# ABSTRACT


Title of dissertation: APPLICATIONS OF NONHARMONIC FOURIER ANALYSIS AND SINGLE PIXEL CAMERA DESIGN

Chenzhi Zhao
Doctor of Philosophy, 2020

Dissertation directed by: Professor John Benedetto
Department of Mathematics



In the first part of this thesis, we focus on the Poisson summation formula in the non-periodic setting. We introduce the concept of quasicrystal and construct new types of Poisson summation formulas on it. We then develop the theory of sampling and interpolation on model sets. We will prove some uniqueness results regarding exact reconstruction on model sets.

In the second part of this thesis, we focus on compressed sensing and image super-resolution. We develop a technique to reconstruct an image with few measurements but high resolution. Experimental results will be shown.

# APPLICATIONS OF NONHARMONIC FOURIER ANALYSIS AND SINGLE PIXEL CAMERA DESIGN

by

Chenzhi Zhao

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2020

Advisory Committee:
Professor John J. Benedetto, Chair/Advisor
Professor Radu V. Balan
Professor Wojciech Czaja
Professor Kasso A. Okoudjou
Dean's Representative:
Professor Yanne K. Chembo

# Dedication

To my parents

# Acknowledgments

I owe my gratitude to all the people who have made this thesis possible and because of whom my graduate experience has been one that I will cherish forever.

First and foremost I'd like to thank my advisor, Professor John Benedetto for giving me an invaluable opportunity to work on this challenging and extremely interesting topic over the past four years. He has provided me with a lot of deep insights and helpful discussions throughout graduate school. It has been a pleasure to work with and learn from such an extraordinary individual.

I would also like to thank Dr. Alfredo Nava-Tudela. Without his extraordinary theoretical ideas and computational expertise, the second part of this thesis would have been a distant dream. I would also like to thank ARO grant W911NF-17-1-0014, which supported part of my research.

I owe my deepest thanks to my wife Yongzheng Qi who gave me her unconditional support and love.

It is impossible to remember all, and I apologize to those I've inadvertently left out.

# Table of Contents

# Chapter 1: INTRODUCTION

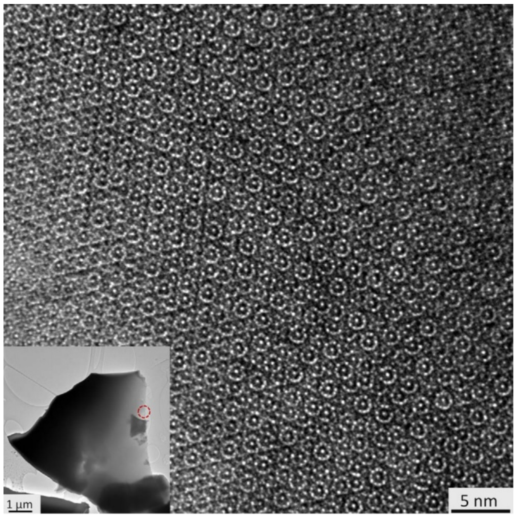
## 1.1 Introduction

Harmonic analysis is one of the most important area in mathematics. It is central in partial differential equations, signal processing and many other parts of science and technology. Initially, harmonic analysis is the analysis and synthesis of functions in terms of harmonics or basis functions $\{e^{2\pi int}\}, n \in \mathbb{Z}$. The frequencies belonging to the integer group $\mathbb{Z}$ is quite essential. For example, the uniqueness of Fourier series says that if we know the Fourier coefficients of a function over all the integers, then the function is uniquely determined.

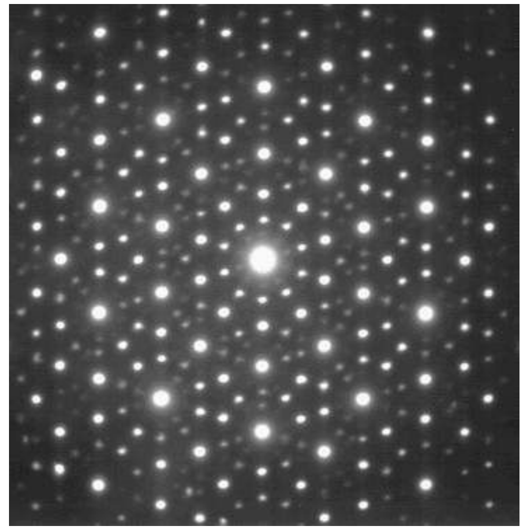
Later on, people realized that the integer group $\mathbb{Z}$ can be generalized to the setting of Euclidean space and even more generally to locally compact abelian groups. Thus, we have the emergence of abstract harmonic analysis. However another line of investigation began with the attempt to replace the group $\mathbb{Z}$ by other non-group structures. This is so called non-harmonic Fourier analysis. The essential question that most research revolves around is: for what set of frequencies $\{\lambda_n\}, n \in \mathbb{Z}$ does the collection $\{e^{2\pi i\lambda_n t}\}, n \in \mathbb{Z}$ have similar properties as the collection $\{e^{2\pi int}\}, n \in \mathbb{Z}$. Commonly, such properties are concerned with the expansion properties of the exponentials over certain Hilbert spaces. For example, when do they form an orthonormal basis, a Riesz basis, a Riesz sequence, a frame, or a set of uniqueness. Various examples have been studied by several authors. Duffin and Schaeffer, who first introduced the concept of frames, gave a density condition in the one dimensional case for which the collection $\{e^{2\pi int}\}, n \in \mathbb{Z}$ forms a frame over an interval [13]. Later on, Jaffard [17] gave a formula for the frame radius, that is the maximal length of the interval for which $\{e^{2\pi int}\}, n \in \mathbb{Z}$ forms a frame, in terms of the lower Beurling density. Landau [19] gave necessary conditions in terms of Beurling lower and upper densities for the collection $\{e^{2\pi int}\}, n \in \mathbb{Z}$ to form a frame, respectively, a Riesz sequence in higher dimensions. The sufficient conditions in one dimension was given by Beurling [5] using the theory of balayage. Among the various examples, we are particularly interested in what is called the quasicrystal, or its mathematical term, model set. The work was led by research of Yves Meyer, Basarab Matei, Alexander Olevskii, Nir Lev.

## 1.2 Background

It has long been believed that crystals can possess only two, three, four, and six-fold rotational symmetries. Dan Shechtman first observed ten-fold electron diffraction patterns in 1982, as described in his notebook. Shechtman and Blech jointly wrote a paper entitled "The Microstructure of Rapidly Solidified Al6Mn" [32] and sent it for publication around June 1984 to the Journal of Applied Physics (JAP). A further study of Khatyrka meteorites revealed micron-sized grains of another natural quasicrystal, which has a ten-fold symmetry and a chemical formula of Al71Ni24Fe5. See figure 3.1.




(a) $Al_{71}Ni_{24}Fe_5$



(b) Ho–Mg–Zn

Figure 1.1: (a) Atomic image of a micron-sized grain of the natural $Al_{71}Ni_{24}Fe_5$ quasicrystal from a Khatyrka meteorite [4]. (b) Electron diffraction pattern of an icosahedral Ho–Mg–Zn quasicrystal.

In 1992, the International Union of Crystallography altered its definition of a crystal, broadening it as a result of Shechtman's findings, reducing it to the ability to produce a clear-cut diffraction pattern and acknowledging the possibility of the ordering to be either periodic or aperiodic.

The introduction of Meyer's mathematical model of quasicrystals in 1972 [25] was even earlier than the discovery of quasicrystals in nature by Dan Shechtman in 1984 [32], who on won the Nobel Prize in Chemistry in 2011 due to his discovery. There are several ways to mathematically define quasicrystalline patterns. One definition, proposed by Meyer [25], the "cut and project" construction, is based on the work of Harald Bohr on almost periodic functions. A crystalline sample is by definition periodic; a crystal is composed of many unit cells repeated indefinitely in three independent directions. Such periodic systems have a Fourier transform that is concentrated at periodically repeating points in reciprocal space known as Bragg peaks; the Bragg peaks correspond to the reflection spots observed in the diffraction image. The mathematical counterpart of physical diffraction is the Fourier transform and the qualitative description of a diffraction picture as 'clear cut' or 'sharp' means that singularities are present in the Fourier spectrum. The discovery of a pure atomic measure with aperiodic atomic diffraction seemed to have predicted the existence of quasicrystals. An atomic measure $\sum_{\lambda\in\Lambda}\delta_\lambda$ with atomic spectrum $S$ will give rise to a Poisson summation formula,

$$\Big(\sum_{\lambda\in\Lambda}\delta_\lambda\Big)\widehat{} = \sum_{y\in S}\delta_y.$$

This motivated Meyer to start looking for non-traditional Poisson summation formulas [27]. We know that the classical sampling theorem can be proved via the classical Poisson summation formula, e.g. [2]. We do not know if this can be generalized in the same way to the new Poisson summation formulas with non-periodic support and spectrum. However, it is proved by Matei and Meyer [24] that quasicrystals defined by one dimensional windows are sets of stable sampling and interpolation for Paley Wiener spaces. In other words, it is an important question to ask what spanning property does the Fourier exponentials $\{e^{2\pi i\lambda t}\}, \lambda\in\Lambda$ possess when $\Lambda$ is a quasicrystal. Do they form a basis, a frame, or a Riesz sequence. The aim of subsections 1.2.1 and 1.2.2 is to give a quick review of the theory of frames and

Riesz basis.

**Definition 1.2.1.** Let $f \in L^1(\mathbb{T})$. The Fourier series of $f$ is the series

$$S(f)(t) = \sum_{n\in\mathbb{Z}} \hat{f}(n)e^{2\pi int}, \tag{1.1}$$

where

$$\hat{f}(n) = \int_{\mathbb{T}} f(t)e^{-2\pi int}dt. \tag{1.2}$$

**Definition 1.2.2.** Let $f \in L^1(\mathbb{R})$. The Fourier transform of $f$ is the function $\hat{f}$ defined as

$$\hat{f}(\gamma) = \int_{\mathbb{R}} f(t)e^{-2\pi it\gamma}dt. \tag{1.3}$$

An important construction of the theory of classical Fourier series is that the frequencies of the basic waves all belong to the group of integers $\mathbb{Z}$, which enables us to have the following two powerful results in harmonic analysis.

(1) The Poisson summation formula, which says for functions satisfying certain decay conditions we have $\sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} \hat{f}(n)$.

(2) The collection $\{e^{2\pi int}, n \in \mathbb{Z}\}$ forms an orthonormal basis of $L^2(\mathbb{T})$.

The Poisson summation formula has very important applications in number theory. It is used to derive a variety of of functional equations including the functional equation for the Riemann zeta function [33]. The ONB property has many applications in signal processing, such as the classical sampling theorem.

However, as the classical theory continues to impact various fields of mathematics, it becomes a growing interest for people to investigate the possibility of building the theory when the frequency set is significantly different from $\mathbb{Z}$.

The Poisson summation formula can also be written as $(\sum_{n\in\mathbb{Z}} \delta_n)\widehat{} = \sum_{n\in\mathbb{Z}} \delta_n$. More generally we want to know whether there exists a summation formula $(\sum_{\lambda\in\Lambda} \delta_\lambda)\widehat{} = \sum_{y\in S} \delta_y$, where the sets $\Lambda$ and $S$ both do not have any group structure. More precisely, we want both $\Lambda$ and $S$ to be locally finite, and each finite subset of them is linearly independent over $\mathbb{Q}$. One of the earliest result was hidden in an almost forgotten paper by Guinand [15]. Some seminal constructions were given by Meyer [27], Kolountzakis [18], Lev and Olevskii [21]. The supports of such measures also have an interesting connection with quasicrystals. In particular, it has been shown that for any quasicrystal, we can associate a Poisson summation formula.

### 1.2.1 Frame theory

The effort to generalize (2) was pioneered by Duffin and Schaeffer. They defined the notion of frame which generalizes orthonormal basis [13]. They showed that for a set of frequencies $\{\lambda_n, n \in \mathbb{Z}\}$ that satisfies a certain density condition, we were able to have a frame of exponentials $\{e^{2\pi i\lambda_n t}, n \in \mathbb{Z}\}$ for $L^2(\mathbb{T})$. The following is a brief introduction to frame theory. More details can be found in [8].

We denote by $\mathcal{H}$ a separable Hilbert space with inner product $\langle\cdot,\cdot\rangle$.

**Definition 1.2.3.** A sequence $\{f_k\}, k \in \mathbb{N}$ of elements in $\mathcal{H}$ is a frame for $\mathcal{H}$ if there exist constants $A, B > 0$ such that

$$A\|f\|^2 \leq \sum_{k=1}^{\infty} |\langle f, f_k\rangle|^2 \leq B\|f\|^2, \quad \forall f \in \mathcal{H}. \tag{1.4}$$

The numbers $A, B$ are called frame bounds. A frame is tight if we can choose $A = B$ as frame bounds; a tight frame with bound $A = B = 1$ is called a Parseval frame. If a frame ceases to be a frame when an arbitrary element is removed, it is called an exact frame.

Since a frame $\{f_k\}, k \in \mathbb{N}$ is a Bessel sequence, the operator

$$T : l^2(\mathbb{N}) \to \mathcal{H},\ T\{c_k\} = \sum_{k=1}^{\infty} c_k f_k \tag{1.5}$$

is bounded. T is called the synthesis operator or the pre-frame operator. The adjoint operator is given by

$$T^* : \mathcal{H} \to l^2(\mathbb{N}),\ T^* f = \{\langle f, f_k\rangle\}_{k=1}^{\infty}. \tag{1.6}$$

The operator $T^*$ is called the analysis operator. By composing $T$ and $T^*$, we obtain the frame operator

$$S : \mathcal{H} \to \mathcal{H},\ Sf = TT^* f = \sum_{k=1}^{\infty} \langle f, f_k\rangle f_k. \tag{1.7}$$

Below are some important properties of $S$:

**Lemma 1.2.1.** *Let $\{f_k\}, k \in \mathbb{N}$ be a frame with frame operator $S$ and frame bounds*

*$A, B$. Then the following hold:*

*(1) $S$ is bounded, invertible, self-adjoint, and positive.*

*(2) $\{S^{-1}f_k\}, k \in \mathbb{N}$ is a frame with bounds $B^{-1}, A^{-1}$; if $A, B$ are optimal obounds for $\{f_k\}, k \in \mathbb{N}$, then $B^{-1}, A^{-1}$ are optimal for $\{S^{-1}f_k\}, k \in \mathbb{N}$. The frame operator for $\{S^{-1}f_k\}, k \in \mathbb{N}$ is $S^{-1}$.*

Proof: (1) S is bounded as a composition of two bounded operators. Since $S^* = (TT^*)* = TT^* = S$, $S$ is self-adjoint. The inequality (1.4) means that $A\|f\|^2 \leq \langle Sf, f\rangle \leq B\|f\|^2$ for all $f \in \mathcal{H}$, or, equivalently, $AI \leq S \leq BI$, thus $S$ is positive. Furthermore, $0 \leq I - B^{-1}S \leq \frac{B-A}{B}I$, and consequently

$$\|I - B^{-1}S\| = \sup_{\|f\|=1} |\langle (I - B^{-1}S)f, f\rangle| \leq \frac{B-A}{B} < 1,$$

thus $S$ is invertible.

(2) Note that for $f \in \mathcal{H}$,

$$\sum_{k=1}^{\infty} |\langle f, S^{-1}f_k\rangle|^2 = \sum_{k=1}^{\infty} |\langle S^{-1}f, f_k\rangle|^2 \leq B\|S^{-1}f\|^2 \leq B\|S^{-1}\|^2\|f\|^2.$$

That is, $\{S^{-1}f_k\}, k \in \mathbb{N}$ is a Bessel sequence. It follows that the frame operator for $\{S^{-1}f_k\}, k \in \mathbb{N}$ is well defined. By definition, it acts on $f \in \mathcal{H}$ by

$$\sum_{k=1}^{\infty} \langle f, S^{-1}f_k\rangle S^{-1}f_k = S^{-1}\sum_{k=1}^{\infty} \langle S^{-1}f, f_k\rangle f_k = S^{-1}SS^{-1}f = S^{-1}f. \tag{1.8}$$

This shows that the frame operator for $\{S^{-1}f_k\}, k \in \mathbb{N}$ equals $S^{-1}$. The operator

$S^{-1}$ commutes with both $S$ and $I$, thus we can multiply the inequality $AI \leq S \leq BI$ with $S^{-1}$ this gives

$$B^{-1}iI \leq S - 1 \leq A^{-1}I,$$

i.e.,

$$B^{-1}\|f\|^2 \leq \langle S^{-1}f, f\rangle \leq A^{-1}\|f\|^2, \ \forall f \in \mathcal{H}.$$

Via (1.8),

$$B^{-1}\|f\|^2 \leq \sum_{k=1}^{\infty} |\langle f, S^{-1}f_k\rangle|^2 \leq A^{-1}\|f\|^2, \ \forall f \in \mathcal{H},$$

thus $\{S^{-1}f_k\}, k \in \mathbb{N}$ is a frame with frame bounds $B^{-1}, A^{-1}$. The argument for the optimality follows easily from the above equations. $\square$

The frame $\{S^{-1}f_k\}, k \in \mathbb{N}$ is called the canonical dual frame of $\{f_k\}, k \in \mathbb{N}$ because it plays the same role in frame theory as the dual of a basis. The frame decomposition, stated below, is the most important frame result. It shows that if $\{f_k\}, k \in \mathbb{N}$ is a frame for $\mathcal{H}$, then every element in $\mathcal{H}$ has a representation as a superposition of the frame elements. Thus it is natural to view a frame as some kind of a generalization of basis.

**Theorem 1.2.2.** *Let $\{f_k\}, k \in \mathbb{N}$ be a frame with frame operator $S$. Then*

$$f = \sum_{k=1}^{\infty} \langle f, S^{-1}f_k\rangle f_k, \quad \forall f \in \mathcal{H}, \tag{1.9}$$

*and*

$$f = \sum_{k=1}^{\infty} \langle f, f_k\rangle S^{-1}f_k, \quad \forall f \in \mathcal{H}. \tag{1.10}$$

*Both series converge unconditionally for all $f \in \mathcal{H}$.*

Proof: Let $f \in \mathcal{H}$. We have

$$f = SS^{-1}f = \sum_{k=1}^{\infty} \langle S^{-1}f, f_k \rangle f_k = \sum_{k=1}^{\infty} \langle f, S^{-1}f_k \rangle f_k.$$

Since $\{f_k\}, k \in \mathbb{N}$ is a Bessel sequence and $\{\langle f, S^{-1}f_k \rangle\} \in l^2(\mathbb{N})$, it follows that the series convervges unconditionally. The second expansion (1.10) follows similarly by using $f = S^{-1}Sf$. $\square$

In general, it is very difficult to directly apply the frame decompositions (1.9) and (1.10) since the operator $S^{-1}$ is unknown. However, for tight frames, this problem is circumvented.

**Corollary 1.2.3.** *If $\{f_k\}, k \in \mathbb{N}$ is a tight frame with frame bound $A$, then the canonical dual frame is $\{A^{-1}f_k\}, k \in \mathbb{N}$ and*

$$f = \frac{1}{A} \sum_{k=1}^{\infty} \langle f, f_k \rangle f_k, \quad \forall f \in \mathcal{H}.$$

### 1.2.2 Riesz basis

The set of all orthonormal bases can be characterized by all unitary operators acting on a single orthonormal basis. The definition of a Riesz basis is given by weakening the condition on the operator from being unitary to being bounded.

**Definition 1.2.4.** A Riesz basis for $\mathcal{H}$ is a family of the form $\{Ue_k\}, k \in \mathbb{N}$, where $\{e_k\}, k \in \mathbb{N}$ is an orthonormal basis for $\mathcal{H}$ and $U : \mathcal{H} \to \mathcal{H}$ is a bounded bijective operator. The dual basis associated to a Riesz basis is also a Riesz basis as the following theorem shows.

**Theorem 1.2.4.** *If $\{f_k\}, k \in \mathbb{N}$ is a Riesz basis for $\mathcal{H}$, there exists a unique sequence $\{g_k\}, k \in \mathbb{N}$ in $\mathcal{H}$ such that*

$$f = \sum_{k=1}^{\infty} \langle f, g_k \rangle f_k, \ \forall f \in \mathcal{H}. \tag{1.11}$$

*$\{g_k\}, k \in \mathbb{N}$ is also a Riesz basis, and $\{f_k\}, k \in \mathbb{N}$ and $\{g_k\}, k \in \mathbb{N}$ are biorthogonal. Moreover, the series (1.11) converges unconditionally for all $f \in \mathcal{H}$.*

Proof: By definition, we can write $f_k = Ue_k$, where $\{e_k\}, k \in \mathbb{N}$ is an orthonormal basis. Now let $f \in \mathcal{H}$. By expanding $U^{-1}f$ in terms of the orthonormal basis $\{e_k\}, k \in \mathbb{N}$, we have

$$U^{-1}f = \sum_{k=1}^{\infty} \langle U^{-1}f, e_k \rangle e_k = \sum_{k=1}^{\infty} \langle f, (U^{-1})^* e_k \rangle e_k.$$

Therefore, with $g_k = (U^{-1})^* e_k$,

$$f = UU^{-1}f = \sum_{k=1}^{\infty} \langle f, (U^{-1})^* e_k \rangle Ue_k = \sum_{k=1}^{\infty} \langle f, g_k \rangle f_k.$$

Since $(U^{-1})^*$ is bounded and bijective, $\{g_k\}, k \in \mathbb{N}$ is a Riesz basis by definition. For $f \in \mathcal{H}$,

$$\sum_{k=1}^{\infty} |\langle f, f_k \rangle|^2 = \sum_{k=1}^{\infty} |\langle f, Ue_k \rangle|^2 = \|U^* f\|^2 \leq \|U^*\|^2 \|f\|^2. \tag{1.12}$$

This proves that a Riesz basis is a Bessel sequence. Thus, the series converges unconditionally. $\square$

The following theorem shows that Riesz basis not only satisfies the Bessel inequality, it also satisfies some kind of opposite inequality.

**Theorem 1.2.5.** *If $\{f_k\} = \{Ue_k\}, k \in \mathbb{N}$ is a Riesz basis for $\mathcal{H}$, then there exist constants $A, B > 0$ such that*

$$A\|f\|^2 \leq \sum_{k=1}^{\infty} |\langle f, f_k\rangle|^2 \leq B\|f\|^2, \ \forall f \in \mathcal{H}. \tag{1.13}$$

*The largest possible value for the constant $A$ is $\frac{1}{\|U^{-1}\|^2}$, and the smallest possible value for $B$ is $\|U\|^2$.*

Proof: That a Riesz basis $\{Ue_k\}, k \in \mathbb{N}$ is a Bessel sequence with optimal bound $\|U\|$ follows already from the estimate in (1.12). The result about the lower bound follows from

$$\|f\| = \|(U^*)^{-1}U^*f\| \leq \|(U^*)^{-1}\|\|U^*f\| = \|U^{-1}\|\|U^*f\|. \quad \square$$

The last part of this section is devoted to the characterization of Riesz basis. For this purpose, we need a technical result about operators.

**Lemma 1.2.6.** *Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces, and let $\{h_k\}, k \in \mathbb{N}$ be a sequence in $\mathcal{H}$, $\{g_k\}, k \in \mathbb{N}$ a sequence in $\mathcal{K}$. Assume that $\{g_k\}, k \in \mathbb{N}$ is a Bessel sequence with bound $B$, that $\{h_k\}, k \in \mathbb{N}$ is complete in $\mathcal{H}$, and that there exists a constant $A > 0$*

*such that*

$$A \sum_{k=1}^{\infty} |c_k|^2 \le \left\| \sum_{k=1}^{\infty} c_k h_k \right\|^2, \tag{1.14}$$

*for all finite scalar sequences* $\{c_k\}, k \in \mathbb{N}$. *Then*

$$U\Big( \sum_{k=1}^{\infty} c_k h_k \Big) = \sum_{k=1}^{\infty} c_k g_k$$

*defines a linear bounded operator from span*$\{h_k\}$ *into span*$\{g_k\}$, *and* $U$ *has a unique extension to a bounded operator from* $\mathcal{H}$ *into* $\mathcal{K}$*; the norm of* $U$ *as well as its extension is at most* $\sqrt{\frac{B}{A}}$.

Proof: By assumption (1.14), every $h \in span\{h_k\}$ has a unique representation $h = \sum_{k=1}^{\infty} c_k h_k$ with $\{c_k\}$ finite. It follows that $U$ is well defined and linear. Given a finite sequence $\{c_k\}, k \in \mathbb{N}$,

$$\begin{aligned} \left\| U\Big( \sum_{k=1}^{\infty} c_k h_k \Big) \right\|^2 &= \left\| \sum_{k=1}^{\infty} c_k g_k \right\|^2 \\ &\le B \sum_{k=1}^{\infty} |c_k|^2 \le \frac{B}{A} \left\| \sum_{k=1}^{\infty} c_k h_k \right\|^2. \end{aligned}$$

Thus $U$ is bounded. Since $\{h_k\}, k \in \mathbb{N}$ is assumed to be complete in $\mathcal{H}$, $U$ has an extension to a bounded operator on $\mathcal{H}$.

The next theorem gives an equivalent definition of a Riesz basis.

