mathematics
Georgia Institute of Technology

Professor Luca Dieci
School of Mathematics
Georgia Institute of Technology

Professor Hao Min Zhou
School of Mathematics
Georgia Institute of Technology

Professor Renato Monteiro
Department of Industrial and System Engineering
Georgia Institute of Technology

Date Approved: 1 July 2010

To my family

ACKNOWLEDGEMENTS

During my PhD adventure, I have experienced sensational oscillations in my mood time to time. Sometimes I found myself questioning the goal of life, sometimes questioning whether doing PhD was a good idea and usually questioning my mistakes on the proofs. I was able to survive in math and life with the help of especially two persons; One is my advisor, Christopher Heil and the other is my wife, Havva.

I would like to present my deepest gratitude to Dr. Christopher Heil. He not only helped me survive in math also helped me as a life mentor. Even in the very desperate moments of my PhD, I found him very patient with me and most of the time our meetings turned out to be great comebacks for me. I learnt very much from him like a novice learning from his master. He is the coolest person I have ever known.

I would like to thank all my committee members Hao Min Zhou, Luca Dieci, Doron Lubinsky and Renato Monteiro who agreed to serve on my committee despite their tight schedules.

I am lucky for having many nice friends around me in Atlanta.

I would like to thank all my friends and their families without whom Atlanta would be a very boring place for me.

I especially thank Adem Cakmak, Turkay Yolcu, Selma Yıldırım, Suleyman Ulusoy, Hasan Palta, Mehmet Fatih Kılınc, Salih Tekin, Huseyin Kurtuldu, Mustafa Alevli, Derrick Hart, Bradley Green and their families for opening their hearts to me and my family.

I would like to thank all my friends in the department. I especially thank Marc Sedjro, my officemate Liangda Huang, Georgios Amanatidis, Benjamin Webb, Alexander Grigo, Kun Zhao, Jorge Viveros, Bülent Tosun for making this department a place of friends.

I am also grateful to Genola Turner, Sharon McDowell and Justin Filoseta for their various forms of support during my graduate study.

Last, I present my deepest gratitude to my wife, Havva and my daughters, Ayse Sude and Hatice Merve and my parents. Without their love, encouragement and constant support I would not have been able to finish any of these. I offer my regards and blessings to all of those who supported me in any respect during the completion of my thesis. Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This thesis consists of two parts. In the first part, we give some results on modulation spaces, which are function spaces based on time-frequency behaviors of functions. First the relationship between the classical spaces and the modulation spaces is established. It is proved that certain modulation spaces defined on $\mathbb{R}^2$ lie in the BMO space. Another result is that the Zak transform, a discrete time-frequency transform, maps a modulation space into a higher dimensional modulation space. This result is important because it uses Gabor analysis methods instead of using classical Fourier analysis. It generalizes Gautam's result [25] and yields a Balian-Low type theorem for modulation spaces.

In the second part, we deal with optimization of an increasing positively homogeneous functions on the unit simplex. The class of increasing positively homogeneous functions is one of the function classes obtained via min-type functions in the context of abstract convexity. The cutting angle method, which is similar to the cutting plane method in classical convex optimization, was given by A. Rubinov et al. ([3],[4]) for the minimization of this type functions. The most important step of this method is the minimization of a function which is the maximum of a number of min-type functions on the unit simplex. We propose a numerical algorithm for the minimization of such functions on the unit simplex and we mathematically prove that this algorithm finds the exact solution of the minimization problem. Some experiments have been carried out and the results of the experiments have been presented.

CHAPTER I

INTRODUCTION

1.1 Part I Introdution

In engineering, signals are considered to be functions of time, and the Fourier transform of a signal is the frequency representation of the signal. Time and frequency are the main ingredients of signal processing in engineering.

Time-frequency analysis is a branch of modern harmonic analysis. It is comprised of the divisions of mathematics that use the structure of translations and modulations (or time-frequency shifts) for the analysis of functions and operators.

The main tool in time-frequency analysis is the short-time Fourier Transform (also called the Windowed Fourier Transform). Time-frequency analysis interprets the short-time Fourier transform as a measure of simultaneous time and frequency information.

Let f and $g \neq 0$ be functions. The short-time Fourier transform of f with respect to g (called the window function) is defined as

$$V_g f(x, w) = \int_{\mathbb{R}^d} f(t)\overline{g(t-x)}e^{-2\pi i t \cdot w} dt$$

for $x, w \in \mathbb{R}^d$. The short time Fourier transform of a function inherits both the time and frequency properties of the function. In the first section of this chapter, a brief background will be presented about the short time Fourier transform. One can find detailed information about the short time Fourier transform in [27] with regards to both its theoretical and applied aspects.

To achieve a quantitative time-frequency analysis, Feichtinger introduced a class of Banach spaces, called modulation spaces, which measure concentration in terms of a weighted mixed norm on the short-time Fourier transform [19, 27].

Definition 1.1.1. Let $g \in \mathcal{S}(\mathbb{R}^d)(\mathbb{R}^d)$. The modulation space $M_m^{p,q}(\mathbb{R}^d)$ $(1 \leq p < \infty)$ consists of all tempered distributions $f \in \mathcal{S}'(\mathbb{R}^d)$ such that the norm

$$\|f\|_{M_m^{p,q}(\mathbb{R}^d)} = \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |V_g f(x,w)|^p m(x,w)^p dx \right)^{\frac{q}{p}} dw \right)^{\frac{1}{q}}$$

is finite. If $p = \infty$ or $q = \infty$, then the corresponding p-norm is replaced by the essential supremum.

Different windows g yield equivalent norms on $M_m^{p,q}(\mathbb{R}^d)$. If $|m(z)| \leq C(1+|z|)^N$ for some N, then $M_m^{p,q}(\mathbb{R}^d)$ is a Banach space $1 \leq p, q \leq \infty$ and the class of Schwartz functions forms a dense subspace of $M_m^{p,q}(\mathbb{R}^d)$ for $1 \leq p, q < \infty$. The weights $m(z)$ are generally considered to be polynomial weights, but exponential weights and other types of weights are also used in the literature. The modulation spaces with $p < 1$ or $q < 1$ also arise naturally in many settings, for example in connection with sparsity issues, but require a more technical viewpoint.

In the second chapter, some detailed information about modulation spaces and related spaces is presented. The interplay between these spaces is summarized.

The modulation spaces are the "right" spaces for time-frequency analysis and they occur in many problems in the same way as Besov spaces are attached to wavelet theory and issues of smoothness. Many classical spaces are well-described by wavelets, but recently the modulation spaces have been successfully used to address problems not suited to traditional techniques, e.g., the analysis of pseudodifferential operators with nonsmooth symbols [29, 15], or the modeling of narrowband wireless communication channels [47, 48, 46].

For $p, q \neq 2$, the modulation spaces $M_m^{p,q}(\mathbb{R}^d)$ do not coincide with any of Besov spaces $B_s^{p,q}(\mathbb{R}^d)$ or the Triebel-Lizorkin spaces $F_s^{p,q}(\mathbb{R}^d)$. The question of which classical function spaces embed into the modulation spaces or vice versa is a natural one. The Sobolev embedding theorem $H^s(\mathbb{R}^d) \subseteq C^k(\mathbb{R}^d)$ for $s > k + \frac{d}{2}$ is an example of such a result, as $H^s(\mathbb{R}^d)$ is a modulation space (with $p = 2$) and $C^k(\mathbb{R}^d)$ is a Besov space. The first systematic study of embeddings of Besov spaces into modulation spaces was by Okoudjou [39], with further results by Toft [49], [50].

One of the classical spaces in harmonic analysis is the BMO space (the space of functions of bounded mean oscillation). In the literature there is no result making a connection between BMO and modulation spaces. Here we present the relation between modulation spaces and BMO. The following theorem is the one of the main results related to the modulation spaces in this thesis.

Theorem 1.1.2. *Let $1 \leq p \leq 2$ and $\widetilde{m}$ be a v-moderate weight such that $v(z) \leq (1+|z|)^{N_0}$ for some $N_0 \in \mathbb{N}$ and $\widetilde{m}$ satisfies $\widetilde{m}(x,w) \geq C'\sqrt{1+|w|^2}$ for some $C' > 0$. Then $M^p_{\widetilde{m}}(\mathbb{R}^2) \subset VMO(\mathbb{R}^2)$, with*

$$\|f\|_{BMO} \leq C\|f\|_{M^p_{\widetilde{m}}(\mathbb{R}^2)}.$$

This embedding theorem is important not only because it is in essence a new uncertainty principle for Gabor frames but also because it implies a relation between modulation spaces and BMO. In harmonic analysis, BMO, which was introduced by Nirenberg (1961), plays the same role in the theory of Hardy spaces that the space L^∞ of bounded functions plays in the theory of L^p-spaces.

The space VMO of functions of vanishing mean oscillation is the closure in BMO of the continuous functions that vanish at infinity. It can also be defined as the space of functions whose "mean oscillations" on balls Q are not only bounded, but also tend to zero uniformly as the radius of the ball Q tends to 0 or infinity. The space VMO is a sort of Hardy space analogue of the space of continuous functions vanishing at infinity, and in particular the Hardy space H^1 is the dual of VMO.

Another result in this chapter is about the Zak transform. One of the important tools in time-frequency analysis is the Zak transform, also known as Weil-Brezin map (first introduced by Gelfand [26]). The Zak transform of a function f is the function on $\mathbb{R}^{2d}$ defined (almost everywhere) by

$$Zf(x,w) = \sum_{k\in\mathbb{Z}^d} f(x-k)e^{2\pi i k\cdot w},$$

where the series converges unconditionally in $L^2(\mathbb{R}^d)$. The Zak transform of a function is quasi-periodic, and the Zak transform is a unitary transform from $L^2(\mathbb{R}^d)$ to $L^2([0,1)^{2d})$ where $[0,1)^{2d}$ is the unit cube. The Zak transform of a function provides a useful time-frequency presentation of a function.

The Zak transform plays an important role in uncertainty principles for Gabor systems. Given $f \in L^2(\mathbb{R})$ and positive constants α, β, the associated Gabor system is

$$\mathcal{G}(f,\alpha,\beta) = \{e^{2\pi i m\beta x} f(x-n\alpha)\}_{m,n\in\mathbb{Z}} \subset L^2(\mathbb{R}).$$

These systems are useful, for instance, in analyzing audio signals with challenges such as extracting different instruments from a recording of a symphony. It is possible to analyze and synthesize a signal without much loss using a Gabor system with certain extra properties, especially if it is a frame. A frame for a Hilbert space $\mathcal{H}$ is a collection $\{f_n\} \subset \mathcal{H}$ satisfying

$$A\|x\|^2_{\mathcal{H}} \leq \sum_n |\langle x, f_n\rangle|^2 \leq B\|x\|^2_{\mathcal{H}}$$

for all $x \in \mathcal{H}$ and some $A, B > 0$. The necessary and sufficient conditions on f, α, and β for $\mathcal{G}(f,\alpha,\beta)$ to be a frame have been investigated extensively. A central question is which function's time-frequency shifts form a frame for $\alpha = \beta = 1$. This particular case is important because the redundancy of a Gabor system is eliminated when $\alpha = \beta = 1$. Some criteria have been given for this (see [17]). For a function to generate a Gabor frame, $|Zf|$ should be bounded below and above by nonzero constants. To classify the functions which generate certain Gabor systems or do not generate certain Gabor systems, the behaviour of Zak Transform on the function spaces should be clarified. For example, the functions whose Zak Transform is continuous are important because of the fact that if the Zak transform of a function is continuous, then the Zak transform of the function has a zero. One of the spaces which allows the continuity of tha Zak transform is the amalgam spaces and another one is the modulation space $M^{1,1}(\mathbb{R}^d)$. Gautam [25] gave a result which tells us that the Zak transform of functions having certain decay and smoothness properties lies in some inhomogeneous Sobolev space which can be embedded into the space $\mathrm{BMO}(\mathbb{R}^d)$. Surprisingly, however, not much is known about the behavior of the other modulation spaces under the Zak transform. One of the results presented in this thesis is the following theorem which gives an idea how the Zak transform acts on a modulation space (Also see [30]).

Theorem 1.1.3. *Let $f \in L^2(\mathbb{R}^d)$ and let $f \in M_m^{p,p}(\mathbb{R}^d)$ be given where $1 \leq p < \infty$, m is a v-moderate weight and $|v(z)| \leq (1+|z|)^{N_0}$ for some $N_0 \in \mathbb{N}$. Let $\psi \in C_c^\infty(\mathbb{R}^d)$ have compact support. Then $\psi Zf \in M_{m^*}^p(\mathbb{R}^{2d})$ where for each N there exists a $C_N > 0$ such that $m^*(N, M) \leq C_N m(M)$ for all $M \in \mathbb{Z}^{2d}$.*

This result provides a new insight in the general theory of modulation spaces in time-frequency analysis, and it is another result showing the relation between modulation spaces.

The Balian-Low Theorem can be regarded as a cornerstone of time-frequency analysis and is the first uncertainty principle for Gabor frames.

Theorem 1.1.4. *[Balian-Low-Coifman-Semmes] Let $f \in L^2(\mathbb{R})$. If $f \in \mathbb{H}^1(\mathbb{R})$ and $\hat{f} \in \mathbb{H}^1(\mathbb{R})$, then $\mathcal{G}(f, 1, 1)$ is not a frame for $L^2(\mathbb{R})$.*

In the literature one can find different versions of uncertainty principles for certain Gabor systems ([9, 37, 14, 12, 13, 25]). The underlying fact behind this theorem is that the Zak transform of a function satisfying the assumption in the theorem lies in VMO space and this forces the Zak transform have a zero. Our first result about modulation spaces and the second result about the Zak transform imply an uncertainty principle for Gabor frames via modulation spaces.

Theorem 1.1.5. *Let $f \in L^2(\mathbb{R}) \cap M^p_m(\mathbb{R})$ where $1 \le p \le 2$ and m is v-moderate weight where $|v(z)| \le (1+|z|)^{N_0}$ for some $N_0 \in \mathbb{N}$ and $m(x) \ge C\sqrt{1+|x|^2}$ for some C. Then $\mathcal{G}(f, 1, 1)$ is not a frame for $L^2(\mathbb{R})$.*

1.2 Outline of Part I

The first part is organized as follows.

Chapter 2 begins with a section about the short time Fourier transform, which is the main tool used in the area of modulation spaces. In the second section, the modulation spaces as well as some related spaces are defined. The first and second sections are mostly expository. The embedding of modulation spaces into BMO, which is the first main result of the thesis, is given in the third section.

In Chapter 3, the Zak transform and its properties are given as well as the characterization of modulation spaces via the Zak transform and vice versa is obtained. The first section of Chapter 3 has the second main result, which states that the Zak transform maps modulation spaces into the modulation spaces of higher dimension. In the second section, the main results given in the second chapter and in the previous section are used to give an uncertainty result for Gabor frames via modulation spaces.

1.3 Part II Introduction

The simplest and most well-known area of optimization is convex optimization. The fundamental tool in the study of convex optimization problems is the subgradient. The subgradient permits the construction of an affine function, which does not exceed f over the entire space and coincides with f at a point. This affine function is called a support function and this support function is less than or equal to f at every point. The existence of such a function follows from the well-known separation theorem for convex sets: each point which does not belong to a convex closed set can be separated from this set by a linear function. This theorem also leads to the following fundamental result of convex analysis ([40]).

Theorem 1.3.1. *A lower semicontinuous convex function is the pointwise supremum of the collection of all affine functions h such that $h \leq f$.*

This presentation of convexity has been extended to non-convex functions underlying the theory of convexity without linearity, known also as abstract convexity (see [**?**], [44], [42]).

Definition 1.3.2. [42] Let H be a set of finite functions defined on a set X. A function $f : X \to \bar{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$ is called H-convex (abstract convex) with respect to H if f can be written in the following form:

$$f(x) = \sup\{h(x)|h \in H,\ h \leq f\},\ \forall x \in X.$$

According to this definition, a lower semicontinuous convex function is abstract convex with respect to affine functions. Many results from convex analysis related to subdifferentials can be transformed into abstract convex environments.

Abstract convexity has found many applications in the study of problems of optimization. For applications of abstract convexity to nonconvex optimization, the objective function should be related to some class of elementary functions via abstract convexity just as lower semicontinuous convex functions are related to affine functions. This problem leads to the notion of a supremal generator [35].

Definition 1.3.3. A supremal generator of a set X of functions is a subset of H of X such that each function from X is abstract convex with respect to H.

Although there are general results about supremal generators, we will mention certain types of functions whose supremal generators are shifted min-type functions: ICAR functions.

Definition 1.3.4. Let $Q \subset \mathbb{R}^d$ be a conic set of positive octant. A function $f : Q \to \mathbb{R}^d_+$ is called increasing convex-along-rays (ICAR) if

i) f is increasing i.e. $x \geq y$ implies $f(x) \geq f(y)$ ($x \geq y$ iff $x_i \geq y_i$ for $i = 1, \cdots, n$)

ii) for each $x \in Q$ the function of one variable $f_x(t) = f(tx)$, $t \in [0, \infty)$, is convex.

ICAR functions are abstract convex with respect to shifted min-type functions.

Proposition 1.3.5. *Let $H = \{h : h(x) = \min\limits_{i=1,\cdots,n} a_i x_i - c, a_i > 0, c \in \mathbb{R}\}$. A finite function f is abstract convex with respect to H if and only if f is ICAR.*

The class of ICAR functions is very broad. In particular, each Lipschitz function defined on the unit simplex $S = \{x \in \mathbb{R}^d_+ : \sum\limits_{i=1}^{n} x_i = 1\}$ can be extended to a finite ICAR function defined on $\mathbb{R}^d_+$ ([41]) .

Theorem 1.3.6. *Let f be a Lipschitz and strictly positive function defined on the simplex S and $K = \sup\limits_{\substack{x \neq y \in S \\ x,y \in S}} \frac{|f(x)-f(y)|}{\|x-y\|}$ where the norm is ℓ_1 norm. Let $c = \min\limits_{x \in S} f > 0$. If $p \geq \max\{1, \frac{2K}{c}\}$ then the function*

$$g(x) = \begin{cases} f\left(\frac{x}{\sum\limits_{i=1}^{n} x_i}\right)\left(\sum\limits_{i=1}^{n} x_i\right)^p & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

is an ICAR function.

This fact allows one to use ICAR functions in the minimization of Lipschitz functions over the unit simplex. The methods and some examples of this is given in [42].

One subclass of ICAR functions is the class of increasing positively homogeneous (IPH) functions.

Definition 1.3.7. A function f defined on $\mathbb{R}^n_+$ is called increasing positively homogeneous of degree one (for short IPH), if

a) for $x, y \in \mathbb{R}^n_+$, $x \geq y$ implies $f(x) \geq f(y)$;

b) $f(\lambda x) = \lambda f(x)$ for all $x \in \mathbb{R}^n_+$ and $\lambda > 0$.

An IPH function is abstract convex with respect to min-type functions. Min-type functions are the functions $f(x_1, \cdots, x_n) = \min_{i=1,\cdots,n} a_i x_i$ where $a_i > 0$ are constants. The class of the IPH functions is fairly large and has many different application areas [42].

In minimizing IPH functions over the unit simplex, cutting angle ([3, 4]) method is applied by taking advantage of abstract convexity. The cutting angle method reduces an original global optimization problem to a sequence of subproblems. The main subproblem is to minimize a function which is the maximum of a finite collection of min-type functions, i.e.,

$$\min \max_{1\leq j\leq k} \{ \min_{1\leq i\leq n} \{ \frac{x_i}{a_{ij}} \} \} \quad \text{subject to} \quad x \in S$$

where $a_{ij} > 0$.