**Theorem 1.2.7.** *The sequence* $\{f_k\}, k \in \mathbb{N}$ *is a Riesz basis for* $\mathcal{H}$ *if and only if* $\{f_k\}, k \in \mathbb{N}$ *is complete in* $\mathcal{H}$ *and there exist constants* $A, B > 0$ *such that for every* *finite scalar sequence* $\{c_k\}$

$$A\sum_{k=1}^{\infty}|c_k|^2 \leq \left\|\sum_{k=1}^{\infty}c_k f_k\right\|^2 \leq B\sum_{k=1}^{\infty}|c_k|^2. \tag{1.15}$$

Proof: The only if part. Assume that $\{f_k\}, k \in \mathbb{N}$ is a Riesz basis, and write it in the form $\{Ue_k\}, k \in \mathbb{N}$. Note that $\{f_k\}, k \in \mathbb{N}$ is compelete. Given any finite scalar sequence $\{c_k\}, k \in \mathbb{N}$,

$$\left\|\sum_{k=1}^{\infty}c_k f_k\right\|^2 = \left\|U\Big(\sum_{k=1}^{\infty}c_k e_k\Big)\right\|^2 \leq \|U\|^2\left\|\sum_{k=1}^{\infty}c_k e_k\right\|^2 = \|U\|^2\sum_{k=1}^{\infty}|c_k|^2,$$

and

$$\left\|\sum_{k=1}^{\infty}c_k e_k\right\|^2 = \left\|U^{-1}U\Big(\sum_{k=1}^{\infty}c_k e_k\Big)\right\|^2 \leq \|U^{-1}\|^2\left\|\sum_{k=1}^{\infty}c_k f_k\right\|^2.$$

Thus we have,

$$\frac{1}{\|U^{-1}\|^2}\sum_{k=1}^{\infty}|c_k|^2 \leq \left\|\sum_{k=1}^{\infty}c_k e_k\right\|^2 \leq \|U\|^2\sum_{k=1}^{\infty}|c_k|^2.$$

For the if part, the right hand side of inequality in (1.15) implies that $\{f_k\}, k \in \mathbb{N}$ is a Bessel sequence with bound $B$. Choose an orthonormal basis $\{e_k\}, k \in \mathbb{N}$ for $\mathcal{H}$ and extend by Lemma 1.2.6 $Ue_k = f_k$ to a bounded operator on $\mathcal{H}$. In the same way, extend $Vf_k = e_k$ to a bounded operator on $\mathcal{H}$. Then $UV = VU = I$, so $U$ is invertible; thus, $\{f_k\}, k \in \mathbb{N}$ is a Riesz basis.

## 1.3 Summary of results

In this subsection, we list the results of the thesis.

The main contribution of Chapter 2 is the proof of the non-harmonic counterpart of the Fourier uniqueness theorem on quasicrystals. We first introduce Meyer's concept of crystalline measures and new versions of the Poisson summation formula on quasicrystals. We give detailed proofs of several technical results, e.g., Lemma 2.1.2, Corollary 2.1.4, that the original paper [27] omitted. We then give a detailed introduction to the theory of Meyer's model sets. We describe the notion of density used by Duffin and Schaeffer, and the density used by Beurling. We show that the former implies the latter in Proposition 2.2.5 and Proposition 2.2.6. Finally, we present Meyer's Poisson summation formula for model sets. Using this result, we prove a uniqueness theorem for discrete measures whose supports are finite dimensional over $\mathbb{Q}$ in Theorem 2.2.10.

In Chapter 3, we first introduce the theory of sampling and interpolation and its relationship with frames and Riesz sequences. We then introduce a result of Beurling [5]. He was the first one to derive a sufficient density condition for exponential frames using the technique of balayage. Finally, we present a theorem by Matei on exact reconstruction of discrete positive measures on $\mathbb{T}^2$. We extend his result by removing the restriction of positiveness and prove a similar result for signed measures in Theorem 3.3.6.

Chapter 4 details the design of a type of single pixel camera. We show how to use a single pixel to reconstruct an image with few measurements. The heart of the technique lies in compressed sensing. We introduce a way to take multiple low resolution pictures and using the intertwine of the information of neighboring pixels. We are able to reconstruct a picture with two times, or three times of the original resolution in each direction. We did several experiments to support our theory. In particular, our experiments show that basis pursuit outperforms orthogonal matching pursuit in this setup.

# Chapter 2: POISSON SUMMATION FORMULAS

In this chapter, we start with the classical Poisson summation formula and its variations. We then state several irregular Poisson summation formulas constructed by several authors. Finally, we present an application of one type of Poisson summation formula to prove a uniqueness theorem for certain types of measures.

**Theorem 2.0.1.** *Let* $\Gamma$ *be a lattice in* $\mathbb{R}^d$*, let* $\Gamma^* = \{y \in \mathbb{R}^d |\ y \cdot x \in \mathbb{Z}\ for\ all\ x \in \Gamma\}$ *be the dual lattice, then for every* $f$ *in the Schwartz class* $\mathcal{S}(\mathbb{R}^d)$

$$vol(\Gamma) \sum_{\gamma \in \Gamma} f(\gamma) = \sum_{\gamma^* \in \Gamma^*} \hat{f}(\gamma^*), \tag{2.1}$$

*or equivalently*

$$\Big(vol(\Gamma) \sum_{\gamma \in \Gamma} \delta_\gamma\Big)\hat{} = \sum_{\gamma^* \in \Gamma^*} \delta_{\gamma^*}. \tag{2.2}$$

Proof: Let $\phi(x) = \sum_{n \in \mathbb{Z}^d} f(x+n)$, then $\phi$ is $\mathbb{Z}^d$ periodic. The decay of $f$ indicates that $\phi \in L^1(\mathbb{R})$. Thus

$$\begin{aligned}
\hat{\phi}(n) &= \int\limits_{\mathbb{T}^d} \phi(x) e^{-2\pi i n x} dx \\
&= \int\limits_{\mathbb{T}^d} \Big( \sum_{k \in \mathbb{Z}^d} f(x+k) e^{-2\pi i n (x+k)} \Big) dx \\
&= \int_{\mathbb{R}^d} f(x) e^{-2\pi i n x} dx = \hat{f}(n).
\end{aligned}$$

Since $\sum_{n \in \mathbb{Z}^d} |\hat{f}(n)| < \infty$, $\phi$ has absolutely convergent Fourier series $\phi(x) = \sum_{n \in \mathbb{Z}^d} \hat{f}(n) e^{2\pi i n x}$. Thus we have

$$\sum_{n \in \mathbb{Z}^d} f(x+n) = \sum_{n \in \mathbb{Z}^d} \hat{f}(n) e^{2\pi i n x}.$$

Now let $\Gamma = A\mathbb{Z}^d$, and $\Gamma^* = (A^{-1})^T \mathbb{Z}^d$ be the dual lattice,

$$\begin{aligned}
\sum_{\gamma \in \Gamma} f(x+\gamma) &= \sum_{n \in \mathbb{Z}^d} (f \circ A)(A^{-1}x + n) \\
&= \sum_{n \in \mathbb{Z}^d} (f \circ A)\hat{}(n) e^{2\pi i n A^{-1} x} \\
&= |det(A)|^{-1} \sum_{n \in \mathbb{Z}^d} \hat{f}((A^{-1})^T n) e^{2\pi i (A^{-1})^T n x} \\
&= vol(\Gamma)^{-1} \sum_{\gamma^* \in \Gamma^*} \hat{f}(\gamma^*) e^{2\pi i \gamma^* x}.
\end{aligned}$$

By setting $x = 0$, we get the desired result. □

The form $\sum_{\gamma \in \Gamma} \delta_\gamma$ is called a Dirac comb. To get variations of the Poisson summation formula, the fundamental things we can do are translation, modulation and finite summation. This leads to the following definition:

**Definition 2.0.1.** A *generalized lattice Dirac comb* (GLDC) is a sum $\mu = \mu_1 + ... + \mu_N$ where $(a)$ $\mu_j = g_j \sigma_j$, $1 \leq j \leq N$ $(b)$ $\sigma_j = \sum_{\gamma \in x_j + \Gamma_j} \delta_\gamma$ is a lattice Dirac comb supported by a coset $x_j + \Gamma_j$ of a lattice $\Gamma_j \subset \mathbb{R}^d$ $(c)$ $g_j(x) = \sum_{k \in F_j} c(j,k) exp(2\pi i \omega_{j,k} x)$ is a finite trigonometric sum.

By applying the Poisson summation formula repeatedly, we can see that the Fourier transform of a GLDC is still a GLDC. However, this generalization is too trivial in the sense that the support of a GLDC is finite dimensional over $\mathbb{Q}$. If $\Lambda$ is the support of a GLDC, then $\Lambda_{\mathbb{Q}} = span_{\mathbb{Q}} \Lambda$ has finite dimension as a $\mathbb{Q}$ vector subspace of $\mathbb{R}^d$. The more interesting question is can we find new types of Poisson summation formulas such that the support of the measures on both sides are infinite dimensional over $\mathbb{Q}$.

## 2.1 Crystalline measures

To further generalize GLDC, Meyer [27] proposed the following definition:

**Definition 2.1.1.** An atomic signed measure $\mu$ on $\mathbb{R}^d$ is a crystalline measure if

$(a)$ the support $\Lambda$ of $\mu$ is a locally finite set (the intersection of $\Lambda$ with any compact set is finite)

$(b)$ $\mu$ is a tempered distribution

$(c)$ the distributional Fourier transform $\widehat{\mu}$ of $\mu$ is also a discrete measure supported by a locally finite set $S$ (the spectrum of $\mu$)

### 2.1.1 Guinand's measure

One of the first examples of such measures was constructed by Guinand [15], and was named by Meyer as the Guinand distribution.

**Theorem 2.1.1.** *The Fourier transform of the one dimensional odd distribution*

$$\sigma = -2\frac{d}{dt}\delta_0 + \sum_{n=1}^{\infty} r_3(n) n^{-1/2}(\delta_{\sqrt{n}} - \delta_{-\sqrt{n}}) \tag{2.3}$$

*is* $\hat{\sigma} = -i\sigma$. *Here,* $r_3(n)$ *is the number of decompositions of the integer* $n \geq 0$ *into a sum of three squares.*

For the proof of Theorem 2.1.1, we'll need the following technical lemma:

**Lemma 2.1.2.** *The collection* $\{f_x(t) = te^{-\pi x t^2} : x > 0\}$ *is total in the space of odd Schwartz functions, which is denoted as* $\mathcal{S}_o(\mathbb{R})$.

Proof: We will avoid the issue of dealing with odd functions by showing a stronger statement:

$$\{f_x(t) = te^{-\pi x t^2},\ g_x(t) = e^{-\pi x t^2} : x > 0\} \text{ is total in } \mathcal{S}(\mathbb{R}).$$

To see this implies Lemma 2.1.2, let $d$ denote the metric on $\mathcal{S}(\mathbb{R})$ induced by the family of semi-norms. Given $f(t) \in \mathcal{S}_o(\mathbb{R})$ and $\epsilon > 0$, there exists a finite set

$\{x_j\}_{j=1}^N$ of positive real numbers such that

$$d\Big(f(t), \sum_{j=1}^{M} a_n e^{-\pi x_j t^2} + \sum_{j=M+1}^{N} b_n t e^{-\pi x_j t^2}\Big) < \epsilon.$$

Replacing $t$ by $-t$ to get

$$d\Big(f(-t), \sum_{j=1}^{M} a_n e^{-\pi x_j t^2} - \sum_{j=M+1}^{N} b_n t e^{-\pi x_j t^2}\Big) < \epsilon.$$

By triangle inequality and oddness of $f$, we have

$$d\Big(f(t), \sum_{j=M+1}^{N} b_n t e^{-\pi x_j t^2}\Big) < \epsilon.$$

Next, by using the dual characterization of closed linear span, [20] chapter 8.2 theorem 8, we only need to show the following: let $T \in \mathcal{S}'(\mathbb{R})$ be a tempered distribution, suppose $T(e^{-\pi x t^2}) = T(te^{-\pi x t^2}) = 0, \forall x > 0$, then $T(f) = 0, \forall f \in \mathcal{S}(\mathbb{R})$. Now by taking $\frac{d}{dx}$, we see that

$$T(e^{-\pi x t^2}) = T(te^{-\pi x t^2}) = T(t^2 e^{-\pi x t^2}) = T(t^3 e^{-\pi x t^2}) = ... = 0.$$

By linearity, $T(p(t)e^{-\pi x t^2}) = 0$ for any polynomial $p(t)$. Since $\mathcal{C}_c^\infty(\mathbb{R})$ is dense (with respect to the *weak* $*$ topology $\sigma(\mathcal{S}', \mathcal{S})$) in $\mathcal{S}'(\mathbb{R})$, we can take a sequence $\{T_n(t)\}_{n=1}^\infty$ of $\mathcal{C}_c^\infty(\mathbb{R})$ functions such that

$$T_n \to T, \quad weak\ *.$$

Let $\epsilon > 0$, then for $f \in \mathcal{S}(\mathbb{R})$, we have:

$$\langle T, f \rangle = \langle T - T_n, f \rangle + \langle T_n(t), f(t) - p(t)e^{-\pi x t^2} \rangle + \langle T_n, p(t)e^{-\pi x t^2} \rangle,$$

where $p(t)$ is a polynomial left to be specified. Observe that both the first and third term on the right hand side of the above identity can be made less than$\epsilon$ if we choose a sufficiently large $n$, say $n'$. For such an $n'$, we will choose $p(t)$ so that the second term is also less than $\epsilon$. Indeed,

$$|\langle T_{n'}(t), f(t) - p(t)e^{-\pi x t^2} \rangle| < \int_{-A}^{A} |T_{n'}(t)\Big(\frac{f(t)}{e^{-\pi x t^2}} - p(t)\Big)e^{-\pi x t^2}|dt,$$

where $A$ is such that the support of $T_{n'}$ is contained in $[-A, A]$.

Then, on $[-A, A]$, by Weierstrass approximation theorem, choose a polynomial $p(t)$ such that $||\frac{f(t)}{e^{-\pi x t^2}} - p(t)||_{L^\infty_{[-A,A]}} < C\epsilon$, where $C = \int_{-A}^{A} |T_{n'}(t)e^{-\pi x t^2}|dt$, then

$$\int_{-A}^{A} |T_{n'}(t)\Big(\frac{f(t)}{e^{-\pi x t^2}} - p(t)\Big)e^{-\pi x t^2}|dt < \epsilon.$$

Finally, combining the above results, we have $T(f) = 0$ for all $f \in \mathcal{S}(\mathbb{R})$. Thus Lemma 2.1.2 is proved. □

Now we are ready to prove Theorem 2.1.1:

Proof of Theorem 2.1.1: for $x > 0$, raise the following PSF

$$\sum_{k \in \mathbb{Z}} exp(-\pi k^2 x) = x^{-1/2} \sum_{k \in \mathbb{Z}} exp(-\pi k^2 / x) \tag{2.4}$$

to cubic power yields

$$1+\sum_{n=1}^{\infty} r_3(n)exp(-\pi nx) = x^{-3/2} + x^{-3/2}\sum_{n=1}^{\infty} r_3(n)exp(-\pi n/x). \tag{2.5}$$

Let $f_x(t) = texp(-\pi xt^2), t \in \mathbb{R}, x > 0$. Then $f_x$ is odd and its Fourier transform is $\hat{f}_x(\gamma) = -ix^{-3/2}\gamma exp(-\pi\gamma^2/x)$, then (2.5) can be written as

$$\langle \sigma, f_x\rangle = i\langle \sigma, \hat{f}_x\rangle. \tag{2.6}$$

The collection $\{f_x(t) = texp(-\pi xt^2), x > 0\}$ is total (linear span being dense) in the subspace of odd Schwartz class by Lemma 2.1.2. Thus (2.6) implies $\langle \sigma, \phi\rangle = i\langle \hat{\sigma}, \phi\rangle$ holds for every odd Schwartz function $\phi$. However, since $\sigma$ itself is odd, thus $\langle \sigma, \phi\rangle = i\langle \hat{\sigma}, \phi\rangle$ is automatically true for even Schwartz functions $\phi$. Thus, (2.6) is true for all Schwartz functions, since every Schwartz function can be written as a sum of an odd one and an even one. Thus we have $\hat{\sigma} = -i\sigma$.$\square$

The Guinand distribution $\sigma$ is not a measure due to the first term $\frac{d}{dt}\delta_0$. However a simple modification yields a crystalline measure [27]:

**Theorem 2.1.3.** *Let $\alpha \in (0,1)$, the Fourier transform of the crystalline measure*

$$\tau_\alpha = \left(\alpha^2 + \frac{1}{\alpha}\right)\sigma(t) - \alpha\sigma(\alpha t) - \sigma(t/\alpha) \tag{2.7}$$

*is $\hat{\tau}_\alpha = -i\tau_\alpha$*

Proof: In order to show that the derivative of delta disappears from the linear combination, it suffices to show

$$\frac{d}{dt}\delta_0(\alpha t) = \frac{1}{\alpha^2}\frac{d}{dt}\delta_0(t). \tag{2.8}$$

Using the fact that $\widehat{\delta'} = 2\pi i t$, we have for every $f \in \mathcal{S}(\mathbb{R})$:

$$\begin{aligned}
&\langle \delta'(\alpha t), \hat{f}(t)\rangle = \langle \widehat{\delta'(\alpha t)}, f(\gamma)\rangle = \langle \frac{1}{\alpha}\hat{\delta}'(\frac{t}{\alpha}), f(\gamma)\rangle \\
&= \langle \frac{1}{\alpha^2}2\pi i t, f(\gamma)\rangle = \langle \frac{1}{\alpha^2}\widehat{\delta'(t)}, f(\gamma)\rangle = \langle \frac{1}{\alpha^2}\delta'(t), \hat{f}(t)\rangle.
\end{aligned}$$

The fact that $\tau_\alpha$ is a tempered distribution follows from the arithmetic property of the sum of $r_3(n)$ [7]:

$$\sum_{0\le n\le x} r_3(n) = \frac{4}{3}\pi x^{3/2} + \mathcal{O}(x^{3/4+\epsilon}), \quad x \to \infty. \tag{2.9}$$

Finally, the Fourier transform of $\tau_\alpha$ is given by

$$\hat{\tau}(\gamma) = (\alpha^2 + \frac{1}{\alpha})\hat{\sigma}(\gamma) - \hat{\sigma}(\frac{\gamma}{\alpha}) - \alpha\hat{\sigma}(\alpha\gamma) = -i\tau_\alpha \square$$

Let's take an look at what this measure looks like when $\alpha = \frac{1}{2}$.

**Corollary 2.1.4.** *The Fourier transform of the measure*

$$\tau = \sum_{n=1}^{\infty} \chi(n) r_3(n) n^{-1/2}\big(\delta_{\sqrt{n}/2} - \delta_{-\sqrt{n}/2}\big)$$

*is $-i\tau$, where $\chi(n) = -1/2$, if $n \in \mathbb{N}\backslash 4\mathbb{N}$, $\chi(n) = 4$, if $n \in 4\mathbb{N}\backslash 16\mathbb{N}$, and $\chi(n) = 0$,* *if* $n \in 16\mathbb{N}$

Proof: It suffices to verify that $\tau = \tau_{\frac{1}{2}}$. Since the derivative of Dirac at 0 was canceled in the expression of $\tau_{\frac{1}{2}}$, we have

$$\begin{aligned}
\tau_{\frac{1}{2}} &= (\frac{1}{4}+2)\sum_{n=1}^{\infty} r_3(n)n^{-1/2}\Big(\delta_{\sqrt{n}}(t) - \delta_{-\sqrt{n}}(t)\Big) \\
&- \frac{1}{2}\sum_{n=1}^{\infty} r_3(n)n^{-1/2}\Big(\delta_{\sqrt{n}}(t/2) - \delta_{-\sqrt{n}}(t/2)\Big) \\
&- \sum_{n=1}^{\infty} r_3(n)n^{-1/2}\Big(\delta_{\sqrt{n}}(2t) - \delta_{-\sqrt{n}}(2t)\Big) \\
&= (\frac{1}{4}+2)\sum_{n=1}^{\infty} r_3(n)n^{-1/2}\Big(\delta_{\sqrt{n}}(t) - \delta_{-\sqrt{n}}(t)\Big) \\
&- \sum_{n=1}^{\infty} r_3(n)n^{-1/2}\Big(\delta_{2\sqrt{n}}(t) - \delta_{-2\sqrt{n}}(t)\Big) \\
&- \frac{1}{2}\sum_{n=1}^{\infty} r_3(n)n^{-1/2}\Big(\delta_{\sqrt{n}/2}(t) - \delta_{-\sqrt{n}/2}(t)\Big) \\
&= (\frac{1}{4}+2)\sum_{n=1}^{\infty} r_3(n)n^{-1/2}\Big(\delta_{\sqrt{4n}/2}(t) - \delta_{-\sqrt{4n}/2}(t)\Big) \\
&- \sum_{n=1}^{\infty} r_3(n)n^{-1/2}\Big(\delta_{\sqrt{16n}/2}(t) - \delta_{-\sqrt{16n}/2}(t)\Big) \\
&- \frac{1}{2}\sum_{n=1}^{\infty} r_3(n)n^{-1/2}\Big(\delta_{\sqrt{n}/2}(t) - \delta_{-\sqrt{n}/2}(t)\Big) \\
&= (\frac{1}{2}+4)\sum_{n\in 4\mathbb{N}} r_3(n)n^{-1/2}\Big(\delta_{\sqrt{n}/2}(t) - \delta_{-\sqrt{n}/2}(t)\Big) \\
&- 4\sum_{n\in 16\mathbb{N}} r_3(n)n^{-1/2}\Big(\delta_{\sqrt{n}/2}(t) - \delta_{-\sqrt{n}/2}(t)\Big) \\
&- \frac{1}{2}\sum_{n\in \mathbb{N}} r_3(n)n^{-1/2}\Big(\delta_{\sqrt{n}/2}(t) - \delta_{-\sqrt{n}/2}(t)\Big) = \tau. \square
\end{aligned}$$

### 2.1.2 Kolountzakis's Measure

Inspired by Meyer's work, Kolountzakis gave an example of a measure whose support and spectrum are both not contained in any finite union of arithmetic progressions. A measure $\mu$ is said to be translation bounded if the set $\{|\mu|(K+x) : x \in \mathbb{R}^d\}$ is bounded for each compact set $K \subset \mathbb{R}^d$.A translation bounded measure will always be a tempered distribution.

**Theorem 2.1.5.** *There is a translation bounded measure $\nu$ of the form $\nu = \sum_{\lambda\in\Lambda} c_\lambda \delta_\lambda$, ($c_\lambda \neq 0$) such that $\Lambda \subset \mathbb{R}$ is a locally finite set and such that $\hat{\nu}$ is also a translation bounded measure of the form $\hat{\nu} = \sum_{s\in S} d_s \delta_s$, ($d_s \neq 0$) where $S$ is also a locally finite set and such that both $\Lambda$ and $S$ are not contained in finite unions of arithmetic progressions. Therefore this Fourier pair cannot be derived by finitely many applications of the PSF.*

We need the following two lemmas:

**Lemma 2.1.6.** *There is a function $f\colon \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$, not identically zero, such that both the function and its Fourier transform $\hat{f}\colon \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ vanish in the interval*

$$I = \left\{x \in \mathbb{Z}/N\mathbb{Z} : |x| \leq \tfrac{N}{10}\right\}.$$

Proof: We search for $f\colon \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ which is 0 on $I$ such that $\hat{f}$ also vanishes on $I$. This is a homogeneous linear system (the unknowns are the values of $f$ off $I$) with more unknowns ($\sim 4N/5$ of them) then equations ($\sim N/5$ of them) so there is a non-zero solution. $\square$

**Lemma 2.1.7.** *Suppose $M > 1$ is an integer. Then there is a non-zero measure $\mu$ of the form $\mu = \sum_{n\in\mathbb{Z}} a_n\delta_{An}$, whose Fourier transform is a measure $\hat{\mu}$ of the form $\hat{\mu} = \sum_{n\in\mathbb{Z}} b_n\delta_{Bn}$, where $A$, $B$ are positive real numbers, and such that both $\mu$ and $\hat{\mu}$ are $0$ in the interval $(-M, M)$. Furthermore the measures $\mu$ and $\hat{\mu}$ can be taken to be periodic and the numbers $A$ and $B$ may be chosen to be rational.*

Proof: Let us start with the function $f$: $\mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ of Lemma (2.1.6), with $N = 100M^2$. Define first the measure $\tau = \sum_{n\in\mathbb{Z}} f(nmodN)\delta_n$. Then the measure $\tau$ is $N$-periodic whose Fourier transform given by $\hat{\tau} = \sum_{n\in\mathbb{Z}} \hat{f}(nmodN)\delta_{n/N}$, where $\hat{f}(n) = \frac{1}{N}\sum_{k=0}^{N-1} f(k)e^{-2\pi ink/N}$ is the discrete Fourier transform. To see this, write

$$\tau = \sum_{n\in\mathbb{Z}} f(nmodN)\delta_n = \sum_{k=0}^{N-1}\sum_{n\in\mathbb{Z}} f(k)\delta_{nN+k}.$$

Thus by PSF, we have

$$\hat{\tau} = \sum_{k=0}^{N-1} f(k)\big(\sum_{n\in\mathbb{Z}}\delta_{nN+k}\big)^{\wedge} = \sum_{k=0}^{N-1} f(k)\tfrac{1}{N}e^{-2\pi ink/N}\sum_{n\in\mathbb{Z}}\delta_{\frac{n}{N}} = \sum_{n\in\mathbb{Z}}\hat{f}(nmodN)\delta_{\frac{n}{N}}.$$

It follows that $\tau$ vanishes in the interval $(-\frac{N}{10}, \frac{N}{10})$ and $\hat{\tau}$ vanishes in the interval $(-\frac{1}{10}, \frac{1}{10})$. $\square$

Now we are ready to prove Theorem 2.1.5:

Proof of Theorem 2.1.5: Take a sequence $M_n \to \infty$ and apply repeatedly Lemma 3 to obtain a sequence of periodic measures $\mu_n$ of discrete support, having also $\widehat{\mu_n}$ periodic and of discrete support and such that both $\mu_n$ and $\widehat{\mu_n}$ vanish in the interval $(-M_n, M_n)$.