The solution of this subproblem is central to the execution of the method. Since the efficiency of the method strongly depends on the technique for the solution of the auxiliary problem, some studies have been done on this subproblem [1, 5, 6, 38]. We present a new algorithm for the solution of this subproblem [2]. The new approach is based on the geometric observation of the behavior of $h_j(x)$. The approach is first described in $\mathbb{R}^2$ then is extended to higher dimensions. Then we develop Algorithm 5.1.11, an algorithm for solving the subproblem. This algorithm is based on Theorem 5.1.7. Theorem 5.1.7 is proved by establishing characterizations of local minima of $h_j(x)$ over the unit simplex. Some experiments have been demonstrated using test problems with IPH objective functions. These functions are also used in [42, 6]. The following theorem is the main result of Part II. It is the theoretical basis of the proposed algorithm

Theorem 1.3.8. *[2] Let $S \subset \mathbb{R}^n_+$ be the unit simplex and $P(x) = \max_{1\leq j\leq m} \{ \min_{1\leq i\leq n} \{ \frac{x_i}{a_{ij}} \} \}$ be defined on $\mathbb{R}^n_+$ where $m \geq n$. The point x^* in the relative interior of S is a local minimum of $P(x)$ if and only if there exist n indices $k_1, \cdots, k_n \in \{1, \cdots, m\}$ such* *that*

a) $a_{ik_i} > \max\limits_{\substack{1 \geq j \geq n \\ j \neq i}} a_{ik_j}$ *for all* $i = 1, \cdots, n,$

and

b) for $j \neq k_i$ *there exists* $1 \geq r \geq n$ *such that* $a_{rj} > a_{sk_s}$ *for some* $1 \geq s \geq n.$

Moreover, the local minimum x^* *satisfies the following equality;*

$$\frac{x_1^*}{a_{1k_1}} = \cdots = \frac{x_n^*}{a_{nk_n}}$$

This theorem gives a criteria for finding all local minima. The algorithm based on this theorem basically finds n min-type functions whose coefficients satisfy a) and b) in the theorem, then finds a local minimum. The global minimum is decided after finding all local minima.

The method suggested by this theorem in minimization of IPH functions over the unit simplex can also be adapted to solve the problem of minimizing increasing convex-along-rays (ICAR) functions over the unit simplex, which by the previous theorem leads to a new algorithm for Lipschitz programming over the unit simplex. However, since ICAR functions are abstract convex with respect to shifted min-type functions, the adaptation is not easy.

1.4 Outline of Part II

In Chapter IV, min-type functions and IPH functions are defined and some of their properties and some examples are given. The main relation between IPH functions and min-type functions is that IPH is abstract convex with respect to min-type functions. Then the main problem is stated and the cutting angle method is established. The most important step in the implementation of this algorithm is the minimization of a max-min function. The solution of this subproblem is given in Chapter V. In the first section of Chapter V, the necessary and sufficient conditions for a local minimum point is established in Theorem 5.1.7 which enables us to develop an algorithm for finding the minimum of the subproblem. The proof of this theorem is given by a sequence of propositions in the first section. In the second section the results of some numerical experiments are established and the third section concludes the chapter.

PART I

Modulation Spaces, BMO and the Zak Transform

CHAPTER II

MODULATION SPACES AND BMO

2.1 Notation and Preliminary Concepts

$\mathbb{R}^d$ denotes the usual d dimensional Euclidean space, and $\mathbb{Z}^d$ ($\mathbb{Z}^d_+$) denotes the set of vectors in $\mathbb{R}^d$ whose components are integers (positive integers). Q denotes the unit cube

$$[0,1)^d = \underbrace{[0,1) \times \cdots \times [0,1)}_{d \text{ times}}.$$

Depending on the context Q will also denote $[0,1)^{2d}$.

For $x, y \in \mathbb{R}^d$, $x \cdot y$ denotes the usual inner product $x \cdot y = x^{(1)}y^{(1)} + \cdots + x^{(d)}y^{(d)}$. For any $x \in \mathbb{R}^d$, $x^{(r)}$ denotes the r^{th}component of x.

The notation $|\cdot|$ will stand for different meanings depending on the context. For a vector $x \in \mathbb{R}^d$, $|x| = \sqrt{x \cdot x}$. For a measurable set $E \subseteq \mathbb{R}^d$, $|E|$ denotes the Lebesgue measure of E. For a complex number a, $|a|$ denotes the magnitude of a complex number. For a multi-index $\alpha = (\alpha_1, \cdots, \alpha_d)$ we write $|\alpha| = \sum_{i=1}^d \alpha_i$.

For $1 \leq p < \infty$, p' will denote the conjugate of p, i.e, $\frac{1}{p} + \frac{1}{p'} = 1$. $L^p(\mathbb{R}^d)$ denotes the Banach space of complex valued functions f on $\mathbb{R}^d$ with norm

$$\left(\|f\|_p = \int_{\mathbb{R}^d} |f(x)|^p \right)^{\frac{1}{p}}$$

for $1 \leq p \leq \infty$. For $p = \infty$, the norm is given by

$$\|f\|_\infty = ess \sup_{x \in \mathbb{R}^d} |f(x)|.$$

For a set S, χ_S denotes the characteristic function of S.

If $f\chi_K \in L^p(\mathbb{R}^d)$ for any compact set $K \subset \mathbb{R}^d$, we write $f \in L^p_{loc}(\mathbb{R}^d)$.

For the case $p = 2$, $L^2(\mathbb{R}^d)$ turns into a Hilbert Space with the inner product

$$\langle f, g\rangle = \int_{\mathbb{R}^d} f(x)\overline{g(x)}dx.$$

For a positive function $m(x)$, $L^p_m(x)$ denotes the weighted $L^p(\mathbb{R}^d)$ space with the norm

$$\|f\|_{L^p_m} = \left(\int_{\mathbb{R}^d} |f(x)m(x)|^p\right)^{\frac{1}{p}}.$$

We use the notation $f^*(t) = \overline{f(-t)}$.

The convolution of f and g is denoted by

$$(f * g)(x) = \int_{\mathbb{R}^d} f(x-t)g(t)dt.$$

The Fourier transform of a function $f \in L^1(\mathbb{R}^d)$ is defined as

$$\widehat{f}(w) = \int_{\mathbb{R}^d} f(x)e^{-2\pi i x\cdot w}dx$$

for $w \in \mathbb{R}^d$. The Fourier transform is extended to $L^2(\mathbb{R}^d)$ by the density of $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ in $L^2(\mathbb{R}^d)$. When we want to emphasize that the Fourier transform is a linear operator acting on a function space, we write $\mathcal{F}f$ instead of $\widehat{f}$.

In signal processing, the variable $x \in \mathbb{R}^d$ often signifies "time". $f(x)$ is the amplitude of the signal, and the variable w in $\widehat{f}(w)$ is called "frequency".

The Fourier coefficients of a function $f \in L^1(Q)$ are defined as

$$\widehat{f}(n) = \textstyle\int_Q f(x)e^{-2\pi i n\cdot x}dx, \text{ for } n \in \mathbb{Z}^d.$$

For $x, w \in \mathbb{R}^d$, we define T_x to be translation by x,

$$T_x f(t) = f(t-x)$$

and M_w is modulation by w,

$$M_w f(t) = e^{2\pi i w\cdot t} f(t).$$

$M_w T_x f(t)$ or $T_x M_w f(t)$ are called time-frequency shifts of $f(t)$. The following equality follows easily:

$$T_x M_w f(t) = e^{-2\pi i x \cdot w} M_w T_x f(t).$$

Let X be a Banach space. Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent on X, if there exist constants $C_1, C_2 > 0$ such that

$$C_1 \|u\|_1 \leq \|u\|_2 \leq C_2 \|u\|_1$$

for any $u \in X$.

The dual of X (the set of continuous linear functionals on X) is denoted by X'. Given a linear map $h \in X'$ ($h : X \to \mathbb{R}$) its action on $g \in X$ is written $\langle h, g \rangle$ where the symbol $\langle h, g \rangle$ stands for $h(g)$ (and also the usual inner product whenever it makes sense).

The support of a function f on $\mathbb{R}^d$ is defined as $supp(f) = \{x \in \mathbb{R}^d : f(x) \neq 0\}$. $C_c^\infty(\mathbb{R}^d)$ is the space of functions on $\mathbb{R}^d$ which are infinitely differentiable and have compact support.

We say the derivative of a function f exists in the weak sense if there exists a function g such that

$$\langle f, \psi' \rangle = -\langle g, \psi \rangle \text{ for all } \psi \in C_c^\infty(\mathbb{R}^d).$$

The differentiation operator D^α and the multiplication operator X^β are defined as

$$D^\alpha f(x) = \prod_{i=1}^d (\partial_{x_i})^{\alpha_i} f(x) \text{ and } X^\beta f(x) = \sum_{i=1}^d x_i^{\beta_i} f(x)$$

where α and β are multi-indices.

$\mathcal{S}(\mathbb{R}^d)$ is the Schwartz space of all infinitely differentiable functions f on $\mathbb{R}^d$ such that $\sup_{x \in \mathbb{R}^d} |D^\alpha X^\beta f(x)| < \infty$ for all $\alpha, \beta \in \mathbb{Z}_+^d$. The elements of its topological dual, $\mathcal{S}'(\mathbb{R}^d)$ are called the tempered distributions. The Fourier transform $\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \to$

$\mathcal{S}(\mathbb{R}^d)$ is a continuous bijection. By duality the Fourier transform can be extended to $\mathcal{S}'(\mathbb{R}^d)$ by

$$\langle \widehat{f}, \widehat{\phi} \rangle = \langle f, \phi \rangle \text{ for } \phi \in \mathcal{S}(\mathbb{R}^d), f \in \mathcal{S}'(\mathbb{R}^d).$$

Given an open set where $U \subseteq \mathbb{R}^d$, is an open set. The Sobolev space $W^{1,p}(U)$ consists of all locally integrable functions $f : U \to \mathbb{R}$ such that each partial derivative of f exists in the weak sense and belongs to $L^p(U)$

The Bessel potential spaces are defined as

$$H^s(\mathbb{R}^d) = \left\{ f \in \mathcal{S}'(\mathbb{R}^d) : \|f\|_{H^s} = \left(\int_{\mathbb{R}^d} |\widehat{f(w)}|^2 (1 + |w|^2)^s dw \right)^{\frac{1}{2}} < \infty \right\}.$$

If $s \geq 0$, then $H^s(\mathbb{R}^d)$ is a subspace of $L^2(\mathbb{R}^d)$ that consists of smooth functions in $L^2(\mathbb{R}^d)$. For $s < 0$, $L^2(\mathbb{R}^d) \subseteq H^s(\mathbb{R}^d)$, and in general $H^s(\mathbb{R}^d)$ also contains measures and distributions when s is negative.

To avoid dealing with too many intermediate constants, sometimes we will write $A(x) \lesssim B(x)$ to mean that $A(x) \leq CB(x)$ for all $x \in \mathbb{R}^d$ where C is a constant independent of x.

More details on the basic properties of the Fourier transform and more generally, the parts of real analysis and functional analysis which is required to understand this part of the thesis can be found in many standard texts, e.g., [23, 33].

2.2 Short Time Fourier Transform

In order to get information about local properties of a function f, in particular about its local frequency spectrum, the short time Fourier transform (or windowed Fourier transform) is defined. Taking a (usually compactly supported) function g and restricting f with this function g allows this localization. Then the Fourier transform is taken so that the frequency information for the localization of f is obtained.

Definition 2.2.1. Let $g \neq 0$ (called the window function). Then the short time Fourier transform of a function f with respect to g is defined as

$$V_g f(x,w) = \int_{\mathbb{R}^d} f(t)\overline{g(t-x)}e^{-2\pi i t\cdot w}dt$$

for $x, w \in \mathbb{R}^d$ whenever this is wel-defined

When $f, g \in L^2(\mathbb{R}^d)$, $V_g f$ satisfies some properties.

Lemma 2.2.2. *If $f, g \in L^2(\mathbb{R}^d)$ then $V_g f$ is uniformly continuous on $\mathbb{R}^{2d}$ and*

$$\begin{aligned}
V_g f(x,w) &= \widehat{(fT_x\overline{g})}(w) \\
&= \langle f, M_w T_x g\rangle \\
&= \langle \hat{f}, T_w M_{-x}\hat{g}\rangle \\
&= e^{-2i\pi x\cdot w}\widehat{(\hat{f}T_w\overline{\hat{g}})}(-x) \\
&= e^{-2i\pi x\cdot w}V_{\hat{g}}\hat{f}(w,-x) \\
&= e^{-2i\pi x\cdot w}(f * M_w g^*)(x) \\
&= (\hat{f} * M_{-x}\hat{g}^*)(w) \\
&= e^{-i\pi x\cdot w}\int_{\mathbb{R}^d} f(t+\frac{x}{2})\overline{g(t-\frac{x}{2})}e^{-2i\pi t\cdot w}dt.
\end{aligned}$$

The symmetric variation $\int_{\mathbb{R}^d} f(t+\frac{x}{2})\overline{g(t-\frac{x}{2})}e^{-2i\pi t\cdot w}dt$ is often called the cross-ambiguity function. It plays an important role in radar and in optics ([16, 53]). Except for the phase factor $e^{-i\pi x\cdot w}$, it coincides with the short time Fourier transform.

The formula

$$V_g f(x,w) = e^{-2i\pi x\cdot w}V_{\hat{g}}\hat{f}(w,-x)$$

is the fundamental identity of time-frequency analysis. It combines both f and $\hat{f}$ into a joint time-frequency representation. In this representation the Fourier transform amounts to a rotation of the time-frequency plane by an angle of $\frac{\pi}{2}$.

In the previous lemma, f, g are assumed to be in $L^2(\mathbb{R}^d)$. However, the definition of short time Fourier transform accepts more generality on f and g as long as the definition makes sense. In this case the following property holds.

Lemma 2.2.3. *Whenever $V_g f$ is defined, we have*

$$V_g(T_u M_v f)(x,w) = e^{-2\pi i u\cdot w} V_g f(x-u, w-v)$$

for $x, w, u, v \in \mathbb{R}^d$.

In a more general setting, one can take $f \in \mathcal{S}'$ and $g \in \mathcal{S}$.

The short time Fourier transform enjoys several properties similar to those satisfied the ordinary Fourier Transform. The following theorem on inner products of the short time Fourier transforms corresponds to Parseval's formula.

Theorem 2.2.4. *Let $f_1, f_2, g_1, g_2 \in L^2(\mathbb{R}^d)$. Then $V_{g_1} f_1, V_{g_2} f_2 \in L^2(\mathbb{R}^{2d})$ and*

$$\langle V_{g_1} f_1, V_{g_2} f_2 \rangle_{L^2(\mathbb{R}^{2d})} = \langle f_1, f_2 \rangle \overline{\langle g_1, g_2 \rangle}.$$

The following corollary is immediate.

Corollary 2.2.5. *If $f, g \in L^2(\mathbb{R}^d)$, then*

$$\|V_g f\|_2 = \|f\|_2 \|g\|_2.$$

In particular, if $\|g\|_2 = 1$ then

$$\|f\|_2 = \|V_g f\|_2$$

for all $f \in L^2(\mathbb{R}^d)$. Thus, in this case the short time Fourier transform $f \longrightarrow V_g f$ is an isometry from $L^2(\mathbb{R}^d)$ into $L^2(\mathbb{R}^{2d})$.

It is possible to reconstruct a function from its short time Fourier transform. However, the following inversion formula holds in a weak sense, i.e., $f = g$ means $\langle f, h \rangle = \langle g, h \rangle$ for all $h \in L^2(\mathbb{R}^d)$.

Corollary 2.2.6. *Suppose that $g, \gamma \in L^2(\mathbb{R}^d)$ and $\langle g, \gamma \rangle \neq 0$. Then for all $f \in L^2(\mathbb{R}^d)$,*

$$f = \frac{1}{\langle \gamma, g \rangle} \int \int_{\mathbb{R}^{2d}} V_g f(x,w) M_w T_x \gamma \, dw \, dx.$$

Another fact about the short-time Fourier transform is that a function cannot be concentrated on small sets in the time frequency plane, no matter which time-frequency representation is used.

Proposition 2.2.7. *Suppose that* $\|f\|_2 = \|g\|_2 = 1$ *and that* $U \subseteq \mathbb{R}^{2d}$ *and* $\epsilon \geq 0$ *are such that*

$$\iint_U |V_g f|^2 \, dx \, dw \geq 1 - \epsilon.$$

Then $|U| \geq 1 - \epsilon$.

A much deeper and stronger inequality for the short time Fourier Transform was proved by E. Lieb [36]:

Theorem 2.2.8. *If* $f, g \in L^2(\mathbb{R}^d)$ *and* $2 \leq p < \infty$*, then*

$$\int \int_U |V_g f|^p dx dw \leq \left(\frac{2}{p}\right)^d (\|f\|_2 \|g\|_2)^p.$$

A careful analysis of the minimizing functions in the sharp version of Young and Hausdorff-Young shows that equality in Lieb's uncertainty principle is obtained if and only if f and g are time-frequency shifts of Gaussian functions [36].

Lieb's inequality improves the Proposition 2.2.7 and yields a sharper estimate for the essential support of $V_g f$.

Theorem 2.2.9. *Suppose that* $\|f\|_2 = \|g\|_2 = 1$*. If* $U \subseteq \mathbb{R}^d$ *and* $\epsilon \geq 0$ *are such that*

$$\int \int_U |V_g f(x, w)|^2 \, dx \, dw \geq 1 - \epsilon,$$

then

$$|U| \geq (1 - \epsilon)^{\frac{p}{p-2}} \left(\frac{p}{2}\right)^{\frac{2d}{p-2}}$$

for all $p > 2$.

Conversely, it can be also shown that if $|supp(V_g f)| < \infty$, then either $f = 0$ or $g = 0$ (See [31]).

2.3 Modulation Spaces and related spaces

To get a quantitative analysis of the distribution of the short time Fourier transform, the modulation spaces were defined. The modulation spaces were invented in 1983 by H. Feichtinger [19] and were subsequently investigated in [20, 22, 51].

To give a precise mathematical setting for the modulation spaces we begin with the definitions of weights and mixed-norm spaces.

Definition 2.3.1. A weight function v on $\mathbb{R}^{2d}$ is called submultiplicative if

$$v(z_1 + z_2) \leq v(z_1)v(z_2)$$

for all $z_1, z_2 \in \mathbb{R}^{2d}$.

A weight function m on $\mathbb{R}^{2d}$ is v-moderate if $m(z_1 + z_2) \leq Cv(z_1)m(z_2)$, for all $z_1, z_2 \in \mathbb{R}^{2d}$.

Two weights m_1, m_2 are equivalent if $C^{-1}m_1(z) \leq m_2(z) \leq Cm_1(z)$ for all $z \in \mathbb{R}^{2d}$.

We will require weight functions to be symmetric, i.e,

$$m(x, y) = m(-x, y) = m(x, -y) = m(-x, -y)$$

for $x, y \in \mathbb{R}^d$.

Example 2.3.2. The standard class of weights on $\mathbb{R}^{2d}$ are weights of polynomial type:

$$v_s(z) = (1 + |z|)^s$$

where $z = (x, w) \in \mathbb{R}^{2d}$ and $s \geq 0$. We note that $v_s(z)$ is equivalent to the weights $(1 + |x| + |w|)^s$ and $(1 + |z|^2)^{\frac{s}{2}}$.

One can use the exponential weights $v_s(z) = e^{\alpha|z|^\beta}$ for some $\alpha > 0$ and $0 \leq \beta < 1$ to quantify a faster decay.

Next we give the definition and properties of $L^{p,q}_m(\mathbb{R}^{2d})$, which is a generalization of $L^p(\mathbb{R}^{2d})$.

Definition 2.3.3. Let m be a weight function on $\mathbb{R}^{2d}$ and let $1 \leq p, q < \infty$. Then the weighted mixed-norm space $L_m^{p,q}(\mathbb{R}^{2d})$ consists of all (Lebesgue) measurable functions on $\mathbb{R}^{2d}$ such that the norm

$$\|F\|_{L_m^{p,q}(\mathbb{R}^{2d})} = \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |F(x,w)|^p m(x,w)^p dx \right)^{\frac{q}{p}} dw \right)^{\frac{1}{q}}$$

is finite. If $p = \infty$ or $q = \infty$, then the corresponding p-norm is replaced by the essential supremum.

In the above definition, one finds the classical $L^p(\mathbb{R}^{2d})$ when $p = q$. The mixed-norm spaces have the same basic properties as the classical L^p spaces.

Lemma 2.3.4. *Let m be v-moderate weight and let $1 \leq p, q \leq \infty$.*

(a) $L_m^{p,q}(\mathbb{R}^{2d})$ *is a Banach space.*

(b) $L_m^{p,q}(\mathbb{R}^{2d})$ *is invariant under translations T_z, $z \in \mathbb{R}^{2d}$ and*

$$\|T_z F\|_{L_m^{p,q}(\mathbb{R}^{2d})} \leq v(z) \|F\|_{L_m^{p,q}(\mathbb{R}^{2d})}.$$

(c) Hölder's inequality: If $F \in L_m^{p,q}(\mathbb{R}^{2d})$ and $H \in L_{1/m}^{p',q'}$, where $\frac{1}{p} + \frac{1}{p'} = 1$ and $\frac{1}{q} + \frac{1}{q'} = 1$, then $F \cdot H \in L^1(\mathbb{R}^{2d})$ and

$$\left| \int_{\mathbb{R}^{2d}} F(z) \overline{H(z)} dz \right| \leq \|F\|_{L_m^{p,q}(\mathbb{R}^{2d})} \|H\|_{L_{1/m}^{p',q'}}.$$

(d) Duality: If $1 \leq p, q < \infty$, then $(L_m^{p,q}(\mathbb{R}^{2d}))' = L_{1/m}^{p',q'}$. The duality is given by

$$\langle F, H \rangle = \int_{\mathbb{R}^{2d}} F(u) \overline{H(u)} du$$

where $F \in L_m^{p,q}(\mathbb{R}^{2d})$ and $H \in L_{1/m}^{p',q'}$.

The convolution properties in classical L^p spaces holds for $L_m^{p,q}(\mathbb{R}^{2d})$ spaces. The convolution of an $L_m^{p,q}(\mathbb{R}^{2d})$ function with an $L_v^1(\mathbb{R}^{2d})$ function results a function in $L_m^{p,q}(\mathbb{R}^{2d})$ again.

Proposition 2.3.5. *(a) If m is v-moderate weight, $F \in L^1_v(\mathbb{R}^{2d})$, and $G \in L^{p,q}_m(\mathbb{R}^{2d})$, then*

$$\|F * G\|_{L^{p,q}_m(\mathbb{R}^{2d})} \leq C\|F\|_{L^1_v}\|G\|_{L^{p,q}_m(\mathbb{R}^{2d})}.$$

That is, $L^1_v * L^{p,q}_m(\mathbb{R}^{2d}) \subseteq L^{p,q}_m(\mathbb{R}^{2d})$.

(b) If $s > 2d$, then $L^\infty_{v_s} * L^\infty_{v_s} \subseteq L^\infty_{v_s}$ and

$$\|F * G\|_{L^\infty_{v_s}} \leq C_s\|F\|_{L^\infty_{v_s}}\|G\|_{L^\infty_{v_s}}.$$

Another space which arises in the discussion of modulation spaces is the discrete mixed-norm space.

Definition 2.3.6. $\ell^{p,q}_m(\mathbb{Z}^{2d})$ consists of all complex valued sequences $a = (a_{k,n})_{k,n\in\mathbb{Z}^d}$ for which the norm

$$\|a\|_{\ell^{p,q}_m} = \left(\sum_{n\in\mathbb{Z}^d}\left(\sum_{k\in\mathbb{Z}^d}|a_{kn}|^p m(k,n)^p\right)^{q/p}\right)^{1/q}$$

is finite, where m is a v-moderate weight.

We will need the following version of Young's Inequality for convolution in weighted sequence spaces (proved in [27]).

Proposition 2.3.7. *Let m be v-moderate weight. Given $a = (a_{k,n}) \in \ell^1_v(\mathbb{Z}^{2d})$ and $b = (b_{k,n}) \in \ell^{p,q}_m(\mathbb{Z}^{2d})$, we have*

$$\|a * b\|_{\ell^{p,q}_m} \lesssim \|a\|_{\ell^1_v}\,\|b\|_{\ell^{p,q}_m}.$$

Definition 2.3.8. A measurable function F on $\mathbb{R}^{2d}$ belongs to the amalgam space $W(L^{p,q}_m(\mathbb{R}^{2d}))$, if the sequence of local suprema

$$a_{kn} = \operatorname*{ess\,sup}_{x,w\in[0,1]^d}|F(x+k, w+n)| = \|F \cdot T_{(k,n)}\chi\|_\infty$$

belongs to $\ell^{p,q}_m(\mathbb{Z}^{2d})$. The norm on $W(L^{p,q}_m(\mathbb{R}^{2d}))$ is

$$\|F\|_{W(L^{p,q}_m(\mathbb{R}^{2d}))} = \|a\|_{\ell^{p,q}_m}.$$

The subspace of continuous functions is denoted by $W_0(\mathbb{R}^d)$.

The amalgam spaces amalgamate (mix) the local boundedness of functions with a global property and they have convenient sampling properties, as expressed in the following proposition.

Proposition 2.3.9. *If $F \in W(L_m^{p,q}(\mathbb{R}^{2d}))$ is continuous, then for all $\alpha, \beta > 0$ the restriction $F|_{\alpha\mathbb{Z}^d \times \beta\mathbb{Z}^d}$ is in $\ell_{\tilde{m}}^{p,q}$, where $\tilde{m}(k,n) = m(\alpha k, \beta n)$, and*

$$\|F|_{\alpha\mathbb{Z}^d \times \beta\mathbb{Z}^d}\|_{\ell_{\tilde{m}}{p,q}} \leq C_{\alpha,\beta}\|F\|_{W(L_m^{p,q}(\mathbb{R}^{2d}))}.$$

Now we give the definition of the modulation spaces.

Definition 2.3.10. Let $g \in \mathcal{S}(\mathbb{R}^d)$, m be a v-moderate weight function on $\mathbb{R}^{2d}$ and $1 \leq p, q \leq \infty$. Then the modulation space $M_m^{p,q}(\mathbb{R}^d)$ consists of all tempered distributions $f \in \mathcal{S}'(\mathbb{R}^d)$ such that $V_g(f) \in L_m^{p,q}(\mathbb{R}^{2d})$. The norm on $M_m^{p,q}(\mathbb{R}^d)$ is

$$\|f\|_{M_m^{p,q}(\mathbb{R}^d)} = \|V_g f\|_{L_m^{p,q}(\mathbb{R}^{2d})}.$$

If $p = q$ then we write M_m^p instead of $M_m^{p,p}$, and if $m(z) = 1$ on $\mathbb{R}^{2d}$, then we write $M^{p,q}$ and M^p for $M_m^{p,q}$ and M_m^p.

In this definition, we require $1 \leq p, q \leq \infty$. However, the modulation spaces with p or $q < 1$ are defined in the same way, but their treatment requires much more technicality. We only note that they are quasi-Banach spaces and the window function should be chosen carefully. See [34] for further information about them.

The modulation spaces coincides with some classical spaces for different choices of weights.

Proposition 2.3.11. *Let $g \in \mathcal{S}(\mathbb{R}^d) \setminus \{0\}$.*

(a) If $|f(x)| \leq C(1+|x|)^{-s}$ and $s > d$, then $|V_g f(x,w)| \leq C'(1+|x|)^{-s}$. If $|\hat{f}(w)| \leq C(1+|w|)^{-s}$ and $s \geq d$, then $|V_g f(x,w)| \leq (1+|w|)^{-s}$.

(b) If $m(x,w) = m(x)$, then $M_m^2(\mathbb{R}^d) = L_m^2(\mathbb{R}^d)$.

(c) If $m(x,w) = m(w)$, than $M^2_m(\mathbb{R}^d) = \mathcal{F}L^2_m(\mathbb{R}^d)$. In particular if $m_s(w) = (1+|w|)^{\frac{s}{2}}$ for some $s \in \mathbb{R}$, then $M^2_{m_s}$ coincides with the Bessel potential space $H^s(\mathbb{R}^d)$.

(d) Writing $v_s(z) = (1+|z|)^s$, $z \in \mathbb{R}^{2d}$, we have $\mathcal{S}(\mathbb{R}^d) = \bigcap_{s\geq 0} M^\infty_{v_s}$ and $\mathcal{S}'(\mathbb{R}^d) = \bigcup_{s\geq 0} M^\infty_{\frac{1}{v_s}}$.

For a choice of appropriate weights we can identify the space of functions with certain smoothness and decay properties as a modulation space.

Proposition 2.3.12. *Fix s_1, $s_2 > 0$.*

(a) Assume $1 + m_{s_1}(\omega)^2 + m_{s_2}(x)^2 \gtrsim m(x,\omega)^2$. If $f \in H^{s_1}(\mathbb{R}^d)$ and $\widehat{f} \in H^{s_2}(\mathbb{R}^d)$, then $f \in M^2_m(\mathbb{R}^d)$.

(b) Assume $m(x,\omega) \gtrsim \max\{m_{s_1}(\omega), m_{s_2}(x)\}$. If $f \in M^2_m(\mathbb{R}^d)$, then $f \in H^{s_1}(\mathbb{R}^d)$ and $\widehat{f} \in H^{s_2}(\mathbb{R}^d)$.

Proof. (a) The weight $m_{s_1}(\xi)$ is a weight on $\mathbb{R}^d$. If we think of it as a weight on $\mathbb{R}^{2d}$ that depends only on ω, then we have $H^{s_1}(\mathbb{R}^d) = M^2_{m_{s_1}(\omega)}(\mathbb{R}^d)$, where we write $m_{s_1}(\omega)$ to indicate that this weight depends on the frequency variable. Using a similar notation for m_{s_2}, if $f \in H^{s_1}(\mathbb{R}^d) = M^2_{m_{s_1}(\omega)}(\mathbb{R}^d)$ and $\widehat{f} \in H^{s_2}(\mathbb{R}^d) = M^2_{m_{s_2}(\omega)}(\mathbb{R}^d)$, then we have

$$\begin{aligned}
&\iint |V_g f(x,\omega)|^2\, m(x,\omega)^2\, dx\, d\omega \\
&\leq \iint |V_g f(x,\omega)|^2\, m_{s_1}(\omega)^2\, dx\, d\omega + \iint |V_{\widehat{g}} \widehat{f}(\omega,-x)|^2\, m_{s_2}(-x)^2\, dx\, d\omega \\
&= \iint |V_g f(x,\omega)|^2\, m_{s_1}(\omega)^2\, dx\, d\omega + \iint |V_g f(x,\omega)|^2\, m_{s_2}(x)^2\, dx\, d\omega \\
&\leq \iint |V_g f(x,\omega)|^2\, (1 + m_{s_1}(\omega)^2 + m_{s_2}(x)^2)\, dx\, d\omega.
\end{aligned}$$

(b) Given $f \in M^2_m(\mathbb{R}^d)$, we have

$$\iint |V_g f(x,\omega)|^2\, m_{s_1}(\omega)^2\, dx\, d\omega \leq \iint |V_g f(x,\omega)|^2\, m(x,\omega)^2\, dx\, d\omega,$$

so $f \in M^2_{m_{s_1}(\omega)}(\mathbb{R}^d) = H^{s_1}(\mathbb{R}^d)$. Also,

$$\iint |V_{\widehat{g}}\widehat{f}(\omega,-x)|^2\, m_{s_2}(-x)^2\, dx\, d\omega \le \iint |V_{\widehat{g}}\widehat{f}(\omega,-x)|^2\, m(x,\omega)^2\, dx\, d\omega$$
$$= \iint |V_g f(x,\omega)|^2\, m(x,\omega)^2\, dx\, d\omega,$$

so $\widehat{f} \in M^2_{m_{s_2}(x)}(\mathbb{R}^d) = H^{s_2}(\mathbb{R}^d)$. □

Corollary 2.3.13. *Fix $s_1,\ s_2 > 0$, and set $m(x,\omega) = \big(1 + m_{s_1}(\omega)^2 + m_{s_2}(x)^2\big)^{1/2}$. Then the following are equivalent:*

i. $f \in H^{s_1}(\mathbb{R}^d) \cap L^2_{s_2}(\mathbb{R}^d)$

ii. $f \in H^{s_1}(\mathbb{R}^d),\ \widehat{f} \in H^{s_2}(\mathbb{R}^d)$

iii. $f \in M^2_m(\mathbb{R}^d)$.

Proposition 2.3.14. *If $|m(z)| \le C(1+|z|)^N$ for some N, then $\mathcal{S}(\mathbb{R}^d)$ is a dense subspace of $M^{p,q}_m(\mathbb{R}^d)$ for any $1 \le p, q < \infty$.*

This proposition allows only polynomial weights in order to use Schwartz functions in the theory of modulation spaces. One should consider more general spaces, if an exponential weight is desired.

Theorem 2.3.15. *Let m be a v-moderate weight.*

(a) $M^{p,q}_m(\mathbb{R}^d)$ *is a Banach space for* $1 \le p, q \le \infty$.

(b) $M^{p,q}_m(\mathbb{R}^d)$ *is invariant under time-frequency shifts and*

$$\|T_x M_w f\|_{M^{p,q}_m(\mathbb{R}^d)} \le C v(x,w) \|f\|_{M^{p,q}_m(\mathbb{R}^d)}.$$

(c) If $p = q$ and $m(w,-x) \le C m(x,w)$, then $M^p_m(\mathbb{R}^d))$ is invariant under the Fourier transform.

Theorem 2.3.16. *If* $g \in M_v^1(\mathbb{R}^d)$ *and* $f \in M_m^{p,q}(\mathbb{R}^d)$, *then* $V_g f \in W(L_m^{p,q}(\mathbb{R}^{2d}))$ *and*

$$\|V_g f\|_{W(L_m^{p,q}(\mathbb{R}^{2d}))} \leq \|V_g g\|_{W(L_v^1)} \|f\|_{M_m^{p,q}(\mathbb{R}^d)}.$$

The estimation of the $W(L_m^{p,q}(\mathbb{R}^{2d}))$-norm of the short time Fourier transform of f by its $M_m^{p,q}(\mathbb{R}^d)$-norm makes the embedding between modulation spaces look like the embedding between $\ell^{p,q}(\mathbb{Z}^{2d})$ spaces. Further information can be found in [21].

Theorem 2.3.17. *If* $p_1 \leq p_2, q_1 \leq q_2$, *and* $m_2 \leq Cm_1$, *then* $M_{m_1}^{p_1,q_1}(\mathbb{R}^d) \subseteq M_{m_2}^{p_2,q_2}(\mathbb{R}^d)$.

2.4 Embedding of Modulation Spaces into BMO in $\mathbb{R}^2$

The space of bounded mean oscillations (BMO) plays an important role in harmonic analysis.

Definition 2.4.1. BMO($\mathbb{R}^d$) is the space of functions that have bounded mean oscillation on $\mathbb{R}^d$:

$$\mathrm{BMO}(\mathbb{R}^d) = \left\{ f \in L^1_{\mathrm{loc}}(\mathbb{R}^d) : \|f\|_{\mathrm{BMO}} = \sup_Q \left(\fint_Q \left| f(x) - \fint_Q f \right| dx \right) < \infty \right\},$$

where the supremum is taken over all cubes Q in $\mathbb{R}^d$, and

$$\fint_E g = \frac{1}{|E|} \int_E g$$

denotes the average of a function over a Lebesgue measurable set E.

Note that in the definition above, the functions which differ by a constant are represented by only one function in BMO($\mathbb{R}^d$).

BMO($\mathbb{R}^d$) is a Banach space. The functions in BMO($\mathbb{R}^d$) belong to L^p_{loc} if $0 < p < \infty$, but need not be locally bounded. We have $L^\infty(\mathbb{R}^d) \subset \mathrm{BMO}(\mathbb{R}^d)$.

The space VMO($\mathbb{R}^d$) of functions having vanishing mean oscillation on $\mathbb{R}^d$ consists of those functions $f \in \mathrm{BMO}(\mathbb{R}^d)$ such that

$$\lim_{a \to 0} \sup_{|Q| \leq a} \left(\fint_Q \left| f(x) - \fint_Q f \right| dx \right) = 0. \tag{1}$$

Equivalently, VMO($\mathbb{R}^d$) is the closure of the uniformly continuous functions in BMO-norm. We will work mostly with a local version of VMO; given a compact set K we define VMO(K) to be the set of functions satisfying the limit condition in equation (1) for cubes $Q \subset K$. For details on VMO, we refer to [43].

The importance of BMO in classical harmonic analysis comes from the fact that the dual of the Hardy space with $p = 1$ is BMO. In the same way, the Hardy space with $p = 1$ is the dual of VMO. See [18], [45] for a detailed information.

We will need an embedding of a certain class of functions defined on $\mathbb{R}^2$ into BMO($\mathbb{R}^2$). This embedding is a consequence of Poincare's inequality for functions defined on $\mathbb{R}^2$. The Sobolev space $W^{1,p}(U)$ consists of all locally integrable functions $f : U \to \mathbb{R}$ such that each partial derivative of f exists in the weak sense and belongs to $L^p(U)$.

Theorem 2.4.2 (Poincare's Inequality)**.** *If* $1 \le p \le \infty$ *then*

$$\left\| u - \fint_{B(x,r)} u \right\|_{L^p(B(x,r))} \lesssim r \, \|Du\|_{L^p(B(x,r))}$$

for each ball $B(x,r) \subseteq \mathbb{R}^d$ *and each function* $u \in W^{1,p}(B(x,r))$*, where* $W^{1,p}(B(x,r))$ *is the usual* L^p *Sobolev space defined on* $B(x,r)$*.*

Corollary 2.4.3. *For all* $u \in W^{1,2}(\mathbb{R}^2) \cap L^1(\mathbb{R}^2)$ *we have*

$$\|u\|_{\mathrm{BMO}(\mathbb{R}^2)} \lesssim \|Du\|_{L^2(\mathbb{R}^2)}.$$

Proof. If we take $p = 1$ and $d = 2$ in Poincare's Theorem and apply Hölder's Inequality, we obtain

$$\begin{aligned} \fint_{B(x,r)} \left| u(y) - \fint_{B(y,r)} u \right| dy &\lesssim r \fint_{B(x,r)} |Du| \\ &= \frac{r}{|B(x,r)|} \int_{B(x,r)} |Du| \\ &\le \frac{r}{|B(x,r)|} |B(x,r)|^{1/2} \left(\int_{B(x,r)} |Du| \right)^{1/2} \\ &\lesssim \left(\int_{\mathbb{R}^2} |Du|^2 \right)^{1/2}. \end{aligned}$$

□

Now we can prove our first main theorem, on the embedding of weighted modulation spaces into $\mathrm{VMO}(\mathbb{R}^2)$.