Denote by $T_r$ the translation by $r$ and by $M_a$ the modulation operator by $a$. Let $\epsilon_n \to 0$ be a $\mathbb{Q}$-linearly independent sequence. Each measure $\mu_n$ or $\widehat{\mu_n}$ has bounded total variation in any interval of unit length (since they are periodic), say by $V_n$. Define $D_n = V_n n^2$. Consider the measure

$$\nu = \sum_{n\geq 1} \tfrac{1}{D_n} M_{\epsilon_n} T_{\epsilon_n} \mu_n,$$

whose Fourier transform is the measure

$$\hat{\nu} = \sum_{n\geq 1} \tfrac{1}{D_n} T_{\epsilon_n} M_{-\epsilon_n} \widehat{\mu_n}.$$

It follows that $\nu$ and $\hat{\nu}$ have bounded total variation in any interval of unit length. We need to show that the support of both $\mu$ and $\hat{\mu}$ are locally finite. Let $J = (a,b)$ be any interval. Then there is an index $n_0$ such that for $n \geq n_0$ we have $(a,b) \subset (-M_n+1, M_n-1)$, therefore the support of $\nu$ or $\hat{\nu}$ in $J$ comes only from the contributions of the measures $\mu_1$, $\mu_2$,..., $\mu_{n_0}$ or $\widehat{\mu_1}$, $\widehat{\mu_2}$,..., $\widehat{\mu_{n_0}}$ and consists therefore of a finite number of points. Hence both *supp*$\nu$ and *supp*$\hat{\nu}$ are locally finite. The fact that both *supp*$\nu$ and *supp*$\hat{\nu}$ are infinite dimensional over $\mathbb{Q}$ follows from our choice of the numbers $\epsilon_n$. □

### 2.1.3 Meyer's Measure

The following construction was due to Meyer [27]:

**Theorem 2.1.8.** *Let $\alpha$, $\beta \in \mathbb{R}^3 \backslash \mathbb{Z}^3$. Then the Fourier transform of the measure*

$$\sigma_{(\alpha,\beta)} = \sum_{k\in\mathbb{Z}^3} \frac{exp(2\pi i k\beta)}{\|k+\alpha\|} \big(\delta_{\|k+\alpha\|} - \delta_{-\|k+\alpha\|}\big) \sigma_{(\alpha,\beta)} \tag{2.10}$$

*is*

$$\mathcal{F}\big(\sigma_{(\alpha,\beta)}\big) = -iexp(-2\pi i\alpha\beta)\sigma_{(\beta,\alpha)}.$$

To prove Theorem 2.1.8, we prove a more general result:

**Theorem 2.1.9.** *Let $\mu$ be a crystalline measure on $\mathbb{R}^3$. We then have $\mu = \sum_{\lambda\in\Lambda} a(\lambda)\delta_\lambda$ and $\hat{\mu} = \sum_{y\in S} b(y)\delta_y$, where $S$ is the spectrum of $\mu$. Let us assume that $0 \notin \Lambda$, $0 \notin S$, and consider the one dimensional measure*

$$\sigma_\Lambda = \sum_{\lambda\in\Lambda} \frac{a(\lambda)}{\|\lambda\|}(\delta_{\|\lambda\|} - \delta_{-\|\lambda\|}).$$

*Then $\sigma_\Lambda$ is a crystalline measure and the distributional Fourier transform of $\sigma_\Lambda$ is $-i\sigma_S$, where $\sigma_S = \sum\limits_{y\in S} \frac{b(y)}{\|y\|}(\delta_{\|y\|} - \delta_{-\|y\|})$.*

Proof: Since the measure $\sigma_\Lambda$ and $\sigma_S$ are odd measures, it suffices to prove

$$\langle \sigma_\Lambda, \hat{\phi}\rangle = -i\langle \sigma_S, \phi\rangle \tag{2.11}$$

for every odd Schwartz function $\phi$. Let $\omega = \hat{\phi}$ be the 1D Fourier transform of $\phi$. Then $\omega$ is also and odd Schwartz function and the left hand side of (2.11) is

$$s(\omega) = 2\sum_{\lambda\in\Lambda} a(\lambda)\frac{\omega(\|\lambda\|)}{\|\lambda\|}.$$

Introduce the radial function $\Phi(x) = \omega(\|x\|)/\|x\|$, which belongs to $\mathcal{S}(\mathbb{R}^3)$. Then

$$s(\omega) = 2\sum_{\lambda\in\Lambda} a(\lambda)\Phi(\lambda).$$

We have for every test function $F$,

$$\sum_{\lambda\in\Lambda} a(\lambda)F(\lambda) = \langle \mu, \hat{F}\rangle = \langle \hat{\mu}, F\rangle = \sum_{y\in S} b(y)F(y).$$

**Lemma 2.1.10.** *The 3D Fourier transform of the radial function $F(y) = -i\phi(\|y\|)/\|y\|$ is $\hat{F}(x) = \Phi(x)$.*

Indeed, the 3D Fourier transform of a radial function $\hat{F} \in L^1(\mathbb{R}^3)$ is given by:

$$\begin{aligned}
\hat{F}(x) &= \int\int\int F(y)e^{-2\pi ixy}dy \\
&= \int_0^\infty \int_0^\pi \int_0^{2\pi} F(r)e^{-2\pi i\|x\|rcos\phi}r^2 sin\phi d\theta d\phi dr \\
&= 2\pi \int_0^\infty \int_{-1}^1 F(r)e^{-2\pi i\|x\|ru}r^2 dudr \\
&= 4\pi \int_0^\infty F(r)\frac{sin(2\pi\|x\|r)}{2\pi\|x\|r}dr.
\end{aligned}$$

Applying this to $F(y) = -i\phi(\|y\|)/\|y\|$, we have

$$\hat{F}(x) = \frac{-2i}{\|x\|}\int_0^\infty \phi(r)sin(2\pi\|x\|r)dr.$$

On the other hand, sine $\phi$ is odd, the 1D Fourier transform of $\phi$ is

$$\hat{\phi}(\|x\|) = \int_{-\infty}^\infty \phi(r)e^{-2\pi ir\|x\|}dr = -2i\int_0^\infty \phi(r)sin(2\|x\|)dr.$$

Thus, we have

$$\hat{F}(x)\frac{\hat{\phi}(\|x\|)}{\|x\|} = \frac{\omega(\|x\|)}{\|x\|}.$$

Finally, we have

$$\begin{aligned} s(\omega) &= 2\sum_{\lambda\in\Lambda} a(\lambda)\Phi(\Lambda) = 2\sum_{\lambda\in\Lambda} \hat{F}(\lambda) \\ &= 2\sum_{y\in S} b(y)F(s) = -2i\sum_{y\in S} \frac{\phi(\|y\|)}{\|y\|} = -i\langle \sigma_S, \phi\rangle. \quad \square \end{aligned}$$

To see Theorem 2.1.9 implies Theorem 2.1.8, we only need to observe that the Fourier transform of the measure $\mu = \sum_{k\in\mathbb{Z}^3} exp(2\pi ik\beta)\delta_{k+\alpha}$ is the measure $\hat{\mu} = exp(-2\pi i\alpha\beta)\sum_{k\in\mathbb{Z}^3} exp(-2\pi ik\alpha)\delta_{k+\beta}$. Indeed we only need to apply the Poisson summation formula (2.1) to the function $g(\gamma) = e^{2\pi i\beta(\gamma+\alpha)}\hat{f}(\gamma+\alpha)$.$\square$

## 2.2 Model set

Although we will mostly focus on model sets with Euclidean internal space, it is more natural to introduce model sets in the general setting.

### 2.2.1 Construction and Examples

Let $D \subset \mathbb{R}^n \times H$ be an oblique lattice, where $H$ is a locally compact abelian group. This means that the two natural projections $p_1 : \mathbb{R}^n \times H \to \mathbb{R}^n$, $p_2 : \mathbb{R}^n \times H \to H$, once restricted to $D$, are injective with dense images.

**Definition 2.2.1.** A subset $\Lambda$ of the real space $\mathbb{R}^n$ is a model set if there is a locally compact abelian group $H$ and a relatively compact window set $\Omega$ of $H$ with non-empty interior such that

$$\Lambda(\Omega) = \{\lambda = p_1(d) : d \in D, p_2(d) \in \Omega\}. \tag{2.12}$$

Let $M = p_1(D)$ and denote by $*$ the mapping $p_2 \circ (p_1|_D)^{-1}$, we have

$$* : \; M \to H.$$

Here are some examples of model sets:

(1)The Fibonacci sequence is a 1d model set.

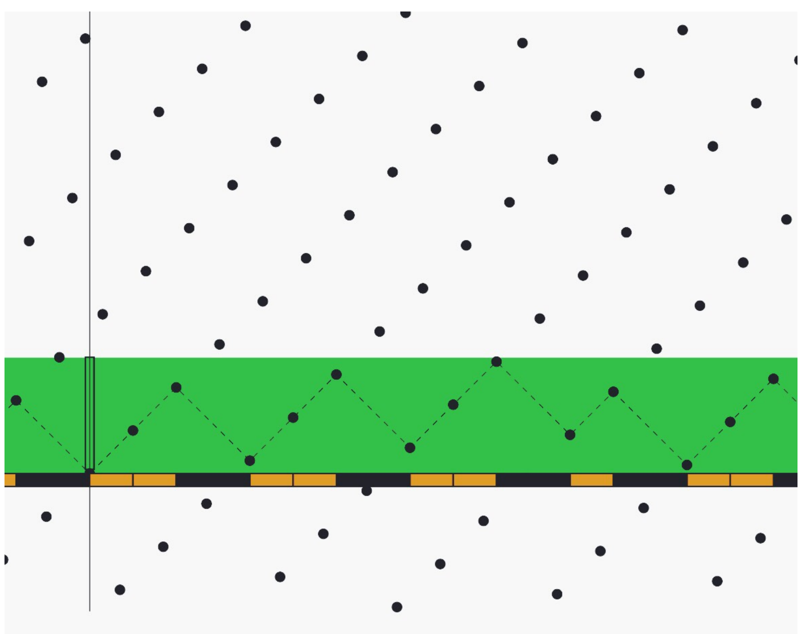


Figure 2.1: The Fibonacci sequence

We begin with the ring $\mathbb{Z}[\tau] = \mathbb{Z} + \mathbb{Z}\tau \subset \mathbb{R}$, where $\tau = (1+\sqrt{5})/2$. This is the ring of integers of the quadratic field $\mathbb{Q}[\tau] = \mathbb{Q}[\sqrt{5}]$. Let $'$ denote the automorphism of $\mathbb{Q}[\tau]$ that maps $\sqrt{5} \to -\sqrt{5}$. The set

$$D = \{(x, x') \mid x \in \mathbb{Z}[\tau]\}$$

is a lattice in $\mathbb{R} \times \mathbb{R}$, and the Fibonacci sets are the sets of the form

$$\{x \in \mathbb{Z}[\tau] \mid x' \in I, \}$$

where $I$ is some open (or closed) interval in $\mathbb{R}$.

(2)The Penrose tiling is a 2d model set.

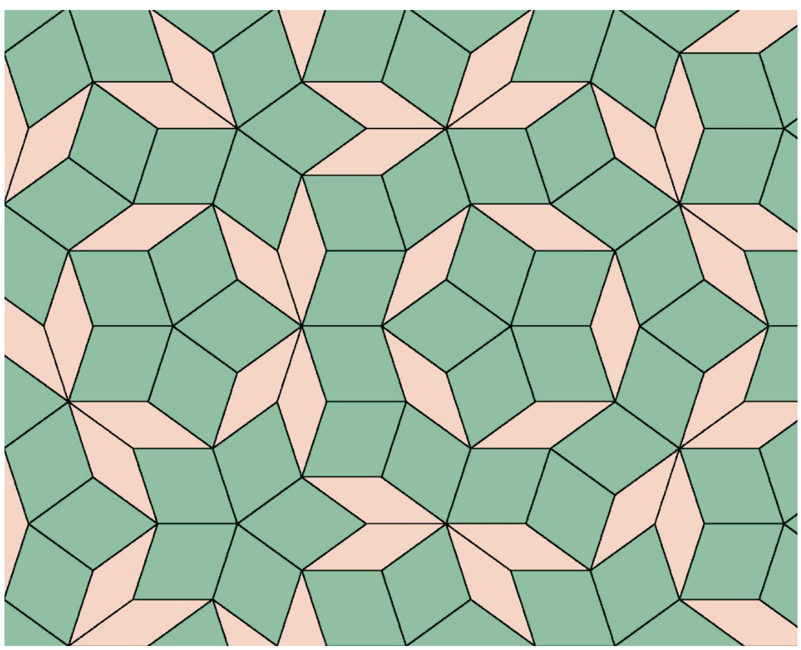


Figure 2.2: The Penrose tiling

Let $\zeta = e^{2\pi i/5} \in \mathbb{C}$. The ring $\mathbb{Z}[\zeta]$ generated by $\zeta$ over $\mathbb{Z}$ is the ring of integers of the number field $\mathbb{Q}[\zeta]$. Let $'$ denote the automorphism of $\mathbb{Q}[\zeta]$ that maps $\zeta$ into

$\zeta^2$. The mapping $\nu : \mathbb{Q}[\zeta] \to \mathbb{Z}/5\mathbb{Z}$ defined by

$$\sum_{i=0}^{4} a_i \zeta^i \mapsto \sum a_i \ (mod\ 5)$$

is a ring homomorphism and for all $x \in \mathbb{Z}[\zeta]$, $\nu(x) = \nu(x')$. Define

$$^* : \mathbb{Z}[\zeta] \to \mathbb{Z}[\zeta] \times (\mathbb{Z}/5\mathbb{Z}) \subset \mathbb{C} \times (\mathbb{Z}/5\mathbb{Z})$$

by $x \mapsto (x', \nu(x))$. We obtain a cut and project scheme from $\mathbb{C} \times (\mathbb{Z}/5\mathbb{Z})$ and the lattice $D = \{(x, x^*) \mid x \in \mathbb{Z}[\zeta]\} \subset \mathbb{C} \times \mathbb{C} \times (\mathbb{Z}/5\mathbb{Z})$.

Let $P_1$ be the convex hull (a closed pentagon) of the fifth roots of 1 in $\mathbb{C}$ and let

$$\Omega = (P_1 \times \{\bar{1}\}) \cup (-\tau P_1 \times \{\bar{2}\}) \cup (\tau P_1 \times \{\bar{3}\}) \cup (-P_1 \times \{\bar{4}\}),$$

where $\bar{j}$ is the congruence class of $j$ modulo 5.

Let $\gamma \in \mathbb{C}$. Then the sets

$$\Lambda(\Omega, \gamma) = \{x \in \mathbb{Z}[\zeta] \mid x^* \in \Omega + \gamma\}$$

are the vertex sets of the Penrose tilings (parametrized by $\gamma$). This observation was constructed by de Bruijn [9].

## 2.2.2 Properties of model sets

Below we introduce some basic properties of model sets. Most of these properties are also shared by lattices. If we want to generalize harmonic analysis to the non-harmonic setting, we should pay our attention to these essential properties shared by lattices and model sets.

**Definition 2.2.2.** Let $G$ be a locally compact abelian group.

(1) A subset $\Lambda$ of $G$ is relatively dense if there is a compact subset $K$ of $G$ so that $G = \Lambda + K$.

(2) A subset $\Lambda$ of $G$ is uniformly discrete if there is an open neighbourhood $U$ of 0 in $G$, so that $(\Lambda - \Lambda) \cap U = \{0\}$.

(3) A subset $\Lambda$ of $G$ is a Delone set if $\Lambda$ is both relatively dense and uniformly discrete.

**Definition 2.2.3.** A subgroup $D$ of a locally compact abelian group $G$ is a lattice if either of the following two equivalent conditions is satisfied:

(1) $D$ is discrete and $G/D$ is compact.

(2) $D$ is Delone.

**Lemma 2.2.1.** *Let $(\mathbb{R}^n \times H, D)$ be a cut and project scheme. Let $U \subset H$ be a non-empty open set. Then there is a compact set $K$ in $\mathbb{R}^n$ so that $\mathbb{R}^n \times H = D + (K \times U)$.*

Proof: Since $D$ is a lattice there is a compact set $C$ in $\mathbb{R}^n \times H$ so that $D + C = \mathbb{R}^n \times H$. The projections of $C$ determine compact sets $K_1$ and $K_2$ so that $\mathbb{R}^n \times H = D + (K_1 \times K_2)$.

Since $p_2(D)$ is dense in $H$. $\cup_{d\in D}(p_2(d)+U) = H$, and by compactness there is a finite set $F$ in $D$ so that $\cup_{f\in F}(p_2(f)+U) \supset K_2$. Now $K = K_1 - p_1(F)$ is compact in $\mathbb{R}^n$. For any $x \in \mathbb{R}^n \times H$ there is a $d \in D$ so that $x - d \in K_1 \times K_2$, and then an $f \in F$, so that $p_2(x-d-f) \in U$. Then

$$p_1(x-d-f) \in K_1 - p_1(F) = K,$$

and

$$x = d + f + (x-d-f) \in D + (K \times U). \quad \square$$

The next theorem shows that model set is essentially similar to a lattice, except for not being periodic.

**Proposition 2.2.2.** *A model set is Delone.*

Proof: By Lemma 2.2.1, there is a compact set $K$ in $\mathbb{R}^n$ so that $\mathbb{R}^n \times H = D + K \times (-\Omega)$. For $x \in \mathbb{R}^n$,

$$(x, 0) = (d, d^*) + (k, -\omega)$$

for some $d \in M$, $k \in K$, $\omega \in \Omega$. Then $d^* = \omega \in \Omega$ gives $d \in \Lambda$, and $x = d+k \in \Lambda+K$. Thus $\mathbb{R}^n = \Lambda + K$ and $\Lambda$ is relatively dense.

For each $r > 0$, $K_r \bar{B}_r \times (\bar{\Omega} - \bar{\Omega})$ is compact in $\mathbb{R}^n \times H$. For small enough $r$, $K_r \cap D = \{0\}$, for otherwise, 0 would have been a limit point for $D$. For such an $r$, if $x, y \in \Lambda$ and $|x - y| \leq r$, then $(x - y, x^* - y^*) \in K_r$, so $x = y$. Thus

$(\Lambda - \Lambda) \cap B_r = \{0\}$ thus $\Lambda$ is uniformly discrete.

Among the many nice properties of model sets, we find the following one most interesting.

**Definition 2.2.4.** Let $\Lambda \in \mathbb{R}^n$ be any subsets and let $[\Lambda]$ be the subgroup of $\mathbb{R}^n$ generated by $\Lambda$ under $+$. A character on $\Lambda$ is a mapping $\chi : \Lambda \to \mathbb{T}$ that is the restriction of a character $\chi : [\Lambda] \to \mathbb{T}$.

**Definition 2.2.5.** A subset $\Lambda \subset \mathbb{R}^n$ is harmonious if for each character $\chi_0 : \Lambda \to \mathbb{T}$ and each $\epsilon > 0$ there is a continuous character $\chi : \mathbb{R}^n \to \mathbb{T}$ such that $\chi$ is an $\epsilon-$uniform approximation of $\chi_0$. That is for all $x \in \Lambda$, $|\chi(x) - \chi_0(x)| < \epsilon$.

Observe from the definition, we can easily see that any finite set is harmonious. Also any subset of a harmonious set is harmonious. Below is a trivial example of a harmonious set in $\mathbb{Z}^2$.

**Example 2.2.1.** $\mathbb{Z}^2 \subset \mathbb{R}^2$ is harmonious. Indeed, $[\mathbb{Z}^2] = \mathbb{Z}^2$ and $\chi_0 \in Hom(\mathbb{Z}^2, \mathbb{T})$ is of the form

$$\chi_0(m,n) = \chi_0(1,0)^m \chi_0(0,1)^n = e^{2\pi i(m\alpha + n\beta)},$$

where we designate $\chi_0(1,0) = e^{2\pi i\alpha}$ and $\chi_0(0,1) = e^{2\pi i\beta}$. Thus we can simply define $\chi(x,y) = e^{2\pi i(x\alpha + y\beta)}$ to get the approximation property.

An amazing fact is that model sets are harmonious. This was proved by Meyer [25].

**Proposition 2.2.3.** *Let $\Lambda$ be a model set in $\mathbb{R}^n$. Then $\Lambda$ is harmonious.*

Proof: Using the inclusion mapping $[\Lambda] \hookrightarrow D$ and the lifting property of group characters we obtain successively $\chi_D$ and $\chi_c$ (continuous) making the following commutative diagram

$$\begin{tikzcd}
\mathbb{R}^n \times H \arrow[r, "\chi_c"] & \mathbb{T} \arrow[d, equal] \\
D \arrow[u] \arrow[r, "\chi_D"] & \mathbb{T} \arrow[d, equal] \\
\bar{\Lambda} \arrow[u, leftrightarrow] \arrow[r, "\chi_0"] & \mathbb{T}
\end{tikzcd}$$

where $\bar{\Lambda}$ is the group generated by $\Lambda$. Let $Z$ be the annihilator of $D$ in $(\mathbb{R}^n \times H)\hat{}$. Choose $\epsilon > 0$ and take the open neighbourhood

$$N = N(\bar{\Omega}, \epsilon) = \{\psi \in H \mid |\psi(x) - 1| < \epsilon \; on \; \bar{\Omega}.\}$$

Using Lemma2.2.1 it follows that $\widehat{\mathbb{R}^n \times H} = Z + (\hat{\mathbb{R}^n} \times N)$. In particular, we can write

$$\chi_c = \chi_Z + (\chi, \chi_N)$$

in the obvious notation. We show that $\chi \in \hat{\mathbb{R}^n}$ is the desired $\epsilon$-approximation to $\chi_0$. For all $x \in \Lambda \in [\Lambda]$:

$$\begin{aligned}
\chi_0(x) &= \chi_D(x, x^*) = \chi_c(x, x^*) \\
&= \chi_Z(x, x^*) \cdot \chi(x)\chi_N(x^*) = \chi(x)\chi_N(x^*),
\end{aligned}$$

since $\chi_Z \in Z$ annihilates $D$. Since $x^* \in \Omega \in \bar{\Omega}$,

$$|\chi(x) - \chi_0(x)| = |\chi(x) - \chi(x)\chi_N(x^*)|$$

$$= |1 - \chi_N(x^*)| < \epsilon. \quad \square$$

Finally we point out the fact that any harmonious set is a subset of a certain model set, see Meyer [25]. Unions of harmonious sets need not be harmonious. Also unions of model sets need not be model sets. In fact, unions of model sets need not even be harmonious. Before we give counterexamples, we need the following lemma.

**Lemma 2.2.4.** *If $\Lambda$ is harmonious, then so is $\Lambda - \Lambda$.*

Proof: Given $\epsilon > 0$, for any weak character $\chi$ on $\Lambda$, there is a strong character $h$ on $\Lambda$ such that

$$\sup_{\lambda \in \Lambda} |\chi(\lambda) - h(\lambda)| \leq \epsilon.$$

Now

$$\sup_{\lambda_1,\lambda_2} |\chi(\lambda_1 - \lambda_2) - h(\lambda_1 - \lambda_2)| = \sup_{\lambda_1,\lambda_2} |\chi(\lambda_1)\chi(\lambda_2)^{-1} - h(\lambda_1)h(\lambda_2)^{-1}|$$

$$\leq \sup_{\lambda_1,\lambda_2} |\chi(\lambda_1)\chi(\lambda_2)^{-1} - h(\lambda_1)\chi(\lambda_2)^{-1}| + \sup_{\lambda_1,\lambda_2} |h(\lambda_1)\chi(\lambda_2)^{-1} - h(\lambda_1)h(\lambda_2)^{-1}|$$

$$= \sup_{\lambda_1} |\chi(\lambda_1) - h(\lambda_1)| + \sup_{\lambda_2} |\chi(\lambda_2)^{-1} - h(\lambda_2)^{-1}| \leq 2\epsilon. \quad \square$$

Now we are ready to give counter examples. Given $\Lambda_1$ and $\Lambda_2$ both harmonious, if $\Lambda_1 \cup \Lambda 2$ is also harmonious, then $(\Lambda_1 \cup \Lambda 2) - (\Lambda_1 \cup \Lambda 2)$ would also be harmonious by the above theorem. Then $\Lambda_1 - \Lambda_2$ as a subset of $(\Lambda_1 \cup \Lambda 2) - (\Lambda_1 \cup \Lambda 2)$ is also harmonious. But this is not the case if $\Lambda_1 = \mathbb{Z}$ and $\Lambda_2 = \sqrt{2}\mathbb{Z}$, since $\mathbb{Z} - \sqrt{2}\mathbb{Z}$ is not uniformly discrete, thus can not be harmonious.

The union of two model sets is in general not a model set. For example if $\Lambda_1 = \{n + \{\sqrt{2}n\}\}$, and $\Lambda_2 = \{n + \{\sqrt{3}n\}\}$, where $\{\cdot\}$ denotes the fractional part of a real number. Observe that $\Lambda_1$ is the model set defined by

$$\Lambda_1 = \{p_1(\gamma) \ : \ \gamma \in \Gamma, \ p_2(\gamma) \in [0, 1]\},$$

where $\Gamma = \begin{pmatrix} -1 & 1+\sqrt{2} \\ -1 & \sqrt{2} \end{pmatrix} \mathbb{Z}^2$, and similarly $\Lambda_2$. The union $\Lambda_1 \cup \Lambda_2$ is not uniformly discrete, since $\{\sqrt{2}n\} - \{\sqrt{3}n\}$ is dense in $[0, 1]$ by Kronecker's theorem. Hence the union is not a model set.

### 2.2.3 Density

From now on we will consider model sets with Euclidean internal spaces. Also we put the extra condition on the window set $\Omega$ to be Riemann measurable. That is the boundary of $\Omega$ has measure zero. This condition guarantees that the density of a model set is proportional to the measure of $\Omega$. First, we introduce two types of densities. They were proposed by Duffin Schaeffer [13] and Beurling [5] respectively.

**Definition 2.2.6.** A set $\Lambda = \{\lambda_n\}$ in $\mathbb{R}$ is said to have DS density $d$ if there exists a constant $L$ such that for all $n$,

$$|\lambda_n - \frac{n}{d}| \leq L.$$

**Definition 2.2.7.** A set $\Lambda = \{\lambda_n\}$ is said to have lower Beurling density $D^-(\Lambda)$ if

$$D^-(\Lambda) = \lim_{R\to\infty} \inf_{x\in\mathbb{R}} \frac{\#\{\Lambda \cap (x, x+R)\}}{R}$$

exists. It is said to have upper Beurling density $D^+(\Lambda)$ if

$$D^+(\Lambda) = \lim_{R\to\infty} \sup_{x\in\mathbb{R}} \frac{\#\{\Lambda \cap (x, x+R)\}}{R}$$

exists. It is said to have uniform Beurling density $d$ if $D^+(\Lambda) = D^-(\Lambda) = d$.