Theorem 2.4.4. *Fix $1 \le p \le 2$. Assume that:*

(a) v is a submultiplicative weight on $\mathbb{R}^4$ with $v(z) \lesssim (1+|z|)^{N_0}$ for some $N_0 \in \mathbb{N}$,

(b) $\widetilde{m}$ is a v-moderate weight on $\mathbb{R}^4$ such that $\widetilde{m}(x,\xi) \gtrsim (1+|\xi|^2)^{1/2}$.

Then $M^p_{\widetilde{m}}(\mathbb{R}^2) \subseteq \mathrm{VMO}(\mathbb{R}^2)$ with

$$\|f\|_{\mathrm{BMO}} \lesssim \|f\|_{M^p_{\widetilde{m}}(\mathbb{R}^2)}.$$

Proof. The Schwartz class $\mathcal{S}(\mathbb{R}^2)$ is dense in $M^p_{\widetilde{m}}(\mathbb{R}^2)$, so consider a fixed $f \in \mathcal{S}(\mathbb{R}^2)$. Setting $m(x) = (1+|x|^2)^{1/2}$ and applying Corollary 2.4.3 and the Plancherel Equality, we compute that

$$\begin{aligned}
\|f\|_{\mathrm{BMO}(\mathbb{R}^2)} &\lesssim \|Df\|_{L^2(\mathbb{R}^2)} \\
&\lesssim \|\, |\xi|\, \widehat{f}(\xi)\,\|_2 \\
&\lesssim \big\|(1+\|\xi|^2)^{1/2}\, \widehat{f}(\xi)\,\big\|_2 \\
&= \|\widehat{f}\,\|_{L^2_m}.
\end{aligned}$$

If we write $m(x,\xi) = m(x)$ then by Proposition 2.3.11(b) we have

$$\|\widehat{f}\,\|_{L^2_m} \lesssim \|\widehat{f}\,\|_{M^2_m}.$$

Now let $m'(x,w) = \widetilde{m}(-w,x)$. By assumption we have $m \lesssim m'$, which by Theorem 2.3.17 implies that

$$\|\widehat{f}\,\|_{M^2_m} \lesssim \|\widehat{f}\,\|_{M^p_{m'}} = \|f\|_{M^p_{\widetilde{m}}}.$$

An extension by density argument establishes that $\|f\|_{\mathrm{BMO}(\mathbb{R}^2)} \lesssim \|f\|_{M^p_{\widetilde{m}}}$ for all $f \in M^p_{\widetilde{m}}(\mathbb{R}^2)$. Moreover, since VMO is the closure of the uniformly continuous functions in BMO-norm, this also implies that $M^p_{\widetilde{m}}(\mathbb{R}^2) \subseteq \mathrm{VMO}(\mathbb{R}^2)$. $\square$

CHAPTER III

ZAK TRANSFORM AND MODULATION SPACES

3.1 Zak Transform on Modulation Spaces

In signal processing, there are practical situations where time-continuous signals are sampled at a certain uniform rate and at an unknown or uncontrolled sampling frequency. The fact that this sampling frequency is unspecified usually presents no problems when the sampling frequency is well above the Nyquist rate. However, for undersampled signals, the particular sampling frequency might play a significant role. To get detailed information about the sampling of time frequency, the Zak transform of a function $f \in L^2(\mathbb{R}^d)$ is defined.

Definition 3.1.1. For a given parameter $\alpha > 0$ the Zak transform $Z_\alpha f$ of f is the function defined on $\mathbb{R}^{2d}$ by

$$Z_\alpha f(x,\omega) = \sum_{k \in \mathbb{Z}^d} f(x - \alpha k)\, e^{2\pi i \alpha k \cdot \omega}.$$

In the sequel we will take $\alpha = 1$.

The Zak transform of a signal can be considered a mixed time-frequency representation of f. It is seen that for a fixed $x \in \mathbb{R}^d$, the Zak transform is a Fourier series where the coefficients are the samples of f, and for a fixed $w \in \mathbb{R}^d$ it behaves as a discrete Fourier transform of $f(x-t)$. On the other hand, the Zak Transform can be regarded as a version of the Poisson summation formula.

The Zak transform has many properties. Before giving the main properties, it is important to choose a suitable class of functions for which the Zak transform is well-defined.

Lemma 3.1.2. *(a) If $f \in L^1(\mathbb{R}^d)$, then $Zf \in L^1(Q \times Q)$.*

(b) If $f \in W(\mathbb{R}^d)$ then $Zf \in L^\infty(\mathbb{R}^{2d})$.

(c) If $f \in W_0(\mathbb{R}^d)$ then Zf is continuous on $\mathbb{R}^{2d}$.

(d) If $f \in L^2(\mathbb{R}^d)$,then Zf is defined almost everywhere and $Zf(x,w) \in L^2(Q,dw)$ for almost $x \in \mathbb{R}^d$.

The following properties of Zf follows from its definition.

i. (Quasiperiodicity of the Zak transform) For any $n \in \mathbb{Z}^d$,

$$Zf(x+n,\omega) = e^{2\pi i\omega\cdot n}\, Zf(x,\omega),$$
$$Zf(x,\omega+n) = Zf(x,\omega).$$

Hence, Zf is completely determined by its values on the cube $Q \times Q \subseteq \mathbb{R}^{2d}$.

ii. (Time-frequency shifts) For $(u,\nu) \in \mathbb{R}^{2d}$,

$$Z(T_u M_\nu f)(x,w) = e^{2\pi i\nu\cdot(x-u)} Zf(x-u,w-\nu).$$

iii. (Inversion formulas)

$$\int_Q Zf(x,w)dw = f(x),$$
$$\int_Q Zf(x,w)e^{-2\pi i x\cdot w}dx = \widehat{f}(w).$$

Theorem 3.1.3. *If $f \in W(\mathbb{R}^d)$, then $\int_Q \int_Q |Zf(x,w)|^2 dxdw = \|f\|_2^2$. Consequently, Zf extends to a unitary operator from $L^2(\mathbb{R}^d)$ onto $L^2(Q \times Q)$.*

Schwartz functions can also be characterized by the Zak transform. The Zak transform maps a function in $\mathcal{S}(\mathbb{R}^d)$ to an infinitely differentiable function on $\mathbb{R}^{2d}$.

Theorem 3.1.4. *If $f \in \mathcal{S}(\mathbb{R}^d)$, then $Zf \in C^\infty(\mathbb{R}^{2d})$. Conversely, if $F \in C^\infty(\mathbb{R}^{2d})$ is quasiperiodic, then $F = Zf$ for a (unique) $f \in \mathcal{S}(\mathbb{R}^d)$.*

The Zak transform has the following important and interesting property due to quasiperiodicity.

Lemma 3.1.5. *If Zf is continuous on $\mathbb{R}^{2d}$, then Zf has a zero in every unit square in $\mathbb{R}^{2d}$.*

Theorem 3.1.6. *If $f \in M^1(\mathbb{R}^d)$, then $f \in W_0(\mathbb{R}^d)$. Consequently, by Lemma 3.1.5, if $f \in M^1(\mathbb{R}^d)$ then Zf has a zero in every unit square in $\mathbb{R}^{2d}$.*

A useful property of Zak Transform is the following equality which describes the Fourier coefficients of $Zf \cdot \overline{Zg}$ as samples of the STFT of f with respect to the window g: Given f, $g \in L^2(\mathbb{R}^d)$,

$$\int_{[0,1)^d} \int_{[0,1)^d} Zf(x,\omega)\, \overline{Zg(x,\omega)}\, e^{-2\pi i n\cdot x}\, e^{-2\pi i m\cdot\omega}\, dx\, d\omega = V_g f(-m,n). \tag{2}$$

This fact allows a characterization of the modulation spaces via the Zak transform. The following result is due to Janssen [32], and we also mention that a different characterization of the modulation spaces was obtained earlier by Walnut [52].

Theorem 3.1.7. *Fix $1 \le p,\ q \le \infty$ and let $g^{(r)} \in M^1_v(\mathbb{R}^d)$ for $r = 1, \dots, N$ be such that the functions $Zg^{(r)}$ have no common zeros. If $f \in L^2(\mathbb{R}^d)$, then $f \in M^{p,q}_m(\mathbb{R}^d)$ if and only if each $Zf \cdot \overline{Zg^{(r)}}$ has a Fourier series*

$$\sum_{k,\ell} c^{(r)}_{k,\ell}\, e^{2\pi i k\cdot x + 2\pi i \ell\cdot\omega}$$

such that $(c_{\ell,-k})_{k,\ell\in\mathbb{Z}^d} \in \ell^{p,q}_m(\mathbb{Z}^{2d})$.

It is necessary to take $N \ge 2$ in Theorem 3.1.7 because the functions $Zg^{(r)}$ will be continuous and therefore have zeros.

Our second main theorem describes the mapping properties of the Zak transform acting on the modulation spaces.

Theorem 3.1.8. *Fix $1 \le p < \infty$. Assume that:*

(a) *v is a submultiplicative weight on $\mathbb{R}^{2d}$ such that $|v(z)| \lesssim (1+|z|)^{N_0}$ for some $N_0 \in \mathbb{N}$,*

(b) *m is a v-moderate weight on $\mathbb{R}^{2d}$,*

(c) *$\psi \in C_c^\infty(\mathbb{R}^{2d})$.*

Let m^* be any weight function satisfying $m^*(N,M) \le C_N\, m(M)$ for all M, $N \in \mathbb{Z}^{2d}$. Then for every $f \in L^2(\mathbb{R}^d) \cap M^p_m(\mathbb{R}^d)$ we have $\psi Z f \in M^p_{m^*}(\mathbb{R}^{2d})$.

Proof. First we note that $\psi Zf \in L^2(\mathbb{R}^{2d})$.

Let $g \in C_c^\infty(\mathbb{R}^d)$ be supported on $[0,1]^d$ and nonzero on the interior of this cube. Define

$$\xi = \underbrace{(1, \cdots, 1)}_{d \text{ times}} \in \mathbb{R}^d$$

and

$$g_j(x) = g\Big(x + \frac{j-1}{2d+1}\xi\Big), \qquad j = 1, \dots, 2d+1.$$

We have $g_j \in M^1_v(\mathbb{R}^d)$ for each j. Taking the support of g_j into consideration, if $k \in \mathbb{Z}^d$ then

$$Zg_j(x,\omega) = e^{2\pi i k\cdot\omega} g_j(x-k), \qquad x \in [0,1)^d + k - \frac{j-1}{2d+1}\xi.$$

The zero set of Zg_j is

$$\Big\{(x,\omega) \in \mathbb{R}^{2d} : x \in k - \Big(\frac{j-1}{2d+1}\Big)\xi + \partial[0,1)^d \cap [0,1)^d,\ k \in \mathbb{Z}^d,\ \omega \in \mathbb{R}^d\Big\},$$

where ∂A denotes the boundary of a set A.

Claim 1. The functions Zg_j have no common zeros.

To see this, assume that $Zg_j(x_0, \omega_0) = 0$ for each $j = 1, \dots, 2d+1$. Then for each j there exists a $k_j \in \mathbb{Z}^d$ such that

$$x_0 \in k_j - \Big(\frac{j-1}{2d+1}\Big)\xi + \partial[0,1)^d \cap [0,1)^d.$$

Let $\theta_j \in \partial[0,1)^d \cap [0,1)^d$ be such that $x_0 = k_j - (\frac{j-1}{2d+1})\xi + \theta_j$. For $j \neq j'$ we have

$$x_0 = k_j - \Big(\frac{j-1}{2d+1}\Big)\xi + \theta_j = k_{j'} - \Big(\frac{j'-1}{2d+1}\Big)\xi + \theta_{j'},$$

so $k_j - k_{j'} = (\frac{j-j'}{2d+1})\xi + \theta_{j'} - \theta_j$. Each of θ_j and $\theta_{j'}$ must have a zero component. If they both have a zero in the mth coordinate, then $k_j^{(m)} - k_{j'}^{(m)} = (\frac{j-j'}{2d+1})\xi^{(m)}$. This is a contradiction since $k_j^{(m)} - k_{j'}^{(m)}$ is integer. Hence no θ_j and $\theta_{j'}$ can have a zero in the same component. However, this is impossible as there are $2d+1$ vectors θ_j each with d components. This proves the claim.

Consequently, Theorem 3.1.7 implies that the Fourier coefficients of $Zf\,\overline{Zg_j}$ belong to $\ell^p_m(\mathbb{Z}^{2d})$ for each $j = 1, \ldots, 2d+1$.

Now let $\phi \in C_c^\infty(\mathbb{R}^{2d})$ be supported in the unit square $[0,1]^{2d}$ and strictly positive on its interior. For $j = 1, \ldots, 2d+1$ define

$$\phi_j(x,\omega) = \phi\Big((x,\omega) + \Big(\frac{j-1}{2d+1}\Big)(\xi,\xi)\Big) \qquad \text{and} \qquad G_j = \phi_j Z g_j.$$

Since each g_j is Schwartz-class, we have $Zg_j \in C^\infty(\mathbb{R}^{2d})$. Therefore, for any $N_1 \in \mathbb{N}$ and any weight v_2 satisfying $v_2(z) \lesssim (1+|z|)^{N_1}$ we have

$$G_j = \phi_j Z g_j \in C_c^\infty(\mathbb{R}^{2d}) \subseteq M^1_{v_2}(\mathbb{R}^{2d}).$$

For the remainder of this proof, let v_2 be any such weight.

Claim 2. The zero set of Zg_j is contained in the zero set of ϕ_j.

This follows from the fact that the zero set of Zg_j is

$$\bigcup_{k\in\mathbb{Z}^d} \Big(\Big(-\frac{j-1}{2d+1}\Big)\xi + k + \partial[0,1)^d \cap [0,1)^d\Big) \times \mathbb{R}^d$$

while the zero set of ϕ_j is

$$\mathbb{R}^{2d} \setminus \Big(-\frac{j-1}{2d+1}(\xi,\xi) + (0,1)^{2d}\Big).$$

Now,

$$ZG_j(x,\omega,p,s) = \sum_{(m,n)\in\mathbb{Z}^{2d}} e^{2\pi i (p,s)\cdot(m,n)}\, \phi_j(x-m,\omega-n)\, Zg_j(x-m,\omega-n).$$

Since ϕ_j is compactly supported, for $(x,\omega) \in (m,n) - \left(\frac{j-1}{2d+1}\right)(\xi,\xi) + [0,1)^{2d}$ we have that

$$ZG_j(x,\omega) = e^{2\pi i(p,s)\cdot(m,n)}\,\phi_j(x-m,\omega-n)\,Zg_j(x-m,\omega-n).$$

The zero set of ZG_j is

$$\bigcup_{(m,n)\in\mathbb{Z}^{2d}} \left(\left(-\frac{j-1}{2d+1}\right)(\xi,\xi) + (m,n) + \partial[0,1)^{2d} \cap [0,1)^{2d}\right) \times \mathbb{R}^{2d}.$$

Claim 3. The functions ZG_j have no common zeros.

To see this, suppose $(x,\omega,p,s) \in \mathbb{R}^{2d} \times \mathbb{R}^{2d}$ is a common zero point for the functions ZG_j. Then for each $j = 1,\dots,2d+1$ there exists $(m_j,n_j) \in \mathbb{Z}^{2d}$ and $\theta_j \in (-\frac{j-1}{2d+1})(\xi,\xi)+(m,n)+\partial[0,1)^{2d}\cap[0,1)^{2d}$ such that $(x,\omega) = (m_j,n_j)-(\frac{j-1}{2d+1})(\xi,\xi)+\theta_j$. Each θ_j must have a component that is zero.

If $j \neq j'$ and the sth component of θ_j and $\theta_{j'}$ is both zero then, as in the proof of Claim 1, we obtain that $(m_j,n_j)^{(s)} - (\frac{j-1}{2d+1})(\xi,\xi)^{(s)} = (m_{j'},n_{j'})^{(s)} - (\frac{j'-1}{2d+1})(\xi,\xi)^{(s)}$. This implies $\frac{j-j'}{2d+1}$ is an integer, which is a contradiction. Hence no θ_j and $\theta_{j'}$ can have a zero in the same component, which is impossible since the number of j's is $2d+1$ while the dimension of θ_j is $2d$.

Claim 4. There exists a v_2-moderate weight $m^* : \mathbb{R}^{2d} \times \mathbb{R}^{2d} \to (0,\infty)$ such that the sequence $c_{N,-M}$ belongs to $\ell^p_m(\mathbb{Z}^{2d})$, where $c_{M,N}$ are the Fourier coefficients of $Z(\psi Zf) \cdot \overline{ZG_j}$

To see this, fix j and let $K \in \mathbb{Z}^{2d}$ and $\alpha \in \mathbb{N}$ be such that $\mathrm{supp}(\psi) \subseteq K + [0,\alpha)^{2d}$. Write $M = (m_1,m_2)$, $N = (n_1,n_2) \in \mathbb{Z}^{2d}$ and $X = (x_1,x_2)$, $Y = (y_1,y_2) \in \mathbb{R}^{2d}$, etc. Recalling that $G_j = \phi_j\, Zg_j$, we use equation (2) to compute that the Fourier coefficient $c_{N,-M}$ of $Z(\psi Zf) \cdot \overline{ZG_j}$ is

$$\begin{aligned}
c_{N,-M} &= V_{G_j}(\psi\, Zf)(M,N) \\
&= \int_{\mathbb{R}^{2d}} \psi(X)\, Zf(X)\, \overline{G_j(X-M)}\, e^{-2\pi i N\cdot X}\, dX \\
&= \int_{K+[0,\alpha)^{2d}} \psi(X)\, Zf(X)\, \phi_j(X-M)\, \overline{Zg_j(X-M)}\, e^{-2\pi i N\cdot X}\, dX
\end{aligned}$$

$$
\begin{aligned}
&= \alpha^{2d} \int_{[0,1)^{2d}} \psi(\alpha Y + K)\, Zf(\alpha Y + K)\, \phi_j(\alpha Y + K - M) \\
&\qquad \times \overline{Zg_j(\alpha Y + K - M)}\, e^{-2\pi i N\cdot(\alpha Y + K)}\, dY \\
&= \alpha^{2d} \int_{[0,1)^{2d}} \psi(\alpha Y + K)\, \phi_j(\alpha Y + K - M)\, Zf(\alpha y_1 + k_1, \alpha y_2 + k_2) \\
&\qquad \times \overline{Zg_j(\alpha y_1 + k_1 - m_1, \alpha y_2 + k_2 - m_2)}\, e^{-2\pi i N\cdot(\alpha Y + K)}\, dY \\
&= \alpha^{2d} \int_{[0,1)^{2d}} \psi(\alpha Y + K)\, \phi_j(\alpha Y + K - M)\, Zf(\alpha y_1 + k_1, \alpha y_2) \\
&\qquad \times \overline{Zg_j(\alpha y_1 + k_1 - m_1, \alpha y_2)}\, e^{-2\pi i N\cdot(\alpha Y + K)}\, dY \\
&= \alpha^{2d} \int_{[0,1)^{2d}} \psi(\alpha Y + K)\, \phi_j(\alpha Y + K - M)\, e^{2\pi i \alpha y_2 \cdot k_1}\, Zf(\alpha y_1, \alpha y_2) \\
&\qquad \times e^{-2\pi i \alpha y_2\cdot(k_1 - m_1)}\, \overline{Zg_j(\alpha y_1, \alpha y_2)}\, e^{-2\pi i (n_1,n_2)\cdot(\alpha y_1 + k_1, \alpha y_2 + k_2)}\, dY \\
&= \alpha^{2d} \int_{[0,1)^{2d}} \psi(\alpha Y + K)\, \phi_j(\alpha Y + K - M)\, Zf(\alpha Y)\, \overline{Zg_j(\alpha Y)} \\
&\qquad \times e^{2\pi i(\alpha y_2\cdot k_1 - \alpha y_2 \cdot k_1 + \alpha y_2\cdot m_1 - n_1\cdot \alpha y_1 - n_2 \cdot \alpha y_2)}\, dY \\
&= \alpha^{2d} \int_{[0,1)^{2d}} \psi(\alpha X + K)\, \phi_j(\alpha X + K - M)\, Zf(\alpha X)\, \overline{Zg_j(\alpha X)} \\
&\qquad \times e^{-2\pi i (n_1, n_2 - m_1)\cdot(\alpha x_1 + k_1, \alpha x_2 + k_2)}\, dX,
\end{aligned}
$$

where we have applied the quasiperiodicity of the Zak transform and the fact that M, N, K belong to $\mathbb{Z}^{2d}$.