Definition 2.2.7 extends to higher dimensions in the obvious way. One simply replace the interval of length $R$ by a ball (with respect to some norm) with radius $R$.

**Proposition 2.2.5.** *In 1D, DS density implies the same Beurling density.*

Proof: Assume an interval $(a, b)$ contains exactly $n - m$ points $\{\lambda_{m+1}, \cdots \lambda_n\}$. We estimate the upper and lower bound of $n - m$ respectively.

(1)Upper bound: $\lambda_n \in (a, b)$ implies that $\frac{n}{d} - L \leq b$ and $\lambda_{m+1} \in (a, b)$ implies that $\frac{m+1}{d} + L \geq a$. Together, they imply that $n - m \leq d(b - a) + 2dL + 1$.

(2)Lower bound: $\lambda_{n+1} \notin (a, b)$ implies that $\frac{n+1}{d} + L \geq b$ and $\lambda_m \notin (a, b)$ implies that $\frac{m}{d} - L \leq a$. Together, they imply that $n - m \geq d(b - a) - 2dL - 1$.

Thus

$$d - \frac{2dL+1}{b-a} \leq \frac{\#\{\Lambda \cap (a,b)\}}{b-a} = \frac{n-m}{b-a} \leq d + \frac{2dL+1}{b-a.}$$

Let $b - a \to \infty$ we get the desired result. $\square$

We can easily see that the same result can be derived in higher dimensions. Below we give a proof of the two dimensional case, the higher dimensional proof is similar.

**Definition 2.2.8.** A set $\Lambda = \{\lambda_{m,n}\}$ is said to have DS density d if there exists a constant $L$ such that for all $n$,

$$|\lambda_{m,n} - \frac{(m,n)}{d}| \leq L,$$

where $|\cdot|$ denotes the $l_1$ norm.

**Definition 2.2.9.** A set $\Lambda = \{\lambda_{m,n}\}$ is said to have Beurling density $d$ if

$$d = \lim_{R\to\infty} \frac{\#\{\Lambda \cap [0,R]^2\}}{R^2}.$$

**Remark**: In Definition 2.2.8, using different equivalent norms of $\mathbb{R}^2$ will only affect the constant $L$, but not the DS density. We use $l_1$ norm since it is easier to count points in a square. In Definition 2.2.9, the position of the square $[0,R]^2$ is also not important as can be seen from the following proof. However, in 2D, one cannot define the Beurling density using a long thin rectangle since the limit

$$\lim_{R\to\infty} \frac{\#\{\Lambda \cap [0,R] \times [\epsilon, 2\epsilon]\}}{\epsilon R}$$

is apparently zero if $\Lambda \subset \mathbb{Z}^2$. In fact, the most general definition of Beurling uniform density is

$$\lim_{R\to\infty} \frac{\#\{\Lambda \cap B(x,R)\}}{|B(x,R)|},$$

where $B(x,R)$ is the ball of radius $R$ under some norm that is equivalent to the Euclidean norm.

**Proposition 2.2.6.** *In 2D, DS density implies the same Beurling density.*

Proof: For simplicity, we assume WLOG that d=1 in Definition 2.2.8. That is

$$|\lambda_{m,n} - (m,n)| \leq L.$$

Next, we estimate the upper and lower bound of $\#\{\Lambda \cap [0,R]^2\}$. Denote by $S_{m,n}$ the square centered at $(m,n)$ with side length $2L$. Then by assumption $\lambda_{m,n} \in S_{m,n}$.

(1)Upper bound: Let $S_{m+1,n+1}$ be the most bottom left square that has nonempty intersection with $[0,R]^2$. Then we have $m+1+L \geq 0, n+1+L \geq 0, m+L < 0, n+L < 0$. Let $S_{m+p,n+1}$ be the most bottom right square that has nonempty intersection with $[0,R]^2$. Then we have $m+p-L \leq R, m+p+1-L > R$. Let $S_{m+1,n+q}$ be the most top left square that has nonempty intersection with $[0,R]^2$. Then $m+q-L \leq R, m+q+1-L > R$. Then we have

$$\#\{\Lambda \cap [0,R]^2\} = pq \leq (R+L-m)(R+L-n).$$

(2)Lower bound: Let $S_{m+1,n+1}$ be the most bottom left square that is entirely contained in $[0,R]^2$. Then $m+1-L \geq 0, n+1-L \geq 0, m-L < 0, n-L < 0$. Let $S_{m+p,n+1}$ be the most bottom right square that is entirely contained in $[0,R]^2$. Then $m+p+L \leq R, m+p+1+L > R$. Let $S_{m+p,n+q}$ be the most top left square that is entirely contained in $[0,R]^2$. Then $n+q+L \leq R, n+q+1+L > R$. Then we have

$$\#\{\Lambda \cap [0,R]^2\} = pq \geq (R-L-m-1)(R-L-n-1).$$

Thus,

$$\lim_{R\to\infty} \frac{\#\{\Lambda \cap [0,R]^2\}}{R^2} = 1.$$

Another intrinsic property that model sets share with lattices is that they both have a certain uniform density property. It is clear by the definition of DS density that lattices have uniform DS density. It is an amazing fact that model sets have uniform Beurling density. This provides the theoretical guarantee that model sets can be used in sampling and interpolation as will be shown in the following sections. Below we give a proof from Matei and Meyer [24] of the density formula of model sets.

**Theorem 2.2.7.** *Given a model set $\Lambda_K = \{p_1(\gamma) : \gamma \in \Gamma, p_2(\gamma) \in K\}$ where $\Gamma \in \mathbb{R}^n \times \mathbb{R}^m$ is a lattice, and $K$ a compact Riemann measurable set. The Beurling density of $\Lambda_K$ can be computed by the following:*

$$dens\Lambda_K = |K|/|\Gamma|.$$

*Proof: Since $K$ is Riemann integrable, which means $|\partial K| = 0$, for every $\epsilon > 0$ one can find two non-negative smooth and compactly supported functions $g_\epsilon$ and $h_\epsilon$ such that $g_\epsilon \leq \chi_K \leq h_\epsilon$ and $\int |h_\epsilon - g_\epsilon|(x)dx \leq \epsilon$, where $\chi_K$ is the indicator function of $K$. Define:*

$$\mu_\epsilon = \sum_{\gamma \in \Gamma} g_\epsilon(p_2(\gamma))\delta_{p_1(\gamma)},$$

$$\nu_\epsilon = \sum_{\gamma \in \Gamma} h_\epsilon(p_2(\gamma))\delta_{p_1(\gamma)},$$

$$\sigma = \sum_{\gamma \in \Gamma} \chi_K(p_2(\gamma))\delta_{p_1(\gamma)} = \sum_{\lambda \in \Lambda_K} \delta_\lambda.$$

*Then clearly we have $\mu_\epsilon \leq \sigma \leq \nu_\epsilon$. We now compute the mean value of $\mu_\epsilon$:*

$$\mathcal{M}(\mu_\epsilon) = \lim_{R \to \infty} \frac{1}{|B(x,r)|} \int_{B(x,R)} d\mu_\epsilon.$$

*To do this, we take an approximate identity $\{\phi_n\}$, such that $\phi_n \in C_c^\infty(\mathbb{R}^n)$ and that $\int \phi_n = 1$. Define $f_{\epsilon,n} = \mu_\epsilon * \phi_n$then the mean value of $f_{\epsilon,n}$ will converge to the mean value of $\mu_\epsilon$. But*

$$f_{\epsilon,n}(x) = \sum_{\gamma \in \Gamma} \phi_n(x - p_1(\gamma))g_\epsilon(p_2(\gamma)) = |\Gamma|^{-1} \sum_{\gamma^* \in \Gamma^*} e^{-2\pi i x \cdot p_1(\gamma^*)} \widehat{\phi_n}(p_1(\gamma^*))\widehat{g_\epsilon}(p_2(\gamma^*)),$$

*where the second equality is the Poisson summation formula. Since both $\widehat{\phi_n}$,$\widehat{g_\epsilon}$ have rapid decay at infinity, the sum on the right hand side is absolutely continuous. Thus $f_{\epsilon,n}$ is an almost periodic function. Its mean value is the $0$'s coefficient of the Fourier expansion. Thus*

$$\mathcal{M}(\mu_\epsilon) = \lim_{n \to \infty} \mathcal{M}(f_{\epsilon,n}) = |\Gamma|^{-1} \int g_\epsilon.$$

*Similarly, we have* $\mathcal{M}(\nu_\epsilon) = |\Gamma|^{-1} \int h_\epsilon$. *Now*

$$\int h_\epsilon = \int g_\epsilon + \int (h_\epsilon - g_\epsilon) \leq |K| + \epsilon.$$

*Similarly,*

$$\int g_\epsilon \geq |K| - \epsilon,$$

*thus*

$$|K| - \epsilon \leq \int g_\epsilon \leq \int h_\epsilon \leq |K| + \epsilon.$$

*Divide by* $|\Gamma|$ *we have*

$$\frac{|K| - \epsilon}{|\Gamma|} \leq |\Gamma|^{-1} \int g_\epsilon = \mathcal{M}(\mu_\epsilon) \leq \mathcal{M}(\sigma) \leq \mathcal{M}(\nu_\epsilon) = |\Gamma|^{-1} \int h_\epsilon \leq \frac{|K| + \epsilon}{|\Gamma|}.$$

*But*

$$\mathcal{M}(\sigma) = \lim_{R \to \infty} \frac{\sigma(B(x,R))}{|B(x,R)|} = \lim_{R \to \infty} \frac{\#\{\Lambda_K \cap B(x,R)\}}{|B(x,R)|} = dens\Lambda_K.$$

*Thus* $dens\Lambda_K = |K|/|\Gamma|$. □

### 2.2.4 A Model Set Poisson Summation Formula

Due to its special form of construction by a window set in the internal space, each model set associates with it a Poisson summation formula, provided that a weight factor defined by compactly supported (by the window) function is present at one side of the equality. Let $\Gamma^*$ be the dual lattice of $\Gamma$ defined by

$$\Gamma^* = \{\gamma^* \in \mathbb{R}^n \times \mathbb{R}^m : \ (\gamma^*, \gamma) \in \mathbb{Z}, \ \forall \gamma \in \Gamma\}.$$

Let $p_1^*$ and $p_2^*$ be defined the same as $p_1$ and $p_2$ and define the map $(\cdot)^\star = p_2 \circ p_1^{-1}(\cdot)$. Let $\psi \in C_0^\infty(\Omega)$, then Meyer [26] proved the following Poisson summation formula for model sets:

**Theorem 2.2.8.** *For every $F \in W(L^\infty, l^1)$ such that $\widehat{F} \in W(L^\infty, l^1)$,*

$$\sum_{\lambda \in \Lambda(\Omega)} \psi(\lambda^\star) F(\lambda) e^{-2\pi i \lambda \cdot t} = vol(\Gamma)^{-1} \sum_{\gamma^* \in \Gamma^*} \hat{\psi}(-p_2^*(\gamma^*)) \widehat{F}(t - p_1^*(\gamma^*)).$$

*The identity holds pointwise for all $t \in \mathbb{R}^n$, and both sums converge uniformly and absolutely for all $t \in \mathbb{R}^n$.*

**Remark**: We can make our choice of $\psi$ such that $\hat{\psi} \geq 0$ by taking $\psi = \phi * \phi$. Furthermore, since $\hat{\psi}$ has isolated zeros, by rescaling if necessary, we can assume $\hat{\psi}(-p_2^*(\gamma^*)) > 0$.

### 2.2.5 A Uniqueness Theorem

Define the model set $\Lambda(\Omega)$ using the following lattice:

$$\Gamma = \begin{bmatrix} \sqrt{3} & -\sqrt{2} \\ -\sqrt{2} & \sqrt{3} \end{bmatrix} \begin{bmatrix} m \\ n \end{bmatrix}, \ \Gamma^* = \begin{bmatrix} \sqrt{3} & \sqrt{2} \\ \sqrt{2} & \sqrt{3} \end{bmatrix} \begin{bmatrix} m \\ n \end{bmatrix},$$

where $(m,n)\in\mathbb{Z}^2$. Let $\Omega$ be a compact Riemann integrable window set. Then we have

**Theorem 2.2.9.** *Suppose $\mu,\nu\in M_b(\mathbb{R})$ are discrete measures with support contained in $\mathbb{Z}$. If $\hat{\mu}(\lambda)=\hat{\nu}(\lambda)$, for $\lambda\in\Lambda(\Omega)$. Then $\mu=\nu$.*

Proof: Let $\rho=\mu-\nu=\sum\limits_{k\in\mathbb{Z}} c_k\delta_k$, let $\{f_j\}_{j=1}^{\infty}\in C_c^{\infty}(\mathbb{R})$ be an approximate identity. So then $\rho * f_j(x)=\sum_k c_k\delta_k * f_j=\sum_k c_k f_j(x-k)\in W(L^{\infty},l^1)$. On the other hand, $\|\hat{\rho}\|_{\infty}\le\|\rho\|_1=\sum_k|c_k|<\infty$, and $\hat{f}_j\in\mathcal{S}(\mathbb{R})$, thus $\hat{\rho}\hat{f}_j\in W(L^{\infty},l^1)$. By the above PSF for model sets, we have

$$\begin{aligned}
0 &= \sum_{\lambda\in\Lambda(\Omega)}\psi(\lambda^{\star})\hat{\rho}(\lambda)\hat{f}_j(\lambda)e^{-2\pi i\lambda t}\\
&= vol(\Gamma)^{-1}\sum_{\gamma^*\in\Gamma^*}\hat{\psi}(-p_2^*(\gamma^*))\rho * f_j(t-p_1^*(\gamma^*))\\
&= vol(\Gamma)^{-1}\sum_{\gamma^*\in\Gamma^*}\hat{\psi}(-p_2^*(\gamma^*))\sum_{k\in\mathbb{Z}}c_k f_j(t-p_1^*(\gamma^*)-k) \qquad (2.13)\\
&= vol(\Gamma)^{-1}\sum_{m,n\in\mathbb{Z}}\hat{\psi}(\sqrt{2}m+\sqrt{3}n)\sum_{k\in\mathbb{Z}}c_k f_j(t-\sqrt{3}m-\sqrt{2}n-k)\\
&\to vol(\Gamma)^{-1}\sum_{m,n,k\in\mathbb{Z}}\hat{\psi}(\sqrt{2}m+\sqrt{3}n)c_k\delta_{\sqrt{3}m+\sqrt{2}n+k}(t),
\end{aligned}$$

where the last limit is taking in the *weak$*$* sense. Indeed, for any test function

$\phi \in \mathcal{C}_0(\mathbb{R})$,we have

$$\begin{aligned}
&\lim_{j\to\infty} \sum_{m,n,k\in\mathbb{Z}} \hat{\psi}(\sqrt{2}m+\sqrt{3}n)c_k\big\langle f_j(t-\sqrt{3}m-\sqrt{2}n-k),\phi(t)\big\rangle \\
&= \sum_{m,n,k\in\mathbb{Z}} \hat{\psi}(\sqrt{2}m+\sqrt{3}n)c_k \lim_{j\to\infty} \big\langle f_j(t-\sqrt{3}m-\sqrt{2}n-k),\phi(t)\big\rangle \\
&= \sum_{m,n,k\in\mathbb{Z}} \hat{\psi}(\sqrt{2}m+\sqrt{3}n)c_k\phi(\sqrt{3}m+\sqrt{2}n+k) \\
&= \sum_{m,n,k\in\mathbb{Z}} \hat{\psi}(\sqrt{2}m+\sqrt{3}n)c_k\big\langle \delta_{\sqrt{3}m+\sqrt{2}n+k},\phi(t)\big\rangle,
\end{aligned} \tag{2.14}$$

where in the first equality we can switch limit and summation since the summation is uniformly convergent in $t$. Indeed, $\forall\epsilon > 0$, we can find an integer $N > 0$ that does not depend on $t$ such that

$$\begin{aligned}
&\sum_{|m|,|n|,|k|>N} \big|\hat{\psi}(\sqrt{2}m+\sqrt{3}n)c_k\big\langle f_j(t-\sqrt{3}m-\sqrt{2}n-k),\phi(t)\big\rangle\big| \\
&\leq \sum_{|m|,|n|,|k|>N} |\hat{\psi}(\sqrt{2}m+\sqrt{3}n)|\|\phi\|_\infty\|f_j\|_1 c_k \leq \epsilon.
\end{aligned} \tag{2.15}$$

Since $1, \sqrt{2}, \sqrt{3}$ are linearly independent over $\mathbb{Z}$, thus the support on the limit measure does not overlap for $m, n, k \in \mathbb{Z}$. Since by our choice, $\hat{\psi} > 0$, thus we can conclude $c_k = 0$ for all $k \in \mathbb{Z}$.

The above theorem can be more general. In fact, if the support of $\rho$ is finite dimensional over $\mathbb{Q}$, then we can construct $\Gamma$ in a way such that $p_1^*(\gamma^*)$ together with the support of $\rho$ is linearly independent over $\mathbb{Q}$. Then we can still conclude the same thing.

**Theorem 2.2.10.** *Suppose a discrete measure $\rho$ has the form*

$$\rho = \sum_{(n_1,\ldots,n_k)\in\mathbb{Z}^k} c_{n_1,\ldots,n_k}\delta_{\alpha_1 n_1+\cdots+\alpha_k n_k},$$

*where $\alpha_1,\ldots,\alpha_k$ are linearly independent over $\mathbb{Q}$. Then there exist a model set $\Lambda(\Omega)$ such that if $\hat{\rho}=0$ on $\Lambda(\Omega)$, then $\rho=0$.*

Proof: Define the lattice $\Gamma^*$ by

$$\Gamma^* = \begin{bmatrix} \beta_1 & \beta_2 \\ \beta_3 & \beta_4 \end{bmatrix} \begin{bmatrix} m \\ n \end{bmatrix},$$

where $(m,n)\in\mathbb{Z}^2$, such that $\alpha_1,\ldots,\alpha_k,\beta_1,\beta_2$ are linearly independent over $\mathbb{Q}$. Then we have

$$\begin{aligned}
0 &= \sum_{\lambda\in\Lambda(\Omega)} \psi(\lambda^\star)\hat{\rho}(\lambda)\hat{f}_j(\lambda)e^{-2\pi i\lambda t} \\
&= vol(\Gamma)^{-1}\sum_{\gamma^*\in\Gamma^*}\hat{\psi}(-p_2^*(\gamma^*))\rho * f_j(t-p_1^*(\gamma^*)) \\
&= vol(\Gamma)^{-1}\sum_{\gamma^*\in\Gamma^*}\hat{\psi}(-p_2^*(\gamma^*))\sum_{(n_1,\ldots,n_k)\in\mathbb{Z}^k} c_{n_1,\ldots,n_k} f_j(t-p_1^*(\gamma^*)-\alpha_1 n_1-\cdots\alpha_k n_k) \\
&= vol(\Gamma)^{-1}\sum_{m,n\in\mathbb{Z}}\hat{\psi}(-p_2^*(\gamma^*))\sum_{(n_1,\ldots,n_k)\in\mathbb{Z}^k} c_{n_1,\ldots,n_k} f_j(t-\alpha_1 n_1-\cdots\alpha_k n_k-\beta_1 m-\beta_2 n) \\
&\to vol(\Gamma)^{-1}\sum_{m,n,n_1,\ldots,n_k\in\mathbb{Z}^{k+2}}\hat{\psi}(-p_2^*(\gamma^*))c_{n_1,\ldots,n_k}\delta_{\beta_1 m+\beta_2 n+\alpha_1 n_1+\cdots+\alpha_k n_k}(t),
\end{aligned}$$

where the last limit is taking in the *weak*$*$ sense. The justification is similar to that of Theorem 2.2.9. By assumption, $\psi > 0$ and the supports of the delta's are not overlapping. Thus we conclude that the coefficients $c_{n_1,\ldots,n_k} = 0$. $\square$

# Chapter 3: SAMPLING AND EXACT RECONSTRUCTION

## 3.1 Introduction to sampling

We first fix some notations. Let $\Lambda \subset \mathbb{R}^d$, $S \subset \widehat{\mathbb{R}^d}$. We consider the case where $\Lambda$ is a sequence and $S$ is bounded measurable. Let $PW_S$ be the Paley-Wiener space consisting of all $f \in L^2(\mathbb{R}^d)$ whose Fourier transform $\hat{f}(\gamma) = \int e^{-2\pi i x\cdot\gamma} f(x)dx$ is supported by $S$. Let $E(\Lambda) = \{e^{2\pi i\lambda\cdot t}\}_{\lambda\in\Lambda}$ denote the sequence of exponential functions. When considered as functions defined on $S$, it is implicitly assumed that each one is multiplied by the characteristic function of $S$. A set $\Lambda$ is called uniformly discrete (u.d.) if

$$\inf_{\lambda_1,\lambda_2\in\Lambda} |\lambda_1 - \lambda_2| = \delta > 0.$$

**Definition 3.1.1.** $\Lambda$ is a set of sampling for $PW_S$ if there exists a constant $c$ such that for all $f \in PW_S$

$$||f||_2^2 \leq c \sum_{\lambda\in\Lambda} |f(\lambda)|^2. \tag{3.1}$$

The sampling problem is concerned with recovering a signal $f$ from a sequence of samples $f(\lambda), \lambda \in \Lambda$. Inequality (3.1) ensures that the reconstruction is stable in the sense that if the $l^2$ norm of the samples $f(\lambda)$ is small, then the $l^2$ norm of the reconstruction signal $f$ is small too. The converse of 3.1 holds under rather mild conditions as the next lemma shows.

**Lemma 3.1.1.** *(Bessel's inequality) If $S$ is bounded and $\Lambda$ is uniformly discrete. Then there exists a constant $C$ such that for all $f \in PW_S$*

$$\sum_{\lambda \in \Lambda} |f(\lambda)|^2 \leq C||f||_2^2. \tag{3.2}$$

Proof: We prove (3.2) with the constant

$$C = \frac{2\pi}{\sigma}$$

where $\sigma > 0$ is any number satisfying $S \subset [-\pi/4\sigma, \pi/4\sigma]$ and $\delta(\Lambda) > 2\sigma$, where $\delta(\Lambda) = \inf\limits_{\lambda, \lambda' \in \Lambda} |\lambda - \lambda'|$. Set

$$h(x) = \sqrt{\frac{1}{2\sigma}} \cdot \mathbb{1}_{[-\sigma,\sigma](x)},$$

then

$$\hat{h}(\gamma) = \sqrt{\frac{\sigma}{\pi}} \frac{sin\ \sigma\gamma}{\sigma\gamma}.$$

Since $\delta(\lambda) > 2\sigma$, the translates $\{h(x - \lambda), \lambda \in \Lambda\}$ form an orthonormal system in $L^2(\mathbb{R})$. Hence the same is true for the system $\{e^{i\lambda\gamma}\hat{h}(\gamma), \lambda \in \Lambda\}$. One may also verify that $\inf_{\gamma \in S} \hat{h}(\gamma) \geq \hat{h}(\frac{\pi}{4}) = \sqrt{\sigma/2\pi}$. Then, using Bessel's inequality for orthogonal systems, for every $f \in PW_S$ we have

$$\|f\|_2^2 = \|\hat{f}\|_2^2 \geq \frac{\sigma}{2\pi}\|\frac{\hat{f}}{\hat{h}}\|_2^2 \geq \frac{\sigma}{2\pi}\sum_{\lambda\in\Lambda} | < \frac{\hat{f}}{\hat{h}, e^{i\lambda\gamma}\hat{h}(\gamma)} > |^2 = \frac{\sigma}{2\pi}\sum_{\lambda\in\lambda} |f(\lambda)|^2. \quad \square$$

Using Plancherel's theorem and Lemma 3.1.1, it is easy to prove the following.

**Theorem 3.1.2.** *Let $S$ be bounded and measurable, and let $\Lambda$ be a u.d. set. Then $\Lambda$ is a set of sampling for $PW_S$ if and only if $E(\Lambda)$ is a frame in $L^2(S)$.*

**Definition 3.1.2.** $\Lambda$ is a set of interpolation for $PW_S$ if for every $l^2$ sequence $\{c_\lambda\}_{\lambda\in\Lambda}$, there exists $f \in PW_S$ with $f(\lambda) = c_\lambda$.

We have the following equivalent characterization of interpolation set.

**Theorem 3.1.3.** *Let $S$ be bounded and measurable, and let $\Lambda$ be a u.d. set. Then $\Lambda$ is a set of interpolation for $PW_S$ if and only if $E(\Lambda)$ is a Riesz sequence in $L^2(S)$, which means that there are positive constants $A$ and $B$ such that for every $\{c_\lambda\}_{\lambda\in\Lambda} \in l^2(\Lambda)$*

$$A||\sum_{\lambda\in\Lambda} c_\lambda e^{2\pi i\lambda t}||_{L^2(S)} \leq \sum_{\lambda\in\Lambda} |c_\lambda|^2 \leq B||\sum_{\lambda\in\Lambda} c_\lambda e^{2\pi i\lambda t}||_{L^2(S)}. \tag{3.3}$$

*Remark.* Note that the left inequality in (3.3) is just the dual inequality of the Bessel's inequality (3.2). It only depends on $S$ being bounded and $\Lambda$ being uniformly discrete. Indeed, inequality (3.2) means that the restriction operator

$$R : f \mapsto f|_\Lambda$$

is a bounded operator from $PW_S$ to $l^2(\Lambda)$. This operator can be identified with the operator

$$T : F \mapsto f|_\Lambda, \quad F = \hat{f},$$

acting from $L^2(S)$ to $l^2(\Lambda)$. The conjugate operator $T^* : l^2(\Gamma) \to L^2(S)$ has the form

$$T^*\{c_\lambda\} \mapsto \sum_{\lambda \in \Lambda} c_\lambda e^{2\pi i \lambda t}.$$

Thus the left inequality follows. So it is the right inequality in (3.3) that essentially characterizes the interpolation property.