Note that there are only a finite number of M such that $\phi_j(\alpha X - M + K)$ is not identically zero on $[0,1)^{2d}$, and hence a finite number of M such that $c_{N,-M}$ is nonzero. Let B be the set of those $M \in \mathbb{Z}^{2d}$ such that $c_{N,-M} \neq 0$. Let us fix an $M \in B$ and define the following functions:

$$
\begin{aligned}
\Phi^{(M)}(X) &= \psi(\alpha X + K)\, \phi_j(\alpha X + K - M)\, Zf(\alpha X)\, \overline{Zg_j(\alpha X)}\, e^{-2\pi i m_1\cdot x_2}, \\
U^{(M)}(X) &= \psi(\alpha X + K)\, \phi_j(\alpha X + K - M)\, e^{-2\pi i m_1 \cdot x_2}, \\
W(X) &= Zf(\alpha X)\, \overline{Zg_j(\alpha X)}.
\end{aligned}
$$

We note some facts about these functions.

i. $\Phi^{(M)} = U^{(M)} W$.

ii. We have

$$c_{N,-M} = \alpha^{2d} \int_{[0,1)^{2d}} \Phi^{(M)}(X)\, e^{-2\pi i \alpha N\cdot X}\, dX$$
$$= \alpha^{2d}\, \widehat{\Phi^{(M)}}(\alpha N) = \alpha^{2d}\, (\widehat{U^{(M)}} * \widehat{W})(\alpha N).$$

iii. $\widehat{W}(\alpha N) = \int_{[0,1)^{2d}} Zf(\alpha X)\, \overline{Zg_j(\alpha X)}\, e^{-2\pi i \alpha N\cdot X}\, dX = (Zf\, \overline{Zg_j})^{\wedge}(N).$

iv. $U^{(M)}$ is a compactly supported infinitely differentiable function whose support lies inside the unit cube $[0,1]^{2d}$. Consequently, $\big\|\big\{\widehat{U^M}(\alpha N)\big\}\big\|_{\ell^1_v(\mathbb{Z}^{2d})} < \infty$ for any polynomial weight function v defined on $\mathbb{R}^{2d}$.

v. By Theorem 3.1.7, the hypothesis $f \in M^p_m(\mathbb{R}^{2d})$ implies that the Fourier coefficients of $Zf\, \overline{Zg_j}$ belong to $\ell^p_m(\mathbb{Z}^{2d})$, and therefore $\big\{\widehat{W}(\alpha N)\big\} \in \ell^p_m(\mathbb{Z}^{2d})$.

Fix now any weight function $m^* : \mathbb{R}^{2d} \times \mathbb{R}^{2d} \to \mathbb{R}$ such that:

- there exists a $C > 0$ such that $m^*(M,N) \le C m(N)$ for all $N \in \mathbb{Z}^{2d}$ and $M \in B$, and
- m^* is v_2-moderate with respect to some submultiplicative function v_2.

For example, we can simply take $m^*(M,N) = m(N)$. Define $m_M(N) = m^*(M,N)$.

Using Proposition 2.3.7, we compute that

$$\begin{aligned}
\big\|\big\{c_{N,-M}\big\}_{N\in\mathbb{Z}^{2d}}\big\|_{\ell^p_{m_M}(\mathbb{Z}^{2d})} &= \Bigg(\sum_{N\in\mathbb{Z}^{2d}} |c_{N,-M}|^p\, m^*(M,N)^p\Bigg)^{1/p} \\
&\le C\Bigg(\sum_{N\in\mathbb{Z}^{2d}} |c_{N,-M}|^p\, m(N)^p\Bigg)^{1/p} \\
&= C\alpha^{2d}\, \big\|\big\{(\widehat{U^{(M)}} * \widehat{W})(\alpha N)\big\}\big\|_{\ell^p_m(\mathbb{Z}^{2d})} \\
&\le CC'\alpha^{2d} \big\|\widehat{U}(\alpha N)\big\|_{\ell^1_v(\mathbb{Z}^{2d})}\, \big\|\widehat{W}(\alpha N)\big\|_{l^p_m(\mathbb{Z}^{2d})} \\
&< \infty.
\end{aligned}$$

From above, we obtain

$$\begin{aligned}\left\|\left\{c_{N,-M}\right\}_{M,N\in\mathbb{Z}^{2d}}\right\|_{\ell^p_{m^*}(\mathbb{Z}^{2d}\times\mathbb{Z}^{2d})} &= \left(\sum_{N\in\mathbb{Z}^{2d}}\sum_{M\in B}|c_{N,-M}|^p\, m^*(M,N)^p\right)^{1/p} \\ &= \left(\sum_{M\in B}\sum_{N\in\mathbb{Z}^{2d}}|c_{N,-M}|^p\, m^*(M,N)^p\right)^{1/p} \\ &\le \left(\sum_{M\in B}\left\|\left\{c_{N,-M}\right\}_{N\in\mathbb{Z}^{2d}}\right\|^p_{l^p_{m_M}(\mathbb{Z}^{2d})}\right)^{1/p} < \infty,\end{aligned}$$

since B is a finite set. □

As an application of this theorem, we recover the following result, which is [25, Lemma 2.3].

Corollary 3.1.9. *Assume $f \in H^{s_1}(\mathbb{R}^d)$ and $\widehat{f} \in H^{s_2}(\mathbb{R}^d)$ where s_1, $s_2 > 0$. Then for any smooth, compactly supported function $\psi \in C_c^\infty(\mathbb{R}^{2d})$ we have $\psi Zf \in \mathcal{F}L^2_m(\mathbb{R}^{2d})$ where $m(x,y) = (1+|x|^{2s_1}+|y|^{2s_2})^{1/2}$.*

Proof. Note that $m(x,y)$ satisfies

$$m(x,y) \le \left(1+m_{s_1}(x)^2+m_{s_2}(y)^2\right)^{1/2}.$$

Hence $f \in M^2_m(\mathbb{R}^d)$ by Proposition 2.3.12. Now let

$$v(x,y) = \left(1+(|x|^2+|y|^2)^{1/2}\right)^{\max\{s_1,s_2\}}.$$

Then m is a v-moderate weight. By Theorem 3.1.8, $\psi Zf \in M^2_{m^*}(\mathbb{R}^{2d})$. Taking

$$m^*(x,y,u,v) = m(u,v),$$

by Proposition 2.3.11(b) we obtain $\psi Zf \in \mathcal{F}L^2_m(\mathbb{R}^{2d})$. □

Another implication of Theorem 3.1.8 is that the Zak transform embeds certain modulation spaces into a local VMO space.

Corollary 3.1.10. *Let v be a submultiplicative weight satisfying $|v(z)| \le (1+|z|)^{N_0}$ for some $N_0 \in \mathbb{N}$, and let m be a v-moderate weight such that $m(x) \lesssim \left(1+|x|^2\right)^{1/2}$. If $f \in L^2(\mathbb{R}) \cap M^p_m(\mathbb{R})$ where $1 \le p \le 2$, then $Zf \in \mathrm{VMO}(K)$ for any compact set $K \subset \mathbb{R}^2$.*

3.2 *Application: A Balian-Low Type Theorem for Gabor Frames via Modulation Spaces*

In applications, rather than the continuous representation of signals, sampling of signals is used. One of the tools related to sampling of signals in signal processing is frames.

Definition 3.2.1. A sequence $\{f_j : j \in J\}$ in a (separable) Hilbert space $\mathcal{H}$ is called a frame if there exist positive constants $A, B > 0$ such that for all $g \in \mathcal{H}$,

$$A\|g\|^2 \leq \sum_{j\in J} |\langle g, f_j\rangle|^2 \leq B\|g\|^2.$$

Any two constants satisfying the inequality above are called frame bounds and $\langle g, f_j\rangle$ are called frame coefficients. If we can take $A = B$, then $\{f_j : j \in J\}$ is called a tight frame.

Frames can be regarded as a generalization of orthonormal bases, obviously an orthonormal basis is a frame with $A = B = 1$. The main difference between orthonormal bases and frames is that frames can have redundancy whereas orthonormal bases are not redundant. One can easily form a frame by taking the union of two orthonormal bases.

It is possible to recover a signal from the frame coefficients of a function.

Proposition 3.2.2. *If $\{f_j\}$ is a frame with the frame bounds A and B, then there exists a frame $\tilde{f}_j$ with frame bounds A^{-1}, B^{-1} such that for any $g \in \mathcal{H}$,*

$$g = \sum_j \langle g, f_j\rangle \widetilde{f_j} = \sum_j \langle g, \tilde{f}_j\rangle f_j.$$

In particular, if $\{f_j\}$ is a tight frame then $\widetilde{f_j} = \frac{1}{A} f_j$.

In time-frequency analysis of a function, the samples of the short time Fourier transform of a function are collected. The samples of the short time Fourier transform of a function f for a given window g satisfy

$$V_g f(\alpha k, \beta n) = \langle f, M_{\beta n} T_{\alpha k} g \rangle = e^{-2\pi i \alpha k \cdot n} \langle f, T_{\alpha k} M_{\beta n} g \rangle.$$

It is reasonable to form a frame by using the functions $T_{\alpha k} M_{\beta n} g$. Indeed, the time frequency shifts of a window g is called a Gabor system. Given $g \in L^2(\mathbb{R})$ and constants α, $\beta > 0$, the associated Gabor system is

$$\mathcal{G}(g, \alpha, \beta) = \left\{ e^{2\pi i m \beta x} \, g(x - n\alpha) \right\}_{m,n \in \mathbb{Z}}.$$

It is natural to consider under what conditions $\mathcal{G}(g, \alpha, \beta)$ will generates a frame. The conditions on α and β determine the density of sampling of the short time Fourier transform of f under the window g. The following theorem gives an idea about the density of sampling.

Theorem 3.2.3. *Let $g \in L^2(\mathbb{R}^d)$. If $\mathcal{G}(g, \alpha, \beta)$ is a frame for $L^2(\mathbb{R}^d)$, then $\alpha\beta \leq 1$.*

In signal analysis it has become customary to call $\alpha\beta < 1$ oversampling, $\alpha\beta = 1$ critical sampling, and $\alpha\beta > 1$ undersampling.

The classical Balian–Low Theorem addresses the question of what conditions are required for g in order $\mathcal{G}(g, \alpha, \beta)$ to be a frame for the case $\alpha\beta = 1$ (the "critical density"), which by a change of variables can be further reduced to the case $\alpha = \beta = 1$. Other Balian–Low-type theorems can be found in [28], [11], [24], [12], [13], [25].

Theorem 3.2.4. *[Classical BLT] Given $g \in L^2(\mathbb{R})$, if $g \in H^1(\mathbb{R})$ and $\widehat{g} \in H^1(\mathbb{R})$ then $\mathcal{G}(g, 1, 1)$ is not a frame for $L^2(\mathbb{R})$.*

The approach to the proof is based on the Zak transform, which completely determines whether the Gabor system associated with a function g is a frame at the critical density.

Proposition 3.2.5. *Fix $g \in L^2(\mathbb{R})$. Then $\mathcal{G}(g, 1, 1)$ is a frame for $L^2(\mathbb{R})$ if and only if $0 < A^{1/2} \leq |Zf| \leq B^{1/2} < \infty$ a.e.*

If a function belongs to the modulation space $M^1(\mathbb{R})$, which is a subspace of $L^2(\mathbb{R})$, then the associated Gabor system cannot be a frame for $L^2(\mathbb{R})$. The reason for this is given in the following proposition, whose proof can be found in [27].

Proposition 3.2.6. *(a) If $f \in M^1(\mathbb{R})$ then Zf is continuous.*

(b) If $f \in L^2(\mathbb{R})$ and Zf is continuous then Zf must have a zero.

The following result is due to Gautam [25]. The original proof of the classical Balian-Low Theorem did not mention VMO($\mathbb{R}$) space. To show how $VMO(\mathbb{R})$ forces a quasiperiodic function to have a zero, we will give the proof of the following theorem. Without assuming the continuity of Zf, when applied to $F = Zf$ it gives a sufficient condition for Zf to have a zero inside the unit square.

Theorem 3.2.7. *Suppose $F \in \mathrm{VMO}(K) \cap L^\infty(\mathbb{R}^2)$ for a compact set K that contains the unit square in $\mathbb{R}^2$. If*

$$F(x+1, y) = e^{2\pi i y} F(x, y) \ \textit{a.e.},$$

$$F(x, y+1) = F(x, y) \ \textit{a.e.},$$

then $\operatorname{ess\,inf} |F| = 0$.

Proof. Let K contain the unit square. Since $|F|$ is periodic, it is enough to analyze F on the unit square. Assume that $\operatorname{ess\,inf} |F| \neq 0$. By scaling, we may assume also $|F| \leq 1$. Now there exists a $d > 0$ such that

$$d \leq |F| \leq 1$$

almost everywhere. Let $Q_\epsilon(x, y)$ denote the cube of side length ϵ centered at $(x, y) \in K$, and define

$$F_\epsilon(x, y) = \fint_{Q_\epsilon(x,y)} F(t, v)dtdv,$$

the average of F over $Q_\epsilon(x, y)$. F_ϵ is continuous and satisfies the modified quasi-periodicity relations

$$F_\epsilon(x+1,y) = e^{2\pi i y}\, F_\epsilon(x,y) + \Phi_\epsilon(x,y) \text{ a.e.},$$
$$F_\epsilon(x,y+1) = F_\epsilon(x,y) \text{ a.e.},$$

where the error term $\Phi_\epsilon(x,y)$ satisfies

$$|\Phi_\epsilon(x,y)| \lesssim \epsilon.$$

Moreover, for ϵ sufficiently small, F_ϵ is also bounded from below, as we now show. Since $F \in \mathrm{VMO}(K)$, we may choose $\epsilon_0$so that $⨍_{Q_\epsilon(x,y)} |F - F_\epsilon(x,y)| \leq \frac{d}{2}$ for all $(x,y) \in K$ and $\epsilon < \epsilon_0$. Now by triangle inequality, we have

$$\begin{aligned} |F_\epsilon(x,y)| &= ⨍_{Q_\epsilon(x,y)} |F(x,y)| \\ &\geq ⨍_{Q_\epsilon(x,y)} |F| - ⨍_{Q_\epsilon(x,y)} |F - F_\epsilon(x,y)| \\ &\geq \frac{d}{2}, \end{aligned}$$

since we assume $|F| \geq d$ almost everywhere. Thus $\frac{d}{2} \leq |F_\epsilon| \leq 1$ almost everywhere for $\epsilon < \epsilon_0$.

Now since F_ϵ is continuous and $\frac{d}{2} \leq |F_\epsilon| \leq 1$, we can define a continuous branch γ_ϵ of $\log F_\epsilon$. From the modified quasi-periodicity conditions we have

$$\gamma_\epsilon(x+1,y) = \gamma_\epsilon(x,y) + 2\pi i j + 2\pi i y + \Psi_\epsilon(x,y),$$
$$\gamma_\epsilon(x,y+1) = \gamma_\epsilon(x,y) + 2\pi i k$$

for all x, y in some simply connected neighborhood U of Q_0. Here $j, k \in \mathbb{Z}$ are constant on U by continuity of γ_ϵ, and

$$|\Psi_\epsilon| \leq -\log\left(1 - \frac{|\Phi_\epsilon|}{|F_\epsilon|}\right) \lesssim \frac{|\Phi_\epsilon|}{|F_\epsilon|}$$

provided that $\frac{|\Phi_\epsilon|}{|F_\epsilon|}$ is sufficiently small. This can be arranged by taking ϵ sufficiently small, since $|\Phi_\epsilon| \lesssim \epsilon$ and $|F_\epsilon| \geq \frac{d}{2}$; thus for ϵ small we have $|\Psi_\epsilon| < 1$ on U. To obtain a contradiction, we simply compute

$$\begin{aligned}0 &= \gamma_\epsilon(1,0) - \gamma_\epsilon(0,0) + \gamma_\epsilon(1,1) - \gamma_\epsilon(1,0)\\ &\quad + \gamma_\epsilon(0,1) - \gamma_\epsilon(1,1) + \gamma_\epsilon(0,0) - \gamma_\epsilon(0,1)\\ &= \Psi_\epsilon(0,0) - \Psi_\epsilon(0,1) - 2\pi i\\ &\neq 0,\end{aligned}$$

since $|\Psi_\epsilon| < 1$. Thus $|F| \geq d$ a.e. is impossible and hence $\operatorname{ess\,inf} |F| = 0$ as desired. $\square$

Now we give a Balian–Low-type theorem in terms of modulation spaces.

Theorem 3.2.8. *Let v be a submultiplicative weight with $|v(z)| \lesssim (1+|z|)^{N_0}$ for some $N_0 \in \mathbb{N}$, and let m be a v-moderate weight such that $m(x) \gtrsim (1+|x|^2)^{1/2}$. If $f \in L^2(\mathbb{R}) \cap M^p_m(\mathbb{R})$ where $1 \leq p \leq 2$, then $\mathcal{G}(f,1,1)$ is not a frame for $L^2(\mathbb{R})$.*

Proof. If $\mathcal{G}(f,1,1)$ is a frame then $Zf \in L^\infty(\mathbb{R}^2)$. Combining this with Corollary 3.1.10 and Theorem 3.2.7 implies that $\operatorname{ess\,inf} |Zf| = 0$, which is a contradiction. $\square$

Letting $p = 2$ and $m(x,\omega) = (1+|x|^{2s_1} + |\omega|^{2s_2})^{1/2}$ where s_1, $s_2 \geq 1$ recovers the usual Balian–Low Theorem.

Corollary 3.2.9. *If $f \in L^2(\mathbb{R})$ satisfies $f \in H^{s_1}(\mathbb{R})$ and $\widehat{f} \in H^{s_2}(\mathbb{R})$ where s_1, $s_2 \geq 1$, then $\mathcal{G}(f,1,1)$ is not a frame for $L^2(\mathbb{R})$.*

PART II

Minimizing IPH Functions over the Unit Simplex

CHAPTER IV

MINIMIZATION OF IPH FUNCTIONS VIA MIN-TYPE FUNCTIONS: CUTTING ANGLE METHOD

4.1 Preliminaries and Notation

Consider an $n-$dimensional linear space $\mathbb{R}^n$. We shall use the following notations:

$I = \{1, .., n\}$;

x_i is the ith coordinate of a vector $x \in \mathbb{R}^n$;

$[l, x] = \sum\limits_{i \in I} l_i x_i$ is the inner product of vectors $l \in \mathbb{R}^n$ and $x \in \mathbb{R}^n$;

$e_m \in \mathbb{R}^n$ is the unit vector whose mth coordinate is 1;

For $x, y \in \mathbb{R}^n$, $x \geq y \Leftrightarrow x_i \geq y_i$ for all $i \in I$;

For $x, y \in \mathbb{R}^n$, $x \gg y \Leftrightarrow x_i > y_i$ for all $i \in I$;

x and y are said to be comparable if $x \geq y$ or $x \leq y$ for $x, y \in \mathbb{R}^n$, otherwise they are called incomparable;

$\mathbb{R}^n_+ = \{x = (x_1, ..., x_n) \in \mathbb{R}^n : x_i \geq 0 \text{ for all } i \in I\}$ (nonnegative orthant);

$S = \left\{ x \in \mathbb{R}^n_+ : \sum\limits_{i \in I} x_i = 1 \right\}$ (unit simplex);

$riS = \left\{ x \in \mathbb{R}^n_+ : \sum\limits_{i \in I} x_i = 1,\ x_i \neq 0 \text{ for all } i \in I \right\}$ (relative interior of S);

$C\binom{w}{k}$ denotes the number of k-combinations from w elements.

A feasible direction $h \in \mathbb{R}^n$ for a point $x \in S$ is the nonzero vector such that $x + h \in S$. Note that a feasible direction $h = (h_1, \cdots, h_n)$ for a point in S satisfies $\sum\limits_{i}^{n} h_i = 0$.

4.2 Min-type functions and IPH functions

Definition 4.2.1. A function f defined on $\mathbb{R}^n_+$ is called increasing positively homogeneous of degree one (for short IPH), if

a) for $x, y \in \mathbb{R}^n_+$, $x \geq y$ implies $f(x) \geq f(y)$;

b) $f(\lambda x) = \lambda f(x)$ for all $x \in \mathbb{R}^n_+$ and $\lambda > 0$.

Example 4.2.2. Let D be either $\mathbb{R}^d_+$. The following functions defined on D are IPH:

1. $a(x) = \sum_{i \in I} a_i x_i$ with $a_i \geq 0$;

2. $p_k(x) = \left(\sum_{i \in I} x_i^k\right)^{\frac{1}{k}}$ with $k > 0$;

3. $f(x) = \sqrt{\langle Ax, x \rangle}$, where A is a matrix with nonnegative entries;

4. $f(x) = \prod_{j \in J} x_j^{t_j}$, where $J \subset I$, $t_j > 0, \sum_{j \in J} t_j = 1$.

Proposition 4.2.3. *The following statements hold for IPH functions.*

i. The sum of two IPH functions is also an IPH function.

ii. If f is IPH, then the function γf is IPH for all $\gamma > 0$.

iii. T is an arbitrary index set and $\{f_t\}_{t \in T}$ is a family of IPH functions, then the functions $h(x) = \inf_{t \in T} f_t(x)$ and $g(x) = \sup_{t \in T} f_t(x)$ are IPH.

iv. The pointwise limit of a sequence (and more general, of a directed set) of IPH functions is IPH.

Example 4.2.4. The following functions are IPH by the previous proposition:

1. $f(x) = \max_{k \in K} \min_{j \in J} \sum_{i \in I} a_i^{jk} x_i$ where $a_i^{jk} \geq 0, k \in K, j \in J, i \in I$ where J, K are finite sets of indices,

2. $f(x) = \max_{k \in K} \min_{j \in J_k} \sum_{i \in I} a_i^j x_i$, where $a_i^j \geq 0, j \in J_k, k \in K$, K and J_k are finite sets of finite sets of indices.

We note (see [10]) that an arbitrary piecewise linear function f generated by a collection of linear functions $a^1, \cdots, a^m$ can be represented as in statement 2; hence an arbitrary piecewise linear function defined on D and generated by nonnegative vectors is IPH.

Example 4.2.5. Let f be a Lipschitz function defined on the unit simplex S. Let $\min_{x \in S} f(x) \geq c$ for some $c > 0$, and let L be a Lipschitz constant of the function f in ℓ_1 norm, i.e.,

$$|f(x) - f(y)| \leq L \textstyle\sum_{i \in I} |x_i - y_i| \text{ for all } x, y \in S.$$

Assume that $\frac{2L}{c} \leq 1$. Then the function g defined on $\mathbb{R}^n_+$ by

$$g(x) = \left(\sum_{i \in I} x_i\right) f\left(\frac{x}{\sum_{i \in I} x_i}\right) \tag{3}$$

is IPH. This follows from the Theorem 1.3.6 with $p = 1$. Note that $g(x) = f(x)$ for $x \in S$.

Remark 4.2.6. In applications, one should find or estimate L and c in order to determine whether a function is IPH or not on S. This problem can be solved practically in minimization problems in the following way. Consider an arbitrary Lipschitz function ϕ defined on S. Consider the function $f(x) = \phi(x) + M$ where M is a very large number. Since the Lipschitz constant of the function f coincides with the Lipschitz constant of ϕ, we can obtain the inequality $\frac{2L}{c} \leq 1$ by choosing a very large M. Using (3) we can construct the IPH function $g(x)$ which coincides with $\phi(x) + M$ for $x \in S$. Note that the functions ϕ, f and g have the same global minimizers on S. Hence global minimization of a Lipschitz function over S can be accomplished by global minimization of IPH functions.

Definition 4.2.7. Consider the function $l : \mathbb{R}^n_+ \to \mathbb{R}_+ \cup \{+\infty\}$ as follows:

$$l(x) = \min_{i \in I} \frac{x_i}{l_i}, \;\; l \in \mathbb{R}^n_+.$$

We call this function a min-type function generated by the vector l. We shall denote the function by the same symbol l.

Let us give some examples of min-type functions.

Example 4.2.8. $l^{(1)} = (1,2,3)$ then $l^{(1)}(x,y,z) = \min\{\frac{x}{1}, \frac{y}{2}, \frac{z}{3}\}$.

$l^{(2)} = (0,2,1)$ then $l^{(2)}(x,y,z) = \min\{\infty, \frac{y}{2}, z\} = \min\{\frac{y}{2}, z\}$.

$l^{(3)} = (0,1,0)$ then $l^{(3)}(x,y,z) = \min\{\infty, y, \infty\} = y$.

A min-type function is clearly IPH.

The following theorem establishes the relation between IPH functions and min-type functions.

Theorem 4.2.9. *[42]*

1. *A finite function f defined on $\mathbb{R}^n_+$ is IPH if and only if f is abstract convex with respect to min-type functions, i.e.,*

$$f(x) = \sup\{l(x) : l \in \mathbb{R}^n_+, l(x) \leq f(x)\}.$$

2. *Let $x^0 \in \mathbb{R}^n_+$ be a vector such that $f(x^0) > 0$ and $l = \frac{x^0}{f(x^0)}$. Then $l(x) \leq f(x)$ for all $x \in \mathbb{R}^n_+$ and $l(x^0) = f(x^0)$.*

4.3 Statement of the Main Problem and the Cutting Angle Method

We consider the following problem:

Let f be an IPH function and $f(x) > 0$ over S.

$$\begin{aligned} &\min f(x) \\ &x \in S \end{aligned} \tag{4}$$

For this problem, by using the condition that f is abstract convex with respect to the set of min-type functions, cutting angle method has been introduced in [4]. The cutting angle method is a generalization of cutting plane method from convex programming.

This method is as follows.

Algorithm 4.3.1. ***Step 0:*** *Take points* $x^m = e_m$, $m = 1, \cdots, n$ *and* $j = n+q$. *Choose arbitrary points* $x^{n+i} \in S$ *such that* $x_r^{n+i} \gg (0, \cdots, 0)$, $i = 1, \cdots, q$. *Let* $l_i^k = x_i^k / f(x^k)$, $k = 1, \cdots, j$. *Define the function* h_j *by*

$$h_j(x) = \max_{k=1,\cdots,j} \min_{i \in I} \frac{x_i^k}{l_i^k}.$$

Step 1: *Solve the problem* $\min_{x \in S} h_j(x)$.

Step 2: *Let* y^* *be a solution of the problem in Step 1. Set* $j = j+1$, $x^j = y^*$. *Find* $l_i^j = x_i^j / f(x^j)$ *and set*

$$h_j(x) = \max(h_{j-1}(x), \min_i \frac{x_i^j}{l_i^j}) = \max_{k=1,\cdots,j} \min_i \frac{x_i^k}{l_i^k}.$$

Then go to Step 1.

The algorithm provides lower and upper estimates of the global minimum f_* for Problem (4).

Let $\lambda_j = \min_{x \in S} h_j(x)$. It follows from Theorem 4.2.9 that

$$\min\nolimits_{i \in I} l_i^k x_i \leq f(x) \text{ for all } x \in S, k = 1, \cdots, j.$$

Hence $h_j(x) \leq f(x)$ for all $x \in S$ and $\lambda_j \leq \min_{x \in S} f(x)$. Thus λ_j is a lower estimate of the global minimum f_*. Consider the number $f(x^j) = \mu_j$. Clearly $\mu_j \geq f_*$. It can be shown [3] that λ_j is an increasing sequence and $\mu_j - \lambda_j \to 0$ as $j \to \infty$. This allows us to obtain an approximate solution with an arbitrary given tolerance.

Remark 4.3.2. Let f_* be the value of the global minimum of a function f over the simplex S. The precision δ_r of the current point of x^j is defined as follows:

$$\delta_r(x^j) = \min \left(f(x^j) - f_*, \frac{f(x^j) - f_*}{f_*} \right).$$

Generally, f_* is unknown, so we shall consider the following number:

$$\delta(x^j) = \min \left(f(x^j) - \lambda_j, \frac{f(x^j) - \lambda_j}{\lambda_j} \right)$$

as an estimate of the precision. Note that $\delta(x^j) > \delta_r(x^j)$. Numerical experiments show that $\delta_r(x^j)$ is substantially less than $\delta(x^j)$ in many instances. Thus very often we have a substantially more precise solution indicated by the estimate δ of the precision.

The algorithm is stopped at the desired precision.

Step 1, to find a global minimizer of the function h_j, is the most difficult and important step of the cutting angle method. Different algorithms have been developed to solve this problem (see, for example, [1], [5],[6],[38]). The next chapter is mostly devoted to the solution of this subproblem. We give necessary and sufficient conditions on the solution of this problem.

CHAPTER V

AN ALGORITHM FOR THE SUBPROBLEM IN CUTTING ANGLE METHOD

5.1 The Subproblem and an algorithm for solving it

As was mentioned in the previous section, the subproblem in the cutting angle method is expressed as follows:

$$\begin{aligned} &P(x) \to \min \\ &x \in S \end{aligned} \tag{5}$$

where

$$P(x) = \max_{1 \le j \le k} \{l^{(j)}(x)\} \text{ and } l^{(j)}(x) = \min_{i \in I} \{\frac{x_i}{l_i^{(j)}}\} j = 1, \cdots, k.$$

Before giving the algorithm for this subproblem, we make some observations. The following proposition tells us where to look for minimum points of $P(x)$ over the unit simplex. We refer [6] for the proof.

Proposition 5.1.1. *Let $m > n$ and $l^k = c_k e_k$ where c_k are positive numbers for $k \in I$ and $l^k \gg (0, \cdots, 0)$ for $n + 1 \le k \le m$. Then the function $P(x) = \max_{1 \le i \le m} l^i(x)$ has a local minimizer over S and each local minimizer of $P(x)$ lies in riS.*

This proposition tells us to look for the global minimizer in the interior of the unit simplex if $P(x)$ has at least n min-type functions whose vectors are multiples of unit vectors. Since Algorithm 4.3.1 begins with unit vectors and $f(x) > 0$ over the unit simplex, we have this type of function at step 2 in every iteration of Algorithm 4.3.1. Knowing this, we will focus on finding the local minimizers of this function in riS, because once the local minimizers are located, then the global minimum is one of them.

For the sake of simplicity we start with the problem in $\mathbb{R}^2_+$. We give the following easy and important result about min-type functions.

Lemma 5.1.2. *Let $l^{(1)}$ and $l^{(2)}$ be comparable vectors in $\mathbb{R}^n_+$ and let*

$$P(x) = \max\left\{l^{(1)}(x), l^{(2)}(x)\right\}.$$

Then

$$P(x) = \begin{cases} l^{(1)}(x), & \text{if } l^{(1)} \leq l^{(2)}, \\ l^{(2)}(x), & \text{if } l^{(1)} \geq l^{(2)}. \end{cases}$$

Lemma 5.1.2 gives a very important property about min-type functions. It tells us that if the vectors of two min-type functions are comparable then their maximum will be automatically the one which has smaller vector. This helps writing $P(x)$ as the maximum of incomparable min-type functions. To do so, one can find all comparable pairs of min-type functions in the expression of $P(x)$ and delete the bigger ones in light of Lemma 5.1.2.

Now we give the following definition, which will be quite useful in stating our theorems and reducing the work in the proofs.

Definition 5.1.3. If no pair of $l^{(i)} \in \mathbb{R}^n_+$ in the function $P(x) = \max\limits_{1\leq i\leq k}\left\{l^{(i)}(x)\right\}$ are comparable, then $P(x)$ is called a uniformly stated function.

The following proposition gives the local minimum of a uniformly stated function which consists of two min-type functions defined on $\mathbb{R}^2_+$.

Proposition 5.1.4. *Let $P(x) = \max\{l^{(1)}(x), l^{(2)}(x)\}$ where $l^{(1)} = (l^{(1)}_1, l^{(1)}_2)$ and $l^{(2)} = (l^{(2)}_1, l^{(2)}_2)$. Then $x^* = (x^*_1, x^*_2) \in riS$ is the local minimum of $P(x)$ if and only if the equality $\frac{x^*_1}{l_1} = \frac{x^*_2}{l_2}$ is satisfied where $l_1 = \max\{l^{(1)}_1, l^{(2)}_1\}$ and $l_2 = max\{l^{(1)}_2, l^{(2)}_2\}$.*

Let us give the theorem which is the basis for the algorithm for $n = 2$, and state the generalized case of Proposition 5.1.4. One can easily see this result geometrically.

Theorem 5.1.5. *Let $S \subset \mathbb{R}^2_+$ be the unit simplex and let*

$$P(x) = \max_{1 \le j \le k} \left\{ l^{(j)}(x) \right\}$$

be a uniformly stated function. A point $x^* \in riS$ is a local minimum if and only if there exist two vectors $l^{(k_1)}$ and $l^{(k_2)}$ such that $l_1^{(k_1)} > l_1^{(k_2)}$ and $l_2^{(k_2)} > l_2^{(k_1)}$ and the vector $l = (l_1^{(k_1)}, l_2^{(k_2)})$ is incomparable to all $l^{(j)}$ except $l^{(k_1)}$ and $l^{(k_2)}$. Moreover, the local minimum $x^* \in riS$ satisfies the equality $\frac{x_1^*}{l_1^{(k_1)}} = \frac{x_2^*}{l_2^{(k_2)}}$.

Using the theorem above, we give the algorithm for the solution of the problem (5) on $\mathbb{R}^2_+$ as follows.

Algorithm 5.1.6. ***Step 0***. *State $P(x)$ uniformly. Let L be the set of vectors composing the uniformly stated function $P(x)$. Find the number of the elements of the set L, call w to this number, and compute $h = C\binom{w}{2}$, then set $g = 1$ and $z = 1$.*

Step 1. *If $g \neq h + 1$, take any two vectors $l^{(i)}, l^{(j)}$ of L which were not taken before and go to step 2, otherwise $y^{(z-1)}$ is the global minimum, stop the algorithm.*

Step 2. *Set the vector $l = (l_1, l_2)$ where $l_1 = \max\left\{l_1^{(i)}, l_1^{(j)}\right\}$, $l_2 = \max\left\{l_2^{(i)}, l_2^{(j)}\right\}$ and $g = g + 1$. Compare this vector with the other vectors of L (except $l^{(i)}, l^{(j)}$). If there exists a vector smaller than this vector, go to Step 1, otherwise go to Step 3.*

Step 3. *Find the solution $x^* = (x_1, x_2)$ of $\dfrac{x_1}{l_1} = \dfrac{x_2}{l_2}$ on the unit simplex and set $y^{(z)} = x^*$. If $z = 1$, then $z = z + 1$, otherwise, if $P(y^{(z-1)}) < P(y^{(z)})$ then set $y^{(z)} = y^{(z-1)}$, $z = z + 1$. Go to Step 1.*

Now we can state the following theorem, which is a generalized version of Theorem 5.1.5. The algorithm for the general case completely depends on this theorem. The proof of the theorem will be given after a sequence of useful propositions.

Theorem 5.1.7. *Let $S \subset \mathbb{R}^n_+$ be the unit simplex and*

$$P(x) = \max_{1 \le j \le m} \left\{ l^{(j)}(x) \right\}$$

be a uniformly stated function defined on $\mathbb{R}^n_+$ where $m \geq n$. The point $x^* \in riS$ is local minimum of $P(x)$ if and only if there exist n vectors $l^{(k_1)}, \cdots, l^{(k_n)}$ among the vectors $l^{(1)}, \cdots, l^{(m)}$ such that

a)

$$l_i^{(k_i)} > \max_{\substack{j \in I \\ j \neq i}} l_i^{(k_j)} \quad i = 1, \cdots, n.$$

and

b) for $j \neq k_i$ no $l^{(j)}$ satisfies $l \gg l^{(j)}$ where $l = (l_1^{(k_1)}, \cdots, l_n^{(k_n)})$.

Moreover, the local minimum $x^* \in riS$ satisfies the following equality;

$$\frac{x_1^*}{l_1^{(k_1)}} = \cdots = \frac{x_n^*}{l_n^{(k_n)}}. \tag{6}$$

This theorem gives necessary and sufficient conditions for a point to be a local minimum of the given function. By using this theorem all local minima can be obtained. To do so, first take n vectors from $l^{(1)}, \cdots, l^{(m)}$. Then check if they satisfy properties a) and b). Checking condition a) can be implemented by forming an $n \times n$ matrix whose rows are the selected n vectors. Condition a) turns out to be that the biggest entry of each column of the matrix will be strictly larger than the other entries of the column and the maximum entries of the columns will lie on the diagonal. However, to check this condition, one should consider all the matrices made up by different permutations of these vectors. To avoid this difficulty, this condition can be changed to the the following condition: there will be no two biggest elements in one column and the biggest entries of any two different columns will not lie on the same row. Forming the matrices by using the vectors as rows can be used in checking condition b), too.

If these n vectors satisfy the conditions a) and b) then we can find x^* by using the equality given in the theorem and the fact that $x^* \in riS$. Also, to get all local points in riS, one has to make sure that all n combinations of the vectors $l^{(1)}, \cdots, l^{(m)}$ are checked by this procedure.

When any of these conditions is not satisfied, the function constructed by the considered combination with n vectors does not have a local minimum on riS. These cases can be seen in some examples in $\mathbb{R}^3$ geometrically. Since the feasible is the unit simplex, we take $z = 1 - x - y$, henceforth $P(x, y, z) = f(x, y)$ i.e., we can reduce the function on riS in $\mathbb{R}^3$ to a function on $\mathbb{R}^2$.