Landau [19] proved necessary conditions for sampling and interpolation for $PW_S$ in terms of Beurling upper and lower densities.

**Theorem 3.1.4.** *(Landau) If $\Lambda$ is a set of sampling for $PW_S$, then*

$$D^-(\Lambda) \geq |S|.$$

*If $\Lambda$ is a set of interpolation for $PW_S$, then*

$$D^+(\Lambda) \leq |S|.$$

*Remark.* Landau's necessary conditions holds in any dimension.

In 1d, if one removes the equal signs, then the density conditions become sufficient, e.g. Seip[8]

**Theorem 3.1.5.** *Let $\lambda$ be uniformly discrete. For $E(\Lambda)$ to be a frame in $L^2([0,1])$,* *it is necessary that* $D^-(\Lambda) \geq 1$*, it is sufficient that* $D^-(\Lambda) > 1$*. For* $E(\Lambda)$ *to be a Riesz sequence in* $L^2([0,1])$*, it is necessary that* $D^+(\Lambda) \leq 1$*, it is sufficient that* $D^-(\Lambda) < 1$*.*

However, in higher dimensions, density condition alone cannot give sufficient conditions, since for any fixed bounded $S$, there is a lattice $\Lambda_1$ with arbitrarily large density that is not a set of sampling for $PW_S$. Similarly, there is a lattice $\Lambda_2$ with arbitrarily small density that is not a set of interpolation for $PW_S$. e.g. Olevskii & Ulanovskii [6], p.p.54, corollary 5.25. A concrete counter example is constructed via the following lemma.

**Lemma 3.1.6.** *Let* $E = T\mathbb{Z}^d$ *be a lattice,* $E^*$ *be the dual lattice, and* $\Omega \in \widehat{\mathbb{R}^d}$*.* $\mathcal{E}(E)$ *is a frame for* $L^2(\Omega)$ *if and only if*

$$|\Omega \cap (\Omega + e^*)| = 0, \quad e^* \in E^*, \ e^* \neq 0.$$

**Counterexample**. Let $E = \mathbb{Z}^2$ and $\Omega = [0, \frac{3}{2}] \times [0, \frac{1}{2}]$, then $1 = D(E) > |\Omega| = \frac{3}{4}$. However, $|\Omega \cap \Omega + (1,0)| = \frac{1}{4} > 0$. Thus $\mathcal{E}(E)$ is not a frame for $L^2(\Omega)$ even though the density condition is satisfied.

Even though Landau's necessary conditions are in no way sufficient, we could still ask ourselves the following two questions.

(1) Is there any specific set $\Lambda$ with $D^-(\Lambda) = D^+(\Lambda) = D(\Lambda)$, such that

$$D(\Lambda) > |S| \implies \textit{sampling} \tag{3.4}$$

and/or

$$D(\Lambda) < |S| \implies interpolation \tag{3.5}$$

(2) For a fixed set $S$, e.g. the unit ball, can we characterize the set $\Lambda$ such that the sampling or interpolation condition holds. In other words, can we characterize $\Lambda$ such that $E(\Lambda)$ forms a frame or a Riesz sequence for $L^2(S)$.

There are currently several answers to question (1), for example, simple quasicrystals, constructed by Matei & Meyer [24], and small perturbation of lattice, Olevskii [28]. Such a $\Lambda$ satisfying (3.4), resp. (3.5) is called universal sampling set, resp. universal interpolation set. In the first example simple quasicrystal cannot be replaced by harmonious set as the counterexample above shows. However, Model sets are harmonious and subsets of a harmonious set are still harmonious. Thus if we remove a finite subset $F$ from a simple quasicrystal $E$ with $D(E) > |\Omega|$, the remaining set $E\backslash F$ will be a harmonious set with $D(E\backslash F) > |\Omega|$. It can be shown that $\mathcal{E}(E\backslash F)$ is a frame for $L^2(\Omega)$ [23]. To prove this, recall the fact that the removal of one element from a frame will either result in a frame or a non-complete sequence.

**Theorem 3.1.7.** *(Matei) Let $E$ be a simple quasicrystal with $D(E) > |\Omega|$. Let $F \subset E$ be a finite set. Then $\mathcal{E}(E\backslash F)$ is still a frame for $L^2(\Omega)$.*

Proof: This will be proved by induction on the cardinality $m$ of $F$. More precisely we denote by $P_m$ the following property: For every finite set $F$ of cardinality not exceeding $m$ and every compact set $\Omega$ with $|\Omega| < D(E)$, $\mathcal{E}(E\backslash F)$ is a frame for $L^2(\Omega)$. $P_0$ is true by the theorem of Matei and Meyer. We now prove $P_m \implies P_{m+1}$. Assume the cardinality of $F$ is $m+1$. Translating $E$ if necessary we can assume

$0 \in F$. Let $F' = F\backslash\{0\}$. Then $\mathcal{E}(E\backslash F')$ is a frame for $L^2(\Omega)$ as long as $|\Omega| < D(E)$. Now, we need only to show $\mathcal{E}(E\backslash F)$ is a complete sequence in $L^2(\Omega)$ with $|\Omega| < D(E)$. Assume $g \in L^2(\Omega)$ is orthogonal to each $e^{2\pi i x\cdot\gamma} \cdot \mathbb{1}(\gamma)$, $x \in E\backslash F$. Consider $h_\epsilon(\gamma) = g(\gamma) - g(\gamma - \epsilon)$. Then $h_\epsilon \in L^2(\Omega_\epsilon)$, where $\Omega_\epsilon = \Omega + B(0, \epsilon)$. We choose $\epsilon$ small enough such that $|\Omega_\epsilon| < D(E)$. Then by construction $h_\epsilon$ is orthogonal to each $e^{2\pi i x\cdot\gamma} \cdot \mathbb{1}(\gamma)$, $x \in E\backslash F'$. However, by induction hypothesis, $\{e^{2\pi i x\cdot\gamma} \cdot \mathbb{1}(\gamma)\}$, $x \in E\backslash F'$ is a frame for $L^2(\Omega_\epsilon)$. Thus $h_\epsilon = 0$, which implies $g$ is periodic. But $g$ is compactly supported, so $g = 0$ as desired. $\square$

The above theorem says that if we have an Fourier frame with frequencies coming from a simple quasicrystal, then it is still a frame if we remove a finite set from it. A natural question to ask is that can $F$ be an infinite set. The answer is yes. Recall that the definition of the simple quasicrystal $E$ is that:

$$E = \{x \mid (x, y) \in \Gamma,\ y \in I\},$$

where $\Gamma \in \mathbb{R}^d \times \mathbb{R}$ is a lattice in general position(the canonical projection are one to one with dense range) and $I$ is an interval. The Beurling density of $E$ is given by $D(E) = |I|/vol(\Gamma)$. We can slightly shrink the size of $I$ by taking $I' \subset I$ so that $|I'|/vol(\Gamma) > |\Omega|$. Then $F = E\backslash E'$ is an infinite set and $E'$ is a simple quasicrystal defined by

$$E' = \{x \mid (x, y) \in \Gamma,\ y \in I'\}$$

with $D(E') > |\Omega|$. Thus $\mathcal{E}(E')$ is still a frame for $L^2(\Omega)$.

A reasonable question to ask is that can $F$ be arbitrary as long as $E \backslash F$ satisfies the necessary condition of Landau, i.e. $D^-(E\backslash F) \geq |\Omega|$. We know density condition alone cannot guarantee frame. But what if we restrict ourselves to subsets of a model set. This remains an open question.

An answer to question (2) in the case when $S$ is the unit ball was given by Beurling [3].

**Theorem 3.1.8.** *(Beurling) Let* $\rho = \rho(\Lambda) = \sup_{\xi \in \mathbb{R}^d} dist(\xi, \Lambda)$. *If*

$$\rho < \frac{1}{4},$$

*then* $\Lambda$ *is a set of sampling for* $PW_B$*, where* $B$ *is the unit ball in* $\mathbb{R}^d$.

*Remark.* Note that for a bounded set $S \in \mathbb{R}^d$, construction of an exponential frame $E(\Lambda)$ is easy. Just take $R$ large enough such that $R[-\frac{1}{2}, \frac{1}{2}]^d$ contains $S$. Then take $E(\Lambda)$ to be $\{e^{2\pi i n \cdot t/R}\}_{n \in \mathbb{Z}^d}$, which is an ONB for $R[-\frac{1}{2}, \frac{1}{2}]^d$, thus a tight frame with frame bounds 1 when restricted to the subset $S$.

In the case when $\Lambda$ has uniform density $D(\Lambda)$, if $\Lambda$ is both a sampling set and an interpolation set for $PW_S$, $E(\Lambda)$ will be both a frame and a Riesz sequence for $PW_S$, hence a Riesz basis for $PW_S$. In this case, Landau's theorem implies that we must have

$$D(\Lambda) = |S|.$$

We mentioned in the last remark that it is easy to construct an exponential frame for the unit disk. However it is an open question whether $L^2(B)$ admit an exponential

Riesz basis? Simple quasicrystals are sets of universal sampling and interpolation. One might ask for a simple quasicrystal $\Lambda$, for what set $S$ with $|S| = D(\Lambda)$, is $E(\Lambda)$ a Riesz basis for $L^2(S)$. This was answered by Grepsted & Lev in[3].

## 3.2 Beurling's balayage

The notion of balayage originated from Poincaré balayage process in potential theory. He showed that if $E$ is a set of spectral synthesis and of strict multiplicity, then balayage is possible for $(\Lambda, E)$ if and only if for every weak limit of translates of $\Lambda$ is a uniqueness sampling set for bounded continuous function whose spectrum is contained in $E$. This is also a necessary condition for stable sampling in the $L^\infty$ norm:

$$\sum_{x \in \mathbb{R}^d} |f(x)| \leq \ k \ \sup_{\xi \in \Lambda} |f(\xi)|$$

The connection between balayage and Fourier frames is detailed in the next subsection.

### 3.2.1 Introduction to balayage

In this subsection, we introduce the work of Beurling on Fourier frames and balayage. Let $E \in \mathbb{R}^d$ be a set. Denote by $M(E)$ the set of Radon measures with support in $E$. Let $\Lambda \in \hat{\mathbb{R}}^d$ be a compact set. Let $\mathcal{A}(\Lambda)$ be the restriction algebra with the usual quotient norm

$$\|f\|_\Lambda = \inf\{\int |d\alpha|, \hat{\alpha}(\xi) = f(\xi) \text{ on } \Lambda\}.$$

We say that balayage is possible for $(\Lambda, E)$ if there is for every $f \in \mathcal{A}(\Lambda)$ a measure $\beta \in M(E)$ such that

$$f(\xi) = \hat{\beta}(x), \quad \text{for } \xi \in \Lambda.$$

Balayage is possible implies that the mapping:

$$T : M(E) \to \mathcal{A}(\Lambda)$$

$$\nu \mapsto \hat{\nu}|_{\Lambda}$$

is onto. Thus by open mapping theorem, $T$ is an open map. Thus there exists a constant $K$ such that

$$\inf_{\beta \in M(E), \beta \sim \alpha} \int_E |d\beta| \leq K \int_{\mathbb{R}^d} |d\alpha|,$$

where the equivalent relationship $\beta \sim \alpha$ means $\hat{\beta} = \hat{\alpha}$ on $\Lambda$. The smallest such constant $K$ is denoted $K(\Lambda, E)$ and we set $K(\Lambda, E) = \infty$ is balayage is not possible.

Let $\phi$ be a bounded continuous function on $\mathbb{R}^d$, then $\phi$ has Fourier transform in the sense of distributions. Then the spectrum of $\phi$, denoted by $S_\phi$, is defined as the support of the Fourier transform of $\phi$. Define what's called the Bernstein space

$$\mathcal{C}(\Lambda) = \{\phi : \ \phi \text{ is bounded and continuous with } S_\phi \subset \Lambda\}$$

Such a function $\phi$ extends to an entire function of exponential type on $\mathbb{C}^d$. In the one dimensional case, if $S_\phi \subset [-\sigma, \sigma]$, then $\phi$ can be extend to a function $\Phi$ on $\mathbb{C}$

such that

$$|\Phi(x+iy)| \leq Ce^{\sigma|y|}.$$

$\mathcal{C}(\Lambda)$ equipped with the $\|\cdot\|_\infty$ is a Banach space. The Bernstein space is related to the Paley-Wiener space in the following way:

$$PW_\Lambda = B(\Lambda) \cap L^2(\mathbb{R}^d)$$

Below we give several properties for functions in the Bernstein space $\mathcal{C}(\Lambda)$.

**Proposition 3.2.1.** *If $\Lambda \subset \hat{\mathbb{R}}^d$ is contained in $|\xi| \leq R$, then*

$$|grad\phi(x_0)| \leq R \sup_{x\in\mathbb{R}^d} |\phi(x)|, \quad \textit{for}\ x \in \mathbb{R}^d. \tag{3.6}$$

**Proposition 3.2.2.** *Any bounded sequence of $\mathcal{C}(\Lambda)$ contains a subsequence that converges uniformly on compact subsets of $\mathbb{R}^d$ to some function in $\mathcal{C}(\Lambda)$.*

This follows from Montel's theorem which states that a uniformly locally bounded sequence of holomorphic functions is a normal family.

We shall need two conditions on the set $\Lambda$.

*Condition*$(\alpha)$: (strict multiplicity). For each $\xi_0 \in \Lambda$ and each $\epsilon > 0$, there exists a probability measure $\mu_\epsilon$ with support in $\{\xi \mid \xi \in \Lambda, |\xi - \xi_0| \leq \epsilon\}$ so that $\hat{\mu}_\epsilon(\gamma) \to 0$, $\gamma \to \infty$.

A condition equivalent to $(\alpha)$ is the following:

*Condition*$(\alpha_1)$: There exists a probablity measure $\mu$ with $S_\mu = \Lambda$, such that

$\hat{\mu}(\gamma) \to 0$, $\gamma \to \infty$.

*Condition*$(\beta)$: (spectral synthesis). If $\phi \in \mathcal{C}(\Lambda)$ and $\alpha \in M(\mathbb{R}^d)$ and $\hat{\alpha}(\gamma) = 0$ on $\Lambda$, then

$$\int \phi(x) d\alpha(x) = 0.$$

**Lemma 3.2.3.** *Assume* $(\alpha)$ *holds. Fix* $f \in \mathcal{A}(\Lambda)$. *Then there exists a measure* $\gamma$ *so that*

$$\hat{\gamma}(\xi) = f(\xi), \quad \int |d\gamma| = \|f\|_\Lambda.$$

*If* $K = K(\Lambda, E) < \infty$, $\gamma$ *exists with support in* $E$, $\int |d\gamma| \le K \|f\|_\Lambda$.

Proof: By definition of the quotient norm, let $\gamma_n$ be a minimizing sequence of measures with $\hat{\gamma}_n = f$ on $\lambda$, so that $\int |d\gamma_n| \to \|f\|_\Lambda$. Since $\|\gamma_n\|$ is bounded, by Banach-Alaoglu theorem, $\gamma_n \to \gamma$ *weakly*$^*$ for some $\gamma \in M(\mathbb{R})^d$. Take some $\mu_\epsilon$ given by condition $(\alpha)$ and fix $\xi_0 \in \Lambda$, we have

$$\int f(\xi) d\mu_\epsilon(\xi) = \int \hat{\gamma}_n(\xi) d\mu_\epsilon(\xi) = \int \hat{\mu}_\epsilon(x) d\gamma_n(x) \to \int \hat{\mu}_\epsilon d\gamma = \int \hat{\gamma} d\mu_\epsilon.$$

Letting $\epsilon \to 0$, we get $f(\xi_0) = \hat{\gamma}(\xi_0)$. Similar argument gives the second result. □

For a given closed set $Q$ and for $t > 0$, let $Q(t)$ denote the set of points with distance $\le t$ from $Q$. The Frechet distance $[Q, R]$ between two closed sets $Q$ and $R$ is the smallest number $t$ so that

$$Q \subset R(t), \quad R \subset Q(t)$$

Now let $Q_n$ be a sequence of closed sets.

**Definition 3.2.1.** $Q_n$ converges strongly to $Q$, denoted $Q_n \to Q$, if $[Q_n, Q] \to 0$. $Q_n$ converges weakly to $Q$, denoted $Q_n \rightharpoonup Q$, if for every compact set $L$, $Q_n \cap L \to Q \cap L$.

**Theorem 3.2.4.** *If condition $(\alpha)$ holds, then $E_n \rightharpoonup E$ implies*

$$K(\Lambda, E) \leq \liminf K(\Lambda, E_n).$$

Proof: Without loss of generality, we may assume that $\liminf K(\Lambda, E_n) < \infty$. Now given $\phi \in \mathcal{A}(\Lambda)$, by Lemma 3.2.3, there exists a measure $\nu_n$ supported on $E_n$ such that

$$\hat{\nu}_n(\xi) = \phi(\xi), \quad \text{for all } \xi \in \Lambda,$$

and

$$\int |d\nu_n| \leq K(\Lambda, E_n) \|\phi\|_\Lambda.$$

By passing to a subsequence, we may assume with out loss of generality that

$$\int |d\nu_n| \to \liminf K(E_n, \Lambda) \|\phi\|_\Lambda.$$

By Banach-Alaoglu theorem, without loss of generality, $\nu_n \rightharpoonup \nu$ weakly*, so then

$$\hat{\nu}(\xi) = \phi(\xi), \text{for } \xi \in \Lambda$$

and

$$\int |d\nu| = \liminf K(\Lambda, E_n)\|\phi\|_\Lambda.$$

Since $E_n \rightharpoonup E$, *supp* $\nu = E$. Thus

$$\inf_{\beta \in M(E), \hat{\beta}=\phi} \int |d\beta| \leq \int |d\nu| = \liminf K(\Lambda, E_n)\|\phi\|_\Lambda$$

By definition of $K(\Lambda, E)$, we have $K(\Lambda, E) \leq \liminf K(\Lambda, E_n)$. □

**Definition 3.2.2.** $k(\Lambda, E)$ is the smallest number $k$ so that for all $\phi \in \mathcal{C}(\Lambda)$

$$\sup_{x \in \mathbb{R}^d} |\phi(x)| \leq k \sup_{x \in E} |\phi(x)|.$$

If such a $k$ does not exist, we simply say $k = \infty$.

**Lemma 3.2.5.** *Condition* $(\alpha)$ *implies* $K \leq k$. *Condition* $(\beta)$ *implies* $k \leq K$.

Proof: Assume $k \leq \infty$ and that $(\alpha)$ holds. Let $\mathcal{C}_0(\Lambda)$ denote $\{\phi \in \mathcal{C}_\Lambda \mid \lim_{|x|\to\infty} \phi(x) = 0\}$. Given $\alpha \in M(\mathbb{R}^d)$, define

$$L(\phi) = \int_{\mathbb{R}^d} \phi(x) d\alpha(x).$$

$L$ is a linear functional on $\mathcal{C}_0(\Lambda)$ with

$$\|L\| = \int |d\alpha|.$$

Let $\mathcal{C}_0(\Lambda)|_E$ be the space of restrictions of functions in $\mathcal{C}_0(\Lambda)$ to $E$, endowed with the $\|\cdot\|_\infty$ norm. Due to the fact that $\sup_{x\in\mathbb{R}^d}|\phi(x)| \le k\sup_{x\in E}|\phi(x)|$, each function $\psi \in \mathcal{C}_0(\Lambda)|_E$ is the restriction to a unique function $\phi \in \mathcal{C}_0(\Lambda)$, i.e., $\psi = \phi|_E$. Thus we can define a linear functional $\tilde{L}$ on $\mathcal{C}_0(\Lambda)|_E$ by

$$\tilde{L}(\psi) = L(\phi),$$

where $\phi$ is the unique extension of $\psi$ to $\mathcal{C}_0(\Lambda)$. Clearly,

$$\begin{aligned}
|\tilde{L}(\psi)| &= |L(\phi)| \\
&\le \|L\| \sup_{x\in\mathbb{R}^d}|\phi(x)| \\
&\le \|L\| k \sup_{x\in E}|\phi(x)| \\
&= \|L\| k \|\psi\|_\infty.
\end{aligned}$$

Thus $\|\tilde{L}\| \le k\|L\|$. By Riesz representation theorem, there exists a measure $\beta \in M(E)$ such that

$$\tilde{L}(\psi) = \int_E \psi d\beta.$$

Therefore, for every $\phi \in \mathcal{C}_0(\Lambda)$, we have

$$\int_{\mathbb{R}^d} \phi(x)d\alpha(x) = L(\phi) = \tilde{L}(\phi|_E) = \int_E \phi(x)d\beta(x),$$

and

$$\int |d\beta| = \|\tilde{L}\| \le k\|L\| = k\int |d\alpha|.$$

Next, $K \leq k$ follows from the fact that $\hat{\alpha} = \hat{\beta}$ on $\Lambda$. Indeed, fix $\xi_0 \in \Lambda$, choose $\mu_\epsilon$ as in condition $(\alpha)$, then

$$\int \hat{\mu}_\epsilon d\alpha = \int \hat{\mu}_\epsilon d\beta.$$

By letting $\epsilon \to 0$, we get $\hat{\alpha}(\xi_0) = \hat{\beta}(\xi_0)$. This finishes the proof of the first statement.

Now assume that $(\beta)$ holds and that $K < \infty$. Fix $x_0 \in \mathbb{R}^d$. Then $\exists \beta_{x_0} \in M(E)$ with $\hat{\delta}_{x_0}(\xi) = e^{-2\pi i \langle x_0, \xi \rangle} = \int_E e^{-2\pi i \langle x, \xi \rangle} d\beta_{x_0}(x)$, $\xi \in \Lambda$ and

$$\int |d\beta_{x_0}| \leq K.$$

Hence by condition $(\beta)$ if $\phi \in \mathcal{C}(\Lambda)$

$$\int \phi(d\delta_{x_0} - d\beta_{x_0}) = 0, \quad \text{or } \phi(x_0) = \int \phi d\beta_{x_0}. \tag{3.7}$$

It follows that

$$|\phi(x_0)| \leq K \sup_E |\phi(x)|.$$

Hence $k \leq K$. □

The corollary below follows immediately.

**Corollary 3.2.6.** *Assume condition $(\alpha)$ and $(\beta)$ holds. Assume $K(\Lambda, E) < \infty$. If $\phi \in \mathcal{C}(\Lambda)$ and $\phi = 0$ on $E$, then $\phi = 0$*

**Theorem 3.2.7.** *Assume $(\alpha)$ and $\beta$. Let $E_1$, $E_2$ be two closed sets. Then*

$$|K(\Lambda, E_1)^{-1} - K(\Lambda, E_2)^{-1}| \leq diam(\Lambda)[E_1, E_2].$$

Proof: Let $r = diam(\Lambda)$. We may assume that $\Lambda \subset \{|x| < R\}$. Choose $\phi \in \mathcal{C}(\Lambda)$ with $\sup_{\mathbb{R}^d} |\phi(x)| = 1$. Then

$$\sup_{E_1} |\phi(x)| \geq k(\Lambda, E_1)^{-1} = k_1^{-1}.$$

Choose $x_0 \in E_1$ so that $|\phi(x_0)| > k_1^{-1} - \epsilon$.

If $|x - x_0| \leq t$, by Bernstein's inequality,

$$|\phi(x)| > k_1^{-1} - \epsilon - rt.$$

If we take $t = [E_1, E_2]$, then there exists $x \in E_2$ with $|x - x_0| \leq t$. Hence

$$\sup_{E_2} |\phi(x)| > k_1^{-1} - r[E_1, E_2].$$

By definition of $k_2 = k(\Lambda, E_2)$, we have

$$k_2^{-1} \geq k_1^{-1} - r[E_1, E_2]$$

By reverting $E_1$, $E_2$ and using the fact that $k = K$, the theorem follows. $\square$

**Corollary 3.2.8.** *(1)If $(\alpha)$ and $(\beta)$ hold and $K(\Lambda, E) < \infty$, then there is a uniformly discrete subset $E'$ of $E$ so that*

$$K(\Lambda, E') < K(\Lambda, E) + \epsilon.$$

*(2)If $(\alpha)$ and $(\beta)$ hold and balayage is possible for $(\Lambda, E)$, then there exists a uniformly discrete subset $E^{'}$ of $E$ so that balayage is possible for $(\Lambda, E^{'})$.*

**Definition 3.2.3.** For a closed set $E$, let $W(E)$ be the collection of weak limits of translates $E^{\tau} = E + \tau$. That is if $E' \in W(E)$, there exists a sequence $\{\tau_n\}$ such that

$$E^{\tau_n} \rightharpoonup E' \quad \text{as } n \to \infty.$$

**Theorem 3.2.9.** *Assume $(\alpha)$ and $(\beta)$ hold. Then $K(\Lambda, E) < \infty$ if and only if, for every $E_0 \in W(E)$, $\phi \in \mathcal{C}(\Lambda)$ and $\phi(x) = 0$ on $E_0$ implies $\phi \equiv 0$.*

Proof: First, assume $K(\Lambda, E) < \infty$ and $E_0 \in W(E)$. There exists $\{\tau_n\}$ such that $E^{\tau_n} \rightharpoonup E_0$, then by Theorem 3.2.4, we have

$$K(\Lambda, E_0) \leq \liminf_{n\to\infty} K(\Lambda, E^{\tau_n}) = K(\Lambda, E) < \infty.$$

Then by corollary 3.2.6, we have $\phi \equiv 0$.

Conversely, assume $K(\Lambda, E) = \infty$. Then by Lemma 3.2.5, there exists a sequence $\phi_n \in \mathcal{C}(\Lambda)$ so that $\sup |\phi_n(x)| = 1$, $\sup_E |\phi_n(x)| \to 0$. Choose $x_n$ so that $|\phi_n(x_n)| = \frac{1}{2}$ and define $\psi_n(x) = \phi_n(x+x_n)$. Then $|\psi_n(0)| = \frac{1}{2}$. Setting $E_n = E - x_n$, we have

$$\sup_{E_n} |\psi_n(x)| \to 0.$$

Denote by $E_0$ a weak limit (possibly empty) of $E - x_n$. By compactness property, we may assume $\psi_n$ converges pointwise to some $\psi \in \mathcal{C}(\Lambda)$. This implies that $\psi = 0$ on $E_0$. However $\psi(0) = \frac{1}{2}$, which is a contradiction. $\square$

**Theorem 3.2.10.** *Assume* $(\alpha)$ *and* $(\beta)$ *hold. Let* $\Lambda_\epsilon = \{x \mid dist(x, \Lambda) \leq \epsilon\}$. *If* $K(\Lambda, E) \leq \infty$, *then there exists* $\epsilon_0 > 0$ *such that* $K(\Lambda_\epsilon, E) < \infty$ *for* $\epsilon < \epsilon_0$.

Proof: Suppose $K(\Lambda_\epsilon, E) = \infty$ for arbitrarily small $\epsilon$. Clearly, $\Lambda_\epsilon$ satisfies $(\alpha)$. By Lemma 3.2.5, $k(\Lambda_\epsilon, E) = \infty$. So there exists $\phi_\epsilon \in \mathcal{C}(\Lambda_\epsilon)$ with $\sup |\phi_\epsilon(x)| = 1$ and $|\phi_\epsilon(x)| \leq \epsilon$ on some $E_\epsilon$ in $W(E)$. We choose $x_\epsilon$ so that $|\phi_\epsilon(x_\epsilon)| = \frac{1}{2}$. The translates $E - x_\epsilon$ converge weakly to some $E_0 \in W(E)$. Also $\psi_\epsilon(x) = \phi_\epsilon(x + x_\epsilon) \to \psi$ for some $\psi \in \mathcal{C}(\Lambda)$ with $\psi = 0$ on $E_0$, and $|\psi(0)| = \frac{1}{2}$. By theorem 3.2.9, $K(\Lambda, E) = \infty$ which is a contradiction. $\square$

### 3.2.2 Balayage for an interval

Now we focus on the case when $\Lambda$ is an interval $(-a, a)$ on the real line. In this case, $(\alpha)$ and $(\beta)$ both hold. We also assume that $E \subset \mathbb{R}$ is a uniformly discrete set $\{x_n\}$. In this case, Beurling showed that the possibility of Balayage is equivalent to the density condition that $D^-(E) > \frac{a}{\pi}$. The details of the proof of the following theorem can be founded in [5].

**Theorem 3.2.11.** $K(E, a) < \infty$ *if and only if* $D^-(E) > \frac{a}{\pi}$.

Next, we show that Balayage is a sufficient condition for exponential frame, c.f. [1]. First we need the Ingham inequality [16].

**Theorem 3.2.12.** *Let* $\epsilon > 0$ *and let* $\Omega : [0, \infty) \to (0, \infty)$ *be a continuous function that increases to infinity. Assume the following conditions:*

$$\int_1^\infty \Omega(r) \frac{dr}{r^2} < \infty,$$

$$\int_{\mathbb{R}^d} e^{-\Omega(\|x\|)}dx < \infty,$$

*and $\Omega(r) > r^a$ on some interval $[r_0, \infty)$ and for some $a < 1$. Then there exists $h \in L^1(\mathbb{R}^d)$ for which $h(0) = 1$, supp$(\hat{h}) \subset B(0, \epsilon)$, and $|h(x)| \leq Ce^{-\Omega(\|x\|)}$.*

Even though the equivalence of balayage and lower Beurling density being greater than the size of the interval only holds for $\mathbb{R}$, the following theorem holds in any dimension.

**Theorem 3.2.13.** *Assume condition $(\alpha)$ and $(\beta)$ hold for $\Lambda \subset \widehat{\mathbb{R}}^d$, and that $E \subset \mathbb{R}^d$ is a uniformly discrete sequence. If balayage is possible for $(E, \Lambda)$, then $\{e^{-2\pi i x\gamma} \mid x \in E\}$ is a frame for $L^2(\Lambda)$.*

Proof: By Theorem 3.2.10, there exists $\epsilon > 0$ such that balayage is possible for $(E, \Lambda_\epsilon)$. That is $K = K(\Lambda_\epsilon, E) < \infty$. For this $\epsilon$, take $h$ from Ingham's Theorem 3.2.12. By definition of balayage, there exists a sequence $\{a_x(y) : x \in E\}$ such that

$$(\delta_y)\widehat{\ } = (\sum_{x \in E} a_x(y)\delta_x)\widehat{\ }, \quad \text{on } \Lambda_\epsilon$$

where

$$\sum_{x \in E} |a_x(y)| \leq K(\Lambda_\epsilon, E)$$

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

Now fix $y \in \mathbb{R}^d$, consider the measure

$$\eta_y(w) = h_y(w)\Big(\delta_y - \sum_{x \in E} a_x(y)\delta_x\Big),$$

where $h_y(w) = h(w-y)$. Then we have

$$
\begin{aligned}
(\eta_y)\widehat{\ }(\gamma) &= \Big[\, (h_y)\widehat{\ } * \Big(\delta_y - \sum_{x\in E} a_x(y)\delta_x\Big)\widehat{\ }\,\Big] \\
&= \int \hat{h}(\gamma-\lambda)e^{-2\pi i y\cdot(\gamma-\lambda)}\Big(\delta_y - \sum_{x\in E} a_x(y)\delta_x\Big)\widehat{\ }(\lambda)d\lambda \\
&= \int_{(\Lambda_\epsilon)^c} \hat{h}(\gamma-\lambda)e^{-2\pi i y\cdot(\gamma-\lambda)}\Big(\delta_y - \sum_{x\in E} a_x(y)\delta_x\Big)\widehat{\ }(\lambda)d\lambda = 0
\end{aligned}
$$

whenever $y \in \mathbb{R}^d$ and $\gamma \in \Lambda$. Thus $\forall f \in \mathcal{C}(\Lambda)$, $\langle f, \eta_y\rangle = 0$. That is

$$f(y) = \sum_{x\in E} f(x)a_x(y)h(x-y).$$

In particular,

$$e^{2\pi i y\gamma} = \sum_{x\in E} a_x(y)h(x-y)e^{2\pi i x\gamma}.$$

Now we are ready to prove the frame inequality. It suffices to prove the lower frame bound. (The upper frame bound follows from uniformly discreteness and Bessel's inequality.) Let $F \in L^2(\Lambda)$ and $\hat{f} = F$ so that $f \in L^2(\mathbb{R}^d)$.

$$
\begin{aligned}
\|F\|^2_{L^2(\Lambda)} &= \int_\Lambda \overline{F(\lambda)}\Big(\int f(y)e^{-2\pi i\lambda y}dy\Big)d\lambda \\
&= \int_\Lambda \overline{F(\lambda)}\Big(\int f(y)\Big(\sum_{x\in E} a_x(y)h(x-y)e^{-2\pi i x\gamma}\Big)dy\Big)d\lambda \\
&= \sum_{x\in E}\overline{f(x)}\Big(\int a_x(y)h(x-y)f(y)dy\Big) \\
&\leq \Big(\sum_{x\in E}|f(x)|^2\Big)^{1/2}\Big(\sum_{x\in E}\Big|\int a_x(y)h(x-y)f(y)dy\Big|^2\Big)^{1/2} \\
&\leq \Big(\sum_{x\in E}|f(x)|^2\Big)^{1/2}\Big(\sum_{x\in E}\int |a_x(y)||h(x-y)|^2dy\int |a_x(y)||f(y)|^2dy\Big)^{1/2} \\
&\leq C^{1/2}\|h\|\Big(\sum_{x\in E}|f(x)|^2\Big)^{1/2}\Big(\int(\sum_{x\in E}|a_x(y)|)|f(y)|^2dy\Big)^{1/2} \\
&\leq C^{1/2}\|h\|K(\Lambda_\epsilon, E)^{1/2}\|f\|_2\Big(\sum_{x\in E}|f(x)|^2\Big)^{1/2}
\end{aligned}
$$

Divide by $\|f\|_2$ on both sides, we get the desired estimate. □

## 3.3 Exact reconstruction

In image processing, the problem of the exact reconstruction of a positive measure appears in several applications as image compression, superresolution problem and image denoising. The pioneering basis pursuit algorithm is used for the exact reconstruction of sparse finite dimensional vectors. The basis pursuit was introduced to the statistics community by Chen, Donoho and Saunders [6] and by earlier works of Donoho and Stark [10]. P. Doukhan, E. Gassiat and P. Gamboa considered in [14] and in [11] the exact reconstruction of a nonnegative measure in relation with super-resolution problem. It is worth mentioning that in both problems, uniqueness play a central role because it will guarantee that any numerical approximation process will converge to the right answer.

In the this subsection, we present a result of Matei [22]. He uses the arithmetical properties of simple quasicrystals to reconstruct a positive measure in a deterministic way. In particular, he reconstructs measures with small supports by using an irregular sampling in the Fourier domain defined by simple quasicrystals.

### 3.3.1 Construction of sampling sets

$\mathbb{Z}^d$ can be embedded into $\mathbb{T}$ by the mapping $\gamma^* : \mathbb{Z}^d \to \mathbb{T}$ defined by:

$$\gamma^*(n_1, n_2, ..., n_d) = (n_1 q_1 + ... + n_d q_d) \quad mod\ 1,$$

where $q_1, ..., q_d$ are $d$ irrational numbers such that $1, q_1, ..., q_d$ are linearly independent over $\mathbb{Q}$.

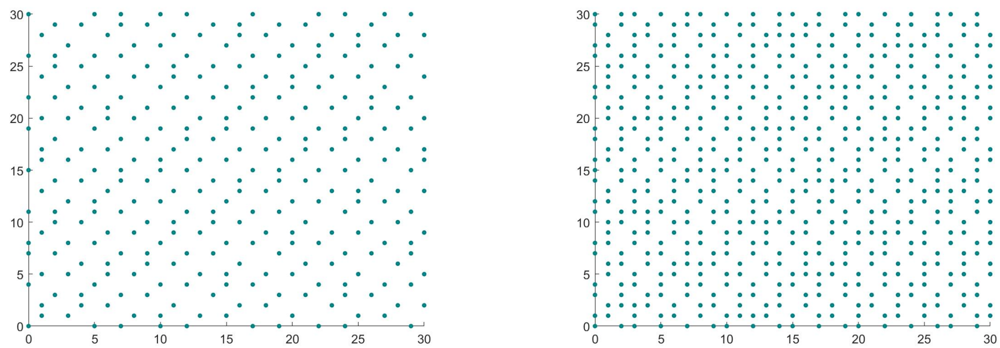
This mapping $\gamma^*$ is injective (by linear independence) and has dense range in $\mathbb{T}$ (by Dirichlet Theorem).

Let $I \subset \mathbb{T}$ be an interval. Define the model set $\Lambda_I \subset \mathbb{Z}^d$ by letting

$$\Lambda_I = \left\{(n_1, n_2, ..., n_d) \in \mathbb{Z}^d : \ \gamma^*(n_1, n_2, ..., n_d) \ \in \ I\right\}$$

If $I = [-\alpha, \alpha]$, $\alpha < 1/2$, we write $\Lambda_\alpha$ instead of $\Lambda_I$. Here are some examples of what the set looks like in 2 dimensional space.

One remark we would like to make is that these model sets are subsets of $\mathbb{Z}^d$. However, when we talk about their properties such as being harmonious, uniformly




(a) $alpha = 0.15$ (b) $alpha = 0.3$

Figure 3.1: Model set defined by $\{(p,q) \in \mathbb{Z}^2 \mid \exists r \in \mathbb{Z}\ s.t.\ |p\sqrt{2} + q\sqrt{3} - r| < \alpha\}$ with (a)$\alpha = 0.15$ and (b)$\alpha = 0.3$.

discrete and relatively dense, they should really be treated as subsets of $\mathbb{R}^d$. For example, the property of being harmonious is trivial if we consider these model set in the group of $\mathbb{Z}^d$, since every algebraic character on a discrete group is automatically continuous, thus every subset of $\mathbb{Z}^d$ is harmonious. However, when we talk about exact reconstruction of a function defined on $\mathbb{T}^d$, its spectrum lies in $\mathbb{Z}^d$. This may cause ambiguity since we are sampling on a subset of the spectrum $\mathbb{Z}^d$, whose certain properties rely on being viewed as a subset of $\mathbb{R}^d$. In the next section, we will see that the Theorem 3.3.1 relies only on the arithmetic property of model set directly related to its definition. Thus it makes sense to use model set in this specific setting where the ambient group $\mathbb{R}^d$ is ignored.

### 3.3.2 Exact reconstruction for positive discrete measures

Let $\alpha$ be a fixed constant in (0,1/2), and $\nu$ be a measure in $\mathcal{M}_{\mathcal{N}}$, that is $\nu = \sum_{j=1}^{N} \omega_j \delta_{x_j}$, where the weights $\omega_j \geq 0$, $x_j \in F = \{x_1, ..., x_N\} \subset \mathbb{T}^d$

**Theorem 3.3.1.** *Let $\nu$ be a measure defined above and $\mu \geq 0$ be a positive measure on the torus $\mathbb{T}^d$ such that $\hat{\mu}(\lambda) = \hat{\nu}(\lambda)$, $\lambda \in \Lambda_\alpha$. Then $\mu = \nu$.*

Proof: We define $\rho = \mu - \nu$, then $\hat{\rho}(\lambda) = 0$, $\lambda \in \Lambda_\alpha$. We begin with the following lemma:

**Lemma 3.3.2.** *Let $\theta$ be the triangle function supported on $[-\alpha, \alpha]$. Then*

$$\theta(x) = \sum_{k=-\infty}^{\infty} p_k exp(2\pi ikx), \text{ where } p_k \geq 0.$$

*Let $\tau$ be the measure $\sum_{k=-\infty}^{\infty} p_k \delta_{y_k}$, where $y_k = (kq_1, ..., kq_d)$ mod $\mathbb{Z}^d$. Then $\tau * \rho = 0$.*

Proof: By definition of the measure $\tau$, for all $(n_1, ..., n_d) \in \mathbb{Z}^d$, we have

$$\hat{\tau}(n_1, ..., n_d) = \sum_{k=-\infty}^{\infty} p_k exp(2\pi i(n_1q_1 + ... + n_dq_d)k) = \theta(n_1q_1 + ... + n_dq_d)$$

Note that $(n_1q_1 + ... + n_dq_d)$ *mod* $1 \in J = \mathbb{T} \backslash I$ for all $(n_1, ..., n_d) \notin \Lambda_\alpha$. Since $\theta$ vanishes on $J$ it follows that $\hat{\tau}(n_1, ...n_d) = 0$ whenever $(n_1, ..., n_d) \notin \Lambda_\alpha$. Now $\hat{\rho}(\lambda) = 0$ whenever $\lambda \in \Lambda_\alpha$. Then $\hat{\tau} \cdot \hat{\rho} = 0$, which implies $\tau * \rho = 0$.

Lemma 3.3.2 shows that $\tau * (\mu - \nu) = 0$ over $\mathbb{T}^d$. More precisely,

$$(\tau * \rho)(x) = \sum_{k=-\infty}^{\infty} p_k(\mu - \nu)(x - y_k) = 0.$$

Consider the measure $\widetilde{\mu} = \tau * \mu$. Then $\widetilde{\mu}$ satisfies the following identity

$$\widetilde{\mu}(x) = \sum_{k=-\infty}^{\infty} p_k(\mu)(x - y_k) = \sum_{k=-\infty}^{\infty} \sum_{j=1}^{N} p_k \omega_j \delta_{y_k + x_j}.$$

Note that the right hand side of the above equality is an atomic measure supported by $\Gamma + F$ where $\Gamma = \{(kq_1, ..., kq_d) : k \in \mathbb{Z}\}$ and $F = \{x_1, ..., x_N\}$.

The set of points $\Gamma + F$ may be written as $\Gamma_1 \cup \Gamma_2 \cup ... \cup \Gamma_m$ where $\Gamma_j = \Gamma + x_j$ where $\Gamma_j$ are disjoint cosets after relabeling of $F$ if necessary. It follows that the measure $\mu$ is absolutely continuous with respect to the measure $\widetilde{\mu}$ (since$\widetilde{\mu} \geq p_0\mu$). So, $\mu$ is also an atomic measure supported on the set of points $\Gamma + F$, which implies $\rho = \mu - \nu$ is also supported on $\Gamma + F$. We decompose the measure $\rho$ as follows

$$\rho = \rho_1 + ... + \rho_m, \text{ where } \mathrm{supp}(\rho_j) \subset \Gamma + x_j, 1 \leq j \leq m.$$

It follows that

$$\tau * \rho = \tau * \rho_1 + ... + \tau * \rho_m,$$

and $\mathrm{supp}(\tau * \rho_j) \subset \Gamma + x_j$,$1 \leq j \leq m$.

Since $\Gamma_j$ are disjoint sets of points, it follows that $\tau * \rho_j = 0$ for all $1 \leq j \leq m$. Let us consider $\mu_j(x) = \rho_j(x + x_j)$. Then $\tau * \mu_j = 0$ and $\mu_j$ is a measure supported on $\Gamma$. Since $\nu$ has finite support, it follows that $\mu_j$ has positive weights except for a finite number of terms.

**Lemma 3.3.3.** *Let $\sigma$ be an atomic measure supported on $\Gamma$, excepting a finite number, all weights in the definition of $\sigma$ are nonnegatives and also we assume that $\tau * \sigma = 0$. Then $\sigma = 0$.*

Proof: Let us consider $\sigma = \sum_{k=-\infty}^{\infty} a_k \delta_{y_k}$ and also

$$g(x) = \sum_{k=-\infty}^{\infty} a_k exp(2\pi i k x.)$$

The hypothesis $\tau * \sigma = 0$ is equivalent to $\hat{\tau} \cdot \hat{\sigma} = 0$. Therefore

$$\hat{\sigma}(n_1, ..., n_d) = 0, \quad \text{ if } n_1 q_1 + ... + n_d q_d \in [-\alpha, \alpha].$$

But $\hat{\sigma}(n_1, ..., n_d) = g(n_1 q_1 + ... + n_d q_d)$, thus by the denseness of $(n_1 q_1 + ... + n_d q_d)$ *mod* 1 in $\mathbb{T}$, g=0 on $[-\alpha, \alpha]$. Let $\phi$ be a compactly supported function such that $\phi \in \mathcal{C}_0^\infty(\mathbb{T})$ and $\text{supp}(\phi) \subset [-\frac{\alpha}{8}, \frac{\alpha}{8}]$, $\phi > 0$ over $(-\frac{\alpha}{8}, \frac{\alpha}{8})$. We define $\Phi = \phi * \phi$. It follows that $\hat{\Phi}(k) \geq 0$. By the following Lemma 3.3.4, we get a more precise result, namely $\hat{\Phi}(k) > 0, k \in \mathbb{Z}$. We replace $g$ by $G = g * \Phi$ and we get

$$G(x) = \sum_{k=-\infty}^{\infty} \alpha_k exp(2\pi ikx), \text{ where } \alpha_k = a_k \hat{\Phi}(k).$$

Now, $G \in \mathcal{C}_0^\infty(\mathbb{T})$, excepting a finite number all $\alpha_k \geq 0$. Moreover $G$ vanishes over $[-\frac{\alpha}{4}, \frac{\alpha}{4}]$ by construction. By the following Lemma 3.3.5 in Appendix, we get $G = 0$. But $\hat{G} = \hat{g} \cdot \hat{\Phi}$, which implies $\hat{g} = 0$, i.e., $g = 0$. So $a_k = 0$ for all $k$, thus $\sigma = 0$.

**Lemma 3.3.4.** *Let $\Phi$ be a compatly supported function such that $\hat{\Phi}(\xi) \geq 0$. If $\Phi \not\equiv 0$, then there exists $\lambda > 1$ such that $\widehat{\Phi_\lambda}(k) > 0$ for all $k \in \mathbb{Z}$, where $\Phi_\lambda(t) = \lambda\Phi(\lambda t)$.*

Proof: Note that the function $\hat{\Phi}(\xi)$ is the restriction to $\mathbb{R}$ of an entire function increasing exponentially. Consequently, the zeros of this function are isolated and form a sequence $\{\xi_i\}$, $j \in \mathbb{Z}$. Note that this sequence can be finite or empty. Then there exist $\lambda > 1$ such that $\lambda\xi_j \notin \mathbb{Z}$, $j \in \mathbb{Z}$. It follows that $\widehat{\Phi_\lambda}(k) \neq 0$, $k \in \mathbb{Z}$.

**Lemma 3.3.5.** *Consider $G(x) = \sum_{k=-\infty}^{\infty} \alpha_k exp(2\pi ikx)$. Assume that $G \in \mathcal{C}^\infty(\mathbb{T})$, $\alpha_k \geq 0$ excepting a finite number of them and also $G = 0$ over $[-\frac{\alpha}{4}, \frac{\alpha}{4}]$. Then $G = 0$.*

Proof: From the definition of the function $G$ we get $G^{(2m)}(0) = 0$. Hence

$$\sum_{k=-\infty}^{\infty} \alpha_k k^{2m} = 0.$$

By hypothesis, $\alpha_k \geq 0$ if $|k| > N_0$. Let

$$S(m) = \sum_{|k|>N_0} \alpha_k k^{2m}.$$

We denote by $k_1$ and index satisfying $|k_1| > N_0$ and $\alpha_{k_1} > 0$. It follows that

$$S(m) \geq \alpha_{k_1} k_1^{2m}.$$

Now let

$$R(m) = \sum_{|k| \leq N_0} \alpha_k k^{2m}.$$

Therefore

$$|R(m)| \leq C N_0^{2m}.$$

Since by construction $R(m) + S(m) = 0$, we deduce that

$$S(m) = -R(m) \leq C N_0^{2m}.$$

Then we deduce the following estimate

$$\alpha_{k_1} k_1^{2m} \leq C N_0^{2m}.$$

This is a contradiction if we let $m \to \infty$ since $k_1 > N_0$. Now, if such an index $k_1$ does not exist, we obtain that $G(x) = \sum_{|k| \leq N_0} \alpha_k exp(2\pi i k x)$ is an entire function that vanishes over an interval, which is again a contradiction.

### 3.3.3 Exact reconstruction for signed discrete measures

Matei's theorem in the previous subsection deals with positive measures. More precisely, it says that exact reconstruction is possible for a positive discrete measure with finite support when we know the spectral information on a model set. In this subsection, we try to extend this result to signed measures. Below, we show that a discrete measure whose support contains only finitely many translates of a subgroup of $\mathbb{R}^2$ is uniquely defined by its spectral information on a model set.

**Theorem 3.3.6.** *If $\rho$ be a discrete measure on $\mathbb{T}^2$ with finite support, whose Fourier coefficients vanishes on the model set*

$$\Lambda_\alpha = \{(m,n) \in \mathbb{Z}^2 : \ (m\sqrt{2} + n\sqrt{3}) \ mod \ 1 \in [-\alpha, \alpha]\},$$

*then $\rho = 0$.*

In order to prove Theorem 3.3.6, we need the following lemma:

**Lemma 3.3.7.** *Let $\theta$ be the triangle function on $\mathbb{T}$ supported on $[-\alpha, \alpha]$ defined by $\theta(0) = 1$, $\theta(\alpha) = \theta(-\alpha) = 0$, $\theta$ being affine on $[-\alpha, 0]$ and $[0, \alpha]$. Then*

$$\theta(x) = \sum_{k \in \mathbb{Z}} p_k exp(2\pi i k x), \quad where \ p_k \geq 0.$$

*Let $\tau$ be the measure $\sum_{k \in \mathbb{Z}} p_k \delta_{y_k}$, where $y_k = (k\sqrt{2}, k\sqrt{3}) \ mod \ \mathbb{Z}^2$. Then $\tau * \rho = 0$.*

Proof: By definition of the measure $\tau$, for all $(p, q) \in \mathbb{Z}^2$, we have

$$\hat{\tau}(p,q) = \sum_{k \in \mathbb{Z}} p_k exp(2\pi i (p\sqrt{2} + q\sqrt{3})k) = \theta(p\sqrt{2} + q\sqrt{3}).$$

So $\hat{\tau}(p,q) = 0$ when $(p,q) \notin \Lambda_\alpha$. But $\hat{\rho}(p,q) = 0$ when $(p,q) \in \Lambda_\alpha$. Thus $\hat{\tau} \cdot \hat{\rho} = 0$, which implies $\tau * \rho = 0$. $\square$

The proof of Theorem 3.3.6 is split into two parts:

$(i)$ We assume $\rho$ is of the form $\rho = \sum_{j=1}^{N} \alpha_j \delta_{x_j}$ where $x_j \in \mathbb{T}^2$, $j = 1, 2, ...N$. Thus

$$0 = \tau * \rho = \sum_{j=1}^{N} \alpha_j \ \tau * \delta_{x_j}.$$

Denote by $\Gamma$ the support of the measure $\tau$. Then $\Gamma = \{(k\sqrt{2}, k\sqrt{3}) : \ k \in \mathbb{Z}\}$ is a subgroup of $\mathbb{T}^2$. The support of $\tau * \delta_{x_j}$ is the coset $\Gamma + x_j$. Thus we can rearrange the above sum according to cosets as follows:

$$0 = \tau * \rho = \tau * \rho_1 + \cdots + \tau * \rho_K,$$

where $supp(\tau * \rho_i), \ i = 1, ..., K$ are disjoint cosets and each $\rho_i$ is of the form

$$\rho_i = \alpha_{j_1}\delta_{x_{j_1}} + \cdots + \alpha_{j_i}\delta_{x_{j_i}},$$

such that the pairwise differences between $x_{j_1}, ..., x_{j_i}$ all belong to $\Gamma$. Then it follows that $\tau * \rho_i = 0, \ i = 1, .., K$. So it suffices to prove that each $\rho_i = 0$, given that $\tau * \rho_i = 0$.