Example 5.1.8. Consider the function

$$\begin{aligned}\varphi(x, y) \quad = \quad & \max\left\{\min\left\{0.1x, 0.3y, 0.4(1 - x - y)\right\}, \min\left\{0.1x, 0.6y, 0.2(1 - x - y)\right\},\right. \\ & \left.\min\left\{0.5x, 0.1y, 0.7(1 - x - y)\right\}\right\}\end{aligned}$$

The corresponding vectors for min-type functions are $l^{(1)} = (10, \frac{10}{3}, \frac{5}{2})$, $l^{(2)} = (10, \frac{5}{3}, 5)$ and $l^{(3)} = (2, 10, \frac{10}{7})$. Now we check the condition a). Pick three vectors (there are already three vectors). Now form $l = (\max\{10, 10, 2\}, \max\left\{\frac{10}{3}, \frac{5}{3}, 10\right\}, \max\left\{\frac{5}{2}, 5, \frac{10}{7}\right\})$, (ith component of l is the maximum of the ith components of $l^{(1)}, l^{(1)}$ and $l^{(3)}$. so $l = (10, 10, 5)$. Since $l_1 = 10$ exists in the first component of both the vectors $l^{(1)}$ and $l^{(2)}$, i.e., $l_1 = 10$ is not only maximum with respect to the first component, therefore condition a) is not satisfied. Then as is seen from Figure 1, the function $\varphi(x, y)$ cannot have local minimum on riS.

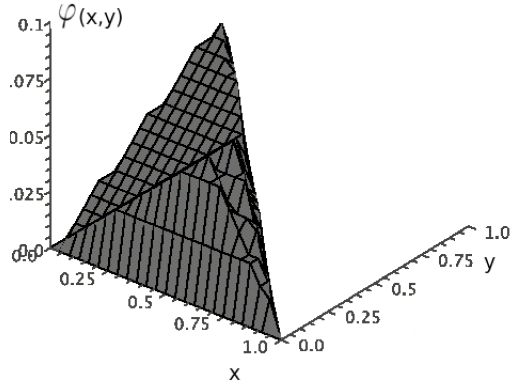

Figure 1: $\varphi(x, y)$ does not have a local minimum in riS.

Example 5.1.9. Consider the function

$$\begin{aligned}\varphi(x,y) &= \max\{\min\{6x,4y,1-x-y\},\min\{2x,10y,5(1-x-y)\},\\ &\qquad \min\{10x,2y,3(1-x-y)\},\min\{3x,y,10(1-x-y)\}\}\end{aligned}$$

The corresponding vectors for min-type functions are $l^{(1)} = (\frac{1}{6}, 0.25, 1)$, $l^{(2)} = (0.5, 0.1, 0.2)$, $l^{(3)} = (0.1, 0.5, \frac{1}{3})$ and $l^{(4)} = (\frac{1}{3}, 1, 0.1)$. We can choose four 3-combinations of these vectors. Let us analyze all 3-combinations.

Pick $l^{(1)}, l^{(2)}, l^{(3)}$. Form the vector $l = (0.5, 0.5, 1)$ as in previous example. l satisfies condition a). Compare it to the vector $l^{(4)} = (\frac{1}{3}, 1, 0.1)$. Since $l \gg l^{(4)}$ does not hold, it satisfies condition b). So this combinations yields a local minimum. The minimum point can be found by Equation (6). The local minimum is $(\frac{1}{4}, \frac{1}{4}, \frac{1}{2})$.

Pick $l^{(1)}, l^{(2)}, l^{(4)}$. Form the vector $l = (0.5, 1, 1)$. l satisfies condition a). Compare it to the vector $l^{(3)} = (\frac{1}{3}, 1, 0.1)$. Since $l \gg l^{(3)}$, it does not satisfy condition b). So this combinations does not yields a local minimum.

Pick $l^{(2)}, l^{(3)}, l^{(4)}$. Form the vector $l = (0.5, 1, \frac{1}{3})$ as in previous example. l satisfies condition a). Compare it to the vector $l^{(1)} = (0.25, 0.4, 1)$. Since $l \gg l^{(1)}$ does not hold, it satisfies condition b). So this combinations yields a local minimum. The minimum point can be found by Equation (6). The local minimum is $(\frac{3}{11}, 0.5, \frac{1}{6})$.

Pick $l^{(1)}, l^{(3)}, l^{(4)}$. Form the vector $l = (\frac{1}{3}, 1, 1)$. l does not satisfy condition a) because the first and second components of l is taken from $l^{(4)}$. So this combinations does not yields a local minimum.

Thus, as is seen from Figure 2, the function $\varphi(x,y)$ have two local minima on riS, the mimimum points are shown on the graph as A and B.

Example 5.1.10. Consider the function

$$\begin{aligned}\varphi(x,y) &= \max\{\min\{x,3y,9(1-x-y)\},\min\{9x,y,3(1-x-y)\},\\ &\qquad \min\{3x,9y,1-x-y\}\}.\end{aligned}$$

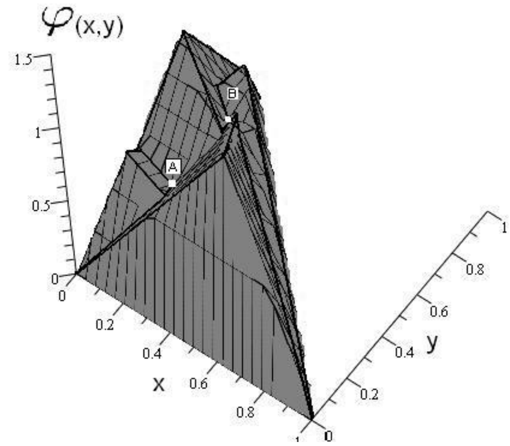

Figure 2: $\varphi(x, y)$ have the local minima A and B in riS.

The corresponding vectors for min-type functions are $l^{(1)} = (1, \frac{1}{3}, \frac{1}{9})$, $l^{(2)} = (\frac{1}{9}, 1, \frac{1}{3})$ and $l^{(3)} = (\frac{1}{3}, \frac{1}{9}, 1)$. $l = (1, 1, 1)$ is derived and conditions a) and b) are satisfied. As is seen from Figure 3, $\varphi(x, y)$ has a local minimum (the point A) on riS.

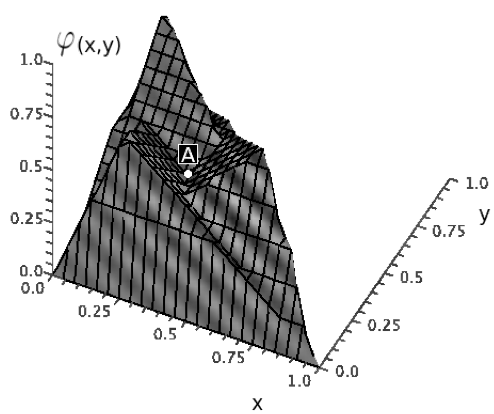

Figure 3: A is the local minimum of $\varphi(x, y)$ in riS.

Now we describe the algorithm. Note that in the cutting angle method, the number of initial points k is bigger than n.

Algorithm 5.1.11. ***Step 0*** *State $P(x)$ uniformly. Let L be the set of vectors composing the uniformly stated function $P(x)$. Find the number of the elements of the set L, call w this number and compute $h = C\binom{w}{n}$, then set $g = 1$ and $z = 1$.*

Step 1 *If $g \neq h+1$, take one of the vector combinations of the set L with n elements which is not taken before, let us denote this combination as $l^{(j_1)}, \cdots, l^{(j_n)}$ and go to step 2, otherwise $y^{(z-1)}$ is the global minimum, and the algorithm terminates.*

Step 2 *Set $g = g+1$. Set a matrix whose rows are the vectors $l^{(j_1)}, \cdots, l^{(j_n)}$. If there are no two biggest elements in any column and the biggest entries of any two different columns do not lie on the same row in this matrix, go to Step 3, otherwise go to Step 1.*

Step 3 *Set the vector $l = (l_1, \cdots, l_n)$ where*

$$l_1 = \max_{i \in I} \left\{ l_1^{(j_i)} \right\}, \cdots, l_n = \max_{i \in I} \left\{ l_n^{(j_i)} \right\}.$$

Compare l with the other vectors of L except $l^{(j_1)}, \cdots, l^{(j_n)}$ If there is a vector strictly smaller than this vector, go to Step 1, otherwise go to Step 4.

Step 4 *Find the point $x^* = (x_1, \cdots, x_n)$ satisfying the equalities $\frac{x_1}{l_1} = \cdots = \frac{x_n}{l_n}$ and $\sum_{i=1}^{n} x_i = 1$. Set $y^{(z)} = x^*$. If $z = 1$, then $z = z+1$, otherwise, if $P(y^{(z-1)}) < P(y^{(z)})$ then set $y^{(z)} = y^{(z-1)}$ and $z = z+1$. Go to Step 1.*

The following proposition allows us to simplify the objective function to easier find its local minima. It will allow us more easily find local minima of $P(x)$ using the fact that $P(x)$ is the maximum of exactly n min-type functions in a neighborhood of a local minimum.

Proposition 5.1.12. *Let $x^* \in riS \subset \mathbb{R}_+^n$ be a local minimum of a uniformly stated function $P(x) = \max_{1 \leq j \leq m} l^{(j)}(x)$ where $m \geq n$. Then there exists a set $\{k_1, \cdots, k_n\} \subset \{1, \cdots, m\}$ and a neighborhood of x^* such that $P(x) = \max_{j \in I} l^{(k_j)}(x)$ on this neighborhood of x^*.*

Proof. Let $x^* \in riS$ be a local minimum of a uniformly stated function $P(x) = \max_{1 \leq j \leq m} l^{(j)}(x)$. Let

$$M = \{1 \leq i \leq m : P(x^*) = l^{(i)}(x^*)\}.$$

Let $K = \{1, \cdots, m\} \setminus M$. Now since $P(x^*) > l^{(j)}(x^*)$ for $j \in K$ and all functions in $P(x)$ are continuous, in some neighborhood of x^*, $P(x) = \max_{j \in M} l^{(j)}(x)$. Let the cardinality of M be bigger than n. Now define

$$M_k = \{i \in M : \tfrac{x_k^*}{l_k^{(i)}} = P(x^*)\} \text{ for } k \in I.$$

First we notice $M_k \neq \emptyset$. This follows from the fact that x^* is a local minimum. Indeed, if $M_k = \emptyset$, one can find a feasible direction h sufficiently small in norm whose k^{th} component h_k is negative so that $P(x^* + h) - P(x^*) < 0$, this contradicts with the fact that x^* is a local minimum. Now since $P(x^*) = \max_{j \in M} l^{(j)}(x^*)$ and the cardinality of M is bigger than n, we can take such a k that the cardinality of M_k is bigger than 2. Since min-type functions are piecewise linear and continuous then in some neighborhood of x^*, the functions $l^{(i)}(x)$ for $i \in M_k$ are same. So in this neighborhood of x^* all min-type functions $l^{(i)}(x)$, $i \in M_k$ in the expression of $P(x)$ can be represented by one function for each k. Repeating this for each M_k, the number of min-type functions in the expression of $P(x)$ can be reduced to some number less than or equal to n. Since x^* is a local minimum of $P(x)$ the number of min-type functions can not be less than n. In fact, it is easy to see this because in this case one can easily find a feasible direction h such that $P(x^* + h) - P(x^*) < 0$. So $P(x)$ can be written as the maximum of n min-type functions, say $l^{(k_1)}, \cdots, l^{(k_n)}$ among $l^{(1)}, \cdots, l^{(m)}$. $\square$

Proposition 5.1.13. *Let $S \subset \mathbb{R}^n_+$ be the unit simplex and*

$$P(x) = \max_{j \in I} \left\{ l^{(j)}(x) \right\}$$

be a uniformly stated function defined on $\mathbb{R}^n_+$. If the point $x^* = (x_1^*, \cdots, x_n^*) \in riS$ is a local minimum of $P(x)$, then

$$P(x^*) = l^{(1)}(x^*) = \cdots = l^{(n)}(x^*).$$

Moreover, there exists a permutation σ such that

$$\frac{x_1^*}{l_1^{(\sigma(1))}} = \cdots = \frac{x_n^*}{l_n^{(\sigma(n))}} = P(x^*).$$

Proof. Let the point $x^* = (x_1^*, \cdots, x_n^*) \in riS$ be a local minimum of $P(x)$. Assume $l^{(1)}(x^*) = \cdots = l^{(n)}(x^*)$ is not true. So $\max_{i \in I} l^i(x^*) > \min_{i \in I} l^i(x^*)$. Let

$$R = \{1 \leq s \leq n : l^s(x^*) = \max_{i \in I} l^i(x^*) = P(x^*)\}.$$

By assumption the number of elements in R is less than n. Let $T = \{1, \cdots, n\} \setminus R$. Now

$$P(x^*) = \max\{\max_{i \in R} l^{(i)}(x^*), \max_{i \in T} l^{(i)}(x^*)\} \geq 0.$$

Since $P(x^*) = l^{(i)}(x^*) > l^{(j)}(x^*)$ for $i \in R$ and $j \in T$ and min-type functions are continuous, for a feasible direction h whose norm is sufficiently small $l^{(i)}(x^* + h) - P(x^*) < 0$ for $i \in T$. So

$$P(x^* + h) - P(x^*) = \max_{i \in R} l^{(i)}(x^* + h) - l^{(i)}.$$

For $i \in R$, there exists a j (say $j(i)$) such that $l^{(i)}(x^*) = \frac{x^*_{j(i)}}{l^{(i)}_{j(i)}}$. So for $i \in R$,

$$l^{(i)}(x^* + h) - l^{(i)}(x^*) = \min\{\cdots, h_{j(i)} \frac{x^*_{j(i)}}{l^{(i)}_{j(i)}}, \cdots\}.$$

By choosing a feasible direction h whose $j(i)^{th}$ component $h_{j(i)}$ is less than zero for all $i \in R$, (which is possible because the number of elements in R is less than n), we have

$$P(x^* + h) - P(x^*) = \max_{i \in R} l^{(i)}(x^*) < 0.$$

This is a contradiction. So the number of the elements in R is equal to n which proves the first part. So $R = \{1, \cdots, n\} = I$

For the second part, let us define $R_k = \{i \in R : \frac{x_k^*}{l_k^i} = P(x^*)\}$ for $k \in I$. Assume $R_{k'} = \emptyset$ for some $1 \leq k' \leq n$. So $\frac{x^*_{k'}}{l^i_{(k')}} > P(x^*)$ for all $i \in R$. Now we choose a feasible direction h whose norm is sufficiently small and $h_{k'} > 0$ and all the other components are negative. In this case $\frac{x^*_{k'} + h_{k'}}{l_i^{(k')}} - P(x^*) > 0$ for any $i \in I$. Now for this h, we get

$$P(x^* + h) - P(x^*) = \max_{i \in I} l^i(x^* + h) - l^{(i)}(x^*) < 0.$$

This shows $R_k \neq \emptyset$ for any k and since the number of the elements in R is n, for every $k \in I$ the number of the elements in R_k is 1. Thus, we define $\sigma(i)$ as the corresponding element in R_i. $\square$

The following theorem plays a very important role in developing the algorithm. It helps to prove the necessity part of Theorem 5.1.7.

Theorem 5.1.14. *Let the vectors $l^{(1)}, \cdots, l^{(n)} \in \mathbb{R}^n_+$ satisfy the following condition*

$$l_i^{(i)} > \max_{\substack{j \in I \\ j \neq i}} l_i^{(j)} \quad i = 1, \cdots, n. \tag{7}$$

In this case, $x^* = (x_1^*, \cdots, x_n^*) \in \mathbb{R}^n_+$ is the local minimum of the function

$$P(x) = \max_{k \in I} \left\{ l^{(k)}(x) \right\}$$

in the unit simplex if and only if the equality

$$\frac{x_1^*}{l_1^{(1)}} = \cdots = \frac{x_n^*}{l_n^{(n)}} \tag{8}$$

holds.

Proof. ($\Leftarrow$) Let us prove the sufficiency first. Suppose that $x^* = (x_1^*, \cdots, x_n^*)$ satisfies the equality (8). Then it can be easily shown by the hypothesis of the theorem that following equality holds:

$$P(x_1^*, \cdots, x_n^*) = \frac{x_1^*}{l_1^{(1)}} = \cdots = \frac{x_n^*}{l_n^{(n)}}.$$

Now, let us find such an $\epsilon > 0$ that for each feasible direction $h = (h_1, \cdots, h_n)$ which satisfies $\|h\| < \epsilon$, the inequality

$$P(x_1^* + h_1, \cdots, x_n^* + h_n) - P(x_1^*, \cdots, x_n^*) \geq 0$$

holds. Notice that by the equality above we have

$$P(x_1^*+h_1,\cdots,x_n^*+h_n)-P(x_1^*,\cdots,x_n^*)=\max_{1\le k\le n}\left\{\min_{1\le i\le n}\left\{\frac{x_i^*+h_i}{l_i^{(k)}}-\frac{x_i^*}{l_i^{(i)}}\right\}\right\}.$$

From the hypothesis (7) of the theorem $\frac{x_i^*}{l_i^{(k)}}-\frac{x_i^*}{l_i^{(i)}}>0$ for all $k\neq i$. In this case we can choose sufficiently small h_i's $(i\neq k)$ such that

$$\frac{x_i^*+h_i}{l_i^{(k)}}-\frac{x_i^*}{l_i^{(i)}}>0$$

holds.

Since h is a feasible direction $(\sum_{i=1}^{n}h_i=0)$, at least one of the h_i $(i=1,\cdots,n)$ must be positive. Without loss of generality, let us suppose $h_1>0$ and it is small enough to choose h_i's $(i\neq 1)$ sufficiently small such that

$$\frac{x_i^*+h_i}{l_i^{(1)}}-\frac{x_i^*}{l_i^{(i)}}>0.$$

In this case we have

$$l^{(1)}(x^*+h)-l^{(1)}(x^*)=\min\{\frac{h_1}{l_1^{(1)}},\{\min_{i\neq 1}\frac{x_i^*+h_i}{l_i^{(1)}}-\frac{x_i^*}{l_i^{(i)}}\}\}>0.$$

Since the function $P(x)$ is the maximum of min-type functions, for any feasible direction h, $P(x^*+h)-P(x^*)>0$. So x^* is a local minimum of $P(x)$ over riS.