Now each $\rho_i$ is of the form $\sigma = \alpha_1\delta_{(\xi_1, \eta_1)} + \cdots + \alpha_M\delta_{(\xi_M, \eta_M)}$, where

$$\begin{cases} \xi_1 = \xi_M + k_1\sqrt{2} \\ \eta_1 = \eta_M + k_1\sqrt{3} \end{cases}, \cdots, \quad \begin{cases} \xi_{M-1} = \xi_M + k_{M-1}\sqrt{2} \\ \eta_{M-1} = \eta_M + k_{M-1}\sqrt{3} \end{cases}$$

and $k_1, ..., k_{M-1}$ are distinct integers that are nonzero. The condition that $\tau * \sigma = 0$ is equivalent to $\hat{\tau} \cdot \hat{\sigma} = 0$. But $\hat{\tau}(p, q) = \theta(p\sqrt{2} + q\sqrt{3}) \neq 0$ when $(p, q) \in \Lambda_\alpha$, so it follows that $\hat{\sigma}(p, q) = 0$ whenever $(p, q) \in \Lambda_\alpha$. Thus

$$0 = \hat{\sigma}(p, q) = \alpha_1 e^{-2\pi i(p\xi_1 + q\eta_1)} + \cdots + \alpha_M e^{-2\pi i(p\xi_M + q\eta_M)}, \quad if \ (p, q) \in \Lambda_\alpha.$$

Substitute using $k_1, ..., k_{M-1}$ we get

$$\alpha_1 e^{-2\pi i(p\sqrt{2}+q\sqrt{3})k_1} + \cdots + \alpha_{M-1} e^{-2\pi i(p\sqrt{2}+q\sqrt{3})k_(M-1)} + \alpha_M = 0,$$

whenever $(p, q) \in \Lambda_\alpha$.

$(ii)$ The result follows if we can show that there exist pairs $(p_1, q_1), ..., (p_{M-1}, q_{M-1}), (p_M, q_M) = (0, 0)$ such that $(p_i\sqrt{2} + q_i\sqrt{3})\ mod\ 1 \in (-\alpha, \alpha)$ and that the coefficient matrix

$$D = \begin{bmatrix} e^{-2\pi i(p_1\sqrt{2}+q_1\sqrt{3})k_1} & e^{-2\pi i(p_1\sqrt{2}+q_1\sqrt{3})k_2} & \ldots & e^{-2\pi i(p_1\sqrt{2}+q_1\sqrt{3})k_{M-1}} & 1 \\ e^{-2\pi i(p_2\sqrt{2}+q_2\sqrt{3})k_1} & e^{-2\pi i(p_2\sqrt{2}+q_2\sqrt{3})k_2} & \ldots & e^{-2\pi i(p_2\sqrt{2}+q_2\sqrt{3})k_{M-1}} & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ e^{-2\pi i(p_{M-1}\sqrt{2}+q_{M-1}\sqrt{3})k_1} & e^{-2\pi i(p_{M-1}\sqrt{2}+q_{M-1}\sqrt{3})k_2} & \ldots & e^{-2\pi i(p_{M-1}\sqrt{2}+q_{M-1}\sqrt{3})k_{M-1}} & 1 \\ 1 & 1 & \ldots & 1 & 1 \end{bmatrix}.$$

is non-singular.

To this end, let $\gamma_i = p_i\sqrt{2} + q_i\sqrt{3}$ so that the matrix $D$ becomes

$$D = \begin{bmatrix} e^{-2\pi i\gamma_1 k_1} & e^{-2\pi i\gamma_1 k_2} & \ldots & e^{-2\pi i\gamma_1 k_{M-1}} & 1 \\ e^{-2\pi i\gamma_2 k_1} & e^{-2\pi i\gamma_2 k_2} & \ldots & e^{-2\pi i\gamma_2 k_{M-1}} & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ e^{-2\pi i\gamma_{M-1} k_1} & e^{-2\pi i\gamma_{M-1} k_2} & \ldots & e^{-2\pi i\gamma_{M-1} k_{M-1}} & 1 \\ 1 & 1 & \ldots & 1 & 1 \end{bmatrix}.$$

We prove by induction from the bottom row. First, our choice of $(p_M, q_M) =$

$(0,0)$ makes the last row of $D$ being a row of constant 1. Next, it is clear that we can choose $\gamma_{M-1} = p_{M-1}\sqrt{2} + q_{M-1}\sqrt{3}$ such that $(p_{M-1}, q_{M-1}) \in \Lambda_\alpha$ and that the $Mth$ and $(M-1)th$ row being linearly independent. Indeed, any choice of $(p_{M-1}, q_{M-1}) \in \mathbb{Z}^2$ will work since in order for the last two row to be linearly dependent, each $\gamma_{M-1}k_j$, $j = 1, 2, ..., M-1$ must be an integer, which is a contradiction since $\gamma_{M-1}$ is irrational.

Next, suppose that we've already picked $\gamma_{M-1}, \gamma_{M-2}, ..., \gamma_{s+1}$ such that rows $M, M-1, ..., s+1$ are linearly independent. Now we are in the place of picking $\gamma_s = p_s\sqrt{2}+q_s\sqrt{3}$ in order that $(p_s, q_s) \in \Lambda_\alpha$ and that row $M, (M-1), ..., s+1, s$ are linearly independent. If this can not be done, then for any value of $\gamma_s = p_s\sqrt{2}+q_s\sqrt{3}$ such that $(p_s, q_s) \in \Lambda_\alpha$, the $(M-s+1) \times M$ matrix

$$D^* = \begin{bmatrix} e^{-2\pi i\gamma_s k_1} & e^{-2\pi i\gamma_s k_2} & \dots & e^{-2\pi i\gamma_s k_{M-1}} & 1 \\ e^{-2\pi i\gamma_{s+1} k_1} & e^{-2\pi i\gamma_{s+1} k_2} & \dots & e^{-2\pi i\gamma_{s+1} k_{M-1}} & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ e^{-2\pi i\gamma_{M-1} k_1} & e^{-2\pi i\gamma_{M-1} k_2} & \dots & e^{-2\pi i\gamma_{M-1} k_{M-1}} & 1 \\ 1 & 1 & \dots & 1 & 1 \end{bmatrix}$$

has row rank $M-s$, and by induction hypothesis, $D^*$ without the first row also has rank $M-s$. It follows that the $(M-s+1)\times(M-s+1)$ matrix

$$D^{**} = \begin{bmatrix} e^{-2\pi i\gamma_s k_1} & e^{-2\pi i\gamma_s k_2} & \dots & e^{-2\pi i\gamma_s k_{M-s+1}} \\ e^{-2\pi i\gamma_{s+1} k_1} & e^{-2\pi i\gamma_{s+1} k_2} & \dots & e^{-2\pi i\gamma_{s+1} k_{M-s+1}} \\ \vdots & \vdots & \vdots & \vdots \\ e^{-2\pi i\gamma_{M-1} k_1} & e^{-2\pi i\gamma_{M-1} k_2} & \dots & e^{-2\pi i\gamma_{M-1} k_{M-s+1}} \\ 1 & 1 & \dots & 1 \end{bmatrix}$$

is singular, for any $\gamma_s = p_s\sqrt{2}+q_s\sqrt{3}$ such that $(p_s,q_s)\in\Lambda_\alpha$. Let

$$f(\gamma_s) = det(D^{**}) = A_1 e^{-2\pi i\gamma_s k_1} + A_2 e^{-2\pi i\gamma_s k_2} + \cdots + A_{M-s+1} e^{-2\pi i\gamma_s k_{M-s+1}},$$

where the last equality is the Laplacian expansion of the matrix $D^{**}$ along the first row. Thus $A_j : j=1,2,...,M-s+1$ are the cofactors. But $D^{**}$ without the first row has full row rank, thus at least one $A_j$ is nonzero. Indeed, at least two of them must be nonzero, since $f(0)=0$. Assume after relabeling that

$$f(\gamma_s) = A_1 e^{-2\pi i\gamma_s k_1} + A_2 e^{-2\pi i\gamma_s k_2} + \cdots + A_l e^{-2\pi i\gamma_s k_l} = 0,\ if (p_s,q_s)\in\Lambda_\alpha,$$

where $A_1, A_2, ..., A_l$ are nonzero. It follows from Kroncker's theorem that $\{\gamma_s : (p_s,q_s)\in\Lambda_\alpha\}$ is dense in $(-\alpha,\alpha)$. It follows that $f(\gamma_s)\equiv 0$, for $\gamma_s\in(-\alpha,\alpha)$. Take the $nth$ derivative of $f(\gamma_s)$ and plug in $\gamma_s=0$, we get

$$A_1 k_1^n + A_2 k_2^n + \cdots + A_l k_l^n = 0, \quad \forall n\in\mathbb{N}.$$

Since by assumption, $k_1, k_2, ..., k_l$ are distinct and nonzero. Assume WLOG that $k_1 = argmax\{|k_j| : j = 1, 2, ..., l\}$. We divide by $k_1$ on both sides to get

$$A_1 + A_2\Big(\frac{k_2}{k_1}\Big)^n + \cdots + A_l\Big(\frac{k_l}{k_1}\Big)^n = 0, \quad \forall n \in \mathbb{N}.$$

Let $n \to \infty$, we get $A_1 = 0$, which is a contradiction. Thus the theorem is proved.

□

From the proof, we see that the support of $\rho$ need not be finite. Indeed, we only need that there are only finitely many points belonging to the same coset of $\Gamma$. Also, the definition of $\Lambda_\alpha$ is not unique. $\sqrt{2}$ and $\sqrt{3}$ can be changed to any pair of irrational numbers that are not linearly dependent over $\mathbb{Q}$. Thus Theorem 3.3.6 can be further refined into the following:

**Theorem 3.3.8.** *Let $\rho$ be a discrete measure (not necessarily positive) on $\mathbb{T}^2$. Suppose we can find a pair of irrational numbers $a, b \in \mathbb{R}$ such that the support of $\rho$ does not contain any infinite coset of the group $\{(ka, kb), k \in \mathbb{Z}\}$. If the Fourier coefficients of $\rho$ vanishes on the model set*

$$\Lambda_\alpha = \{(m, n) \in \mathbb{Z}^2 : \ (ma + nb) \ mod\ 1 \in [-\alpha, \alpha]\},$$

*then $\rho = 0$.*

Unfortunately, we cannot recover a signed measure via a the similar variational method as in the case of positive measures. Here is a one dimensional counter example. Let $\theta$ be the triangle function on $\mathbb{T}$ supported on $[\alpha, 1 - \alpha]$ defined by

$\theta(1/2) = 1$, $\theta(\alpha) = \theta(1-\alpha) = 0$, $\theta$ being affine on $[\alpha, 1/2]$ and $[1/2, 1-\alpha]$. Then

$$\theta(x) = \sum_{k=-\infty}^{\infty} (-1)^k \beta_k e^{2\pi i k x},$$

where $\beta_k > 0$. Define the finite measure

$$\nu_N = \sum_{|k| \leq N} (-1)^k \beta_k \delta_{y_k}, \quad \rho_N = \sum_{|k|>N} (-1)^k \beta_k \delta_{y_k},$$

where $y_k = (k\sqrt{2}, k\sqrt{3}) mod\ \mathbb{Z}^2$. Then by construction, for $\lambda = (p, q) \in \mathbb{Z}^2$, we have

$$\begin{aligned}
\hat{\nu}_N(\lambda) + \hat{\rho}_N(\lambda) &= \sum_{k=-\infty}^{\infty} (-1)^k \beta_k \widehat{\delta_{y_k}}(p, q) \\
&= \sum_{k=-\infty}^{\infty} (-1)^k \beta_k e^{2\pi i k (p\sqrt{2} + q\sqrt{3})} = \theta(p\sqrt{2} + q\sqrt{3}) = 0,
\end{aligned}$$

when $\lambda \in \Lambda_\alpha$. Thus, $\hat{\nu}_N(\lambda) = -\hat{\rho}_N(\lambda)$, $\lambda \in \Lambda_\alpha$. But the norm of $\rho_N$ goes to zero as $N \to \infty$. It follows that for $N$ large enough, $\|\rho_N\| < \|\nu_N\|$. So $\nu_N$ is not the solution of

$$argmin\{\|\mu\| \ : \ \hat{\mu}(\lambda) = \hat{\nu}_N(\lambda), \lambda \in \Lambda_\alpha\}.$$

# Chapter 4: SINGLE PIXEL CAMERA

## 4.1 Background

Nowadays, we can cheaply manufacture a sensor with millions of pixels that is sensitive to the visible light spectrum. But is more desirable to use fewer sensors when it comes to light spectra where the sensors are much more expensive, e.g. infrared or ultra-violet sensors.

Richard G. Baraniuk et al. have constructed a single-pixel camera using a digital micromirror device (DMD) and compressed sensing techniques to produce grayscale images, see [12]. In order to go beyond DMD technology, Dr. David Bowen of LPS has introduced the concept and is designing the experiment for the construction of a single-pixel camera that utilizes a liquid crystal display (LCD). The idea is to simulate a pixel grid with an LCD filter and ”sum up” the resulting light that comes through it with only a single-pixel light sensor.

In this chapter, we conduct an experiment to verify that the mathematical model behind single-pixel camera actually works, that is, the measurements taken from the sum of the light through the LCD can provide a good reconstruction. In addition, we also propose a way of shifting the LCD to enhance the resolution of the reconstruction.

## 4.2 Mathematical modeling

The experimental design that we use is outlined in Figure 4.1. Ambient light reflects off the target and passes through the front lens. This lens focuses the light into a beam which is directed at an LCD. The LCD is a grid of squares, say 1024 by 768, which should be though of as pixels in a typical sensor. Each square can be switched on or off, letting light to pass through, or not, according our configuration. The total admitted light is captured by the back lens, which then concentrates the light into a single-pixel CCD (charge coupled device) or CMOS (complementary metal oxide semiconductor) sensor. We record different measurements according to different configurations of the LCD. In general, if we wish to reconstruct an $m \times n$ image, we would need to take $m \times n$ measurements. However, compressed sensing theory says that, if the image vector is sparsely generated, we can take less measurements and look for the sparest solution.

Suppose we have an image of size $N = m \times n$, which is denoted by the matrix $(\mathbf{a}_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n}$ We transform this image into a column vector $\mathbf{y} \in \mathbb{R}^N$ by stacking the columns of the matrix, that is $\mathbf{a}_{i,j} = \mathbf{y}_{i+m(j-1)}$. The LCD can be viewed as a matrix $(\mathbf{b}_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n}$ of the same size, made of 1's and 0's, which corresponds to whether or not letting the light pass through the LCD at that location. We also transform the sampling mask by stacking the column of the sampling matrix $\mathbf{b}$ to create a sampling vector $\mathbf{p} \in \mathbb{R}^N$, i.e. $\mathbf{b}_{i,j} = \mathbf{p}_{i+m(j-1)}$. The measured total pass of light is then given by the dot product $\mathbf{P}^{\mathrm{T}}\mathbf{y}$. Now We can take several different configuration of the LCD mask, so that we can get different measurements. This corresponds to combining several different vectors $\mathbf{p}$. If we want to take $k$ measurements, we would have a sampling matrix $\mathbf{P} \in \mathbb{R}^{N\times k}$. The measured light pass will be given by $\mathbf{P}^{\mathrm{T}}\mathbf{y} \in \mathbb{R}^k$. Instead of trying to find the solution $\mathbf{y}$, which represents the pixel value of the original image, we apply the discrete cosine transform, and try to find the image representation under the cosine bases. That is, let $\mathbf{y} = \mathbf{D}\mathbf{x}$, where $D$ is the discrete cosine transform matrix. $\mathbf{x}$ is the coefficients of the original image under the cosine bases. Thus, if $\hat{\mathbf{y}}$ is the measured light pass, then we want to solve

$$\hat{\mathbf{y}} = \mathbf{P}^{\mathrm{T}}\mathbf{D}\mathbf{x}$$

for $\mathbf{x}$. However, if the number of total measurements $k$ is less then the number of unknown variables $N$, this is an under-determined system. We wish the image is sparse under the discrete cosine basis so that so that the solution found by compressed sensing is close to the actual solution.

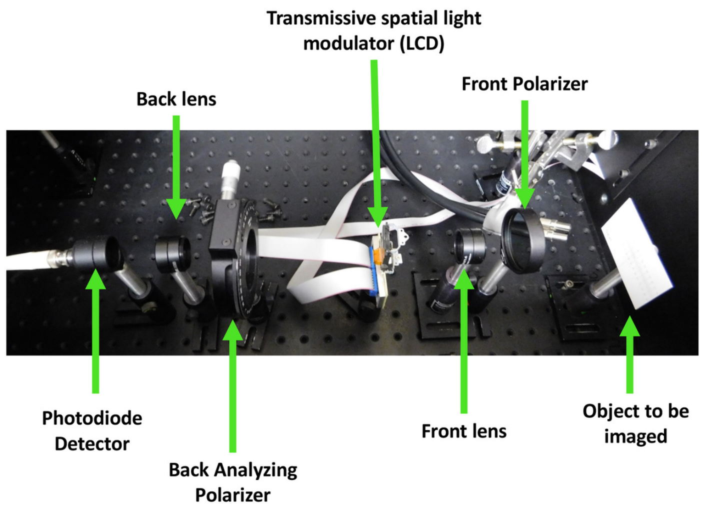



Figure 4.1: Physical experimental design. Photo modified from an original provided courtesy of David Bowen, Laboratory for Physical Sciences.

## 4.3 Compressed sensing

We can formulate the problem of reconstructing $\mathbf{y}$ as a compressed sensing problem with noise:

Given $\hat{\mathbf{y}} \in^{k}$, $\mathbf{P} \in^{N\times k}$, and $\epsilon > 0$, find $\mathbf{x}^* \in^{N}$ solving

$$\min_{\mathbf{x}\in^{N}} \|\mathbf{x}\|_0, \quad \text{subject to } \left\|\mathbf{y} - \mathbf{P}^{\mathrm{T}}\mathbf{D}\mathbf{x}\right\|_2 < \epsilon, \tag{4.1}$$

where $\|\mathbf{x}\|_0 = \#\{x_i : x_i \neq 0,\ \mathbf{x} = (x_0, x_1, \ldots, x_{N-1})^{\mathrm{T}}\}$.

If image $\mathbf{y}$ is close to a sparse representation in the basis or frame $\mathbf{D}$ with at most $s$ nonzero elements, then the theory of compressed sensing can guarantee that we can find a solution $\mathbf{x}^*$ to problem (4.1) provided $k = O(s\log(N/s))$. Algorithms for finding such minimizing $\mathbf{x}^*$ include the Orthogonal Matching Pursuit (OMP) algorithm [30] and Basis Pursuit (BP) algorithm [6], for example.

Let $\Phi \in^{K \times N}$. Thinking of $\Phi$ as $\mathbf{P}^{\mathrm{T}}\mathbf{D}$,

**Definition 4.3.1.** Let $\Phi \in \mathbb{R}^{K \times N}$ and denote the $i-th$ column of $\Phi$ as $\phi_i$. The *mutual coherence* of $\Phi$ is given by

$$\mu(\Phi) = \max_{1 \leq i,j \leq N, i \neq j} \frac{|\langle \phi_i, \phi_j \rangle|}{\|\phi_i\|_2 \|\phi_j\|_2}.$$

It is known that for full-rank matrices of size $K \times N$, the mutual coherence is bounded below by

$$\mu \geq \sqrt{\frac{N-K}{K(N-1)}}.$$

**Theorem 4.3.1.** *If a system of linear equations $\Phi\mathbf{x} = \mathbf{y}$ has a solution $\mathbf{x}$ obeying*

$$\|\mathbf{x}\|_0 < \frac{1}{2}\left(1 + \frac{1}{\mu(\Phi)}\right),$$

*then this solution is necessarily the sparsest possible.*

One of the algorithms which can solve this problem is known as the Orthogonal Matching Pursuit (OMP). This algorithm is detailed below:

**Orthogonal Matching Pursuit Algorithm**

**Task:** Approximate the solution of (4.1): $\min_{\mathbf{x}} \|\mathbf{x}\|_0$ subject to $\Phi\mathbf{x} = \mathbf{y}$.

**Parameters:** We are given $\Phi$, $\mathbf{y}$, and the error threshold, $\epsilon_0$.

**Initialization:** Initialize $k = 0$ and set:

- The initial solution $\mathbf{x}^0 = 0$.
- The initial residue: $\mathbf{r}^0 = \mathbf{y} - \Phi\mathbf{x} = \mathbf{y}$.
- The initial solution support: $\mathcal{S}^0 = \text{supp}(\mathbf{x}^0) = \emptyset$.

**Main Iteration:** Increment $k$ by 1 and perform the following:

- **Sweep:** Compute the errors $\epsilon(j) = \min_{z_j} \|\phi_j z_j - \mathbf{r}^{k-1}\|_2^2$ for all $j$ using the optimal choice $z_j^* = \phi_j^T \mathbf{r}^{k-1}/\|\phi_j\|_2^2$.
- **Update Support:** Find a minimizer $j_0$ of $\epsilon(j) : \forall j \notin \mathcal{S}^{k-1}, \epsilon(j_0) \leq \epsilon(j)$ and update $\mathcal{S}^k = \mathcal{S}^{k-1} \cup \{j_0\}$.
- **Update Provisional Solution:** Compute $\mathbf{x}^k$, the minimizer of $\|\Phi\mathbf{x} - \mathbf{y}\|_2^2$ subject to $\text{supp}(\mathbf{x}) = \mathcal{S}^k$.
- **Update Residual:** Compute $\mathbf{r}^k = \mathbf{y} - \Phi\mathbf{x}^k$.
- If $\|\mathbf{r}^k\|_2 < \epsilon_0$, then stop. Otherwise, apply another iteration.

**Output:** The proposed solution $x^k$ is obtained after $k$ iterations.

If the desired solution $\mathbf{x}$ meets the same sparsity requirements as Theorem 4.3.1, then the following theorem states that OMP will always find this solution.

**Theorem 4.3.2.** *For a system of linear equations $\Phi\mathbf{x} = \mathbf{y}$ ($\Phi \in \mathbb{R}^{K\times N}$ full-rank),* *if a solution* $\mathbf{x}$ *exists obeying*

$$\|\mathbf{x}\|_0 < \frac{1}{2}\left(1 + \frac{1}{\mu(\Phi)}\right),$$

*then OMP run with threshold parameter* $\epsilon = 0$ *is guaranteed to find it exactly.*

We now turn to the problem of allowing error or noise in our setup. To be specific, we wish to solve

$$\min_{\mathbf{x}\in^N} \|\mathbf{x}\|_0, \quad \text{subject to } \|\Phi\mathbf{x} - \mathbf{y}\|_2 < \epsilon. \tag{4.2}$$

With regards to the OMP algorithm, this simply corresponds to running the OMP algorithm with $\epsilon_0 > 0$ as the stopping parameter. Since this leads to an earlier stopping time, this will necessarily lead to a solution at least as sparse as it would find for an exact solution. We also have the following stability result:

**Theorem 4.3.3.** *Consider the problem (4.7). Suppose a sparse vector* $\mathbf{x} \in^N$ *satisfies the sparsity constraint* $\|\mathbf{x}_0\|_0 < \frac{1}{2}(1 + 1/\mu(\Phi))$ *and gives a representation of* $\mathbf{y}$ *to within error tolerance* $\epsilon$ *(that is,* $\|\mathbf{y} - \Phi\mathbf{x}_0\|_2 \leq \epsilon$*). Every solution* $\mathbf{x}_\mathbf{0}^\epsilon$ *to (4.7) must obey*

$$\|\mathbf{x}_0^\epsilon - \mathbf{x}_0\|_2^2 \leq \frac{4\epsilon^2}{1 - \mu(A)(2\|\mathbf{x}_0\|_0 - 1)}. \tag{4.3}$$

## 4.4 Spatial super-resolution

The problem of spatial super-resolution comes in two different forms: single-image and multiple-image super-resolution, see [29]. We shall focus on the multiple-image type of super-resolution. In addition to the single-pixel imaging setup up before, we would like to shift the LCD grid by half of a pixel to try to increase the resolution of our image. Doing this requires either a second LCD, a second sensor, or very careful moving of the current LCD grid or sensor. However, we can circumvent this problem by using the following formulation.

We can view our $M \times N$ LCD grid as a grid of size $M/2 \times N/2$ by grouping $2 \times 2$ blocks of pixels. We treat each block as one single pixel in our $M/2 \times N/2$ image, so every pixel in the $2 \times 2$ grid is switched on and off together. We then use our single-pixel imaging algorithm on this setup. Next, we shift the $2 \times 2$ blocks by one pixel in the $M \times N$ grid, which corresponds to a half-block shift in the $M/2 \times N/2$ grid. With these new blocks, we repeat the experiment. We now have two images of size $M/2 \times N/2$ which are shifted by half of a pixel. We then hope to use known super-resolution techniques to stitch the two together. Furthermore, we can image using the full $M \times N$ grid as a ground truth to see how well we can approximate the $M \times N$ image using two or more $M/2 \times N/2$ images.

With regards to multiple image super-resolution, one question which arises is: Given two sets of evenly spaced points on , $\{0, a_1, \cdots, a_M\}$ and $\{0, b_1, \cdots, b_N\}$, can we construct a new set of evenly spaced points $\{0, c_1, \cdots, c_{N'}\}$ where the original two sets are subsets? This question is a 1-dimensional version of trying to align two images where one of them is a sub-pixel shift of the other that also possibly has warping. Unfortunately, the answer to the question as is, is no. If we combine the sets of $a_i$ and $b_j$ into a new increasing sequence $\{0, x_1, x_2, \cdots, x_{M+N}\}$, then if there was a lattice containing both there would be a $d$ such that $md = x_1$ and $nd = x_2$, where $m$ and $n$ are positive integers. Thus,

$$d = \frac{c_1}{m} = \frac{c_2}{n} \implies \frac{m}{n} = \frac{c_1}{c_2}.$$

However, if $c_1 \in \setminus$ and $c_2 \in$, we obtain a contradiction as the left side is rational and the right side is not.

The single pixel camera solution that we have provided relies on the theory of compressed sensing (CS). This, in turn, can be understood, in particular scenarios, as a generalization of the classical (Nyquist-Shannon) sampling theorem going back to Cauchy in the 1840s. It is with this perspective that we can make the tie-in with super-resolution.

To see this connection, observe first that CS will recover a signal $\mathbf{y} \in \mathbb{R}^n$ from fewer samples than the $n$ normally required if there exists a sparse representation of $\mathbf{y}$ in some basis of $\mathbb{R}^n$. To make this notion more concrete, let that basis have a matrix representation $\mathbf{D} \in \mathbb{R}^{n \times n}$. With this setting, if $\mathbf{y} = \mathbf{D}\mathbf{x}$ and $\|\mathbf{x}\|_0 = \#\{i \in \mathbb{N} : x_i \neq 0, \mathbf{x} = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}\} < n$, we say that $\mathbf{y}$ has a sparse representation in $\mathbf{D}$.

Then, CS guarantees that there is a number $k < n$ such that the solution $\mathbf{x}^*$

to the problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{x}\|_0 \quad \text{subject to } \|\mathbf{P}^{\mathrm{T}}\mathbf{D}\mathbf{x} - \mathbf{P}^{\mathrm{T}}\mathbf{y}\|_2 = 0, \tag{4.4}$$

is such that $\mathbf{y} = \mathbf{D}\mathbf{x}^*$, where $\mathbf{A} = \mathbf{P}^{\mathrm{T}}\mathbf{D} \in \mathbb{R}^{k\times n}$ a full-rank matrix and where the matrix $\mathbf{P} \in \mathbb{R}^{n\times k}$ is called the *sampling matrix*.

In other words, we can find a vector $\mathbf{x}^* \in \mathbb{R}^n$ with fewer nonzero entries than $n$ that recovers $\mathbf{y}$, a vector that generally needs $n$ terms/samples to be defined. We have compressed the sensing with $\mathbf{A}$. The role of the matrix $\mathbf{P}$ is to sample the signal $\mathbf{y}$ by means of the measurement $\mathbf{z} = \mathbf{P}^{\mathrm{T}}\mathbf{y} \in \mathbb{R}^k$, a vector defined by only $k$ elements in $\mathbb{R}$. This is the generalization of the aforementioned classical sampling theorem.

Now suppose that we count with two measurements $\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{R}^k$ of a signal $\mathbf{y} \in \mathbb{R}^{2n}$, and that we happen to know that the measurements are interlocked by half the pixel resolution that each of them would individually give. Then, we aim to recover $\mathbf{y}$ by obtaining the solution $\mathbf{x}^* \in \mathbb{R}^{2n}$ of

$$\min_{\mathbf{x}\in\mathbb{R}^{2n}} \|\mathbf{x}\|_0 \quad \text{subject to } \left\| \mathbf{P}^{\mathrm{T}}\mathbf{D}\mathbf{x} - \begin{pmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{pmatrix} \right\|_2 = 0, \tag{4.5}$$

where in this case we would have a corresponding sampling matrix $\mathbf{P} \in \mathbb{R}^{2n\times 2k}$, for which

$$\mathbf{P}^{\mathrm{T}}\mathbf{y} = \begin{pmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{pmatrix},$$

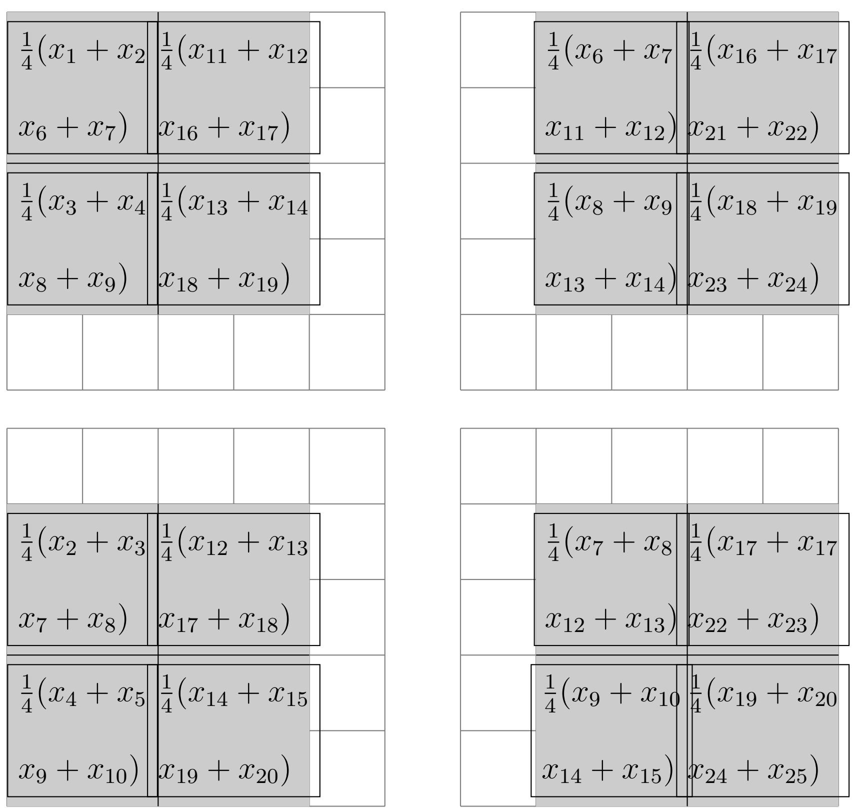
and basis matrix $\mathbf{D} \in \mathbb{R}^{2n \times 2n}$. The reconstruction would then be given by $\mathbf{y} = \mathbf{P}^{\mathrm{T}}\mathbf{D}\mathbf{x}^*$. This translates as having obtained a signal with twice the resolution of the resolution implied by the two original measurements, in which case we can say that we have super-resolved the signal by a factor of two.

## 4.5 An experiment

We want to test our idea of combining spatial super-resolution and compressed sensing mentioned above. Given an $N \times N$ image, where $N$ is of the form $N = 2n+1$, we take four low-resolution pictures of it, each of which has resolution $n \times n$. More specifically, we first form the low-resolution sampling grid by taking $2 \times 2$ blocks of the original (high-resolution) grid. We then choose to place the low resolution sampling grid at the top-left, top-right, bottom-left and bottom-right corner of the image. Thus we have four low-resolution images. By choice of $N = 2n + 1$, the second one is obtained by shifting the first one to the right by a half pixel, the third one is obtained by shifting the first one down by a half pixel, and the last one is obtained by shifting the first one both to the right and down by a half pixel. These half-pixel shifts are measured in the low-resolution grid, which is equivalent to a one pixel shift in the original (high-resolution) grid. We then sample each pixel of the four low-resolution images. This means that we take $4n^2$ measurements, from which we wish to recover the original image of size $N \times N = 4n^2 + 4n + 1$. Thus the sampling matrix $\mathbf{P} \in \mathbb{R}^{(4n^2+4n+1)\times 4n^2}$ and $\mathbf{D} \in \mathbb{R}^{(4n^2+4n+1)\times(4n^2+4n+1)}$. Let's illustrate this idea in the simplest case when $n = 2$ and $N = 2n + 1 = 5$. Thus we have the original image of resolution $5 \times 5$, suppose the image is vectorized, via stacking the columns, into a $25 \times 1$ vector $\mathbf{x} = (x_1, x_2, \dots, x_{25})^T$.

| | | | | |
|---|---|---|---|---|
| $x_1$ | $x_6$ | $x_{11}$ | $x_{16}$ | $x_{21}$ |
| $x_2$ | $x_7$ | $x_{12}$ | $x_{17}$ | $x_{22}$ |
| $x_3$ | $x_8$ | $x_{13}$ | $x_{18}$ | $x_{23}$ |
| $x_4$ | $x_9$ | $x_{14}$ | $x_{19}$ | $x_{24}$ |
| $x_5$ | $x_{10}$ | $x_{15}$ | $x_{20}$ | $x_{25}$ |

The following four picture shows how that we sample four low resolution pictures by taking the average of the $2 \times 2$ blocks. Notice that the half-pixel intertwine between the four low resolution images is essential. Their differences carry information about the original high resolution image.



Each of the 4 low resolution images contains 4 pixels. Thus we have taken $4 \times 4$ measurements on our high resolution image vector $\mathbf{x} = (x_1, x_2, \ldots, x_{25})^T$. This corresponds to a $25 \times 16$ sampling matrix $P$. Now assume the high resolution pixel values have coefficients in DCT basis denoted (with an abuse of notation) again by $\mathbf{x}$. We have

$$\mathbf{y} = \mathbf{P}^{\mathrm{T}}\mathbf{D}\mathbf{x}$$

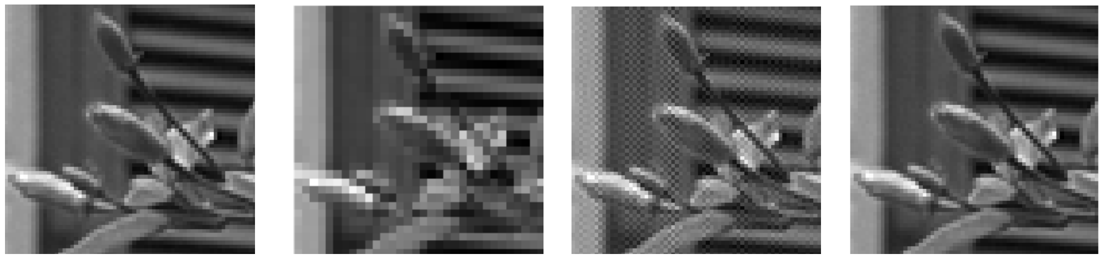
where $\mathbf{y} \in \mathbb{R}^{16}$, $\mathbf{x} \in \mathbb{R}^{25}$. Solving

$$\min_{\mathbf{x} \in ^N} \|\mathbf{x}\|_0, \quad \text{subject to } \|\mathbf{P}^{\mathrm{T}}\mathbf{D}\mathbf{x} - \mathbf{y}\|_2 < \epsilon. \tag{4.6}$$

However, the reconstruction is unsatisfactory when we try to minimize the $l_0$ norm. This is mainly because most of the real-world images are not sparsely generated under the DCT basis, while OMP algorithm will always find the sparsest solution. To solve this issue, we switch to $l_1$ minimization. If the image representation is sparse in DCT basis, BP will still find it as OMP does. However, for images that are not sparsely generated, BP will give a better reconstruction. Therefore, we turn to

$$\min_{\mathbf{x} \in ^N} \|\mathbf{x}\|_1, \quad \text{subject to } \|\mathbf{P}^{\mathrm{T}}\mathbf{D}\mathbf{x} - \mathbf{y}\|_2 < \epsilon. \tag{4.7}$$

Figure 4 shows a comparision between original image, one of the low resolution sample, the reconstruction via OMP and the reconstruction via BP. It is visually clear that the OMP reconstruction has a lot of noise, whereas the BP reconstruction



(a) Original. (b) Low resolution image. (c) OMP. (d) BP.

Figure 4.2: (j) A $65 \times 65$ pixel sub-image of a grayscale version of image I08.BMP in [31], (k) is one of the four low-resolution images , (l) is the reconstruction using OMP, and (d) is the reconstruction using BP. The quality as measured by the signal-to-noise ratio, peak signal-to-noise ratio, and mean structural similarity index for OMP reconstruction is, respectively, SNR = 21.4629 dB, PSNR = 30.6525 dB, and MSSIM = 0.8670. The signal-to-noise ratio, peak signal-to-noise ratio, and mean structural similarity index for BP reconstruction is, respectively, SNR = 43.1449 dB, PSNR = 51.4621 dB, and MSSIM = 0.9985.

is almost indistinguishable with the original image by human eyes.

Using BP, we try the same experiment on the TID2008 image date set which contains 24 images of size $384 \times 512$. Each image is partitioned into $32 \times 32$ sub-images. Since our method requires half-pixel shifts, we need to add two artificial edges at the right and bottom of the original image. The choice of the pixel values is given by alternating between the maximum and minimum pixel value of the $32 \times 32$ sub-image which is adjacent to that edge. For color images with $(R, G, B)$ channels, we deal with the luminance, which is defined by

$$Y = 0.299R + 0.587G + 0.114B$$

Here are two examples taken from the experimental results.


(a) Original.


(b) Low resolution image.

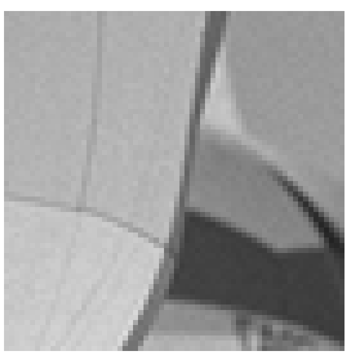

(c) Reconstruction.

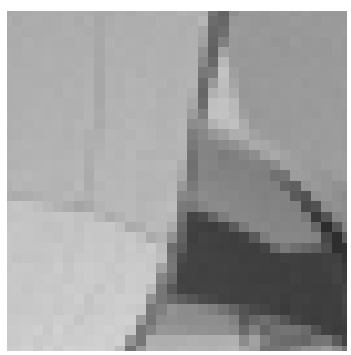

(d) Original.

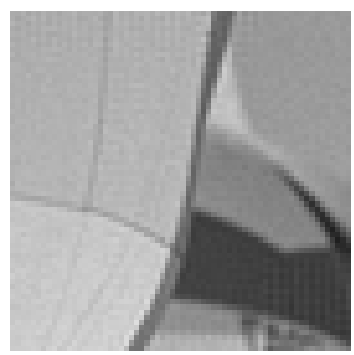

(e) Low resolution image.

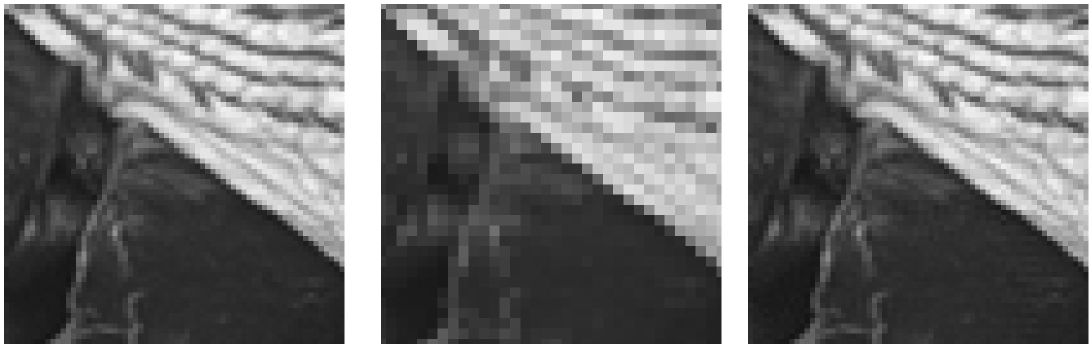

(f) Reconstruction.

To have a better look, we zoom in by taking a look at this algorithm on a $65 \times 65$ sub-image.


(g) Original.


(h) Low resolution image.


(i) Reconstruction.



(j) Original. (k) Low resolution image. (l) Reconstruction.

It can be seen that the reconstruction is very satisfactory. Indeed, we measured the SNR, PSNR and MSSIM on all 24 images for the reader's reference.

| Image# | PSNR | MSSIM | PSNR3×3 | MSSIM3×3 |
|---|---|---|---|---|
| I01 | 41.3193 | 0.9881 | 31.8555 | 0.9167 |
| I02 | 45.8413 | 0.9883 | 37.4531 | 0.9342 |
| I03 | 45.8399 | 0.9872 | 38.0074 | 0.9425 |
| I04 | 45.8368 | 0.9882 | 38.9321 | 0.9503 |
| I05 | 40.6353 | 0.9877 | 35.5134 | 0.9251 |
| I06 | 40.4479 | 0.9834 | 31.8997 | 0.9171 |
| I07 | 44.6227 | 0.9854 | 36.5344 | 0.9357 |
| I08 | 38.1665 | 0.9830 | 30.2109 | 0.9211 |
| I09 | 43.3458 | 0.9853 | 35.3490 | 0.9328 |
| I10 | 44.2966 | 0.9984 | 36.1120 | 0.9396 |
| I11 | 41.6080 | 0.9845 | 33.0221 | 0.9178 |
| I12 | 42.5323 | 0.9723 | 34.7063 | 0.8948 |
| I13 | 38.1261 | 0.9878 | 29.5492 | 0.9240 |
| I14 | 42.3648 | 0.9883 | 33.4042 | 0.9285 |
| I15 | 42.3730 | 0.9723 | 36.1456 | 0.9260 |
| I16 | 43.7124 | 0.9880 | 36.1287 | 0.9394 |
| I17 | 43.7148 | 0.9882 | 35.6682 | 0.9441 |
| I18 | 41.7711 | 0.9883 | 33.6720 | 0.9451 |
| I19 | 41.2583 | 0.9870 | 32.1446 | 0.9373 |
| I20 | 42.4007 | 0.9763 | 35.1158 | 0.9246 |
| I21 | 40.9993 | 0.9848 | 32.5454 | 0.9328 |
| I22 | 42.1259 | 0.9822 | 34.4933 | 0.9231 |
| I23 | 44.2328 | 0.9825 | 37.0566 | 0.9422 |
| I24 | 42.7809 | 0.9894 | 34.3414 | 0.9403 |

Figure 4.3: The reconstruction error of the TID2008 image data set measured by SNR, PSNR and MSSIM

The reconstruction is almost perfect, which leads us to thinking if we can require higher down-sampling rate. Namely, if we take the average of each 3 by 3 group of the original image, will our method still give a good reconstruction. In this case, we assume the original image has size $N \times N$, where $N = 3n + 2$ and our low resolution images have size $n \times n$. In this case, we take 9 low resolution images. The first one is taken by placing the sampling grid on the top-left corner of the original image. The rest low resolution images are derived by shifting the sampling grid to the right and down by a one-third pixel and two-thirds pixel measured in the low resolution grid level. So we have a sampling matrix $P \in \mathbb{R}^{(3n+2)^2 \times 9n^2}$ and the DCT matrix $D \in \mathbb{R}^{(3n+2)^2 \times (3n+2)^2}$. In this case, solving (11) will give us a reconstruction that triple the resolution in each direction. Below is the picture of the experiment on the same examples under the 3 by 3 case.

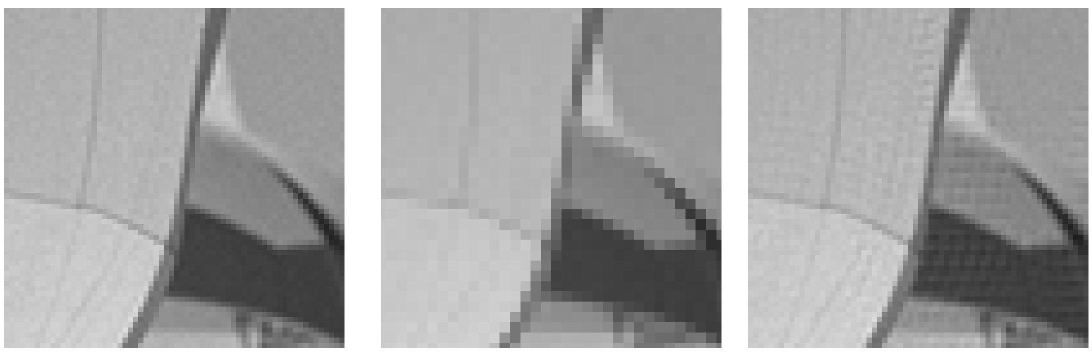

(a) Original.


(b) Low resolution image.

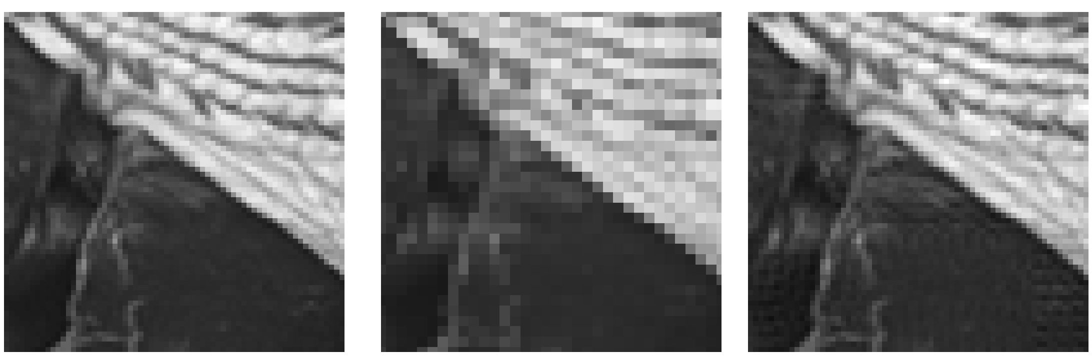

(c) Reconstruction.

(d) Original. (e) Low resolution image. (f) Reconstruction.

We also provide a closer look by looking at a $65 \times 65$ subimage with the same locations as the 2 by 2 case:



(g) Original. (h) Low resolution image. (i) Reconstruction.



(j) Original. (k) Low resolution image. (l) Reconstruction.

# Bibliography


[1] E. Au-Yeung and J. J. Benedetto, *Generalized fourier frames in terms of balayage*, Journal of Fourier Analysis and Applications, 21 (2015), pp. 472–508.

[2] J. J. Benedetto, *Harmonic analysis and applications*, vol. 23, CRC Press, 1996.

[3] A. Beurling, *Local harmonic analysis with some applications to differential operators*, Some Recent Advances in the Basic Sciences, 1 (1966), pp. 1962–1964.

[4] L. Bindi, N. Yao, C. Lin, L. S. Hollister, C. L. Andronicos, V. V. Distler, M. P. Eddy, A. Kostin, V. Kryachko, G. J. MacPherson, et al., *Natural quasicrystal with decagonal symmetry*, Scientific reports, 5 (2015), p. 9111.

[5] J. Carleson, P. Malliavin, J. Neuberger, and J. Wermer, *The collected works of arne beurling*, vol, 2 (1989), pp. 341–365.

[6] S. S. Chen, D. L. Donoho, and M. A. Saunders, *Atomic decomposition by basis pursuit*, SIAM review, 43 (2001), pp. 129–159.

[7] S. K.-K. Choi, A. V. Kumchev, and R. Osburn, *On sums of three squares*, International Journal of Number Theory, 1 (2005), pp. 161–173.

[8] O. Christensen et al., *An introduction to frames and Riesz bases*, Springer, 2016.

[9] N. G. De Bruijn, *Algebraic theory of penrose's non-periodic tilings of the plane*, Kon. Nederl. Akad. Wetensch. Proc. Ser. A, 43 (1981), pp. 1–7.

[10] D. L. Donoho and P. B. Stark, *Uncertainty principles and signal recovery*, SIAM Journal on Applied Mathematics, 49 (1989), pp. 906–931.

[11] P. Doukhan and F. Gamboa, *Superresolution rates in prokhorov metric*, Canadian Journal of Mathematics, 48 (1996), pp. 316–329.

[12] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, *Single-pixel imaging via compressive sampling*, IEEE signal processing magazine, 25 (2008), pp. 83–91.

[13] R. J. Duffin and A. C. Schaeffer, *A class of nonharmonic fourier series*, Transactions of the American Mathematical Society, 72 (1952), pp. 341–366.

[14] F. Gamboa and E. Gassiat, *Sets of superresolution and the maximum entropy method on the mean*, SIAM journal on mathematical analysis, 27 (1996), pp. 1129–1152.

[15] A. Guinand, *Concordance and the harmonic analysis of sequences*, Acta Mathematica, 101 (1959), pp. 235–271.

[16] A. Ingham, *A note on fourier transforms*, Journal of the London Mathematical Society, 1 (1934), pp. 29–32.

[17] S. Jaffard et al., *A density criterion for frames of complex exponentials.*, The Michigan Mathematical Journal, 38 (1991), pp. 339–348.

[18] M. N. Kolountzakis, *Fourier pairs of discrete support with little structure*, Journal of Fourier Analysis and Applications, 22 (2016), pp. 1–5.

[19] H. Landau et al., *Necessary density conditions for sampling and interpolation of certain entire functions*, Acta Mathematica, 117 (1967), pp. 37–52.

[20] P. D. Lax, *Functional Analysis*, Wiley-Interscience, 2002.

[21] N. Lev and A. Olevskii, *Quasicrystals with discrete support and spectrum*, arXiv preprint arXiv:1501.00085, (2014).

[22] B. Matei, *Exact reconstruction of the nonnegative measures using model sets*, Asymptotic Analysis, 97 (2016), pp. 291–299.

[23] ——, *Model sets and new versions of shannon sampling theorem*, in New Trends in Applied Harmonic Analysis, Springer, 2016, pp. 215–279.

[24] B. Matei and Y. Meyer, *Simple quasicrystals are sets of stable sampling*, Complex Variables and Elliptic Equations, 55 (2010), pp. 947–964.

[25] Y. Meyer, *Algebraic numbers and harmonic analysis, vol. 2*, 1972.

[26] Y. Meyer, *Mean-periodic functions and irregular sampling*, Det Kongelige Norske Videnskabers Selskab, 2018.

[27] Y. F. Meyer, *Measures with locally finite support and spectrum*, Proceedings of the National Academy of Sciences, 113 (2016), pp. 3152–3158.

[28] Olevskiĭ, *Universal sampling and interpolation of band-limited signals.*

[29] S. C. Park, M. K. Park, and M. G. Kang, *Super-resolution image reconstruction: a technical overview*, IEEE Signal Processing Magazine, (2003), pp. 21–36.

[30] Y. Pati, R. Rezaiifar, and P. Krishnaprasad, *Orthogonal matching pursuit: recursive function approximation with application to wavelet decomposition*, in 27th Asilomar Conference on Signals, Systems and Computers, 1993, 1993, pp. 40–44.

[31] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti, *TID2008 - A database for evaluation of full-reference visual quality assessment metrics*, Advances of Modern Radioelectronics, 10 (2009), pp. 30–45.

[32] D. Shechtman, I. Blech, D. Gratias, and J. W. Cahn, *Metallic phase with long-range orientational order and no translational symmetry*, Physical review letters, 53 (1984), p. 1951.

[33] A. 'e. Weil, *On the explicit formulas of number theory*, Mathematics of the USSR-Izvestiya, 6 (1972), p. 1.