($\Rightarrow$) Now, let us prove the necessity part. Let x^* be a local minimum point of P over riS. There exist such positive numbers $d_1,\cdots,d_n$ that the following equality holds:

$$P(x_1^*,\cdots,x_n^*)=\frac{x_1^*}{d_1}=\cdots=\frac{x_n^*}{d_n}.$$

Since x^* is the local minimum point, by Proposition 5.1.13 there exists a permutation σ such that $\frac{x_i^*}{l_i^{(\sigma(i))}}=P(x^*)$. So $l_i^{(\sigma(i))}=d_i$ for $i\in I$. Now assume that $\sigma(j)\neq j$ for some $j\in I$. Then because of the assumption $\frac{x_j^*}{l_j^{(j)}}<\frac{x_j^*}{l_j^{(\sigma(j))}}$ for a feasible direction h whose j^{th} component is positive i.e., $h_j>0$ we have

$$\frac{x_j^*+h_j}{l_j^{(j)}}-\frac{x_j^*}{l_j^{(\sigma(j))}}<0$$

, hence $l^{(\sigma(j))}(x^*+h)-l^{(\sigma(j))}(x^*)<0$. Now choose a feasible direction h_j such that $h_j>0$ and $h_i<0$ for $i\neq j$. In this case for $i\neq j$, $l^{(\sigma(i))}(x^*+h)-l^{(\sigma(i))}(x^*)<0$ because

$$l^{(\sigma(i))}(x^*+h)-l^{(\sigma(i))}(x^*)=\min\{\cdots,\frac{h_i}{l_i^{(\sigma(i))}},\cdots\}$$

where $h_i<0$. So for this feasible direction, we have $P(x^*+h)-P(x^*)<0$ which contradicts with the fact that x^* is a local minimum. So $\sigma(i)=i$ for all $i\in I$. Since $l_i^{(\sigma(i))}=d_i$ for $i\in I$, we get $l_i^{(i)}=d_i$ for $i\in I$. Thus the proof is completed. □

Proposition 5.1.15. *Let $x^*\in riS$ be a local minimum of a uniformly stated function $P(x)=\max_{j\in I} l^{(j)}(x)$. Then there exists a permutation θ such that the following condition is satisfied;*

$$l_i^{(\theta(i))}>\max_{\substack{j\in I\\ j\neq\theta(i)}} l_i^{(j)}\quad i=1,\cdots,n.$$

Proof. It is proved by contradiction. Let $x^*\in riS$ be a local minimum of a uniformly stated function $P(x)=\max_{j\in I} l^{(j)}(x)$. Assume such a permutation does not exist. In this case either (1) there exist at least two k and m such that $\max_{j\in I} l_k^{(j)}=l_k^{(r)}$ and $\max_{j\in I} l_m^{(j)}=l_m^{(r)}$ for some r or (2) there exist at least two indices j_1 and j_2 such that $\max_{j\in I} l_i^{(j)}=l_i^{(j_1)}=l_i^{(j_2)}$ for some i. We give the proof of the first case, the other case can be proved in a similar way.

Assume there exist at least two k and m such that $\max_{j\in I} l_k^{(j)}=l_k^{(r)}$ and $\max_{j\in I} l_m^{j}=l_m^{r}$ for some r. By Proposition 5.1.13 there exists a permutation σ such that

$$l_1^{(\sigma(1))}x_1^*=\cdots=l_n^{(\sigma(n))}x_n^*=P(x^*).$$

Let $r=\sigma(s)$ for some s. Now we get

$$\frac{x_m^*}{l_m^{(r)}}\leq\frac{x_m^*}{l_m^{(\sigma(m))}}=\frac{x_s^*}{l_s^{(r)}}\leq\frac{x_m^*}{l_m^{(r)}}$$

from the assumption and Proposition 5.1.13. This means $\frac{x_s^*}{l_s^{(r)}}=\frac{x_m^*}{l_m^{(r)}}$. Now let us consider

$$l^{(r)}(x^*+h)-l^{(r)}(x^*)=\min\{\cdots,\frac{h_m}{l_m^r},\cdots,\frac{h_s}{l_s^r},\cdots\}$$

for a feasible direction $h = (h_1, \cdots, h_n)$. In this case if we let $h_s > 0$ and all the others negative, then $P(x^* + h) - P(x^*) < 0$. So this means $r = \sigma(m)$. Rewriting the inequalities above we get $r = \sigma(k)$, too. This is a contradiction because σ is a permutation. $\square$

Now we can give the proof of Theorem 5.1.7.

Proof. Let us prove the necessity first. Let $S \subset \mathbb{R}^n_+$ be the unit simplex and

$$P(x) = \max_{1 \le j \le m} \left\{ l^{(j)}(x) \right\}$$

be a uniformly stated function defined on $\mathbb{R}^n_+$ where $m \ge n$ and let the point $x^* \in riS$ be a local minimum of $P(x)$. By Proposition 5.1.12, in a neighborhood of x^*, $P(x)$ can be written as

$$P(x) = \max_{j \in I} l^{(k_j)}(x)$$

where $1 \le k_j \le m$. By Proposition 5.1.15 we get the condition a). Now let l be defined as in part b). Let us consider the negation of part b). This means l is strictly greater than $l^{(j)}$ for $j \ne k_i$ where $1 \le i \le n$. So $l(x) < l^{(i)}(x)$ hence $P(x^*) = l(x^*) < l^{(i)}(x^*)$. But by the definition of $P(x)$, $P(x^*) \ge l^{(i)}(x^*)$, so we get a contradiction. So part b) holds. The equality given at the end of Theorem 5.1.7 follows from Theorem 5.1.14.

Let us prove the sufficiency now. There exists n vectors $l^{(k_1)}, \cdots, l^{(k_n)}$ among $l^{(1)}, \cdots, l^{(m)}$ such that a), b) are satisfied. In this case the existence of a local minimum of the function $\max_{i \in I} l^{(k_i)}(x)$ follows from Theorem 5.1.14. Let this point be x^*. We need to prove that a local minimum of $\max_{i \in I} l^{(k_i)}(x)$ is also a minimum of the function $P(x)$. For this, first we notice that the min-type function which is generated by the vector $l = (l_1^{(k_1)}, \cdots, l_n^{(k_n)})$ has its maximum value at x^*. For this reason if a vector l' is incomparable with l then $l(x^*) > l'(x^*)$. From part b) we know l is incomparable to some of the vectors in $\left\{ l^{(1)}, \cdots, l^{(n)} \right\}$ and for the rest of the vectors at least one of its components is equal to one component of the vector. Let

$$T = \{ i \in I : l^{(i)} \text{ is incomparable with } l \}$$

and

$$R = \{i \in I : \text{ at least one component of } \ l^{(i)} \text{ is equal to one component of } l\}.$$

Now for $i \in T$ $l(x^*) > l^{(i)}(x^*)$. So in some neighborhood of x^* we can write

$$\max_{i \in I} l^{(k_i)}(x) = \max\{\max_{i \in I} l^{(k_i)}(x), \max_{i \in T} l^{(i)}(x)\}.$$

Now let us show x^* is also a local minimum of $\max\limits_{i \in \{k_1, \cdots, k_n\} \cup R} l^{(i)}(x^*)$. For any feasible direction h, we have

$$\begin{aligned}
&\max_{i \in \{k_1, \cdots, k_n\} \cup R} l^{(i)}(x^* + h) - l^{(i)}(x^*) \\
&= \max\{\max_{i \in I}\{l^{(k_i)}(x^* + h) - l^{(k_i)}(x^*)\}, \max_{i \in R}\{l^{(i)}(x^* + h) - l^{(i)}(x^*)\}\} \\
&\geq 0
\end{aligned}$$

because $\max\limits_{i \in I}\{l^{(k_i)}(x^* + h) - l^{(k_i)}(x^*)\} \geq 0$. This proves that x^* is also a local minimum of $P(x)$.

□

5.2 The results of numerical experiments

To show the efficiency of Algorithm 4.3.1 supported with Algorithm 5.1.11, some numerical experiments have been carried out. We applied this algorithm to three global minimization problems with different increasing positively homogeneous objective functions. These problems are described in [6].

The code for Algorithm is written in Matlab. Numerical experiments are carried on Intel(R) Quad 2 Core CPU Q6600 2.40 GHz computer.

The following notations are used in all examples:

- $f = f(x)$ is objective function;
- n is the number of variables;

Table 1: Results for $f_1(x)$

n	k	t
2	9	0.59
3	40	5.55
4	55	49.26
5	60	776

- k is the number of iterations;
- t is the computational time.

The stopping criteria is expressed as in the Algorithm 4.3.1. Since the required accuracy in [6] is 10^{-2}, the examples are solved in this accuracy.

Example 5.2.1. Let

$$a_k^i = \tfrac{20i}{k(1+|i-k|)}, k = 1, 2, \cdots, n; i = 1, 2, \cdots, 40$$

and

$$b_k^j = 5\left|\sin(j)\sin(k)\right|, k = 1, 2, \cdots, n; j = 1, 2, \cdots, 20.$$

$$f_1(x) = \max\left\{[a^i, x] : i = 1, 2, \cdots, 40\right\} + \min\left\{[b^j, x] : j = 1, 2, \cdots, 20\right\},$$

Example 5.2.2. Let

$$a_k^{ij} = \tfrac{10j}{k(1+|k-j|)}\left|\cos(i-1)\right|, i = 1, 2, \cdots, 20, j = 1, 2, \cdots, n, k = 1, \cdots, n.$$

$$f_2(x) = \max_{1\le i\le 20}\min_{j\in I}[a^{ij}, x],$$

Table 2: Results for $f_2(x)$

n	k	t
2	6	0.5
3	12	2.06
4	16	6.43
5	22	19.61

Table 3: Results for $f_3(x)$

n	k	t
2	6	1.16
3	10	4.03
4	16	10.45
5	20	33.44

Example 5.2.3. Let

$$a_{ij} = \begin{cases} 12 + \frac{n}{i} & \text{if } i = j \\ 0 & \text{if } i = j + 1 \\ 0 & \text{if } j = i + 2 \\ \frac{15}{i + 0.1j} & \text{Otherwise.} \end{cases}$$

$$f_3(x) = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \right)^{\frac{1}{2}},$$

Comments on experiments

In the experiments, we observed the combinatorial nature of the subproblem. When compared to the algorithms proposed in [6, 42] in terms of the number of iterations, the algorithm proposed in this thesis uses less iterations. This is because the subproblem is solved exactly. The exact solution of the subproblem makes it necessary to check all combinations of the vectors. This requirement causes the algorithm slow down as the number of variable increases.

5.3 Conclusion

The most important step of the Cutting Angle Method ([3, 4]) suggested for global optimization problems with IPH functions is that in each iteration, an optimization problem with minimum type functions should be solved. A new algorithm for this problem is presented (Algorithm 3). The algorithm is mainly based on Theorem 5.1.7. The presented algorithm can be commented different from others in that the algorithm can be expressed geometrically and it is proved that it solves the subproblem exactly.

REFERENCES

[1] G. Adilov, G. Tinaztepe and S. Kemali, *On the Solution of an Optimization Problem Expressed via Increasing Positively Homogeneous Functions*, *Proceedings of The 16th Jangjeon Mathematical Society*, 4-6 July, pp. 1-6, 2005.

[2] G. Adilov, G. Tinaztepe and R. Tinaztepe, *On the global minimization of IPH functions over the unit simplex*, *Int. J. Comput. Math.* (2009) (accepted).

[3] M.Yu. Andromonov, A.M. Rubinov, and B.M. Glover, *Cutting Angle Method for Minimizing Increasing Convex-along-rays Functions*, *Research Report 97/7, SITMS, University of Ballarat*, Ballarat, 1997.

[4] M.Yu. Andromonov, A.M. Rubinov, and B.M. Glover, *Cutting Angle Methods in Global Optimization*, *Appl. Math. Lett.*, Vol.12 (1999), pp. 95-100.

[5] A. DJ. Babayev, *An Exact Method for Solving the Subproblem of the Cutting Angle Method of Global Optimization*, *Optimization and Related Topics, in Kluwer Academic Publishers*, ser. Applied Optimization, Dordrecht, Vol. 47, pp. 15-26, 2000.

[6] A. Bagirov and A.M. Rubinov, *Global Minimization of Increasing Positively Homogeneous Function Over the Unit Simplex*, *Ann. Oper. Res.*, 98 (2000), pp. 171-187.

[7] A. Bagirov and A.M. Rubinov, *Modified Versions of the Cutting Angle Method*, In: *Nonconvex optimization and its applications, Vol. 54, Advances in Convex Analysis and Global Optimization, N. Hadjisavvas and P.M. Pardalos (eds.)*, Kluwer Publishers, Dordrecht, pp. 245-268, 2001.

[8] A. Bagirov and A.M. Rubinov, *Cutting Angle Method and a Local Search*, *J. Global Optim.*, 27 (2003), pp.193-213.

[9] R. Balian, *Un principle d'incertitude fort en theorie du signal ou en mecanique quantique*, *C. R. Acad. Sci. Paris Ser. II Mec. Phys. Chim. Sci. Univers Sci. Terre*, 292(20) (1981), pp. 1357–1362.

[10] S. G. Bartels and B. M. Glover, *Continuous selections of linear functions and nonsmooth critical point theory* *Nonlinear Anal.*, TMA 24 (1995), pp. 385-407.

[11] J. J. Benedetto, W. Czaja, P. Gadziński, and A. M. Powell, *The Balian–Low Theorem and regularity of Gabor systems*, *J. Geom. Anal.*, 13 (2003), pp. 239–254.

[12] J. J. Benedetto, W. Czaja, and A. M. Powell, *An optimal example for the Balian–Low uncertainty principle*, *SIAM J. Math. Anal.*, 38 (2006), pp. 333–345.

[13] J. J. Benedetto, W. Czaja, A. M. Powell, and J. Sterbenz, *An endpoint* $(1, \infty)$ *Balian–Low theorem*, *Math. Res. Lett.*, 13 (2006), pp. 467–474.

[14] J. J. Benedetto, C. Heil and D. F. Walnut, *Differentiation and Balian–Low Theorem*, *J. Fourier Anal. Appl.*, 1(4) (1995), pp. 355–402.

[15] A. Bényi and K. Okoudjou, *Bilinear pseudodifferential operators on modulation spaces*, *J. Fourier Anal. Appl.*, 10 (2004), pp. 301–313.

[16] C.E. Cook and M. Bernfeld, *Radar Signals–An introduction to Theory and Applications*, Academic Press, New York, 1967.

[17] I. Daubechies, *The wavelet transform, time-frequency localization and signal analysis*, *IEEE Trans. Inform. Theory*, 39 (1990), pp. 961–1005.

[18] C. Fefferman, *Characterizations of bounded mean oscillation*, *Bull. Amer. Math. Soc.*, 77 (1972), pp. 587–588.

[19] H. G. Feichtinger, *Modulation spaces on locally compact Abelian groups*, *Technical Report*, University of Vienna, 1983.

[20] H. G. Feichtinger, *Atomic characterizations of modulation spaces through Gabor-type representations*, In *Proc. Conf. Constructive Function Theory*,Edmonton, pp. 113–126, July 1986.

[21] H. G. Feichtinger, K. Gröchenig and D. Walnut, *Wilson Bases and modulation Spaces*, *Math. Nachr.*, 155 (1992), pp. 7–17.

[22] H. G. Feichtinger and K. Gröchenig, *Banach spaces related to integrable group of representations and their atomic decompositions*, *Monatsh. Math.*,108(2-3) (1989), pp. 129–148.

[23] G. B. Folland, *Real Analysis: Modern Techniques and Their Applications*, Pure and Applied Mathematics, Wiley, New York, second ed., 1999.

[24] J.-P. Gabardo and D. Han, *Balian–Low phenomenon for subspace Gabor frames*, *J. Math. Phys.*, 45 (2004), pp. 3362–3378.

[25] S. Z. Gautam, *A critical-exponent Balian–Low Theorem*, *Math. Res. Letters*, 15 (2008), pp. 471–483.

[26] I. M. Gel'fand, *Expansion in characteristic functions of an equation with periodic coefficients (Russian)*, *Doklady Akad. Nauk SSSR (N.S.)* 73 (1950), pp. 1117–1120.

[27] K. Gröchenig, *Foundations of Time-Frequency Analysis*, Birkhäuser, Boston, 2001.

[28] K. Gröchenig, D. Han, C. Heil and G. Kutyniok, *The Balian–Low Theorem for symplectic lattices in higher dimensions*, *Appl. Comput. Harmon. Anal.*, 13 (2002), pp. 169–176.

[29] K. Gröchenig and C. Heil, *Modulation spaces and pseudodifferential operators*, *Integral Equations Operator Theory*, 34 (1999), pp. 439–457.

[30] C. Heil and R. Tinaztepe, *Embedding of Modulation Spaces into BMO*,(preprint) (2010).

[31] A. J. E. M. Janssen, *Proof of a conjecture on the supports of Wigner distributions*, *J. Fourier Anal. Appl.*, 4 (6) (1998), pp. 723–726.

[32] A. J. E. M. Janssen, *Zak transform characterization of S_0*, *Sampl. Theory Signal Image Process.*, 5 (2006), pp. 141–162.

[33] Y. Katznelson, *An introduction to Harmonic Analysis*, Dover, New York, 1976.

[34] M. Kobayashi, *Modulation Spaces $M^{p,q}$ for $0 < p, q \leq \infty$* ,*J. Funct. Spaces and Appl.*, 4 (3) (2006), pp. 329-341.

[35] S. S. Kutateladze and A. M. Rubinov, *Minkowski Duality and its applications*, *Russian Mathematics Surveys*, 27 (1972), pp. 137-191.

[36] H. E. Lieb, *Integral Bounds for Radar Ambiguity Functions and Wigner Distributions*, *J. Math. Phys.*, 31 (1990), pp.594-599.

[37] F. Low, *Complete sets of wave packets*, in C. DeTar, editor, *A passion for physics-Essay in honor of Geoffrey Chew* (1985), World Scientific, Singapore, pp. 17–22.

[38] U.G. Nuriyev, *An Approach to the Subproblem of the Cutting Angle Method of Global Optimization*, *J. Global Optim.*, 31 (3) (2005), pp. 353-370.

[39] K. A. Okoudjou, *Embedding of some classical Banach spaces into modulation spaces*, *Proc. Amer. Math. Soc.*, 132 (2004), pp. 1639–1647.

[40] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, New Jersey, 1970.

[41] A. M. Rubinov and M. Andramonov, *Lipschitz Programming via Convex-along-rays Functions*, *Optim. Methods and Softw.*, 10 (1999), pp. 763-781.

[42] A. M. Rubinov, *Abstract Convexity and Global Optimization*, Dordrecht, Kluwer Academic publishers, 2000.

[43] D. Sarason, *Functions of vanishing mean oscillation*, *Trans. Amer. Math. Soc.*, 207 (1975), pp. 391–405.

[44] I. Singer, *Abstract Convex Analysis*, Wiley Interscience Publication, New York, 1997.

[45] E. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals.*, *Princeton Mathematical Series*, 43., Princeton, NJ,: Princeton University Press, 1993.

[46] T. Strohmer, *Pseudodifferential operators and Banach algebras in mobile communications*, *Appl. Comput. Harmon. Anal.*, 20 (2006), pp. 237–249.

[47] T. Strohmer and S. Beaver, *Optimal OFDM system design for time-frequency dispersive channels*, *IEEE Trans. Communications*, 51 (2003), pp. 1111–1122.

[48] T. Strohmer and R. W. Heath, Jr., *Grassmannian frames with applications to coding and communication*, *Appl. Comput. Harmon. Anal.*, 14 (2003), pp. 257–275.

[49] J. Toft, *Continuity properties for modulation spaces, with applications to pseudo-differential calculus I*, *J. Funct. Anal.*, 207 (2004), pp. 399–429.

[50] J. Toft, *Convolutions and embeddings for weighted modulation spaces*, in: *"Advances in Pseudo-Differential operators, R. Ashino et al.*, eds., Birkhäuser, Basel (2004), pp. 165–186.

[51] H. Triebel, *Modulation spaces on the Euclidean n-space*, *Z. Anal. Anwendungen*, 2(5) (1983), pp. 443–457.

[52] D. F. Walnut, *Lattice size estimates for Gabor decompositions*, *Monatsh. Math.* 115 (1993), pp. 245–256.

[53] P.M. Woodward, *Probability and Information Theory, with applications to radar*, Pergamon Press, Oxford, Second Edition, 1964.

VITA

Ramazan Tinaztepe was born in Antalya in 1978. He got his B.S. in Math Education from Middle East Technical University in 1999. He got M.S. in Mathematics from Akdeniz University in 2001. He worked as a research assistant in Akdeniz University. Then he began Ph.D. in mathematics at Georgia Institute of Technology in 2004. He is married and has two children.